Discrete Dynamics in Nature and Society

Discrete and Dynamic Optimization Problems in Operations Management 2014

Guest Editors: Xiang Li, Ou Tang, Weihua Liu, and Xiaochen Sun

Discrete and Dynamic Optimization Problems in Operations Management 2014

Discrete and Dynamic Optimization Problems in Operations Management 2014

Guest Editors: Xiang Li, Ou Tang, Weihua Liu, and Xiaochen Sun

Copyright © 2014 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in "Discrete Dynamics in Nature and Society." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

Mustapha A. Rami, Spain
Douglas R. Anderson, USA
Viktor Avrutin, Germany
Stefan Balint, Romania
Kenneth S. Berenhaut, USA
Gabriele Bonanno, Italy
Pasquale Candito, Italy
Jinde Cao, China
Wei-Der Chang, Taiwan
Cengiz Çinar, Turkey
Daniel Czamanski, Israel
Richard H. Day, USA
B.S. Daya Sagar, India
Manuel De la Sen, Spain
Josef Diblík, Czech Republic
Xiaohua Ding, China
Adam Doliwa, Poland
Zengji Du, China
E. Elabbasy, Egypt
H. A. El-Morshedy, Egypt
D. Fournier-Prunaret, France
Vladimir Gontar, Israel
Luca Guerrini, Italy
Taher S. Hassan, Egypt
Xue Z. He, Australia
Weihong Huang, Singapore
Bratislav Iricanin, Serbia
Giuseppe Izzo, Italy
Sarangapani Jagannathan, USA
Lee-Chae Jang, Korea
Jun Ji, USA
Zhen Jin, China
N. I. Karachalios, Greece
Hamid R. Karimi, Norway
Eric R. Kaufmann, USA
Recai Kilic, Turkey
Ryusuke Kon, Japan
Victor S. Kozyakin, Russia
Mustafa Kulenoviffj, USA
Jügen Kurths, Germany
Wei Nian Li, China
Xiaojie Lin, China
Xiaohui Liu, UK
Ricardo López-Ruiz, Spain
Qing-hua Ma, China
Akio Matsumoto, Japan
Rigoberto Medina, Chile
Eleonora Messina, Italy
Ugurhan Mugan, Turkey
Pham H. A. Ngoc, Vietnam
P. Niamsup, Thailand
Juan J. Nieto, Spain
P. K. Panigrahi, India
G. Papashinopoulos, Greece
B. Paternoster, Italy
Mingshu Peng, China
Carlo Piccardi, Italy
Peter Plath, Germany
Chuanxi Qian, USA
Vladimir S. Rabinovich, Mexico
Irena Rachůnková, Czech Republic
Morteza Rafei, The Netherlands
Aura Reggiani, Italy
Pavel Rehak, Czech Republic
Marko Robnik, Slovenia
Yuriy Rogovchenko, Norway
Ventsi G. Rumchev, Australia
R. Sahadevan, India
Leonid Shaikhet, Ukraine
Vimal Singh, Turkey
Seenith Sivasundaram, USA
Francisco Solis, Mexico
G. Solís-Perales, Mexico
Yihong Song, China
Michael Sonis, Israel
Jian-Ping Sun, China
Stepan Tersian, Bulgaria
Gerald Teschl, Austria
Tetsuji Tokihiro, Japan
Delfim F. M. Torres, Portugal
Firdaus Udwadia, USA
Antonia Vecchio, Italy
Qi-Ru Wang, China
Gongnan Xie, China
Ibrahim Yalcinkaya, Turkey
Xiang P. Yan, China
Xiaofan Yang, China
Bo Yang, USA
Her-Terng Yau, Taiwan
Jianshe Yu, China
Agacik Zafer, Turkey
Zhengqiu Zhang, China
Guang Zhang, China
Binggen Zhang, China
Lu Zhen, China
Baodong Zheng, China
Yong Zhou, China
Zhan Zhou, China
Zuo-nong Zhu, China

Contents

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 597980, 2 pages
http://dx.doi.org/10.1155/2014/597980

Editorial

Discrete and Dynamic Optimization Problems in Operations Management 2014

Xiang Li,[1] Ou Tang,[2] Weihua Liu,[3] and Xiaochen Sun[4]

[1] *College of Economic and Social Development, Nankai University, Tianjin 300071, China*
[2] *Department of Management and Engineering, Linköping University, 58183 Linköping, Sweden*
[3] *School of Management, Tianjin University, Tianjin 300072, China*
[4] *Department of Mathematics, School of Science, Tianjin University, Tianjin 300072, China*

Correspondence should be addressed to Xiang Li; xiangli@nankai.edu.cn

Received 9 December 2014; Accepted 9 December 2014; Published 31 December 2014

Copyright © 2014 Xiang Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This special issue in 2014, in the spirit of its predecessor published in 2013, continues with the aim of presenting the original research and review articles on the latest theoretical, numerical, and practical achievements on discrete and dynamic optimization problems in operation management. The issue contains 18 papers as follows.

X. Bao studies the influence of two types of overconfident behavior, overestimation, and overprecision, on decision of capacity recovery when power system's critical capacity is seriously damaged. A newsvendor model is used to prove that increasing regulatory punishment for electricity shortage and providing subsidy for capacity recovery are conducive measures to calibrate insufficient service level caused by an overconfident manager.

L. Geng et al. focus on the self-organization of supply chain with a MAS-based supply chain resilience model established. The local fitness function and neighborhood structure as well as interaction rules applicable to supply chain system are designed through viewing the enterprise as agent. It is found that there is agglomeration effect and SOC characteristic in supply chain and the evolution of supply chain is controlled by parameters of MAS.

X. Li introduces a review paper which proposes some hot issues in the current research by examining the most related existing literature from the perspective of operations management. Some insights and future research directions in this field are generated.

Y. Li et al. investigate the actuarial models of defined contribution pension plan. Through assumptions and calculations, the expected replacement ratios of three different defined contribution pension plans are compared. Particularly, more significant considerable factors are put forward in the further cost and risk analyses. In order to get an assessment of current status, the paper finds a relationship between the replacement ratio and the pension investment rate using econometrics method.

Q.-C. Meng et al. model the centralization ordering problem of retailers who face stochastic demands when the suppliers offer free shipping, in which limited distributional information such as known mean, support, and some deviation measures of the random data is needed only. The optimal order strategies of retailers are explored based on the linear decision rule mainly for stochastic programming. Further, the core allocation is also presented among all retailers.

C. G. Monyei et al. examine the characterization of six oil wells and the allocation of gas considering limited and unlimited case. A successive application of modified artificial neural network (MANN) is presented combined with a mild intrusive genetic algorithm (MIGA) to the oil well characteristics.

J. Qin and W. Liu develop an EOQ model under trade credit financing with ramp type demand and the demand dependent production rate. Subsequently, the algorithms are proposed to decide the optimal replenishment cycle and the optimal order quantity for the retailer. Finally, the numerical analysis is demonstrated to illustrate the models and the sensitively analysis is carried out to give some managerial insights.

D.-L. Sheng and X. Rong analyze the mean-variance insurers with the return of premiums clauses to study the optimal time-consistent investment strategy for the DC pension merged with an annuity contract. Both accumulation phase before retirement and distribution phase after retirement are studied. The numerical analysis is also conducted.

L. Xia and H. Zhi study the Stackelberg game between a retailer and a manufacturer considering the cap and trade system and analyse the impact of system parameters on the participants' decision making. It is also shown that the side-payment self-enforcing contract can resolve the arguments that the existing research overemphasizes on spontaneity of participation in side-payment contracts design.

W. Xue et al. establish the decision making model for a firm procuring multiple suppliers and spot markets. The suppliers are unreliable and provide different types of option-type supply contracts which should be made before demand realization, while the spot market can only be used after demand realization and has both the price and liquidity risks. The optimal portfolio policy for the firm is developed with conditions to find the qualified suppliers.

L. Zhang et al. focus on optimal investment strategy for a dual risk model under the condition that the company can invest into a risk-free asset and a risky asset. The precommitted strategy and time-consistent strategy are compared with the following results: the former can make value function maximized at the original time $t = 0$ and the latter strategy is time-consistent for the whole time horizon.

L. Zhang et al. establish a double objectives path optimization model with the consideration of carbon emissions cost and economy cost. The DNA-ant colony algorithm is used to solve the problem and its performance is shown to be better than that of the basic ant colony algorithm.

H. Zhang et al. propose the macroscopic model of air traffic flow in airport terminal area and carried out a series of simulation experiments with the NetLogo platform. Through both of the theoretical and practical discussions, the basic interrelationships and influential factors of air traffic flow characteristic parameters are revealed.

J.-H. Zhang et al. study a two-level and four-party supply chain consisting of a supplier, an e-commerce platform, the third-party logistics, and a demander. The existence of maximum profit of supply chain is obtained and the supply chain coordination issue is considered.

Y. Zhou et al. use an evolutionary ant colony algorithm based on RFID and knowledge refinement to solve the routing optimization problem of warehouse intelligent vehicle in complex environments. The algorithm is shown to be more effective than the traditional ant colony algorithm and genetic algorithm with higher convergence speed, and it can jump out of the U-type or V-type obstacle traps easily.

Some theoretical researches are also included in this special issue.

L. F. Lu et al. propose a new decomposition model based on accelerated proximal gradient method for packet-level traffic data and present the iterative scheme of the algorithm for network anomaly detection problem, which is termed as NAD-APG.

L. Shi et al. propose an approach to assess the gauge capability when a simple linear profile is used to reflect product quality. This method can simplify the problem when the measured values are multidimensioned. The example of spring measurement is presented in this paper which shows how to implement the proposed method.

L. Xu et al. introduce a new difference system from a differential competition system using different discrete methods. The paper gives theoretical analysis for local bifurcation of the fixed points and derives the conditions under which the local bifurcations such as flip occur at the fixed points. Furthermore, one- and two-dimensional diffusion systems are given when diffusion terms are added.

Xiang Li
Ou Tang
Weihua Liu
Xiaochen Sun

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 741654, 15 pages
http://dx.doi.org/10.1155/2014/741654

Research Article

Macroscopic Model and Simulation Analysis of Air Traffic Flow in Airport Terminal Area

Honghai Zhang, Yan Xu, Lei Yang, and Hao Liu

National Key Laboratory of Air Traffic Flow Management, Nanjing University of Aeronautics & Astronautics, Nanjing 211106, China

Correspondence should be addressed to Honghai Zhang; zhh0913@163.com and Yan Xu; ixuyan@hotmail.com

Received 31 March 2014; Accepted 2 July 2014; Published 25 August 2014

Academic Editor: Xiang Li

Copyright © 2014 Honghai Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We focus on the spatiotemporal characteristics and their evolvement law of the air traffic flow in airport terminal area to provide scientific basis for optimizing flight control processes and alleviating severe air traffic conditions. Methods in this work combine mathematical derivation and simulation analysis. Based on cell transmission model the macroscopic models of arrival and departure air traffic flow in terminal area are established. Meanwhile, the interrelationship and influential factors of the three characteristic parameters as traffic flux, density, and velocity are presented. Then according to such models, the macro emergence of traffic flow evolution is emulated with the NetLogo simulation platform, and the correlativity of basic traffic flow parameters is deduced and verified by means of sensitivity analysis. The results suggest that there are remarkable relations among the three characteristic parameters of the air traffic flow in terminal area. Moreover, such relationships evolve distinctly with the flight procedures, control separations, and ATC strategies.

1. Introduction

Air traffic management in terminal area is a knotty problem for controllers as this place is considered to be the air congestion, flight delay, and aviation accident-prone area. Researching on the fundamental operating features of terminal area traffic flow and deducing the macro emergence of traffic flow evolution may contribute to revealing the parameters with mutual relations, as well as the mechanism of spatiotemporal evolution, in terms of the traffic flow characteristic elements. By these means, we can move forward to exploring the objective law in air traffic in order to enrich the air traffic flow theory and to provide scientific basis for air traffic dispersion, which may have very important theoretical value and practical significance.

Traffic flow parameters are the physical variables that represent traffic flow characteristics including qualitative and quantitative features of operating states [1]. Basic theories of vehicle traffic flow have developed for decades and many results have been made by scholars. Lighthill and Whitham proposed the simulated dynamic model of traffic flow after researching on the evolution pattern of traffic flow under high traffic density circumstances [2]. Meanwhile, Richards proposed a first order continuum model of traffic flow, which has been integrated as the LWR theory [3]. Biham studied urban traffic flow based on a two-dimensional cellular automaton [4], while Daganzo researched into dynamic traffic problems with a cellular transmission model [5–8]. Compared to the vehicle traffic, less research has been devoted to air traffic flow theory so far, not to mention that many studies focusing just on modeling. A simplified Eulerian network model of air traffic flow was proposed by Menon et al. [9, 10], and Bayen et al. studied the liner control problems derived from Eulerian network model [11–13]. Laudeman et al. noted a quantitative mathematical model on dynamic density of air traffic flow [14]. Complexity model based on traffic flow disturbance was advocated by Lee et al. [15]. Liu et al. proposed a one-dimensional cellular transmission model specifically applicable to air route [16]. Wang et al. studied the microscopic plane-following performance and built the air freeway flow model [17]. Primary discussion for the stability of air traffic flow operating system was advocated by Zhang and Wang, while some basic characteristics of air traffic flow were also involved in this paper [18]. Such research findings have made great foundations for further study on air traffic flow theory. However, no detailed studies have investigated the characteristic parameters with their objective evolution law of air traffic flow. In this paper, we will combine traditional mathematical formula derivation with modern simulation techniques. A macroscopic model of air traffic flow in airport terminal area will be proposed, with which we will simulate and analyze the interrelation and influential evolvement law of the characteristic parameters with the goal of providing theoretical basis for scientific air traffic management.

2. Macroscopic Model

2.1. Definitions. There are many kinds of definitions for air traffic flow parameters, because of different research intentions and different methods that can be used. Since we focus on airport terminal area, the definitions of velocity, density, and traffic flux on a segment of air route are as follows.

Traffic flux (q) is the number of aircraft passing a reference profile of the observation segment per unit of time. $q = N/T$, where T is the observation time and N represents the number of aircraft passing in T.

Density (ρ) is the number of aircraft per unit length of the observation segment. $\rho = N/P$, where P means length of the observation segment, while N is the number of aircraft in P.

Velocity can be divided into micro and macro definitions. Microscopic definitions including instantaneous velocity (v) and average velocity ($\overline{v}$) focus on some point or profile of the observation segment; macroscopic ones including the space mean velocity ($\overline{v}^s$) and time mean velocity ($\overline{v}^t$) focus on some area extents or time ranges:

$$\begin{gathered} v = \frac{dx}{dt}, \qquad \overline{v} = \frac{1}{N}\sum_{i=1}^{N} v_i, \\ \overline{v}^s = D \cdot \left(\frac{1}{N}\sum_{i=1}^{N} t_i\right)^{-1}, \qquad \overline{v}^t = \frac{1}{N}\sum_{i=1}^{N}\frac{s_i}{t_1 - t_0}. \end{gathered} \tag{1}$$

One flight (f_i) of all (N) uses t_i time to pass the observation segment, where the actual distance is D. And s_i is the flying distance of f_i in the time period $t_1 - t_0$.

Historically, the first macroscopic traffic flow model is a continuity equation, called the Lighthill-Whitham-Richards (LWR) equation [2, 3]. Like the vehicle traffic flow, air traffic flux (q) and linear density (ρ) may also satisfy the corresponding relations for some given functions $f(\cdot)$ and $g(\cdot)$ with location x and time t, as follows:

$$\begin{gathered} \rho(x,t) = f(q(x,t), x), \\ q(x,t) = g(\rho(x,t), x), \end{gathered} \tag{2}$$

$$\frac{\partial q(x,t)}{\partial x} + \frac{\partial \rho(x,t)}{\partial t} = s. \tag{3}$$

On the right side of (3), s denotes the number of aircraft that enter or exit the observation segment. With the basic equation $q = \rho v$, where v denotes space mean velocity $\overline{v}^s$ and assuming every flight has the same flying case, we can derive

$$v(x)\frac{\partial q(x,t)}{\partial x} + \frac{\partial q(x,t)}{\partial t} = v(x)s. \tag{4}$$

There are two variables as density and velocity but only one equation in the LWR model. It cannot be solved as the differential equation is not closed. In response to this problem, a balancing velocity-density functional relationship, as $v(x) = v_e(\rho(x,t))$, was introduced into LWR theory by assuming traffic flow invariably in equilibrium state. To plug this into the equation, a hyperbolic equation of density can be derived, as follows:

$$\frac{\partial q_e(\rho(x,t))}{\partial x} + \frac{\partial \rho(x,t)}{\partial t} = s. \tag{5}$$

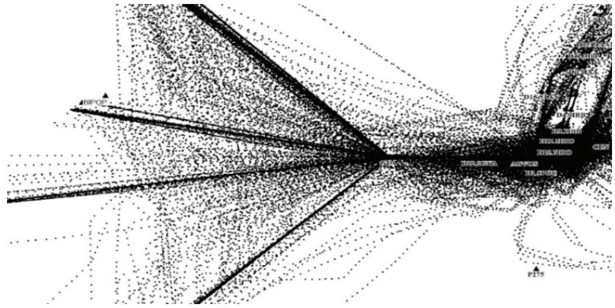


Figure 1: Radar track plot in airport terminal area (source: ZGGG TMA).

2.2. Arrival Traffic Flow Model. LWR model described the propagation characteristic of the nonlinear density wave to find the evolution rule of traffic shock wave and rarefaction wave by characteristics method or numerical simulation [19]. However, air traffic flow differs from vehicle traffic, especially in airport terminal area (Figure 1). First, the arrival and departure traffic in terminal area will follow the designed STAR/SID (standard terminal arrival route/standard instrument departure), which means that aircraft in every route position must act in accordance with the operational flight program, that is, flying within a certain scope of designed flight level and speed. It is usually a small scope and can be reassigned by controllers. As a result of that, some balancing velocity-density functional relationships in vehicle traffic may not exist in air traffic flow. Second, the density of air traffic flow is generally much lower than the vehicle traffic in real operations, which causes the interactions among aircraft to be weaker compared to vehicles. Therefore, it will be difficult to make statistical fit of the relationships according to the available radar data.

To solve the partial differential equation of the LWR theory, we use cell transmission model to discretize the continuity equation of macroscopic traffic flow. The method of discretization is applying a series of interconnected one-dimensional cells to denote the air route and using the difference equation of time discretization to describe aircraft passing through every adjoining cell. It should be feasible to

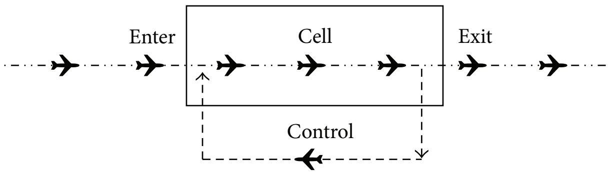


Figure 2: One-dimensional cell transmission model of air traffic flow in single direction.

model any air traffic flow scenarios by using interconnected cells, as shown in Figure 2, putting all the aircraft flying along the arrival route while in several flight levels onto the same imaginary plane, hence the air traffic flow in one route can be approximately seen as one-dimensional continuous flow. In addition, one air route will be divided into several segments by series of unit cells. To simplify the model we assume that there is a unified variable which we will introduce later to represent different kinds of control measures in one cell, such as speed control, maneuvering actions, or circling, that is, adjusting flows by changing speed or flight path of certain aircraft in the cell.

In the matter of arrival route, let $N_i^a(t)$ be the number of arrival aircraft in cell i at time t; then the changes in the number of arrival aircraft in one cell can be described by the following difference equation of time discretization:

$$N_i^a(t+1) = N_i^a(t) + \tau_i\left[q_{i-1}^a(t) - q_i^a(t)\right]. \tag{6}$$

In the equation above, $N_i^a(t+1)$ is the number of arrival aircraft in cell i at time $t+1$, $q_{i-1}^a(t)$ represents the flux of arrival aircraft entering cell i from cell $i-1$ at time t, while $q_i^a(t)$ means the flux of arrival aircraft exiting cell i at time t, and τ_i is the time step.

Note. The number, flux et al. mentioned in this section, is specific to arrival aircraft, while departures will be included in following parts.

Air traffic flow in air route segment still satisfies the basic equation $q_i = \rho_i v_i$, where ρ_i and v_i stand, respectively, for segment liner density and space mean velocity of cell i. The traffic flux of cell i will be $q_i(t) = \alpha_i q_i$. Coefficient α_i is the rate of outflow per unit time, which reflects the saturation level of cell i. From Section 2.1., we know segment liner density $\rho_i = N_i/\Omega_i$, where Ω_i is the length of cell i. Since most of the time arrival aircraft in airport terminal area are in a deceleration process, we assume aircraft entering cell i from some certain arrival route positions with an initial velocity v_i' that comes from the STAR and then uniformly decelerating along the cell (route) with a rate a_i^a, so we can get the space mean velocity of cell i, as follows:

$$v_i = \frac{a_i^a\Omega_i}{v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}}. \tag{7}$$

To plug (7) into $q_i = \rho_i v_i$, we can derive

$$q_i^a(t) = \frac{\alpha_i a_i^a N_i^a(t)}{v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}}. \tag{8}$$

However, since the aforementioned situation is an ideal condition, it is necessary to take control measures for part of the arrival aircraft into consideration, such as speed control, maneuvering actions, or circling, because of the existence of traffic congestion, safety interval, and so forth in real conditions. For simplicity, we assume that there is a unified variable to stand for different kinds of control measures, since whichever measures controllers took, the actual effects caused by changing speed or flight path and so forth can all be seen as there to be $N_i^{a\mathrm{ATC}}$ aircraft flying along the cell (route) with a new constant velocity v_i^{ATC} in cell i. To be clear, the variable v_i^{ATC} here represents the displacement velocity that is lower than the STAR designed velocity generally.

Considering these problems, we derived (9) after integrating control measures into (8), as follows:

$$\begin{aligned} q_i^a(t) &= \frac{\alpha_i a_i^a\left[N_i^a(t) - N_i^{a\mathrm{ATC}}(t)\right]}{v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}} + \frac{v_i^{\mathrm{ATC}} N_i^{a\mathrm{ATC}}(t)}{\Omega_i} \\ &= \frac{\alpha_i a_i^a}{v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}} N_i^a(t) \\ &\quad - \left(\frac{\alpha_i a_i^a}{v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}} - \frac{v_i^{\mathrm{ATC}}}{\Omega_i}\right) N_i^{a\mathrm{ATC}}(t). \end{aligned} \tag{9}$$

After substituting $q_i^a(t)$ in the difference equation (6) with (9), we got

$$\begin{aligned} N_i^a(t+1) &= \left(1 - \frac{\alpha_i a_i^a \tau_i}{v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}}\right) N_i^a(t) \\ &\quad + \left(\frac{\alpha_i a_i^a \tau_i}{v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}} - \frac{v_i^{\mathrm{ATC}}\tau_i}{\Omega_i}\right) N_i^{a\mathrm{ATC}}(t) \\ &\quad + \tau_i q_{i-1}^a(t). \end{aligned} \tag{10}$$

Let $A_i^a = 1 - (\alpha_i a_i^a \tau_i/(v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}))$; $B_i^a = (\alpha_i a_i^a/(v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i})) - (v_i^{\mathrm{ATC}}/\Omega_i)$; $C_i^a = \alpha_i a_i^a/(v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i})$; we can have the simplified equation sets, as follows:

$$\begin{gathered} N_i^a(t+1) = A_i^a N_i^a(t) + B_i^a \tau_i N_i^{a\mathrm{ATC}}(t) + \tau_i q_{i-1}^a(t), \\ q_i^a(t) = C_i^a N_i^a(t) - B_i^a N_i^{a\mathrm{ATC}}(t). \end{gathered} \tag{11}$$

The aircraft in cell are taken as evenly distributed in the macroscopic traffic flow theory, so the number of aircraft should be equivalent number. Let $d_i^a(t)$ be the mean nose interval of adjacent aircraft at the same direction in cell i; therefore the number of aircraft in the cell is $N_i(t) = \Omega_i/d_i^a(t)$. If $d_i^a(t)$ becomes less than the control separation standard denoted by $d_i^{a\mathrm{ATC}}$, that is to say, aircraft density exceeds the threshold level, the exceeded aircraft should be arranged to

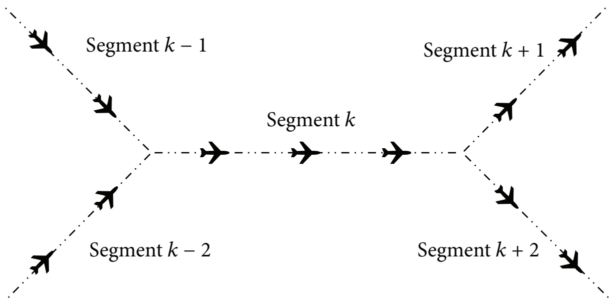


Figure 3: Converging and diverging of air traffic flow.

take control measures. In addition, arrival aircraft in airport terminal area also have to obey the safety separation d_i^{safe} according to STAR; that is, $d_i^a(t) \geq d_i^{\text{safe}}$. The safety separation is evaluated by air traffic security department and usually gets more stringent compared to control separation as $d_i^{\text{safe}} < d_i^{a\text{ATC}}$. The number of controlled aircraft at time t in cell i satisfies

$$N_i^{a\text{ATC}}(t) = \begin{cases} \left(\dfrac{1}{d_i^a(t)} - \dfrac{1}{d_i^{a\text{ATC}}}\right)\Omega_i & d_i^a(t) < d_i^{a\text{ATC}} \\ 0 & d_i^a(t) \geq d_i^{a\text{ATC}}. \end{cases} \tag{12}$$

We have discussed the one-dimensional cell transmission model of air traffic flow before, while there may be multiple air routes in terminal area and some of them may cross with each other, in the way, in the same imaginary plane, as shown in Figure 3. So the converging and diverging situations are as shown in Figure 3.

According to the conservation of number, in converging situation the number of aircraft satisfies

$$N_k = N_{k-1} + N_{k-2} + \cdots + N_{k-n}. \tag{13}$$

On the contrary, in diverging situation we have

$$\begin{aligned} N_{k+1} &= k_{k+1} N_k, \\ N_{k+2} &= k_{k+2} N_k, \\ &\vdots \\ N_{k+n} &= k_{k+n} N_k. \end{aligned} \tag{14}$$

The coefficient k stands for the proportion of traffic flow to different air route segments; meanwhile $0 \leq k_n \leq 1$, $\sum k_n = 1$.

2.3. Departure Traffic Flow Model. The model of departure flow in airport terminal area consists of two phases: taking off from runway and flying along air routes. In the take-off phase, departure air traffic flow will make the most use of runway time slots based on the arrival priority, in order to guarantee a smooth arrival and landing process. Flight departure schedules are generated in stochastic cases, while the departure flights that disagree with the operation time interval of runway will be delayed to ground holding procedures until enough time slots come [20–22]. In the flying stage, departure aircraft diverge in various directions from the runway center, which differs from the arrival flow as there would be both diverging and converging situations in arrival processes. Under the condition of fully isolation between arrival and departure routes, flying stage will dispense with flow control if the runway interval problem is solved in the taking-off stage. And under the condition of semi-isolation between arrival and departure routes, which means that some air routes may be overlapped in several segments (in the same imaginary plane), it is necessary to take both departure aircraft and arrival ones into consideration simultaneously to focus on the average nose interval of all aircraft in the overlapped segments (cells). If the average interval cannot meet safety needs, departure flow should be adjusted prior to the arrivals.

Let $N^{\text{ground}}(t)$ be the number of aircraft holding for departure on ground in time t, and its change can be described as follows:

$$N^{\text{ground}}(t+1) = N^{\text{ground}}(t) + \tau_i \left[q_{\text{plan}}^d(t) - q_s^d(t)\right]. \tag{15}$$

In (15), $N^{\text{ground}}(t+1)$ is the number of aircraft holding for departure on ground in time $t+1$. The demand of departure per unit time that produced by flight schedules is represented by $q_{\text{plan}}^d(t)$, while $q_s^d(t)$ represents the actual number of aircraft taking off from the runway per unit time. The time step is still τ_i, and $N^{\text{ground}}(t)$ cannot be negative, so we can get

$$q_s^d(t) = \begin{cases} \dfrac{N^{\text{ground}}(t)}{\tau_i} + q_{\text{plan}}^d(t) & N^{\text{ground}}(t) + \tau_i \left[q_{\text{plan}}^d(t) - C_s^d(t)\right] < 0 \\ C_s^d(t) & N^{\text{ground}}(t) + \tau_i \left[q_{\text{plan}}^d(t) - C_s^d(t)\right] \geq 0. \end{cases} \tag{16}$$

Consider maximizing the use of runway time slots; variable $C_s^d(t)$ represents the maximum take-off rate of the runway, which depends on the time interval of runway operation. If the time interval of arrival landing aircraft gets large, influence such as wake vortex on the runway will have no effect on take-off aircraft, in which cases departure flows may take off by the standard time interval separately. Otherwise if arrival time interval gets intense, mutual interference between departures and arrivals will be strong, so that departure flows need to make use of the interspace of the time slots under the circumstances of an affected runway operation, as follows:

$$C_s^d(t) = \begin{cases} \dfrac{dt_s^a(t) - dT_m^{\text{runway}}}{dT_m^{\text{runway}} \cdot dt_s^a(t)} & dt_s^a(t) < dT_s^{\text{runway}} \\ \dfrac{1}{dT_s^d} & dt_s^a(t) \geq dT_s^{\text{runway}}. \end{cases} \tag{17}$$

In (17), this $dt_s^a(t)$ is the actual time interval of arrival landing and dT_s^d is the standard take-off time interval of departures, while dT_s^{runway} and dT_m^{runway}, respectively, stand for the irrelevant runway operation time interval and the modified operation time interval according to the mutual interference between landing and taking off. Among this, the irrelevant runway operation time dT_s^{runway} will be a critical point where interference occurred.

Similar to the Arrival Model, we divide air routes into several segments by a series of unit-interconnected cells in the stage of flying along departure routes. Let $N_i^d(t)$ be the number of departure aircraft in cell i at time t and its changing process can be described as (18). On the left side of the equation $N_i^d(t+1)$ is the number of departures in cell i at time $t+1$; $q_{i-1}^d(t)$ represents the flux of departure aircraft entering cell i from cell $i-1$ at time t, while $q_i^d(t)$ means that the flux of departure aircraft exiting cell i at time t, τ_i is the time step.

Since most of the time departure aircraft in airport terminal area are in an acceleration process, we assume aircraft entering cell i from some certain departure route positions with an initial velocity u' that comes from the SID and then uniformly accelerating along the cell (route) with a rate a_i^d. So according to the same modeling principle from the Arrival Model, we can get the following similar equations:

$$\begin{aligned} q_i^d(t) &= \frac{\beta_i a_i^d\left[N_i^d(t) - N_i^{d\text{ATC}}(t)\right]}{\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u_i'} + \frac{u_i^{\text{ATC}} N_i^{d\text{ATC}}(t)}{\Omega_i} \\ &= \frac{\beta_i a_i^d}{\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u'} N_i^d(t) \\ &\quad - \left(\frac{\beta_i a_i^d}{\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u'} - \frac{u_i^{\text{ATC}}}{\Omega_i}\right) N_i^{d\text{ATC}}(t), \end{aligned} \tag{18}$$

$$\begin{aligned} N_i^d(t+1) &= \left(1 - \frac{\beta_i a_i^d \tau_i}{\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u_i'}\right) N_i^d(t) \\ &\quad + \left(\frac{\beta_i a_i^d \tau_i}{\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u_i'} - \frac{u_i^{\text{ATC}}\tau_i}{\Omega_i}\right) N_i^{d\text{ATC}}(t) \\ &\quad + \tau_i q_{i-1}^d(t). \end{aligned} \tag{19}$$

In the above equations, coefficient β_i stands for the rate of departures outflow per unit time. The number of aircraft that need to be arranged to take departure flow control measures is denoted by the variable $N_i^{d\text{ATC}}(t)$. As mentioned before, in circumstances of fully isolation between arrival and departure route segments there will be no extra control to the departures; that is, $N_i^{d\text{ATC}}(t) = 0$. On the other hand, in the routes overlapped condition part of the departures may take extra controls. We assume that these controlled departure flows would move with a new constant ATC velocity u_i^{ATC} in cell i and be the same with variable v_i^{ATC} in the Arrival Model, in which both represent displacement velocity. Meanwhile the simplification form is

$$\begin{aligned} N_i^d(t+1) &= A_i^d N_i^d(t) + B_i^d \tau_i N_i^{d\text{ATC}}(t) + \tau_i q_{i-1}^d(t), \\ q_i^d(t) &= C_i^d N_i^d(t) - B_i^d N_i^{d\text{ATC}}(t), \\ A_i^d &= 1 - \frac{\beta_i a_i^d \tau_i}{\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u_i'}; \\ B_i^d &= \frac{\beta_i a_i^d}{\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u_i'} - \frac{u_i^{\text{ATC}}}{\Omega_i}; \\ C_i^d &= \frac{\beta_i a_i^d}{\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u'}. \end{aligned} \tag{20}$$

To determine the value of variable $N_i^{d\text{ATC}}(t)$ in overlapped segments, we will bring the average nose interval (not time interval) of arrivals noted by $d_i^a(t)$, the average nose interval of departures noted by $d_i^d(t)$, and the average interval between both arrivals and departures noted by $d_i^{a/d}(t)$ in cell i all into consideration. The fundamental aim of flow control in this stage is to make the entire average interval of arrivals and departures meet the ATC separation requirement d_i^{ATC} by adjusting departure flows, as follows:

$$N_i^{d\text{ATC}}(t) = \begin{cases} \dfrac{\Omega_i}{d_i^d(t)} & d_i^a(t) \le d_i^{\text{ATC}} \\ \left(\dfrac{1}{d_i^{aDEP}(t)} - \dfrac{1}{d_i^{\text{ATC}}}\right) \cdot \Omega_i & d_i^a(t) > d_i^{\text{ATC}}, \\ & d_i^{a/d}(t) < d_i^{\text{ATC}} \\ 0 & d_i^{a/d}(t) \ge d_i^{\text{ATC}}. \end{cases} \tag{21}$$

The average interval of arrivals and departures is $d_i^{a/d}(t) = d_i^a(t) \cdot d_i^d(t)/[d_i^a(t) + d_i^d(t)]$.

3. Parameter Analysis

We focus on one single cell in the model as to deduce and analyze the interrelationship of air traffic flow characteristic parameters including flight flux, linear density, and traffic velocity. In the following sections, we discuss this problem from the two different aspects as arrival and departure, just like the models we had established before.

3.1. Arrival Routes. To plug (12) into (9), we can have

$$q_i^a(t) = \begin{cases} \dfrac{\alpha_i a_i^a \Omega_i}{d_i^{a\text{ATC}}\left(v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}\right)} + v_i^{\text{ATC}}\left(\dfrac{1}{d_i^a(t)} - \dfrac{1}{d_i^{a\text{ATC}}}\right) & d_i^a(t) < d_i^{a\text{ATC}} \\ \dfrac{\alpha_i a_i^a \Omega_i}{d_i^a(t)\left(v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}\right)} & d_i^a(t) \ge d_i^{a\text{ATC}}. \end{cases} \tag{22}$$

According to $\rho_i^a(t) = 1/d_i^a(t)$, we get

$$q_i^a(t) = \begin{cases} v_i^{\mathrm{ATC}} \rho_i^a(t) + \dfrac{\alpha_i a_i^a \Omega_i + v_i^{\mathrm{ATC}}\left(v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}\right)}{d_i^{a\mathrm{ATC}}\left(v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}\right)} & \rho_i^a(t) > \dfrac{1}{d_i^{a\mathrm{ATC}}} \\ \dfrac{\alpha_i a_i^a \Omega_i}{\left(v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}\right)} \rho_i^a(t) & \rho_i^a(t) \le \dfrac{1}{d_i^{a\mathrm{ATC}}}. \end{cases} \quad (23)$$

From (23), we can see that the basic parameters such as flux $q_i^a(t)$ and density $\rho_i^a(t)$ form a piecewise function and the segment point is the reciprocal value of ATC separation $d_i^{a\mathrm{ATC}}$ to the arrival aircraft. Assuming that coefficient α_i, standard deceleration a_i^a, initial speed v_i', and length Ω_i of one cell are all constant values, the slope and intercept of the relation curve will be determined by ATC velocity v_i^{ATC} and the position of inflection point will be determined by $d_i^{a\mathrm{ATC}}$. Let both sides of (23) be divided by $\rho_i^a(t)$:

$$v_i(t) = \begin{cases} \dfrac{\alpha_i a_i^a \Omega_i + v_i^{\mathrm{ATC}}\left(v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}\right)}{d_i^{a\mathrm{ATC}}\left(v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}\right)} \cdot \dfrac{1}{\rho_i^a(t)} + v_i^{\mathrm{ATC}} & \rho_i^a(t) > \dfrac{1}{d_i^{a\mathrm{ATC}}} \\ \dfrac{\alpha_i a_i^a \Omega_i}{\left(v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}\right)} & \rho_i^a(t) \le \dfrac{1}{d_i^{a\mathrm{ATC}}}. \end{cases} \quad (24)$$

If aircraft density comes lower than the critical value, the mean velocity of traffic flow in one cell will stay constant; otherwise, there will be an inverse relation between velocity $v_i(t)$ and density $\rho_i^a(t)$. With an increase of density in cell i, mean velocity will become lower gradually and the speed of reducing just gets more and more slow and eventually tends to the ATC velocity v_i^{ATC}. Using the basic formula of fluid $q = \rho v$, we can get

$$q_i^a(t) = \begin{cases} \dfrac{\alpha_i a_i^a \Omega_i + v_i^{\mathrm{ATC}}\left(v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}\right)}{d_i^{a\mathrm{ATC}}\left(v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}\right)} \cdot \dfrac{v_i(t)}{v_i(t) - v_i^{\mathrm{ATC}}} & v_i^{\mathrm{ATC}} < v_i(t) < \dfrac{\alpha_i a_i^a \Omega_i}{\left(v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}\right)} \\ \exists q_i^a(t) \in \left[0, \dfrac{\alpha_i a_i^a \Omega_i}{d_i^{a\mathrm{ATC}}\left(v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}\right)}\right] & v_i(t) = v_i^{\mathrm{ATC}}, \dfrac{\alpha_i a_i^a \Omega_i}{\left(v_i' - \sqrt{v_i'^2 - 2a_i^a\Omega_i}\right)}. \end{cases} \quad (25)$$

From (25), we can see that the mean velocity of entire traffic flow in cell i ranges from ATC velocity v_i^{ATC} to standard velocity designed by STAR. If $v_i(t)$ is equal to one of the critical values flux $q_i^a(t)$ may be arbitrary-sized data taking from zero to the maximum flux value. With an increase of mean velocity in cell i, the flux value goes down gradually and the speed of reducing gets slow. When this mean velocity goes up to the peak value, it will cause a jump of traffic flux. Variables $d_i^{a\mathrm{ATC}}$ and v_i^{ATC} are still the order parameters of the interrelationship between flux and velocity in cell.

3.2. Departure Routes. In the stage of flying along departure routes, aircraft keep accelerating and diverging along with the varying of orientation of air routes. There may be $N_i^{d\mathrm{ATC}}(t)$ aircraft taken departure flow control measures in some of the overlapped segments. As an exceptional case of this situation, flying stage will dispense with flow control in the case of fully isolation between arrival and departure routes; that is, $N_i^{d\mathrm{ATC}}(t) = 0$. We will focus on the overlapped segments and analyze the traffic flow characteristic parameters of departures in detail. To plug (21) into (18) we can have

$$q_i^d(t) = \begin{cases} \dfrac{u_i^{\mathrm{ATC}}}{d_i^d(t)} & d_i^a(t) \le d_i^{\mathrm{ATC}} \\ \dfrac{u_i^{\mathrm{ATC}}}{d_i^d(t)} + \dfrac{d_i^a(t) - d_i^{\mathrm{ATC}}}{d_i^{\mathrm{ATC}} d_i^a(t)} \times \left(\dfrac{\beta_i a_i^d \Omega_i}{\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u'} - u_i^{\mathrm{ATC}}\right) & d_i^a(t) > d_i^{\mathrm{ATC}}, \; d_i^{a/d}(t) < d_i^{\mathrm{ATC}} \\ \dfrac{\beta_i a_i^d}{d_i^d(t)\left(\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u'\right)} & d_i^t(t) \ge d_i^{\mathrm{ATC}}. \end{cases} \quad (26)$$

Since $\rho_i^d(t) = 1/d_i^d(t)$, $\rho_i^a(t) = 1/d_i^a(t)$, substituting in (26), we get

$$q_i^d(t) = \begin{cases} \dfrac{\beta_i a_i^d}{\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u'}\rho_i^d(t) & \rho_i^a(t) + \rho_i^d(t) < \dfrac{1}{d_i^{\mathrm{ATC}}} \\ u_i^{\mathrm{ATC}}\rho_i^d(t) + \dfrac{1 - d_i^{\mathrm{ATC}}\rho_i^a(t)}{d_i^{\mathrm{ATC}}} \times \left(\dfrac{\beta_i a_i^d\Omega_i}{\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u'} - u_i^{\mathrm{ATC}}\right) & \rho_i^a(t) < \dfrac{1}{d_i^{\mathrm{ATC}}},\ \rho_i^a(t) + \rho_i^d(t) > \dfrac{1}{d_i^{\mathrm{ATC}}} \\ u_i^{\mathrm{ATC}}\rho_i^d(t) & \rho_i^a(t) \geq \dfrac{1}{d_i^{\mathrm{ATC}}}. \end{cases} \tag{27}$$

From (27), we can see that flux $q_i^d(t)$ and density $\rho_i^d(t)$ form a piecewise function. Similar to the arrival segments (cells), we assume that coefficient β_i, standard acceleration a_i^d, initial speed u', and length Ω_i of one cell are all constant values. The density of arrival aircraft denoted by $\rho_i^a(t)$ in cell i changes over time, which has an effect on the intercept value of this linear equation. The ATC velocity of departures denoted by u_i^{ATC} shows the slope while the reciprocal of ATC separation d_i^{ATC} is the critical value of flow density, determining the position of inflection points. Then, letting both sides of (27) be divided by $\rho_i^d(t)$, we got

$$u_i(t) = \begin{cases} \dfrac{\beta_i a_i^d}{\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u'} & \rho_i^a(t) + \rho_i^d(t) < \dfrac{1}{d_i^{\mathrm{ATC}}} \\ u_i^{\mathrm{ATC}} + \dfrac{1 - d_i^{\mathrm{ATC}}\rho_i^a(t)}{d_i^{\mathrm{ATC}}\rho_i^d(t)} \times \left(\dfrac{\beta_i a_i^d\Omega_i}{\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u'} - u_i^{\mathrm{ATC}}\right) & \rho_i^a(t) < \dfrac{1}{d_i^{\mathrm{ATC}}},\ \rho_i^a(t) + \rho_i^d(t) > \dfrac{1}{d_i^{\mathrm{ATC}}} \\ u_i^{\mathrm{ATC}} & \rho_i^a(t) \geq \dfrac{1}{d_i^{\mathrm{ATC}}}. \end{cases} \tag{28}$$

From (28), we can see that mean velocity $u_i(t)$ and density $\rho_i^d(t)$ also form a piecewise function. Two extreme cases are as follows: the entire density of arrivals and departures is lower than the ATC density that is denoted by $1/d_i^{\mathrm{ATC}}$; density only considers arrivals that have already exceeded $1/d_i^{\mathrm{ATC}}$. These two cases will cause mean velocity of departures to be the standard velocity determined by β_i, acceleration a_i^d, initial speed u', length Ω_i, and departure ATC velocity u_i^{ATC}, respectively. In a certain condition between such two extreme cases the slope of linear equation is greater than or equal to u_i^{ATC}, while the intercept value is influenced by both u_i^{ATC} and $\rho_i^a(t)$. According to the basic formula of fluid $q = \rho v$, we get

$$q_i^d(t) = \begin{cases} \dfrac{\left(1 - d_i^{\mathrm{ATC}}\rho_i^a(t)\right)\left[\beta_i a_i^d\Omega_i - u_i^{\mathrm{ATC}}\left(\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u'\right)\right]}{d_i^{\mathrm{ATC}}\left(\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u'\right)} \cdot \dfrac{u_i(t)}{u_i(t) - u_i^{\mathrm{ATC}}} & u_i^{\mathrm{ATC}} < u_i(t) < \dfrac{\beta_i a_i^d}{\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u'} \\ \exists q_i^d(t) \in \left[0, \dfrac{\beta_i a_i^d\Omega_i}{d_i^{d\mathrm{ATC}}\left(\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u'\right)}\right] & u_i(t) = u_i^{\mathrm{ATC}},\ \dfrac{\beta_i a_i^d}{\sqrt{u_i'^2 + 2a_i^d\Omega_i} - u'}. \end{cases} \tag{29}$$

The form of (29) is much like (25) from the arrival parts. If mean velocity of departures lies between the standard velocity and the departure ATC velocity u_i^{ATC} there will be an inverse proportional function between departure flux $q_i^d(t)$ and mean velocity $u_i(t)$. Moreover, the coefficient of this inverse proportional function is determined by β_i, acceleration a_i^d, initial speed u', and cell length Ω_i all together and changes in value of the ATC velocity u_i^{ATC} will make the function curve

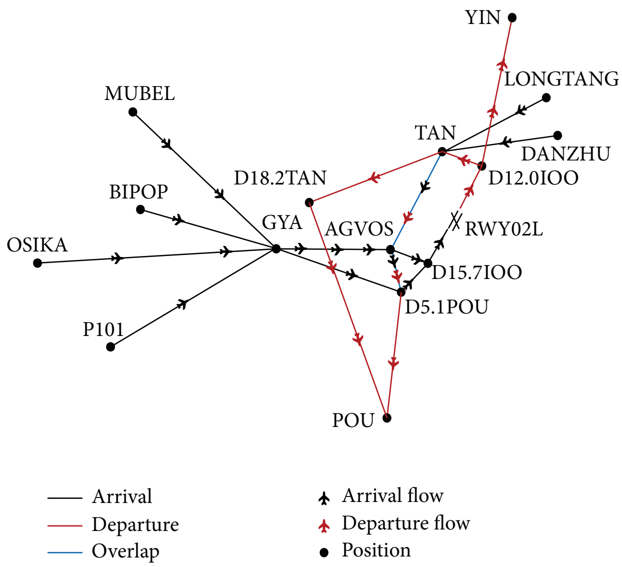


Figure 4: STAR/SID routes to RWY02L ZGGG.

move horizontally. While if mean velocity of departures lies on either of the two critical positions departure flux may vary randomly.

4. Simulation Experiment

4.1. Simulation Sample. Based on the simulation platform NetLogo [23–25] we focused on each cell in the macroscopic model and designed traffic inflow/outflow behaviors to be each agent, so as to take control of the inflow/outflow volumes among all the cells. Applying the STAR/SID procedures to runway 02L of Guangzhou Baiyun International Airport (as shown in Figure 4) into the Netlogo system dynamic simulator (as shown in Figure 5) we can emulate the operation and evolution process of air traffic flow in the airport terminal area.

The network of arrival and departure routes in terminal can be separated into several single direction segments and converging or diverging segments that can be further decomposed into more unit cells. For the purpose of simplicity, we set one cell dimension to be equal to the length of the shortest segment in the network and the other air segments to be integer multiples of unit cell dimension. According to this, we made the following Tables 1 and 2 of the terminal network.

Each yellow rectangle in Figure 4 represents one cell, which means a "stock" of fluid in the simulator system. Each grey pipe represents flowing between adjacent cells, which means "flow" in system, and the direction of arrow is the same as the flow direction. Each black valve represents control measures, which means "strategy" in system, and the inflow and outflow of air traffic will be under control by giving the certain valves some corresponding rules. Take typical overlapped segments "TAN=AGVOS (Segment9&19)" as an example, the cell transmission model of this is like (30). In the model, $N^a_{i/j}$ $A^a_{i/j}$ $B^a_{i/j}$ $C^a_{i/j}$, respectively, represents the number of arrival aircraft and the simplified coefficients of

Table 1: Cell quantities of arrival routes to RWY02L ZGGG.

Segment	Code	Quantity
MUBEL-GYA	Segment1	3 cells
BIPOP-GYA	Segment2	2 cells
OSIKA-GYA	Segment3	3 cells
P101-GYA	Segment4	2 cells
GYA-AGVOS	Segment5	2 cells
GYA-D5.1POU	Segment6	2 cells
LONGTANG-TAN	Segment7	2 cells
DANZHU-TAN	Segment8	2 cells
TAN-AGVOS	Segment9	2 cells
AGVOS-D15.7IOO	Segment10	1 cell
AGVOS-D5.1POU	Segment11	1 cell
D5.1POU-D15.7IOO	Segment12	1 cell
D15.7IOO-RWY02L	Segment13	1 cell

Table 2: Cell quantities of departure routes to RWY02L ZGGG.

Segment	Code	Quantity
RWY02L-D12.0IOO	Segment14	1 cell
D12.0IOO-YIN	Segment15	2 cells
D12.0IOO-TAN	Segment16	1 cell
TAN-D18.2TAN	Segment17	2 cells
D18.2TAN-POU	Segment18	3 cells
TAN-AGVOS	Segment19	2 cells
AGVOS-D5.1POU	Segment20	1 cell
D5.1POU-POU	Segment21	2 cells

cell j in Segment i. $N^{a\mathrm{ATC}}_{i/j}$ denotes the number of arrival aircraft taking control measures in cell j of Segment i. $q^a_{i/j}(t)$ means the number of arrival aircraft that exit cell j in Segment i per unit time. $N^d_{i/j}$ $A^d_{i/j}$ $B^d_{i/j}$ $C^d_{i/j}$ $N^{d\mathrm{ATC}}_{i/j}$ $q^d_{i/j}(t)$, respectively, represents the similar corresponding ones but specially for departures. Coefficient k stands for the proportion of traffic flow into some segment when diverging occurs in air routes. τ is the designed time step for the simulator system:

$$\begin{aligned}
N^a_{9/1}(t+1) &= A^a_{9/1}N^a_{9/1}(t) + B^a_{9/1}\tau N^{a\mathrm{ATC}}_{9/1}(t) + \tau\left[q^a_{7/2}(t) + q^a_{8/2}(t)\right], \\
q^a_{9/1}(t) &= C^a_{9/1}N^a_{9/1}(t) - B^a_{9/1}N^{a\mathrm{ATC}}_{9/1}(t), \\
N^a_{9/2}(t+1) &= A^a_{9/2}N^a_{9/2}(t) + B^a_{9/2}\tau N^{a\mathrm{ATC}}_{9/2}(t) + \tau q^a_{9/1}(t), \\
q^a_{9/2-10/1}(t) &= k_{9/2-10/1}\left[C^a_{9/2}N^a_{9/2}(t) - B_{9/2}N^{a\mathrm{ATC}}_{9/2}(t)\right], \\
q^a_{9/2-11/1}(t) &= k_{9/2-11/1}\left[C^a_{9/2}N^a_{9/2}(t) - B_{9/2}N^{a\mathrm{ATC}}_{9/2}(t)\right],
\end{aligned} \tag{30}$$

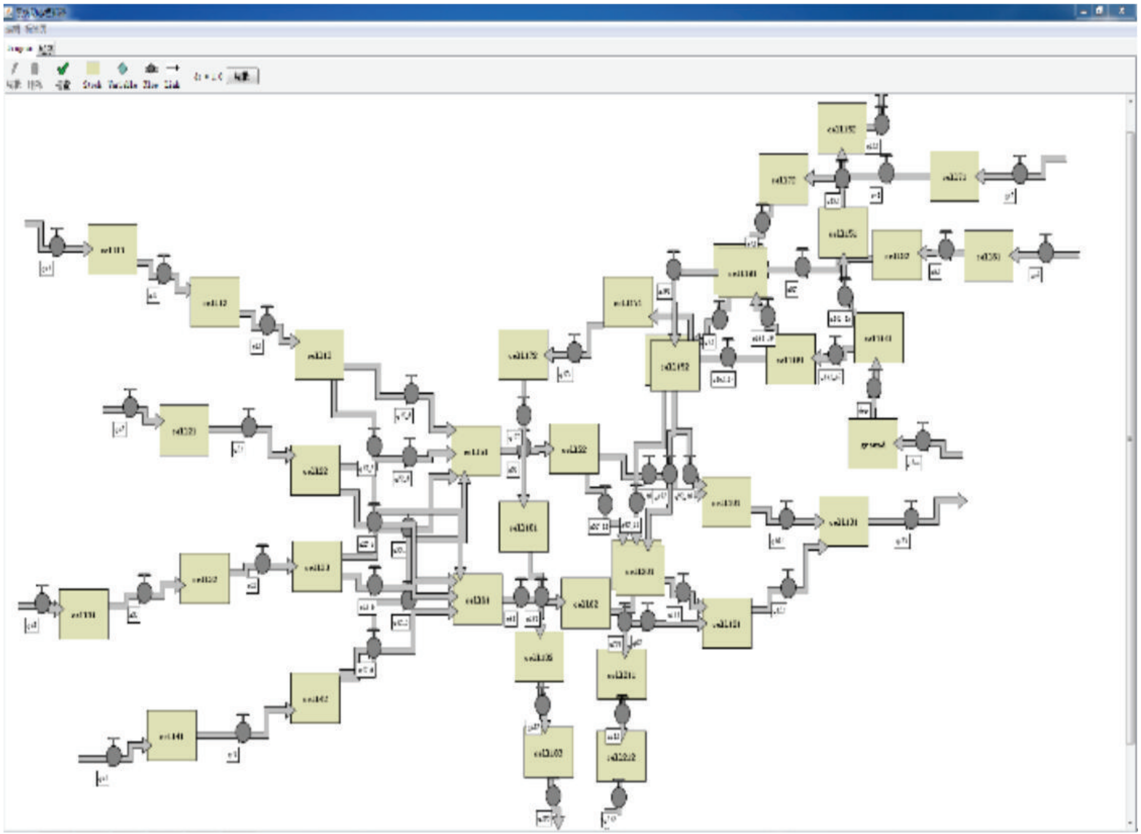


Figure 5: NetLogo system dynamic simulator.

$$\begin{aligned}
&N^d_{19/1}(t+1) \\
&\quad = A^d_{19/1}N^d_{19/1}(t) + B^d_{19/1}\tau N^{d\text{ATC}}_{19/1}(t) + \tau q^d_{16/1}(t), \\
&q^d_{19/1}(t) = C^d_{19/1}N^d_{19/1}(t) - B^d_{19/1}N^{d\text{ATC}}_{19/1}(t), \\
&N^d_{19/2}(t+1) \\
&\quad = A^d_{19/2}N^d_{19/2}(t) + B^d_{19/2}\tau N^{d\text{ATC}}_{19/2}(t) + \tau q^d_{16/2}(t), \\
&q^d_{19/2}(t) = C^d_{19/2}N^d_{19/2}(t) - B^d_{19/2}N^{d\text{ATC}}_{19/2}(t).
\end{aligned} \tag{31}$$

4.2. Simulation Design. In this simulation sample there are 6 entry points for arrivals of the terminal network: MUBEL, OSIKA, BIPOP, P101, LONGTANG, and DANZHU. We assume that the arrival rate of each entry point obeys the negative exponential distribution [26]. To plug the average arrival rate that comes from flight historical statistics as expected value into distribution functions we can obtain the changes of traffic flux over time at each entry point as shown in Figure 6. It should be noted that we use equivalent traffic flow in this simulation and take each simulation time step as 1 minute.

It can be noticed that traffic flux changes significantly over time at each entry point. When aircraft keep going to convergent points, the peak of traffic wave may happen to meet another one, in which situation traffic density nearby will be too high to meet separation requirements and the probability of unsafe events will increase accordingly. Instead, the trough of traffic wave may also meet another trough that may lead to a low traffic density and a large aircraft interval at convergent points, thus reducing the time/space utilization of limited airspace. Based on the situation we designed a "valve". Agent aims at balancing traffic flow in the simulator system to manage the inflow/outflow of adjacent cells. In addition, this is much like the principle of "Cutting peak and filling valley" in real air traffic flow management [27, 28].

The basic strategy of arrival "valve" control is as follows: first determine whether the mean arrival aircraft interval is lower than the arrival ATC separation $d_i^{a\text{ATC}}$ in one cell at any time, which means whether the stock of aircraft in one cell exceeds. If yes then the excessive number of aircraft should operate with the assigned ATC velocity $v_i^{a\text{ATC}}$ and the rest do not change their speed, while if no then all aircraft can still follow the STAR. Since $v_i^{a\text{ATC}}$ is lower than normal velocity, the traffic flux will also be lower compared to no control situations. With outflow decreasing, the stock of aircraft in one cell will increase accordingly. At next time step inflow will be added to the original cell stock. If this total value still exceeds the standards, control measures should be taken like before. These cyclic steps keep going until sometime there comes a small inflow and the total value added with cell stock goes lower than the standards. In this circumstance all aircraft in the cell can operate with STAR and aircraft controlled before can exit cell normally. In the whole process, the total

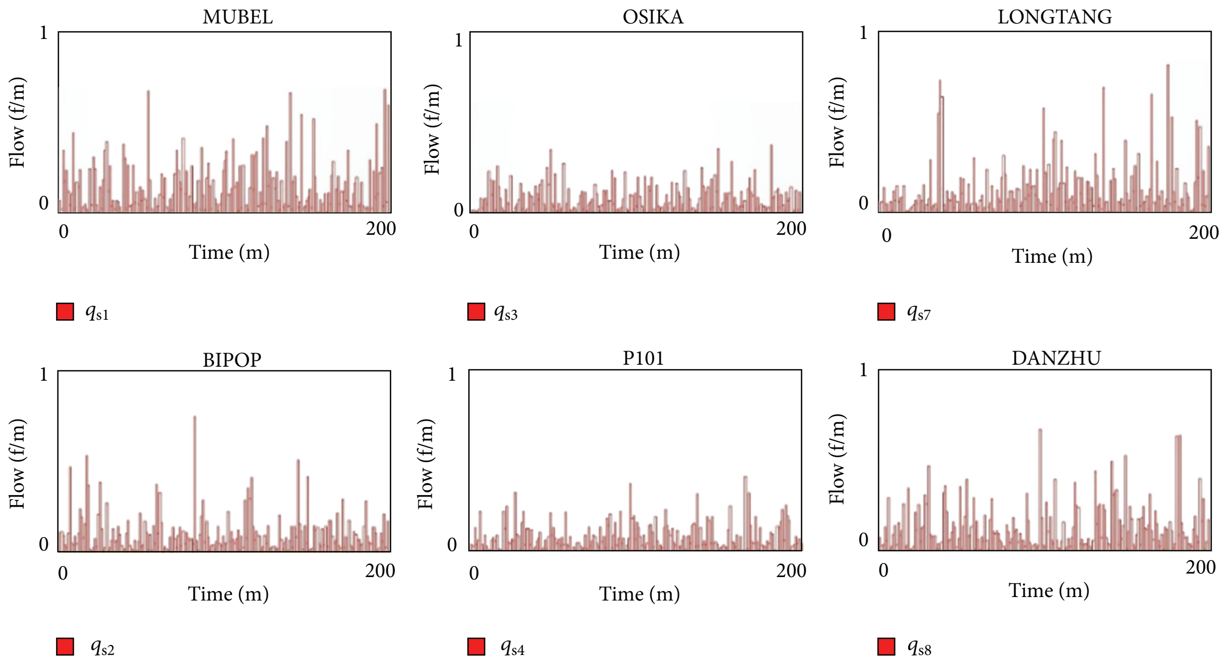


FIGURE 6: Changes of traffic flow over time at entry points.

number of aircraft in one cell cannot exceed the safety value at any time; that is, mean interval should never be smaller than the safety separation.

Take typical arrival segment "GYA-AGVOS (Segment5)" as an example and making a comparison of air traffic flux and stock in cells between before and after the "valve" control, it is easy to find that the inflow and outflow in this segment become smooth and steady when there is the "valve" control and the stock of aircraft in each cell always meets the safety requirement as shown in Figures 7(a) and 7(b). The unit is flights/minute.

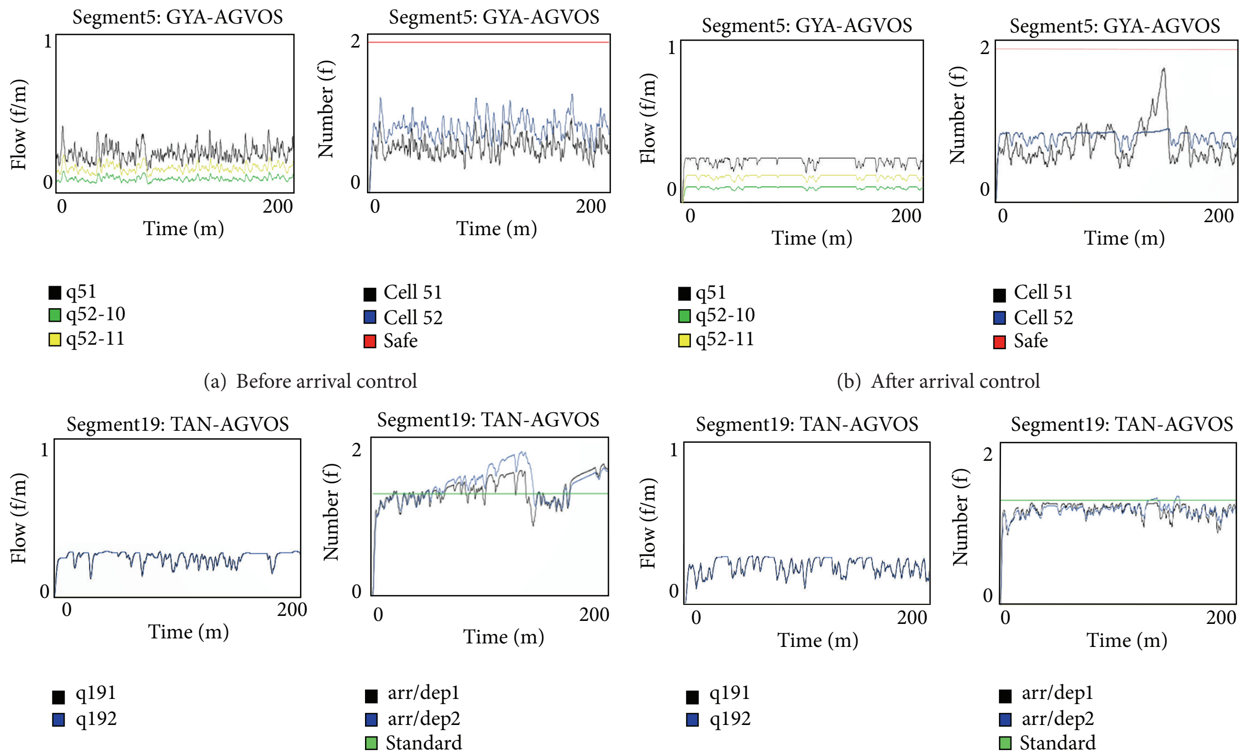
The departure in terminal area consists of two major parts: taking off from runway and flying along departure routes. We designed two successive processes in this simulation accordingly called the airport surface part and the air routes part, which are complementary in departure process. Specifically, the airport surface part mainly consists of departure flights schedule generation, runway time occupation, take-off slots allocation, and aircraft ground holding. We assume that the generation of departure flights schedule obeys a negative exponential distribution. The number of scheduled departure flights per unit time is as shown in Figure 8(a). The runway time occupation and take-off slots allocation are actually how the landing aircraft and taking-off aircraft make the most use of runway slot resources within the limited and dynamic runway operation capacity. Changes in the number of landing and taking-off aircraft per unit time are as shown in Figures 8(b) and 8(c). Since landing aircraft have the priority, taking-off aircraft that are unable to go by flight schedules will be postponed to take ground-holding processes. Changes in the number of ground holding aircraft over time are as shown in Figure 8(d).

In the process of flying along departure routes, there is no need to take control measures under the condition of full isolation between arrival and departure routes. The comparison of departure flux and stock over time before and after control measures is as shown in Figures 7(c) and 7(d). The basic strategy of departure "valve" control is as follows: first determine whether the mean arrival aircraft interval is lower than the ATC separation d_i^{ATC} in one cell at any time. If yes then all the departures in this cell should take control measures, which means passing the segment with constant ATC velocity u_i^{ATC}. If no then determine whether the mean interval of both arrivals and departures is lower than ATC separation d_i^{ATC} in one cell, and if yes departures must be adjusted to increase the entire mean interval until it goes above d_i^{ATC}.

4.3. Result Analysis. According to the export data derived from NetLogo simulation platform we got series of parameter scatter diagrams that could reflect the basic traffic flow characteristics through statistic and analysis, as shown in Figure 9. Based on this, we analyzed the mutual influence relationship among traffic flux q, density k, and velocity v of the arrival and departure routes in airport terminal area. Limited by space this paper only made detailed discussion on $q-k$ relation of arrivals and departures; meanwhile the rest of $v-k$ and $q-v$ relations were listed in the form of statistical diagrams for reference.

Taking typical arrival segment, Segment5, as an example, we derived the basic tendency of traffic flux and density relationship for arrival routes in terminal area as shown in Figure 9(a1). From the diagram we can find that the relationship tendency consists of three main stages.

Stage I is the free flow state in which the number of aircraft in segment is very low and the mean interval exceeds arrival ATC separation $d_5^{a\mathrm{ATC}}$, which means all the aircraft can follow STAR. The traffic flux is directly proportional to density in



(a) Before arrival control

(b) After arrival control

(c) Before departure control

(d) After departure control

Figure 7: Comparison of traffic flux and stock over time before and after control measures.

segment and the proportionality coefficient is equal to the mean velocity of segment. Stage II is the congestion flow state in which the number of aircraft in segment increases and then the mean interval is under arrival ATC separation d_5^{aATC}, which means part of the aircraft should be assigned to take control measures including decelerating, maneuvering or holding, and so forth. The relationship between traffic flux and density occurs a inflection point. The traffic flux in segment is still directly proportional to density; however the new proportionality coefficient is equal to the space mean value of the controlled aircraft velocity and uncontrolled aircraft velocity; that is,

$DN_5^a(\overline{N_5^{aATC}}t_5^s + \sum_{i=1}^{N_5^{aATC}} t_i)^{-1}$, see details in Section 2.1(1).

All of the control measures in the arrival process, including deceleration, maneuvering and holding, will lead to a decline in the displacement velocity. In the diagram it made the tendency of flux and density relationship to be leveling off. While the traffic flux will still increase with density, which differs from normal vehicle traffic since after inflection point the vehicle traffic flux decreases with density. The reason why this difference exists is that the car-following behavior between adjacent vehicles has significant influences when congestion occurs on the road while air traffic normally maintains a larger safe separation. Apart from this, some control measures as hold pattern make aircraft deviate from original air routes, which does not affect the other aircraft. Therefore, after the inflection point traffic flux will not decrease but increase with a low slope. Stage III is the block flow state in which the number of aircraft in segment exceeds safe value with the control adjustment in congestion state and large amount of traffic converging continuously. The mean interval is under the safe separation d_5^{safe} and insecurity factors surge that should be avoided as much as possible.

From the above theoretical derivation we can find that variables related to operational flight program including initial velocity v_i^I, acceleration a_i, and length Ω_i of segment (cell) become constant values as the STAR/SID are established. This paper focuses on the rest of the variables especially flow control variables including ATC separation d_i^{ATC} and ATC velocity v_i^{ATC}. Such variables will become order parameters that influence the mutual relationships among three basic air traffic flow parameters.

As shown in Figure 9(a2), when arrival ATC separation d_5^{aATC} rises to 25% the inflection point of traffic flux and density relationship moves forward and actually the horizontal axis value equals $1/d_5^{aATC}$. The rise of ATC standard will lead to a decrease of free flow state in Stage I. More aircraft need to take control measures; meanwhile the frequency of high traffic density even block flow in Stage III also increases. But in whatever Stage I or Stage III, the slope of traffic flux and density relationship maintains constant; that is, before inflection point the increase tendency coincides and after inflection point the increase tendency still keeps parallel. As shown in Figure 9(a3), when arrival ATC separation d_5^{aATC} remains the same and arrival ATC velocity v_5^{ATC} turns down, the inflection point stays the same but aircraft pass

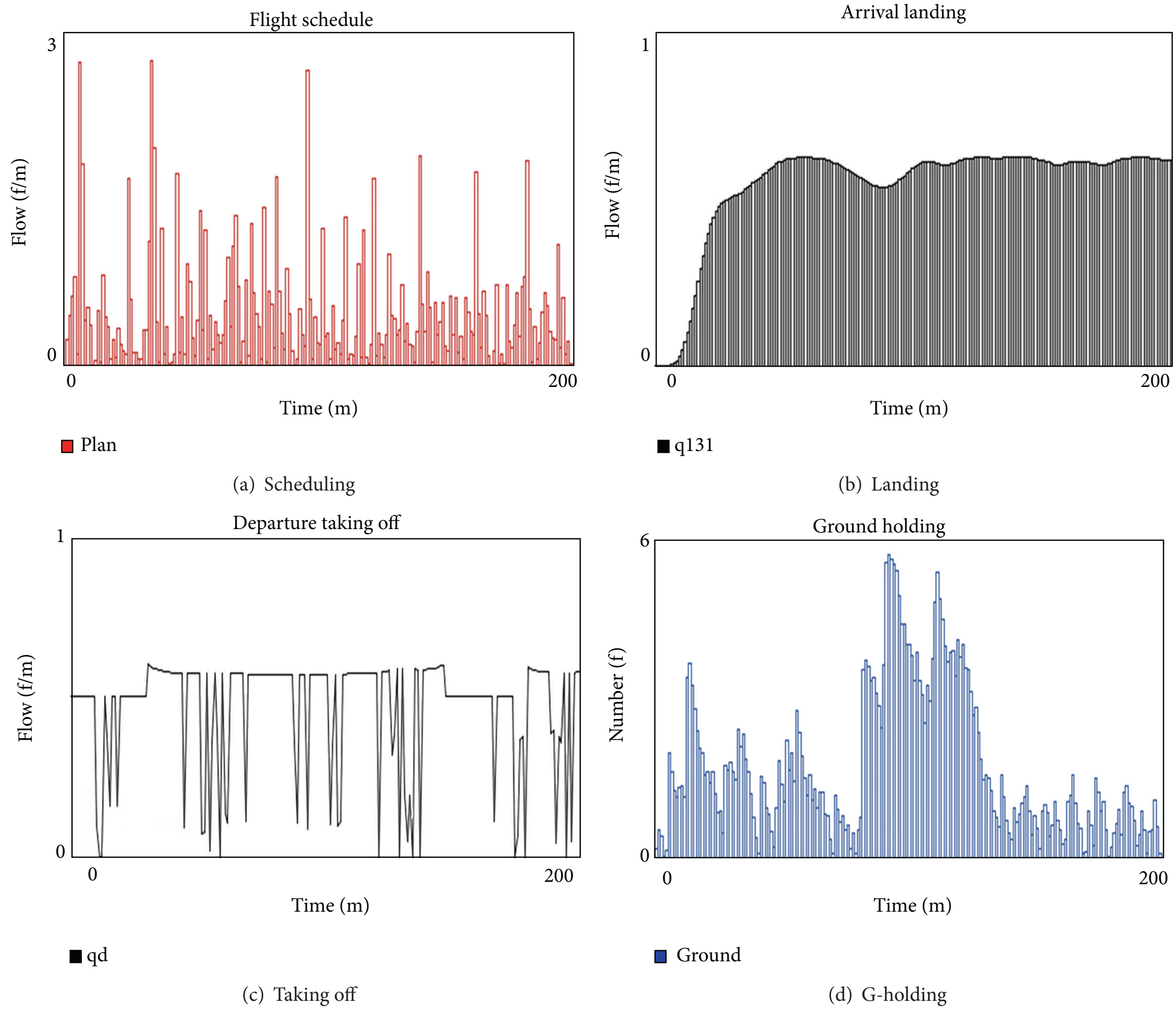
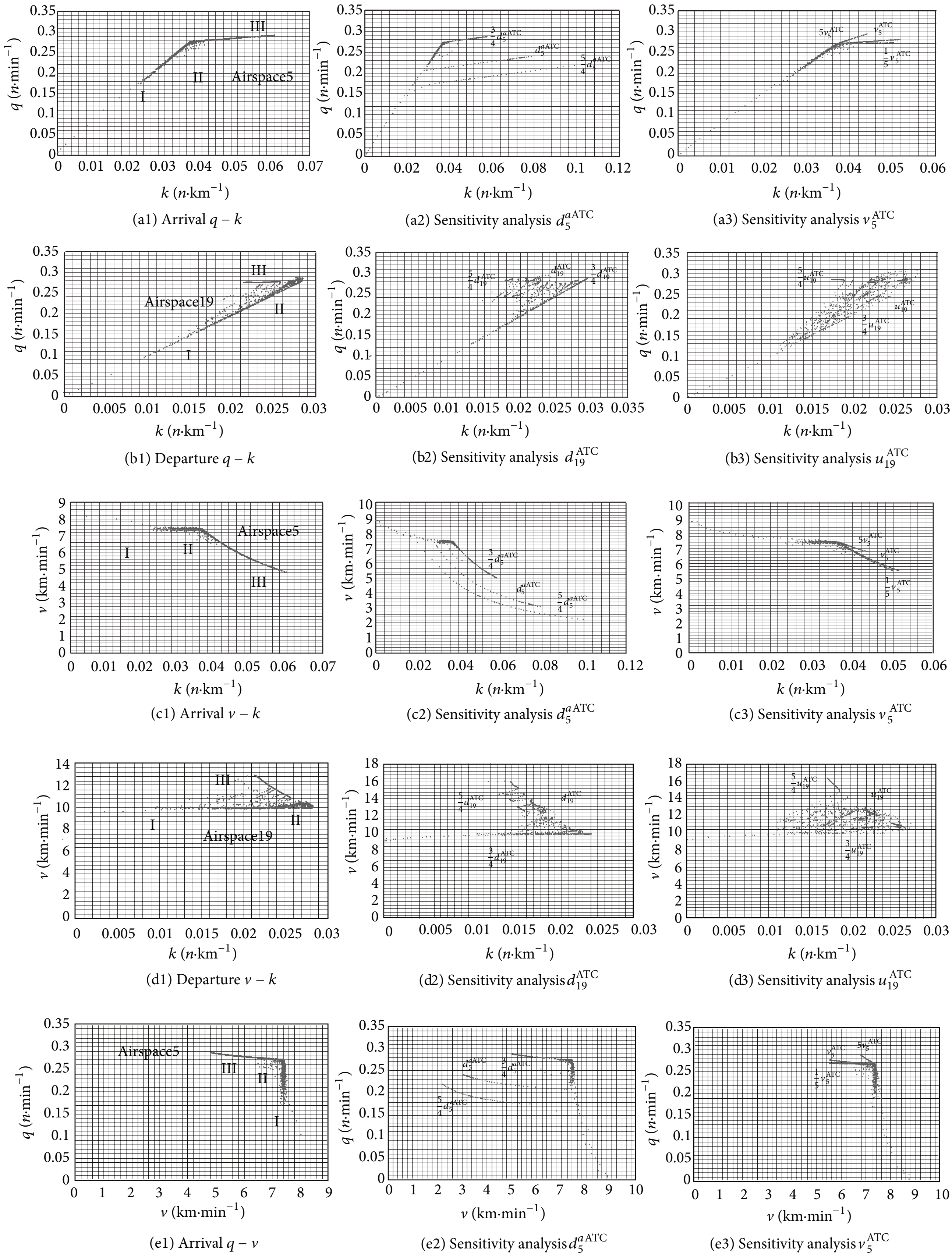


Figure 8: Changes of traffic flux and stock over time in airport surface part simulation.

the segment with a lower mean velocity after this point. Therefore, traffic flux has a decelerated growth with the density. Conversely, if it needs to let traffic flux reach the level before changing of ATC velocity v_5^{ATC}, the density of air segment should be higher, which may lead to an early arrival of the block flow state in Stage III. The free flow state in Stage I has no change now since it is not affected by control measures.

Taking typical departure segment, Segment19, as an example, we derived the basic tendency of traffic flux and density relationship for departure routes in terminal area as shown in Figure 9(b1). The relationship tendency also consists of three main stages. Similar to arrival parts, Stage I is the free flow state in which the traffic flux of departures is directly proportional to density in segment and the proportionality coefficient is equal to the mean velocity of segment from SID. Stage II is the congestion flow state which is unlike arrival parts, since in departure routes we focus on the entire mean interval of both arrivals and departures, not just departures. With the density of departures rising up in segment, the entire mean interval can still meet the ATC separation d_{19}^{ATC} if arrivals density is small enough. Then the tendency of traffic flux and density relationship lies on the extension line of Stage I. Stage III is the block flow state in which the number of arrivals and the number of departures increase simultaneously. Accordingly the entire mean interval becomes lower than the ATC separation d_{19}^{ATC}. Considering the principle of arrivals priority part of the departures will accelerate to leave the heavy-traffic segment in order to release the space resources for arrivals. The new slope of traffic flux and density relationship equals the departure ATC velocity u_{19}^{ATC}. The number of adjusted departures is determined by arrivals, which is shown as a series of scatter values between the tendency line of Stage III and extension line of Stage II. The vertical distance from tendency line of Stage III to the scatter values varies inversely to the density of arrivals, that is, $((1 - d_{19}^{\text{ATC}}\rho_{19}^{a}(t))/d_{19}^{\text{ATC}})K$ (K is constant coefficient); see details in (27).

As shown in Figure 9(b2), after raising up 25% of the ATC separation d_{19}^{ATC}, there appear more departure aircraft that need to take control measures. Thus the Stage III also appears



(a1) Arrival $q-k$

(a2) Sensitivity analysis $d_5^{a\mathrm{ATC}}$

(a3) Sensitivity analysis v_5^{ATC}

(b1) Departure $q-k$

(b2) Sensitivity analysis d_{19}^{ATC}

(b3) Sensitivity analysis u_{19}^{ATC}

(c1) Arrival $v-k$

(c2) Sensitivity analysis $d_5^{a\mathrm{ATC}}$

(c3) Sensitivity analysis v_5^{ATC}

(d1) Departure $v-k$

(d2) Sensitivity analysis d_{19}^{ATC}

(d3) Sensitivity analysis u_{19}^{ATC}

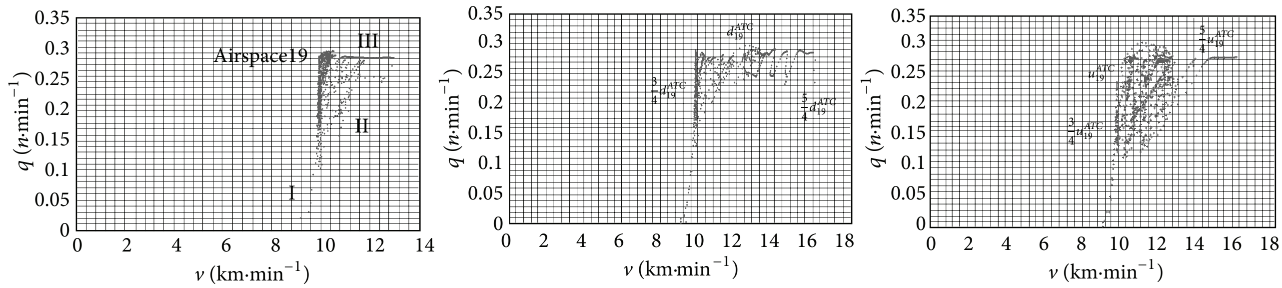
(e1) Arrival $q-v$

(e2) Sensitivity analysis $d_5^{a\mathrm{ATC}}$

(e3) Sensitivity analysis v_5^{ATC}

(a)

Figure 9: Continued.



(f1) Departure $q - v$

(f2) Sensitivity analysis d_{19}^{ATC}

(f3) Sensitivity analysis u_{19}^{ATC}

(b)

Figure 9: Scatter diagrams of basic arrival/departure flow characteristic parameters in terminal area.

earlier while the slopes of tendency lines in congestion and block flow states have no change. However, there are more scatter values between these two lines. The Stage II congestion flow state in which entire mean interval (high departure, low arrival) still exceeds the ATC separation will reduce its frequency-of-occurrence. More scatter values lie in the Stage III block flow state. Conversely reducing the ATC separation d_{19}^{ATC} departures may not need to be adjusted in most cases, which tend to be the Stage I free flow state. More scatter values of traffic flux and density lie more on the extension line of free flow state.

Keeping the ATC separation d_{19}^{ATC} unchanged and changing the departure ATC velocity u_{19}^{ATC} we can get new relationship tendencies as shown in Figure 9(b3). The Stage I free flow state keeps the same. But in Stage III block flow state the variation rate of traffic flux with density increases with the departure ATC velocity u_{19}^{ATC}, which means departure flow has raised its outflow rate of heavy-traffic segment. Meanwhile the impact of arrivals increases by $T\rho_{19}^{a}(t)u_{19}^{\mathrm{ATC}}$ (T is constant coefficient); see details in (27), which is shown as the scatter values spread a larger scope from center line of the free flow state extension direction.

5. Conclusions

Based on CTM we have proposed the macroscopic model of air traffic flow in airport terminal area and carried out a series of simulation experiments with the NetLogo platform. Through both of the theoretical and practical discussions, we could generally reveal the basic interrelationships and influential factors of air traffic flow characteristic parameters. The research findings are as follows.

(1) The CTM could accurately reflect the macroevolution laws of air traffic flow in terminal air route network. Meanwhile, it may also be applied to different kinds of air traffic scenes including airways, sectors, or airspaces by modifying conditions correspondingly. Moreover, it has a remarkable operational efficiency in multiagent simulations.

(2) There are obvious relationships among the three characteristic parameters as flux, density, and velocity of the air traffic flow in terminal area. And such relationships evolve distinctly with the flight procedures, control separations, and ATC strategies. The air traffic flow characteristics may take the specific changes through flight procedure optimization, control separation modification, or ATC strategy regulation, which could be part of the scientific basis for air traffic management in airport terminal area.

(3) The default parameters we used in simulation experiments are from practical ATC rules. Automatic optimization of these parameters for desired traffic flow characteristics should be taken into consideration in further studies. In addition, discussions in this paper focus on macro perspectives, thus the research results seem rough anyway. To obtain more detailed and elaborate traffic flow characteristics it is necessary to combine such macro studies with micro perspectives that give full expressions to the following: overflying or turning and so forth of individual behaviors and interactive effects. It also should be an important direction for further study.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research is supported by "the National Natural Science Foundation of China (NSFC) no. 61104159" and "the Fundamental Research Funds for the Central Universities no. NJ20130019."

References

[1] L. Li, R. Jiang, B. Jia, and X. M. Zhao, *Modern Traffic Flow Theory and Application*, vol. 1, Tsinghua University Press, Beijing, China, 2011.

[2] M. J. Lighthill and G. B. Whitham, "On kinematic waves. I. Floow movement in long rivers," *Proceedings of the Royal Society, London A*, vol. 229, pp. 281–316, 1955.

[3] M. J. Lighthill and G. B. Whitham, "On kinematic waves. II. A theory of traffic flow on long crowded roads," *Proceedings of the Royal Society: London A: Mathematical, Physical and Engineering Sciences*, vol. 229, pp. 317–345, 1955.

[4] O. Biham, A. A. Middleton, and D. Levine, "Self-organization and a dynamical transition in traffic-flow models," *Physical Review A*, vol. 46, no. 10, pp. R6124–R6127, 1992.

[5] C. F. Daganzo, "The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transportation Research B*, vol. 28, no. 4, pp. 269–287, 1994.

[6] C. F. Daganzo, "The cell transmission model, part II: network traffic," *Transportation Research B*, vol. 29, no. 2, pp. 79–93, 1995.

[7] C. F. Daganzo, "A variational formulation of kinematic waves: basic theory and complex boundary conditions," *Transportation Research Part B: Methodological*, vol. 39, no. 2, pp. 187–196, 2005.

[8] C. F. Daganzo, "On the variational theory of traffic flow: well-posedness, duality and applications," *Networks and Heterogeneous Media*, vol. 1, no. 4, pp. 601–619, 2006.

[9] P. K. Menon, G. D. Sweriduk, and K. D. Bilimoria, "Air traffic flow modeling, analysis and control," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference and Exhibit*, pp. 11–14, 2003.

[10] P. K. Menon, G. D. Sweriduk, and K. D. Bilimoria, "New approach for modeling, analysis, and control of air traffic flow," *Journal of Guidance, Control, and Dynamics*, vol. 27, no. 5, pp. 737–744, 2004.

[11] A. M. Bayen, R. L. Raffard, and C. J. Tomlin, "Adjoint-based control of a new Eulerian network model of air traffic flow," *IEEE Transactions on Control Systems Technology*, vol. 14, no. 5, pp. 804–818, 2006.

[12] A. Bayen, R. Raffard, and C. Tomlin, *Hybrid Control of PDE Driven Highway Networks*, Lecture Notes in Computer Science, Springer, New York, NY, USA, 2004.

[13] R. Raffard, S. L. Waslander, A. Bayen, and C. Tomlin, "A cooperative, distributed approach to multi-agent eulerian network control: application to air traffic management," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference and Exhibit*, 2005, AIAA 2005-6050.

[14] I. V. Laudeman, S. G. Shelden, and R. Branstrom, "Dynamic density: an air traffic management metric," NASA/TM219982112226, 1998.

[15] K. Lee, E. Feron, and A. Pritchett, "Air traffic complexity," in *Proceedings of the 44th Annual Allerton Conference*, pp. 1358–1363, 2006.

[16] Q. Liu, C. R. Bai, and Q. Lin, "Linear-quadratic optimal control of air traffic flow," *Computer and Communications*, vol. 26, no. 6, pp. 116–119, 2008.

[17] L. Wang, X. Zhang, and Z. Zhang, "Following phenomenon and air freeway flow model," *Journal of Southwest Jiaotong University*, vol. 47, no. 1, pp. 158–162, 2012.

[18] Z. N. Zhang and L. L. Wang, *Air Traffic Flow Management Theory and Method*, Science Press, Beijing, China, 2009.

[19] P. I. Richards, "Shock waves on the highway," *Operations Research*, vol. 4, no. 1, pp. 42–51, 1956.

[20] T. Vossen and M. Ball, "Optimization and mediated bartering models for ground delay programs," *Naval Research Logistics A: Journal Dedicated to Advances in Operations and Logistics Research*, vol. 53, no. 1, pp. 75–90, 2006.

[21] T. W. M. Vossen and M. O. Bal, "Slot trading opportunities in collaborative ground delay programs," *Transportation Science*, vol. 40, no. 1, pp. 29–43, 2006.

[22] T. Vossen, M. Ball, R. Hoffman, and M. Wambsganss, "A general approach to equity in traffic flow management and its application to mitigating exemption bias in ground delay programs," in *Proceedings of the 5th USA/Europe ATM R&D Seminar*, Budapest, Hungary, 2003.

[23] S. Tisue and U. Wilensky, "NetLogo: design and implementation of a multi-Agent modeling environment," in *Proceedings of the Agent Conference on Social Dynamics: Interaction, Reflexivity and Emergence*, Chicago, Ill, USA, 2004.

[24] W. Rand and U. Wilensky, "Visualization tools for agent-based modeling in NetLogo," in *Proceedings of Agent Chicago*, 2007.

[25] G. Z. Jiao, C. Guo, X. J. Hu, Y. Qin, and X. Ou, "Research on self-governing traffic model based on micro behavior information decision," *Control and Decision*, vol. 25, no. 1, pp. 64–68, 2010.

[26] G. C. Ovuworie, J. Darzentas, and R. C. McDowell, "Free moves, followers and others: a reconsideration of headway distributions," *Traffic Engineering and Control*, vol. 21, no. 8-9, pp. 425–428, 1980.

[27] D. Bertsimas, G. Lulli, and A. Odoni, "The air traffic flow management problem: an integer optimization approach," in *The ATFM Problem: An Integer Optimization Approach*, vol. 5035 of *Lecture Notes in Computer Sciences*, pp. 34–46, Springer, Berlin, Germany, 2008.

[28] D. Bertsimas and S. S. Patterson, "The air traffic flow management problem with enroute capacities," *Operations Research*, vol. 46, no. 3, pp. 406–422, 1998.

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 523451, 12 pages
http://dx.doi.org/10.1155/2014/523451

Research Article

Ananlysis of Carbon Emission Reduction and Power Dominance between Single Manufacturer and Single Retailer in Regulatory Cap and Trade System

Liang-jie Xia[1] and Hua-wei Zhi[2]

[1] *School of Business, Tianjin University of Finance and Economy, No. 25, Zhujiang Road, Tianjin 300222, China*
[2] *School of Management, Tianjin Polytechnic University, No. 399, Binshuixi Road, Tianjin 300387, China*

Correspondence should be addressed to Hua-wei Zhi; zhihw1983@163.com

Received 11 April 2014; Accepted 25 June 2014; Published 11 August 2014

Academic Editor: Xiang Li

Copyright © 2014 L.-j. Xia and H.-w. Zhi. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the cap and trade system, the paper analyses the Stackelberg game between the power asymmetrical retailer and manufacturer and designs a side-payment self-enforcing contract to resolve some arguments that the existing research overemphasizes spontaneity of participation in side-payment contracts design based on supply chain coordination and does not consider rationality and fairness of allocation of profit increment. Also, the numerical analysis was given. The research shows some important conclusions: in the supply chain, the dominant manufacturer is not able to encourage the retailer to improve its promotion level by increasing its carbon cutting level, but the optimal emission reduction level increases with the dominant retailer's promotion level; the optimal promotion level, emission reduction, and product demand in a retailer leading supply chain are higher than those in a supply chain dominated by manufacturer; with the new side-payment self-enforcing contract, decentralized decision according to individual rationality incurs a collective reason effect in the centralized setting.

1. Introduction

With consumer's low-carbon awareness improving, product's carbon emission affects the utility of consumer directly and is becoming one of the most important factors affecting the market demand. All carbon emissions are generated in the process of product manufacturing and delivering to meet the needs of the end consumer, and the consumer's purchase is the fundamental driving force for enterprises' emission cutting [1]. With the influence of awareness of environmental, consumers are willing to pay a higher price for environmentally friendly products [2].

In order to enhance the transparency around the carbon emission of products, carbon labels have become an important common choice of many enterprises to provide carbon related information of their products. And it makes it easy for consumers to compare the carbon footprints of different products and make right purchase decision [3]. At the same time, many international companies take measures to improve their images on carbon emission by publishing annual environmental and social responsibility reports. In a word, the impact of carbon emissions on the demand of product is becoming increasingly important; carbon cutting has become an important issue for enterprises' operational decision making.

Cap and trade system has become an important regulation adopted by many countries to achieve the aim of carbon emission reduction. In this system, emission allowance is necessary for company's production. Company should buy allowance to offset the quota gap from the carbon market if its emission volume is larger than its allowance allocated by government. In contrast, it can sell its surplus quotas to gain profit. So the cap and trade system will significantly affect company's production-related decision-making. However, there are only a few of papers dealing directly with the issue of carbon emissions and operations [4, 5].

At the same time, retailers have become the dominant participants in many industries [6], as a matter of fact, they determine the product manufacture and sale of the supply chain [7]. Taking full advantage of their "buyer power," they can pass on the operation risk and cost to the manufacturers by some unequal contracts. It is good for the retailers but really hurts the system profit of the supply chain and results in a deviation from Pareto optimality [8]. However, cooperation is also important for the competing enterprises in supply chain. They should cooperate to improve the market demand. For instance, the manufacturer engages in carbon cutting and the retailer focuses on marketing and promoting related to carbon cutting of the product. It is conducive to maximize their profits.

In this paper, considering the impact of carbon emission and promotion on market demand and the regulation of cap and trade on company operation, the game between manufacturer and retailer is analyzed. In the game, the manufacturer decides the carbon emission reduction of per unit of product to maximize its profit and the retailer maximizes its profit by adjusting the promotion levels related to emission cutting. Also, we compare the decision-making of company in supply chain with a dominant manufacturer with that in a supply chain dominated by a retailer. Finally, the paper introduces a new contract to improve the system profit of the supply chain, while the participants make decisions individually. With the new contract, both of the manufacturer and the retailer can gain more profit than before and the supply chain can gain as much system profit as that in centralized decision-making case. Also, the new contract is rational and fair in allocating the profit increment of the supply chain.

The rest of the paper is organized as follows. Section 2 gives the literature review of the supply chain operation and coordination in low-carbon environment. Section 3 sets forth the characteristics of the concerned issue, assumptions, and notations of this paper. In Section 4, we develop game theoretical model of two parties of the supply chain in cap and trade system and analyze the decision-making of the participants in the supply chain dominated by retailer or manufacturer. The centralized decision analysis will be carried out finally. In Section 5, we will show how to coordinate the supply chain with a side-payment self-enforcing contract. In Section 6, the numerical analysis will be shown. Finally, concluding remark and future research are given in Section 7.

2. Literature Review

Carbon footprint is the cumulative carbon emissions across the life cycle of product. Integrating the supply chain together to manage the carbon footprint can not only reduce carbon emissions, but also create financial value [9]. The global executives increasingly regard the environment problem including climate change and carbon regulation as one of the top concerns [10]. However, research on supply chain operations under a low-carbon environment is still in its infancy [5]. Existing research can be reviewed from the following three aspects.

Firstly, transportation mode selection is an important topic under carbon regulations. Researches discuss how to cut carbon emission by adjusting the transportation mode of supply chain. Carbon emissions arising from transportation account for a significant proportion of the supply chain carbon footprint [11]. Hoen et al. have made groundbreaking researches in this field. Using case studies, they discovered that in an established network, adjusting transportation model can reduce carbon emissions by 10% and increased total logistics costs by only 0.7% [12], but companies should make decision according to specific carbon regulations and practical issues [13].

Secondly, supply chain network design is also an important topic under carbon regulations. Switching logistic model is an effective solution for carbon reducing in established supply chain network. But the network is not always established. Also, companies in different regions or with different ability in carbon cutting are different in marginal cost and marginal profit of carbon cutting. Those differences provide chance for companies to cooperate in emission reduction and optimize carbon emission by redesigning their supply chain network. By this way, they can save logistics cost and enhance carbon cutting level at the same time [14].

Ramudhin et al. [15] introduced a mixed integer mathematical model formulation for green supply chain network design considering carbon trading regulation. They explored solution methodology to evaluate different strategic decisions alternatives and their impact in terms of carbon footprint. Different supply chain network may lead to different carbon intensity in the process of product distribution [16]. Diabat and Simchi-Levi [17] formulated a mixed integer programming model to help companies reveal an optimal strategy to meet their carbon cap and minimize opportunity cost. Supply chain network design would affect the cost of retailers and consumers at the same time, so both the consumers' cost and the retailers' emissions constraints should be considered in supply chain network design [18].

Life cycle assessment (LCA) has been proved as an effective method to reduce the harm to the environment in new products and processes design, but it has been rarely adopted in supply chain management [19]. Considering that, Chaabane et al. [20] integrated LCA into the process of sustainable supply chain designing under carbon trading regulation.

Thirdly, supply chain coordination is another important topic under carbon regulations. Because product carbon footprint is the total carbon emission across its life cycle from raw material used in its manufacturing to the disposal of the finished product, it is difficult for an enterprise to achieve the optimization of carbon footprint by itself, and it is essential for the enterprises in supply chain to cooperate in carbon cutting [21].

In fact, carbon emission can be greatly reduced without significant cost increase by optimizing operation decision and company cooperation across supply chain [4]. For instance, carbon emission in supply chain could be minimized by means of green procurement [22] and order quantity adjustment [23].

Decomposing it into processes of the supply chain is essential for optimizing carbon footprint. Caro et al. [24] introduced a simple but effective model to identify each company's responsibility for carbon emission reduction and provide measures for achieving carbon neutral in supply chain. They found that carbon emission must be allocated based on supply chain system in order to achieve optimal carbon emission level.

On supply chain coordination, Sunand Cao [25] introduced KMRW reputation model into the repeated games of incomplete information on carbon efficient products between retailers and a single manufacturer. Considering the "cap-and-trade" regulation, Du et al. [26] investigated companies' decision-making in the emission dependent supply chain in cap and trade system. Based on it, Du et al. [27] considered the emission cap of emission dependent manufacturer allocated by the government as a kind of environmental policy and formally investigated its influence on decision-making within the supply chain as well as distribution fairness in social welfare. Under carbon trading system, Xia et al. [28] took joint carbon emission reduction as a solution to optimize carbon cutting and improve the profit of the supply chain system. Also, Xia et al. [29] introduced side-payment self-enforcing contract to the emission reduction cooperation between the supplier and the manufacturer. Xie and Zhao [30]analyzed the carbon reduction cooperation mechanism in supply chain under the regulation of CDM. With system dynamics, Yang et al. [31] compared the impact of mandatory emission reduction and carbon tax on carbon cutting and the cost of supply chain members.

Xia and He (2014) analyzed the interactions and decision-making on carbon emission reduction and sales promotion in a dyadic supply chain considering consumers' low-carbon awareness in literature [32]. Although [32] and the present paper study similar games on emission reduction and promotion between single manufacturer and single retailer, they focus on essentially different issues. Firstly in [32], a nonlinear demand function is adopted, while in the present paper, in order to understand whether the results are sensitive to the choice of particular demand structure, a linear demand function is adopted. However, demand characteristics, liner or nonlinear, indeed capture market signals and subsequently affect almost all actions through the channel. This kind of significant difference can also be found in literature [33]. Secondly, the particular difference is that in [32] the regulatory carbon emission policy cap and trade are not yet incorporated as the background scenario for the channel running, which will substantially impacts on the inter-actions choosing as well as channel performance. For example, company's operation decisions are definitely constrained by cap and trade regulation. Furthermore, in [32] consumers' low-carbon awareness is considered but is not incorporated and combined with cap and trade regulatory policy. Compared with [32], the third significant difference is that in present paper we introduce the concept power dominance to research how the shift of power dominance between upstream supplier and downstream retailer influences individual interactions and channel performance in presence of cap and trade regulation and sensitiveness of market low-carbon awareness, which finally impact on both carbon emission reduction and marketing decisions.

3. Problem Characteristics, Notations, and Assumptions

In this paper, the supply chain consists of a manufacturer focusing on improving the consumer utility by carbon cutting and a retailer providing promotion for the low-carbon properties of the product.

Some parameters involved are given as follows.

Decision variables:

v: the retailer's promotion level, $0 < v < 1$;

e: the emission reduction of per unit product, $0 \le e < 1$.

Input and output parameters:

D: market demand of product;

a: market volume without considering the impact of emission reduction and sales promotion, $a > 0$;

b: the coefficient of promotion level's impact on demand, $b > 0$;

r: the coefficient of carbon emission's impact on demand, $r > 0$;

σ: the emission cap of per unit production, $\sigma > 0$;

p_c: the price of per unit of carbon allowance, $p_c > 0$;

u_r: the marginal cost of promotion, $u_r > 0$;

u_m: the marginal cost of emission reduction, $u_m > 0$;

ρ_m: manufacturer's profit of per unit product without considering carbon cutting cost, $\rho_m > 0$;

ρ_r: retailer's profit of per unit product without considering promotion cost, $\rho_r > 0$;

Π_m, Π_r, and Π: the total profit of manufacturer, retailer, and the supply chain system, respectively.

For convenience, this paper is based on the following assumptions. (1) The supply chain provides only one product; (2) the case of out of stock is not considered; (3) both of the retailer and manufacturer earn fixed profits per unit product (without considering the cost of promotion and emission reduction); (4) the retailer has no carbon emission; (5) the bid-ask spreads of carbon emission permit are ignored; (6) the manufacturer's initial emissions of per unit product are 1; (6) the relationship between marginal cost of emission reduction and the price of carbon trade satisfies $u_m > 2rp_c$.

Since the product carbon footprint affects the utility of consumer, carbon emission reduction can enhance product demand. The profit of manufacturing and selling environment-friendly products derives from the demand increment resulting from the improvement of consumer's environmental awareness [2]. In order to enhance product demand, large retail supermarkets often take action to encourage the manufacturers to improve their emission reduction level [34]. For example, Walmart has embraced its responsibility to protect the environment and has paid great attention to reduce emissions. Also, Walmart is profiting from its actions reducing greenhouse gas emissions in its own operations and its supply chain [35]. In addition to providing competitive product quality and price to consumers, carbon cutting and disclosing the information related to carbon emission have become important measures to improve market shares and consumer confidence [7].

Considering the price and emission reduction of product, the product demand can be described as $D = D_0 - kp + bv - r(1 - e)$. Price can be shown as $p = c_r + c_m + \rho_m + \rho_r$. c_m is the marginal production cost of manufacturer without considering emission reduction. c_r is the marginal cost of retailer without considering promotion. Consumers are willing to pay higher price for the environment-friendly products as the improving of their environmental awareness [36]. But the process of public awareness for environment protection enhancing is gradual. Also, the consumers' purchase decisions are constricted by their income. On the current, it is reasonable to suppose that the purpose of companies is to improve market demand by carbon cutting rather than by raising prices. It is similar to the "bonus pack" strategy adopted by companies. So, we assume ρ_m and ρ_r are constant. Let $a = D_0 - k(c + \rho_m + \rho_r) - r$; the product demand can be expressed as

$$D = a + bv + re. \tag{1}$$

Since the costs of emission reduction ($c(e)$) and promotion ($c(v)$) are convex in e and v respectively, and they can be assumed as $c(e) = u_m e^2/2$ and $c(v) = u_r v^2/2$, respectively.

Profit function of manufacturer and retailer can be expressed as following, respectively:

$$\begin{aligned} \Pi_m = {} & \rho_m (a + bv + re) - \frac{u_m e^2}{2} \\ & + (\sigma + e - 1) p_c (a + bv + re), \end{aligned} \tag{2}$$

$$\Pi_r = \rho_r (a + bv + re) - \frac{u_r v^2}{2}. \tag{3}$$

4. Analytical Model

In a supply chain with a dominant manufacturer or a dominant retailer, the manufacturer and retailer, in pursuit of their own maximal profits, make their optimal decisions on emission reduction and promotion related to carbon cutting, respectively, in Stackelberg game.

4.1. Stackelberg Game between the Retailer and the Dominant Manufacturer. In a supply chain with a dominant manufacturer, the manufacturer makes its optimal decisions on emission cutting. Then the retailer makes its optimal decisions on promotion when sufficient information about emission reduction is acquainted.

The optimal promotion level will be determined according to the retailer's profit-maximizing function which can be given by

$$\max_{0<v<1} \Pi_r = \rho_r (a + bv + re) - \frac{u_r v^2}{2}. \tag{4}$$

Since $\partial^2 \Pi_r / \partial v^2 = -u_r < 0$, the optimal promotion level can be expressed as

$$v^* = \frac{\rho_r b}{u_r}. \tag{5}$$

Substituting $v^* = \rho_r b / u_r$ into (2), the optimal carbon emission reduction per unit of production can be gained as

$$e^* = \frac{1}{u_m - 2rp_c} \left[r\rho_m + p_c a - rp_c (1 - \sigma) + \frac{p_c b^2 \rho_r}{u_r} \right]. \tag{6}$$

Since $\partial e^* / \partial p_c > 0$, $\partial e^* / \partial u_m < 0$, $\partial e^* / \partial \rho_r > 0$, $\partial e^* / \partial \rho_m > 0$, $\partial e^* / \partial u_r < 0$, $\partial v^* / \partial \rho_r > 0$, and $\partial v^* / \partial u_r < 0$, thus we can get Conclusion 1 shown as follows.

Conclusion 1. In the supply chain with a dominant manufacturer, (1) the optimal promotion level v^* increases with ρ_r and decreases as u_r increases, but it has nothing to do with ρ_m, p_c, and u_m; (2) the optimal emission reduction level e^* increases with ρ_m, ρ_r, and p_c and decreases as u_m and u_r increase; (3) the manufacturer is not able to influence the retailer's promotion level by adjusting emission reduction level.

Conclusion 1 shows that the dominant manufacturer is not able to encourage the retailer to improve its promotion level by increasing the volume of emission reduction. The reason is that the retailer can obtain more profit than before without improving promotion if the manufacturer increases emission reduction.

The higher ρ_m is, the more profit the manufacturer has to invest in emissions reduction. It results in more demand and more profit in turn. Even it is possible for the manufacturer to sell its surplus quotas to gain more profit by means of emission reduction. The higher u_m is, the higher the marginal cost of carbon cutting is, the lower the incentive for the manufacturer to reduce carbon emission is. On the contrary, the higher p_c is, the more the relative marginal carbon cutting cost is. So the manufacturer's emission reduction volume increases with p_c. It is similar to the fact that the optimal emission reduction level increases with ρ_r and decreases with u_r.

4.2. Stackelberg Game between the Manufacturer and the Dominant Retailer. In the supply chain with a dominant retailer, at first, the retailer makes its optimal decisions on promotion level related to emission reduction. Then the manufacturer makes its optimal decisions on emission reduction when sufficient information about promotion level is acquainted.

The optimal carbon cutting level will be determined according to the manufacturer's profit-maximizing function which can be given by

$$\begin{aligned} \max_{0\le e<1} \Pi_m = {} & \rho_m (a + bv + re) - \frac{u_m e^2}{2} \\ & + (\sigma + e - 1) p_c (a + bv + re). \end{aligned} \tag{7}$$

Since $u_m > 2rp_c$, Π_m is concave in e. Let $\partial \Pi_m / \partial e = 0$; the optimal emission reduction level is gained as

$$e_{rd}{}^* = \frac{r\rho_m + p_c (a + bv + r\sigma - r)}{u_m - 2rp_c}. \tag{8}$$

Substituting ${e_{rd}}^*$ into (3), let $\partial\Pi_r/\partial v = 0$; then the retailer's optimal promotion level can be gained as

$$ {v_{rd}}^* = \frac{b\rho_r}{u_r}\left(1+\frac{rp_c}{u_m-2rp_c}\right). \tag{9} $$

Then the optimal emission reduction is

$$ \begin{aligned} {e_{rd}}^* = {} & \frac{1}{u_r\left(u_m-2rp_c\right)} \\ & \times\left\{u_r\left[r\rho_m+ap_c+rp_c\left(\sigma-1\right)\right]\right. \\ & \left.+b^2\rho_r p_c\left[1+\frac{rp_c}{u_m-2rp_c}\right]\right\}. \end{aligned} \tag{10} $$

Since $\partial {e_{rd}}^*/\partial v > 0$, $\partial {v_{rd}}^*/\partial \rho_r > 0$, $\partial {v_{rd}}^*/\partial \rho_m = 0$, $\partial {v_{rd}}^*/\partial u_r < 0, \partial {v_{rd}}^*/\partial u_m < 0, \partial {v_{rd}}^*/\partial p_c > 0, \partial {e_{rd}}^*/\partial \rho_m > 0$, $\partial {e_{rd}}^*/\partial \rho_r > 0, \partial {e_{rd}}^*/\partial u_m < 0, \partial {e_{rd}}^*/\partial u_r < 0, \partial {e_{rd}}^*/\partial p_c > 0$, ${v_{rd}}^* > v^*$, and ${e_{rd}}^* > e^*$, thus we can get Conclusion 2 shown as follows.

Conclusion 2. In the supply chain with a dominant retailer, (1) the optimal emission reduction level increases with the retailer's promotion level, ρ_m, ρ_r, and p_c, and decreases with u_m and u_r; (2) the retailer's optimal promotion level increases with ρ_r and p_c and decreases with u_m and u_r, but it has nothing to do with ρ_m; (3) the optimal promotion and emission reduction level in the supply chain dominated by retailer are higher than those in the supply chain dominated by manufacturer.

The optimal emission reduction level increases with the retailer's promotion level. It means that the dominant retailer is able to encourage the manufacturer to improve emission reduction level by improving the promotion level. Since the retailer makes decision firstly. Its decision affects the product demand directly and affects the manufacturer's profit or cost gained by carbon trade. So the retailer's promotion affects the manufacturer's carbon cutting decision. It is an important difference insight compared to the supply chain dominated by manufacturer.

The optimal promotion and emission reduction level in the supply chain dominated by retailer are higher than those in the supply chain dominated by manufacturer. There are two reasons. On the one hand, the retailer is closer to market than the manufacturer. It can obtain market information with lower cost and understand consumer preference more clearly than manufacturer. On the other hand, as above, the manufacturer's emission reduction increases with the promotion level. The retailer can encourage the manufacturer to improve carbon cutting effectively. But the dominant manufacturer is not able to encourage the retailer to improve promotion level. So, the promotion and emission reduction level in the retailer leading supply chain are higher than those in the supply chain dominated by manufacturer.

4.3. Centralized Decision. Centralized decision can bring optimal profit for the supply chain system. But a binding agreement is necessary to encourage the manufacturer and the retailer to agree with centralized decision-making. The optimal carbon cutting and promotion level will be determined according to the system's profit-maximizing function which can be given as

$$ \begin{aligned} \max_{0\le e<1, 0<v<1} \Pi\left(v,e\right) = {} & \left[\rho_m+\rho_r+\left(\sigma+e-1\right)p_c\right] \\ & \times\left(a+bv+re\right)-\frac{1}{2}u_m e^2-\frac{1}{2}u_r v^2. \end{aligned} \tag{11} $$

If $-u_r(2rp_c - u_m) - b^2{p_c}^2 \geq 0$, the optimal carbon reduction and promotion level of the centralized decision supply chain can be expressed as

$$ \begin{aligned} e^{**} &= \frac{au_r p_c+\left(b^2p_c+ru_r\right)\left[\rho_m+\rho_r+\left(\sigma-1\right)p_c\right]}{u_m u_r-2ru_r p_c-b^2{p_c}^2}, \\ v^{**} &= \frac{b\left(u_m-rp_c\right)\left[\rho_m+\rho_r+\left(\sigma-1\right)p_c\right]+ab{p_c}^2}{u_m u_r-2ru_r p_c-b^2{p_c}^2}. \end{aligned} \tag{12} $$

5. Supply Chain Coordination Based on Side-Payment Contract Design

The decision (v^{**}, e^{**}) is the optimal decision of the centralized decision supply chain. It achieves the system' Pareto equilibrium point without considering the cost of resource allocation between the manufacturer and the retailer. It is also the benchmark of maximizing the outputs of the supply chain system.

Obviously, because of double marginal effect, it is impossible to achieve the optimal profit point in Stackelberg game without effective contract. In a centralized decision supply chain, the profit of the system can be maximized, and it is possible for the manufacturer and the retailer to gain more profit than in decentralized decision model. It provides impetus for the manufacturer and the retailer to cooperate. At the same time, participation constraint and incentive compatible constraint are two essential constraints for the manufacturer and the retailer to agree with centralized decision-making. The distribution of profit increment must be fair and reasonable from the perspective of both manufacturer and retailer.

Assume that the supply chain is dominated by the manufacturer. In the following section, the paper will analyze how to achieve Pareto Optimality with a side-payment self-executing contracts (SSEC) [37]. The new contract changes the payment function of both sides in the game and makes the supply chain system achieve the same optimal profit as in centralized decision-making model when the manufacturer and the retailer make decision according to individual rationality. At the same time, the profits of manufacturer and retailer satisfy the participation constraint and the incentive compatible constraint, and both of the players agree the deal is fair.

The side-payment self-executing contracts can be designed as $\widetilde{T}(v,e) = T(v,e) + g$. In order to conveniently research, $T(v,e)$ can be assumed as the linear function about v and e; that is, $T(v,e) = xv + ye$. x and y are nonzero constants. The parameter g is constant. In this case, the profit-maximizing functions of the manufacturer and the retailer can be expressed, respectively, as follows (the actual direction of the side-payment is determined according to the sign of $\widetilde{T}(v,e)$):

$$\begin{aligned}\max_{0\le e<1}\Pi'_m &= \Pi_m - \widetilde{T}(v,e)\\ &= \left[\rho_m + (\sigma + e - 1)\, p_c\right](a + bv + re)\\ &\quad - \frac{1}{2}u_m e^2 - T(v,e) - g,\\ \max_{0\le v<1}\Pi'_r &= \Pi_r + \widetilde{T}(v,e)\\ &= \rho_r (a + bv + re) - \frac{1}{2}u_r v^2\\ &\quad + T(v,e) + g.\end{aligned} \tag{13}$$

Since all of $\Pi_m(v,e)$, $\Pi_r(v,e)$, and $\widetilde{T}(v,e)$ are concave in v and e, both of $\Pi'_m(v,e)$ and $\Pi'_r(v,e)$ are concave in v and e. In this case, the manufacturer plays Stackelberg game with the retailer and decides its emission reduction to maximize its profit firstly. Then, the retailer makes decision about the promotion level to maximize its profit. The equilibrium of the game could be gained by backward induction as follows:

$$\begin{aligned}v' &= \frac{\rho_r b + x}{u_r},\\ e' &= \frac{1}{u_m - 2rp_c}\\ &\quad \times \left\{p_c\left(a + b\frac{\rho_r b + x}{u_r}\right) + r\left[\rho_m + (\sigma - 1)\, p_c\right] - y\right\}.\end{aligned} \tag{14}$$

Then, the requirement of the side-payment self-executing contracts can be expressed as follows.

5.1. Incentive Compatible Constraint. The game equilibrium in case of side-payment self-executing contracts is the same as that in case of centralized decision-making; that is, $(v', e') = (v^{**}, e^{**})$.

Under this constraint, since the optimal decisions in the two cases are the same, the profit of the system in the two cases is the same and the supply chain system is able to achieve Pareto Optimality.

5.2. Participation Constraint. Both of the manufacturer and the retailer are able to gain more profit in case of side-payment self-executing contracts implemented not less than that before this contract is adopted; that is, $\Pi'_m \ge \Pi_m{}^*$, $\Pi'_r \ge \Pi_r{}^*$.

The side-payment function is determined once x, y, and g are determined. According to the incentive compatible constraint, x and y can be given as follow:

$$\begin{aligned}x = &-\rho_r b + u_r\Big(\Big(bp_c{}^2(r\sigma - r + a)\\ &+ b(u_m - 2rp_c)\left[\rho_m + \rho_r + (\sigma - 1)\, p_c\right]\Big)\\ &\times \left(u_m u_r - 2rp_c u_r - b^2 p_c{}^2\right)^{-1}\Big),\\ y = &\; r\rho_m.\end{aligned} \tag{15}$$

Then, the side-payment function can be expressed as

$$\begin{aligned}\widetilde{T}(v,e) = &-\Big(\Big(bu_r p_c(\sigma - 1)(u_m - rp_c) + bp_c{}^2\left(au_r - b^2\rho_r\right)\\ &+ bu_r(u_m - 2rp_c)(2\rho_r + \rho_m)\Big)\\ &\times \left(u_m u_r - 2rp_c u_r - b^2 p_c{}^2\right)^{-1}\Big) v\\ &- r\left[(\sigma - 1)\, p_c + \rho_m\right] e + g.\end{aligned} \tag{16}$$

In the following section, we discuss how the value of g can be determined according to the participation constraint.

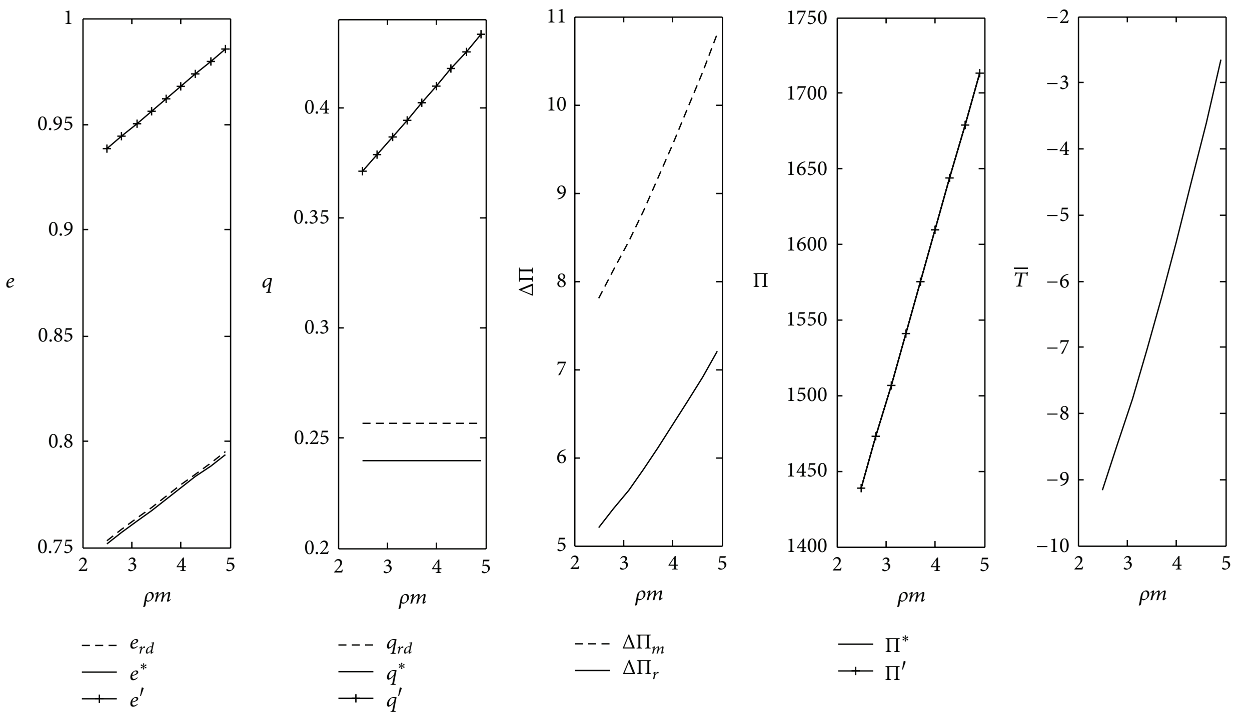
As pointed out in [32], Nash bargainingtheory is often used to resolve the problem of profit allocating in perfect information static game. In fact, Nash bargaining theory is fitful in the case of cooperative game. When Nash bargaining is adopted, a binding agreementis necessary for the dominantplayer would not prefer to cooperative game. So Nash bargaining solution is not easy to achieve. But in real market, transaction is frequently between the manufacturer and the retailer. It is similar to multi-period bargaining game. So we can adopt Rubinstein bargaining game to simulate the process of profit allocation.

Ruhinstein [38] has proved thata unique sub-game refining Nash equilibrium is existed as $(\delta_r{}^*, \delta_m{}^*) = ((1-\theta_r)/(1-\theta_r\theta_m), (\theta_r - \theta_r\theta_m)/(1-\theta_r\theta_m))$ in an indefinite duration alternating offers game. In which θ_r and θ_m is the discount factors of retailer and manufacturer respectively, δ_r and δ_m denote the share of profit increment of retailer and manufacturer in total profit increment because of cooperation respectively ($\delta_r + \delta_m = 1$). And the larger δ is, the more patient the player is, the greater impact it has on the game and less lost it will suffered because of time delay. In this paper,

$$\begin{aligned}\Delta\Pi_m &= \Pi'_m - \Pi_m{}^*\\ &= \Pi_m\left(v', e'\right) - T\left(v', e'\right) - g - \Pi_m\left(v^*, e^*\right)\\ &= \delta_m \Delta\Pi,\\ \Delta\Pi_r &= \Pi'_r - \Pi_r{}^*\\ &= \Pi_r\left(v', e'\right) + T\left(v', e'\right) + g - \Pi_r\left(v^*, e^*\right)\\ &= \delta_r \Delta\Pi,\\ \Delta\Pi &= \Delta\Pi_m + \Delta\Pi_r.\end{aligned} \tag{17}$$

Table 1: The value of some parameters.

	ρ_m	ρ_r	u_m	u_r	p_C
Figure 1	$2 \le \rho_m \le 5$	10	650	500	4
Figure 2	4	$6 \le \rho_r \le 10$	650	500	4
Figure 3	4	10	$650 \le u_m \le 800$	500	4
Figure 4	4	10	650	$400 \le u_r \le 600$	4
Figure 5	4	10	800	600	$3 \le p_c \le 5$



(a) The impact of ρ_m on e (b) The impact of ρ_m on q (c) The impact of ρ_m on $\Delta\Pi$ (d) The impact of ρ_m on Π (e) The impact of ρ_m on side-payment

Figure 1: The impact of ρ_m on decision variables and profits.

Based on the above three equations, the value of g can be determined by

$$g = \delta_r \left[\Pi_m\left(v', e'\right) - T\left(v', e'\right) - \Pi_m\left(v^*, e^*\right)\right] - \delta_m \left[\Pi_r\left(v', e'\right) + T\left(v', e'\right) - \Pi_r\left(v^*, e^*\right)\right]. \quad (18)$$

According to the result of game equilibrium above and (18), the value of g can be determined. Thus, the side-payment function of the new contract, $\widetilde{T}(q, e) = xq^2 + ye^2 + g$, can be identified.

6. Numerical Analysis

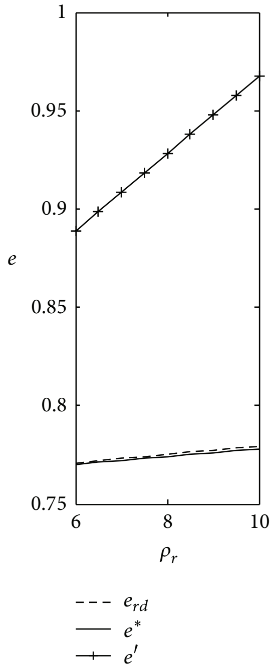
In the following section, assuming the manufacturer is the dominator, we analyze the impact of u_m, u_r, p_c, ρ_r, and ρ_m on the decision and profit of manufacturer and retailer and the profit of the supply chain system. Also, we illustrate the application of the side-payment self-executing contract. The similar analysis can be carried out in the supply chain dominated by retailer.

To ensure the existence of optimum conditions in all kinds of game, the related parameters are set as follows.

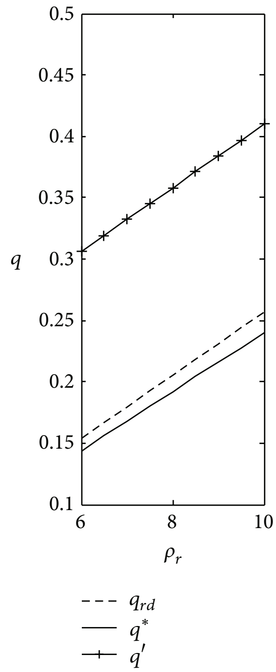
Let $a = 100$, $b = 12$, $\sigma = 0.8$, $r = 10$, $\delta_m = 0.6$, and $\delta_r = 0.4$. The values of u_m, u_r, p_c, ρ_r and ρ_m are shown as Table 1.

The output of numerical analysis, expressed from Figures 1 to 5, verifies Conclusions 1 and 2. At the same time, some important conclusions are gained as follows.

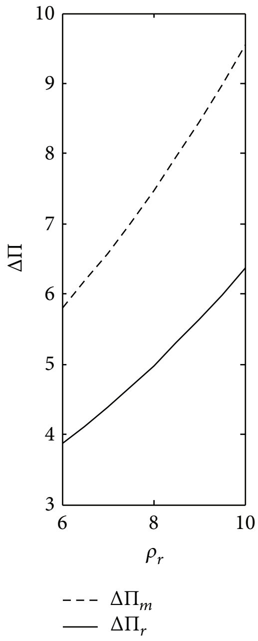
(1) In the case of implementing the side-payment self-executing contract, the dominant manufacturer's optimal emission reduction level increases with ρ_m, ρ_r, and p_c (shown as Figures 1(a), 2(a), and 5(a), resp.), but it decreases with increasing u_m and $u_r p_c$ (shown as Figures 3(a) and 4(a), resp.). It is the same as the discipline before the new contract was implemented. In the supply chain dominated by manufacturer, the main purpose of the contract is to encourage the retailer to improve the promotion level. It does not affect



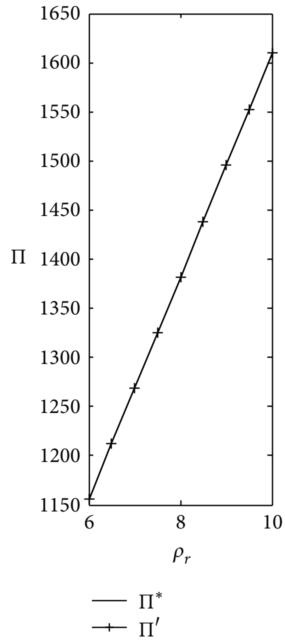
(a) The impact of ρ_r on e



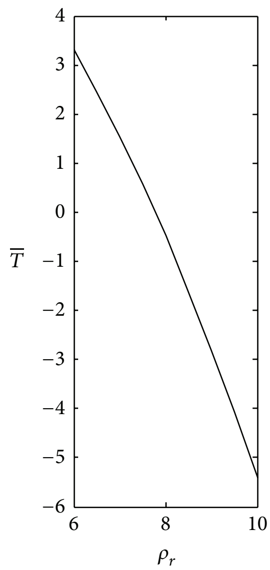
(b) The impact of ρ_r on q



(c) The impact of ρ_r on $\Delta\Pi$



(d) The impact of ρ_r on Π



(e) The impact of ρ_r on side-payment

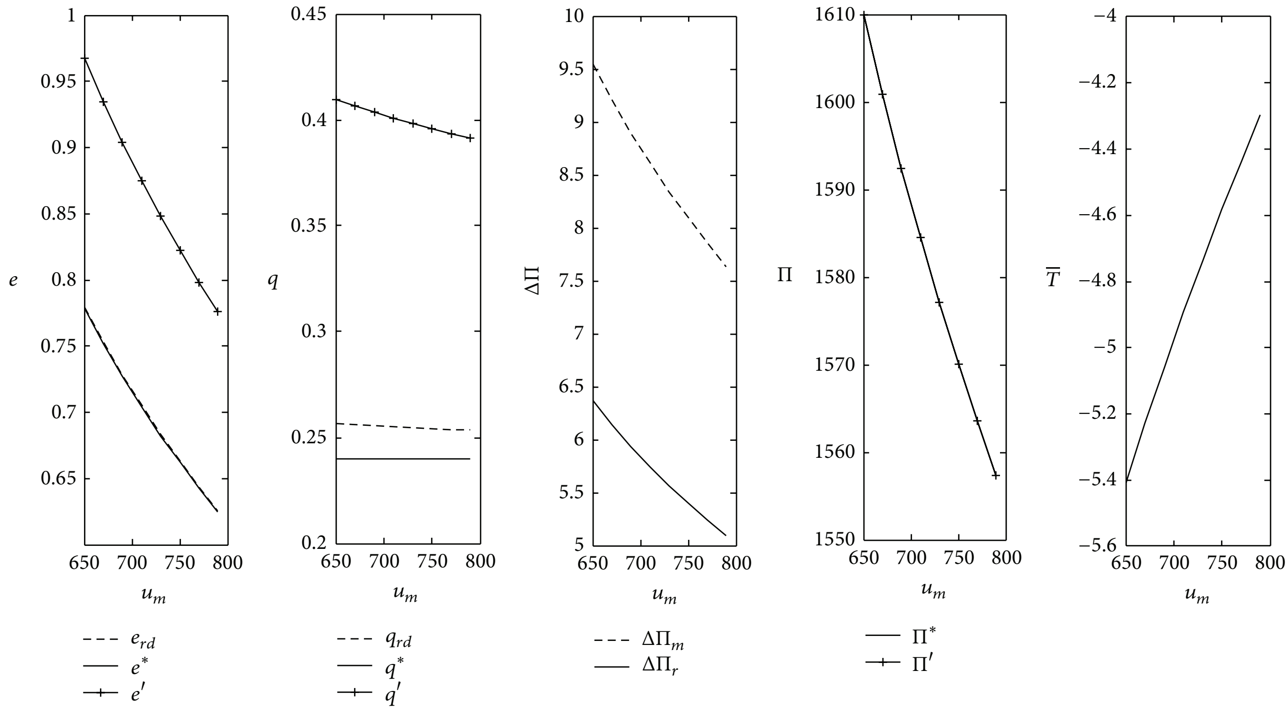
Figure 2: The impact of ρ_r on decision variables and profits.

the impact of ρ_m, ρ_r, p_c, u_m, and u_r on the manufacturer's decision-making.

(2) With the side-payment self-executing contract implemented, the optimal promotion level of the retailer, v', in the supply chain dominated by the manufacturer, increases with the increasing ρ_r, ρ_m, and p_c (shown as Figures 1(b), 2(b), and 5(b), resp.), but it is decreases with the increasing u_r and u_m (shown as Figures 4(b) and 3(b), resp.).

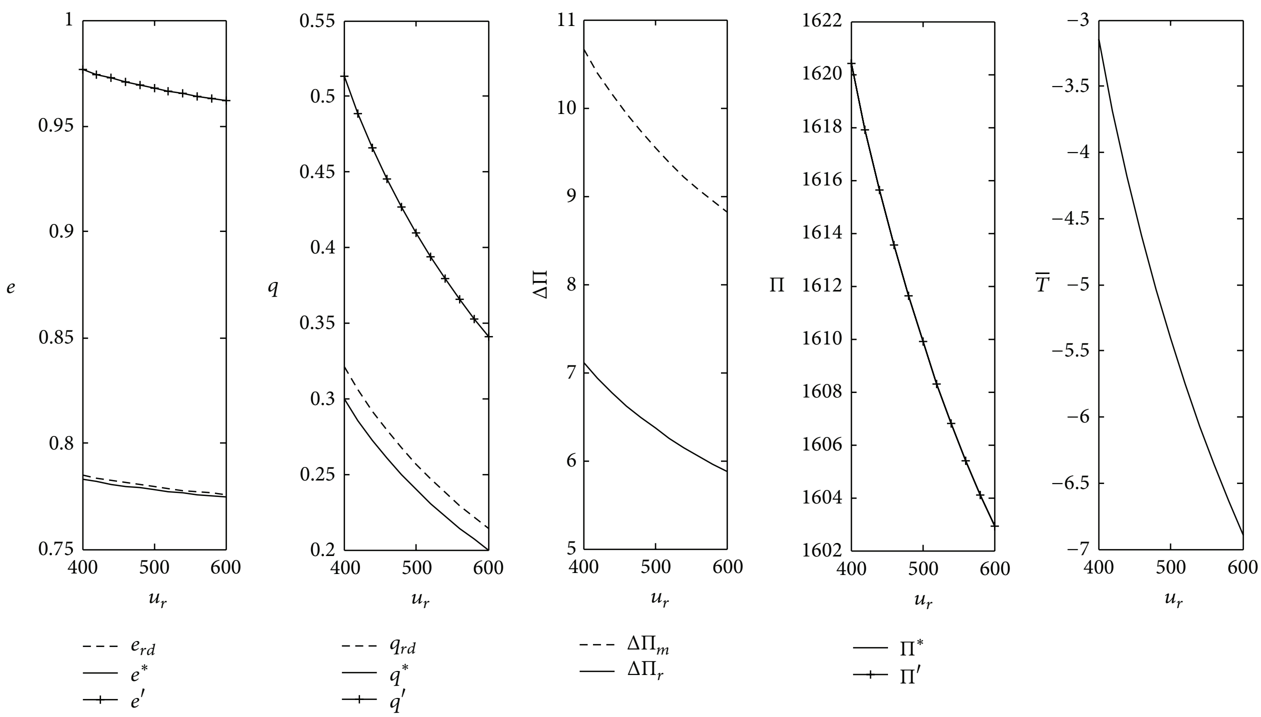
Similar to the analysis in Section 4.1, v' increases with ρ_r and decreases with u_r after the new contract is implemented. Before implementing the new contract, since the retailer can obtain more profit than the case without improving promotion if the manufacturer increases emission reduction, the dominant manufacturer is not able to encourage the retailer to improve its promotion level by increasing emission reduction. In this case, v' has nothing to do with ρ_m and u_m.

In the side-payment self-executing contract, $T(v,e) = xv + ye$ makes the profit increment associate with emission reduction and promotion level, and g makes the allocation of profit increment associate with the participant's discount



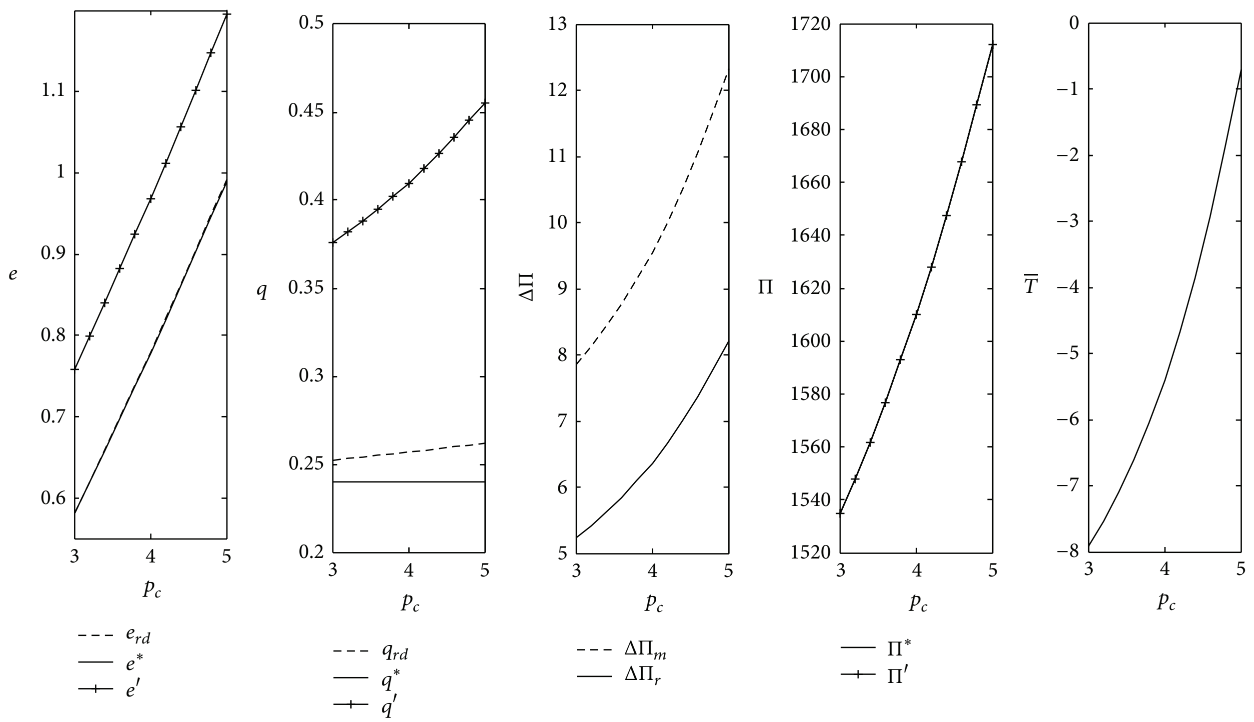
(a) The impact of u_m on e (b) The impact of u_m on q (c) The impact of u_m on $\Delta\Pi$ (d) The impact of u_m on Π (e) The impact of u_m on side-payment

FIGURE 3: The impact of u_m on decision variables and profits.



(a) The impact of u_r on e (b) The impact of u_r on q (c) The impact of u_r on $\Delta\Pi$ (d) The impact of u_r on Π (e) The impact of u_r on side-payment

FIGURE 4: The impact of u_r on decision variables and profits.



(a) The impact of p_c on e (b) The impact of p_c on q (c) The impact of p_c on $\Delta\Pi$ (d) The impact of p_c on Π (e) The impact of p_c on side-payment

Figure 5: The impact of p_c on decision variables and profits.

factors. Because of those, the allocation of incremental profit is fair, and the retailer's promotion level increases with emission reduction level. So the parameters influencing the manufacturer's decision-making affect the retailer's decision-making too.

(3) With the side-payment self-executing contract, the profit of supply chain system when the manufacturer and retailer make decision individually is equal to that in centralized decision-making case; that is, $\Pi' = \Pi^{**}$(shown as Figures 1(d), 2(d), 3(d), 4(d), and 5(d), resp.). And both the promotion level and emission reduction level are higher than before; that is, $e' > {e_{rd}}^* > e^*$ (shown as Figures 1(a), 2(a), 3(a), 4(a), and 5(a)) and $v' > {v_{rd}}^* > v^*$ (shown as Figures 1(b), 2(b), 3(b), 4(b), and 5(b)).

Since the side-payment self-executing contract satisfies both the participation constraint and the incentive compatible constraint, the incremental profit can be allocated fairly, and the contradiction between individuals and system is resolved when the participants make decision individually. Thus, $\Pi' = \Pi^{**}$. At the same time, the manufacturer is willing to improve emission reduction level and so is the retailer. Combined with the previous analysis, ${e_{rd}}^* > e^*$, and ${v_{rd}}^* > v^*$, we can gain the conclusion $e' > {e_{rd}}^* > e^*$ and $v' > {v_{rd}}^* > v^*$.

(4) With the side-payment self-executing contract, all the profit of the supply chain system, manufacturer, retailer, and the profit increment of the system increase with the increasing ρ_m, ρ_r and p_c (shown as Figures 1(c), 2(c), and 5(c), resp.), and decreases with the increasing u_m and u_r (shown as Figures 3(c) and 4(c), resp.).

According to the previous analysis, it is obvious that Π' increases with ρ_m, ρ_r, and p_c and decreases with u_m and u_r. In case that the new contract is executed, either of the manufacturer and the retailer improves its decision value which can improve the profit of the supply chain system and the profit increment will be allocated fairly between the participants. Thus, both of the participants' profits are improved, the increments $\Delta\Pi_m$ and $\Delta\Pi_r$ increase with ρ_m, ρ_r, and p_c and decrease with u_m and u_r.

(5) The sign of the side-payment function $(\widetilde{T})$ is uncertain, shown as Figures 1(e), 2(e), 3(e), 4(e), and 5(e).

Since the side-payment function $(\widetilde{T})$ is designed to encourage the participants to improve emission reduction and promotion level and to balance the allocation of profit increment, its sign is determined by the value of $T(v, e) = xv + ye$ and g.

7. Conclusions

The paper studies the Stackelberg game between a retailer and a manufacturer considering the cap and trade system and analyzes the impact of ρ_r, ρ_m, p_c, u_m, and u_r on the participants' decision-making. Some conclusions are gained: the dominant manufacturer is not able to encourage the retailer to improve its promotion level by increasing emission reduction, but the dominant retailer is able to encourage the manufacturer to increase its emission reduction by improving promotion level; the optimal emission reduction and promotion level in the retailer leading supply chain are higher than those in the supply chain dominated by manufacturer. Further, using the manufacturer leading supply chain for example, the paper designs a side-payment self-enforcing contract to maximize the system profit of supply chain and to make the optimal system profit in case of decentralized decision-making equal to that in case of centralized decision-making. Finally, a numerical analysis is given.

The paper considers the impact of carbon cutting and promotion related to carbon cutting on the product demand under the cap and trade system. The side-payment self-enforcing contract can resolve the arguments that the existing research overemphasizes on spontaneity of participation in side-payment contracts design based on supply chain coordination and does not consider rationality and fairness of allocation of profit increment. It is helpful for the company operation under low-carbon environment. The paper is based on a two-echelon supply chain with deterministic product demand. The research on supply chain coordination in case of supply chain network with stochastic product demand considering carbon regulations is the further research direction.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to express their sincere thanks to the anonymous referees and editors for their time and patience devoted to the review of this paper. This work is partially supported by NSFC Grants (nos. 71202086 and 71302115), the Scientific Research Development Fund of Tianjin University of Finance and Economy (no. Q130204), and the Training Program for Innovation Teams of Universities in Tianjin (no. TD12-5051).

References

[1] Carbon Trust, *The Carbon Emissions Generated in All That We Consume*, Carbon Trust, London, UK, 2006.

[2] Z. Liu, T. D. Anderson, and J. M. Cruz, "Consumer environmental awareness and competition in two-stage supply chains," *European Journal of Operational Research*, vol. 218, no. 3, pp. 602–613, 2012.

[3] M. A. Cohen and M. P. Vandenbergh, "The potential role of carbon labeling in a green economy," *Energy Economics*, vol. 34, no. 1, pp. S53–S63, 2012.

[4] S. Benjaafar, Y. Li, and M. Daskin, "Carbon footprint and the management of supply chains: insights from simple models," *IEEE Transactions on Automation Science and Engineering*, vol. 10, no. 1, pp. 99–116, 2013.

[5] J. Chen, "Study on supply chain management in low-carbon era," *Journal of Systems & Management*, vol. 21, no. 6, pp. 721–729, 2012.

[6] R. Inderst and C. Wey, "Buyer power and supplier incentives," *European Economic Review*, vol. 51, no. 3, pp. 647–667, 2007.

[7] T. Geylani, A. J. Dukes, and K. Srinivasan, "Strategic manufacturer response to a dominant retailer," *Marketing Science*, vol. 26, no. 2, pp. 164–178, 2007.

[8] J. Han, D. Zhao, and J. Qin, "Evolutionary game analysis of bertrand duopoly's behavior to their upstream suppliers," *Journal of Management Sciences*, vol. 22, no. 2, pp. 57–63, 2009.

[9] E. L. Plambeck, "Reducing greenhouse gas emissions through operations and supply chain management," *Energy Economics*, vol. 34, no. 1, pp. S64–S74, 2012.

[10] McKinsey, "How companies think about climate change: a McKinsey global survey," The McKinsey Quarterly, 2008.

[11] E. Masanet, K. J. Kramer, G. Homan et al., "Assessment of household carbon footprint reduction potentials," Tech. Rep., California Energy Commission, California, Calif, USA, 2008.

[12] K. M. R. Hoen, T. Tan, C. J. Fransoo et al., "Switching transport modes to meet voluntary carbon emission targets," *Transportation Science*, pp. 1–17, 2013.

[13] K. M. R. Hoen, T. Tan, J. C. Fransoo, and G. J. van Houtum, "Effect of carbon emission regulations on transport mode selection under stochastic demand," *Flexible Services and Manufacturing Journal*, vol. 26, no. 1-2, pp. 170–195, 2014.

[14] S. Elhedhli and R. Merrick, "Green supply chain network design to reduce carbon emissions," *Transportation Research D: Transport and Environment*, vol. 17, no. 5, pp. 370–379, 2012.

[15] A. Ramudhin, A. Chaabane, M. Kharoune, and M. Paquet, "Carbon market sensitive green supply chain network design," in *Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management (IEEM '08)*, pp. 1093–1097, Singapore, December 2008.

[16] S. Cholette and K. Venkat, "The energy and carbon intensity of wine distribution: a study of logistical options for delivering wine to consumers," *Journal of Cleaner Production*, vol. 17, no. 16, pp. 1401–1413, 2009.

[17] A. Diabat and D. Simchi-Levi, "A carbon-capped supply chain network problem," in *Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management (IEEM '09)*, pp. 523–527, December 2009.

[18] G. P. Cachon, "Supply chain design and the cost of greenhouse gas emissions," Working Paper, University of Pennsylvania, 2011.

[19] S. Seuring and M. Müller, "Core issues in sustainable supply chain management: a Delphi study," *Business Strategy and the Environment*, vol. 17, no. 8, pp. 455–466, 2008.

[20] A. Chaabane, A. Ramudhin, and M. Paquet, "Design of sustainable supply chains under the emission trading scheme," *International Journal of Production Economics*, vol. 135, no. 1, pp. 37–49, 2012.

[21] K. H. Lee, "Integrating carbon footprint into supply chain management: the case of Hyundai Motor Company (HMC) in the automobile industry," *Journal of Cleaner Production*, vol. 19, no. 11, pp. 1216–1223, 2011.

[22] T. Abdallah, A. Diabat, and D. Simchi-Levi, "A carbon sensitive supply chain network problem with green procurement," in *Proceedings of the 40th International Conference on Computers and Industrial Engineering (CIE '10)*, pp. 1–6, Awaji, Japan, July 2010.

[23] X. Chen, S. Benjaafar, and A. Elomri, "The carbon-constrained EOQ," *Operations Research Letters*, vol. 41, no. 2, pp. 172–179, 2013.

[24] F. Caro, C. Corbett, T. Tan, and R. Zuidwijk, "Carbon-optimal and carbon-neutral supply chains," Working Paper, Anderson Graduate School of Management-Decisions, Operations, and Technology Management, UC Los Angeles, 2011.

[25] F. Sun and J. Cao, "Research on cooperation and incentive mechanism in low-carbon supply chain based on KMRW model," *Research on Financial and Economic Issues*, no. 12, pp. 45–49, 2011.

[26] S. Du, F. Ma, Z. Fu et al., "Game-theoretic analysis for an emission-dependent supply chain in a "cap-and- trade" system," *Annals of Operations Research*, pp. 1–15, 2011.

[27] S. Du, L. Zhu, L. Liang, and F. Ma, "Emission-dependent supply chain and environment-policy-making in the "cap-and-trade" system," *Energy Policy*, vol. 57, pp. 61–67, 2013.

[28] L. Xia, D. Zhao, and Y. Li, "Game on jointly emission reduction based on transfer payment contract between the supplier and the manufacturer," *Systems Engineering*, vol. 31, no. 8, pp. 39–46, 2013.

[29] L. Xia, D. Zhao, L. He et al., "Game on carbon emission reduction and coordination based on side-payment self-enforcing contract between supplier and manufacturer," *Chinese Journal of Management*, vol. 11, no. 5, pp. 750–757, 2004.

[30] X. Xie and D. Zhao, "Research on cooperation strategy of enterprises' carbon emission reduction in low carbon supply chain," *Journal of Management Science*, vol. 26, no. 3, pp. 108–119, 2013.

[31] J. Yang, J. Li, and W. Lu, "The impact of emission policies on supply chain based on system dynamics," *Industrial Engineering and Management*, vol. 17, no. 4, pp. 21–30, 2012.

[32] L. Xia and L. He, "Game theoretic analysis of carbon emission reduction and sales promotion in dyadic supply chain in presence of consumers' low-carbon awareness," *Discrete Dynamics in Nature and Society*, vol. 2014, Article ID 837376, 13 pages, 2014.

[33] C. D. Ozgun, Y. Chen, and J. Li, "Channel coordination under fairness concerns and nonlinear demand," *European Journal of Operational Research*, vol. 207, no. 3, pp. 1321–1326, 2010.

[34] S. Jing, "The impact of green marketing on consumer psychology and behavior," *Management World*, vol. 5, no. 5, pp. 145–146, 2004.

[35] E. L. Plambeck and L. Denend, "The greening of Wal-Mart's supply chain," *Supply Chain Management Review*, vol. 11, no. 5, pp. 18–25, 2007.

[36] M. Laroche, J. Bergeron, and G. Barbaro-Forleo, "Targeting consumers who are willing to pay more for environmentally friendly products," *Journal of Consumer Marketing*, vol. 18, no. 6, pp. 503–520, 2001.

[37] L. He, D. Zhao, and Y. Liu, "Dynamic game of supply chain coordination based on automatically perform contract," *System Engineering Theory and Practice*, vol. 31, no. 10, pp. 1864–1878, 2011.

[38] A. Rubinstein, "Perfect equilibrium in a bargaining model," *Econometrica*, vol. 50, no. 1, pp. 97–109, 1982.

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 621341, 14 pages
http://dx.doi.org/10.1155/2014/621341

Research Article

Research on MAS-Based Supply Chain Resilience and Its Self-Organized Criticality

Liang Geng,[1,2] Renbin Xiao,[1] and Xing Xu[3]

[1] *School of Automation, Huazhong University of Science and Technology, Wuhan 430074, China*
[2] *School of Science, Hubei University of Technology, Wuhan 430068, China*
[3] *School of Mechanical and Automotive Engineering, Zhejiang University of Science and Technology, Hangzhou 310023, China*

Correspondence should be addressed to Renbin Xiao; rbxiao@hust.edu.cn

Received 11 April 2014; Revised 11 June 2014; Accepted 28 June 2014; Published 23 July 2014

Academic Editor: Weihua Liu

Copyright © 2014 Liang Geng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Building resilient supply chain is an effective way to deal with uncertain risks. First, by analyzing the self-organization of supply chain, the supply chain resilience is described as a macroscopic property that generates from self-organizing behavior of each enterprise on the microlevel. Second, a MAS-based supply chain resilience model is established and its local fitness function, neighborhood structure, and interaction rules that are applicable to supply chain system are designed through viewing the enterprise as an agent. Finally, with the help of a case, we find that there is an agglomeration effect and a SOC characteristic in supply chain and the evolution of supply chain is controlled by parameters of MAS. Managers can control the supply chain within the resilient range and choose a good balance between interest and risk by controlling enterprises' behavior.

1. Introduction

Although the supply chain has been improved in efficiency during the globalization of economy, it has also become riskier. A great deal of research has been conducted on supply chain flexibility, robustness, and vulnerability in recent years. The risk in different aspects of the supply chain is nevertheless unresolved. In order to overcome the vulnerability, the supply chain should possess comprehensive advantages in robust and flexible strategies. A concept of establishing a resilient supply chain is proposed. Supply chain resilience can be regarded as a comprehensive integration of the supply chain flexibility, robustness, and vulnerability.

Supply chain flexibility is a concept that focuses on the capability to adapt to the external changes. And it also reflects the adaptability of the system, which is a "change responding to change" strategy. Compared with the supply chain network resilience, the definition of flexibility is relatively narrower, which mainly refers to the ability to respond to the changes in customers' demand. This depends on the coordination of information flow, logistics, and capital flow between the upstream and downstream members of the supply chain.

The supply chain robustness is the ability to maintain operational function when the system is subject to uncertain interference from the internal operations and external events, which characterizes the robustness of the system, and is a rigid "maintaining the status quo" strategy [1]. Although the supply chain robustness cannot adapt, the supply chain flexibility can do it. This adaptive capacity refers to the ability to quickly restore its original state or a state that benefits the supply chain operations when an interruption occurs. In other words, robustness is the ability of anti-impact, while resilience is the ability to quickly recover from the impact.

Supply chain vulnerability reflects the maximum impact that the supply chain can resist under network interruption, which means the maximum deviation from the normal performance after an interruption occurs. The duration of interruption is not a significant factor of vulnerability, while the maximum negative impact of the interruption is significant. Vulnerability is one of the aspects that measure resilience.

The supply chain flexibility strategy and robustness strategy have their own strengths as well as limitations. Supply chain resilience, which contains adaptability without losing robustness, puts more emphasis on recovery capability.

The reason why supply chain has such strong competitive ability is owing to the fact that the supply chain is a virtual organization which has self-organizing ability and consists of enterprises with different functions and different types. In actual operation, every enterprise is an organization that possesses independent power to make decisions, and any enterprise cannot fully control the behavior of other enterprises. Besides, there is no internal centralized control within supply chain system. Each enterprise follows certain rules, fulfills its duties, and mutually coordinates with each other, and the system automatically forms an ordered structure, which reflects the self-organizing ability of the supply chain. In a word, the microlevel self-organization of supply chain is the root of reflecting the macrolevel supply chain resilience.

The way to make in-depth analysis of supply chain resilience is to start from microlevel self-organization of the supply chain and adopt bottom-up modeling method to study the interaction rules of each individual in supply chain system, which generates macrolevel supply chain resilience. From the system point of view, this paper studies the root of generating supply chain resilience and its properties. Through taking each enterprise as an agent and regarding supply chain local relationships as the environment, it builds multiagent systems that are similar to physical system and designs the unique local fitness function, neighborhood structure, and mutual interaction rules that supply chain owns. Then, it uses local information to guide the interactions between the agents as well as between the agents and the environment by agents' self-organizing ability, which ultimately generates supply chain resilience.

2. Literature Review

Resilience was first introduced by Holling [2], who proposed that "resilience determines the persistence of relationships within a system and is a measure of the ability of these systems to persist and absorb changes of state variables, driving variables, and parameters." Since then, the concept of resilience has been studied in many disciplines as shown in Table 1.

Table 1: Resilience studies in many disciplines.

Authors	Disciplines
Folke et al. [3]	Ecology
Adger [4]	Sociology
Bonanno [5]	Psychology
Callaway et al. [6]	Network theory
Starr et al. [7]	Risk management
Sheffi [8]	Supply chain

Resilience of supply chain was defined as the ability to return to its original state or an even better state after the disruptive event. Resilience of supply chain encompasses conducting preparation, response, and recovery activities for a disaster event.

There are some quantitative studies which attempt to create a framework for assessing supply chain resilience performance (see Table 2).

These literatures studied supply chain resilience from certain aspects. But there is no literature on the root of supply chain resilience and its characteristic, which is what we want to study.

3. Generation of Supply Chain Resilience

3.1. Supply Chain Resilience. Supply chain is an open system. As an economic system, supply chain system is essentially the result of communicating and integrating resources and ability between enterprises or between enterprises and the external environment. In the supply chain system, the distribution of resources, such as capital, technology, and brand, is different and unbalanced. Besides, the ownership and acquisition of resources, the obtaining of synergy theory and synergy results, the distribution of the interests between allies, and so on are also unbalanced. Competitive environment makes collaborative organizing activities of the supply chain members far from equilibrium. The nonlinear interaction of supply chain system determines its self-organizing evolution. Nonlinear interaction leads to the existence of both cooperation and competition between all members, and fluctuation of the individual is able to scale into the overall behavior. Due to random fluctuations, supply chain system becomes orderly and even forms new structure. For example, the consequences of information distortion and bullwhip effect lead to the fluctuations of manufacturers' state, and some minor changes in market factors may accumulate to fierce fluctuation of manufacturers' state, which may even make the whole system collapse. Thus, supply chain is a self-organizing system. Only when supply chain is in an open environment can it input the material, energy, and information and reduce the entropy of the system. Furthermore, only when supply chain stays far from equilibrium can it form intrasystem fluctuations and promote its spontaneous activity under nonlinear interactions. Such self-organizing ability gives the supply chain strong innovation ability and adaptability, which is the root of its strong competition.

On the microlevel, each enterprise in the supply chain is deemed as an agent. The agent can change with the environment and self-organize to make decision. Every time the environment factors change, it prompts each agent in the supply chain to change its own executing ability under the action of self-organizing ability and further changes its status in the supply chain. Such process goes on continuously, which forms the evolution process of supply chain system. The supply chain evolution process is in fact a self-organizing process, which generates supply chain resilience. In a self-organizing system, it forms the macroform or order from bottom to top through the "local rules" of each individual behavior, such property of the whole while part doesn't possess, is called "emergence." Supply chain resilience generates from self-organizing behavior of each member, as shown in Figure 1.

On the macrolevel, resilience management ability analyzes the effect on resilience from the following aspects as a whole: adaptation and coordination between node members,

Table 2: Quantitative studies on supply chain resilience.

Authors	Methods	Main works and findings
Priya Datta et al. [9]	Agent-based computational modelling	This literature studied complex multiproduct and multicountry supply chain subject to demand variability and production and distribution capacity constraints, which aimed to improve operational resilience.
Ratick et al. [10]	Linear programming	The literature used set cover location modeling to show that it is important to take into account potential exposure of facilities when designing supply chains.
Colicchia et al. [11]	Simulation applied to real scenario	This literature identified a set of approaches to manage risks to enhance supply chain resilience. Mitigation strategies do not influence lead-time variability but can reduce lead-time average, which will lead to resilience.
Zhao et al. [12]	The simulation analysis of military security network	This literature analyzed resilience of supply network topology structure under random attacks and attempted attacks from the aspects of availability, connectivity, accessibility, and so forth.
Geng et al. [13]	The simulation analysis of cluster supply chain	This literature analyzed dynamic evolution process based on cascading effect mode when cluster supply chain failure happens, which helped to illustrate that the root of vulnerability lies in cascading failure, while self-organization is the key to resilient recovery.
Pettit et al. [14]	Questionnaire survey	This literature developed a measurement tool titled the Supply Chain Resilience Assessment and Management. Critical linkages are uncovered between the inherent vulnerability factors and controllable capability factors.
Francis and Bekera [15]	The simulation analysis of electric power network	This literature proposed a framework which is focused on the achievement of three resilience capacities: adaptive capacity, absorptive capacity, and recoverability.

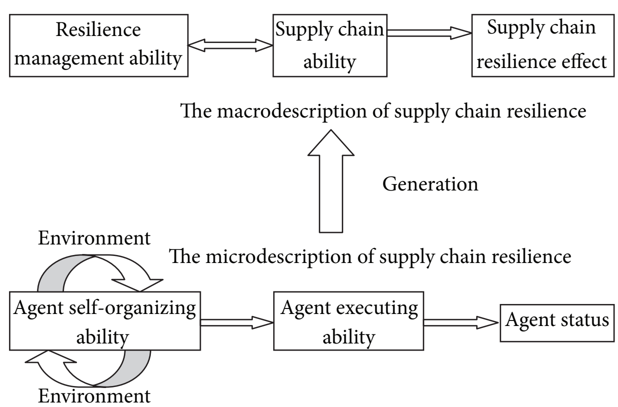


Figure 1: Micro- and macrodescription of supply chain resilience.

prediction of future events, recovery situation from disaster, and so forth. Reflected into the microlevel, resilience management ability is the self-organizing ability of each agent in supply network. Supply chain ability is an important ability that maintains the normal operation of entire supply chain network, which is the executing ability of each agent in supply network when reflected into the microlevel. Reflected into the microlevel, supply chain resilience effect is each agent's status in the supply network, which represents each enterprise's activity ability and network status, and is represented by its kinetic energy and potential energy. Micro- and macrocontrast of supply chain are shown in Table 3.

Resilience management ability has positive influence on supply chain resilience. When supply chain risk happens, some node members may not be able to survive together with the other members in the supply chain. At this time, it needs some enterprises with similar resources to join in; if the new enterprise can quickly integrate into the supply chain, then it can improve customer satisfaction and business performance of the supply chain, which means that it can improve supply chain resilience. If enterprises can predict future events well, they can take corresponding prevention measures to reduce the effects that risk brings, and even if the risk happens, they can take positive measures to restore normal state of supply chain and improve supply chain resilience. And if the recovery ability of node members is strong enough, the supply chain can meet normal operation conditions of enterprises as soon as possible and decrease stock-out rate and then improve supply chain network resilience. In addition, after the risk happens, if enterprises can coordinate well, they can ride out the storm together, maintain the normal operation of supply chain, improve customer satisfaction, ensure enterprise performance, and improve supply chain resilience.

Supply chain ability has positive impact on supply chain resilience and resilience management ability. The stronger the supply chain ability is, the stronger the ability of supply chain node enterprises to deal with various emergencies is. Even if supply chain risk happens, supply chain can recover quickly, and supply chain resilience will be strengthened. Thus it can be seen that supply chain ability is a key factor that

TABLE 3: Micro- and macrocontrast of supply chain.

Macrocontrast		Microcontrast	
Resilience management ability	Adaptability, predictive ability, recovery ability, coordinating ability	Agent self-organizing ability	Learn from each other, exchange experiences, communicate and coordinate, share risk, get warning information as early as possible, reduce unnecessary loss, and recover function quickly.
Supply chain ability	Information management ability, logistics management ability, resource support ability, target combination ability, fund procurement ability, risk management ability, reconstruction transformation ability, study and absorptive ability	Agent executing capability	Collect, analyze, and sort the information and finally obtain effective information. Deal with logistics facilities establishment and location and provide optimized logistics protection. Find node members that fit for their own development and make effective combination to reach the company target. Complete fund procurement and maintain the normal operation of enterprises after the occurrence of risk. Find new resources to meet the changing market. Improve the defects in the supply chain and build a new supply chain network.
Supply chain resilience effect	Wavy curve asymptotically stable	Agent status	Activity ability, network status.

influences supply chain resilience, and flexible use of various supply chain abilities can improve the ability of enterprises to face risks, improve supply chain resilience, and then improve enterprise performance. Under the effect of resilience management ability that formed by self-organization, supply chain ability can manage various risks effectively and help enterprises survive through a variety of difficulties [16].

Supply chain resilience effect is an asymptotically stable wave curve. It changes with time, showing the process that successively reflects response delay, reflects destruction spread, reflects destruction recovery, and eventually achieves asymptotically stable state. Such description of supply chain resilience effect not only clearly lays out the self-organizing evolution of supply chain resilience that changes with the time, but also meets several major features that supply network resilience should have, such as self-recovery, antidisruption, and active response. Besides, it can evaluate absorption ability, adaptability, and recovery ability of the supply chain. For details, please see our previous study in [13].

Definition of Resilience. The supply chain resilience is defined as follows: when dealing with unexpected interrupts, supply chain system restores a new stable state through microconstant evolution of enterprises' self-organizing ability; such emergence on the macrolevel is shown as supply chain resilience. Supply chain system with resilience can eliminate the influence brought by fluctuations, release energy, and then slow down the occurrence of large-scale collapse on the whole.

3.2. Self-Organized Criticality of Supply Chain Resilience. An important concept related to the self-organization is self-organized criticality (SOC) [17–19]. The SOC theory refers to an open and dynamic complex system that is far from equilibrium state and consists of multiple cells, which can evolve to a critical state through self-organizing process. When such system stays in the critical state, a small local disturbance may be amplified by mechanism similar to "domino effect," and the effect may be extended to the entire system, forming a large avalanche. System in the critical state will bring about "avalanche" events of various sizes, and both the avalanche's time dimension and its space dimension obey power-law distribution. The power-law distribution is an important characteristic of self-organized criticality, which also owns scale-free property, openness, robustness, and other characteristics.

Some literatures show that supply chain network has scale-free, high integration, and other characteristics. The average outgoing and incoming degree distribution of supply chain network under steady state obeyed power-law distribution [20–22].

This paper argues that supply chain resilience possesses self-organized criticality. The process that supply chain evolves towards the critical state is not affected by any external factors. Owing to the dynamic interactions between individual elements in the system, the critical state can be set up and such critical state is self-organizing. Within the critical point of supply chain resilience, core enterprises can control the changes of system external parameters at proper time based on understanding and grasping the operation mechanism and laws, so they can constrain system's operation process, making the system produce small internal fluctuations. Such gradually varied internal fluctuation process of the system could eliminate conflicts within the system, making the system stay in balance, which means that the system possesses resilience. At the critical point of supply chain resilience, it may be amplified into a huge fluctuation through nonlinear interactions and cascading effects within the system, leading to a rapid collapse after the system loses its balance.

This paper will expound such a view as follows. The microlevel interactions between enterprises in the supply chain result in the macrolevel self-organized criticality generation of supply chain resilience. If the supply chain is within the critical value, small disturbance will only lead to tiny changes and system will not meet any great disaster. On the contrary, if the environment variables reach the critical value, small disturbance may lead to various scales of avalanche.

4. Supply Chain Resilience Control Model Based on MAS

The paper is intended to build supply chain MAS system and interpret and evaluate supply network resilience through self-organizing ability of its intelligent agents. In this section, we will define the agent, environment, local fitness function, interactive strategy, and reaction behavior between the agents of MAS, build the MAS system of supply network, put forward supply network resilience index, and study the characteristics of resilient generation from the point of self-organization.

4.1. Basic Concepts of MAS. An agent is able to survive and take part in activities in the environment, which can sense its local environment and has certain goal as well as state variable of reaction behavior. MAS system is a system that achieves a certain goal by the interactions and effects between many internal agents. MAS system consists of computing entities that possess autonomy, reactivity, initiative, and social attributes. By microscopic or local (i.e., agents and the interactions between them) behavior or mechanism, MAS desires to obtain necessary macroglobal (i.e., the entire multiagent system) function [23, 24].

The definition of MAS system is as follows.

Definition (MAS system). An agent system is a system that contains the following elements:

(1) the environment E, in which the agents survive and conduct activities;

(2) a set of agents, $A = (a_1, a_2, \ldots, a_n)$;

(3) a set of reaction rules, which controls the interaction behavior between agents as well as between agents and the environment and is the law of the agent society.

Supply chain individuals own self-organizing properties, including intelligence and adaptability, follow certain rules, and adjust their states and behavior according to the environment and received information, making the system show a higher level of order as a whole. In other words, the supply chain is a MAS system. Supply chain system is based on the idea of decentralized control and each individual makes corresponding decision-making behavior in the current neighborhood according to the local information and local fitness function. The overall behavior of the supply chain generates through competition, cooperation, and other local behaviors between individuals.

The study on resilient self-organizing emergence of supply chain MAS is faced with 4 key problems as follows [25].

(1) The first problem is how to reasonably map the supply chain system into a MAS system, namely, the definition of the agent.

(2) The second problem is how to define the local fitness function. The local fitness function is based on the idea of distributed control, using limited local information to evaluate the state of a single agent and guide the overall evolution. Although the definition of local fitness function is not unique, reasonable local fitness function is the key to the effect of self-organization emergence.

(3) The third problem is how to define the neighborhood. The definition of neighborhood is a very important concept for optimizing the neighborhood, which guides how to generate a new state from the current state. Furthermore, how to choose a reasonable neighborhood structure is an important factor to the self-organizing optimization.

(4) The fourth problem is how to define the interaction rules between agents. MAS system achieves overall self-organization evolution through competition and cooperation between the agents. The reasonable and effective interaction rules between the agents can quickly lead the system to generate stable structure.

4.2. Mapping the Supply Network into MAS. To map the supply network into MAS system, now we define the following concepts.

4.2.1. Agent. Take each enterprise in the supply chain as an agent [26] and the specific descriptions of the agents in supply chain MAS are shown in Figure 2.

A supply network system that consists of N agents is given and represented as $S = \langle X, P, V, A\rangle$, where $X = (x_1, x_2, \ldots, x_i, \ldots, x_N)$, $x_i \in R^n$, is the position vector of agent i, which represents the status of this member in supply chain; $P = \{c_1, c_2, c_3, c_4, c_5, d, r, A_{\max}, V_{\max}\}$ is the control parameter set of the agents; $V = (v_1, v_2, \ldots, v_i, \ldots, v_N)$, $v_i \in R^n$, is the velocity vector of agent i, which represents the executing ability of the member; and $A = (a_1, a_2, \ldots, a_i, \ldots, a_N)$, $a_i \in R^n$, is the acceleration vector, which represents the self-organizing ability of the member. In other words, we assume that each member owns such attributes as self-organizing ability, executing ability, and network status, as shown in Figure 3. The executing ability of a member is its self-organizing ability multiplied by time, and the status of this member is its executing ability multiplied by time. The motion of an agent is controlled by the following equation:

$$\dot{x}_i = v_i, \qquad \dot{v}_i = a_i. \tag{1}$$

4.2.2. Local Fitness Function. Assume that the decision of each enterprise in supply chain is determined by the enterprise status V_1 in the neighborhood, the core enterprise of the overall supply chain has status V_2, the executing ability of the enterprise within neighborhood is V_3, the core enterprise within the neighborhood has status V_4, and random perturbation is V_5. Affected by local influence, self-organizing

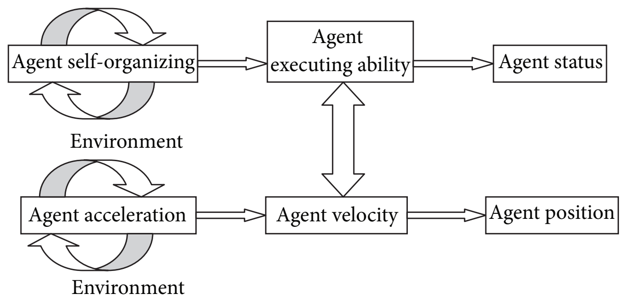


Figure 2: MAS mapping of supply chain.

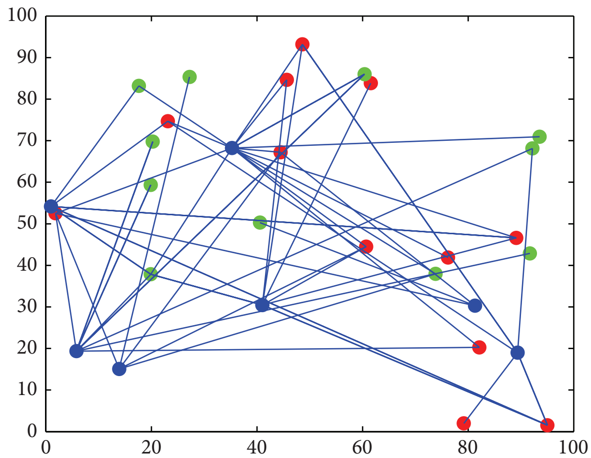


Figure 3: The initial supply network of FAW Xiali.

ability and executing ability of each member in the supply chain are changing with the time, and self-organizing ability causes changes in executing ability. In the model, local fitness function of the agent represents the agent's self-organizing ability at each time step, which means the instantaneous acceleration vector of the agent.

The local fitness function, namely, the calculating formula of instantaneous acceleration, is as follows:

$$V = c_1 V_1(d) + c_2 V_2 + c_3 V_3 + c_4 V_4 + c_5 V_5, \quad A = A_{\max} \frac{V}{|V|}. \tag{2}$$

V_1 is a vector that points to the neighbors within a distance of d; V_2 is a vector describing the center of the simulated world; V_3 is the average of the neighbors' velocities; V_4 is a vector that points toward the center of gravity of the neighbors; V_5 is a random unit-length vector; V is the summation of behavioral tendencies with weights; A is an agent's instant acceleration vector at each time step ($|A| < A_{\max}$); $A_{\max}$ is the maximum acceleration of an agent; $V_{\max}$ is the maximum velocity of an agent; c_1–c_5 are the weights of vectors V_1–V_5; d is the desired distance that an agent tries to maintain with its neighbors; the combination of c_1 and parameter d determines the relaxation degree of model shape; c_2 is used to control the motion range of all agents in the analog world; c_3 represents the consistency degree that makes certain individual keep the same velocity with other individuals in its neighborhood; c_4 indicates the distance degree that makes certain individual keep away from other individuals in its neighborhood; and c_5 is used to control the arbitrary degree of individual's acceleration change.

4.2.3. Environment. The agent in supply chain mainly self-organizes and adjusts its behavior according to the neighbors' behavior. An agent that relates to any other agent must exist within certain range, and the relationship between agents can be divided into two states: cooperation and noncooperation, which is shown as attraction and rejection behavior. Define the neighborhood of agent i as follows:

$$\mathrm{NEB}_i = \left\{ x_j \mid \left\| x_i - x_j \right\| \le d, i \ne j; i, j = 1, 2, \ldots, N \right\}, \tag{3}$$

where $\|x_i - x_j\|$ is Euclidean distance and d is neighborhood radius, which is the optimal distance between every two agents.

Define the rejection region of agent i as follows:

$$\mathrm{REP}_i = \left\{ x_j \mid \left\| x_i - x_j \right\| \le r, i \ne j; i, j = 1, 2, \ldots, N \right\}, \tag{4}$$

where r is the rejection distance, which is the maximum operating distance of the rejection function, $r \le d$. If the distance between two agents is less than r, then the rejection occurs.

4.2.4. Response Rules. In this paper, we divide the neighborhood range according to the distance between each agent and its neighbor agents. An agent establishes or disconnects the interaction between it and other agents within the neighborhood, which makes the local topology structure of agents in the system change, leading to the entire system topology evolution.

Response rules are as follows. The member that enters into the neighborhood from somewhere outside the neighborhood establishes new contacts. And the members within rejection region have competitive relationship and repulsive force towards other agents within rejection region, while the members within neighborhood have cooperative relationship and attractive force towards other agents within the neighborhood. By the local fitness function, we can obtain new self-organizing ability and executing ability of each member and then get the new network status.

4.3. The Resilient Index of Supply Chain MAS. Supply chain MAS is a self-organizing system, and its resilience generates from the microlevel and is revealed in macrolevel. The concept of energy in system dynamics can be an effective interpretation of the relationship between behavior and internal mechanism. Positive and negative feedback mechanisms are formed between every two agents, which is shown as the transmission of system internal energy. Therefore, we can show the influence of dynamic process on overall system resilience through quantifying system internal energy.

Assume that each supply chain member is an agent with resilience; then we can determine whether the relationship between it and other members is strengthened or weakened according to the environment, such as order quantity. The relation change between two agents can be seen as elastic force of mutual attraction or mutual repulsion, and the elastic force causes the movement of the agent, which leads to changes of the agent status in the entire supply chain. For example, certain enterprise will rise in the competition and become the core business, and core enterprise may gradually lose dominance. Take the agent status in the supply network as the elastic potential energy and the agent activity in the supply network as the elastic kinetic energy, which together constitute the resilient system of supply chain.

$U_i(t)$ and $E_i(t)$ are used to, respectively, represent the potential energy and kinetic energy of agent i, and $U_i(t)$ is the relative potential energy of the agent compared to MAS center; $U_i(t) = \|\text{pos}_i - \overline{\text{pos}}\|^2$. Kinetic energy $E_i(t)$ is the relative kinetic energy of the agent compared to MAS center; $E_i(t) = \|v_i - \overline{v}\|^2$.

The resilient index of supply chain MAS is

$$Q(t) = \sum_{i=1}^{N} \left(U_i(t) + E_i(t)\right). \tag{5}$$

4.4. Algorithm Implementation

Step 1. Initialize the state of each agent in supply chain and corresponding weights of the objects.

Step 2. With regard to the current state, do the following.

(1) Calculate the neighborhood of current scheme and judge whether it belongs to neighbors.

Because each agent only communicates with adjacent agents and local environment, we need to determine whether the distance between each agent and other agents is within the given value of the neighbor; namely, we need to judge whether the two agents are neighbors or not. And this is the basis of calculating subsequent variables. If the distance between two agents is less than the value of neighbor, then the two agents are neighbors and their behaviors affect each other. Otherwise, the two agents do not affect each other. Given that the distance value between two agents is $d(i, j)$, the judge is as follows:

if $d \leq$ neighbor, then i and j are neighbors.

(2) Calculate the values of variables V_1–V_5.

V_1 is the average vector of certain agent pointing to all other agents within a distance of d, named as Spacing$U$. If two agents i and j are neighbors but their distance is less than a given value, then the two will produce repulsion and adjust the vector V_1 as follows:

Spacing$U(i,:)$ = Spacing$U(I,:)$ + [abs($x(i) - x(j)$)abs($y(i) - y(j)$)].

V_2 is a vector pointing to the center of analog world and its name is decided by the center of analog world, named as Wcenter$U$, where $i \neq j$ and N is the number of agents. Consider the following:

Wcenter$U(i,:)$ = [Wcenter$U(i,:) + x(j)W$center$U(i,:) + y(j)$],

Wcenter$U(i,:)$ = Wcenter$U(i,:)/N$.

Moreover, we need to limit the range of the vector, in order to make entire motions of the system within the range, and adjust the vector as follows:

If $|W$center$U(i,:)|$ > bound,

Wcenter$U(i,:)$ = [Wcenter$U(i,:) - x(j)W$center$U(i,:) - y(j)$].

V_3 is the average speed vector of all agents that is adjacent to certain agent, named as Velocity$U$, and its value is determined by the speed of the agent and its neighbors:

Velocity$U(i,:)$ = Velocity$U(i,:)$ + Velocity$U(j,:)$,

Velocity$U(i,:)$ = Velocity$U(i,:)$/number(neighbor).

V_4 is a vector of certain agent pointing to the center of all of its adjacent agents, which is determined by the average vector of the agent and its adjacent agents, named as Center$U$. Its value is determined by the speed of the agent and its neighbors:

Center$U(i,:)$ = [Center$U(i,1) + x(j)$Center$U(j,:) + y(j)$],

Center$U(i,:)$ = Center$U(i,:)$/number(neighbor).

V_5 is a vector of random unit length, named as Wander$U$, the initial value of which is randomly given as follows:

Wander$U(i,:) = 2 * \text{rand}(1, 2)$.

(3) Calculate the local fitness value of each agent according to local fitness function, which means the calculation of agent's acceleration.

(4) Guide the interactions between all agents according to local fitness function, which means the calculation of agent's new position.

The interactions between all agents make the system evolve from initial interrupts to the equilibrium state of each node. Under the new obtained equilibrium state, change the agent state according to system local energy information, which makes the system achieve a more stable equilibrium state again and optimize the system resilience.

(5) Return the current value of system resilience.

Step 3. Repeat Step 2 until termination.

5. Case Simulation and Analysis

Take the Tianjin FAW Xiali Automobile Group's supply chain as a prototype for simulation. According to the transactions in 2011, we extract and compare FAW Xiali's supply data of recent years and then select the enterprises that make transactions with Tianjin FAW Group of more than 500,000 yuan as the nodes to constitute the supply chain we study, which includes 12 major suppliers and 7 principal subsidiaries as well as 11 major vendors of Tianjin FAW Group [27]. The initial network is shown in Figure 3.

5.1. Evolution Characteristics of Supply Chain. We select 5 sets of parameters to regulate the behavior of each supply chain member and simulate the process of supply chain evolution,

Table 4: The simulation coefficient.

Number	c_1	c_2	c_3	c_4	c_5	V_{max}	A_{max}	d
1	5	8	7	8	5	13	38	0.14
2	6	15	10	8	7	5	38	0.14
3	6	9	1	5	8	36	11	0.59
4	3	5	7	7	4	13	37	0.40
5	3	5	7	7	4	13	37	0.10

as shown in Table 4. In particular, the network diagram of part time period in evolution process of number 1 is selected to represent the evolution process of supply chain, as shown in Figure 4. In Figure 4, vertical and horizontal coordinates represent agent's position in MAS, which reflects enterprise status in the supply chain. As we can see from Figure 4, each agent's vertical and horizontal coordinates' value gradually increases, which means that the ability of each enterprise in supply chain is constantly increasing. And the distance between all agents is smaller and smaller, which means that the cooperation between enterprises is closer and closer. The supply chain shows a combined effect.

Figures 5, 6, 7, and 8, respectively, correspond to the evolution results of the 4 sets of parameters from number 2 to number 5. These evolution results all reflect the combined effect of supply chain and the parameters have a certain impact on it. We find that the parameter of V_{max} directly influences the speed of combined effect. Comparing number 1 with number 2, when we decrease the value of V_{max} from 13 to 5 and keep the other parameters unchanged, the speed of combined effect is significantly faster. Because the enterprise executive ability becomes stronger, mutual cooperation is more efficient. In addition, the parameter d also has a direct impact on the speed of combined effect. Comparing number 3 with number 5, the value of d is larger in number 3, which means that each enterprise has larger neighbor range and can cooperate with more enterprises, leading to more mutual cooperation and faster combined effect speed, while the value of d in number 5 is smaller and the situation is just the opposite.

Supply chain vulnerability is caused by cascading failures, while the combined effect happens to be the hindrance of cascading failures. Therefore, the combined effect is a factor that shows supply chain resilience.

The evolution of supply chain lies in affecting and constraining mutual behavior through interconnections and interactions between individuals. Clustering coefficient can be used to express connectivity, as a quantitative indicator of supply chain MAS connectivity factor. The entire network clustering coefficient C is the average of clustering coefficient C_i for each node i. Consider the following:

$$C_i = \frac{2E_i}{[k_i(k_i-1)]}. \qquad (6)$$

Rate of information flow is established based on the connectability of individuals. The information flow can be shown through measuring the connection path length between individuals. We choose the average path length L as a quantitative indicator of supply chain MAS internal information flow. Consider the following:

$$L = \frac{2}{N(N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} d_{ij}. \qquad (7)$$

Figure 9 shows the clustering coefficient (red) and the average path length (green) of supply chain evolution in number 1. The clustering coefficient keeps rising, indicating that individual connectivity continuously strengthen, which gradually generates system resilience. When the clustering coefficient is 1, it indicates that all agents are within other agents neighborhood and connectivity is optimal. The average path length keeps decreasing, indicating that the interactions of individual information change from indirect process to direct process. With the communication constantly decreasing, local rate of information flow is enhanced. When the average path length is 1, the system has the best overall interactivity. The same characteristics of the two indexes are also shown in the data from number 2 to number 5.

As the above characteristics, the degree of homogeneity in the supply chain enterprise is higher and higher. Such a system can resist well the expected danger, but in the face of some unexpected threat it becomes extremely fragile. Therefore, the supply chain is a robust-yet-fragile system.

5.2. Resilience Analysis of Supply Chain. The average acceleration vector, position center, and average velocity vector are represented as $\overline{a}, \overline{x}, \overline{v}$, respectively, and the calculation is as follows:

$$\overline{a} = \frac{1}{N}\sum_{i=1}^{N} a_i, \qquad \overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i, \qquad \overline{v} = \frac{1}{N}\sum_{i=1}^{N} v_i. \qquad (8)$$

The neighborhood center of agent i and the average velocity vector of all agents within its neighborhood are represented as $\overline{x}_i^*$ and $\overline{v}_i^*$, respectively, and the calculation is as follows:

$$\overline{x}_i^* = \frac{1}{M_i}\sum_{j=1}^{M_i} x_j, \qquad \overline{v}_i^* = \frac{1}{M_i}\sum_{j=1}^{M_i} v_j, \qquad (9)$$

where M_i is the number of agents within the neighborhood of agent i. Consider the following:

$$Q(t) = \frac{1}{2}\sum_{i=1}^{N} \left(U_i(t) + E_i(t)\right),$$

$$E_i(t) = c_3 \left\| v_i - \overline{v} \right\|^2,$$

$$U_i(t) = c_1 \left\| x_i - \overline{x}_i^* \right\|^2 + c_2 \left\| x_i - \overline{x} \right\|^2 + \sum_{j=1, j\neq i} c_4 \psi\left(\left\| x_i - x_j \right\|\right). \qquad (10)$$

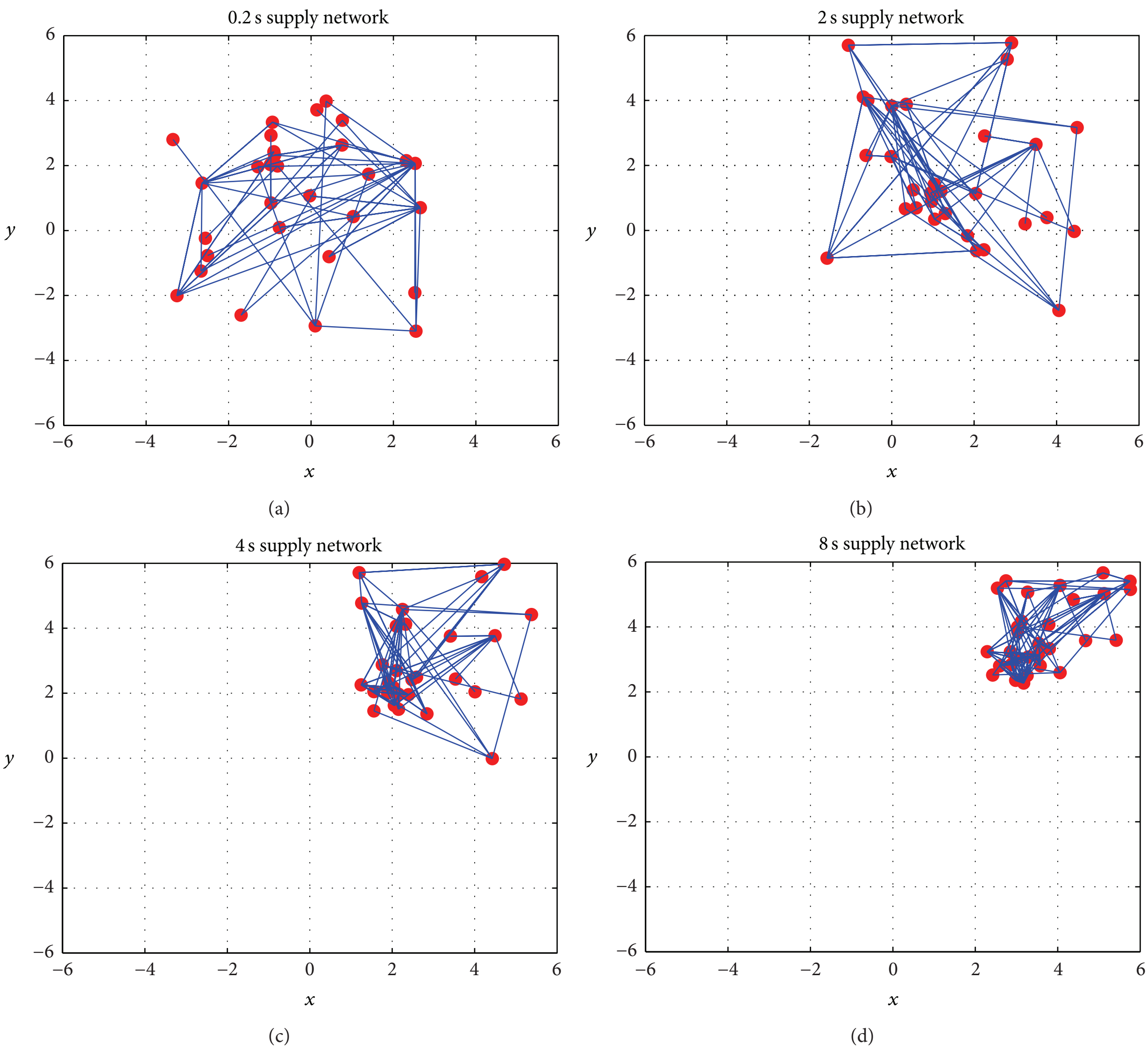


Figure 4: The emergence of supply chain combined effect.

The mean square error of acceleration, position, and velocity are, respectively, represented as σ_a, σ_x, and σ_v. Consider the following:

$$\sigma_a = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left\|a_i - \bar{a}\right\|^2}, \qquad \sigma_x = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left\|x_i - \bar{x}\right\|^2}, \qquad \sigma_v = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left\|v_i - \bar{v}\right\|^2}. \tag{11}$$

Assume that (1), for any agent, $\|x_i - x_j\| \le d$ and (2) the initial system resilience is Q_0, $\forall t > 0$, $Q(t) \le Q_0$.

The characteristic of function with potential energy $\psi(\|x_i - x_j\|) \ge \psi(\rho)$ is given as follows:

$$Q(t) = \frac{1}{2}\sum_{i=1}^{N}\left(U_i(t) + E_i(t)\right),$$

$$\sum_{i=1}^{N} c_1\left\|x_i - \bar{x}_i^*\right\|^2 + \sum_{i=1}^{N} c_2\left\|x_i - \bar{x}\right\|^2 + \sum_{i=1}^{N}\sum_{j=1, j\neq i} c_4\psi\left(\left\|x_i - x_j\right\|_\sigma\right) + \sum_{i=1}^{N} c_3\left\|v_i - \bar{v}\right\|^2 \le 2Q_0,$$

$$\sum_{i=1}^{N}\sum_{j=1, j\neq i} c_4\psi\left(\left\|x_i - x_j\right\|_\sigma\right) \ge c_4 N(N-1)\psi(\rho),$$

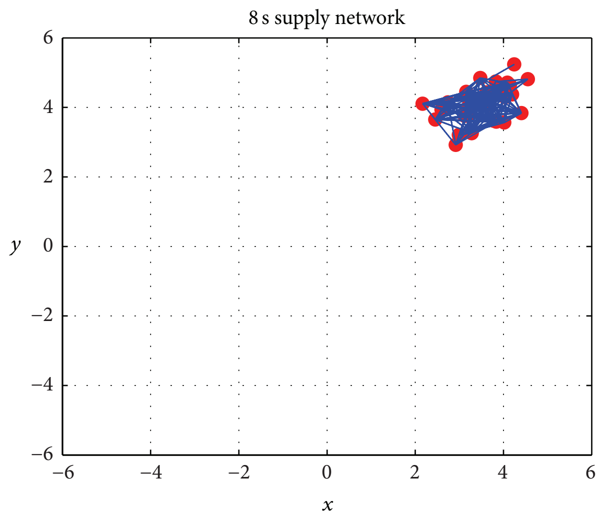


FIGURE 5: Evolution results of number 2.

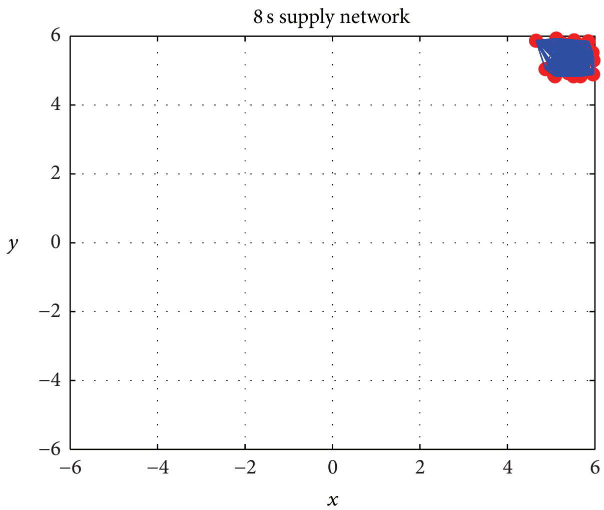


FIGURE 6: Evolution results of number 3.

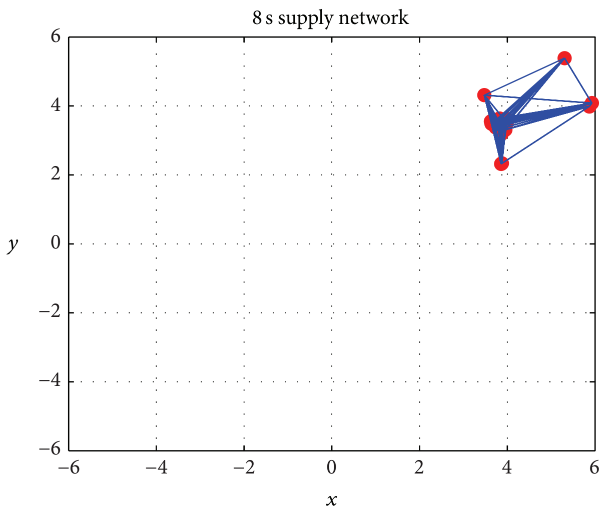


FIGURE 7: Evolution results of number 4.

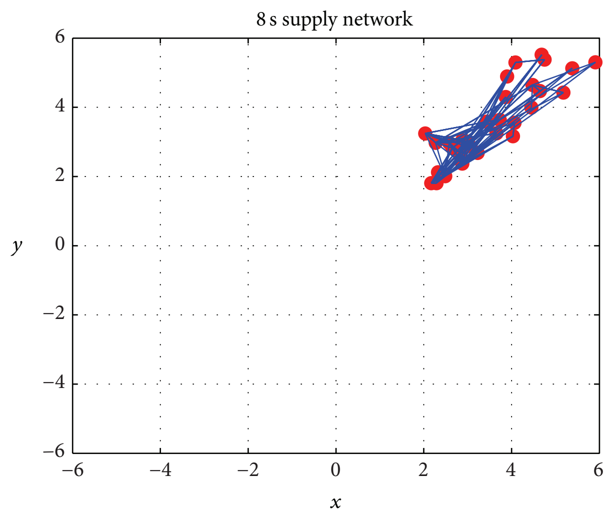


FIGURE 8: Evolution results of number 5.

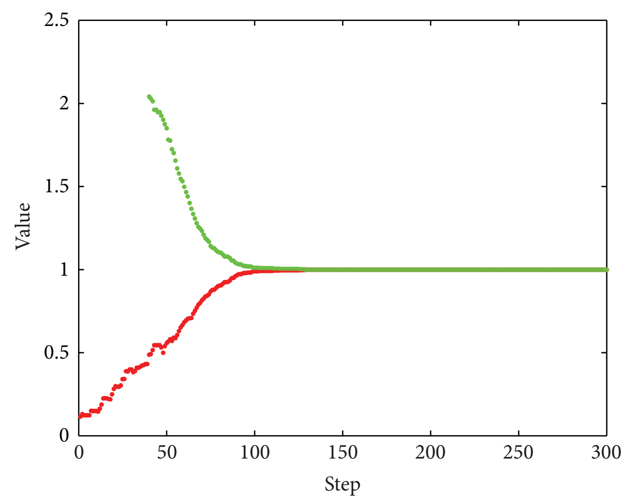


FIGURE 9: The emergence of connectivity and rate of information flow.

$$\sum_{i=1}^{N} c_1 \left\| x_i - \overline{x}_i^* \right\|^2 + \sum_{i=1}^{N} c_2 \left\| x_i - \overline{x} \right\|^2 + \sum_{i=1}^{N} c_3 \left\| v_i - \overline{v} \right\|^2 \le 2Q_0 - c_4 N (N-1) \psi(\rho). \tag{12}$$

Assume that $K = 2Q_0 - c_4 N(N-1)\psi(\rho)$ and

$$\begin{aligned} \sum_{i=1}^{N} c_2 \left\| x_i - \overline{x} \right\|^2 &\le K, & \sum_{i=1}^{N} c_3 \left\| v_i - \overline{v} \right\|^2 &\le K, \\ \sum_{i=1}^{N} \left\| x_i - \overline{x} \right\|^2 &\le \frac{K}{c_2}, & \sum_{i=1}^{N} \left\| v_i - \overline{v} \right\|^2 &\le \frac{K}{c_3}, \end{aligned} \tag{13}$$

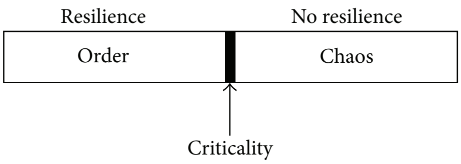


Figure 10: Self-organized criticality.

$$\sigma_x = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left\|x_i - \overline{x}\right\|^2} \le \sqrt{\frac{K}{c_2 N}}$$
$$\sigma_v = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left\|v_i - \overline{v}\right\|^2} \le \sqrt{\frac{K}{c_3 N}}. \quad (14)$$

As shown in (14), the fluctuation of all agents' ability and status will stabilize within a certain range, and the specific range is related to the system initial state and its coefficient. The smaller the value is, the more consistent the agent behavior will be and the more orderly the system will become. Small perturbation will not have a devastating impact on the supply chain, and the system will recover during the dynamic self-organizing process. In the case of small random disturbance, we can constrain the overall behavior of the supply network through setting the value of the initial state and the control parameters.

5.3. Supply Chain Resilience and SOC Emergence. Self-organized criticality can be considered as a characteristic state of criticality which is formed by self-organization in a long transient period at the border of stability and chaos, as shown in Figure 10.

Figure 11 shows the resilience index of evolution in number 1, which is used to display the effect of supply chain resilience. Figure 12 shows the resilience index of evolution in number 1 when the coefficient of Wander$U$ is larger ($V_5 = 20$); namely, the random disturbance is very large, in which the resilience disappears. Figure 11 is an orderly part of Figure 10, while Figure 12 is the other part of Figure 10, and there will be a critical value. The same characteristic is shown in Figures 13 and 14, which corresponds to the situation in number 5.

The simulation verifies the essence of supply chain resilience that we have described in Section 3, in which the supply chain resilience is the macroproperty that generates from microindividual "self-organization." In the process of evolution, supply chain system owns a stable state, which is gradually formed through each enterprise in the supply chain constantly adapting to the internal and external environment, and the dependence on the environment forms contractual relationships between enterprises. Thus, it ensures that the parameters of enterprises can be far away from mutation and obtain a relatively stable operational condition. This stable state is within the range of critical value. Because the random variable V_5 is added from the outside of system and the kinetic energy of each agent is transformed into potential energy through self-organizing interactions, the agent with stronger potential energy will have stronger self-organizing ability,

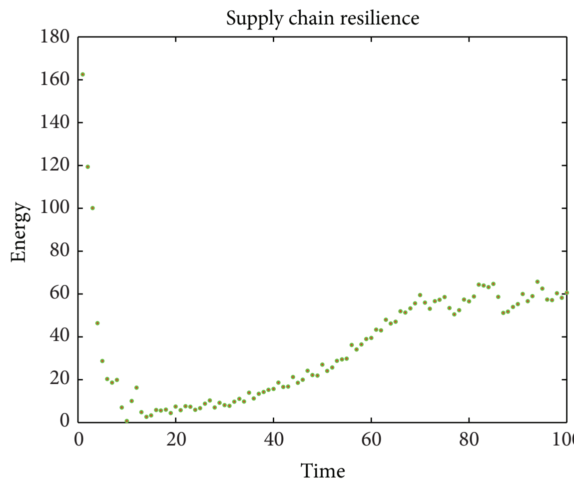


Figure 11: There is resilience within SOC/number 1.

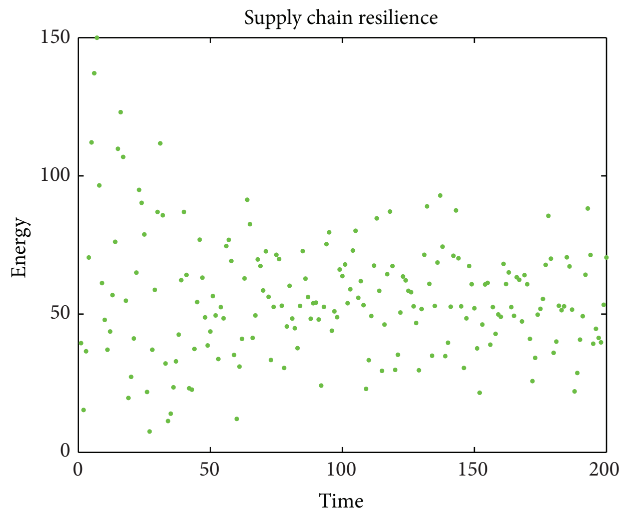


Figure 12: There is no resilience at SOC/number 1.

which can further enhance the agent kinetic energy. Therefore, there is energy flow within whole supply chain system. The reason why the critical point can be achieved is because the energy is inputted from the outside of system in the form of random variable V_5. The impact of fluctuations continues to gather, eventually leading to a large-scale collapse.

In order to further verify the above conclusion, we conduct a test based on the data of chemical integrated supply network in the literature of [28]. The network organization focuses on natural gas and chemical engineering production and owns a cooperative system of up, middle, and down product chain with comprehensive supporting enterprises, including 40 functional units, in which 1–19 are the core production units, 20 and 21 are raw materials processing units, and 22–40 are storage and transportation units of semifinished products and finished products. The initial network is shown in Figure 15. The parameters of number 1 are used in this case. The same characteristics are shown in Figures 16 and 17, which correspond to the situation in the case.

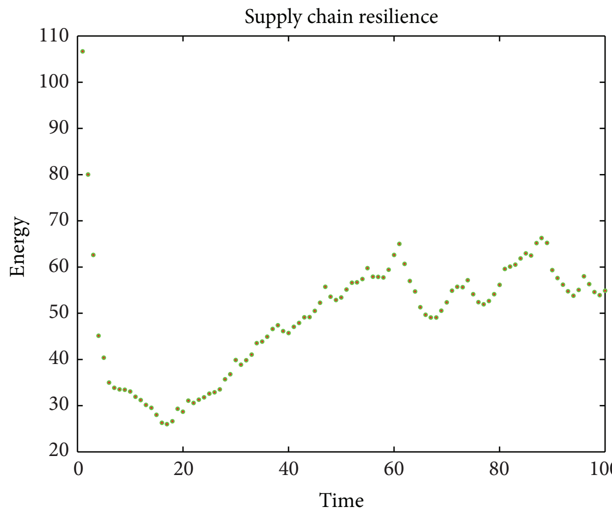


Figure 13: There is resilience within SOC/number 5.

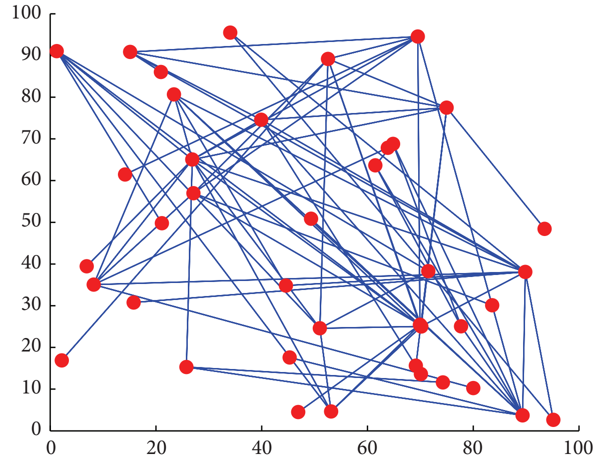


Figure 15: The supply network of chemical group.

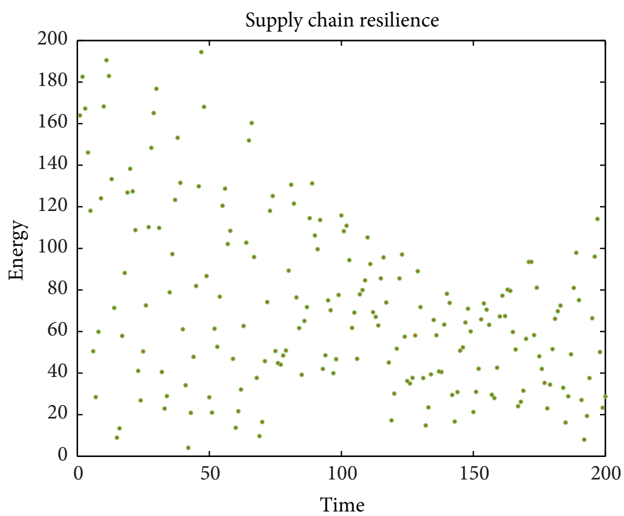


Figure 14: There is no resilience at SOC/number 5.

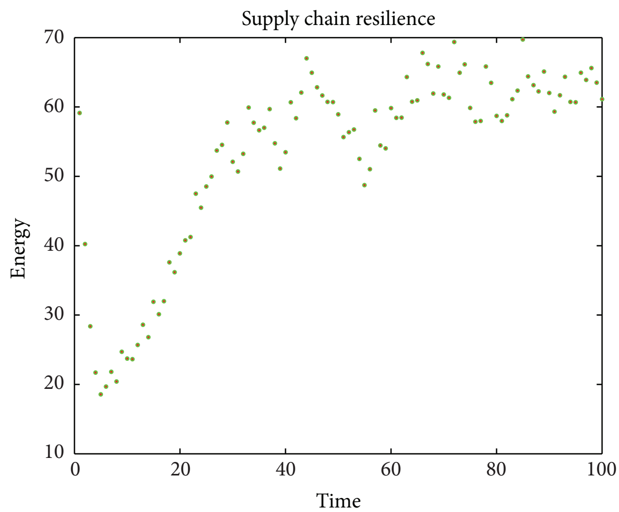


Figure 16: There is resilience within SOC/chemical.

In fact, the disaster is caused by many small highly decentralized decisions, which gather together in a pace that is difficult to detect. Every decision is too tiny to look harmful, but they can slowly erode the resilience of the system. When making decisions, none of the enterprises can understand the impact that their behavior brings to the entire supply chain and, unknowingly, the system becomes more vulnerable. Every enterprise takes rational action itself within local range, which can bring significant individual interest. As time goes on, these decisions slowly change the rules of the system. Since the previous choices may not be punished, each enterprise tends to choose behavior with more interest, which also brings higher risk. As a whole, the supply chain system slowly approaches the possible disaster, showing self-organized criticality.

From resilience to no-resilience, it will cross a critical point. And the critical point for the supply chain system also frequently changes. It needs to balance between interest and

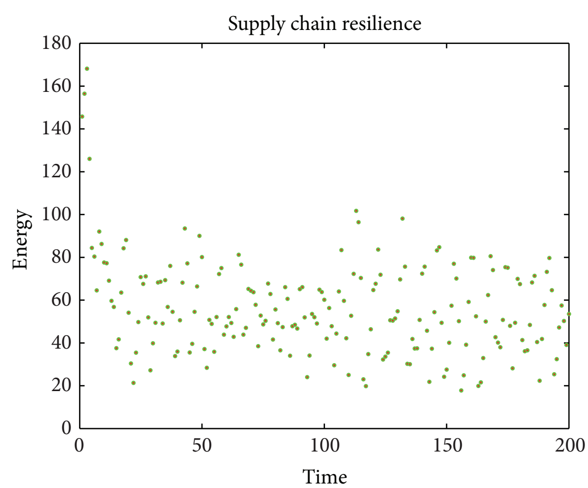


Figure 17: There is no resilience at SOC/chemical.

risk to make choice between the efficient but vulnerable and the inefficient but robust. To control supply chain resilience, it needs to choose a good balance between them.

6. Conclusion

Based on our previous study [11], we map the supply chain system into MAS system through the comparison of supply chain's microscopic individual behavior and its macroscopic property. Through taking the enterprise as agent and designing the local fitness function, neighborhood structure, and interaction rules that are suitable for supply chain system, we find that the macroscopic property of supply chain system generates from its microscopic behavior and the supply chain resilience generates by local self-organizing behavior. Besides, we also study the resilience of supply chain and its self-organized critical characteristic. And we get some meaningful conclusions as follows.

(1) We find the combined effect in the evolution of supply chain through the simulation, which is similar to the actual situation. The connectivity increases as the clustering coefficient increases, and the rate of information flow increases as the average path length decreases. Through closer and closer cooperation and communication of information between enterprises, the degree of homogeneity in the supply chain enterprise becomes higher and higher.

(2) Through analyzing the statistical properties of supply chain resilience model, it is shown that resilience is controlled by parameters of agents, such as the executive ability and the scope of choosing partners, which have a positive effect on the combined effect. Managers can control the supply chain within the resilient range by controlling the behavior of the enterprise. Within the range, supply chain can continuously eliminate effects and release energy due to its self-organizing behavior, to slow down the occurrence of large-scale collapses events.

(3) We find that there exists SOC characteristic in supply chain resilience through the simulation, and the supply chain is a robust but vulnerable system. Most of the time, supply chain stays in a stable state due to its resilience. But when it crosses the self-organization critical point, the supply chain will out of control and an avalanche event may occur by small perturbations, which means that the supply chain is extremely vulnerable in this environment. Managers will find the proper balance between the efficiency and risk by controlling the behavior of the enterprises.

To enhance the supply chain resilience, we should pay close attention to self-organizing ability of each enterprise. As the supply chain has characteristic of self-organized criticality, supply chain should be prevented from crossing self-organized critical point and should be controlled. To control supply chain resilience, we can choose a good balance between the efficient but vulnerable and the inefficient but robust. This study provides theoretical guidance to optimize the resilience of the supply chain through local control. In the future, we will study the factors that affect the self-organized critical point and how to prevent the supply chain from crossing self-organized critical point.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This project is supported by the National Natural Science Foundation of China (no. 71171089) and the Natural Science Foundation of Zhejiang Province of China (no. LY14F020023).

References

[1] X. Y. Huang and N. N. Yan, "Research progress on supply chain robustness," *Chinese Journal of Management*, vol. 4, no. 4, pp. 521–528, 2007.

[2] C. S. Holling, "Resilience and stability of ecological systems," *Annual Review of Ecology and Systematics*, vol. 4, pp. 1–23, 1973.

[3] C. Folke, S. Carpenter, T. Elmqvist, L. Gunderson, C. S. Holling, and B. Walker, "Resilience and sustainable development: building adaptive capacity in a world of transformations," *Ambio*, vol. 31, no. 5, pp. 437–440, 2002.

[4] W. N. Adger, "Social and ecological resilience: are they related?" *Progress in Human Geography*, vol. 24, no. 3, pp. 347–364, 2000.

[5] G. A. Bonanno, "Loss, trauma, and human resilience: have we underestimated the human capacity to thrive after extremely aversive events?" *The American Psychologist*, vol. 59, no. 1, pp. 20–28, 2004.

[6] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, "Network robustness and fragility: percolation on random graphs," *Physical Review Letters*, vol. 85, no. 25, pp. 5468–5471, 2000.

[7] R. Starr, J. Newfrock, and M. Delurey, "Enterprise resilience: managing risk in the networked economy," *Strategy and Business*, no. 30, pp. 70–79, 2003.

[8] Y. Sheffi, "Building a resilient supply chain," *Harvard Business Review*, vol. 1, no. 8, pp. 1–4, 2005.

[9] P. P. Datta, M. Christopher, and P. Allen, "Agent-based modelling of complex production/distribution systems to improve resilience," *International Journal of Logistics Research and Applications: A Leading Journal of Supply Chain Management*, vol. 10, no. 3, pp. 187–203, 2007.

[10] S. Ratick, B. Meacham, and Y. Aoyama, "Locating backup facilities to enhance supply chain disaster resilience," *Growth and Change*, vol. 39, no. 4, pp. 642–666, 2008.

[11] C. Colicchia, F. Dallari, and M. Melacini, "Increasing supply chain resilience in a global sourcing context," *Production Planning and Control: The Management of Operations*, vol. 21, no. 7, pp. 680–694, 2010.

[12] K. Zhao, A. Kumar, T. P. Harrison, and J. Yen, "Analyzing the resilience of complex supply network topologies against random and targeted Disruptions," *IEEE Systems Journal*, vol. 5, no. 1, pp. 28–39, 2011.

[13] L. Geng, R. B. Xiao, and S. S. Xie, "Research on self-organization in resilient recovery of cluster supply chains," *Discrete Dynamics in Nature and Society*, vol. 2013, Article ID 758967, 11 pages, 2013.

[14] T. J. Pettit, K. L. Croxton, and J. Fiksel, "Ensuring supply chain resilience: development and implementation of an assessment tool," *Journal of Business Logistics*, vol. 34, no. 1, pp. 46–76, 2013.

[15] R. Francis and B. Bekera, "A metric and frameworks for resilience analysis of engineered and infrastructure systems," *Reliability Engineering and System Safety*, vol. 121, pp. 90–103, 2014.

[16] J. G. Liu, G. W. Shi, B. Lu et al., "The three factor model of the supply chain elasticity," *Chinese Journal of Management Science*, vol. 20, pp. 528–535, 2012.

[17] P. Bak, "How nature works: the science of self-organized criticality," *Nature*, vol. 383, no. 6603, pp. 772–773, 1996.

[18] P. Bak, *How Nature Works*, Oxford University Press, Oxford, UK, 1997.

[19] J. P. Da Cruz and P. G. Lind, "Self-organized criticality in a network of economic agents with finite consumption," *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 4, pp. 1445–1452, 2012.

[20] J. L. Guo, "Bilateral power-law distribution model of supply chain networks," *Acta Physica Sinica*, vol. 55, no. 8, pp. 3916–3921, 2006 (Chinese).

[21] X. Gabaix, P. Gopikrishnan, V. Plerou, and H. E. Stanley, "A theory of power-law distributions in financial market fluctuations," *Nature*, vol. 423, no. 6937, pp. 267–270, 2003.

[22] A. Clauset, C. R. Shalizi, and M. E. . Newman, "Power-law distributions in empirical data," *SIAM Review*, vol. 51, no. 4, pp. 661–703, 2009.

[23] L. Lu and G. Wang, "A study on multi-agent supply chain framework based on network economy," *Computers & Industrial Engineering*, vol. 54, no. 2, pp. 288–300, 2008.

[24] A. Smith and J. M. Vidal, "A practical multiagent model for resilience in commercial supply networks," *Agent-Mediated Electronic Commerce*, vol. 2010, article 169, 2010.

[25] Z. Huang and R. Xiao, "An emergent computation approach to the problem of polygon layout with performance constraints," *Physica A: Statistical Mechanics and Its Applications*, vol. 392, no. 20, pp. 5074–5088, 2013.

[26] Y. Wu, J. Su, H. Tang, and H. Tianfield, "Analysis of the emergence in swarm model based on largest Lyapunov exponent," *Mathematical Problems in Engineering*, vol. 2011, Article ID 745257, 21 pages, 2011.

[27] http://www.cninfo.com.cn/finalpage/2012-04-21/60870836.PDF.

[28] F. Zhang, Y. Yang, J. Jia, and J. Wang, "Correlation analysis of failure modes and vulnerability in collaborative production networked organizations," *Computer Integrated Manufacturing Systems, CIMS*, vol. 18, no. 6, pp. 1236–1245, 2012 (Chinese).

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 906739, 7 pages
http://dx.doi.org/10.1155/2014/906739

Research Article

Options Procurement Policy for Option Contracts with Supply and Spot Market Uncertainty

Weili Xue,[1] Xiaolin Xu,[2] and Lijun Ma[3]

[1] *School of Economics and Management, Southeast University, Nanjing 210096, China*
[2] *School of Business, Nanjing University, Nanjing 210093, China*
[3] *Department of Management Science, Shenzhen University, Shenzhen 518060, China*

Correspondence should be addressed to Xiaolin Xu; xuxl@nju.edu.cn

Received 11 April 2014; Accepted 11 June 2014; Published 1 July 2014

Academic Editor: Xiang Li

Copyright © 2014 Weili Xue et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplier's reliability is a major issue in procurement management. In this paper, we establish a decision making model from the perspective of the firm who will procure from the multiple suppliers and the spot markets. The suppliers are unreliable and provide different types of option-type supply contracts which should be made before demand realization, while the spot market can only be used after demand realization and has both the price and liquidity risks. We establish the optimal portfolio policies for the firm with conditions to find the qualified suppliers. By defining a new function which contains the demand risk, the supplier's risk, and the liquidity risk, we find that the optimal policy is to allocate different curves of this function to different suppliers. We also study some special cases to derive some managerial insights. At last, we numerically study how the various risks affect the choice of suppliers and the value of the option contract.

1. Introduction

Procurement is one of the major elements of a firm's operations management. In most industries, procurement cost usually consists of a large part of the firm's total costs and thus has large influence on the firm's total operation cost and its competitive advantage. To manage the procurement cost, it is a common practice for firms to maintain a list of suppliers with different cost structures to secure their supply. However, the resulting complex procurement structure leads to some complex management problems, for example, how to choose the appropriate suppliers and how to allocate the demand to each of these suppliers. These management problems ask for the practitioners and scholars to develop or provide tools to procurement managers in decision making, to maintain a stable production and a high customer service level under a reasonable procurement cost. In practice, long term contract between a firm and its supplier, with fixed or flexible terms in price or quantity, is still the major way. With such a long term contract, a firm can always secure the quantity and maintains a reasonable procurement cost in the meantime. However, as reported in a global investigation by Mckinsey [1], with the rapid changing of the business environment, the reliability risk of the supplier has also become one of the major issues in the procurement practice and has been paid attention to by the practitioners and scholars. For example, in 2000, Sony's stock price plummets 9% because his supplier cannot supply the LCD and flash memories in time, which leads to Sony's incapability to meet his customer's demand for PlayStation. In 2005, one of the major private-owned coal companies, that is, Peabody Energy, has announced a loss of 3.4$ in his semiannual report because his supplier cannot fulfill the supply contract.

Besides the traditional long term contract, spot market (and B2B market) has played a more and more important role in a firm's procurement strategy as a solution to those mismatches between supply and demand, especially with the development of information technology. It has become a supplement to the traditional long term contract, with its time and quantity flexibility in procurement. Currently, the products traded in the spot market have expanded and include semiconductors, memories, and CPUs besides the traditional commodities such as oil, metal, and agricultural products. However, spot market has its own deficiency as compared to the long term contract. On the one hand, the price in spot market is volatile. For example, although the price for semiconductors has a decreasing tend in the long term, the short term price increase can arrive to 200% or 300%. On the other hand, the spot market exits liquidity problem, which means that the firm cannot buy the required quantity at the spot price in some times. Although the liquidity issue is not a problem for the liquid market such as the copper and zinc, it is still a problem for iron and paper pulp [2].

To mitigate the procurement risk (i.e., reliability risk, price risk, and liquidity risk) discussed above, HP launches a PRM (proactive risk management) project. The major idea of such project is to segment the demand curve with different risk and then outsource them to the supplier with different cost structures. Our paper is based on the idea proposed by such project. Specifically, we establish a decision making model from the perspective of the firm who will procure from the multiple suppliers and the spot markets. The suppliers are unreliable and provide different types of option-type supply contracts which should be made before demand realization, while the spot market can only be used after demand realization and has both the price and liquidity risks. Under such a decision-making framework, we establish the optimal portfolio policies for the firm with conditions to find the qualified suppliers. Moreover, we show how the various risks affect the choice of suppliers and the value of the option contract. Our contribution has twofold: first, we provide a procurement portfolio policy to a more general environment; second, we provide some insights on how the environment parameters affect the buyer's cost.

The paper is organized as follows. In Section 2, we will review the related literature and position our research among them. In Section 3, we will establish the model and develop the optimal policy. Moreover, we will study two special cases to dig out more insights. In Section 4, the numerical study will be conducted. And last, in Section 5, a conclusion will be made with some discussions for future research.

2. Review of the Literature

One research stream related to our current research is those works on optimal coordination policies with long term contracts, including the quantity flexible contract, price flexible contracts, and time flexible contracts. However, most previous works do not consider the spot market and do not include the supplier's reliability risk. Readers can refer to Cachon [3] for a brief review on these models. Based on these works, Wu et al. [4], Kleindorfer and Wu [5], and Wu and Kleindorfer [6] study the coordination mechanism with the presence of spot market. Specifically, Wu et al. [4] and Kleindorfer and Wu [5] consider the coordination strategy and the options' pricing strategy under two-party tariff contract. Wu and Kleindorfer [6] extend the model to include multiple suppliers and investigate the coordination strategy and the contract's pricing problem. However, different to the model in our paper, in those models, the spot prices are always assumed to be a deterministic function of the demand and, moreover, the suppliers are all reliable.

Another related research stream is those works which consider the optimal procurement policies from the buyer's perspective, to minimize his total procurement cost. Ritchken and Tapiero [7] study the role of the option contracts in the risk management, when the spot market is stochastic. Mattock [8] studies the optimal capacity planning problem when the demand and the spot price are related and proposes an algorithm to solve the corresponding problem. However, these models only consider one supplier. When there are multiple suppliers, Schummer and Vohra [9] consider the optimal procurement policy and find that the problem can be converted to a linear programming problem. They thus designed an incentive compatible auction mechanism for the option-type procurement contract. With a similar framework, Martínez-de-Albéniz and Simichi-Levi [10] study the procurement problems for planning horizon with multiple periods. They derive the optimal procurement policy under the condition that the contract types and numbers in each period are the same. Fu et al. [11] further propose an algorithm for the optimal ordering policy when the stochastic demand and the spot price are correlated. However, all those works neglect the reliability risk and the spot market liquidity risk.

Our current research also contributes to the literature on operations management in the presence of random yield. Yano and Lee [12] and Minner [13] make some conclusions on these works. In general, most of these works only establish the optimal policy under supply risk and the supply chain typically includes only one supplier. With the newsvendor framework, Dada et al. [14] extend previous research to include multiple suppliers with different reliability risks. In their model, the supply risk for each supplier is independent of each other and there are no spot markets. Moreover, the contracts between the suppliers and the buyer are wholesale-price contracts. Our paper considers the random yield problems with multiple suppliers; however, the difference lies in twofold: first, the long term contract we considered is option-type contract; second, the reliability risks for each supplier are totally correlated. Although the second characteristic is strict, it has its own practice background. Moreover, it makes our model more tractable.

3. Model Formulation

We consider a single period procurement model under a supply chain with a buyer and N suppliers (see Figure 1), in which the ith supplier is called supplier i, $i = 1, 2, \ldots, N$. Similar to the settings in Fu et al. [11], the buyer can procure from each of these suppliers before demand realization or can buy from the spot market after demand realization. Thus, we assume the lead time for procurement from suppliers as one and that from the spot market as zero. This is a typical assumption in literature considering both the long term supplier and the spot market. However, different to the models of Fu et al. [11], we consider two more realistic

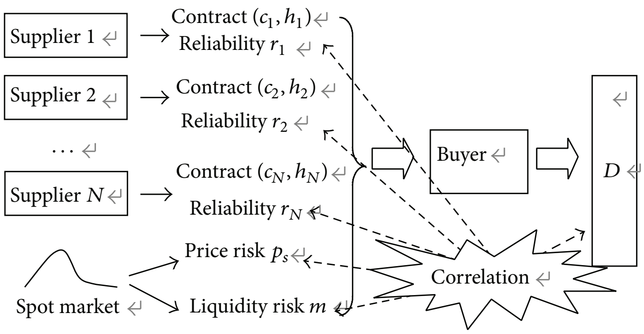


Figure 1: Model description.

settings. First, the procurement from the spot market will face the liquidity risk and thus has uncertainties, even after spot price realization. We denote the probability that the buyer can procure from the spot market at the prevailing spot price by m, $m \in [0, 1]$. The value of m is a stochastic variable. Such probability measures the uncertainty that the buyer can find a proper seller in the spot market. For a discussion on the liquidity risk, readers can refer to Wu and Kleindorfer [6]. Second, the suppliers are unreliable. That is, if the buyer procures Q_i from the ith supplier, he can only get $r_i Q_i$ after demand realization, in which r_i is a random variable realized after demand realization (or equivalently, after products arrive). Such multiplicative type of random yield is also studied in Yano and Lee [12].

Then, the decision sequences of the model are as follows.

(1) At the beginning of the planning horizon, denoted by time $t = 0$, supplier i, $i = 1, 2, \ldots, N$, will provide an option-type supply contract (c_i, h_i) to the buyer. Here, c_i is the unit reservation price which the buyer should pay when making the contract for reserving a fixed quantity, h_i is the unit exercise price which will be paid by the buyer on the actual quantity of the products received. The different pairs of (c_i, h_i) represent the different cost structure and risk attitude of the suppliers. From the perspective of the buyer, these pairs provide him different flexibilities. Thus, the buyer should decide the quantity Q_i to reserve from each of the suppliers.

(2) Then, the selling season starts and the products from the supplier are delivered at time $t = 1$. Depending on the realization of the customer demand D, the spot price p_s, the market liquidity m, and the suppliers' supply risk r_i's, the buyer should decide the exercise quantity from each of these suppliers and the spot market to satisfy the demand. We denote the order quantities from the supplier i after demand realization by x_i when the spot market has enough liquidity and denote that from the supplier i by $\overline{x}_i$ when the spot market has no liquidity to satisfy the buyer's demand.

(3) At last, all unsatisfied demands will incur a shortage cost s to the buyer, and the buyer calculates his total procurement cost. We assume $s \geq \max_i\{h_i, i = 1, 2, \ldots, N\}$ without loss of generality.

To make our model tractable, we propose the following assumption on the random yield r_i for each supplier.

Assumption 1. $r_i = k_i \eta$, in which η is a random variable with mean $\overline{\eta}$.

Although this assumption is required mainly by technical reason, it has its own practical support. This assumption implies that the supply risks of these suppliers are all associated with a common environment variable η. For example, if those suppliers all have the same upstream supplier with disruption risk, or if all these suppliers' risks associated with some common variable (e.g., location, macroeconomic environment, spot price, etc.), we can regard that the supply risks of those suppliers have the form which this assumption asks for. Here, k_i measures the magnitude that the supplier i's supply risk affected by the environment variable. The higher the value of k_i is, the higher the mean supply quantity and the associated variance are. However, the coefficient of variant for each supplier's random yield r_i remains the same.

Without loss of generality, we assume the suppliers are ordered by the values of c_i. That is, $c_1 > c_2 > \cdots > c_N$. Then, we should have $h_1 < h_2 < \cdots < h_N$, as other contract pairs unsatisfying such condition will be ruled out. Moreover, we assume $h_1 = 0$. Then, the contract provided by supplier i is a wholesale price contract with wholesale price c_1, while other contracts are option contracts. We assume that even when the realized quantity from any supplier is less than the quantity reserved by the buyer, the supplier will not refund those reservation prices to the buyer. The other notation is as follows. Denote the cumulative distribution function and the probability density function of m, p_s, D, η by $F(m, p_s, D, \eta)$ and $f(m, p_s, D, \eta)$, respectively. Denote the marginal cumulative distribution function and the marginal probability density function of random variable X by $F_X(X)$ and $f_X(X)$, respectively. Denote the joint cumulative distribution function and the joint probability density function of random variables X and Y by $F_{XY}(X, Y)$ and $f_{XY}(X, Y)$, respectively. Denote the marginal cumulative distribution function and the marginal probability density function of random variable X, when $Z = z$, by $F_X(X, z)$ and $f_X(X, z)$, respectively.

With the model description (see Notation Section), we can establish the optimization model at time $t = 0$ for the buyer as follows:

$$\min_{Q_i, i=1,2,\ldots,N} \sum_{i=1}^{N} c_i Q_i + E_{D,m,p,\eta}\left[(1-m)\,\Phi_1(\mathbf{Q}, D, p, \eta) + m\Phi_2(\mathbf{Q}, D, p, \eta)\right], \tag{1}$$

in which $\mathbf{Q} = [Q_1, \ldots, Q_N]$ and $\Phi_1(\mathbf{Q}, D, p, \eta)$ and $\Phi_2(\mathbf{Q}, D, p, \eta)$ are the decision models at time $t = 1$, after realization of

D, m, p, η, depends on whether the buyer can or cannot buy from the spot market. Specifically,

$$\begin{aligned}\Phi_1(\mathbf{Q}, D, p, \eta) = \min_{\overline{x}_i, i=1,2,\ldots,N} \quad & \left(\sum_{i=1}^{N} h_i \overline{x}_i + s\overline{y}\right) \\ \text{s.t.} \quad & \overline{x}_i \leq r_i Q_i, \quad i = 1, 2, \ldots, N \\ & \sum_{i=1}^{N} \overline{x}_i + \overline{y} = D \\ & \overline{x}_i \geq 0, \quad i = 1, 2, \ldots, N\end{aligned} \tag{2}$$

is the decision on the procurement quantity from each supplier when the spot market has liquidity problem, and

$$\begin{aligned}\Phi_2(\mathbf{Q}, D, p, \eta) = \min_{x_i, i=1,2,\ldots,N} \quad & \left(\sum_{i=1}^{N} h_i x_i + \min\{p_s, s\} y\right) \\ \text{s.t.} \quad & x_i \leq r_i Q_i, \quad i = 1, 2, \ldots, N \\ & \sum_{i=1}^{N} x_i + y = D \\ & x_i \geq 0, \quad i = 1, 2, \ldots, N\end{aligned} \tag{3}$$

is the decision on the procurement quantities from each supplier and the spot market when the spot market has ample supply.

We now first consider the decision problems in the time $t = 1$. The policy in this stage is rather simple as all market uncertainties have realized. Specifically, the buyer satisfies his demand from the supplier with the lowest exercise price to the one with the largest exercise price and bears the shortage cost at last when the spot market is unavailable (or procures from the spot market if the spot market is available and the spot price is lower than the shortage cost). In the following part, we will only focus on the decision problems on the optimal reservation quantity Q_i from supplier i at time $t = 0$. Specifically, we first analyze the case, the general case, when all stochastic factors are considered. Then, we consider two special cases. One case only considers the demand risk and assuming $r_i = 1$, for all $i = 1, 2, \ldots, N$, while the other case assumes that the spot price p_s and the market liquidity probability m are deterministic. The purpose of these two special cases is to derive some managerial insights.

Before going further to the analysis, we first introduce the definition of active contract as defined in Fu et al. [11]. We use Q_i^* to denote the optimal reservation quantity to supplier i.

Definition 2. Contract i is called active, if $Q_i^* > 0$. Contracts i and j are called consecutive active contracts, if

$$Q_i^* > 0, \quad Q_j^* > 0, \quad Q_r^* = 0 \quad \text{for } i < r < j. \tag{4}$$

Contract i is called the last active contract, if

$$Q_i^* > 0, \quad Q_r^* = 0, \quad r > i. \tag{5}$$

Moreover, we define the effective exercise price h_i', $i = 1, 2, \ldots, N$, and the effective shortage cost s' as

$$\begin{aligned} h_i' &= \begin{cases} h_i, & \text{if } h_i \leq p_s; \\ (1-m) h_i + m p_s, & \text{if } h_i > p_s, \end{cases} \\ s' &= \begin{cases} s, & \text{if } s \leq p_s; \\ (1-m) s + m p_s, & \text{if } s > p_s. \end{cases} \end{aligned} \tag{6}$$

3.1. Analysis for the General Setting. For the general setting, we can define a series of functions, $\{g_n\}_{n=1}^{N}$, as follows:

$$g_n(Q_1, Q_2, \ldots, Q_n) = \eta F_D\left(m, p_s, \eta\left(\sum_{i=1}^{n} k_i Q_i\right), \eta\right). \tag{7}$$

Then, we have the following property associated with the optimal solutions to the procurement problem.

Theorem 3. *There exists a series of active contracts $\{Q_i^*\}$ for the optimization problem* (1). *For any consecutive active contracts i and j, $(j > i)$, one has the following.*

(1) *If i is not the last active contract, then one has*

$$\begin{aligned}\iiint_{\eta, p_s, m} & \left(\eta - g_i\left(Q_1^*, Q_2^*, \ldots, Q_i^*\right)\right)\left(h_j' - h_i'\right) \\ & \times f_{m,p_s,\eta}(m, p_s, \eta)\, d\eta\, dp_s\, dm = \frac{c_i}{k_i} - \frac{c_j}{k_j}.\end{aligned} \tag{8}$$

(2) *If i is the last active contract, then one has*

$$\begin{aligned}\iiint_{\eta, p_s, m} & \left(\eta - g_i\left(Q_1^*, Q_2^*, \ldots, Q_i^*\right)\right)\left(s' - h_i'\right) \\ & \times f_{m,p_s,\eta}(m, p_s, \eta)\, d\eta\, dp_s\, dm = \frac{c_i}{k_i}.\end{aligned} \tag{9}$$

Proof. By expanding the optimization problem, given the reservation quantity Q_i form supplier i, we get that the objective function has the form of

$$\begin{aligned}\sum_{i=1}^{N} c_i Q_i + \int_{\eta} & \left\{\sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \left(\sum_{t=1}^{i-1} h_t' r_t Q_t + h_i'\left(D - T_{i-1}\right)\right) f_{D,\eta}\, dD \right. \\ & \left. + \int_{T_N}^{\infty} \left(\sum_{t=1}^{N} h_t' r_t Q_t + s'\left(D - T_k\right)\right) f_{D,\eta}\, dD\right\} d\eta\end{aligned} \tag{10}$$

in which $T_i = \sum_{t=1}^{i} r_t Q_t$, $T_0 = 0$, and k is the last contract whose exercise price is less than the spot price.

Taking the first derivative with respect to Q_i and equaling it to zero, we get

$$\begin{aligned} c_i + \int_{\eta} & \left\{\sum_{j=i}^{N-1} \int_{T_j}^{T_{j+1}} r_i\left(h_i' - h_j'\right) f_{D,\eta}\, dD \right. \\ & \left. + \int_{T_N}^{\infty} r_i\left(h_i' - s'\right) f_{D,\eta}\, dD\right\} d\eta = 0.\end{aligned} \tag{11}$$

It is obvious that the Hessian Matrix of the objective function is positive definite and thus there exist unique solutions for Q_i, for $i = 1, 2, \ldots, N$, within the restriction that $Q_i \geq 0$.

For any consecutive active contracts i, j, and the last active contract l, $(l > j > i)$, we have

$$\begin{aligned} c_i + \int_\eta \Bigg\{ \sum_{t=i}^{l-1} \int_{T_t}^{T_{t+1}} r_i \left(h_i' - h_t'\right) f_{D,\eta} dD \\ + \int_{T_l}^{\infty} r_i \left(h_i' - s'\right) f_{D,\eta} dD \Bigg\} d\eta = 0, \\ c_j + \int_\eta \Bigg\{ \sum_{t=j}^{l-1} \int_{T_t}^{T_{t+1}} r_j \left(h_j' - h_t'\right) f_{D,\eta} dD \\ + \int_{T_l}^{\infty} r_j \left(h_j' - s'\right) f_{D,\eta} dD \Bigg\} d\eta = 0, \\ c_l + \int_\eta \left\{ \int_{T_l}^{\infty} r_l \left(h_l' - s'\right) f_{D,\eta} dD \right\} d\eta = 0. \end{aligned} \tag{12}$$

Rearranging these terms yields the result. □

Thus, this result shows how to allocate the various risks to each supplier by the new function $g_n(Q_1, Q_2, ..., Q_n)$. As $g_n(Q_1, Q_2, ..., Q_n) > 0$ and $h_j' > h_i'$, we can get the following relationship between two consecutive active contracts $i, j, (j > i)$:

$$\frac{c_j}{k_j} + E\left[\eta h_j'\right] > \frac{c_i}{k_i} + E\left[\eta h_i'\right]. \tag{13}$$

This means that the adjusted cost of the active option contract, that is, $c_j/k_j + E[\eta h_j']$, has an increasing trend.

3.2. Special Case with Only Demand Uncertainty. In this case, there is only demand risk, and there are no reliable risks from the supplier and without loss of generality, we assume that $r_i = 1$, for all i and m and p_s are constants. Then, we have the following result.

Proposition 4. *For any consecutive active contracts i and j, (j > i), one has the following.*

(1) *If i is not the last active contract, then one has*

$$1 - F_D\left(\sum_{t=1}^{i} Q_i^*\right) = -\frac{c_j - c_i}{h_j' - h_i'}. \tag{14}$$

(2) *If i is the last active contract, then one has*

$$1 - F_D\left(\sum_{t=1}^{i} Q_i^*\right) = -\frac{c_i}{h_i' - s'}. \tag{15}$$

Note that the active contract with exercise price less than the spot market is independent of the market liquidity m, while other active contracts depend on m; we have the following corollary concerning the difference between the optimal policies when the spot market is available or not.

Corollary 5. *Given the spot price p_s, denote the first active contract with exercise price higher than p_s by j^* when the spot market exists and j^o when there is no spot market. Then, $j^* \geq j^o$. Moreover, if $j^* = j^o$, then all the active contracts are the same whenever there is a spot market.*

This corollary implies that it is possible for the buyer to buy from those suppliers with exercise price higher than the spot price when there is market liquidity risk. Furthermore, as compared to the optimal policy without the spot market, the existence of the spot market will induce the buyer to choose those suppliers with higher exercise price and lower reservation price (i.e., $j^* \geq j^o$).

When $j^* = j^o$, if we use $T_i^o = \sum_{t=1}^{i} Q_t$ and $T_i^* = \sum_{t=1}^{i} Q_t^*$ to represent the optimal order up to level from the active supplier i. Denote the last active contract with exercise price lower than p_s as k. Then, when $i \leq k$, $T_i^* = T_i^o$ and when $i > k$, we have

$$\frac{1 - F_D\left(T_i^*\right)}{1 - F_D\left(T_i^o\right)} = \frac{1}{1 - m}. \tag{16}$$

We have the following result concerning how the order up to level from supplier i changes as the market liquidity varies.

Corollary 6. *Consider $T_i^* \leq T_i$ and $\partial T_i^*/\partial m = -(1 - F_D(T_i^o))/f_D(T_i^*)(1 - m)^2 < 0$.*

This result implies that the order up to level from supplier i when there is a spot market will be lower than that when there is no spot market. Moreover, such order up to level is increasing in the market liquidity risk, that is, $1 - m$.

3.3. Special Case with Only Supply Uncertainty and Demand Uncertainty. In this case, the spot price p_s and the market liquidity probability m are deterministic; we can define a series of functions, $\{G_n\}_{n=1}^{N}$, as follows:

$$G_n\left(Q_1, Q_2, \ldots, Q_n\right) = E\left[\eta F_D\left(\eta\left(\sum_{i=1}^{n} k_i Q_i\right), \eta\right)\right]. \tag{17}$$

Then, we have the following property associated with these functions. This property is an application of the stochastic small property and we neglect the proof of this result here.

Lemma 7. *Consider $G_n(\mathbf{Q}) \leq G_{n+1}(\mathbf{Q})$.*

With the assistant of this property, one can establish the following result on the optimal reservation policies at time $t = 0$.

Proposition 8. *For any consecutive active contracts i and j, (j > i) one has the following.*

(1) *If i is not the last active contract, then one has*

$$\overline{\eta} - G_i\left(Q_1^*, Q_2^*, \ldots, Q_i^*\right) = -\frac{c_j/k_j - c_i/k_i}{h_j' - h_i'}. \tag{18}$$

Table 1: Parameters for numerical studies.

Demand D	Spot price p_s	Market liquidity m	Shortage cost s
$N(100,\ 30)$	$N(20,\ 6)$	$N(0.5,\ 0.15)$	30

(2) *If i is the last active contract, then one has*

$$\overline{\eta} - G_i\left(Q_1^*, Q_2^*, \ldots, Q_i^*\right) = \frac{c_i/k_i}{h_i' - s'}. \tag{19}$$

This result, combined with the analysis in Fu et al. [11], also implies that the algorithm to choose the active contracts is as follows: (1) find the contract with the lowest adjusted cost $c_i/\widehat{\eta}k_i$ as the first active contract; (2) given the previous active contract i, find the next active contract as the one that has the largest value of $(c_i/k_i\overline{\eta} - c_j/k_j\overline{\eta})/(h_j' - h_i')$. Moreover, we have the following result.

Corollary 9. *If $k_1 = k_2 = \cdots = k_N$, then, the active contracts are independent of η.*

This result is interesting and shows that the active contracts are independent of the supply uncertainty when the suppliers all have the same yield risk. However, the order up to level from each active contract will depend on the distribution of η.

4. Numerical Analysis

In this section, we numerically study how different risks affect the buyer's optimal reservation quantity. We only focus on the case when there is no supply risk, that is, $r_i = 1$ for $i = 1, 2, \ldots, N$.

We assume that the demand, the spot price, and the market liquidity have joint normal distributions with parameters on their marginal distribution as showed in Table 1.

We consider three contracts with parameters as $(10, 0), (5.3237, 6.7810), (1.1580, 16.7230)$. Thus, the first contract is the wholesale price contract and the other two are option-type contracts. The purpose of this study is twofold. First, we will analyze how the correlation between the spot market price and spot market liquidity affects the optimal policy. Second, we will study the effect of the correlation between the market liquidity and the customer demand. In both cases, we will vary the correlation tested while keeping other correlations as zero.

4.1. Effect of the Correlation between Market Liquidity and Spot Price. Keeping other parameters unchanged, we vary the correlation between the market liquidity and spot price from −0.8 to 0.8. The high correlation implies that the event that the spot price is high and the market is available has a large probability. The corresponding results are listed in Table 2.

From Table 2, we find that the higher the correlation, the larger the quantity reserved by the option-type contract. This implies that when the correlation is high, the buyer will rely more on the long-term contract, and when the correlation is negative, he will rely more on the spot market. This is reasonable as the large correlation implies that the buyer can only buy from spot market at high spot price with a high probability. Moreover, the quantity procured from the wholesale contract is insensitive to such correlation, while the option-type contract with the lowest reservation price and the highest exercise price is the most sensitive to the correlation. This is because those option-type contracts play a role mainly as a risk hedging tool and thus are more sensitive to the risk.

Table 2: Effect of correlation between spot market price and liquidity.

The correlation	Optimal solution			
	Wholesale contract	Option-type contract		
	Q_1^*	Q_2^*	Q_3^*	Total
0.8	85.0244	19.7656	29.4336	49.1992
0.4	85.0244	19.7656	29.4336	49.1992
0	85.0244	19.2285	29.3262	48.5574
−0.4	85.0244	19.2285	28.7891	48.0176
−0.8	85.0244	18.5840	28.7891	47.3731

Table 3: Effect of correlation between market liquidity and demand.

The correlation	Optimal solution			
	Wholesale contract	Option-type contract		
	Q_1^*	Q_2^*	Q_3^*	Total
0.8	85.0244	18.5840	25.7812	44.3652
0.4	85.0244	19.2285	27.5000	46.7285
0	85.0244	19.2285	29.3262	48.5547
−0.4	85.0244	19.2285	31.1534	50.3809
−0.8	85.0244	19.7656	32.4414	52.2070

4.2. Effect of the Correlation between Market Liquidity and Demand. Keeping other parameters unchanged, we vary the correlation between the market liquidity and customer demand from −0.8 to 0.8. The high correlation implies that the event that the customer demand is high and the spot market is unavailable has a high probability. The corresponding results are listed in Table 3.

From Table 3, we find similar effect of the correlation on the sensitivity of the contracts, with the possible same reasons. However, the effect of the correlation on the optimal reservation quantity by the option-type contract is reversed. This is also intuitive as the higher the correlation, the higher the probability that the buyer can use the spot market to satisfy his demand. Thus, this results in the low reservation quantity by the option-type contract, and the buyer will rely more on the spot market procurement.

5. Conclusion

In this paper, we establish a decision making model from the perspective of the firm who will procure from the multiple suppliers and the spot markets. The suppliers are unreliable and provide different types of option-type supply contracts which should be made before demand realization, while the spot market can only be used after demand realization and has both the price and liquidity risks. Thus, compared to existing literature, one major feature of our model is that we take the supply uncertainty and spot market liquidity risk into consideration. From our analysis, we develop the optimal portfolio policies for the firm with conditions to find the qualified suppliers. Specifically, we find that we can define a new series of functions that consist of all the associated risks and then allocate different curves of these functions to the suppliers with different contract structures. We also analyze two special cases of the model to dig out some managerial insights. Moreover, we numerically study how the various risks affect the choice of suppliers and the value of the option contract.

Notations

N:	The number of suppliers
m:	The probability the spot market is illiquid
$\mathbf{Q} = [Q_1, \ldots, Q_N]$:	Q_i is the order quantity from supplier i
$r_i = k_i\eta$:	The random yield for supplier i
(c_i, h_i):	The reservation price and exercise price provided by supplier i
s:	The shortage penalty cost
p:	The spot market price
D:	The stochastic demand
x_i:	The exercise quantity from supplier i when spot market is liquid
$\overline{x}_i$:	The exercise quantity from supplier i when spot market is illiquid.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors thank the editor and anonymous referees for their constructive comments. The first author was partially supported by the National Natural Science Foundation of China (no. 71301023), the Humanities and Social Sciences Foundation of Ministry of Education of China (no. 13YJC630197), and the Basic Research Foundation (Natural Science) of Jiangsu Province (no. BK20130582). The second author was partially supported by the National Natural Science Foundation of China (no. 71171105). The third author was partially supported by the National Natural Science Foundation of China (no. 71001073), the Humanities and Social Sciences Foundation of Ministry of Education of China with Project title "Joint Inventory and Pricing Decision with Payment Delay Trade Credit", and the Distinguished University Young Scholar Program of Guangdong Province (no. Yq2013140).

References

[1] R. Muthukrishnan and J. A. Shulman, *Understanding Supply Chain Risk: A McKinsey Global Survey*, The McKinsey Quarterly, 2006.

[2] C. Haksoz and A. Kadam, "Supply portfolio risk," *The Journal of Operational Risk*, vol. 4, no. 1, pp. 59–77, 2009.

[3] G. Cachon, "Supply chain coordination with contracts," in *Handbook of Operations Research and Management Science: Supply Chain Management*, Kluwer Academic Publishers, Amsterdam, The Netherlands, 2003.

[4] D. J. Wu, P. R. Kleindorfer, and J. E. Zhang, "Optimal bidding and contracting strategies for capital-intensive goods," *European Journal of Operational Research*, vol. 137, no. 3, pp. 657–676, 2002.

[5] P. R. Kleindorfer and D. J. Wu, "Integrating long- and short-term contracting via business-to-business exchanges for capital-intensive industries," *Management Science*, vol. 49, no. 11, pp. 1597–1615, 2003.

[6] D. J. Wu and P. R. Kleindorfer, "Competitive options, supply contracting, and electronic markets," *Management Science*, vol. 51, no. 3, pp. 452–466, 2005.

[7] P. H. Ritchken and C. S. Tapiero, "Contingent claims contr acting for purchasing decisions in inventory management," *Operations Research*, vol. 34, no. 6, pp. 864–870, 1986.

[8] M. G. Mattock, *Optimal Commercial Satellite Leasing Strategies*, RAND, 2002.

[9] J. Schummer and R. V. Vohra, "Auctions for procuring options," *Operations Research*, vol. 51, no. 1, pp. 41–51, 2003.

[10] V. Martínez-de-Albéniz and D. Simchi-Levi, "A portfolio approach to procurement contracts," *Production and Operations Management*, vol. 14, no. 1, pp. 90–114, 2005.

[11] Q. Fu, C. Lee, and C. Teo, "Procurement management using option contracts: random spot price and the portfolio effect," *IIE Transactions*, vol. 42, no. 11, pp. 793–811, 2010.

[12] C. A. Yano and H. L. Lee, "Lot sizing with random yields: a review," *Operations Research*, vol. 43, no. 2, pp. 311–334, 1995.

[13] S. Minner, "Multiple-supplier inventory models in supply chain management: a review," *International Journal of Production Economics*, vol. 81-82, pp. 265–279, 2003.

[14] M. Dada, N. C. Petruzzi, and L. B. Schwarz, "A newsvendor's procurement problem when suppliers are unreliable," *Manufacturing and Service Operations Management*, vol. 9, no. 1, pp. 9–32, 2007.

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 862694, 13 pages
http://dx.doi.org/10.1155/2014/862694

Research Article

Optimal Time-Consistent Investment Strategy for a DC Pension Plan with the Return of Premiums Clauses and Annuity Contracts

De-Lei Sheng and Ximin Rong

Department of Mathematics, Tianjin University, Tianjin 300072, China

Correspondence should be addressed to De-Lei Sheng; tjhbsdl@126.com

Received 3 March 2014; Accepted 1 June 2014; Published 24 June 2014

Academic Editor: Xiang Li

Copyright © 2014 D.-L. Sheng and X. Rong. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Defined contribution and annuity contract are merged into one pension plan to study both accumulation phase and distribution phase, which results in such effects that both phases before and after retirement being "defined". Under the Heston's stochastic volatility model, this paper focuses on mean-variance insurers with the return of premiums clauses to study the optimal time-consistent investment strategy for the DC pension merged with an annuity contract. Both accumulation phase before retirement and distribution phase after retirement are studied. In the time-consistent framework, the extended Hamilton-Jacobi-Bellman equations associated with the optimization problem are established. Applying stochastic optimal control technique, the time-consistent explicit solutions of the optimal strategies and the efficient frontiers are obtained. In addition, numerical analysis illustrates our results and also deepens our knowledge or understanding of the research results.

1. Introduction

Annuity contract and defined contribution are merged into one pension plan to study both accumulation phase and distribution phase, which results in such effects that both phases before and after retirement being "defined," making the defined contribution plans even more portable and of great convenience for insurance companies.

Annuity is any financial contract providing continuing payment with a fixed total amount on fixed time interval which usually can be once a year. A lot of research have been done on annuity plans and have gained many good results. For example, Gao [1] investigated annuity contracts in the optimal investment problem under the constant elasticity of variance model in 2009.

Defined contribution (DC) pension plan is a type of retirement plan in which fixed contributions are paid into an individual account, and then the contributions are invested in a financial market and the returns on the investment (positive or negative) are credited to the individual's account. Only contributions to the account are guaranteed, but the future benefits fluctuate on the basis of investment earnings (referring to Cairns et al. [2], etc.). A pension member contributes a predetermined amount of money as premiums before retirement, which lasts for the whole working period of the pension member. From the moment the member retires, the accumulation phase end and the fund will be distributed monthly as old-age pension. Obviously, the distribution per month is not predetermined; however, it is determined by the whole accumulation fund size.

Different from the defined contribution (DC) pension plans, a defined benefit plan is "defined" in the sense that the benefit formula is defined and known in advance, which is based on the earnings history, tenure of service, and age, rather than depending on individual investment returns directly. Because of the cost of administration being fewer than defined benefit plans and ease of determining the plan sponsor's liability in practice, defined contribution plans have been widespread all over the world as the dominant form of plan in many countries.

For the reason that the retirement benefits of the DC pension plan depend on the fund size, the insurer must invest on financial markets to increase the returns, which results in the optimal investment problem becoming so crucial that lots of interests attracted into this field. Dokuchaev and Yu Zhou [3] studied optimal investment strategies with bounded risks, general utilities, and goal achieving. Blanchet-Scalliet et al. [4] investigated optimal investment decisions when time horizon is uncertain. Faggian and Gozzi [5] developed dynamic programming approach for a family of optimal investment models with vintage capital. Blake et al. [6] study the optimal asset allocation problem for DC pension funds.

In the financial market, the price process of stock is described by Heston's stochastic volatility (SV) model. Heston's stochastic volatility (SV) is a classical stochastic volatility model. In the previous literatures, Heston's SV model was very popular for option pricing; however, there are few literatures about the investment problem for insurers. Kraft [7] began to apply the Heston model to study the portfolio problem By maximizing utility from terminal wealth with respect to a power utility function. Li et al. [8] apply the Heston's SV model to investigate the reinsurance and investment problem under the mean-variance criterion.

Mean-variance criterion is first proposed by Markowitz [9] to investigate portfolio selection. But the optimal strategies under the mean-variance criterion are not time consistent, because the mean-variance criterion lacks the iterated expectation property so that the Bellman's principle of optimality does not hold. However, in many situations time consistency of strategies is a basic requirement for rational decision makers. Recently, many researchers paid much attention to time-inconsistent stochastic control problems and aimed at deriving the optimal time-consistent strategies. In 2010, Bjork and Murgoci [10] studied the general theory of Markovian time inconsistent stochastic control problems. Bjork et al. [11] investigated the portfolio optimization with state-dependent risk aversion in the mean-variance framework in 2012.

Some pension plan members may die early during the accumulation phase so that they have no chance to accept pension distribution after retirement. The DC pension plans must have return of premium clauses to protect the rights of them. With this kind of actuarial clause, the dead member can withdraw the premiums she/he contributes or the premiums accumulated by a predetermined interest rate. For this problem, He and Liang [12] studied the DC pension plan for a mean-variance insurer with the return of premiums clauses in 2013. But they focused on the accumulation phase before retirement while there is no serious research on the other phase after retirement.

As far as we know, there is no literature to study both annuity contract and the DC plan with the return of premiums clauses under Heston's SV models under the mean-variance criterion. In this paper, we study a whole pension plan that the DC pension plan with the return of premiums clauses is merged with annuity contract under the mean-variance criterion to find an optimal time consistent strategy under the Heston's stochastic volatility model which can describe the volatility of risk asset more perfectly. Both accumulation phase before retirement and distribution phase after retirement of pension plan are studied in detail.

This paper proceeds as follows. In Section 2, we formulate the model and introduce the actuarial methods of the DC pension plan with the return of premiums clauses. Section 3 solves the time inconsistent problem in the framework of mean-variance criterion. In Section 4, we give some numerical analysis to demonstrate the results. Section 5 concludes the paper.

2. Formulation of the Model

In this paper, the defined contribution in the accumulation phase before retirement is merged with an annuity in the distribution phase after retirement to make one whole pension plan. The contributions are invested in a financial market, which consists of one risk-free asset and a stock, to increase revenues. We try to find the optimal time-consistent investment policy of the DC pension fund for a mean-variance insurer with the return of premiums clauses during the accumulation phase and the benefits of pension fund paid by the form of annuities in distribution phase.

2.1. The Financial Market. Throughout this paper, $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{0\le t\le T})$ denotes a complete probability space satisfying the usual condition, where $T > 0$ is a finite constant representing the investment time horizon; $\mathcal{F}_t$ stands for the information available until time t.

The price $S_0(t)$ of bonds is given by

$$\mathrm{d}S^0(t) = rS^0(t)\,\mathrm{d}t, \quad S^0(0) = 1. \tag{1}$$

The price $S(t)$ of equities obeys the Heston's stochastic volatility model:

$$\begin{aligned} \mathrm{d}S(t) &= S(t)\left[(r + \lambda L(t))\,\mathrm{d}t + \sqrt{L(t)}\mathrm{d}W_1(t)\right], \\ &\qquad S(0) = s_0, \\ \mathrm{d}L(t) &= k(\theta - L(t))\,\mathrm{d}t + \sigma\sqrt{L(t)}\mathrm{d}W_2(t), \\ &\qquad L(0) = l_0, \end{aligned} \tag{2}$$

where $r > 0$ is the risk-free interest rate and $\lambda, k, \theta, \sigma$ are positive constants and the two Brownian motions satisfying $\mathrm{E}[W_1(t), W_2(t)] = \rho \cdot t$, $\rho \in [-1, 1]$ are the correlation coefficients of $W_1(t), W_2(t)$.

2.2. The Accumulation Phase before Retirement. During the period before retirement, the contributions are invested in a risk-free asset and a stock to maximize the pension fund size at retirement. Let $X(t)$ denote the pension wealth at time $t \in [0, T]$, inspired by He and Liang [12], and we formulate the model of DC pension fund with the return of premiums clauses for Heston's stochastic volatility model as follows.

For the convenience of expression, Let us do some symbol descriptions first.

P denotes the premium per unit time, which is a predetermined variable;

ω_0 denotes the accumulation period starting age;

T is the time length; that is, $\omega_0 + T$ is the end age of the pension fund accumulation period;

$\delta_{(1/n),\omega_0+t}$ denotes the mortality rate from time t to time $t + (1/n)$;

tP is the accumulated premium at time t;

$tP\delta_{(1/n),\omega_0+t}$ is the premium returned to the dead member from time t to time $t + (1/n)$.

The plan members who die early can withdraw the premiums she/he contributes or the premiums accumulated at a predetermined interest rate, which is the actuarial return of premiums clause.

To guarantee the interests of pension members, the pension management must invest in equities and bonds to increase the size of pension fund during the accumulation phase.

π is the proportion allocated in the equities, which is the control variable;

$1 - \pi$ is the remaining allocated in bonds.

First, we formulate the fund size $X(t)$ as a differential form. Taking the time interval by $[t, t + (1/n)]$,

$$\begin{aligned} X\left(t + \frac{1}{n}\right) &= X(t)\left(\pi \frac{S_{t+(1/n)}}{S_t} + (1-\pi)\frac{S^0_{t+(1/n)}}{S^0_t}\right) \\ &\quad + P\frac{1}{n} - Pt\delta_{(1/n),\omega_0+t} \\ &= X(t)\left(1 + \pi\frac{S_{t+(1/n)} - S_t}{S_t} + (1-\pi)\frac{S^0_{t+(1/n)} - S^0_t}{S^0_t}\right) \\ &\quad + P\frac{1}{n} - Pt\delta_{(1/n),\omega_0+t} \\ &= X(t)\left(1 + \Delta_{t,(1/n)}\right) + P\frac{1}{n} - Pt\delta_{(1/n),\omega_0+t}, \end{aligned} \tag{3}$$

where $\Delta_{t,(1/n)} = \pi((S_{t+(1/n)} - S_t)/S_t) + (1-\pi)((S^0_{t+(1/n)} - S^0_t)/S^0_t)$.

Using the actuarial formulas to simplify (3), the force function of mortality denoted by $\mu(t)$ and the conditional death probability satisfies

$$\delta_{t,y} = 1 - p_{t,y} = 1 - e^{-\int_0^t \mu(y+s)\mathrm{d}s}. \tag{4}$$

So

$$\delta_{(1/n),\omega_0+t} = 1 - e^{-\int_0^{(1/n)} \mu(\omega_0+t+s)\mathrm{d}s} = \mu(\omega_0 + t)\frac{1}{n} + o\left(\frac{1}{n}\right), \tag{5}$$

as $n \to \infty$, and $\mu(\omega_0 + t)$ is small during the accumulation phase of the pension plan. Thus

$$\begin{aligned} X\left(t + \frac{1}{n}\right) &= X(t)\left(1 + \Delta_{t,(1/n)}\right) + P\frac{1}{n} - Pt\delta_{(1/n),\omega_0+t} \\ &= X(t)\left(1 + \Delta_{t,(1/n)}\right) + P\frac{1}{n} - Pt\mu(\omega_0 + t)\frac{1}{n} \\ &\quad + o\left(\frac{1}{n}\right), \end{aligned} \tag{6}$$

$$\begin{aligned} X\left(t + \frac{1}{n}\right) - X(t) &= X(t)\Delta_{t,(1/n)} + P\frac{1}{n} \\ &\quad - Pt\mu(\omega_0 + t)\frac{1}{n} + o\left(\frac{1}{n}\right). \end{aligned} \tag{7}$$

Let $n \to \infty$, and then

$$\begin{gathered} \frac{S_{t+(1/n)} - S_t}{S_t} \longrightarrow \frac{\mathrm{d}S_t}{S_t}, \\ \frac{S^0_{t+(1/n)} - S^0_t}{S^0_t} \longrightarrow \frac{\mathrm{d}S^0_t}{S^0_t}, \\ \Delta_{t,(1/n)} = \pi\frac{\mathrm{d}S_t}{S_t} + (1-\pi)\frac{\mathrm{d}S^0_t}{S^0_t}. \end{gathered} \tag{8}$$

And (7) becomes

$$\begin{aligned} \mathrm{d}X(t) &= X(t)\left[\pi\frac{\mathrm{d}S_t}{S_t} + (1-\pi)\frac{\mathrm{d}S^0_t}{S^0_t}\right] \\ &\quad + P\mathrm{d}t - Pt\mu(\omega_0 + t)\mathrm{d}t. \end{aligned} \tag{9}$$

Plugging (1) and (2) into (9)

$$\begin{aligned} \mathrm{d}X(t) &= X(t)\left[\pi\left((r + \lambda L(t))\mathrm{d}t + \sqrt{L(t)}\mathrm{d}W_1(t)\right) + (1-\pi)r\mathrm{d}t\right] \\ &\quad + P\mathrm{d}t - Pt\mu(\omega_0 + t)\mathrm{d}t \\ &= \left[X(t)(r + \pi\lambda L(t)) + P\left(1 - t\mu(\omega_0 + t)\right)\right]\mathrm{d}t \\ &\quad + X(t)\pi\sqrt{L(t)}\mathrm{d}W_1(t). \end{aligned} \tag{10}$$

If we choose the mortality force function as the following form:

$$\mu(t) = \frac{1}{\omega - t}, \quad 0 \le t < \omega, \tag{11}$$

where ω is the maximal age of the life table. Then the SDE (10) becomes

$$\begin{aligned} \mathrm{d}X(t) &= \left[X(t)(r + \pi\lambda L(t)) + P\frac{\omega - \omega_0 - 2t}{\omega - \omega_0 - t}\right]\mathrm{d}t \\ &\quad + X(t)\pi\sqrt{L(t)}\mathrm{d}W_1(t), \quad 0 \le t \le T. \end{aligned} \tag{12}$$

The pension management's optimization problem could be described as follows:

$$\sup_{\pi\in\prod} \{\mathrm{E}_{t,x,l}X^{\pi}(T) - \mathrm{Var}_{t,x,l}X^{\pi}(T)\}, \tag{13}$$

where $\prod = \{\pi \mid \pi \in [0,\infty)\}$, which means that a short sell of the bonds is permitted.

2.3. The Distribution Phase after Retirement. Inspired by Gao [1], the whole accumulation fund will purchase a paid-up annuity at retirement time $t = T$ and the purchase rate of annuity will calculate on a predetermined interest rate. The part of the fund used to purchase an annuity of N periods is denoted as D, where $D \le X(T)$. The surplus at the end of the fixed period can be used again in a similar way or paid back to the participants. The contributions benefit to pay between T and $T+N$ are given by

$$\zeta = \frac{D}{\overline{a}_{\overline{N|}}}, \tag{14}$$

where $\overline{a}_{\overline{N|}} = 1 - e^{-\xi N}/\xi$, ξ is a continuous technical rate.

During the period after retirement $t \in [T, T+N]$, the insurer also invests in one risk-free asset and a risk asset. In addition, he has to pay the guaranteed annuity to pension members. The evolution of the pension fund during $[T, T+N]$ is described by the following equation:

$$\mathrm{d}X(t) = [X(t)(r + \pi\lambda L(t)) - \zeta]\,\mathrm{d}t + X(t)\pi\sqrt{L(t)}\mathrm{d}W_1(t). \tag{15}$$

The objective of the optimization problem for a mean-variance pension management could be described as follows:

$$\sup_{\pi\in\prod} \{\mathrm{E}_{t,x,l}[X^{\pi}(T+N)] - \mathrm{Var}_{t,x,l}[X^{\pi}(T+N)]\}, \tag{16}$$

where $\prod = \{\pi \mid \pi \in [0,\infty)\}$, which means that a short sell of the bonds is permitted.

3. The Time Consistent Solution in the Framework of Mean-Variance Criterion

3.1. The Accumulation Phase before Retirement. According to the recent research paper, such as Bjork and Murgoci [10] and so forth, the mean-variance optimal control problem is equivalent to the following Markovian time inconsistent stochastic optimal control problem:

$$\begin{aligned} J(t,x,l,\pi) &= \mathrm{E}_{t,x,l}[X^{\pi}(T)] - \frac{\gamma}{2}\mathrm{Var}_{t,x,l}[X^{\pi}(T)], \\ &= \mathrm{E}_{t,x,l}[X^{\pi}(T)] - \frac{\gamma}{2}\{\mathrm{E}_{t,x,l}[X^{\pi}(T)^2] \\ &\quad - (\mathrm{E}_{t,x,l}[X^{\pi}(T)])^2\}, \\ V(t,x,l) &= \sup_{\pi\in\prod} J(t,x,l,\pi). \end{aligned} \tag{17}$$

Denote

$$\begin{aligned} y^{\pi}(t,x,l) &= \mathrm{E}_{t,x,l}[X^{\pi}(T)], \\ z^{\pi}(t,x,l) &= \mathrm{E}_{t,x,l}[X^{\pi}(T)^2], \end{aligned} \tag{18}$$

and the value function

$$V(t,x,l) = \sup_{\pi\in\prod}\{f(t,x,l,y^{\pi}(t,x,l),z^{\pi}(t,x,l))\}, \tag{19}$$

where

$$f(t,x,l,y,z) = y - \frac{\gamma}{2}(z - y^2). \tag{20}$$

Theorem 1 (verification theorem). *If there exist three real functions $F, G, H : [0,T]\times R\times R \to R$ satisfying the following extended HJB equations:*

$$\begin{aligned} &\sup_{\pi}\Big\{F_t - f_t + (F_x - f_x)\Big[rx + \lambda lx\pi + P\frac{\omega-\omega_0-2t}{\omega-\omega_0-t}\Big] \\ &\quad + (F_l - f_l)k(\theta - l) + \frac{1}{2}(F_{xx} - U^{\pi}_{xx})x^2 l\pi^2 \\ &\quad + \frac{1}{2}(F_{ll} - U^{\pi}_{ll})\sigma^2 l + (F_{xl} - U^{\pi}_{xl})\rho\sigma x l\pi\Big\} = 0, \\ &F(T,x,l) = f(T,x,l,x,x^2), \end{aligned} \tag{21}$$

where

$$\begin{aligned} &G_t + G_x\Big[rx + \lambda lx\pi + P\frac{\omega-\omega_0-2t}{\omega-\omega_0-t}\Big] + k(\theta - l)G_l \\ &\quad + \frac{1}{2}x^2 l\pi^2 G_{xx} + \frac{1}{2}\sigma^2 l G_{ll} + \rho\sigma x l\pi G_{xl} = 0, \\ &G(T,x,l) = x, \end{aligned} \tag{22}$$

$$\begin{aligned} &H_t + H_x\Big[rx + \lambda lx\pi + P\frac{\omega-\omega_0-2t}{\omega-\omega_0-t}\Big] + k(\theta - l)H_l \\ &\quad + \frac{1}{2}x^2 l\pi^2 H_{xx} + \frac{1}{2}\sigma^2 l H_{ll} + \rho\sigma x l\pi H_{xl} = 0, \\ &H(T,x,l) = x^2, \\ &U^{\pi}_{xx} := f_{xx} + 2f_{xy}y^{\pi}_x + 2f_{xz}z^{\pi}_x + f_{yy}(y^{\pi}_x)^2 \\ &\qquad + f_{zz}(z^{\pi}_x)^2 + 2f_{yz}y^{\pi}_x z^{\pi}_x, \\ &U^{\pi}_{ll} := f_{ll} + 2f_{yl}y^{\pi}_l + 2f_{zl}z^{\pi}_l + f_{yy}(y^{\pi}_l)^2 \\ &\qquad + f_{zz}(z^{\pi}_l)^2 + 2f_{yz}y^{\pi}_l z^{\pi}_l, \\ &U^{\pi}_{xl} := f_{xl} + f_{xy}y^{\pi}_l + f_{xz}z^{\pi}_l + f_{yl}y^{\pi}_x + f_{zl}z^{\pi}_x + f_{yy}y^{\pi}_x y^{\pi}_l \\ &\qquad + f_{yz}y^{\pi}_x z^{\pi}_l + f_{yz}y^{\pi}_l z^{\pi}_x + f_{zz}z^{\pi}_x z^{\pi}_l. \end{aligned} \tag{23}$$

Then $V(t,x,l) = F(t,x,l)$, $y^{\pi^*}(t,x,l) = G(t,x,l)$, $z^{\pi^*}(t,x,l) = H(t,x,l)$ for the optimal investment strategy π^*.

Proof. The way to prove the theorem is completely similar to Li et al. [8], so we omit the details here. □

Theorem 2. *For the optimal control problem* (17), *there exist unique optimal time-consistent strategy*

$$\pi^* = \frac{\lambda}{x\gamma e^{r(T-t)}} \cdot \left[1 - \rho\sigma\lambda \frac{\left(1 - e^{(t-T)(k+\lambda\rho\sigma)}\right)}{k + \lambda\rho\sigma}\right] \tag{24}$$

and the optimal value function

$$F(t,x,l) = e^{r(T-t)}x + \frac{B(t)}{\gamma} + \frac{C(t)}{\gamma}, \tag{25}$$

where $B(t)$ and $C(t)$ are given by (40) *and* (41) *explicitly. $\gamma > 0$ denotes the risk aversion coefficient.*

Proof. According to (20), we have

$$\begin{gathered} f_t = f_x = f_l = f_{xx} = f_{xy} = f_{xz} = f_{yz} \\ = f_{xl} = f_{yl} = f_{zl} = f_{zz} = f_{ll} = 0, \\ f_y = 1 + \gamma y, \qquad f_{yy} = \gamma, \qquad f_z = -\frac{\gamma}{2}. \end{gathered} \tag{26}$$

Plugging (26) into $U^{\pi}_{xx}, U^{\pi}_{lx}, U^{\pi}_{ll}$, respectively,

$$U^{\pi^*}_{xx} = \gamma G_x^2, \qquad U^{\pi^*}_{ll} = \gamma G_l^2, \qquad U^{\pi^*}_{xl} = \gamma G_x G_l. \tag{27}$$

According to (21)

$$\begin{aligned} \pi^* &= -\frac{\lambda(F_x - f_x) + \rho\sigma\left(F_{xl} - U^{\pi^*}_{xl}\right)}{x\left(F_{xx} - U^{\pi^*}_{xx}\right)} \\ &= -\frac{\lambda F_x + \rho\sigma(F_{xl} - \gamma G_x G_l)}{x\left(F_{xx} - \gamma G_x^2\right)}. \end{aligned} \tag{28}$$

Equation (21) turns into

$$\begin{aligned} &F_t + F_x\left[rx + P\frac{\omega - \omega_0 - 2t}{\omega - \omega_0 - t}\right] + k(\theta - l)F_l \\ &\quad + \frac{1}{2}\left(F_{ll} - \gamma G_l^2\right)\sigma^2 l - \frac{l}{2}\frac{\left[\lambda F_x + \rho\sigma(F_{xl} - \gamma G_x G_l)\right]^2}{F_{xx} - \gamma G_x^2} = 0, \\ &F(T,x,l) = f\left(T,x,l,x,x^2\right), \end{aligned} \tag{29}$$

and (22) becomes

$$\begin{aligned} &G_t + G_x\left[rx + P\frac{\omega - \omega_0 - 2t}{\omega - \omega_0 - t}\right] + k(\theta - l)G_l + \frac{1}{2}\sigma^2 l G_{ll} \\ &\quad - (\lambda l G_x + \rho\sigma l G_{xl})\left[\frac{\lambda F_x + \rho\sigma(F_{xl} - \gamma G_x G_l)}{F_{xx} - \gamma G_x^2}\right] \\ &\quad + \frac{1}{2} l G_{xx}\left[\frac{\lambda F_x + \rho\sigma(F_{xl} - \gamma G_x G_l)}{F_{xx} - \gamma G_x^2}\right]^2 = 0, \\ &G(T,x,l) = x. \end{aligned} \tag{30}$$

The remainder of this section focuses on solving (29) w.r.t. F and (30) w.r.t. G. Since the two equations are linear in x and l, it is quite natural to conjecture the following forms of $F(t,x,l)$ and $G(t,x,l)$:

$$\begin{gathered} F(t,x,l) = A(t)x + \frac{B(t)}{\gamma}l + \frac{C(t)}{\gamma}, \\ A(T) = 1, \quad B(T) = 0, \quad C(T) = 0, \\ G(t,x,l) = \alpha(t)x + \frac{\beta(t)}{\gamma}l + \frac{\Delta(t)}{\gamma}, \\ \alpha(T) = 1, \quad \beta(T) = 0, \quad \Delta(T) = 0, \end{gathered} \tag{31}$$

and the corresponding partial derivatives are

$$\begin{aligned} &F_t = A_t x + \frac{B_t}{\gamma}l + \frac{C_t}{\gamma}, \qquad F_x = A(t), \qquad F_l = \frac{B(t)}{\gamma}, \\ &F_{xx} = 0, \qquad F_{ll} = 0, \qquad F_{xl} = 0, \\ &G_t = \alpha_t x + \frac{\beta_t}{\gamma}l + \frac{\Delta_t}{\gamma}, \qquad G_x = \alpha(t), \qquad G_l = \frac{\beta(t)}{\gamma}, \\ &G_{xx} = 0, \qquad G_{ll} = 0, \qquad G_{xl} = 0. \end{aligned} \tag{32}$$

Plugging the above partial derivatives (32) correspondingly into (29), (30), and (28), we obtain

$$\begin{aligned} &A_t x + \frac{B_t}{\gamma}l + \frac{C_t}{\gamma} + A(t)\left[rx + P\frac{\omega - \omega_0 - 2t}{\omega - \omega_0 - t}\right] \\ &\quad + k(\theta - l)\frac{B(t)}{\gamma} - \frac{\sigma^2 l}{2\gamma}\beta^2(t) \\ &\quad + \frac{l\left[\lambda A(t) - \rho\sigma\alpha(t)\beta(t)\right]^2}{2\gamma\alpha^2(t)} = 0, \\ &A(T) = 1, \quad B(T) = C(T) = 0, \end{aligned} \tag{33}$$

$$\begin{aligned} &\alpha_t x + \frac{\beta_t}{\gamma}l + \frac{\Delta_t}{\gamma} + \alpha(t)\left[rx + P\frac{\omega - \omega_0 - 2t}{\omega - \omega_0 - t}\right] \\ &\quad + k(\theta - l)\frac{\beta(t)}{\gamma} + \frac{\lambda l\left[\lambda A(t) - \rho\sigma\alpha(t)\beta(t)\right]}{\gamma\alpha(t)} = 0, \\ &\qquad \alpha(T) = 1, \quad \beta(T) = \Delta(T) = 0, \end{aligned} \tag{34}$$

and the optimal strategy

$$\begin{aligned} \pi^* &= -\frac{\lambda F_x + \rho\sigma(-\gamma G_x G_l)}{x\left(-\gamma G_x^2\right)} = \frac{\lambda F_x - \rho\sigma\gamma G_x G_l}{x\gamma G_x^2} \\ &= \frac{\lambda A(t)}{x\gamma\alpha^2(t)} - \frac{\rho\sigma\beta(t)}{x\gamma\alpha(t)}. \end{aligned} \tag{35}$$

Equation (33) splits into three equations:

$$
\begin{aligned}
&[A_t + rA(t)]\,x = 0, \quad A(T) = 1, \\
&\left[\frac{B_t}{\gamma} - \frac{kB(t)}{\gamma} + \frac{\lambda^2}{2\gamma}\frac{A^2(t)}{\alpha^2(t)} - \frac{\sigma^2\beta^2(t)}{2\gamma} + \frac{\rho^2\sigma^2\beta^2(t)}{2\gamma}\right. \\
&\left.\quad - \frac{\rho\sigma\lambda A(t)\beta(t)}{\gamma\alpha(t)}\right] l = 0, \quad B(T) = 0, \\
&\frac{C_t}{\gamma} + A(t)\,P\frac{\omega - \omega_0 - 2t}{\omega - \omega_0 - t} + \frac{k\theta B(t)}{\gamma} = 0, \quad C(T) = 0.
\end{aligned}
\tag{36}
$$

Equation (34) splits into three equations:

$$
\begin{aligned}
&[\alpha_t + r\alpha(t)]\,x = 0, \quad \alpha(T) = 1, \\
&\left[\frac{\beta_t}{\gamma} - \frac{k\beta(t)}{\gamma} + \frac{\lambda^2}{\gamma}\frac{A(t)}{\alpha(t)} - \frac{\rho\sigma\lambda}{\gamma}\beta(t)\right] l = 0, \quad \beta(T) = 0, \\
&\frac{\Delta_t}{\gamma} + \alpha(t)\,P\frac{\omega - \omega_0 - 2t}{\omega - \omega_0 - t} + \frac{k\theta\beta(t)}{\gamma} = 0, \quad \Delta(T) = 0.
\end{aligned}
\tag{37}
$$

To solve the above equations, we have

$$A(t) = \alpha(t) = e^{r(T-t)}, \tag{38}$$

$$\beta(t) = \frac{\left(1 - e^{(t-T)(k+\lambda\rho\sigma)}\right)\lambda^2}{k + \lambda\rho\sigma}, \tag{39}$$

$$
\begin{aligned}
&B(t) \\
&= \left[2k\rho(k+\lambda\rho\sigma)^2(k+2\lambda\rho\sigma)\right]^{-1} \cdot e^{-kT}\lambda^2 \\
&\quad \cdot \left[-e^{k(2t-T)+2(t-T)\lambda\rho\sigma}k\lambda^2\rho\left(-1+\rho^2\right)\sigma^2\right. \\
&\quad + e^{kt}(k\rho + 2\lambda\sigma)(k+\lambda\rho\sigma)^2 \\
&\quad + e^{kT}\rho(k-\lambda\sigma)(k+\lambda\sigma)(k+2\lambda\rho\sigma) \\
&\quad \left.- 2e^{kt+(t-T)\lambda\rho\sigma}k(k\rho+\lambda\sigma)(k+2\lambda\rho\sigma)\right],
\end{aligned}
\tag{40}
$$

$$
\begin{aligned}
&C(t) \\
&= \frac{\left(-1+e^{r(-t+T)}\right)P\gamma}{r} + \left(2\rho(k+\lambda\rho\sigma)^2(k+2\lambda\rho\sigma)\right)^{-1} \\
&\quad \times e^{-kT}\theta\lambda^2 \\
&\quad \cdot \left[-\frac{\left(e^{kT} - e^{k(2t-T)+2(t-T)\lambda\rho\sigma}\right)k\lambda^2\rho\left(-1+\rho^2\right)\sigma^2}{2(k+\lambda\rho\sigma)}\right. \\
&\quad + \frac{\left(-e^{kt}+e^{kT}\right)(k\rho+2\lambda\sigma)(k+\lambda\rho\sigma)^2}{k} \\
&\quad + e^{kT}(-t+T)\rho(k-\lambda\sigma)(k+\lambda\sigma)(k+2\lambda\rho\sigma) \\
&\quad \left.+ \frac{2\left(-e^{kT}+e^{kt+(t-T)\lambda\rho\sigma}\right)k(k\rho+\lambda\sigma)(k+2\lambda\rho\sigma)}{k+\lambda\rho\sigma}\right] \\
&\quad - \int_t^T \gamma p\frac{\tau}{\omega-\omega_0-\tau}e^{r(T-\tau)}\mathrm{d}\tau,
\end{aligned}
\tag{41}
$$

$$
\begin{aligned}
&\Delta(t) \\
&= \frac{\left(-1+e^{r(-t+T)}\right)P\gamma}{r} - \int_t^T \gamma p\frac{\tau}{\omega-\omega_0-\tau}e^{r(T-\tau)}\mathrm{d}\tau \\
&\quad + \frac{k\theta\lambda^2\left(-1+e^{(t-T)(k+\lambda\rho\sigma)} + (-t+T)(k+\lambda\rho\sigma)\right)}{(k+\lambda\rho\sigma)^2}.
\end{aligned}
\tag{42}
$$

After some simple calculations, the optimal investment strategy (35) becomes

$$
\begin{aligned}
\pi^* &= \frac{1}{x\gamma}\cdot[\alpha(t)]^{-1}\cdot[\lambda - \rho\sigma\beta(t)] \\
&= \frac{\lambda}{x\gamma e^{r(T-t)}}\cdot\left[1 - \rho\sigma\lambda\frac{\left(1-e^{(t-T)(k+\lambda\rho\sigma)}\right)}{k+\lambda\rho\sigma}\right], \\
G(t,x,l) &= e^{r(T-t)}x + \frac{\left(1-e^{(t-T)(k+\lambda\rho\sigma)}\right)\lambda^2}{\gamma(k+\lambda\rho\sigma)}l + \frac{\Delta(t)}{\gamma}, \\
F(t,x,l) &= e^{r(T-t)}x + \frac{B(t)}{\gamma}l + \frac{C(t)}{\gamma}.
\end{aligned}
\tag{43}
$$

□

Theorems 1 and 2 imply that

$$
\begin{aligned}
F(t,x,l) &= V(t,x,l) = f\left(t,x,l,y^{\pi^*},z^{\pi^*}\right) \\
&= \mathrm{E}_{t,x,l}\left[X^\pi(T)\right] - \frac{\gamma}{2}\left\{\mathrm{E}_{t,x,l}\left[X^\pi(T)^2\right]\right. \\
&\qquad \left.- \left(\mathrm{E}_{t,x,l}\left[X^\pi(T)\right]\right)^2\right\} \\
&= G(t,x,l) - \frac{\gamma}{2}\left[H(t,x,l) - G(t,x,l)^2\right],
\end{aligned}
\tag{44}
$$

$$\mathrm{E}_{t,x,l}\left[X^{\pi^*}(T)\right] = G(t,x,l). \tag{45}$$

Since

$$\mathrm{Var}_{t,x,l}\left[X^{\pi^*}(T)\right] = \mathrm{E}_{t,x,l}\left[X^{\pi^*}(T)\right]^2 - \left(\mathrm{E}_{t,x,l}X^{\pi^*}(T)\right)^2, \tag{46}$$

(44) and (46) imply that

$$\begin{aligned}
&\mathrm{Var}_{t,x,l}\left[X^{\pi^*}(T)\right] \\
&\quad= \frac{2}{\gamma}\left[G(t,x,l) - F(t,x,l)\right] \\
&\quad= \frac{2}{\gamma^2}\left[(\beta(t) - B(t))\,l + (\Delta(t) - C(t))\right] \\
&\quad= \left(2k\gamma^2\rho(k+\lambda\rho\sigma)^3(k+2\lambda\rho\sigma)\right)^{-1} e^{-kT}\lambda^2\cdot\Upsilon,
\end{aligned} \tag{47}$$

where

$$\begin{aligned}
\Upsilon = \Big\{&-2e^{kt}(l-\theta)(k\rho+2\lambda\sigma)(k+\lambda\rho\sigma)^3 \\
&-4e^{kt+(t-T)\lambda\rho\sigma}k\lambda\left(-1+\rho^2\right)\sigma(k+2\lambda\rho\sigma) \\
&\times(k(l-\theta)+l\lambda\rho\sigma) \\
&+e^{k(2t-T)+2(t-T)\lambda\rho\sigma}k\lambda^2\rho\left(-1+\rho^2\right)\sigma^2 \\
&\times(2kl-k\theta+2l\lambda\rho\sigma) \\
&+e^{kT}\rho(k+2\lambda\rho\sigma) \\
&\quad\times\Big[2k^3(l+(-1-kt+kT)\theta) \\
&\qquad+6k^2(l+(-1-kt+kT)\theta)\lambda\rho\sigma \\
&\qquad+k\lambda^2\Big(l\left(2+4\rho^2\right) \\
&\qquad\qquad-\theta\left(5+\rho^2+2k(t-T)\left(1+2\rho^2\right)\right)\Big)\sigma^2 \\
&\qquad+2(l+(-1-kt+kT)\theta)\lambda^3\rho\sigma^3\Big]\Big\}.
\end{aligned} \tag{48}$$

Putting (45) and (47) together, the efficient frontier is rewritten as

$$\begin{aligned}
&\mathrm{E}_{t,x,l}\left[X^{\pi^*}(T)\right] \\
&\quad= e^{r(T-t)}x + \frac{\left(1-e^{(t-T)(k+\lambda\rho\sigma)}\right)\lambda^2}{\gamma(k+\lambda\rho\sigma)}l + \frac{\Delta(t)}{\gamma} \\
&\qquad+\gamma\sqrt{\mathrm{Var}_{t,x,l}\left[X^{\pi^*}(T)\right]} \\
&\qquad-\sqrt{2\left[(\beta(t)-B(t))\,l+(\Delta(t)-C(t))\right]},
\end{aligned} \tag{49}$$

where (t), $B(t)$, $C(t)$, and $\Delta(t)$ are given by (39), (40), (41), and (42), respectively.

3.2. The Distribution Phase after Retirement. According to the recent research paper, such as Bjork and Murgoci [10] and so forth, the mean-variance optimal control problem in the distribution phase is equivalent to the following Markovian time inconsistent stochastic optimal control problem:

$$\begin{aligned}
&\widehat{J}(t,x,l,\pi) \\
&\quad= \mathrm{E}_{t,x,l}\left[X^{\pi}(T+N)\right] - \frac{\gamma}{2}\mathrm{Var}_{t,x,l}\left[X^{\pi}(T+N)\right], \\
&\quad= \mathrm{E}_{t,x,l}\left[X^{\pi}(T+N)\right] - \frac{\gamma}{2}\Big\{\mathrm{E}_{t,x,l}\left[X^{\pi}(T+N)^2\right] \\
&\qquad\qquad-\left(\mathrm{E}_{t,x,l}\left[X^{\pi}(T+N)\right]\right)^2\Big\}, \\
&\widehat{V}(t,x,l) = \sup_{\pi\in\prod}\widehat{J}(t,x,l,\pi).
\end{aligned} \tag{50}$$

Denote

$$\begin{aligned}
y^{\pi}(t,x,l) &= \mathrm{E}_{t,x,l}\left[X^{\pi}(T+N)\right], \\
z^{\pi}(t,x,l) &= \mathrm{E}_{t,x,l}\left[X^{\pi}(T+N)^2\right],
\end{aligned} \tag{51}$$

and the value function

$$\widehat{V}(t,x,l) = \sup_{\pi\in\prod}\left\{\widehat{f}\left(t,x,l,y^{\pi}(t,x,l),z^{\pi}(t,x,l)\right)\right\}, \tag{52}$$

where

$$\widehat{f}(t,x,l,y,z) = y - \frac{\gamma}{2}\left(z-y^2\right). \tag{53}$$

Theorem 3 (verification theorem). *If there exist three real functions $\widehat{F},\widehat{G},\widehat{H} : [0,T]\times R\times R \rightarrow R$ satisfying the following extended HJB equations:*

$$\begin{aligned}
&\sup_{\pi}\Big\{\widehat{F}_t - \widehat{f}_t + \left(\widehat{F}_x - \widehat{f}_x\right)\left[rx+\lambda lx\pi-\zeta\right] \\
&\qquad+\left(\widehat{F}_l-\widehat{f}_l\right)k(\theta-l)+\frac{1}{2}\left(\widehat{F}_{xx}-\widehat{U}^{\pi}_{xx}\right)x^2l\pi^2 \\
&\qquad+\frac{1}{2}\left(\widehat{F}_{ll}-\widehat{U}^{\pi}_{ll}\right)\sigma^2l+\left(\widehat{F}_{xl}-\widehat{U}^{\pi}_{xl}\right)\rho\sigma xl\pi\Big\} = 0, \\
&\widehat{F}(T+N,x,l) = \widehat{f}\left(T+N,x,l,x,x^2\right),
\end{aligned} \tag{54}$$

where

$$\begin{aligned}
&\widehat{G}_t + \widehat{G}_x\left[rx+\lambda lx\pi-\zeta\right]+k(\theta-l)\,\widehat{G}_l \\
&\qquad+\frac{1}{2}x^2l\pi^2\widehat{G}_{xx}+\frac{1}{2}\sigma^2l\widehat{G}_{ll}+\rho\sigma xl\pi\widehat{G}_{xl} = 0, \\
&\widehat{G}(T+N,x,l) = x, \\
&\widehat{H}_t + \widehat{H}_x\left[rx+\lambda lx\pi-\zeta\right]+k(\theta-l)\,\widehat{H}_l \\
&\qquad+\frac{1}{2}x^2l\pi^2\widehat{H}_{xx}+\frac{1}{2}\sigma^2l\widehat{H}_{ll}+\rho\sigma xl\pi\widehat{H}_{xl} = 0, \\
&\widehat{H}(T+N,x,l) = x^2, \\
&\widehat{U}^{\pi}_{xx} := \widehat{f}_{xx}+2\widehat{f}_{xy}y^{\pi}_x+2\widehat{f}_{xz}z^{\pi}_x \\
&\qquad+\widehat{f}_{yy}\left(y^{\pi}_x\right)^2+\widehat{f}_{zz}\left(z^{\pi}_x\right)^2+2\widehat{f}_{yz}y^{\pi}_xz^{\pi}_x,
\end{aligned} \tag{55}$$

$$\begin{aligned}\widehat{U}_{ll}^{\pi} &:= \widehat{f}_{ll} + 2\widehat{f}_{yl}y_l^{\pi} + 2\widehat{f}_{zl}z_l^{\pi} \\ &\quad + \widehat{f}_{yy}(y_l^{\pi})^2 + \widehat{f}_{zz}(z_l^{\pi})^2 + 2\widehat{f}_{yz}y_l^{\pi}z_l^{\pi}, \\ \widehat{U}_{xl}^{\pi} &:= \widehat{f}_{xl} + \widehat{f}_{xy}y_l^{\pi} + \widehat{f}_{xz}z_l^{\pi} + \widehat{f}_{yl}y_x^{\pi} + \widehat{f}_{zl}z_x^{\pi} \\ &\quad + \widehat{f}_{yy}y_x^{\pi}y_l^{\pi} + \widehat{f}_{yz}y_x^{\pi}z_l^{\pi} + \widehat{f}_{yz}y_l^{\pi}z_x^{\pi} + \widehat{f}_{zz}z_x^{\pi}z_l^{\pi}.\end{aligned} \tag{56}$$

Then $\widehat{V}(t,x,l) = \widehat{F}(t,x,l)$, $y^{\pi^*}(t,x,l) = \widehat{G}(t,x,l)$, $z^{\pi^*}(t,x,l) = \widehat{H}(t,x,l)$ *for the optimal investment strategy* π^*.

Proof. The way to prove the theorem is completely similar to Li et al. [8], so we omit the details here. □

Theorem 4. *For the optimal control problem* (50)*, there exist unique optimal time-consistent strategy*

$$\pi^* = \frac{\lambda}{x\gamma e^{r(N+T-t)}} \cdot \left[1 - \rho\sigma\lambda \frac{\left(1 - e^{-(N-t+T)(k+\lambda\rho\sigma)}\right)}{k + \lambda\rho\sigma}\right] \tag{57}$$

and the optimal value function

$$\widehat{F}(t,x,l) = e^{r(N+T-t)}x + \frac{\widehat{B}(t)}{\gamma} + \frac{\widehat{C}(t)}{\gamma}, \tag{58}$$

where $\widehat{B}(t)$ *and* $\widehat{C}(t)$ *are given by* (72) *and* (74)*, explicitly.* $\gamma > 0$ *denotes the risk aversion coefficient.*

Proof. According to (53), we have

$$\begin{aligned}\widehat{f}_t &= \widehat{f}_x = \widehat{f}_l = \widehat{f}_{xx} = \widehat{f}_{xy} = \widehat{f}_{xz} = \widehat{f}_{yz} \\ &= \widehat{f}_{xl} = \widehat{f}_{yl} = \widehat{f}_{zl} = \widehat{f}_{zz} = \widehat{f}_{ll} = 0, \\ \widehat{f}_y &= 1 + \gamma y, \qquad \widehat{f}_{yy} = \gamma, \qquad \widehat{f}_z = -\frac{\gamma}{2}.\end{aligned} \tag{59}$$

Substituting (59) into $\widehat{U}_{xx}^{\pi}, \widehat{U}_{lx}^{\pi}, \widehat{U}_{ll}^{\pi}$

$$\widehat{U}_{xx}^{\pi^*} = \gamma\widehat{G}_x^2, \qquad \widehat{U}_{ll}^{\pi^*} = \gamma\widehat{G}_l^2, \qquad \widehat{U}_{xl}^{\pi^*} = \gamma\widehat{G}_x\widehat{G}_l. \tag{60}$$

Taking derivative for (54) with respect to π, according to the first-order necessary condition, we have

$$\left(\widehat{F}_x - \widehat{f}_x\right)\lambda lx + \left(\widehat{F}_{xx} - \widehat{U}_{xx}^{\pi}\right)x^2 l\pi + \left(\widehat{F}_{xl} - \widehat{U}_{xl}^{\pi}\right)\rho\sigma xl = 0, \tag{61}$$

so

$$\pi^* = -\frac{\left(\widehat{F}_x - \widehat{f}_x\right)\lambda + \left(\widehat{F}_{xl} - \widehat{U}_{xl}^{\pi}\right)\rho\sigma}{\left(\widehat{F}_{xx} - \widehat{U}_{xx}^{\pi}\right)x} = \frac{\lambda\widehat{F}_x - \rho\sigma\gamma\widehat{G}_x\widehat{G}_l}{\gamma x\widehat{G}_x^2}. \tag{62}$$

The remainder of this section focuses on solving (54) w.r.t. $\widehat{F}$ and (55) w.r.t. $\widehat{G}$. Since the two equations are linear in x and l, it is quite natural to conjecture the following forms of $\widehat{F}(t,x,l)$ and $\widehat{G}(t,x,l)$:

$$\begin{aligned}&\widehat{F}(t,x,l) = \widehat{A}(t)x + \frac{\widehat{B}(t)}{\gamma}l + \frac{\widehat{C}(t)}{\gamma}, \\ &\widehat{A}(T+N) = 1, \qquad \widehat{B}(T+N) = 0, \qquad \widehat{C}(T+N) = 0, \\ &\widehat{G}(t,x,l) = \widehat{\alpha}(t)x + \frac{\widehat{\beta}(t)}{\gamma}l + \frac{\widehat{\Delta}(t)}{\gamma}, \\ &\widehat{\alpha}(T+N) = 1, \qquad \widehat{\beta}(T+N) = 0, \qquad \widehat{\Delta}(T+N) = 0,\end{aligned} \tag{63}$$

and the corresponding partial derivatives are

$$\begin{aligned}&\widehat{F}_t = \widehat{A}_t x + \frac{\widehat{B}_t}{\gamma}l + \frac{\widehat{C}_t}{\gamma}, \qquad \widehat{F}_x = \widehat{A}(t), \\ &\widehat{F}_l = \frac{\widehat{B}(t)}{\gamma}, \qquad \widehat{F}_{xx} = 0, \qquad \widehat{F}_{ll} = 0, \qquad \widehat{F}_{xl} = 0, \\ &\widehat{G}_t = \widehat{\alpha}_t x + \frac{\widehat{\beta}_t}{\gamma}l + \frac{\widehat{\Delta}_t}{\gamma}, \qquad \widehat{G}_x = \widehat{\alpha}(t), \\ &\widehat{G}_l = \frac{\widehat{\beta}(t)}{\gamma}, \qquad \widehat{G}_{xx} = 0, \qquad \widehat{G}_{ll} = 0, \qquad \widehat{G}_{xl} = 0.\end{aligned} \tag{64}$$

Plugging the above partial derivatives (64) correspondingly into (54), (55), and (62), we obtain

$$\begin{aligned}&\widehat{A}_t x + \frac{\widehat{B}_t}{\gamma}l + \frac{\widehat{C}_t}{\gamma} + \widehat{A}(t)\left[rx - \zeta\right] + k(\theta - l)\frac{\widehat{B}(t)}{\gamma} \\ &\quad - \frac{\sigma^2 l}{2\gamma}\widehat{\beta}^2(t) + \frac{l\left[\lambda\widehat{A}(t) - \rho\sigma\widehat{\alpha}(t)\widehat{\beta}(t)\right]^2}{2\gamma\widehat{\alpha}^2(t)} = 0, \\ &\widehat{A}(T+N) = 1, \qquad \widehat{B}(T+N) = \widehat{C}(T+N) = 0,\end{aligned} \tag{65}$$

$$\begin{aligned}&\widehat{\alpha}_t x + \frac{\widehat{\beta}_t}{\gamma}l + \frac{\widehat{\Delta}_t}{\gamma} + \widehat{\alpha}(t)\left[rx - \zeta\right] + k(\theta - l)\frac{\widehat{\beta}(t)}{\gamma} \\ &\quad + \frac{\lambda l\left[\lambda\widehat{A}(t) - \rho\sigma\widehat{\alpha}(t)\widehat{\beta}(t)\right]}{\gamma\widehat{\alpha}(t)} = 0, \\ &\widehat{\alpha}(T+N) = 1, \qquad \widehat{\beta}(T+N) = \widehat{\Delta}(T+N) = 0,\end{aligned} \tag{66}$$

and the optimal strategy

$$\begin{aligned}\pi^* &= -\frac{\lambda\widehat{F}_x + \rho\sigma\left(-\gamma\widehat{G}_x\widehat{G}_l\right)}{x\left(-\gamma\widehat{G}_x^2\right)} = \frac{\lambda\widehat{F}_x - \rho\sigma\gamma\widehat{G}_x\widehat{G}_l}{x\gamma\widehat{G}_x^2} \\ &= \frac{\lambda\widehat{A}(t)}{x\gamma\widehat{\alpha}^2(t)} - \frac{\rho\sigma\widehat{\beta}(t)}{x\gamma\widehat{\alpha}(t)}.\end{aligned} \tag{67}$$

Equation (65) splits into three equations

$$\begin{aligned}
&\left[\widehat{A}_t + r\widehat{A}(t)\right] x = 0, \qquad \widehat{A}(T+N) = 1,\\
&\left[\frac{\widehat{B}_t}{\gamma} - \frac{k\widehat{B}(t)}{\gamma} + \frac{\lambda^2}{2\gamma}\frac{\widehat{A}^2(t)}{\widehat{\alpha}^2(t)} - \frac{\sigma^2\widehat{\beta}^2(t)}{2\gamma} + \frac{\rho^2\sigma^2\widehat{\beta}^2(t)}{2\gamma}\right.\\
&\quad \left. - \frac{\rho\sigma\lambda\widehat{A}(t)\widehat{\beta}(t)}{\gamma\widehat{\alpha}(t)}\right] l = 0, \quad \widehat{B}(T+N) = 0,\\
&\frac{\widehat{C}_t}{\gamma} - \zeta\widehat{A}(t) + \frac{k\theta\widehat{B}(t)}{\gamma} = 0, \qquad \widehat{C}(T+N) = 0.
\end{aligned} \tag{68}$$

Equation (66) splits into three equations

$$\begin{aligned}
&\left[\widehat{\alpha}_t + r\widehat{\alpha}(t)\right] x = 0, \qquad \widehat{\alpha}(T+N) = 1,\\
&\left[\frac{\widehat{\beta}_t}{\gamma} - \frac{k\widehat{\beta}(t)}{\gamma} + \frac{\lambda^2}{\gamma}\frac{\widehat{A}(t)}{\widehat{\alpha}(t)} - \frac{\rho\sigma\lambda}{\gamma}\widehat{\beta}(t)\right] l = 0,\\
&\widehat{\beta}(T+N) = 0,\\
&\frac{\widehat{\Delta}_t}{\gamma} - \zeta\widehat{\alpha}(t) + \frac{k\theta\widehat{\beta}(t)}{\gamma} = 0, \qquad \widehat{\Delta}(T+N) = 0.
\end{aligned} \tag{69}$$

To solve the above equations, we have

$$\widehat{A}(t) = \widehat{\alpha}(t) = e^{r(N+T-t)}, \tag{70}$$

$$\widehat{\beta}(t) = \frac{\left(1 - e^{-(N-t+T)(k+\lambda\rho\sigma)}\right)\lambda^2}{k+\lambda\rho\sigma}, \tag{71}$$

$$\begin{aligned}
&\widehat{B}(t)\\
&\quad = \left(2k\rho(k+\lambda\rho\sigma)^2(k+2\lambda\rho\sigma)\right)^{-1}\\
&\qquad \times e^{-k(3N+2t+3T)-3(N+t+T)\lambda\rho\sigma}\lambda^2\\
&\qquad \cdot \left[-e^{k(N+4t+T)+(N+5t+T)\lambda\rho\sigma}k\lambda^2\rho\left(-1+\rho^2\right)\sigma^2\right.\\
&\qquad\quad + e^{k(2N+3t+2T)+3(N+t+T)\lambda\rho\sigma}(k\rho+2\lambda\sigma)\\
&\qquad\quad \times (k+\lambda\rho\sigma)^2\\
&\qquad\quad + e^{k(3N+2t+3T)+3(N+t+T)\lambda\rho\sigma}\rho(k-\lambda\sigma)(k+\lambda\sigma)\\
&\qquad\quad \times (k+2\lambda\rho\sigma) - 2e^{k(2N+3t+2T)+2(N+2t+T)\lambda\rho\sigma}\\
&\qquad\quad \left.\times k(k\rho+\lambda\sigma)(k+2\lambda\rho\sigma)\right],
\end{aligned} \tag{72}$$

$$\begin{aligned}
&\widehat{\Delta}(t)\\
&\quad = \left(r(k+\lambda\rho\sigma)^2\right)^{-1}e^{-rt-(N+T)(k+\lambda\rho\sigma)}\\
&\qquad \cdot \left[e^{t(k+r+\lambda\rho\sigma)}kr\theta\lambda^2 - e^{(N+T)(k+r+\lambda\rho\sigma)}\right.\\
&\qquad\quad \times \gamma\zeta(k+\lambda\rho\sigma)^2 + e^{rt+(N+T)(k+\lambda\rho\sigma)}\\
&\qquad\quad \times \left[k\left(k\gamma\zeta + r(-1+k(N-t+T))\theta\lambda^2\right)\right.\\
&\qquad\qquad + k\lambda\left(2\gamma\zeta + r(N-t+T)\theta\lambda^2\right)\rho\sigma\\
&\qquad\qquad \left.\left. +\gamma\zeta\lambda^2\rho^2\sigma^2\right]\right],
\end{aligned} \tag{73}$$

$$\begin{aligned}
&\widehat{C}(t)\\
&\quad = \left(4kr\rho(k+\lambda\rho\sigma)^3(k+2\lambda\rho\sigma)\right)^{-1}\\
&\qquad \cdot \left\{-2k^5\left(2\left(-1+e^{r(N-t+T)}\right)\gamma\zeta\right.\right.\\
&\qquad\qquad \left. - r(N-t+T)\theta\lambda^2\right)\rho\\
&\qquad\quad + 4e^{-k(N-t+T)}\left(-1+e^{k(N-t+T)}\right)r\theta\lambda^6\rho^3\sigma^4\\
&\qquad\quad + 2k^4\lambda\rho\left[-10\left(-1+e^{r(N-t+T)}\right)\gamma\zeta\rho\sigma\right.\\
&\qquad\qquad + r\theta\lambda\left(-1 - e^{-k(N-t+T)}\right.\\
&\qquad\qquad\quad + 2e^{-(N-t+T)(k+\lambda\rho\sigma)}\\
&\qquad\qquad\quad \left.\left. + 3(N-t+T)\lambda\rho\sigma\right)\right]\\
&\qquad\quad - 2e^{-k(N-t+T)}k\lambda^4\rho^2\sigma^3\\
&\qquad\quad \times \left[-4\left(e^{k(N-t+T)} - e^{(k+r)(N-t+T)}\right)\gamma\zeta\rho^3\sigma\right.\\
&\qquad\qquad + r\theta\lambda\left(6+\rho^2 - e^{k(N-t+T)}\right.\\
&\qquad\qquad\quad \left.\left.\times\left(6+\rho^2 - 2(N-t+T)\lambda\rho\sigma\right)\right)\right]\\
&\qquad\quad + 2k^3\lambda^2\sigma\\
&\qquad\quad \times \left[2e^{-(N-t+T)(k+\lambda\rho\sigma)}r\theta\lambda\left(1+2\rho^2\right)\right.\\
&\qquad\qquad - e^{-k(N-t+T)}r\theta\lambda\left(2+3\rho^2\right)\\
&\qquad\qquad - 18e^{r(N-t+T)}\gamma\zeta\rho^3\sigma\\
&\qquad\qquad + \rho\left(18\gamma\zeta\rho^2\sigma + r\theta\lambda\right.\\
&\qquad\qquad\quad \left.\left.\times\left(-\rho + (N-t+T)\lambda\left(-1+2\rho^2\right)\sigma\right)\right)\right]\\
&\qquad\quad + e^{-(N-t+T)(3k+2\lambda\rho\sigma)}k^2\lambda^3\rho\sigma^2\\
&\qquad\quad \times \left[8e^{(N-t+T)(2k+\lambda\rho\sigma)}r\theta\lambda\right.
\end{aligned}$$

$$
\begin{aligned}
&- 28e^{(N-t+T)(3k+r+2\lambda\rho\sigma)}\gamma\zeta\rho^{3}\sigma \\
&+ e^{k(N-t+T)}r\theta\lambda\left(-1+\rho^{2}\right) \\
&- 6e^{2(N-t+T)(k+\lambda\rho\sigma)}r\theta\lambda\left(2+\rho^{2}\right) \\
&+ e^{(N-t+T)(3k+2\lambda\rho\sigma)} \\
&\times\left(28\gamma\zeta\rho^{3}\sigma+r\theta\lambda\right. \\
&\quad\left.\times\left(5+\rho\left(5\rho-6\left(N-t+T\right)\lambda\sigma\right)\right)\right)\Big]\Big\}.
\end{aligned} \tag{74}
$$

After simple calculation, the optimal investment strategy (67) becomes

$$
\begin{aligned}
\pi^{*} &= \frac{1}{x\gamma}\cdot[\widehat{\alpha}(t)]^{-1}\cdot\left[\lambda-\rho\sigma\widehat{\beta}(t)\right] \\
&= \frac{\lambda}{x\gamma e^{r(N-t+T)}}\cdot\left[1-\rho\sigma\lambda\frac{\left(1-e^{-(N-t+T)(k+\lambda\rho\sigma)}\right)}{k+\lambda\rho\sigma}\right], \\
\widehat{G}(t,x,l) &= e^{r(N-t+T)}x+\frac{\left(1-e^{-(N-t+T)(k+\lambda\rho\sigma)}\right)\lambda^{2}}{\gamma(k+\lambda\rho\sigma)}l+\frac{\widehat{\Delta}(t)}{\gamma}, \\
\widehat{F}(t,x,l) &= e^{r(N-t+T)}x+\frac{\widehat{B}(t)}{\gamma}l+\frac{\widehat{C}(t)}{\gamma}.
\end{aligned} \tag{75}
$$

□

Theorem 3 and Theorem 4 imply that

$$
\begin{aligned}
\widehat{F}(t,x,l) &= \widehat{V}(t,x,l)=\widehat{f}\left(t,x,l,y^{\pi^{*}},z^{\pi^{*}}\right) \\
&= \mathrm{E}_{t,x,l}\left[X^{\pi}(T+N)\right]-\frac{\gamma}{2} \\
&\quad\times\left\{\mathrm{E}_{t,x,l}\left[X^{\pi}(T+N)^{2}\right]-\left(\mathrm{E}_{t,x,l}\left[X^{\pi}(T+N)\right]\right)^{2}\right\} \\
&= \widehat{G}(t,x,l)-\frac{\gamma}{2}\left[\widehat{H}(t,x,l)-\widehat{G}(t,x,l)^{2}\right],
\end{aligned} \tag{76}
$$

$$
\mathrm{E}_{t,x,l}\left[X^{\pi^{*}}(T+N)\right]=\widehat{G}(t,x,l). \tag{77}
$$

Since

$$
\begin{aligned}
&\mathrm{Var}_{t,x,l}\left[X^{\pi^{*}}(T+N)\right] \\
&\quad= \mathrm{E}_{t,x,l}\left[X^{\pi^{*}}(T+N)\right]^{2}-\left(\mathrm{E}_{t,x,l}X^{\pi^{*}}(T+N)\right)^{2},
\end{aligned} \tag{78}
$$

(76) and (78) imply that

$$
\begin{aligned}
&\mathrm{Var}_{t,x,l}\left[X^{\pi^{*}}(T+N)\right] \\
&\quad= \frac{2}{\gamma}\left[\widehat{G}(t,x,l)-\widehat{F}(t,x,l)\right] \\
&\quad= \frac{2}{\gamma^{2}}\left[\left(\widehat{\beta}(t)-\widehat{B}(t)\right)l+\left(\widehat{\Delta}(t)-\widehat{C}(t)\right)\right] \\
&\quad= \left(2k\gamma^{2}\rho(k+\lambda\rho\sigma)^{3}(k+2\lambda\rho\sigma)\right)^{-1} \\
&\qquad\times e^{-2(N+T)(k+\lambda\rho\sigma)}\lambda^{2}\cdot\Theta,
\end{aligned} \tag{79}
$$

where

$$
\begin{aligned}
\Theta = \Big\{&-2e^{k(N+t+T)+2(N+T)\lambda\rho\sigma}(l-\theta)(k\rho+2\lambda\sigma)(k+\lambda\rho\sigma)^{3} \\
&- 4e^{(N+t+T)(k+\lambda\rho\sigma)}k\lambda\left(-1+\rho^{2}\right)\sigma(k+2\lambda\rho\sigma) \\
&\times(k(l-\theta)+l\lambda\rho\sigma) \\
&+ e^{2t(k+\lambda\rho\sigma)}k\lambda^{2}\rho\left(-1+\rho^{2}\right)\sigma^{2}(2kl-k\theta+2l\lambda\rho\sigma) \\
&+ e^{2(N+T)(k+\lambda\rho\sigma)}\rho(k+2\lambda\rho\sigma) \\
&\quad\times\Big[2k^{3}(l+(-1+k(N-t+T))\theta) \\
&\qquad+6k^{2}(l+(-1+k(N-t+T))\theta)\lambda\rho\sigma \\
&\qquad+2(l+(-1+k(N-t+T))\theta)\lambda^{3}\rho\sigma^{3} \\
&\qquad+k\lambda^{2}\left(l\left(2+4\rho^{2}\right)\right. \\
&\qquad\qquad+\theta\left(-5-\rho^{2}+2k(N-t+T)\right. \\
&\qquad\qquad\left.\left.\times\left(1+2\rho^{2}\right)\right)\right)\sigma^{2}\Big]\Big\}.
\end{aligned} \tag{80}
$$

Putting (77) and (79) together, the efficient frontier is rewritten as

$$
\begin{aligned}
&\mathrm{E}_{t,x,l}\left[X^{\pi^{*}}(T+N)\right] \\
&\quad= e^{r(N-t+T)}x+\frac{\left(1-e^{-(N-t+T)(k+\lambda\rho\sigma)}\right)\lambda^{2}}{\gamma(k+\lambda\rho\sigma)}l+\frac{\widehat{\Delta}(t)}{\gamma} \\
&\qquad+\gamma\sqrt{\mathrm{Var}_{t,x,l}\left[X^{\pi^{*}}(T+N)\right]} \\
&\qquad-\sqrt{2\left[\left(\widehat{\beta}(t)-\widehat{B}(t)\right)l+\left(\widehat{\Delta}(t)-\widehat{C}(t)\right)\right]},
\end{aligned} \tag{81}
$$

where $\widehat{\beta}(t)$, $\widehat{B}(t)$, $\widehat{\Delta}(t)$, and $\widehat{C}(t)$ are given by (71), (72), (73), and (74), respectively.

Remark 5. In a defined contribution pension plan, only the contributions in the accumulation phase before retirement are guaranteed, but the future benefits are undetermined. Considering a annuity contract in the distribution phase after retirement, it achieves the effect that both phases before and after retirement are "defined," which makes the pension plan even more portable and of great convenience for insurers.

4. Numerical Analysis

In this section, some numerical analysis and graphics are provided to illustrate our results. The main objectives are two aspects: one is to explain the properties of the optimal strategies derived in Section 2 and Section 3 and the other is to illustrate the efficient frontier.

First, let us analyze the expression of the optimal time-consistent investment strategy π^{*} in the accumulation phase before retirement. Analysis of the optimal time-consistent

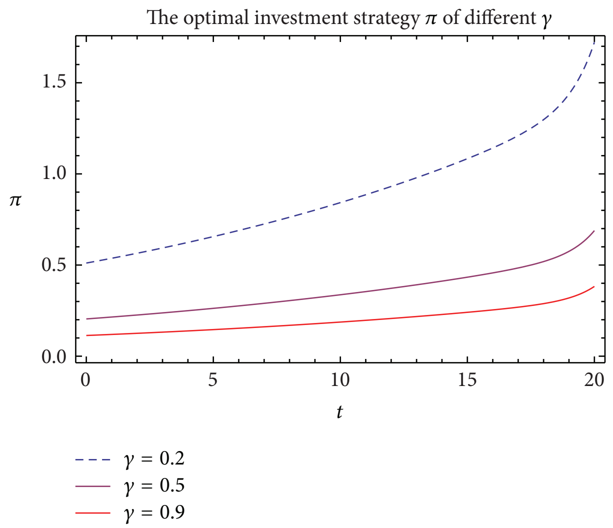


FIGURE 1: Evolutions of optimal investment strategy π^* with different risk aversion coefficients γ.

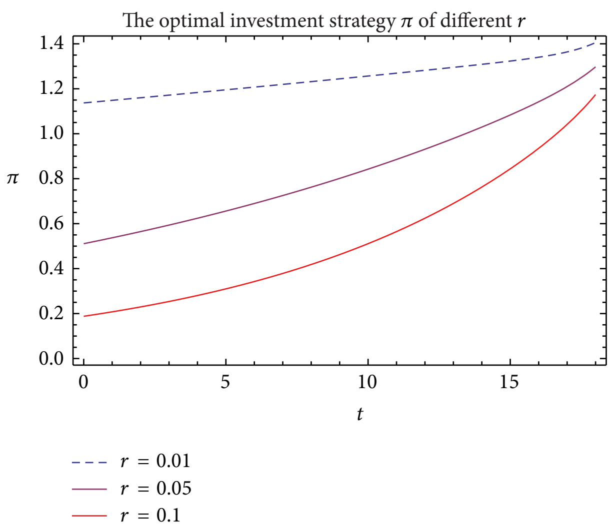


FIGURE 2: Evolutions of optimal investment strategy π^* with different risk-free interest rates r.

investment strategy in the distribution phase after retirement is almost the same and we can get similar results, so it is omitted:

$$\begin{aligned}\pi^* &= \frac{\lambda}{x\gamma e^{r(T-t)}} \cdot \left[1 - \rho\sigma\lambda \frac{\left(1 - e^{(t-T)(k+\lambda\rho\sigma)}\right)}{k + \lambda\rho\sigma}\right] \\ &= \frac{\lambda e^{r(t-T)}}{x\gamma\left(k + \lambda\rho\sigma\right)} \cdot \left(k + \lambda\rho\sigma e^{(t-T)(k+\lambda\rho\sigma)}\right).\end{aligned} \tag{82}$$

The derivative of (82) w.r.t. γ is

$$\frac{\partial\pi^*}{\partial\gamma} = -\frac{\lambda}{x\gamma^2} \cdot \frac{e^{r(t-T)}\left(k + e^{(t-T)(k+\lambda\rho\sigma)}\lambda\rho\sigma\right)}{\left(k + \lambda\rho\sigma\right)}. \tag{83}$$

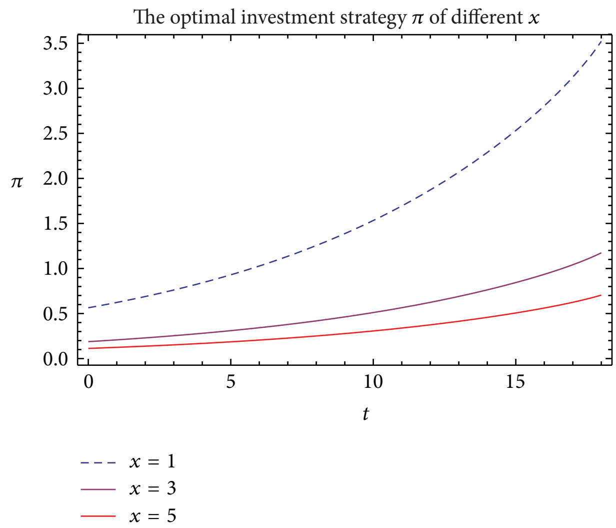


FIGURE 3: Evolutions of optimal investment strategy π^* with different fund size x for the same risk averse level.

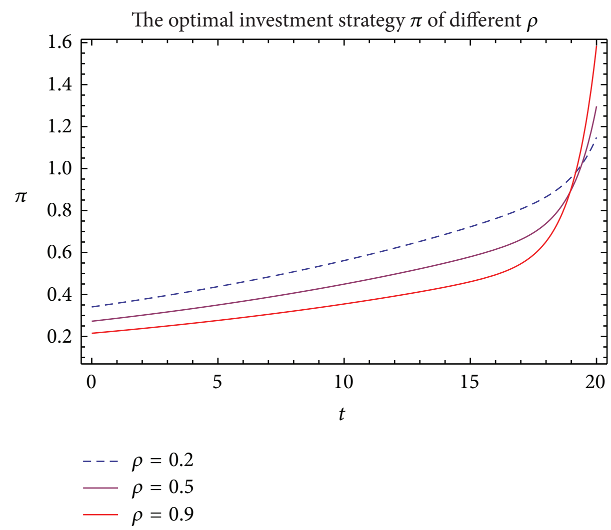


FIGURE 4: Evolutions of optimal investment strategy π^* with different correlation coefficients ρ.

Since

$$\frac{e^{r(t-T)}\left(k + e^{(t-T)(k+\lambda\rho\sigma)}\lambda\rho\sigma\right)}{\left(k + \lambda\rho\sigma\right)} > 0, \quad \frac{\lambda}{x\gamma^2} > 0, \tag{84}$$

then

$$\frac{\partial\pi^*}{\partial\gamma} < 0 \tag{85}$$

which shows that the optimal investment policy decreases with respect to the risk aversion level γ, referring to Figure 1, which is consistent with reality. In fact, the higher degree of risk aversion, people should invest the less cash in risky assets to avoid risk.

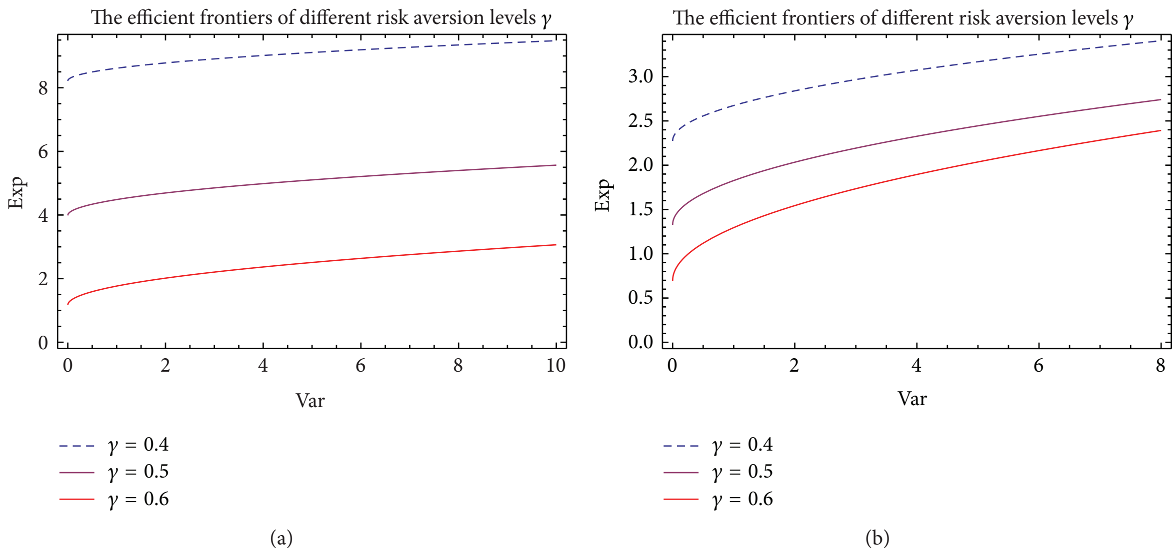


FIGURE 5: (a) The efficient frontiers of different risk averse level γ when $t = 21$. (b) The efficient frontiers of different risk averse level γ when $t = 28$.

Similarly, the derivatives of (82) w.r.t. risk-free interest rate r and starting wealth level x are given by

$$\frac{\partial \pi^*}{\partial r} = -\frac{e^{r(t-T)}(T-t)\lambda\left(k+e^{(t-T)}\lambda\rho\sigma(k+\lambda\rho\sigma)\right)}{x\gamma(k+\lambda\rho\sigma)} < 0,$$
$$\frac{\partial \pi^*}{\partial x} = -\frac{e^{r(t-T)}\lambda\left(k+e^{(t-T)}\lambda\rho\sigma(k+\lambda\rho\sigma)\right)}{x^2\gamma(k+\lambda\rho\sigma)} < 0, \tag{86}$$

which mean that the optimal investment policy also decreases with respect to the risk-free interest rate r and starting wealth level x. See Figures 2 and 3, respectively. Obviously, the higher the risk-free interest rate, people must increase the investment on the risk-free asset because they can get more profit without risk, and risk investment corresponds to reduction. Furthermore, a risk-averse investor must strive to control the investment amount of risk assets, if the initial capital size x increasing, only by reducing the proportion π can make the investment amount of risk assets stay at a relatively low level.

But for the derivative of (82) w.r.t. correlation coefficient ρ, which will be more complicated,

$$\frac{\partial \pi^*}{\partial \rho} = \frac{e^{r(t-T)}\lambda^2\sigma\left(-k+e^{(t-T)}(k+\lambda\rho\sigma)^2\right)}{x\gamma(k+\lambda\rho\sigma)^2}. \tag{87}$$

If $t \in [0,T]$ is close enough to T, then

$$e^{(t-T)}(k+\lambda\rho\sigma)^2 > k,$$
$$-k+e^{(t-T)}(k+\lambda\rho\sigma)^2 > 0, \tag{88}$$

so that

$$\frac{\partial \pi^*}{\partial \rho} > 0 \tag{89}$$

which means the optimal investment strategy π^* increasing with respect to the correlation coefficient ρ. However, as t is close enough to 0, especially for $t \in [0,T]$ small enough, then

$$e^{(t-T)}(k+\lambda\rho\sigma)^2 < k,$$
$$-k+e^{(t-T)}(k+\lambda\rho\sigma)^2 < 0 \tag{90}$$

which shows that the optimal investment strategy π^* decreases with respect to the correlation coefficient ρ, referring to Figure 4 which illustrates our conclusion.

Second, we analyze the efficient frontier in the distribution phase after retirement. Analysis of the efficient frontier in the accumulation phase before retirement is almost the same, so it is omitted. We fix the time t and take appropriate value to x and l for the efficient frontier (49) which is a function of variance as the independent variable. Drawing a picture of the efficient frontier on the rectangular plane coordinate system of variance (Var) and expected return (Exp) for the different risk aversion levels γ, referring to Figure 5, they are convex curves and the expected return (Exp) smaller if the time t takes larger.

Taking derivatives of variance (47) and by similar analysis as above we also have

$$\frac{\partial \mathrm{Var}_{t,x,l}\left[X^{\pi^*}(T)\right]}{\partial \gamma} < 0, \qquad \frac{\partial \mathrm{Var}_{t,x,l}\left[X^{\pi^*}(T)\right]}{\partial t} < 0. \tag{91}$$

The inequalities in (91) show that the variance decreases with respect to the risk aversion coefficient γ and time t when the fund size level and other parameters are fixed. Referring to Figure 6 shows the correctness of our conclusion. Variance as a measure of risk will be smaller if investor be more risk averse. Meanwhile, for more longer time investment, a risk aversion insurer will be more cautious, so the variance also decreases with the investment time t.

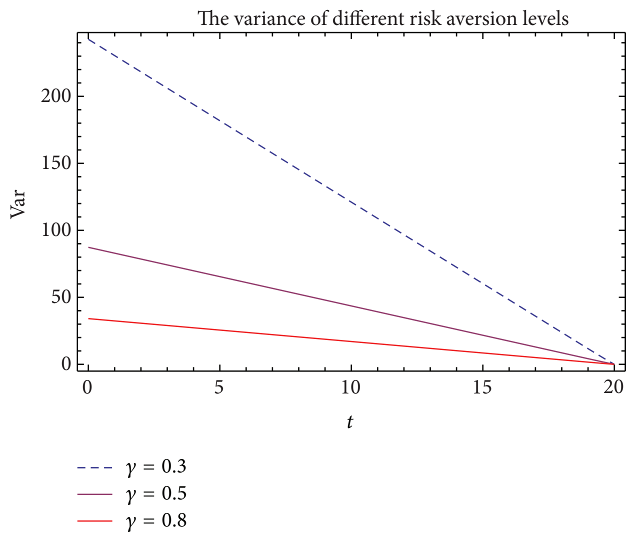


FIGURE 6: Variance with different risk aversion levels γ.

Through detailed comparison of the above figures, we also discover that different risk aversion levels can lead to surprising difference of the fund size variance (seeing Figure 6) and the more strict correlation of Brownian motions may result in fierce variation (seeing Figure 4) of the optimal investment strategy especially when time t is close enough to T. So we should control risk aversion level and correlation coefficient of Brownian motions at an appropriate level for a DC pension plan.

5. Conclusion

The main innovation of this paper is merging defined contribution with annuity contract as a whole pension plan to study both accumulation phase before retirement and distribution phase after retirement in the mean-variance framework with the return of premiums clauses, which achieves the effect that both phases before and after retirement are "defined" to make the defined contribution plans even more portable and great convenience for insurance companies. The return of premiums clauses following the formulation of He and Liang [12], the time-consistent framework according to Li et al. [8] and inspired by the literature of Gao [1], we obtain the time-consistent explicit solution by applying stochastic optimal control techniques under Heston's SV models. Numerical analysis illustrates our results and also deepens our knowledge or understanding of the research results.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant nos. 11201335 and 11301376) and the Humanity and Social Science Foundation of Ministry of Education of China (Grant no. 11YJC910007).

References

[1] J. Gao, "Optimal investment strategy for annuity contracts under the constant elasticity of variance (CEV) model," *Insurance: Mathematics and Economics*, vol. 45, no. 1, pp. 9–18, 2009.

[2] A. J. G. Cairns, D. Blake, and K. Dowd, "Stochastic lifestyling: optimal dynamic asset allocation for defined contribution pension plans," *Journal of Economic Dynamics and Control*, vol. 30, no. 5, pp. 843–877, 2006.

[3] N. Dokuchaev and X. Yu Zhou, "Optimal investment strategies with bounded risks, general utilities, and goal achieving," *Journal of Mathematical Economics*, vol. 35, no. 2, pp. 289–309, 2001.

[4] C. Blanchet-Scalliet, N. El Karoui, M. Jeanblanc, and L. Martellini, "Optimal investment decisions when time-horizon is uncertain," *Journal of Mathematical Economics*, vol. 44, no. 11, pp. 1100–1113, 2008.

[5] S. Faggian and F. Gozzi, "Optimal investment models with vintage capital: dynamic programming approach," *Journal of Mathematical Economics*, vol. 46, no. 4, pp. 416–437, 2010.

[6] D. Blake, D. Wright, and Y. Zhang, "Target-driven investing: optimal investment strategies in defined contribution pension plans under loss aversion," *Journal of Economic Dynamics and Control*, vol. 37, no. 1, pp. 195–209, 2013.

[7] H. Kraft, "Optimal portfolios and Heston's stochastic volatility model: an explicit solution for power utility," *Quantitative Finance*, vol. 5, no. 3, pp. 303–313, 2005.

[8] Z. Li, Y. Zeng, and Y. Lai, "Optimal time-consistent investment and reinsurance strategies for insurers under Heston's SV model," *Insurance: Mathematics and Economics*, vol. 51, no. 1, pp. 191–203, 2012.

[9] H. Markowitz, "Portfolio selection," *Jornal of Finance*, vol. 7, pp. 77–91, 1952.

[10] T. Bjork and A. Murgoci, "A general theory of Markovian time inconsistent stochastic control problems," Working Paper, Stockolm School of Economics, 2010.

[11] T. Bjork, A. Murgoci, and Z. Xunyu, "Mean-variance portfolio optimization with statedependent risk aversion," *Mathematical Finance*, vol. 24, no. 1, pp. 1–24, 2012.

[12] L. He and Z. Liang, "Optimal investment strategy for the DC plan with the return of premiums clauses in a mean-variance framework," *Insurance: Mathematics and Economics*, vol. 53, no. 3, pp. 643–649, 2013.

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 893851, 13 pages
http://dx.doi.org/10.1155/2014/893851

Research Article

The Research on Low Carbon Logistics Routing Optimization Based on DNA-Ant Colony Algorithm

Liyi Zhang,[1] Ying Wang,[2] Teng Fei,[1] and Hongwei Ren[3]

[1] *Information Engineering College, Tianjin University of Commerce, Tianjin 300134, China*
[2] *Economic College, Tianjin University of Commerce, Tianjin 300134, China*
[3] *Waston Engineering Shool, State University of New York at Binghamton, 138 Conklin Avenue, Binghamton, NY 13902, USA*

Correspondence should be addressed to Teng Fei; fei_8825@163.com

Received 11 April 2014; Accepted 15 May 2014; Published 22 June 2014

Academic Editor: Xiang Li

Copyright © 2014 Liyi Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As the energy conservation and emission reduction and sustainable development have become the hot topics in the world, low carbon issues catch more and more attention. Logistics, which is one of the important economic activities, plays a crucial role in the low carbon development. Logistics leads to some significant issues about consuming energy and carbon emissions. Therefore, reducing energy consumption and carbon emissions has become the inevitable trend for logistics industry. Low carbon logistics is introduced in these situations. In this paper, from the microcosmic aspects, we will bring the low carbon idea in the path optimization issues and change the amount of carbon emissions into carbon emissions cost to establish the path optimization model based on the optimization objectives of the lowest cost of carbon emissions. According to different levels of air pollution, we will establish the double objectives path optimization model with the consideration of carbon emissions cost and economy cost. Use DNA-ant colony algorithm to optimize and simulate the model. The simulation indicates that DNA-ant colony algorithm could find a more reasonable solution for low carbon logistics path optimization problems.

1. Introduction

As the energy conservation and emission reduction and sustainable development has become the hot topics in the world, low carbon issues catch more and more attention. Logistics, which is one important economic activity, plays a crucial role in the low carbon development. Logistics leads to some significant issues about consuming energy and carbon emissions. Therefore, reducing energy consumption and carbon emissions have become the inevitable trend for logistics industry. Low carbon logistics is introduced in these situations. Low carbon logistics means that the processes of logistics, based on the goals of low energy consumption, low pollution, and low emissions, use the technology of energy efficiency, renewable energy, and reducing greenhouse gas emissions to restrain the harm to environment, which would be also helpful for the purification of the logistics environment and get the full use of logistics resources. It includes logistics operation part and the whole process of low carbon logistics management [1]. Logistics operation links include logistics system works optimization, such as reasonable planning distribution path, improving logistics efficiency, reducing resources waste, and taking full use of logistics resources. Logistics management includes using all kinds of low carbon technologies to improve logistics system, which will reduce the carbon emissions in the processes of transportation.

Low carbon logistics is one frontier domain of research, which catches more and more attention of scholars. Most low carbon logistics researches are focused on the macroscopic aspects, which are the qualitative researches of low carbon logistics, including low carbon logistics itself analysis [2–5], low carbon logistics system analysis [6–8], and low carbon logistics network analysis [9–11]. However, there are only few researches about microscopic aspects of low carbon logistics.

In this paper, from the microcosmic aspects, we bring the low carbon idea in the path optimization issues and change the amount of carbon emissions into carbon emissions cost to establish the path optimization model based on the optimization objectives of lowest cost of carbon emissions. According to different levels of air pollution, we establish the double objectives path optimization model with the consideration of carbon emissions cost and economy cost. We will use DNA-ant colony algorithm to optimize and simulate the model. The simulation indicates that DNA-ant colony algorithm could find a more reasonable solution for low carbon logistics path optimization problems.

2. Modeling

2.1. Distribution Optimization Model of Minimum Cost of Carbon Emissions. Low carbon logistics is a kind of low energy cost and low pollution logistics whose goal is to achieve the highest efficiency of logistics with the lowest greenhouse gas emissions [12]. Relevant data shows that thermal power emission, vehicle exhaust emissions, and construction are the main sources of carbon dioxide emissions. Bearing the social goods transportation, storage, packaging, processing, distribution, loading, unloading, and other services, logistics has become the big one for the carbon emissions. Looking for the reasonable distribution route to reduce carbon emissions is particularly important. The reduction in carbon emissions model reflects the reduction of carbon emissions cost. Therefore, this paper established low carbon logistics distribution route optimization model to achieve the minimizing carbon emissions.

2.1.1. The Calculation of Cost of Carbon Emissions. As the vehicles pickup or delivery goods in the different order crossing all customers, the car load changes. With the increase of vehicle load, the unit distance fuel consumption rises, leading to the increase of carbon emissions cost [13]. With the increase of vehicle transport distance, carbon emissions cost will also increase. Thus, the cost of carbon emissions does not only has relationship with transport distance but is also related to the vehicles' load.

The unit distance fuel consumption $\varepsilon(q)$ has a linear relationship with the total weight of the vehicles q_{sum}, which is

$$\varepsilon(q) = \delta q_{\text{sum}} + b = \delta(q + q_0) + b, \tag{1}$$

where $\varepsilon(q)$ is the unit distance fuel consumption, δ is the relationship factor between the unit distance fuel consumption and vehicles' weight, and q_{sum} is the total weight of the vehicles including vehicle loading weight q and vehicle's weight q_0.

When vehicles are full, the unit distance fuel consumption is ε^*, which is

$$\varepsilon^* = \delta(Q + q_0) + b, \tag{2}$$

where ε^* is the unit distance fuel consumption when vehicles are full and Q is the vehicles' maximum load.

When vehicles are empty, the unit distance fuel consumption is ε_0, which is

$$\varepsilon_0 = \delta q_0 + b. \tag{3}$$

From formula (2) and formula (3), we can get

$$\delta = \frac{\varepsilon^* - \varepsilon_0}{Q}. \tag{4}$$

Bring formula (4) into formula (1); the unit distance fuel consumption $\varepsilon(q)$ is as follows:

$$\begin{aligned}\varepsilon(q) &= \delta q_{\text{sum}} + b = \delta(q + q_0) + b \\ &= \delta q + (\delta q_0 + b) = \frac{\varepsilon^* - \varepsilon_0}{Q} q + \varepsilon_0.\end{aligned} \tag{5}$$

Thus, the carbon emissions cost $CO_2(q_{ij})$ of the goods with the weight q_i from customer i to customer j is as follows:

$$CO_2(q_{ij}) = c_3 \eta_0 \varepsilon(q_i) d_{ij} = c_3 \eta_0 \left(\frac{\varepsilon^* - \varepsilon_0}{Q} q_i + \varepsilon_0\right) d_{ij}, \tag{6}$$

where c_3 is the unit carbon emissions cost, η_0 is the carbon emissions factor, d_{ij} is the distance from customer i to customer j, and $\eta_0 \varepsilon(q_i)$ is the carbon emissions of the goods with the weight q_i from customer i to customer j.

2.1.2. The Calculation of Cost of Carbon Emissions. For the convenience of model establishment, assume that there is only one distribution center whose location is known. All the vehicles start from the distribution center and return to distribution center after delivery. The vehicles' load is known. The location and the demand of customer are known. One vehicle is for one customer.

According to the above assumption, establish the minimizing carbon emissions cost model:

$$\begin{aligned}\min CO_2 &= \sum_{i=0}^{L} \sum_{j=0}^{L} \sum_{k=1}^{m'} CO_2(q_{ij}) x_{ijk} \\ &= \sum_{i=0}^{L} \sum_{j=0}^{L} \sum_{k=1}^{m'} c_3 \eta_0 \varepsilon(q_i) d_{ij} x_{ijk} \\ &= \sum_{i=0}^{L} \sum_{j=0}^{L} \sum_{k=1}^{m'} c_3 \eta_0 \left(\frac{\varepsilon^* - \varepsilon_0}{Q} q_i + \varepsilon_0\right) d_{ij} x_{ijk} \\ &= \sum_{i=0}^{L} \sum_{j=0}^{L} \sum_{k=1}^{m'} c_3 \eta_0 \varepsilon_0 d_{ij} x_{ijk} \\ &\quad + \sum_{i=0}^{L} \sum_{j=0}^{L} \sum_{k=1}^{m'} c_3 \eta_0 \frac{\varepsilon^* - \varepsilon_0}{Q} q_i d_{ij} x_{ijk},\end{aligned} \tag{7}$$

where L is the number of customers, m' is the number of vehicles, q_i is the demand of the i customer, Q is the maximum load of the vehicles, and k is the vehicle's number.

The number of distribution center is 0. The numbers of customers are $1, 2, 3 \cdots L$. Define variables x_{ijk} and y_{ki} as

$$x_{ijk} = \begin{cases} 1, & \text{vehicle } k \text{ from } i \text{ to } j \\ 0, & \text{others} \end{cases} \quad (8)$$

$$y_{ki} = \begin{cases} 1, & i \text{ is sevice by vehicle } k \\ 0, & \text{others} \end{cases} \quad (9)$$

$$\sum_{i=0}^{L} q_i y_{ki} \le Q, \quad k \in [1, m'] \quad (10)$$

$$\sum_{k=1}^{m'} y_{ki} = 1, \quad i \in [0, L], \; k \in [1, m'] \quad (11)$$

$$\sum_{i=0}^{L} \sum_{k=1}^{m'} x_{ijk} = 1, \quad j \in [0, L], \; k \in [1, m'] \quad (12)$$

$$\sum_{j=0}^{L} \sum_{k=1}^{m'} x_{ijk} = 1, \quad i \in [0, L], \; k \in [1, m'] \quad (13)$$

$$\sum_{i=0}^{L} \sum_{k=1}^{m'} x_{0ik} = \sum_{j=0}^{L} \sum_{k=1}^{m'} x_{j0k}. \quad (14)$$

Formula (7) is the objective function. The first part is the carbon emissions cost caused by the vehicle's weight. The second part is the carbon emissions cost caused by the vehicle's load. Formula (10) shows that the sum of customers' demand cannot be greater than the maximum load of vehicles. Formulas (11), (12), and (13) show that one vehicle is for one customer. Formula (14) shows that vehicles start from distribution center and then return to distribution center.

2.2. Double Target Distribution Optimization Model with the Consideration of Carbon Emissions. Since the increasing pollution problems, the government considers the logistics network planning from the aspect of the whole area low carbon development to minimize the carbon emissions of the whole area. But transportation enterprises will choose the logistics lines based on the logistics network system according to their target and determine the allocation on each line. To achieve the goal of the government and the transport enterprises win-win, this paper establishes the double target distribution optimization model based on "pay attention to carbon emissions and also give consideration to economy." The basic thought of a multiobjective optimization problem is to transfer multiobject into a numerical target evaluation function, which generally uses the linear weighted sum method [14].

Firstly, the goods are delivered to the distribution center. Then, the goods are sent to each customer by through highways, waterways, air transport, and a variety of other ways. In the whole distribution process, the situation is complicated, in which there exists not only the transformation of multiple transportation modes but also multiple layers of distribution network. In a variety of transportation modes, highway transportation is the largest one of carbon dioxide emissions. Thus, for the convenience of model establishment, only consider one level of distribution network, which is from distribution center to customer delivery and only consider the highway transportation mode.

For the convenience of model establishment, make the following assumptions.

(1) Only consider the distribution center whose location is known. All the vehicles start from the distribution center. After the delivery mission, all the vehicles come back to the distribution center waiting for unified deployment.

(2) The delivery goods can be mixed. Each customer's goods will not exceed the maximum load of the vehicle.

(3) The location and demand of customers are known. One vehicle is for one customer.

(4) Load is known.

(5) The vehicle serves for each affected point service, and on the way only discharges without loading.

According to the above assumptions, establish the objective evaluation function:

$$\min W = (W_1, W_2) = \rho_1 W_1 + \rho_2 W_2 \quad (15)$$

$$W_1 = \sum_{i=0}^{L} \sum_{j=0}^{L} \sum_{k=1}^{m'} CO_2 (q_{ij}) x_{ijk} = \sum_{i=0}^{L} \sum_{j=0}^{L} \sum_{k=1}^{m'} c_3 \eta_0 \varepsilon (q_i) d_{ij} x_{ijk} \quad (16)$$

$$W_2 = c_0 m' + \sum_{i=0}^{L} \sum_{j=0}^{L} \sum_{k=1}^{m'} c_1 d_{ij} x_{ijk} + c_2 \sum_{k=1}^{m'} \left(\max \left\{ \sum_{i=1}^{L} (q_i y_{ki} - Q), 0 \right\} \right) \quad (17)$$

$$\sum_{i=0}^{L} q_i y_{ki} \le Q, \quad k \in [1, m'] \quad (18)$$

$$\sum_{k=1}^{m'} y_{ki} = 1, \quad i \in [0, L], \; k \in [1, m'] \quad (19)$$

$$\sum_{i=0}^{L} \sum_{k=1}^{m'} x_{ijk} = 1, \quad j \in [0, L], \; k \in [1, m'] \quad (20)$$

$$\sum_{j=0}^{L} \sum_{k=1}^{m'} x_{ijk} = 1, \quad i \in [0, L], \; k \in [1, m'] \quad (21)$$

$$\sum_{i=0}^{L} x_{ijk} = y_{kj}, \quad i \in [0, L], \; k \in [1, m'] \tag{22}$$

$$\sum_{j=0}^{L} x_{ijk} = y_{ki}, \quad i \in [0, L], \; k \in [1, m'] \tag{23}$$

$$\sum_{i=0}^{L} \sum_{k=1}^{m'} x_{0ik} = \sum_{j=0}^{L} \sum_{k=1}^{m'} x_{j0k}, \tag{24}$$

where ρ_1, ρ_2 are the weight coefficients of the objective function, c_0 is unit vehicle cost, c_1 is unit cost of travelling distance, and c_2 is the overload punishment coefficient. Formula (15) is the objective function. The lower value the objective evaluation function has, the better the path is. Formula (16) is the evaluation function of carbon emissions. Formula (17) is the economic evaluation function including three parts. The first part is the vehicle fixed cost. The second part is the vehicle's transportation cost. The third part is the overload punishment cost. Formulas (18) to (24) are the constraint conditions. Formula (22) shows that the mission of the demand point j is accomplished by the vehicle k through point i. Formula (23) shows that the mission of the demand point i is accomplished by the vehicle k through point j. The other constraint conditions are the same as the minimum carbon emissions cost distribution optimization model.

According to the degree of pollution, the air pollution index is divided into five levels: top grade, good, light pollution, moderate pollution, and high level of pollution. Top level and good level include normal activities. Light pollution includes long-term exposure to this level air, vulnerable groups' symptoms will be slightly worse, and healthy people will have irritation symptoms. Moderate pollution includes contacting with the air after a certain period of time, symptoms of the people with heart disease and pulmonary disease significantly will increase, exercise tolerance decreases, and common symptoms happen in healthy people. In high level of pollution, healthy exercise tolerance is reduced and it has obvious symptoms and diseases. According to different level of air pollution index, the values of weight coefficients of evaluation function ρ_1, ρ_2 are also different. When the air pollution degree is good, ignore the influence of the carbon emissions cost, which is $\rho_1 = 0$, $\rho_2 = 1$. When the air pollution is light pollution, first consider the economic then consider carbon emissions, which is $\rho_1 \leq \rho_2$. When the air pollution is moderately severe pollution, put the emission reduction and energy saving in the first place with carbon emissions as the first goal and economic target as the balance, which is $\rho_1 > \rho_2$.

3. Methods

3.1. Basic Ant Colony Algorithm. Scientists, who study social insect behavior characteristic, found that the insect at the community level of cooperation is basically self-organizing. This kind of collective behavior produced by the social organism, which is a kind of swarm intelligence, catches the eyes of many researchers in the fields of management science and engineering. Ant colony algorithm is a typical example of the use of swarm intelligence to solve combinatorial optimization problems. Ant colony algorithm as a new bionic evolutionary algorithm is published by Dorigo and Gambardella [15]. The algorithm imitates ants foraging behavior, according to the heuristic idea, which is caused by entrainment of pheromones, gradually converging to the global optimal solution of the problem. So far, the ant colony algorithm has been used to solve traveling salesman problem, quadratic assignment problem, vehicle routing problem, dynamic vehicle scheduling problem, and network routing optimization problem.

Ant colony algorithm is a kind of parallel algorithm. The searching process is not starting from a point, but from the multiple points simultaneously. The distributed parallel model greatly improves the whole operation efficiency and quick reaction capability of the algorithm, which not only increases the reliability of the algorithm but also makes the algorithm have a stronger global searching ability. Ant colony algorithm has positive feedback characteristics, which can strengthen the optimal solution of pheromones to speed up the convergence speed of the algorithm. Ant colony algorithm has robustness, whose result is not dependent on the initial route choice, and does not need manual adjustment in the process. Ant colony algorithm is easily combined with other heuristic algorithms to improve the algorithm performance. Although the ant colony algorithm has many advantages, there are still some defects such as long searching time, slow convergence speed, slow evolution, stagnation happening easily, and precocious phenomena.

Assume that there are n cities. The distance between any two cities i and j is d_{ij} $(i, j = 1, 2, \ldots, n)$. $b_i(t)$ is the number of ants at time t at city i. $m = \sum_{i=1}^{n} b_i(t)$ is the total number of ants. $\tau_{ij}(t)$ is the pheromone at time t on the line ij. At time $t = 0$, every path has the same pheromone strength. $\Delta\tau_{ij}(t) = 0$.

With time passing by, the new pheromone is added and old pheromones evaporate. ρ is the pheromone volatilization coefficient which indicates pheromones volatile speed. When all the ants accomplish one travel, pheromone on every path is

$$\tau_{ij}(t+1) = (1-\rho)\,\tau_{ij}(t) + \Delta\tau_{ij}(t) \tag{25}$$

$$\Delta\tau_{ij}(t) = \sum_{k=1}^{m} \Delta\tau_{ij}^{k}(t), \tag{26}$$

where $\Delta\tau_{ij}(t)$ is the pheromone increment on the path ij. At the beginning, $\Delta\tau_{ij}(0) = 0$. $\Delta\tau_{ij}^{k}(t)$ is the pheromone released by ant k on the path ij, which is determined by the ants performance. The more the paths are, the more the pheromone is:

$$\Delta_{ij}^{k}(t) = \begin{cases} \dfrac{Q}{L_k} & \text{section } k \text{ of ants pass edge} \\ & ij \text{ in the course of this tour} \\ 0 & \text{else,} \end{cases} \tag{27}$$

where Q is constant and L_k is the path length of the ant k. The transition probability of ant k from city i to city j is

$$P_{ij}^{k} = \begin{cases} \dfrac{\left[\tau_{ij}(t)\right]^{\alpha} \cdot \left[\eta_{ij}(t)\right]^{\beta}}{\sum_{s\in \text{allowed}_k} \left[\tau_{is}(t)\right]^{\alpha} \cdot \left[\eta_{is}(t)\right]^{\beta}} & j \in \text{allowed}_k \\ 0 & \text{others,} \end{cases} \quad (28)$$

where $\text{allowed}_k = (1, 2, \ldots, n) - \text{tabu}_k$ indicates that the city collection ant k can choose currently. tabu_k $(k = 1, 2, \ldots, m)$ is the taboo table of ant k, which indicates the cities ant k has passed by to show the memory of ants. $\eta_{ij}(t)$ is prior knowledge visibility. In the TSP problem, it is the heuristic information from one city moved to another city, which is generally $\eta_{ij}(t) = 1/d_{ij}$. α is the importance of the residual information on the path ij. β is the importance of the heuristic information.

The basic ant colony algorithm to achieve the process is that m ants start from one certain city at the same time to choose the next city based on Formula (28), which indicates that ants prefer visiting the path with higher intensity of pheromone. The passed cities will be put into tabu_k. After all ants finish one travel, renew the pheromone on each path based on Formula (25) to Formula (27). Repeat the above processes until termination condition is established.

3.2. DNA-Ant Colony Algorithm Solution Model. The numerous studies of ant colony algorithm have shown its significance in the optimization combination problem, but there are still some shortcomings, such as seeking to local optimal solution rather than the global optimal solution and convergence lag. In particular, for the selection of basic ant colony algorithm parameters, there is no theoretical derivation but relying on the results of experiments. The selection of ant colony algorithm parameters is directly related to the effectiveness of the algorithm's solution. If the parameter selection is improper, it will seriously affect the performance of the ant colony algorithm. DNA-ant colony algorithm controls the parameter selection by the crossover and mutation idea of DNA algorithm to optimize the performance of ant colony algorithm, which will overcome the shortcomings of ant colony algorithm to improve the convergence rate and search the global optimal solution.

DNA, the so-called deoxyribonucleic acid, is the most important biological macromolecules of organisms in the nature and the main genetic materials for all creatures. The discovery of DNA double helix structure marks the development of biological science which has entered the phase of molecular biology. DNA is a kind of high molecular compound, which is the basic unit of DNA nucleotides. Each deoxyribonucleotide is composed of a molecular phosphate, molecular DNA nucleotides, and a molecule nitrogenous base. Nitrogenous base includes adenine deoxynucleotide (A), guanine oligodeoxynucleotides (G), cytosine deoxyribonucleotide (C), and thymidine nucleotide (T). Modern molecular biology believes that DNA is the main material basis of biological inheritance which stored the genetic information. It transfers genetic information from parent to offspring by self-copy transfer and generates the RNA transcription (ribonucleic acid) to translate into specific proteins to control the phenomenon of life [16–18]. DNA molecule with the double helix structure is circled by the two parallel deoxynucleotides long chain. Deoxyribose and phosphate in DNA molecule alternately link and arrange on the outer side constituting the basic skeleton. The bases on the two chains are linked together by hydrogen bonds forming base pairs. Base pairs follow the principle of complementary base pairing, namely, purines and pyrimidines matching; that is, adenine (A) and thymine (T) must be matched and guanine (G) and cytosine (C) must be matched. Although deoxyribose and phosphate arrange stably on the deoxyribonucleotide long chains, the order of base pairs is protean on the long chain. There are only four bases and two methods of four types of the four bases to form the DNA molecules, but different orders of base pairs constitute a wealth of information [19].

The crossover and mutation in DNA algorithm is different from genetic algorithm. Crossover and mutation in DNA algorithm is based on gene level with a different encoding method, which is two digits binary encoding instead of unit binary encoding. The crossover operation of DNA algorithm is based on the two-point crossover method with a certain probability p_c, which means that a part of bases swaps randomly in the deoxyribonucleotide long chains to form new deoxyribonucleotide long chains (see Figure 11).

Transform coding method of DNA algorithm into two binary coding method can be recognized by computers, which is A-00, T-01, C-10, G-11, and 00 with 01 and 10 with 11 (see Figure 12).

Mutation of DNA algorithm is based on purine replacing purine and pyrimidine replacing pyrimidine, with A changing into G and C changing into T. Bases correspond to binary machines coding method, with 00 changing with 11, and 10 changing with 01. In the mutation, the base sequences mutation operates in a certain probability p_m (see Figure 13).

DNA-ant colony algorithm optimizes the parameters α, β, ρ of the basic ant colony algorithm based on the crossover and mutation ideas of DNA algorithm to solve the low carbon logistics route optimization problems. The solution processes of low carbon logistics route optimization based on DNA-ant colony algorithm is as shown in Figure 1. The pseudocode of low carbon logistics route optimization based on DNA-Ant colony algorithm is as follows.

(1) The basic ant colony algorithm parameters initialization is as follows:

$NC = 0$ (NC is the number of iterations);
$Load = 0$ ($Load$ is the vehicle load);
$\tau_{ij}(0) = 0$ (the pheromone on branch ij is 0).

(2) The DNA ant colony algorithm parameters initialization is as follows:

$DNA\_NUM = 10$ (DNA_NUM is the number of updating generation of DNA algorithm);
$\alpha(1, NUM)$ (a is importance matrix of the residual information);

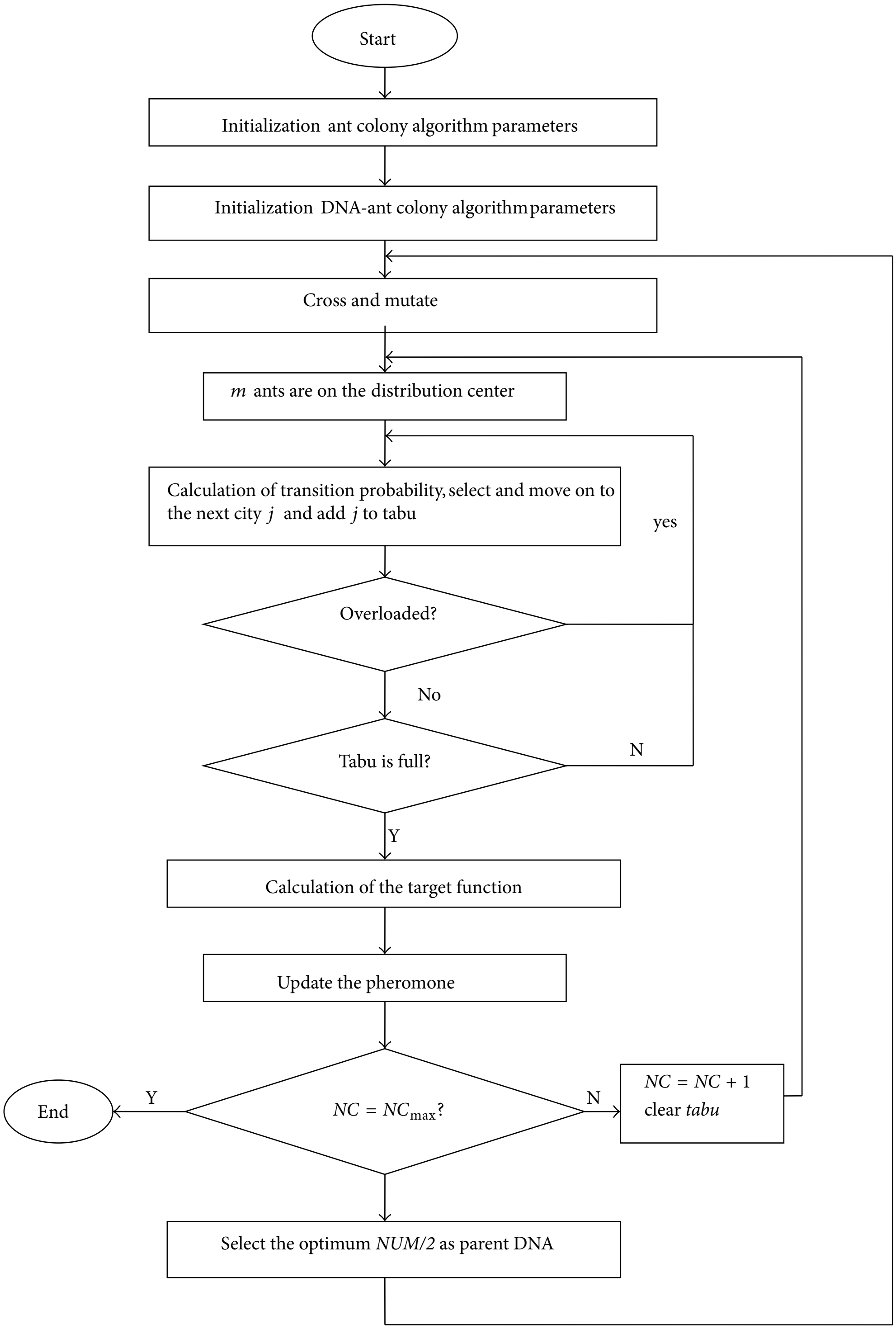


Figure 1: The model solution flow chart of DNA-ant colony algorithm.

$\beta(1, NUM)$ (β is the importance matrix of the inspiring pheromone);

$\rho(1, NUM)$ (ρ is the matrix of pheromone volatilization coefficient);

$p_c = 0.6$ (p_c is the crossover probability of DNA algorithm);

$p_m = 0.08$ (p_m is the mutation probability of DNA algorithm).

(3) Operate crossover and mutation of DNA algorithm are as follows:

for 1 to DNA_NUM do;

repeat Steps (3)~(8);

operate crossover and mutation of DNA algorithm;

for 1 to NUM do;

repeat Steps (4)~(7).

(4) Put m ants at the distribution center. The number of distribution center 1 in the array tabu to form the array tabu(1) is as follows:

while $NC < NC_{\max}$;

for $i = 1$ to m do;

$J = \{1, 2.3 \ldots n\} - \text{tabu}(1)$ (J is the matrix of the cities will be visited).

(5) $k = k + 1$:

calculate the transition probability of ants according to Formula (28);

$Load = Load + Load\_server(k)$;

if $Load \le Load\_$ max;

ant i goes to the next City j. Choose and move to the next City j and put j into array tabu;

else

ant i returns to distribution center and put distribution center with number 1 into array tabu;

$Load = 0$;

repeat Step (5) until all the cities have travelled through and then check whether tabu is full. If not, return to Step (5). Otherwise, go on to Step (6).

(6) for $i = 1$ to m do:

calculate the path length l to solve the objective function (the values of minimum carbon emissions cost function or the objective evaluation function) and record the current best solution. Update the pheromone based on the formula (25).

(7) if $NC < NC_{\max}$:

$$NC = NC + 1 \tag{29}$$

and then clean up tabu and return to Step (4).

(8) Choose the better $NUM/2$ as the parent DNA to operate the crossover and mutation. Get the new α, β, ρ parameter matrices, and return to Step (3).

4. Results and Discussion

Assume the distribution center has 16 customers whose coordinate is (0,0). Table 1 is the demand of each customer. Table 2 is the coordinate of each customer. The maximum load of each vehicle is 75 tons. In the simulation, set $c_0 = 0$, $c_1 = 1, c_2 = \infty$. The unit distance fuel cost of empty car and full car is $\varepsilon_0 = 1$ and $\varepsilon^* = 2$. The unit carbon emissions cost is $c_3 = 0.3$. The emission factor is $\eta = 2.61$ [20]. The distances between each customer and distribution center can be calculated by Formula (30):

$$d_{ij} = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2}. \tag{30}$$

Table 3 shows the 10 results by using the basic ant colony algorithm to solve the minimum carbon emissions cost model. Table 4 shows the 10 results by using the DNA-ant colony algorithm to solve the minimum carbon emissions cost model. Comparing Table 3 with Table 4 we can find the following.

(1) Carbon emissions cost: the minimum carbon emissions cost of basic ant colony algorithm is 426.5, and the cost of DNA-ant colony algorithm is 398.8. The DNA-ant colony algorithm saves 6.4% of minimum carbon emissions cost compared with basic ant colony algorithm. The average carbon emissions cost of basic ant colony algorithm is 546.2. The average cost of DNA-Ant colony algorithm is 470.34. For the average carbon emissions cost, the DNA-ant colony algorithm saves 13.8% compared with basic ant colony algorithm. Thus, for the solution of minimum carbon emissions cost distribution model, DNA-ant colony algorithm can find the path with less carbon emissions cost, which is better for environmental protection.

(2) The number of final generation: the average number of final generation of basic ant colony algorithm is 133.7. The average number of final generation of DNA-ant colony algorithm is 81. From the average number of final generation, we can find that the DNA-ant colony algorithm has a better convergence.

(3) The difference with minimum carbon emissions cost: the average difference between basic ant colony algorithm and minimum carbon emissions cost is 119.7. The average difference between DNA-ant colony algorithm and minimum carbon emissions cost is 71.54. From the difference with minimum carbon emissions cost, we can find that the DNA-Ant colony algorithm is more stable in the process of search minimum carbon emissions.

Figure 2 is the distribution path of basic ant colony based on the minimum carbon emissions cost model. The processes are as follows:

Distribution Center → Customer13 → Customer10 → Customer11 → Distribution Center;

Table 1: Demand of customer.

Customer	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Demand	15	18	20	8	15	10	25	30	17	6	2	24	19	20	7	1

Table 2: Coordinate of customer.

Customer	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
X	2	8	18	34	−25	35	42	−10	−8	−12	−32	28	−18	−9	−18	8
Y	8	33	20	28	15	−5	10	20	10	−32	−40	7	0	25	8	10

Table 3: The results of basic ant colony algorithm (minimum carbon emissions cost).

No.	1	2	3	4	5	6	7	8	9	**10**	Ave.
Carbon emissions cost	572.3	544.1	580.9	592.6	572.5	544.2	520.7	559.9	548.3	**426.5**	546.2
Vehicles	4	4	4	4	4	4	4	4	4	4	4
End iteration	68	180	144	140	188	79	193	182	67	96	133.7
Deviations	145.8	117.6	154.4	166.1	146	117.7	94.2	133.4	121.8	0	119.7

Table 4: The results of DNA-ant colony algorithm (minimum carbon emissions cost).

No.	1	2	3	4	5	6	7	8	9	**10**	Ave.
Carbon emissions cost	516.2	444.6	493.6	489.1	461.3	484.4	456.4	485.3	473.7	**398.8**	470.34
Vehicles	4	4	4	4	4	4	4	4	4	4	4
End iteration	174	144	79	16	16	46	65	175	34	61	81
Deviations	117.4	45.8	94.8	90.3	62.5	85.6	57.6	86.5	74.9	0	71.54

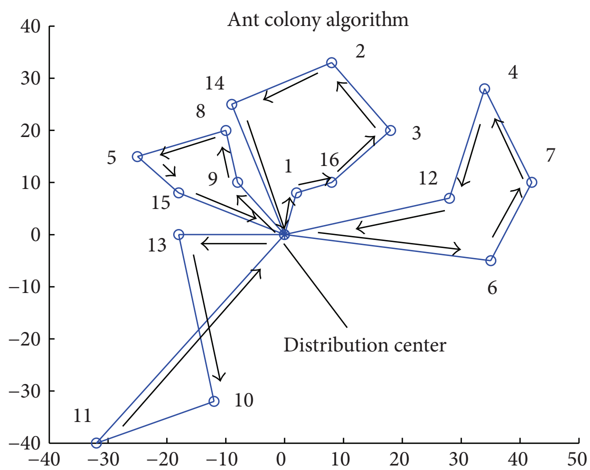


Figure 2: Distribution path basic ant colony algorithm (minimum carbon emissions cost).

Distribution Center → Customer1 → Customer16 → Customer3 → Customer2 → Customer14 → Distribution Center;

Distribution Center → Customer9 → Customer8 → Customer5 → Customer15 → Distribution Center;

Distribution Center → Customer6 → Customer7 → Customer4 → Customer12 → Distribution Center.

Figure 3 is the distribution path of DNA-Ant colony based on the minimum carbon emissions cost model. The processes are as follows:

Distribution Center → Customer10 → Customer11 → Customer13 → Distribution Center;

Distribution Center → Customer6 → Customer7 → Customer4 → Customer12 → Distribution Center;

Distribution Center → Customer8 → Customer14 → Customer5 → Customer15 → Distribution Center;

Distribution Center → Customer1 → Customer16 → Customer3 → Customer2 → Customer9 → Distribution Center.

Figure 4 is the optimization curves of minimum carbon emissions cost. From the figure, we can find that DNA-Ant colony algorithm has a better effectiveness in the process of searching minimum carbon emissions cost. Comparing with basic ant colony algorithm, we can find the lower carbon emissions cost with a faster convergence speed.

When the air pollution level is good level, vehicles are only considered the distribution cost, and the carbon emissions cost is ignored, which means that $\rho_1 = 0$, $\rho_2 = 1$. At this time, the minimum value of objective function is the minimum value of distribution cost.

Table 5 shows the 10 results by using the basic ant colony algorithm to solve the double targets distribution optimization model under the top or good air pollution level condition. Table 6 shows the 10 results by using the DNA-ant colony algorithm to solve the double targets distribution

Table 5: The results of basic ant colony algorithm (top or good air pollution level).

No.	1	2	3	4	5	6	7	8	9	10	Ave.
Cost	403.1	402.7	403.9	404.5	407.5	405.3	**400.7**	405.3	409.5	407.9	405.04
Vehicles	4	4	4	4	4	4	4	4	4	4	4
End iteration	186	159	125	15	178	123	196	5	63	63	111.3
Deviations	2.4	2	3.2	3.8	6.8	4.6	0	4.6	8.8	7.2	4.34

Table 6: The results of DNA-ant colony algorithm (top or good air pollution level).

No.	1	2	3	4	5	6	7	8	9	10	Ave
Cost	400	398	400.7	**398**	**398**	404	**398**	400.7	**398**	407.9	400.33
Vehicles	4	4	4	4	4	4	4	4	4	4	4
End iteration	19	145	77	13	33	16	16	163	27	27	53.6
Deviations	2	0	2.7	0	0	6	0	2.7	0	9.9	2.33

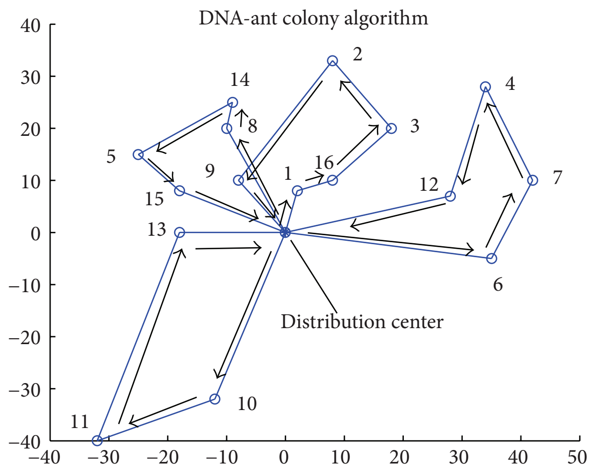


Figure 3: Distribution path DNA-ant colony algorithm (minimum carbon emissions cost).

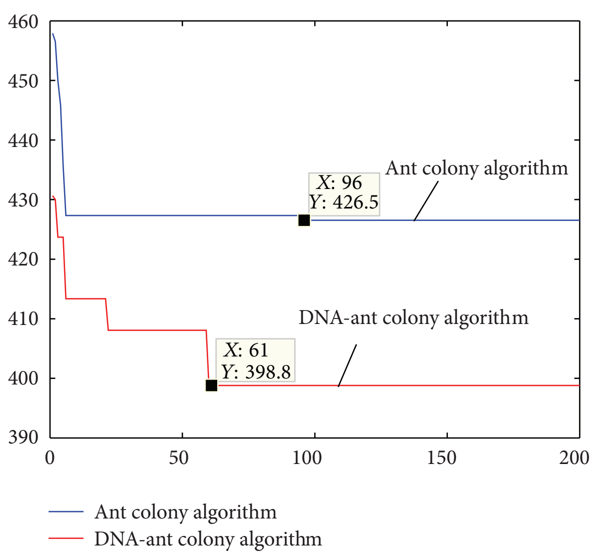


Figure 4: Minimum carbon emissions cost optimization curve.

optimization model under the top or good air pollution level condition. Comparing Table 5 with Table 6 we can find the following.

(1) Distribution cost: the minimum distribution cost of basic ant colony algorithm is 400.7 and the cost of DNA-ant colony algorithm is 398. The DNA-ant colony algorithm saves 0.67% of minimum distribution cost comparing with basic ant colony algorithm. The average distribution cost of basic ant colony algorithm is 405.04. The average cost of DNA-ant colony algorithm is 400.33. For the average distribution cost, the DNA-ant colony algorithm saves 1.16% compared with basic ant colony algorithm. Thus, for the good air pollution level, DNA-ant colony algorithm can find the path with less distribution cost compared with basic ant colony algorithm.

(2) The number of final generation: the average number of final generation of basic ant colony algorithm is 111.3. The average number of final generation of DNA-ant colony algorithm is 53.6. From the average number of final generation, we can find the DNA-ant colony algorithm has a better convergence speed.

(3) The difference with minimum values: the average difference between basic ant colony algorithm and minimum values is 4.34. The average difference between DNA-ant colony algorithm and minimum values is 2.33. From the difference with minimum values, we can find that the DNA-ant colony algorithm is more stable in the process of search the best solution.

Figure 5 is the distribution path figure of basic ant colony algorithm, which is shown in Table 7. Figure 6 is the distribution path figure of DNA-Ant colony algorithm, which is shown in Table 8.

Figure 7 is the distribution cost optimization curves comparing the figure of the two algorithms. From the simulation, we can find that DNA-ant colony algorithm performances is more effective on the solution issue, which can save more distribution cost with a faster convergence speed.

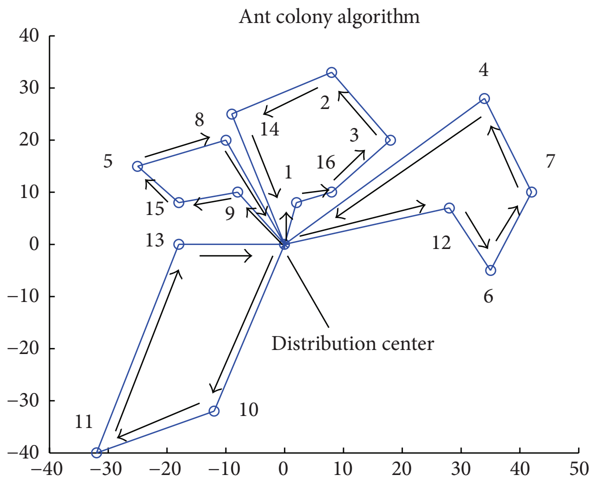


Figure 5: Distribution path basic ant colony algorithm (top or good air pollution level).

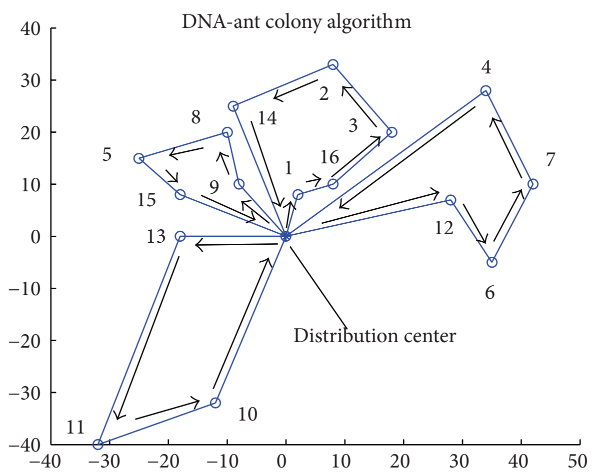


Figure 6: Distribution path DNA-ant colony algorithm (top or good air pollution level).

Table 7: The distribution path figure of basic ant colony algorithm.

Vehicle k	Distribution path
1	Distribution Center–10–11–13–Distribution Center
2	Distribution Center–1–16–3–2–14–Distribution Center
3	Distribution Center–9–15–5–8–Distribution Center
4	Distribution Center–12–6–7–4–Distribution Center

Table 8: The distribution path figure of DNA-ant colony algorithm.

Vehicle k	Distribution path
1	Distribution Center–13–11–10–Distribution Center
2	Distribution Center–1–16–3–2–14–Distribution Center
3	Distribution Center–9–8–5–15–Distribution Center
4	Distribution Center–12–6–7–4–Distribution Center

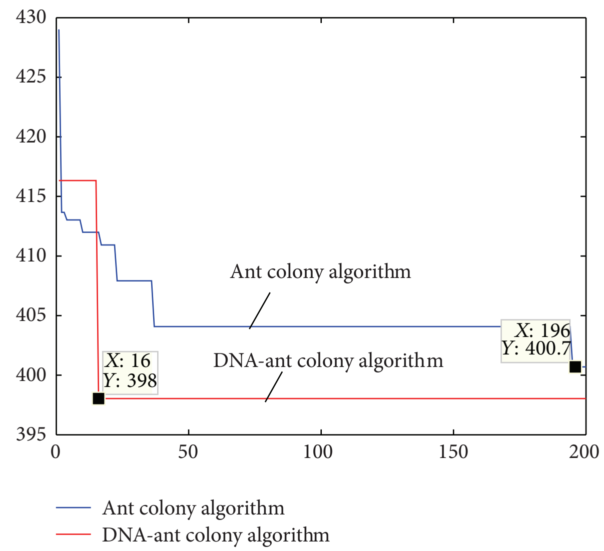


Figure 7: Distribution cost optimization curve (top or good air pollution level).

Assume that the air pollution level is moderate or high level pollution; set $\rho_1 = 0.8$, $\rho_2 = 0.2$. Table 9 shows the 10 results by using the basic ant colony algorithm to solve the double targets distribution optimization model under the moderate or high level pollution condition. Table 10 shows the 10 results by using the DNA-ant colony algorithm to solve the double targets distribution optimization model under the moderate or high level pollution condition. Comparing Table 9 with Table 10 we can find the following.

(1) Objective evaluation function value: the minimum objective evaluation function value of basic ant colony algorithm is 417.5 and the cost of DNA-ant colony algorithm is 402.6. The DNA-ant colony algorithm saves 3.5% of minimum objective evaluation function value compared with basic ant colony algorithm. The average objective evaluation function value of basic ant colony algorithm is 431.22. The average cost of DNA-ant colony algorithm is 414.09. For the average objective evaluation function value, the DNA-ant colony algorithm saves 3.97% compared with basic ant colony algorithm. Thus, for the moderate or high level pollution, DNA-ant colony algorithm can find the path with less objective evaluation function value to find the more effective distribution paths and reduce air pollution and distribution cost.

(2) The number of final generation: the average number of final generation of basic ant colony algorithm is 163.3. The average number of final generation of DNA-ant colony algorithm is 86.9. From the average number of final generation, we can find that the DNA-ant colony algorithm has a better convergence speed.

(3) The difference with minimum values: the average difference between basic ant colony algorithm and minimum values is 13.72. The average difference between DNA-ant colony algorithm and minimum values is

TABLE 9: The results of basic ant colony algorithm (moderate or high level pollution).

No.	1	2	3	4	5	6	7	8	9	10	Ave.
W	430.3	422.4	**417.5**	423.2	458.4	433.2	443.4	437.4	418.4	428	431.22
Vehicles	4	4	4	4	4	4	4	4	4	4	4
End iteration	160	175	193	150	155	162	182	161	163	132	163.3
Deviations	12.8	4.9	0	5.7	40.9	15.7	25.9	19.9	0.9	10.5	13.72

TABLE 10: The results of DNA-ant colony algorithm (moderate or high level pollution).

No.	1	2	3	4	5	6	7	8	9	10	Ave.
W	418.6	**402.6**	**402.6**	409	412.6	428.7	414.8	421.1	415.8	415.1	414.09
Vehicles	4	4	4	4	4	4	4	4	4	4	4
End iteration	126	30	133	68	20	133	26	46	131	156	86.9
Deviations	16	0	0	6.4	10	26.1	12.2	18.5	13.2	12.5	11.49

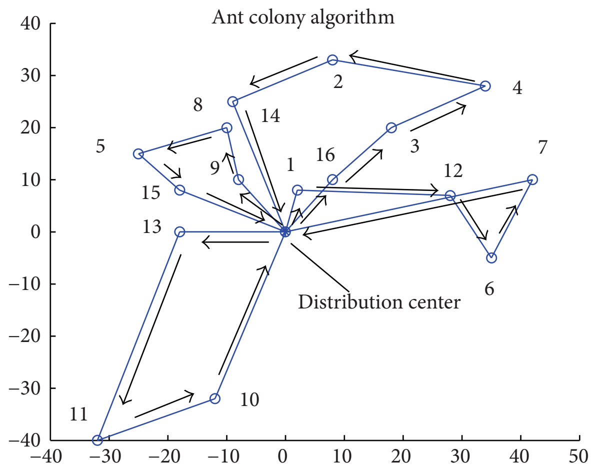


FIGURE 8: Distribution path basic ant colony algorithm (moderate or high level pollution).

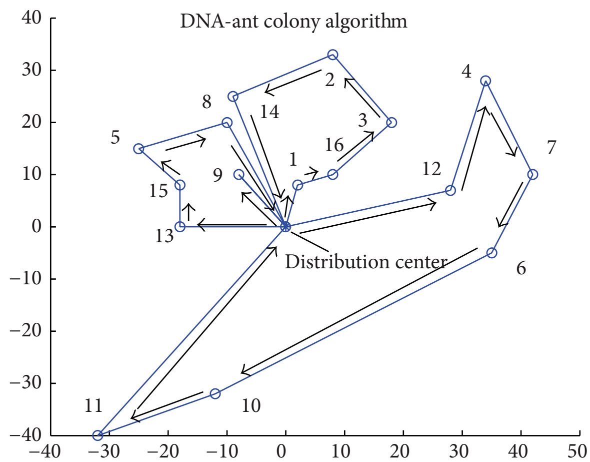


FIGURE 9: Distribution path DNA-ant colony algorithm (moderate or high level pollution).

11.49. From the difference with minimum values, we can find that the DNA-ant colony algorithm is more stable in the process of search the best solution.

Figure 8 is the distribution path of basic ant colony under moderate or high level pollution condition. The processes are as follows:

Vehicle 1: Distribution Center → Customer13 → Customer11 → Customer10 → Distribution Center;

Vehicle 2: Distribution Center → Customer9 → Customer8 → Customer5 → Customer15 → Distribution Center;

Vehicle 3: Distribution Center → Customer16 → Customer3 → Customer4 → Customer2 → Customer14 → Distribution Center;

Vehicle 4: Distribution Center → Customer1 → Customer12 → Customer6 → Customer7 → Distribution Center.

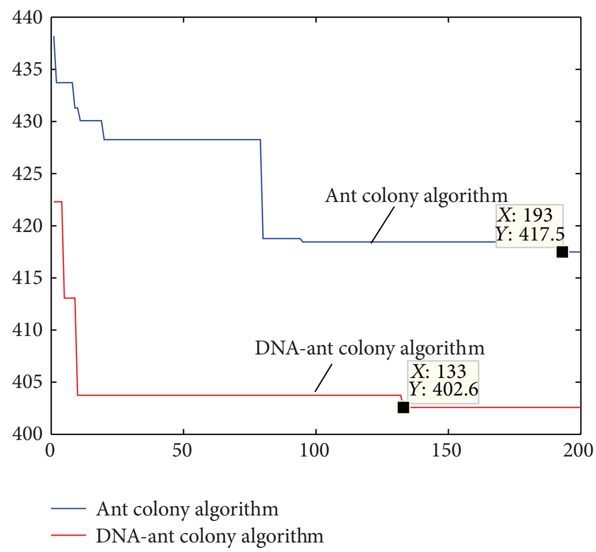


FIGURE 10: Distribution cost optimization curve (moderate or high level pollution).

Ancestral DNA ACTCG||TGCCA GTAC||GTGACA···

GACCT||AGGAT CAGT||TCGAAC···

↓ Cross

Filial generation DNA ACTCG||AGGAT CAGT||TCGAAC···

GACCT||TGCCA GTAC||GTGACA···

Figure 11

Ancestral DNA 00 10 01 10 11||01 11 10 10 00 11 01 00 10||11 01 11 00 10 00···

11 00 10 10 01||00 11 11 00 01 10 00 11 01||01 10 11 00 00 10···

↓ Cross

Filial generation DNA 00 10 01 10 11||00 11 11 00 01 10 00 11 01||01 10 11 00 00 10···

11 00 10 10 01||01 11 10 10 00 11 01 00 10||11 01 11 00 10 00···

Figure 12

Ancestral DNA TAGGC||CAAGT CTGA||CAGATC···

↓ Mutate

Filial generation DNA TAGGC|| CAAGC CTGA||CAGATC···

Ancestral DNA 01 00 11 11 10||10 00 00 11 01 10 01 11 00||10 00 11 00 01 10···

↓ Mutate

Filial generation DNA 01 00 11 11 10||10 00 00 11 10 10 01 11 00||10 00 11 00 01 10···

Figure 13

Figure 9 is the distribution path of DNA-ant colony under moderate or high level pollution condition. The processes are as follows:

Vehicle 1: Distribution Center → Customer1 → Customer16 → Customer3 → Customer2 → Customer14 → Distribution Center;

Vehicle 2: Distribution Center → Customer13 → Customer15 → Customer5 → Customer8 → Distribution Center;

Vehicle 3: Distribution Center → Customer12 → Customer4 → Customer7 → Customer6 → Customer10 → Customer11 → Distribution Center;

Vehicle 4: Distribution Center → Customer9 → Distribution Center.

Figure 10 is the optimization curve of objective evaluation function best value. From the figure, we can find that, under the moderate or high level pollution condition, DNA-ant colony algorithm can find the lower objective evaluation function value to get a faster convergence speed than the basic ant colony algorithm.

5. Conclusions

In this paper, starting from the actual requirements of low carbon logistics, microscopic quantitative analysis was used for low carbon logistics. The minimum cost of carbon emissions model and the double target distribution optimization model with considering the cost of carbon emissions were established to find the reasonable distribution routes to achieve energy conservation and emissions reduction based on the solution of DNA-ant colony algorithm. According to the simulation of MATLAB, DNA-ant colony algorithm had a better effectiveness than the basic ant colony algorithm on the issue of low carbon logistics distribution route optimization. But this is the preliminary research on the low carbon logistics distribution route optimization problem. It is the exploration stage for low carbon logistics distribution route optimization model. The models are established based on the ideal situation without consideration of many complex factors and real situations in the constraints of the models. In future study, there is a to optimize the model to make the model more accord with the actual needs.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported by Soft Science Research Project in Shanxi Province (no. 2010041077-3).

References

[1] X. Xu, "Research on construction and characteristics of low carbon logistics system," *Commercial Times*, no. 10, pp. 23–24, 2011.

[2] X.-Y. Liu, “Analysis of the development of low-carbon logistics based on a low-carbon economy,” in *Proceedings of the International Conference on Low-Carbon Transportation and Logistics and Green Buildings*, pp. 673–679, 2013.

[3] H.-J. Wu and S. Dunn, “Environmentally responsible logisties systems,” *International Journal of Physical Distribution and Logistics Management*, vol. 25, no. 2, pp. 20–38, 1995.

[4] J. P. Rodrigue, B. Slake, and C. Comtois, “Green logistics,” in *The Handbook of Logistics and Supply-Chain Management*, A. M. Brewer, K. J. Button, and D. A. Hensher, Eds., Paragon Elsevier, London, UK, 2001.

[5] A. McKinnon, “Green logistics: the carbon agenda,” *Electronic Scientific Journal of Logistics*, vol. 6, no. 3, pp. 1–9, 2010.

[6] J. Liu and K. Han, “Establishment of logistics system in low-carbon economy,” *Logistics Technology*, no. 13, pp. 77–79, 2013.

[7] X. Lin, “China-ASEAN and Guangxi multiple-regions cooperative low-carbon logistics internet of things service system hidden Markov optimization model,” *Advances in Intelligent and Soft Computing*, vol. 162, pp. 65–72, 2012.

[8] M. Yang and Y. Tian, “Establishment and application of logistics enterprises'low-carbon factors index system,” in *Proceedings of the 2nd International Conference on Logistics, Informatics and Service Science (LISS '12)*, pp. 401–407, July 2012.

[9] S. Cholette and K. Venkat, “The energy and carbon intensity of wine distribution: a study of logistical options for delivering wine to consumers,” *Journal of Cleaner Production*, vol. 17, no. 16, pp. 1401–1413, 2009.

[10] I. Harris, M. Naim, A. Palmer, A. Potter, and C. Mumford, “Assessing the impact of cost optimization based on infrastructure modelling on CO_2 emissions,” *International Journal of Production Economics*, vol. 131, no. 1, pp. 313–321, 2011.

[11] Y. Wang, T. Lu, and C. Zhang, “Integrated logistics network design in hybrid manufacturing/remanufacturing system under low-carbon restriction,” in *Proceedings of the 2nd International Conference on Logistics, Informatics and Service Science (LISS '12)*, pp. 111–121, July 2012.

[12] H. Huang, “A study of developing Chinese low carbon logistics in the new railway period,” in *Proceedings of the International Conference on E-Product E-Service and E-Entertainment (ICEEE '10)*, pp. 635–638, November 2010.

[13] S. Qin, *Research on fuel consumption minimized vehicle routing problem [M.S. thesis]*, Tsinghua University, Beijing, China, 2011.

[14] Y. Le and P. Zhen, “Path design of emergency medical supplies transport,” *Logistics Technology*, vol. 27, no. 11, pp. 84–86, 2008.

[15] M. Dorigo and L. M. Gambardella, “Ant colony system: a cooperative learning approach to the traveling salesman problem,” *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, pp. 53–66, 1997.

[16] C. Wu, *DNA algorithm of the minimum set covering problem in wireless sensor networks research [M.S. thesis]*, East China University of Science and Technology, Shanghai, China, 2010.

[17] X. Chen, *DNA genetic algorithm and its application [Ph.D. thesis]*, Zhejiang University, Zhejiang, China, 2010.

[18] R. J. Lipton, “DNA solution of hard computational problems,” *Science*, vol. 268, no. 5210, pp. 542–545, 1995.

[19] H. Zhu, *Vehicles scheduling problemwith DNA algorithm research [M.S. thesis]*, Harbin University of Science and Technology, Harbin, China, 2003.

[20] L. V. Pin, “Study of logistics network optimization model considering carbon emissions,” *Application Research of Computers*, vol. 30, no. 10, pp. 2977–2980, 2013.

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 839418, 18 pages
http://dx.doi.org/10.1155/2014/839418

Research Article

The Optimal Replenishment Policy under Trade Credit Financing with Ramp Type Demand and Demand Dependent Production Rate

Juanjuan Qin[1] and Weihua Liu[2]

[1] *Business School, Tianjin University of Finance and Economics, Tianjin 300222, China*
[2] *College of Management and Economics, Tianjin University, Tianjin 300222, China*

Correspondence should be addressed to Juanjuan Qin; tjufeqin@163.com

Received 28 February 2014; Accepted 12 May 2014; Published 19 June 2014

Academic Editor: Xiang Li

Copyright © 2014 J. Qin and W. Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper investigates the optimal replenishment policy for the retailer with the ramp type demand and demand dependent production rate involving the trade credit financing, which is not reported in the literatures. First, the two inventory models are developed under the above situation. Second, the algorithms are given to optimize the replenishment cycle time and the order quantity for the retailer. Finally, the numerical examples are carried out to illustrate the optimal solutions and the sensitivity analysis is performed. The results show that if the value of production rate is small, the retailer will lower the frequency of putting the orders to cut down the order cost; if the production rate is high, the demand dependent production rate has no effect on the optimal decisions. When the trade credit is less than the growth stage time, the retailer will shorten the replenishment cycle; when it is larger than the breakpoint of the demand, within the maturity stage of the products, the trade credit has no effect on the optimal order cycle and the optimal order quantity.

1. Introduction

Along with the globalization of the market and increasing competition, the enterprises always take trade credit financing policy to promote sales, increase the market share, and reduce the current inventory levels. As a result, the trade credit financing play an important role in business as a source of funds after Banks or other financial institutions. In the traditional inventory economic order quantity (EOQ) model, it is assumed that the retailer must pay for the products when receiving them. In practice the suppliers often provide the delayed payment time for the payment of the amount owed. Usually, there is no interest charged for the retailer if the outstanding amount is paid in the allowable delay. However, if the payment is unpaid in full by the end of the permissible delay period, interest is charged on the outstanding amount.

The optimal inventory policy is also influenced by the market demand and the production rate with the trade credit. For the market demand, it is always changing fast and influenced by many factors such as the price [1, 2], the inventory level [3, 4], and the stage of the product life time [5, 6]. For the rapid development of the technology and the quickly changing consumer tastes, the products are likely to face the problem of short life cycle, which in return forces the suppliers to frequently promote new products for a competition purpose. A demand model about product introduction is widely used with the ramp type function of time, which assumes that the demand is a linear growth with respect to time (growth stage) and, up to a certain time, reaches a climax and eventually remains to be the peak sales (maturity stage). In addition to the new products introduction, holiday-related products, such as those for Christmas season, are experimentally shown to fit this type of demand rate. On the other hand, when the market demand is better, the supplier will provide a higher production rate; if the market demand is shrinking, the supplier will reduce the production rate. Therefore, the production rate is demand dependent production rate.

The optimal inventory policy is not only related to the market demand but also influenced by the production rate.

As a result, in-depth research is required on the inventory replenishment decisions. Therefore, this study extends the EOQ models in the several ways as follows. First, trade credit financing is introduced to the traditional EOQ models. The retailer is offered by the supplier with a delayed payment time. Second, the ramp type demand is introduced to the EOQ models with the trade credit financing. The demand is nearly a constant in the maturity stage. During the growth time, the demand of the products is increasing with time. Furthermore, the production rate dependent on demand is introduced to the EOQ models with trade credit financing considering the ramp type demand for the first time.

2. Literature Review

2.1. Trade Credit. The trade credit is studied by many scholars from the aspects of finance, accounting, and operations management. Finance and accounting research mainly focused on the study of trade credit enterprise's cash flow, discussing the nature of the trade credit and its impact. In the area of operation management research, mainly from the angle of cash flow and logistics coordination, based on the traditional economic order quantity model framework, weigh the cost of capital cost elements such as fixed ordering cost, storage cost, and profit (or other elements), to analyze the supply chain inventory control and coordination problems. Our research is mainly felt in the operation management flow.

The inventory replenishment policies under trade credit financing have been studied intensively. Most papers discussed the EOQ or EPQ inventory models under trade credit financing all based on the assumptions that the demand rate is a constant and the infinite production rate. The EOQ model with the trade credit financing is put forward for the first time by Goyal [7], which is extended by Chu et al. [8] considering the deteriorated products. The models with the exponential deterioration rate of the items are discussed by Aggarwal and Jaggi [9]. The order strategies are studied by Jamal et al. [10] and Chang and Dye [11] allowing the shortage under delayed payment and deteriorating conditions. The unit selling price is not the same as the unit product costs in the model of Teng [12]. The economic production quantity model is researched by Chung and Huang [13] considering the manufacturer offering the retailer the delayed payment policy. Uthayakumar and Parvathi [14] analyze the supply chain of a single vendor and a single buyer for a single product, taking into consideration the effect of deterioration and credit period incentives with a constant demand and the infinite replenishment rate. In addition, some authors considered the two levels of the delayed payment. Huang [15] first extended Goyal's model to analyze the two levels of the trade credit policy, the manufacturer's credit period available to retailers as M, and the retailers provide the customers a credit period N, $M > N$. And then, Hao et al. [16], Huang [17], and Teng and Chang [18] extend the models of Huang [15]. Recently, Du et al. [19] investigate the coordination of two-echelon supply chains using wholesale price discount and credit option. Wu et al. [20] research the optimal credit period and lot size for deteriorating items with expiration dates under two-level trade credit financing.

2.2. Ramp Type Demand Rate. A number of papers in published literature have extensively studied inventory problem by assuming this ramp type demand rate. Under the assumption, Hwang and Shinn [21] proposed the optimal replenishment policy for perishable seasonal products in a season with ramp type time dependent demand. Panda et al. [22] discussed the inventory models with ramp type demand rate, partial backlogging, and Weibull deterioration rate. Skouri et al. [23] presented an economic production quantity models for deteriorating items with ramp type demand. Manna and Chaudhuri [24] introduced an EOQ inventory model for Weibull distributed deteriorating items under ramp type demand and shortages. Mandal [25] analyzed the supply chain model with stochastic lead time under imprecise partially backlogging and fuzzy ramp-type demand for expiring items. S. R. Singh and C. Singh [26] investigated the inventory model for ramp type demand, time dependent deteriorating items with salvage value, and shortages. Mishra and Singh [27] proposed two-warehouse inventory model with ramp-type demand and partially backlogged shortages. Agrawal and Banerjee [28] discussed a finite time horizon EOQ model with ramp type demand rate under inflation and time discounting with the infinite replenishment rate. Roy and Chaudhuri [29] studied the computational approach to an inventory model with ramp type demand and linear deterioration. Recently, Agrawal et al. [30] analyze a two-warehouse inventory model, with the demand rate being a general ramp type function of time and the shortages are partially backlogged at a constant rate. Ahmed et al. [31] focus on the inventory model with the ramp type demand rate, the partial backlogging, and general deterioration rate. Saha [32] discusses the optimal order quantity of the retailer with quadratic ramp type demand under supplier trade credit financing.

However, the problems of payment delay linked to the ramp type demand rate have not received much attention. The most related literatures to our study are as follows. Mishra and Singh [33] discussed the optimal order policy for single period products under payment delay with ramp type demand rate; Huang et al. [34] made some efforts to build an inventory system for deteriorating products, with ramp type demand rate, under two-level trade credit policy considering the shortages and the partially backlogged unsatisfied demand. But they do not consider the demand dependent production rate.

2.3. Demand Dependent Production Rate. For the trade credit, the researchers attempting to solve the problems mostly assume that the production rate is infinite or a constant value. However, the infinite or a constant replenishment rate of the inventory models is inconsistent with the actual industrial practices. When the market demand is better, the supplier will provide a higher production rate; if the market demand is shrinking, the supplier will reduce the production rate. In the traditional EOQ or EPQ models without the trade credit financing, Darzanou and Skouri [35] presented a production inventory system for deteriorating items with demand rate being a linearly ramp type function of time and production rate being proportional to the demand rate. The two models without shortages and with shortages were discussed. Both models were studied assuming that the time point at which the demand is stabilized occurs before the production stopping time. Manna and Chiang [36] extend this model by considering that (a) for the model with no shortages the demand rate is stabilized after the production stopping time and (b) for the model with shortages the demand rate is stabilized after the production stopping time or after the time when the inventory level reaches zero or after the production restarting time. Recently, Skouri et al. [37] developed a more general integrated supplier retailer inventory model with a demand rate which is sensitive to the retailing price and a demand dependent production rate with two-level of trade credit financing, which did not consider the ramp type demand rate.

Therefore, based on the literatures above, we find that none of the above models explore the optimal replenishment policies of the retailer under trade credit financing with the ramp type demand and demand dependent production rate.

Given the analysis above, this paper developed inventory models under trade credit financing with ramp type demand and demand dependent production rate. In this inventory system, they might attribute to intricate correlations among the period to maturity stage μ, the period of delayed payment M, the production time t_1, and the planning time interval T. We assume that the break point of the demand function μ is less than T. Otherwise, the models in the situation of the ramp type demand are the same as the models with the linear increasing demand function Teng et al. [6], which is not the subjective of this research, and without doubt $t_1 < T$ holds. Still, there exist eight possible subcases discussed in this paper. The remainder of this paper is organized as follows. The notations and assumptions are introduced in Section 3. The inventory levels for $\mu \le t_1$ and $\mu \ge t_1$ are presented in Section 4. Model formulations are developed where we have two inventory models $M \le \mu$ and $M \ge \mu$ including eight subcases in Section 5, for which the algorithms are provided to obtain the optimal decisions. The numerical examples and sensitivity analysis then are conducted in Section 6 in association with managerial insights. Finally, Section 7 closes the paper with a conclusion.

3. Notations and Assumptions

To build the mathematical models, the following notations and assumptions are adopted in this paper.

3.1. Notations

A: Ordering cost per order

t_1: The time at which the inventory level is at maximum for the situation $\mu \le t_1$;

$\bar{t}_1$: The time at which the inventory level is at maximum for the situation $t_1 \le \mu$

S: The maximum inventory level at each scheduling period (cycle)

c_h: The inventory holding cost per unit per unit time (excluding interest charges)

c_p: The purchasing cost per unit

s: The unit sale price

k: The production rate with $k = rf(t)$

R: The market demand with $R = f(t)$

μ: The break point of the demand function

D_0: The rate of demand with the time during the demand increasing stage

$I_i(t)$ $(i = 1, 2, 3)$: The inventory level at time $t \in [0, T]$

I_e: The interest rate earned per unit per unit of time

I_c: The interest rate charged per unit per unit of time

M: The offered credit period by the supplier

T: The replenishment cycle time

$TRC(T)$: The annual total relevant cost.

3.2. Assumptions. The following assumptions are used throughout in this paper.

(1) There is a single supplier and single retailer and they deal with a single product.

(2) The lead time is zero. The planning horizon is infinite. The shortage is not allowed.

(3) The initial and final inventory levels are both zero.

(4) The market demand for the item is assumed to be a ramp type function of time. At the first stage, the demand is increasing with the time, such as the introduction stage of the new products; but as the time increases, the demand will keep a constant, such as the maturity stage of the new product. That is, $R = f(t) = D_0[t-(t-\mu)H(t-\mu)]$ at any time $t \ge 0$, where D_0 and μ are positive constants and $H(t-\mu)$ is the Heaviside's function defined as follows:

$$H(t-\mu) = \begin{cases} 1, & \text{if } t \ge \mu \\ 0, & \text{if } t < \mu. \end{cases} \quad (1)$$

(5) For the production rate, when the market demand is high, the production will be improved; but when the market demand is low, the production rate will be down. Therefore, the production rate is related to the demand closely. In this paper, the production rate is assumed $k = rf(t)$ [38], where $r(>1)$ is a constant. In the production period, when the products have been produced, they will be delivered to the retailer instantaneously. That is, the replenishment rate of the retailer is k.

(6) The retailer would settle the account at $t = M$ and pay for the interest charges on items in stock at the rate I_c over the interval $[M, T]$ as $T \ge M$. Alternatively, the retailer settles the account at $t = M$ and is not

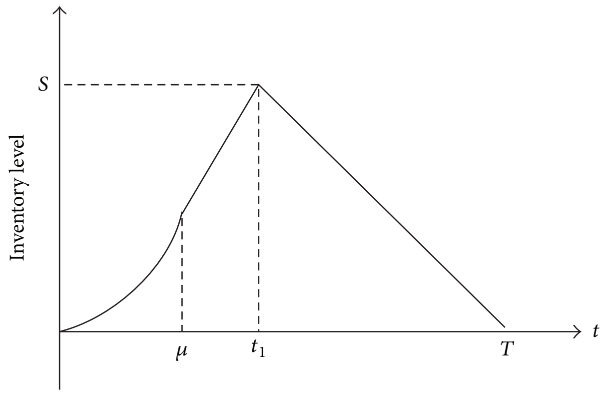


Figure 1: The inventory model for $\mu \le t_1$.

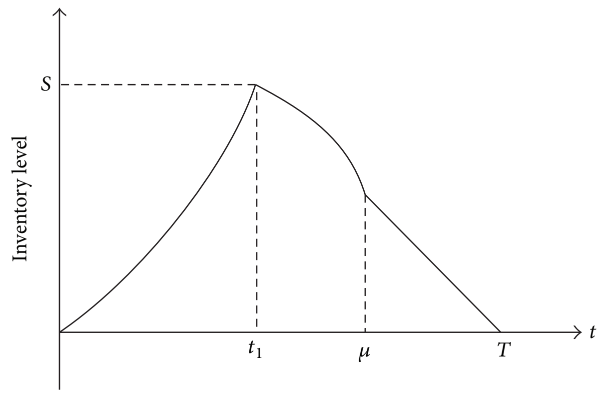


Figure 2: The inventory model for $t_1 \le \mu$.

required to pay any interest charges for items for the stock during the whole cycle as $T \le M$. Before the settlement of the account, the retailer can use sales revenue to earn the interest up to the end of the period M at the rate I_e.

(7) The break point of the demand function μ is less than T.

(8) For the inventory system, we analyze it in the Section 3 for two situations: $\mu \le t_1$ and $t_1 \le \mu$.

Given the above, it is possible to build the mathematical inventory EOQ model with the trade credit financing.

4. The Inventory Level

4.1. Inventory Level for $\mu \le t_1$. Based on the above notations and assumptions, the inventory system can be considered in the following. At the beginning, the stock level is zero. The production starts with zero stock level at time $t = 0$, and the production stops at time $t = t_1$. Then the inventory level gradually diminishes due to demand, till it becomes zero at $t = T$. The whole process is repeated and the behaviour of the inventory system is described in Figure 1. Based on the Figure 1, we can know that the inventory cycle here has the three stages.

Hence, the variation of the inventory level $I(t)$ with respect to the time can be described by the following differential equations.

During the growth stage in the interval $[0, \mu]$, the demand rate is $D_0 t$ and the replenishment rate is $rD_0 t$. Therefore, the inventory level at time t is governed by

$$\frac{dI_1(t)}{dt} = (r-1) D_0 t, \quad 0 \le t \le \mu, \tag{2}$$

with the boundary conditions $I_1(0) = 0$.

Similarly, during the maturity stage $[\mu, t_1]$, the demand rate is $D_0\mu$ and the replenishment rate is $rD_0\mu$. Therefore, the inventory level at time t is governed by

$$\frac{dI_2(t)}{dt} = (r-1) D_0 \mu, \quad \mu \le t \le t_1, \tag{3}$$

with the boundary conditions $I_2(t_1) = S$.

Finally, at the maturity stage $[t_1, T]$, the demand rate is $D_0\mu$ and the production stops. Therefore, the inventory level at time t is governed by

$$\frac{dI_3(t)}{dt} = -D_0 \mu, \quad t_1 \le t \le T, \tag{4}$$

with the boundary conditions $I_3(t_1) = S$ and $I_3(T) = 0$.

Solving (2)–(4), the values of $I_i(t)$ $(i = 1, 2, 3)$ are as follows:

$$\begin{aligned} I_1(t) &= (r-1)\frac{1}{2} D_0 t^2, \quad 0 \le t \le \mu, \\ I_2(t) &= (r-1) D_0 \mu t - (r-1)\frac{1}{2} D_0 \mu^2, \quad \mu \le t \le t_1, \\ I_3(t) &= D_0 \mu (T - t), \quad t_1 \le t \le T. \end{aligned} \tag{5}$$

In addition, using the boundary condition $I_2(t_1) = I_3(t_1) = S$, we obtain

$$(r-1) D_0 \mu t_1 - (r-1)\frac{1}{2} D_0 \mu^2 = D_0 \mu (T - t_1). \tag{6}$$

Solving (6), we can obtain

$$t_1 = \frac{T + (1/2)(r-1)\mu}{r}. \tag{7}$$

4.2. Inventory Level for $t_1 \le \mu$. The inventory system for $t_1 \le \mu$ considered in the following is the same as that in the situations of $t_1 \ge \mu$. The behaviour of the inventory system is described in Figure 2. Based on the Figure 2, we can know that the inventory cycle here still has the three stages. The variation of the inventory level $I(t)$ with respect to the time can be described by the following differential equations.

During the growth stage in the interval $[0, t_1]$, the demand rate is $D_0 t$ and the replenishment rate is $rD_0 t$. Therefore, the inventory level at time t is governed by

$$\frac{dI_1(t)}{dt} = (r-1) D_0 t, \quad 0 \le t \le t_1, \tag{8}$$

with the boundary conditions $I_1(0) = 0$ and $I_1(t_1) = S$.

Similarly, during the maturity stage $[t_1, \mu]$, the demand rate is $D_0 t$ and the replenishment rate is zero. Therefore, the inventory level at time t is governed by

$$\frac{dI_2(t)}{dt} = -D_0 t, \quad t_1 \le t \le \mu, \tag{9}$$

with the boundary conditions $I_2(t_1) = S$.

Finally, at the maturity stage $[\mu, T]$, the demand rate is $D_0\mu$ and there is no production. Therefore, the inventory level at time t is governed by

$$\frac{dI_3(t)}{dt} = -D_0\mu, \quad \mu \le t \le T, \tag{10}$$

with the boundary conditions $I_3(T) = 0$.

Solving (2)–(4), the values of $I_i(t)$ $(i = 1, 2, 3)$ are as follows:

$$\begin{aligned} I_1(t) &= (r-1)\frac{1}{2}D_0 t^2, \quad 0 \le t \le t_1, \\ I_2(t) &= -\frac{1}{2}D_0 t^2 + \frac{1}{2}D_0\mu^2 + D_0\mu(T-\mu), \quad t_1 \le t \le \mu, \\ I_3(t) &= D_0\mu(T-t), \quad \mu \le t \le T. \end{aligned} \tag{11}$$

In addition, using the boundary condition $I_1(t_1) = I_2(t_1) = S$, we obtain

$$(r-1)\frac{1}{2}D_0 t_1^2 = -\frac{1}{2}D_0 t_1^2 + \frac{1}{2}D_0\mu^2 + D_0\mu(T-\mu). \tag{12}$$

Solving (12), we can obtain

$$\bar{t}_1 = \sqrt{\frac{-\mu^2 + 2\mu T}{r}}. \tag{13}$$

5. Mathematical Models and Solutions

The annual total relevant cost consists of the following elements: ordering cost, holding cost, interest payable, and interest earned. The components are evaluated as in the following.

(1) Annual ordering cost: A/T.

(2) Annual stock holding cost (excluding the interest charges).

For $\mu \le t_1$, the annual stock holding cost is

$$\begin{aligned} &\frac{c_h}{T}\left[\int_0^{\mu} I_1(t)\,dt + \int_{\mu}^{t_1} I_2(t)\,dt + \int_{t_1}^{T} I_3(t)\,dt\right] \\ &\quad = \frac{c_h}{T}\left[(r-1)\frac{1}{6}D_0\mu^3 + \frac{1}{2}rD_0\mu t_1^2 \right. \\ &\qquad \left. - (r-1)\frac{1}{2}D_0\mu^2 t_1 + \frac{1}{2}D_0\mu T^2 - D_0\mu t_1 T\right]. \end{aligned} \tag{14}$$

For $t_1 \le \mu$, the annual stock holding cost is

$$\begin{aligned} &\frac{c_h}{T}\left[\int_0^{\bar{t}_1} I_1(t)\,dt + \int_{\bar{t}_1}^{\mu} I_2(t)\,dt + \int_{\mu}^{T} I_3(t)\,dt\right] \\ &\quad = \frac{c_h}{T}\left[\frac{1}{6}rD_0\bar{t}_1^3 - \frac{1}{6}D_0\mu^3 \right. \\ &\qquad \left. + \frac{1}{2}D_0\mu^2\bar{t}_1 - D_0\mu T\bar{t}_1 + \frac{1}{2}D_0\mu T^2\right]. \end{aligned} \tag{15}$$

(3) There are eight cases to occur in interest earned and interest charged for the items kept in stock per year.

From the value of M, we have two possible situations: $M \le \mu$ and $M \ge \mu$. The different mathematical formulations are discussed in the following according to the two possible situations including eight cases.

5.1. Model 1: The Inventory Models for the Case $M \le \mu$

5.1.1. Case 1

Case 1.1 ($t_1 \le M \le \mu \le T$, shown in Figure 3). The annual interest charged will be paid for the inventory not being sold after the due date M. During the time $[M, \mu]$, the inventory level is $I_2(t)$; for the time $[\mu, T]$, the inventory level is $I_3(t)$. Therefore, the annual interest charged for the inventory not being sold after the due date M is given by

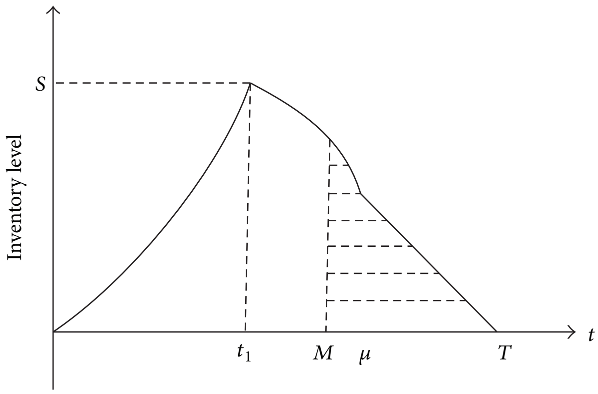
$$\begin{aligned} &\frac{c_p I_c}{T}\left[\int_M^{\mu} I_2(t)\,dt + \int_{\mu}^{T} I_3(t)\,dt\right] \\ &\quad = \frac{c_p I_c}{T}\left[-\frac{1}{6}D_0\mu^3 + \frac{1}{6}D_0 M^3 \right. \\ &\qquad \left. + \frac{1}{2}D_0\mu^2 M - D_0\mu T M + \frac{1}{2}D_0\mu T^2\right]. \end{aligned} \tag{16}$$

During the time $[0, M]$, the retailer can use the sales to accumulate the interest. Therefore, the annual interest earned during the trade credit period is shown in the following:

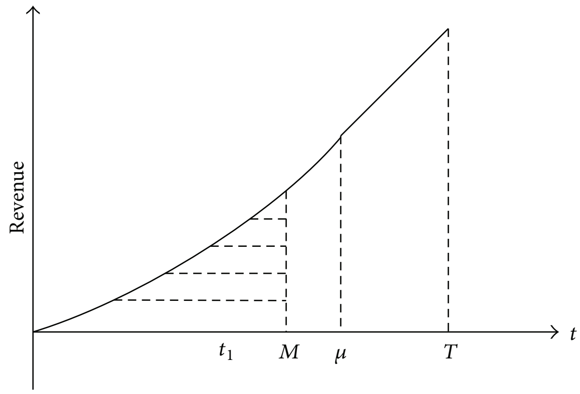
$$\frac{sI_e}{T}\left[\int_0^M (M-t)D_0 t\,dt\right] = \frac{sI_e}{6T}D_0 M^3. \tag{17}$$

Case 1.2 ($M \le t_1 \le \mu \le T$, shown in Figure 4). The annual interest charged will be paid for the inventory not being sold after the due date M. During the time $[M, \bar{t}_1]$, the inventory level is $I_1(t)$; for the time $[\bar{t}_1, \mu]$, the inventory level is $I_2(t)$; for the time $[\mu, T]$, the inventory level is $I_3(t)$. Therefore, the annual interest charged for the inventory not being sold after the due date M is given by

$$\begin{aligned} &\frac{c_p I_c}{T}\left[\int_M^{\bar{t}_1} I_1(t)\,dt + \int_{\bar{t}_1}^{\mu} I_2(t)\,dt + \int_{\mu}^{T} I_3(t)\,dt\right] \\ &\quad = \frac{c_p I_c}{T}\left[\frac{1}{6}rD_0\bar{t}_1^3 - \frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu^2\bar{t}_1 \right. \\ &\qquad \left. - D_0\mu T\bar{t}_1 + \frac{1}{2}D_0\mu T^2 - (r-1)\frac{1}{6}D_0 M^3\right]. \end{aligned} \tag{18}$$

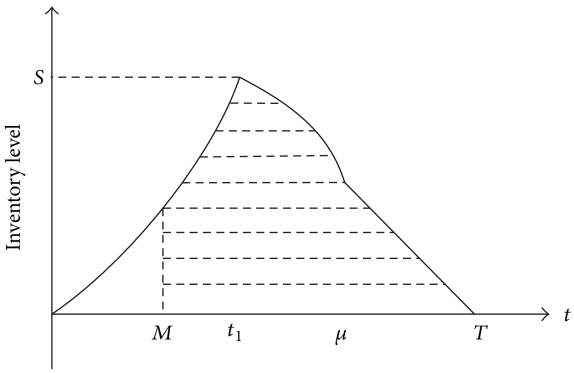


(a) Interest charged for $t_1 \le M \le \mu \le T$

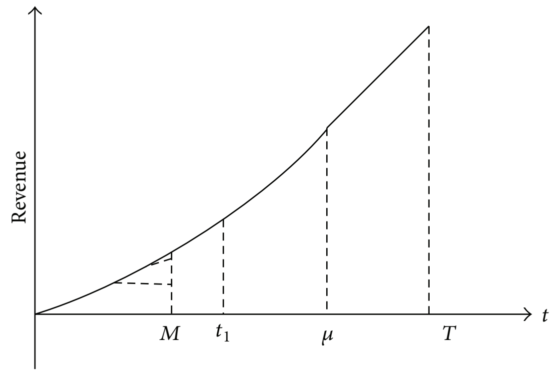


(b) Interest earned for $t_1 \le M \le \mu \le T$

FIGURE 3: Interest charged and interest earned for $t_1 \le M \le \mu \le T$.



(a) Interest charged for $M \le t_1 \le \mu \le T$



(b) Interest earned for $M \le t_1 \le \mu \le T$

FIGURE 4: Interest charged and interest earned for $M \le t_1 \le \mu \le T$.

During the time $[0, M]$, the retailer can use the sales to accumulate the interest. The annual interest earned during the trade credit period is presented in the following:

$$\frac{sI_e}{T}\left[\int_0^M (M-t)\,D_0 t\,dt\right] = \frac{sI_e}{6T}D_0M^3. \tag{19}$$

Case 1.3 ($M \le \mu \le t_1 \le T$, shown in Figure 5). The annual interest charged will be paid for the inventory not being sold after the due date M. During the time $[M, \mu]$, the inventory level is $I_1(t)$; for the time $[\mu, t_1]$, the inventory level is $I_2(t)$; for the time $[t_1, T]$, the inventory level is $I_3(t)$. The annual interest charged for the inventory not being sold after the due date M is shown by

$$\begin{aligned}\frac{c_pI_c}{T}&\left[\int_M^{\mu} I_1(t)\,dt + \int_{\mu}^{t_1} I_2(t)\,dt + \int_{t_1}^{T} I_3(t)\,dt\right]\\ &= \frac{c_pI_c}{T}\left[(r-1)\frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu^2 t_1\right.\\ &\qquad\left. + \frac{1}{2}D_0\mu T^2 - D_0\mu t_1 T - \frac{1}{6}(r-1)D_0M^3\right].\end{aligned} \tag{20}$$

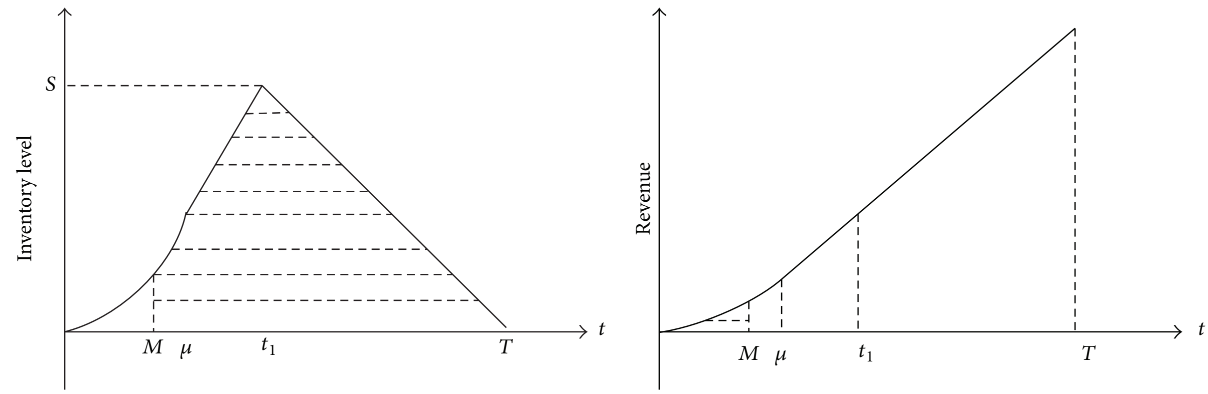
For the interval $[0, M]$, the retailer can use the sales to accumulate the interest. The annual interest earned during the trade credit period is presented in the following:

$$\frac{sI_e}{T}\left[\int_0^M (M-t)\,D_0 t\,dt\right] = \frac{sI_e}{6T}D_0M^3. \tag{21}$$

5.2. The Optimal Replenishment Policy for the Model 1. The results in previous Section 4.1 lead to the following total cost functions.

(1) If $\mu \le (M^2r+\mu^2)/2\mu$, that is, $\mu/\sqrt{r} \le M \le \mu$

$$TRC_1(T) = \begin{cases} TRC_{11}, & t_1 \le M \le \mu \le T, \text{ that is } \mu \le T \le \dfrac{M^2r+\mu^2}{2\mu} \\ TRC_{12}, & M \le t_1 \le \mu \le T, \\ & \text{that is } \dfrac{M^2r+\mu^2}{2\mu} \le T \le \dfrac{1}{2}(r+1)\mu \\ TRC_{13}, & M \le \mu \le t_1 \le T, \text{ that is } \dfrac{1}{2}(r+1)\mu \le T. \end{cases} \tag{22}$$



(a) Interest charged for $M \leq \mu \leq t_1 \leq T$

(b) Interest earned for $M \leq \mu \leq t_1 \leq T$

FIGURE 5: Interest charged and interest earned for $M \leq \mu \leq t_1 \leq T$.

(2) If $\mu \geq (M^2 r + \mu^2)/2\mu$, that is $0 \leq M \leq \mu/\sqrt{r}$

$$TRC_1(T) = \begin{cases} TRC_{12}, & M \leq t_1 \leq \mu \leq T, \text{ that is } \mu \leq T \leq \dfrac{1}{2}(r+1)\mu \\ TRC_{13}, & M \leq \mu \leq t_1 \leq T, \text{ that is } \dfrac{1}{2}(r+1)\mu \leq T, \end{cases} \tag{23}$$

where

$$\begin{aligned} TRC_{11} &= \frac{A}{T} + \frac{c_h}{T}\left(\frac{1}{6}rD_0\bar{t}_1^3 - \frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu^2\bar{t}_1 - D_0\mu T\bar{t}_1 + \frac{1}{2}D_0\mu T^2\right) \\ &\quad + \frac{c_p I_c}{T}\left(-\frac{1}{6}D_0\mu^3 + \frac{1}{6}D_0M^3 + \frac{1}{2}D_0\mu^2 M - D_0\mu T M + \frac{1}{2}D_0\mu T^2\right) - \frac{sI_e}{6T}D_0M^3; \\ TRC_{12} &= \frac{A}{T} + \frac{c_h}{T}\left(\frac{1}{6}rD_0\bar{t}_1^3 - \frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu^2\bar{t}_1 - D_0\mu T\bar{t}_1 + \frac{1}{2}D_0\mu T^2\right) \\ &\quad + \frac{c_p I_c}{T}\left[\frac{1}{6}rD_0\bar{t}_1^3 - \frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu^2\bar{t}_1 - D_0\mu T\bar{t}_1 + \frac{1}{2}D_0\mu T^2 - (r-1)\frac{1}{6}D_0M^3\right] - \frac{sI_e}{6T}D_0M^3; \\ TRC_{13} &= \frac{A}{T} + \frac{c_h}{T}\left[(r-1)\frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu^2 t_1 + \frac{1}{2}D_0\mu T^2 - D_0\mu t_1 T\right] \\ &\quad + \frac{c_p I_c}{T}\left[(r-1)\frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu^2 t_1 + \frac{1}{2}D_0\mu T^2 - D_0\mu t_1 T - \frac{1}{6}(r-1)D_0M^3\right] - \frac{sI_e}{6T}D_0M^3. \end{aligned} \tag{24}$$

The problem is $\min_T TRC_1(T)$. In order to obtain the optimal solutions, we should study each of the branches and then combine the results.

Theorem 1. *(1) If $\alpha < 0$, $TRC_{13}(T)$ is a convex function of T for $T = (0, \infty)$. The optimal value of T_{13} corresponds to $\min\{TRC_{13}(\widetilde{T}_{13}), TRC_{13}((1/2)(r+1)\mu)\}$.*

(2) If $\alpha \geq 0$, $TRC_{13}(T)$ is an increasing function with T for $T = (0, \infty)$.

For $T \in [(1/2)(r+1)\mu, \infty)$, the optimal value T_{13} is obtained when $T_{13} = (1/2)(r+1)\mu$.

It is similar to analyze the optimal solutions for other branches as Theorem 1. Therefore, we can obtain the solution procedure for determining the optimal replenishment cycle and the order quantity of the model 1 in the following algorithm.

Algorithm 1.

Step 1. Input the values for all parameters with $M \leq \mu$.

Step 2. Compare $\mu/\sqrt{r}$ and M. If $\mu/\sqrt{r} \leq M$, go to Algorithm 1.1; if $\mu/\sqrt{r} \geq M$, then go to Algorithm 1.2.

Algorithm 1.1

Step 1. Find the global minimum of $TRC_{11}(T)$, says T_{11}, as follows.

Step 1.1. Compute $\widetilde{T}_{11}$, if $\mu \leq \widetilde{T}_{11} \leq (M^2 r + \mu^2)/2\mu$ and $f''_{11}(\widetilde{T}_{11}) > 0$, find the $\min\{TRC_{11}(\widetilde{T}_{11}), TRC_{11}(\mu), TRC_{11}((M^2 r +$

$\mu^2)/2\mu)\}$. And accordingly set T_{11}. Compute Q_{11} and $TRC_{11}(T_{11})$. Else, go to Step 1.2.

Step 1.2. Find the $\min\{TRC_{11}(\mu), TRC_{11}((M^2r+\mu^2)/2\mu)\}$ and accordingly set T_{11}. Compute Q_{11} and $TRC_{11}(T_{11})$.

Step 2. Find the global minimum of $TRC_{12}(T)$, says T_{12}, as follows.

Step 2.1. Compute $\widetilde{T}_{12}$, if $(M^2r+\mu^2)/2\mu \le \widetilde{T}_{12} \le (r+1)\mu/2$ and $f''_{12}(\widetilde{T}_{12}) > 0$, find the $\min\{TRC_{12}(\widetilde{T}_{12}), TRC_{12}((r+1)\mu/2), TRC_{12}((M^2r+\mu^2)/2\mu)\}$ and accordingly set T_{12}. Compute Q_{12} and $TRC_{12}(T_{12})$. Else, go to Step 2.2.

Step 2.2. Find the $\min\{TRC_{12}((r+1)\mu/2), TRC_{12}((M^2r+\mu^2)/2\mu)\}$ and accordingly set T_{12}. Compute Q_{12} and $TRC_{12}(T_{12})$.

Step 3. Find the global minimum of $TRC_{13}(T)$, says T_{13}, as follows.

Step 3.1. Compute $\widetilde{T}_{13}$, if $(r+1)\mu/2 \le \widetilde{T}_{13}$, and then set $T^*_{13} = T_{13}$. Compute Q_{13} and $TRC_{13}(T_{13})$. Else, go to Step 3.2.

Step 3.2. Set $T_{13} = (r+1)\mu/2$. Compute Q_{13} and $TRC_{13}(T_{13})$.

Step 4. Find the $\min\{TRC_{11}(T_{11}), TRC_{12}(T_{12}), TRC_{13}(T_{13})\}$ and accordingly select the optimal value for T_1. Stop.

Algorithm 1.2

Step 1. Find the global minimum of $TRC_{12}(T)$, says T_{12}, as follows.

Step 1.1. Compute $\widetilde{T}_{12}$, if $\mu \le \widetilde{T}_{12} \le (r+1)\mu/2$ and $f''_{12}(\widetilde{T}_{12}) > 0$, find the $\min\{TRC_{12}(\widetilde{T}_{12}), TRC_{12}((r+1)\mu/2), TRC_{12}(\mu)\}$ and accordingly set T_{12}. Compute Q_{12} and $TRC_{12}(T_{12})$. Else, go to Step 2.2.

Step 1.2. Find the $\min\{TRC_{12}((r+1)\mu/2), TRC_{12}(\mu)\}$ and accordingly set T_{12}. Compute Q_{12} and $TRC_{12}(T_{12})$.

Step 2. Find the global minimum of $TRC_{13}(T)$, says T_{13}, as follows.

Step 2.1. Compute $\widetilde{T}_{13}$, if $(r+1)\mu/2 \le \widetilde{T}_{13}$, and then set $T_{13} = \widetilde{T}_{13}$. Compute Q_{13} and $TRC_{13}(T_{13})$. Else, go to Step 2.2.

Step 2.2. Set $T_{13} = (r+1)\mu/2$. Compute Q_{13} and $TRC_{13}(T_{13})$.

Step 3. Find the $\min\{TRC_{12}(T_{12}), TRC_{13}(T_{13})\}$ and accordingly select the optimal value for T_1. Stop.

5.3. Model 2: The Inventory Models for the Case $\mu \le M$

5.3.1. Case 2 ($\mu \le M \le T$)

Case 2.1 ($t_1 \le \mu \le M \le T$). Similar to Case 1, the annual interest charged for the inventory not being sold after the due date M is given by

$$\frac{c_pI_c}{T}\int_M^T I_3(t)\,dt = \frac{c_pI_c}{T}\left(\frac{1}{2}D_0\ \mu T^2 - D_0\mu TM + \frac{1}{2}D_0\mu M^2\right). \tag{25}$$

The annual interest earned during the trade credit period is shown in the following:

$$\begin{aligned}&\frac{sI_e}{T}\left[\int_0^\mu (M-t)\,D_0t\,dt + \int_\mu^M (M-t)\,D_0\mu\,dt\right]\\&\quad= \frac{sI_e}{T}\left(-\frac{1}{2}MD_0\mu^2 + \frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu M^2\right).\end{aligned} \tag{26}$$

Case 2.2 ($\mu \le t_1 \le M \le T$). The annual interest charged for the inventory not being sold after the due date M is given by

$$\begin{aligned}&\frac{c_pI_c}{T}\left[\int_M^T I_3(t)\,dt\right]\\&\quad= \frac{c_pI_c}{T}\left(\frac{1}{2}D_0\mu T^2 - D_0\mu TM + \frac{1}{2}D_0\mu M^2\right).\end{aligned} \tag{27}$$

The annual interest earned during the trade credit period is shown in the following:

$$\begin{aligned}&\frac{sI_e}{T}\left[\int_0^\mu (M-t)\,D_0t\,dt + \int_\mu^M (M-t)\,D_0\mu\,dt\right]\\&\quad= \frac{sI_e}{T}\left(-\frac{1}{2}MD_0\mu^2 + \frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu M^2\right).\end{aligned} \tag{28}$$

Case 2.3 ($\mu \le M \le t_1 \le T$). The annual interest charged for the inventory not being sold after the due date M is given by

$$\begin{aligned}&\frac{c_pI_c}{T}\left[\int_M^{t_1} I_2(t)\,dt + \int_{t_1}^T I_3(t)\,dt\right]\\&\quad= \frac{c_pI_c}{T}\left(-\frac{1}{3}D_0t_1^3 + \frac{1}{3}D_0M^3 - \frac{1}{2}D_0\mu^2M\right.\\&\qquad\left.- D_0\mu TM + D_0\mu^2M + D_0\mu T^2 - \frac{1}{2}D_0\mu^2T\right).\end{aligned} \tag{29}$$

The annual interest earned during the trade credit period is shown in the following:

$$\begin{aligned}&\frac{sI_e}{T}\left[\int_0^\mu (M-t)\,D_0t\,dt + \int_\mu^M (M-t)\,D_0\mu\,dt\right]\\&\quad= \frac{sI_e}{T}\left(-\frac{1}{2}MD_0\mu^2 + \frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu M^2\right).\end{aligned} \tag{30}$$

5.3.2. Case 3: The Inventory Models for the Case $T < M$

Case 3.1 ($t_1 \le \mu \le T < M$). The annual interest charged for the inventory not being sold after the due date M is zero.

The annual interest earned during the trade credit period is shown in the following:

$$\begin{aligned}&\frac{sI_e}{T}\left[\int_0^T (M-t)\,D_0t\,dt + \int_0^{\bar{t}_1} K(t)\,(M-T)\,dt\right]\\&\quad= \frac{sI_e}{T}\left[\frac{1}{2}MD_0T^2 - \frac{1}{3}D_0T^3 + rD_0\,(M-T)\,\frac{1}{2}\bar{t}_1^2\right].\end{aligned} \tag{31}$$

Case 3.2 ($\mu \le t_1 \le T < M$). The annual interest charged for the inventory not being sold after the due date M is zero.

The annual interest earned during the trade credit period is shown in the following:

$$\frac{sI_e}{T}\left[\int_0^T (M-t)\,D_0 t\,dt + \int_0^{t_1} K(t)(M-T)\,dt\right]$$
$$= \frac{sI_e}{T}\left[\frac{1}{2}MD_0T^2 - \frac{1}{3}D_0T^3 + rD_0(M-T)\frac{1}{2}\mu^2 + rD_0(M-T)(t_1-\mu)\right]. \tag{32}$$

5.4. The Optimal Replenishment Policy for the Model 2. Combine the inventory models of Cases 2 and 3, the results in previous subsections lead to the following total cost function.

(1) If $(1/2)(r+1)\mu \ge M \ge \mu$

$$TRC_2(T) = \begin{cases} TRC_{31}, & t_1 \le \mu \le T \le M, \text{ that is } \mu \le T \le M \\ TRC_{21}, & t_1 \le \mu \le M \le T, \\ & \text{that is } M \le T \le \frac{1}{2}(r+1)\mu \\ TRC_{22}, & \mu \le t_1 \le M \le T, \\ & \text{that is } \frac{1}{2}(r+1)\mu \le T \le Mr - \frac{1}{2}(r-1)\mu \\ TRC_{23}, & \mu \le M \le t_1 \le T, \\ & \text{that is } Mr - \frac{1}{2}(r-1)\mu \le T. \end{cases} \tag{33}$$

(2) If $(1/2)(r+1)\mu \le M$

$$TRC_2(T) = \begin{cases} TRC_{31}, & t_1 \le \mu \le T \le M, \text{ that is } \mu \le T \le \frac{1}{2}(r+1)\mu \\ TRC_{32}, & \mu \le t_1 \le T \le M, \\ & \text{that is } \frac{1}{2}(r+1)\mu \le T \le M \\ TRC_{22}, & \mu \le t_1 \le M \le T, \\ & \text{that is } M \le T \le Mr - \frac{1}{2}(r-1)\mu \\ TRC_{23}, & \mu \le M \le t_1 \le T, \\ & \text{that is } Mr - \frac{1}{2}(r-1)\mu \le T, \end{cases} \tag{34}$$

where

$$TRC_{21} = \frac{A}{T} + \frac{c_h}{T}\left(\frac{1}{6}rD_0\bar{t}_1^3 - \frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu^2\bar{t}_1 - D_0\mu T\bar{t}_1 + \frac{1}{2}D_0\mu T^2\right) + \frac{c_pI_c}{T}\left(\frac{1}{2}D_0\mu T^2 - D_0\mu TM + \frac{1}{2}D_0\mu M^2\right) - \frac{sI_e}{T}\left(-\frac{1}{2}MD_0\mu^2 + \frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu M^2\right);$$

$$TRC_{22} = \frac{A}{T} + \frac{c_h}{T}\left[(r-1)\frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu^2 t_1 + \frac{1}{2}D_0\mu T^2 - D_0\mu t_1 T\right] + \frac{c_pI_c}{T}\left(\frac{1}{2}D_0\mu T^2 - D_0\mu TM + \frac{1}{2}D_0\mu M^2\right) - \frac{sI_e}{T}\left(-\frac{1}{2}MD_0\mu^2 + \frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu M^2\right);$$

$$TRC_{23} = \frac{A}{T} + \frac{c_h}{T}\left[(r-1)\frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu^2 t_1 + \frac{1}{2}D_0\mu T^2 - D_0\mu t_1 T\right] + \frac{c_pI_c}{T}\left(-\frac{1}{3}D_0t_1^3 + \frac{1}{3}D_0M^3 + \frac{1}{2}D_0\mu^2 M - D_0\mu TM + D_0\mu T^2 - \frac{1}{2}D_0\mu^2 T\right) - \frac{sI_e}{T}\left(-\frac{1}{2}MD_0\mu^2 + \frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu M^2\right),$$

$$TRC_{31} = \frac{A}{T} + \frac{c_h}{T}\left(\frac{1}{6}rD_0\bar{t}_1^3 - \frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu^2\bar{t}_1 - D_0\mu T t_1 + \frac{1}{2}D_0\mu T^2\right) - \frac{sI_e}{T}\left(\frac{1}{2}MD_0T^2 - \frac{1}{3}D_0T^3 + rD_0(M-T)\frac{1}{2}\bar{t}_1^2\right);$$

$$TRC_{32} = \frac{A}{T} + \frac{c_h}{T}\left[(r-1)\frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu^2 t_1\right.$$

$$+\frac{1}{2}D_0\mu T^2 - D_0\mu t_1 T\Bigg]$$

$$-\frac{sI_e}{T}\left[\frac{1}{2}MD_0T^2 - \frac{1}{3}D_0T^3 + rD_0(M-T)\frac{1}{2}\mu^2\right.$$

$$\left.+rD_0(M-T)(t_1-\mu)\right]. \quad (35)$$

In order to obtain the optimal replenishment decisions, we should study each branch of the cost functions.

Theorem 2. (1) *If $\beta < 0$, $TRC_{23}(T)$ is a convex function of T for $T = (0,\infty)$. The optimal value of T_{23} corresponds to* $\min\{TRC_{23}(\widetilde{T}_{23}), TRC_{23}(Mr-(1/2)(r-1)\mu)\}$.

(2) *If $\beta \geq 0$, $TRC_{23}(T)$ is an increasing function with T for $T = (0,\infty)$. For $T \in [Mr-(1/2)(r-1)\mu,\infty)$, the optimal value T_{23} is obtained by $T_{23} = Mr-(1/2)(r-1)\mu$.*

It is similar to analyze the optimal solutions for other branches as Theorem 2. Based on the discussion, the solution procedure for determining the optimal replenishment cycle and the order quantity in the model 2 with $M \geq \mu$ can be summarized into the following algorithm.

Algorithm 2.

Step 1. Input the values for all parameters with $M \geq \mu$.

Step 2. Compare $(r+1)\mu/2$ and M. If $(r+1)\mu/2 \geq M \geq \mu$, go to Algorithm 2.1; if $M \geq (r+1)\mu/2$, then go to Algorithm 2.2.

Algorithm 2.1

Step 1. Find the global minimum of $TRC_{31}(T)$, says T_{31}, as following.

Step 1.1. Compute $\widetilde{T}_{31}$, if $\mu \leq \widetilde{T}_{31} \leq M$ and $f''_{31}(\widetilde{T}_{31}) > 0$, find the $\min\{TRC_{31}(\widetilde{T}_{31}), TRC_{31}(\mu), TRC_{31}(M)\}$ and accordingly set T_{31}. Compute Q_{31} and $TRC_{31}(T_{31})$. Else, go to Step 1.2.

Step 1.2. Find the $\min\{TRC_{31}(\mu), TRC_{31}(M)\}$ and accordingly set T_{31}. Compute Q_{31} and $TRC_{31}(T_{31})$.

Step 2. Find the global minimum of $TRC_{21}(T)$, says T_{21}, as follows.

Step 2.1. Compute $\widetilde{T}_{21}$, if $M \leq \widetilde{T}_{21} \leq (r+1)\mu/2$ and $f''_{21}(\widetilde{T}_{21}) > 0$, find the $\min\{TRC_{21}(\widetilde{T}_{21}), TRC_{21}(M), TRC_{21}((r+1)\mu/2)\}$ and accordingly set T_{21}. Compute Q_{21} and $TRC_{21}(T_{21})$. Else, go to Step 2.2.

Step 2.2. Find the $\min\{TRC_{21}(M), TRC_{21}((r+1)\mu/2)\}$ and accordingly set T_{21}. Compute Q_{21} and $TRC_{21}(T_{21})$.

Step 3. Find the global minimum of $TRC_{22}(T)$, says T_{22}, as follows.

Step 3.1. Compute $\widetilde{T}_{22}$, if $(r+1)\mu/2 \leq \widetilde{T}_{22} \leq Mr-(r+1)\mu/2$ and $f''_{22}(\widetilde{T}_{22}) > 0$, find the $\min\{TRC_{22}(\widetilde{T}_{22}), TRC_{22}((r+1)\mu/2), TRC_{22}(Mr-(r+1)\mu/2)\}$ and accordingly set T_{22}. Compute Q_{22} and $TRC_{22}(T_{22})$. Else, go to Step 3.2.

Step 3.2. Find the $\min\{TRC_{22}((r+1)\mu/2), TRC_{22}(Mr-(r+1)\mu/2)\}$ and accordingly set T_{22}. Compute Q_{22} and $TRC_{22}(T_{22})$.

Step 4. Find the global minimum of $TRC_{23}(T)$, says T_{23}, as follows.

Step 4.1. Compute $\widetilde{T}_{23}$, if $(r+1)\mu/2 \leq \widetilde{T}_{23}$, and then set $T_{23} = \widetilde{T}_{23}$. Compute Q_{23} and $TRC_{23}(T_{23})$. Else, go to Step 4.2.

Step 4.2. Set $T_{23} = (r+1)\mu/2$. Compute Q_{23} and $TRC_{23}(T_{23})$.

Step 5. Find the $\min\{TRC_{31}(T_{21}), TRC_{21}(T_{21}), TRC_{22}(T_{22}), TRC_{23}(T_{23})\}$ and accordingly select the optimal value for T_2. Stop.

Algorithm 2.2

Step 1. Find the global minimum of $TRC_{31}(T)$, says T_{31}, as follows.

Step 1.1. Compute $\widetilde{T}_{31}$, if $\mu \leq \widetilde{T}_{31} \leq (r+1)\mu/2$ and $f''_{31}(\widetilde{T}_{31}) > 0$, find the $\min\{TRC_{31}(\widetilde{T}_{31}), TRC_{31}(\mu), TRC_{31}((r+1)\mu/2)\}$ and accordingly set T_{31}. Compute Q_{31} and $TRC_{31}(T_{31})$. Else, go to Step 1.2.

Step 1.2. Find the $\min\{TRC_{31}(\mu), TRC_{31}((r+1)\mu/2)\}$ and accordingly set T_{31}. Compute Q_{31} and $TRC_{31}(T_{31})$.

Step 2. Find the global minimum of $TRC_{32}(T)$, says T_{32}, as follows.

Step 2.1. Compute $\widetilde{T}_{32}$, if $(r+1)\mu/2 \leq \widetilde{T}_{32} \leq M$ and $f''_{32}(\widetilde{T}_{32}) > 0$, find the $\min\{TRC_{32}(\widetilde{T}_{32}), TRC_{32}((r+1)\mu/2), TRC_{32}(M)\}$ and accordingly set T_{32}. Compute Q_{32} and $TRC_{32}(T_{32})$. Else, go to Step 2.2.

Step 2.2. Find the $\min\{TRC_{32}((r+1)\mu/2), TRC_{32}(M)\}$ and accordingly set T_{32}. Compute Q_{32} and $TRC_{32}(T_{32})$.

Step 3. Find the global minimum of $TRC_{22}(T)$, says T_{22}, as follows.

Step 3.1. Compute $\widetilde{T}_{22}$, if $M \leq \widetilde{T}_{22} \leq Mr-(r+1)\mu/2$ and $f''_{22}(\widetilde{T}_{22}) > 0$, find the $\min\{TRC_{22}(\widetilde{T}_{22}), TRC_{22}(M), TRC_{22}(Mr-(r+1)\mu/2)\}$ and accordingly set T_{22}. Compute Q_{22} and $TRC_{22}(T_{22})$. Else, go to Step 3.2.

Step 3.2. Find the $\min\{TRC_{22}(M), TRC_{22}(Mr-(r+1)\mu/2)\}$ and accordingly set T_{22}. Compute Q_{22} and $TRC_{22}(T_{22})$.

Step 4. Find the global minimum of $TRC_{23}(T)$, says T_{23}, as follows.

Step 4.1. Compute $\widetilde{T}_{23}$, if $(r+1)\mu/2 \leq \widetilde{T}_{23}$, and then set $T_{23} = \widetilde{T}_{23}$. Compute Q_{23} and $TRC_{23}(T_{23})$. Else, go to Step 4.2.

Step 4.2. Set $T_{23} = (r+1)\mu/2$. Compute Q_{23} and $TRC_{23}(T_{23})$.

Step 5. Find the $\min\{TRC_{31}(T_{21}), TRC_{32}(T_{32}), TRC_{22}(T_{22}), TRC_{23}(T_{23})\}$ and accordingly select the optimal value for T_2. Stop.

6. Numerical Examples and Sensitivity Analysis

In this section, we carry out some numerical examples to illustrate the algorithms obtained in the previous sections. Additionally, we also provide a sensitivity analysis of the values of most important parameters on the optimal decisions of the retailer and the total cost.

6.1. Numerical Examples

Example 1 ($\mu/\sqrt{r} \le M \le \mu$). The input parameters are $A =$ \$50 per order, $D_0 = 500$ units, $r = 2$, $s =$ \$100 per unit, $c_h =$ \$5 per unit per year, $c_p =$ \$40 per unit, $I_e = 0.15$, $I_c = 0.08$, $M = 0.3$ year, and $\mu = 0.4$ year.

Using Algorithm 1, we can calculate $TRC_{11}(T_{11}) = 351.8$ with $T_{11} = 0.4$, $TRC_{12}(T_{12}) = 393.2$ with $T_{12} = 0.4250$, and $TRC_{13}(T_{13}) = 142.6$ with $T_{13} = 0.6000$. Therefore, the optimal order cycle for the retailer is $T_1 = T_{13} = 0.6000$ year and the optimal order quantity is $Q_1 = 80$ units. The minimum of the total cost is $TRC_1 =$ \$142.6.

Example 2 ($0 \le M \le \mu/\sqrt{r}$). The input parameters are $A =$ \$50 per order, $D_0 = 500$ units, $r = 2$, $s =$ \$100 per unit, $c_h =$ \$5 per unit per year, $c_p =$ \$40 per unit, $I_e = 0.15$, $I_c = 0.08$, $M = 0.2$ year, and $\mu = 0.4$ year.

Using Algorithm 1, we can calculate $TRC_{12}(T_{12}) = 441.6$ with $T_{12} = 0.4000$, and $TRC_{13}(T_{13}) = 190.7$ with $T_{13} = 0.6000$. Therefore, the optimal order cycle for the retailer is $T_1 = T_{13} = 0.6000$ year and the optimal order quantity is $Q_1 = 80$ units. The minimum of the total cost is $TRC_1 =$ \$190.7.

Example 3 ($(r+1)\mu/2 \ge M \ge \mu$). The input parameters are $A =$ \$50 per order, $D_0 = 500$ units, $r = 2$, $s =$ \$40 per unit, $c_h =$ \$5 per unit per year, $c_p =$ \$20 per unit, $I_e = 0.15$, $I_c = 0.08$, $M = 0.6$ year, and $\mu = 0.4$ year.

Using Algorithm 2, we can calculate $TRC_{31}(T_{31}) = 92.3$ with $T_{31} = 0.4000$, $TRC_{21}(T_{21}) = 347.0$ with $T_{21} = 0.6000$, $TRC_{22}(T_{22}) = 107.2$ with $T_{22} = 0.8000$, and $TRC_{23}(T_{23}) = 533.2$ with $T_{23} = 1.4$. Therefore, the optimal order cycle for the retailer is $T_1 = T_{31} = 0.4000$ year and the optimal order quantity is $Q_1 = 40$ units. The minimum of the total cost is $TRC_2 =$ \$92.3.

Example 4 ($(r+1)\mu/2 \le M$). The input parameters are $A =$ \$50 per order, $D_0 = 500$ units, $r = 2$, $s =$ \$40 per unit, $c_h =$ \$5 per unit per year, $c_p =$ \$20 per unit, $I_e = 0.15$, $I_c = 0.08$, $M = 0.9$ year, and $\mu = 0.4$ year.

Using Algorithm 2, we can calculate $TRC_{31}(T_{31}) = 576.1$ with $T_{31} = 0.6367$, $TRC_{32}(T_{32}) = 261.7$ with $T_{32} = 0.8000$, $TRC_{22}(T_{22}) = 416.3$ with $T_{22} = 0.9000$, and $TRC_{23}(T_{23}) = 1140.1$ with $T_{23} = 2.3$. Therefore, the optimal order cycle for the retailer is $T_1 = T_{31} = 0.4000$ year and the optimal order quantity is $Q_1 = 120$ units. The minimum of the total cost is $TRC_2 =$ \$261.7.

6.2. Sensitivity Analysis.

In the model, there are many parameters influencing the decisions of the members. But according to the past research, the effects of some parameters on the decisions are analyzed in many literatures such as the ordering cost per order. Therefore, we choose three parameters, which are directly related to our innovations, to conduct the sensitivity analysis for obtaining interesting management insights. In this subsection, we present the sensitivity analysis of the models mentioned above with respect to the three parameters of r, M, and μ.

6.2.1. The Impact of the Changes of the Parameter r.

Basic parameters' values for the case $M \le \mu$ are $A =$ \$50 per order, $D_0 = 500$ units, $s =$ \$100 per unit, $c_h =$ \$5 per unit per year, $c_p =$ \$40 per unit, $I_e = 0.15$, $I_c = 0.08$, $M = 0.3$ year, and $\mu = 0.6$ year. The impacts of the changes of the parameter r on the retailer' decisions and cost are shown in Figure 6.

Basic parameters' values for the case $M \ge \mu$ are $A =$ \$250 per order, $D_0 = 500$ units, $s =$ \$40 per unit, $c_h =$ \$5 per unit per year, $c_p =$ \$20 per unit, $I_e = 0.15$, $I_c = 0.08$, $M = 0.7$ year, and $\mu = 0.3$ year. The impacts of the changes of the parameter r on the retailer' decisions and its cost are shown in Figure 7.

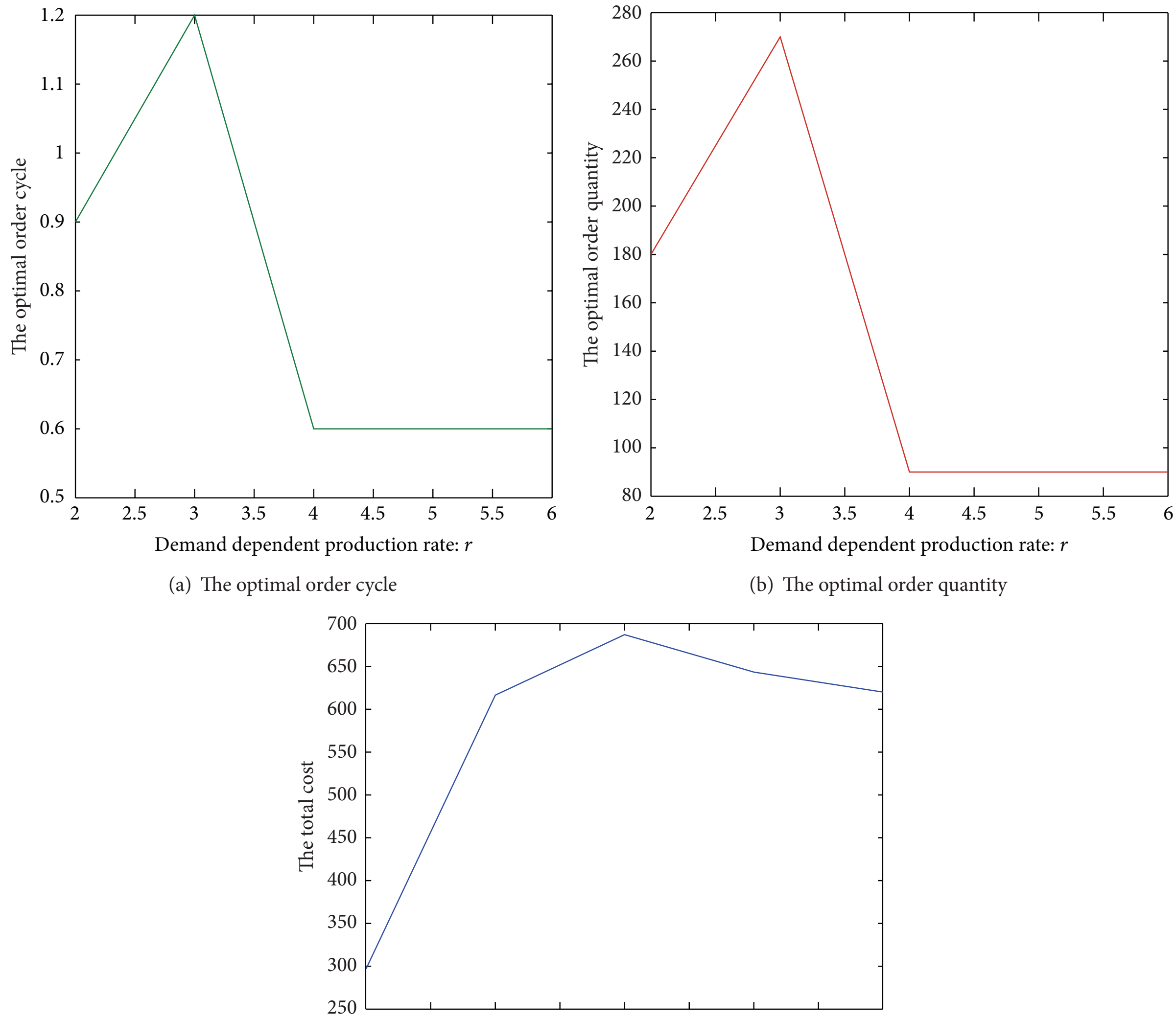
Based on Figures 6 and 7, we can see that if the demand dependent production rate is lower, as the increasing of the parameter r, the optimal order cycle and the optimal order quantity are increasing. It means that if the value of production rate is small, the retailer will lower the frequency of putting the orders to cut down the order cost. If the demand dependent rate is higher, the demand dependent production rate has no effect on the optimal order cycle and the order quantity. For the limit case, when the production rate is infinite, as most papers presented, the production rate has no influence on the members' decisions.

For the total cost, as the increase of the parameter r, the total cost is first increasing and then decreasing with this parameter. It demonstrated that when the value of production rate is small, the on-hand inventory level is low, which results in low total cost. When the demand dependent production rate is much higher, the retailer's profit is decreasing with this rate, because the retailer can get the products as soon as possible to better meet the demand of the market; on the other hand, in general, there is not shortage to occur for the retailer, which will save the shortage cost for the retailer.

6.2.2. The Impact of the Changes of the Parameter M.

Basic parameters' values are $A =$ \$250 per order, $D_0 = 500$ units, $s =$ \$40 per unit, $c_h =$ \$5 per unit per year, $c_p =$ \$20 per unit, $I_e = 0.15$, $I_c = 0.08$, $r = 2$, and $\mu = 0.5$ year.

Based on Figure 8, we can see that as the delayed payment time offered by the suppliers is increasing till $M = \mu$, the optimal order cycle and the optimal order quantity are nonincreasing, which demonstrate that when the delayed payment time is less than the growth stage time of the



(a) The optimal order cycle

(b) The optimal order quantity

(c) The total cost for the retailer

FIGURE 6: The impact of the changes of the parameter r when $M \leq \mu$.

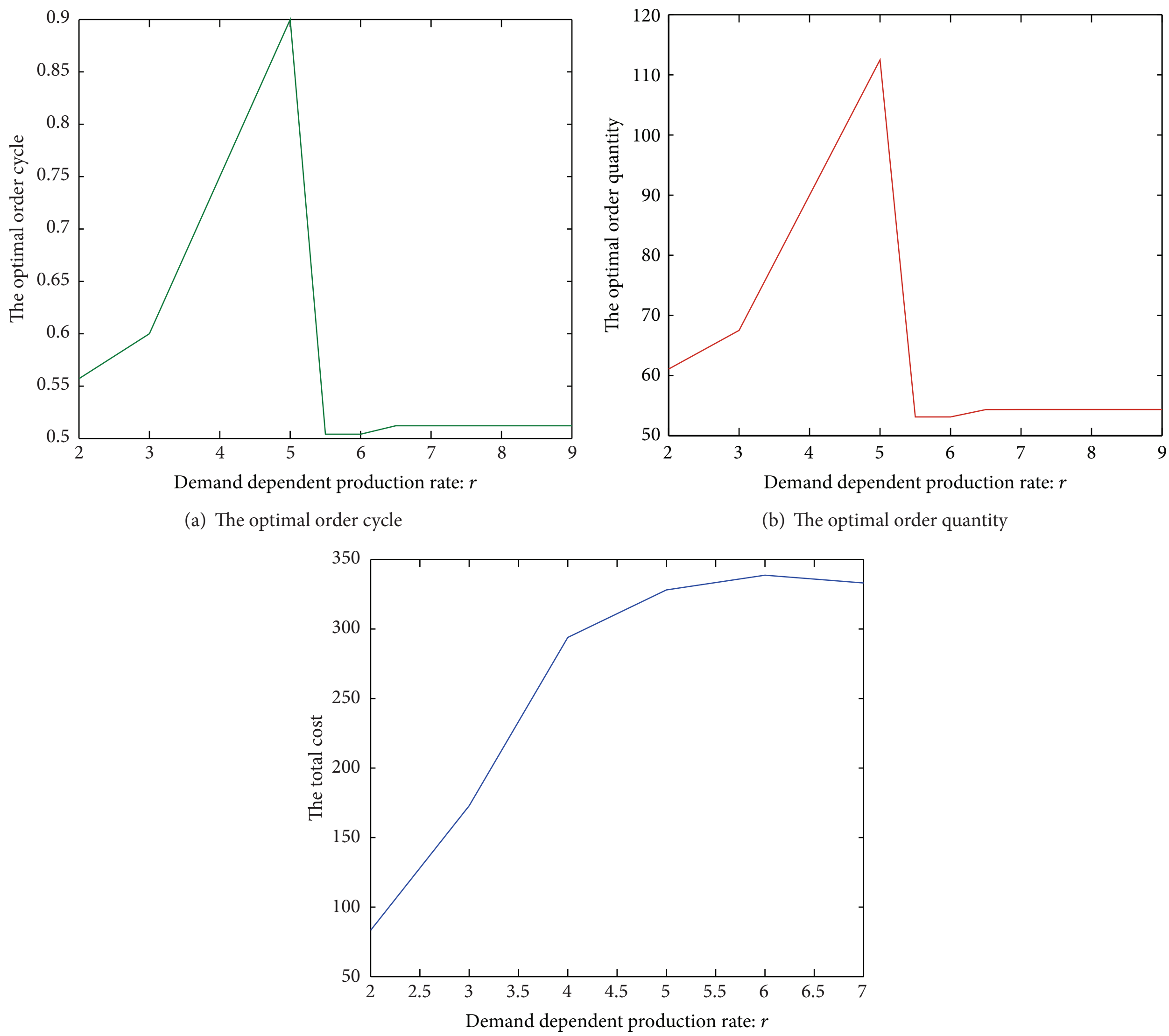
new products' introduction, the retailer will shorten the replenishment cycle to take advantage of the trade credit more frequently for accumulating the interest; on the other hand, the short order cycle can make the retailer adjust its order decisions more quickly for meeting the changeable demand within the growth stage of the products.

When $M > \mu$, that is, the delayed payment time is larger than the breakpoint of the demand, during the maturity stage of the products, the delayed payment time has no effect on the optimal order cycle and the optimal order quantity. Therefore, during the maturity stage of the new products, the retailer's replenishment policies are not influenced by the delayed payment time offered by the supplier.

However, the total cost for the retailer is decreasing with the delayed payment time for the product life cycle. Therefore, the retailer hopes the supplier can offer them the delayed payment time as long as possible to obtain more profits.

6.2.3. The Impact of the Changes of the Parameter μ. Basic parameters' values are A = \$250 per order, D_0 = 500 units, s = \$40 per unit, c_h = \$5 per unit per year, c_p = \$20 per unit, $I_e = 0.15$, $I_c = 0.08$, $r = 2$, and $M = 0.5$ year.

Based on Figure 9, we can see that as the demand breakpoint is increasing; the optimal order cycle and the optimal order quantity are decreasing at short time firstly and then increasing for the last time. That is, when the breakpoint is large, the new products moving to the maturity period at slow speed, the demand is changed slowly. Therefore, in



(a) The optimal order cycle

(b) The optimal order quantity

(c) The total cost for the retailer

Figure 7: The impact of the changes of the parameter r when $M \geq \mu$.

the relative stable market, the retailer can increase the order quantity and replenishment cycle to lower the cost.

The total cost for the retailer is increasing with the demand breakpoint. Hence, for the products, if the growth stage of the product life cycle is longer, the cost of the retailer will be higher. Therefore, the retailer should make some measures, such as advertisement, to induce the products' demand moving to the maturity stage as soon as possible.

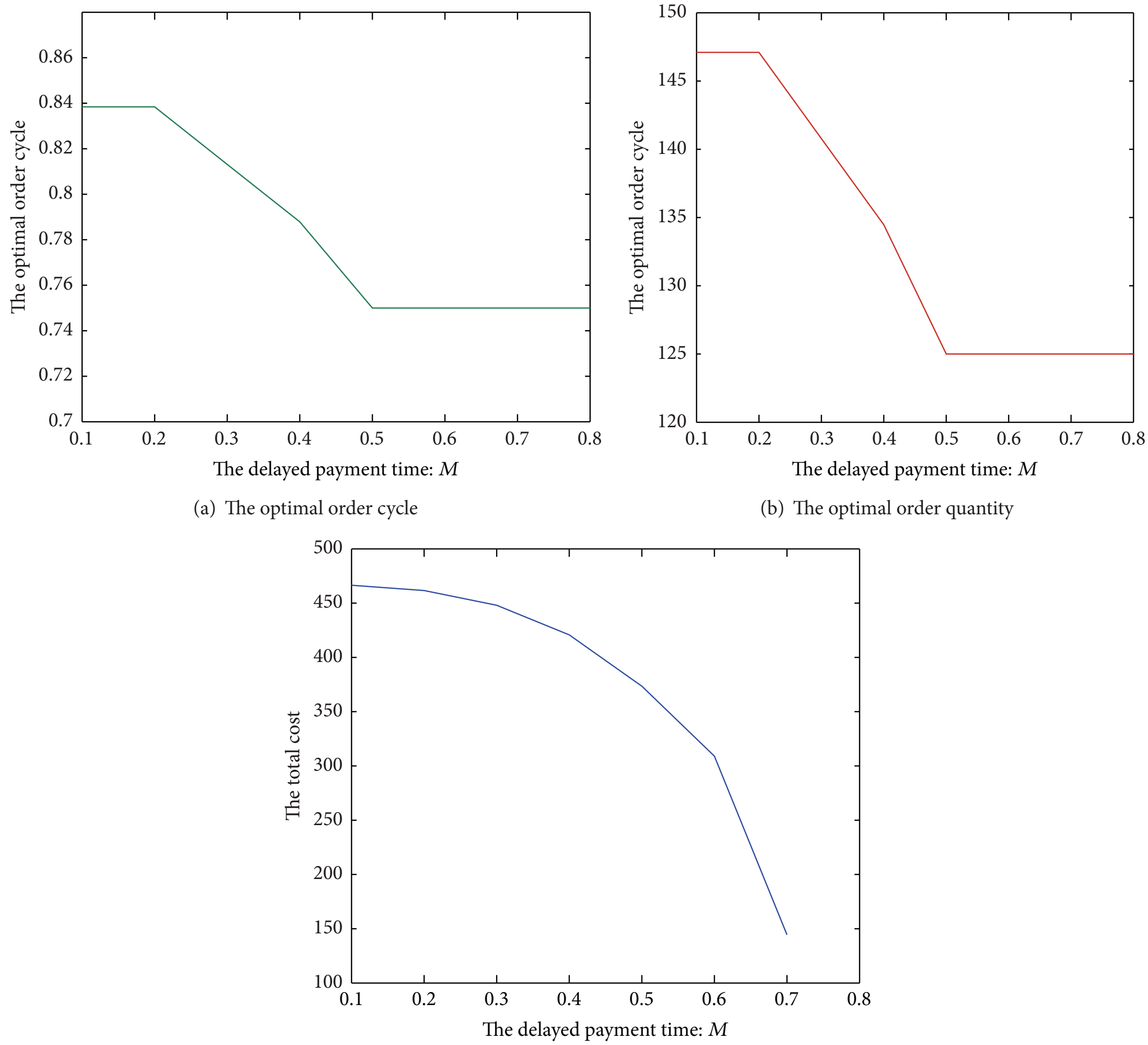
7. Conclusions

Most of the existing inventory models under trade credit financing are assumed that the demand is a constant and the replenishment rate is infinite or a constant. However, in practice, the demand rate is the ramp type function of time for some cases, such as the new products and the holiday related products. On the other hand, the production rate is related to the market demand. When the market is better, the production rate will be improved. When the retailer makes an order, it can be met faster. That is, the replenishment of the retailer is related to the market demand.

Therefore, in this paper, we developed an EOQ model under trade credit financing with ramp type demand and the demand dependent production rate. Subsequently, the algorithms are proposed to decide the optimal replenishment cycle and the optimal order quantity for the retailer. Finally, the numerical analysis is demonstrated to illustrate the models and the sensitively analysis is carried out to give some management insights.

Based on the study, we mainly found the following.

(1) If the value of production rate is small, the retailer will lower the frequency of putting the orders to cut down



(a) The optimal order cycle

(b) The optimal order quantity

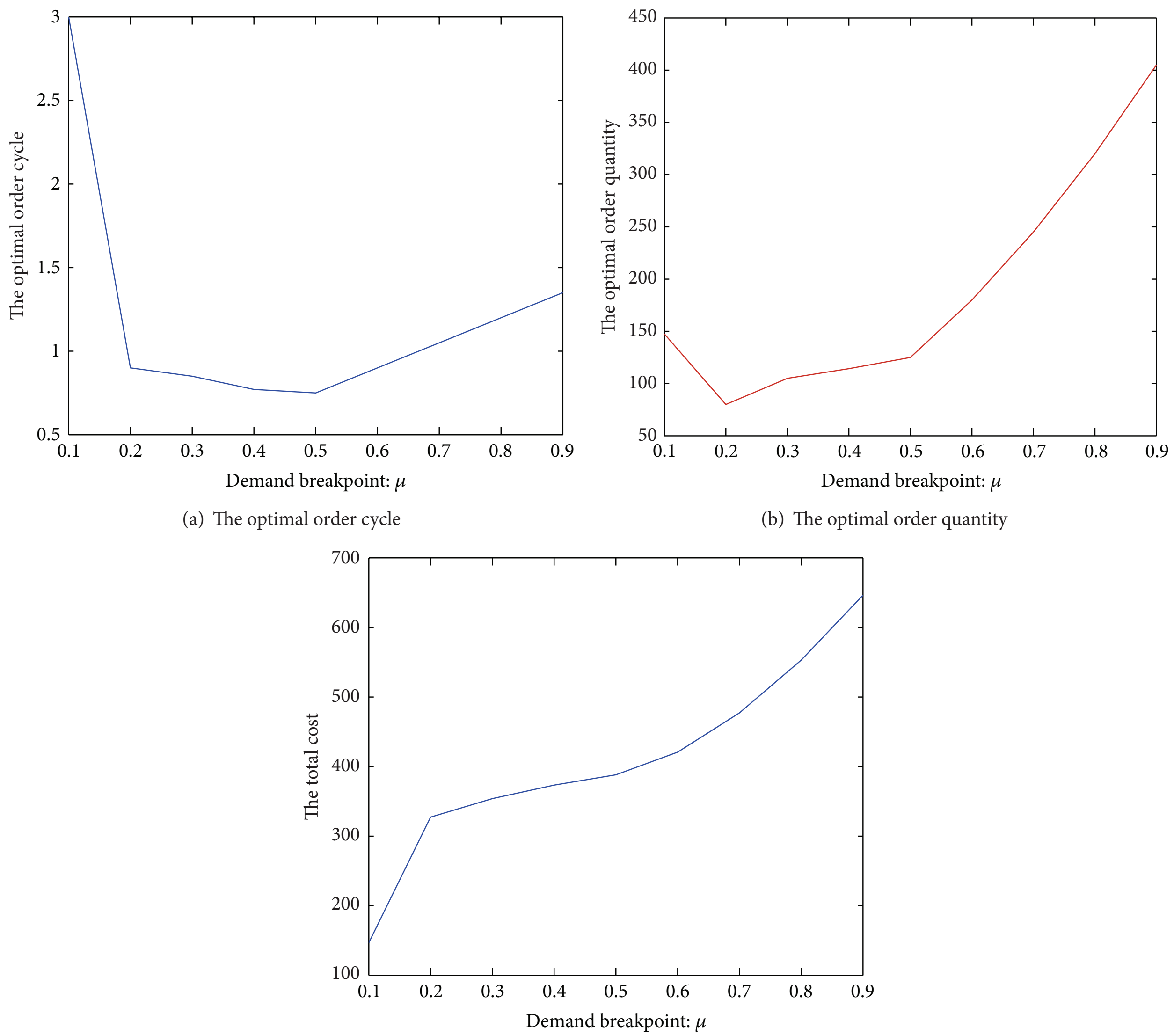
(c) The total cost for the retailer

FIGURE 8: The impact of the changes of the parameter M.

the order cost. If the demand dependent rate is higher, the demand dependent production rate has no effect on the optimal order cycle and the order quantity. For the limit case, when the production rate is infinite, as most papers presented, the production rate has no influence on the members' decisions.

(2) When the delayed payment time is less than the growth stage time of the new products' introduction, the retailer will shorten the replenishment cycle to take advantage of the trade credit more frequently for accumulating the interest; on the other hand, the short order cycle can make the retailer adjust its order decisions more quickly for meeting the changeable demand within the growth stage of the products. When the delayed payment time is larger than the breakpoint of the demand, within the maturity stage of the products, the delayed payment time has no effect on the optimal order cycle and the optimal order quantity. The total cost for the retailer is decreasing with the delayed payment time. Therefore, the retailer hopes the supplier can offer them the delayed payment time as long as possible.

(3) As the demand breakpoint is increasing, the optimal order cycle and the optimal order quantity are decreasing first and then increasing. The total cost for the retailer is increasing with the demand breakpoint. Hence, for the products, if the growth stage of the product life cycle is longer, the cost of the retailer



(a) The optimal order cycle

(b) The optimal order quantity

(c) The total cost for the retailer

Figure 9: The impact of the changes of the parameter μ.

will be higher. Therefore, the retailer should make some measures, such as advertisement, to induce the products' demand moving to the maturity stage as soon as possible.

The research presented in this paper can be extended in several ways. For example, other demand functions can be further discussed considering the demand dependent production rate in the EOQ inventory model with the trade credit financing. Additionally, the models can be generalized to consider the shortage or the partial backlogging. Furthermore, the influence of the poor quality products on the EOQ model can be discussed to obtain some management insights.

Appendices

A. Proof of Theorem 1

The first derivative for a minimum of TRC_{13} is

$$\frac{\partial TRC_{13}}{\partial T} = -\frac{1}{T^2} f_{13}(T) + \frac{1}{T} f'_{13}(T) \quad \text{(A.1)}$$

with

$$f_{13}(T) = A + \left(c_h + c_p I_c\right) \times \left[(r-1)\frac{1}{6} D_0 \mu^3 + \frac{1}{2} D_0 \mu^2 t_1 + \frac{1}{2} D_0 \mu T^2 - D_0 \mu t_1 T\right]$$

$$-\frac{1}{6}c_pI_c(r-1)D_0M^3-\frac{sI_e}{6}D_0M^3,$$
$$f'_{13}(T)=\left(c_h+c_pI_c\right)\times\left[-\frac{1}{2r}(r-1)D_0\mu^2+D_0\mu T\left(\frac{r-1}{r}\right)\right];$$
$$f''_{13}(T)=\left(c_pI_c+c_h\right)D_0\mu\frac{r-1}{r}. \quad (A.2)$$

Based on (A.1), we have

$$\frac{\partial TRC_{13}}{\partial T}=\frac{1}{T^2}\left\{-A-\left(c_h+c_pI_c\right)\times\left[(r-1)\frac{1}{6}D_0\mu^3-\frac{1}{8r}D_0(r-1)^2\mu^3-\frac{r-1}{2r}D_0\mu T^2\right]+\frac{1}{6}c_pI_c(r-1)D_0M^3+\frac{sI_e}{6}D_0M^3\right\}. \quad (A.3)$$

The second derivatives for TRC_{13} with respect to T is

$$\frac{\partial^2 TRC_{13}}{\partial T^2}=\frac{2}{T^3}f(T)-\frac{2}{T^2}f'(T)+\frac{1}{T}f''(T)=\frac{-2}{T^3}\left\{-A-\left(c_h+c_pI_c\right)\times\left[(r-1)\frac{1}{6}D_0\mu^3-\frac{1}{8r}D_0(r-1)^2\mu^3\right]+\frac{1}{6}c_pI_c(r-1)D_0M^3+\frac{sI_e}{6}D_0M^3\right\}. \quad (A.4)$$

We can know that $\lim_{T\to\infty}(\partial TRC_{13}/\partial T)>0$.
Let us set

$$\alpha=-A-\left(c_h+c_pI_c\right)D_0\mu^3(r-1)\left(\frac{1}{24}+\frac{1}{8r}\right)+\frac{1}{6}D_0M^3\left(c_pI_c(r-1)+sI_e\right). \quad (A.5)$$

If $\alpha<0$, then $\lim_{T\to 0}(\partial TRC_{13}/\partial T)<0$. Therefore, the intermediate value theorem yields that there exists $\widetilde{T}_{13}$ as the root of $\partial TRC_{13}/\partial T=0$ for $T=(0,\infty)$. And when $\alpha<0$, we have $\partial^2 TRC_{13}/\partial T^2>0$. This $\widetilde{T}_{13}$not only exits but also is unique. If $\widetilde{T}_{13}$ is feasible, that is $(1/2)(r+1)\mu\le\widetilde{T}_{13}$ for (22) and (23), the optimal replenishment cycle is $T_{13}=\widetilde{T}_{13}$. Otherwise, if $\widetilde{T}_{13}$ is infeasible, the optimal replenishment cycle is $T_{13}=(1/2)(r+1)\mu$. The optimal value of the order level is $Q_{13}=\int_0^{\mu}D_0t\,dt+\int_{\mu}^{T_{13}}D_0\mu\,dt$.

If $\alpha\ge 0$, then $\partial TRC_{13}/\partial T>0$. TRC_{13} is an increasing function with T for $T=(0,\infty)$. For $T\in[(1/2)(r+1)\mu,\infty)$, the optimal value T_{13} is obtained when $T=(1/2)(r+1)\mu$. The optimal value of the order level is $Q_{13}=\int_0^{\mu}D_0t\,dt+\int_{\mu}^{T_{13}}D_0\mu\,dt$.

Based on the analysis above, it is easy to obtain Theorem 1.

B. Proof of Theorem 2

Calculate the first derivatives of the TRC_{23} with respect to T:

$$\frac{\partial TRC_{23}}{\partial T}=-\frac{1}{T^2}f(T)+\frac{1}{T}f'(T); \quad (B.1)$$

with

$$f_{23}(T)=A+c_h\left[(r-1)\frac{1}{6}D_0\mu^3+\frac{1}{2}D_0\mu^2t_1+\frac{1}{2}D_0\mu T^2-D_0\mu t_1T\right]+c_pI_c\left(-\frac{1}{3}D_0t_1^3+\frac{1}{3}D_0M^3-\frac{1}{2}D_0\mu^2M-D_0\mu TM+D_0\mu^2M+D_0\mu T^2-\frac{1}{2}D_0\mu^2T\right)-sI_e\left(-\frac{1}{2}MD_0\mu^2+\frac{1}{6}D_0\mu^3+\frac{1}{2}D_0\mu M^2\right),$$
$$f'_{23}(T)=c_h\left[-\frac{1}{2r}(r-1)D_0\mu^2+D_0\mu T\left(\frac{r-1}{r}\right)\right]+c_pI_c\left(-2D_0t_1\frac{1}{r^2}+2D_0\mu\right);$$
$$f''_{23}(T)=c_hD_0\mu\frac{r-1}{r}-2c_pI_c\frac{1}{r^3}D_0. \quad (B.2)$$

Therefore, we have

$$\frac{\partial TRC_{23}}{\partial T}=\frac{1}{T^2}\left\{-A-c_h\times\left[(r-1)\frac{1}{6}D_0\mu^3-\frac{1}{8r}D_0(r-1)^2\mu^3-\frac{r-1}{2r}D_0\mu T^2\right]\right.$$

$$+ sI_e\left(-\frac{1}{2}MD_0\mu^2 + \frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu M^2\right)$$
$$-c_pI_c\left[\frac{1}{3}D_0M^3 - \frac{1}{2}D_0\mu^2M + D_0\mu^2M\right]\Bigg\}$$
$$+ c_pI_cD_0\mu. \tag{B.3}$$

The second derivatives for TRC_{23} with respect to T is

$$\frac{\partial^2 TRC_{23}}{\partial T^2} = -\frac{2}{T^3}\Bigg\{-A - c_h \times\left[(r-1)\frac{1}{6}D_0\mu^3 - \frac{1}{8r}D_0(r-1)^2\mu^3\right] + sI_e\left(-\frac{1}{2}MD_0\mu^2 + \frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu M^2\right) - c_pI_c\left[\frac{1}{3}D_0M^3 - \frac{1}{2}D_0\mu^2M + D_0\mu^2M\right]\Bigg\}. \tag{B.4}$$

We know that $\lim_{T\to\infty}(\partial TRC_{23}/\partial T) > 0$.
Let us set

$$\beta = -A - c_h \times\left[(r-1)\frac{1}{6}D_0\mu^3 - \frac{1}{8r}D_0(r-1)^2\mu^3\right] + sI_e\left(-\frac{1}{2}MD_0\mu^2 + \frac{1}{6}D_0\mu^3 + \frac{1}{2}D_0\mu M^2\right) - c_pI_c\left[\frac{1}{3}D_0M^3 - \frac{1}{2}D_0\mu^2M + D_0\mu^2M\right]. \tag{B.5}$$

If $\beta < 0$, then $\lim_{T\to 0}(\partial TRC_{23}/\partial T) < 0$. Therefore, the intermediate value theorem yields that there exists $\partial TRC_{23}/\partial T = 0$ for $T = (0,\infty)$. And when $\beta < 0$, we have $\partial^2 TRC_{23}/\partial T^2 > 0$. The optimal value of $\widetilde{T}_{23}$ can be obtained by solving the equation $\partial TRC_{23}/\partial T = 0$. $\widetilde{T}_{23}$ not only exits but also is unique. If $\widetilde{T}_{23}$ is feasible, that is, $Mr - (1/2)(r-1)\mu \le \widetilde{T}_{23}$ for (33) and (34), the optimal replenishment cycle $T_{23} = \widetilde{T}_{23}$. Otherwise, the optimal replenishment cycle is $T_{23} = Mr - (1/2)(r-1)\mu$. The optimal order quantity is $Q_{23} = \int_0^{\mu} D_0 t\,dt + \int_{\mu}^{T_{23}} D_0\mu\,dt$.

If $\beta \ge 0$, $\partial TRC_{23}/\partial T > 0$, so TRC_{23} is an increasing function with T for $T = (0,\infty)$. For $Mr - (1/2)(r-1)\mu \le T$, the optimal value of T_{23} can be obtained by $T = Mr - (1/2)(r-1)\mu$. The optimal order quantity is $Q_{23} = \int_0^{\mu} D_0 t\,dt + \int_{\mu}^{T_{23}} D_0\mu\,dt$.

Based on the analysis above, it is easy to obtain Theorem 2.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The paper is funded by The National Natural Science Foundation of China (71302115, 71172018), The Ministry of education of Humanities, and Social Sciences Project (13YJC630121).

References

[1] K.-J. Chung and J.-J. Liao, "The simplified solution algorithm for an integrated supplier-buyer-inventory model with two-part trade credit in a supply chain system," *European Journal of Operational Research*, vol. 213, no. 1, pp. 156–165, 2011.

[2] Y.-C. Tsao and G.-J. Sheen, "Dynamic pricing, promotion and replenishment policies for a deteriorating item under permissible delay in payments," *Computers & Operations Research*, vol. 35, no. 11, pp. 3562–3580, 2008.

[3] H. Soni and N. H. Shah, "Optimal ordering policy for stock-dependent demand under progressive payment scheme," *European Journal of Operational Research*, vol. 184, no. 1, pp. 91–100, 2008.

[4] J.-T. Teng, I.-P. Krommyda, K. Skouri, and K.-R. Lou, "A comprehensive extension of optimal ordering policy for stock-dependent demand under progressive payment scheme," *European Journal of Operational Research*, vol. 215, no. 1, pp. 97–104, 2011.

[5] B. Sarkar, "An EOQ model with delay in payments and time varying deterioration rate," *Mathematical & Computer Modelling*, vol. 55, no. 3-4, pp. 367–377, 2012.

[6] J. T. Teng, J. Min, and Q. H. Pan, "Economic order quantity model with trade credit financing for non-decreasing demand," *Omega*, vol. 40, no. 3, pp. 328–335, 2012.

[7] S. K. Goyal, "Economic order quantity under conditions of permissible delay in payments," *Journal of the Operational Research Society*, vol. 36, no. 4, pp. 335–338, 1985.

[8] P. Chu, K.-J. Chung, and S.-P. Lan, "Economic order quantity of deteriorating items under permissible delay in payments," *Computers & Operations Research*, vol. 25, no. 10, pp. 817–824, 1998.

[9] S. P. Aggarwal and C. K. Jaggi, "Ordering policies of deteriorating items under permissible delay in payments," *Journal of the Operational Research Society*, vol. 46, no. 5, pp. 658–662, 1995.

[10] A. M. M. Jamal, B. R. Sarker, and S. Wang, "An ordering policy for deteriorating items with allowable shortage and permissible delay in payment," *Journal of the Operational Research Society*, vol. 48, no. 8, pp. 826–833, 1997.

[11] H.-J. Chang and C.-Y. Dye, "An inventory model for deteriorating items with partial backlogging and permissible delay in payments," *International Journal of Systems Science*, vol. 32, no. 3, pp. 345–352, 2001.

[12] J.-T. Teng, "On the economic order quantity under conditions of permissible delay in payments," *Journal of the Operational Research Society*, vol. 53, no. 8, pp. 915–918, 2002.

[13] K.-J. Chung and Y.-F. Huang, "The optimal cycle time for EPQ inventory model under permissible delay in payments," *International Journal of Production Economics*, vol. 84, no. 3, pp. 307–318, 2003.

[14] R. Uthayakumar and P. Parvathi, "A two-stage supply chain with order cost reduction and credit period incentives for deteriorating items," *International Journal of Advanced Manufacturing Technology*, vol. 56, no. 5–8, pp. 799–807, 2011.

[15] Y.-F. Huang, "Optimal retailer's ordering policies in the EOQ model under trade credit financing," *Journal of the Operational Research Society*, vol. 54, no. 9, pp. 1011–1015, 2003.

[16] Q. Hao, L. Liang, Y. Yu-gang, and D. Shao-fu, "EOQ model under three Levels of order-size-dependent delay in payments," *Journal of Systems & Management*, vol. 16, pp. 669–672, 2007.

[17] Y.-F. Huang, "Optimal retailer's replenishment decisions in the EPQ model under two levels of trade credit policy," *European Journal of Operational Research*, vol. 176, no. 3, pp. 1577–1591, 2007.

[18] J.-T. Teng and C.-T. Chang, "Optimal manufacturer's replenishment policies in the EPQ model under two levels of trade credit policy," *European Journal of Operational Research*, vol. 195, no. 2, pp. 358–363, 2009.

[19] R. Du, A. Banerjee, and S.-L. Kim, "Coordination of two-echelon supply chains using wholesale price discount and credit option," *International Journal of Production Economics*, vol. 143, no. 2, pp. 327–334, 2013.

[20] J. Wu, L.-Y. Ouyang, L. E. Cárdenas-Barrón, and S. K. Goyal, "Optimal credit period and lot size for deteriorating items with expiration dates under two-level trade credit financing," *European Journal of Operational Research*, vol. 237, no. 3, pp. 898–908, 2014.

[21] H. Hwang and S. W. Shinn, "Retailer's pricing and lot sizing policy for exponentially deteriorating products under the condition of permissible delay in payments," *Computers & Operations Research*, vol. 24, no. 6, pp. 539–547, 1997.

[22] S. Panda, S. Senapati, and M. Basu, "Optimal replenishment policy for perishable seasonal products in a season with ramp-type time dependent demand," *Computers & Industrial Engineering*, vol. 54, no. 2, pp. 301–314, 2008.

[23] K. Skouri, I. Konstantaras, S. Papachristos, and I. Ganas, "Inventory models with ramp type demand rate, partial backlogging and Weibull deterioration rate," *European Journal of Operational Research*, vol. 192, no. 1, pp. 79–92, 2009.

[24] S. K. Manna and K. S. Chaudhuri, "An EOQ model with ramp type demand rate, time dependent deterioration rate, unit production cost and shortages," *European Journal of Operational Research*, vol. 171, no. 2, pp. 557–566, 2006.

[25] B. Mandal, "An EOQ inventory model for Weibull distributed deteriorating items under ramp type demand and shortages," *Opsearch*, vol. 47, no. 2, pp. 158–165, 2010.

[26] S. R. Singh and C. Singh, "Supply chain model with stochastic lead time under imprecise partially backlogging and fuzzy ramp-type demand for expiring items," *International Journal of Operational Research*, vol. 8, no. 4, pp. 511–522, 2010.

[27] V. K. Mishra and L. S. Singh, "Inventory model for ramp type demand, time dependent deteriorating items with salvage value and shortages," *International Journal of Applied Mathematics & Statistics*, vol. 23, no. 11, pp. 84–91, 2011.

[28] S. Agrawal and S. Banerjee, "Two-warehouse inventory model with ramp-type demand and partially backlogged shortages," *International Journal of Systems Science*, vol. 42, no. 7, pp. 1115–1126, 2011.

[29] T. Roy and K. Chaudhuri, "A finite time horizon EOQ model with ramp-type demand rate under inflation and time-discounting," *International Journal of Operational Research*, vol. 11, no. 1, pp. 100–118, 2011.

[30] S. Agrawal, S. Banerjee, and S. Papachristos, "Inventory model with deteriorating items, ramp-type demand and partially backlogged shortages for a two warehouse system," *Applied Mathematical Modelling*, vol. 37, no. 20-21, pp. 8912–8929, 2013.

[31] M. A. Ahmed, T. A. Al-Khamis, and L. Benkherouf, "Inventory models with ramp type demand rate, partial backlogging and general deterioration rate," *Applied Mathematics & Computation*, vol. 219, no. 9, pp. 4288–4307, 2013.

[32] S. Saha, "Optimal order quantity of retailer with quadratic ramp type demand under supplier trade credit financing," *International Journal of Management Science & Engineering Management*, vol. 9, no. 2, pp. 103–188, 2014.

[33] S. S. Mishra and P. K. Singh, "Computational approach to an inventory model with ramp-type demand and linear deterioration," *International Journal of Operational Research*, vol. 15, no. 3, pp. 337–357, 2012.

[34] Y.-C. Huang, K.-H. Wang, and C.-T. Tung, "Optimal order policy for single period products under payment delay with ramp type demand rate," *Journal of Information & Optimization Sciences*, vol. 31, no. 6, pp. 1337–1360, 2010.

[35] G. Darzanou and K. Skouri, "An inventory system for deteriorating products with ramp-type demand rate under two-level trade credit financing," *Advances in Decision Sciences*, vol. 2011, Article ID 761961, 15 pages, 2011.

[36] S. K. Manna and C. Chiang, "Economic production quantity models for deteriorating items with ramp type demand," *International Journal of Operational Research*, vol. 7, no. 4, pp. 429–444, 2010.

[37] K. Skouri, I. Konstantaras, S. K. Manna, and K. S. Chaudhuri, "Inventory models with ramp type demand rate, time dependent deterioration rate, unit production cost and shortages," *Annals of Operations Research*, vol. 191, no. 1, pp. 73–95, 2011.

[38] H. N. Soni and K. A. Patel, "Optimal strategy for an integrated inventory system involving variable production and defective items under retailer partial trade credit policy," *Decision Support Systems*, vol. 54, no. 1, pp. 235–247, 2012.

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 287925, 7 pages
http://dx.doi.org/10.1155/2014/287925

Research Article

The Profit Distribution of Supply Chain under E-Commerce

Jiang-Hua Zhang,[1] Rui-Rui Yin,[1] Jin Zhang,[2] and Meng-Jiao Nie[1]

[1] *School of Management, Shandong University, Jinan, Shandong 250100, China*
[2] *School of Mathematical Sciences, Shandong Normal University, Jinan, Shandong 250014, China*

Correspondence should be addressed to Jiang-Hua Zhang; zhangjianghua28@163.com

Received 11 April 2014; Accepted 29 May 2014; Published 16 June 2014

Academic Editor: Xiaochen Sun

Copyright © 2014 Jiang-Hua Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the development of e-commerce, its influence on supply chain and supply chain management is becoming increasingly significant too. In this paper, the literature on the supply chain profit is reviewed first, and then a two-level and four-party supply chain which consists of a supplier, an e-commerce platform, third-party logistics, and demander is taken into consideration. The profit function of supply chain under e-commerce is formulated by taking the price of product and the maximum supply amount under certain investment as decision-making variables and taking the expected value of random variables of price as the setting sales quantity. Finally, the existence of maximum profit in the supply chain is proved in the model, and the coordination of supply chain under e-commerce environment can be achieved by setting coordination parameters when the relevant cost parameters of supply chain members satisfy certain conditions.

1. Introduction

Competition in the 21st century no longer takes place between one enterprise and another, but among supply chains. The supply chain connects an enterprise and its suppliers, distributors, and customers together organically to make it a function network which includes information flow, material, and cash flow. The development of e-commerce is becoming more and more mature along with the vigorous development and increasing perfection of network economy. In the meantime, the influence of e-commerce on supply chain and its management also gets increasingly significant. E-commerce can provide customers with direct sales, collecting all kinds of channels of information. Additionally, it can accelerate product to enter market and facilitate the transfer of effective funding, so that a series of opportunities will be created to increase the profit for the enterprise or supply chain. Furthermore, e-commerce has favorable impact on supply chain cost; for example, the cost of product management can decrease as the supply chain is shortened, the cost of inventory can be reduced with centralization, and the coordination of supply chain can be improved via information sharing [1].

In the e-commerce environment, the realization of information sharing and cooperation commerce makes the supply chain respond to market demand quickly and helps improve the operation efficiency of the whole supply chain. What is more, it also brings huge benefits to the supply chain. On the basis of the cooperation of individual firms, the supply chain could operate efficiently through close coordination in production, logistics, inventory, sales, and others so as to improve the competition and profitability of the overall supply chain [2]. However, the chain enterprises are independent economic entity, and each of them aims at maximizing their own profits. Consequently, reasonable profit distribution of supply chain and effective cooperation mechanism play a key role in fostering a long-term and stable cooperative relationship among chain enterprises.

The difference of the supply chain under e-commerce environment and in the traditional condition is embodied not only by the timely updating of product information and the real-time transmission of the need information, but also by the various forms of payment and the resulting complex cash flows. Unlike the traditional supply chain with one level of supplier and retailer, there are at least four party members in the supply chain under e-commerce environment: supplier or manufacturer, e-commerce platform, third-party logistics, and demander. Therefore, the profit distribution of supply chain under the e-commerce environment is different from that in the traditional environment; the former must consider the profit coordination of two-level and four-party members and the relationship among information flow, cash flow, and profit point as well. Based on the established supply chain model under e-commerce environment, the profit distribution of supply chain under e-commerce environment is analysed from two aspects: the existence of maximum profit in supply chain under the e-commerce environment and the role to coordinate supply chain members in this model.

2. Literature Review

According to the analysis of the previous literature, the fact that the study of supply chain is mainly done in the following several fields can be found.

(1) Take the newsboy problem as background to further discuss how to coordinate the supply chain by various contracts, such as buy-back contract, revenue sharing contract, quantity flexibility contract, sales commission contract, and finally maximizing the profit of the whole supply chain. For example, the pieces of literature [3–6] have studied the profit distribution of supply chain based on a variety of supply chain coordination contracts. Literature [3] pointed out three basic requirements for the reasonable profit allocation mechanism against the newsvendor problem, and literature [4] takes the newsvendor problem as background to coordinate the supply chain with the compensation strategy. Additionally, literature [5] and literature [6] have studied the profit distribution of supply chain by revenue sharing contract.

(2) Expand the study of two-echelon supply chain to three-echelon supply chain and the study of supply chain with single retailer and wholesaler to that with several retailers and wholesalers, such as pieces of literature [7–10]. Literature [7] proposed a model of profit sharing and transfer pricing for network companies. Literature [8] studied cooperative behaviors and profit allocation in the supply chain which consists of one supplier and several retailers under the decentralized control on the premise of replenishment allowed. Literature [9] investigated coordination in a three-echelon supply chain, examined the impact of subsupply chain coordination, and pointed out that both the supplier and the retailer would prefer to act alone rather than to coordinate with the manufacturer when subsupply chain coordination is suggested. Literature [10] considered the coordination mechanism with revenue sharing and developed a coordination mechanism for a supply chain made up of one manufacturer and n Cournot competing retailers when the production cost and demands are simultaneously disrupted.

(3) Use Game Theory to analyse profit distribution in supply chains. Literature [11] and Literature [12] discussed how the producer prices and how the retailer decides the order quantity under the Stackelberg Game structure and developed a cooperative game model to implement profit sharing between the manufacturer and the retailer utilizing Nash bargaining model, respectively. Literature [13] analysed the impact of surplus division in supply chains on investment incentives with a biform-game approach, and literature [14] analysed both simultaneous-move and leader-follower games to determine the Nash and Stackelberg equilibrium, respectively, and achieved the globally optimal solution that maximizes the system-wide expected profit.

(4) Take the e-commerce as a decision-making variable to study supply chain. Literature [4] and Literature [15] researched the supply chain by taking the e-commerce as a decision-making variable. Literature [4] studied whether the supplier and the customer are willing to complete the transaction through e-commerce in the perishable product sales which is based on random demand, and literature [15] identified the B2B e-commerce usage patterns in North American small- and medium-sized enterprises (SMEs) in their supply chains, the contextual factors that influence usage patterns, and the subsequent effects of these patterns on firm performance.

Additionally, literature [16] and literature [17] have studied the relationship among supply chain members based on profit distribution. Literature [16] put forward a profit distribution plan among supply chain partnerships according to the maximum profit of supply chain and their respective risk proportion by analyzing supplier-retailer relationship model. Literature [17] studied how an informal, long-term relationship between a manufacturer and a retailer performs in turbulent market environments characterized by uncertain demand.

Based on the analysis we can find that many scholars take the traditional two-echelon or three-echelon supply chain as background to discuss supply chain profit distribution with mathematical model and Game Theory. But it is still not very common to study the supply chain profit distribution under e-commerce environment, and most study is mainly done on a single level supply chain. This paper takes the two-level and four-party supply chain which consists of a supplier, an e-commerce platform, third-party logistics, and a demander into consideration first, and next we take the product price and the maximum supply amount under certain investment as decision-making variables, take the setting sales quality as an expected value of random variables of price to create a profit function, and then study the supply chain profit distribution model under e-commerce environment via the analysis of the profit function.

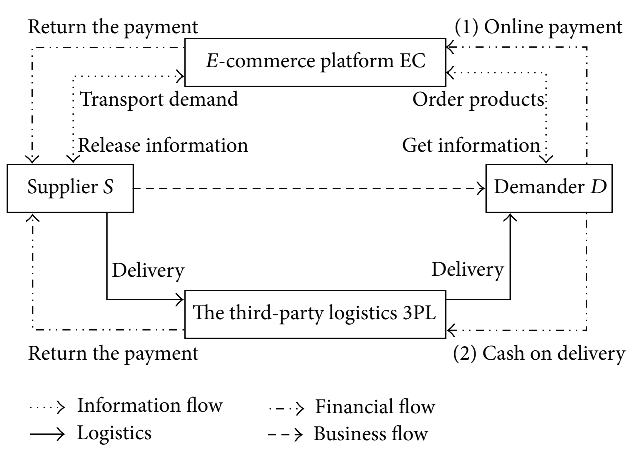


Figure 1: The supply chain model under e-commerce environment.

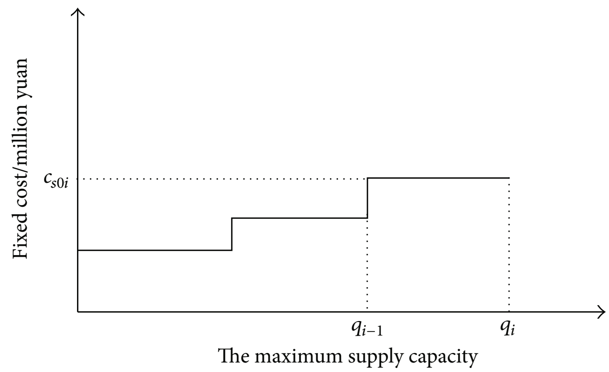


Figure 2: The relationship between fixed cost and the maximum supply capacity.

3. Model Establishment

3.1. Hypothesis

(1) Assume that we only take the supply chain which consists of a supplier S, an e-commerce platform EC, third-party logistics 3PL, and demander D into consideration (as shown in Figure 1).

(2) This model only considers a specified sales cycle and the product is single in the supply chain.

(3) The supplier releases product information in the e-commerce platform, the demander will obtain the required information through the platform, and the order will also be completed in the e-commerce platform. After the order information is received through e-commerce platform, the supplier will manufacture or purchase products and then entrust the third-party logistics to take and deliver goods and finally the logistics will be completed.

(4) There are two approaches to payment for demander: ① online payment, namely, it is paid through online bank and e-commerce platform; ② cash on delivery, namely, the payment is collected by third-party logistics. The demander paid the same in both kinds of payment methods.

(5) For e-commerce platform, it can be profitable only when a cash flow generated the information flow (such as releasing and obtaining product information, order, etc.) which is produced in the deal of both sides of supply and demand does not create profit point. There are two profit points in third-party logistics: the profit of cash flow from collection on delivery; the profit of shipping goods.

(6) There is no condition of asymmetric information in the whole supply chain, the supplier, e-commerce platform investor, and three-party logistics are all risk-free, and their target is to maximize their own profit.

3.2. Symbols and Meanings

(1) S sold the product to EC at unit price w_1 and sold the product to 3PL at unit price w_2; EC and 3PL both sold product to D at unit price $p(w_1 < p,\ w_2 < p)$. ϕ_1 of the EC income for goods as the revenue sharing to S; similarly, ϕ_2 of the 3PL income for goods as the revenue sharing to S.

(2) In a certain sales cycle, the transportation cost which D paid to 3PL shared per unit is t.

(3) The proportion of online payment for payment is α, while the proportion of cash on delivery is $1-\alpha$.

(4) In a certain sales cycle, the relationship between the fixed cost for the product purchase or production and the maximum supply capacity after receiving the need information as shown in Figure 2, the maximum supply capacity (namely, the maximum product number that you can purchase or produce under a certain amount of investment) taken as decision variable is a series of discrete values $\{q_1, q_2, \ldots, q_n\}$. The fixed cost corresponding to the maximum supply capacity q_i is c_{s0i}. The unit variable cost is c_{s1}.

(5) In a certain sales cycle, the construction and maintenance of the fixed cost share of the e-commerce platform by the platform business are c_{e0}, and the variable cost per unit of product that is produced by the transaction is c_{e1}.

(6) In a certain sales cycle, the fixed cost share for transporting goods by 3PL is c_{p0}, and the variable cost per unit of product produced by the transaction is c_{p1}.

(7) In a certain sales cycle, when the sales price of the product is p, the demand for product from user is a random variable with deterministic distribution $D(p)$, and its distribution function and density function, respectively, are $F(q_i \mid p)$ and $f(q_i \mid p)$ $(q_i \geq 0)$. $F(q_i \mid p)$ is differentiable about p, and $\partial F(q_i \mid p)/\partial p > 0$.

(8) $Q(q_i \mid p)$ is the expected value of the number of products in actual trading when the product price is p; then

$$
\begin{aligned}
Q(q_i \mid p) &= E\{\min(q_i, D(p))\} \\
&= \int_0^{\infty} \min(q_i, y) f(y \mid p)\, dy \\
&= \int_0^{q_i} y f(y \mid p)\, dy + q_i \int_{q_i}^{\infty} f(y \mid p)\, dy \\
&= q_i - \int_0^{q_i} F(y \mid p)\, dy.
\end{aligned}
\tag{1}
$$

3.3. Model Establishment. Based on the above assumptions and conditions, we can obtain the profit functions of supply chain.

The profit function of supplier S is

$$
\begin{aligned}
\Omega_s(q_i, p) &= w_1 \alpha Q(q_i \mid p) + \phi_1 p \alpha Q(q_i \mid p) \\
&\quad + w_2(1-\alpha) Q(q_i \mid p) + \phi_2 p (1-\alpha) Q(q_i \mid p) \\
&\quad - c_{s1} Q(q_i \mid p) - c_{s0i} \\
&= [\alpha w_1 + (1-\alpha) w_2] Q(q_i \mid p) \\
&\quad + [\alpha \phi_1 p + (1-\alpha)\phi_2 p] Q(q_i \mid p) \\
&\quad - c_{s1} Q(q_i \mid p) - c_{s0i}.
\end{aligned}
\tag{2}
$$

The profit function of e-commerce platform is

$$
\begin{aligned}
\Omega_e(q_i, p) &= (1-\phi_1) p \alpha Q(q_i \mid p) - w_1 \alpha Q(q_i \mid p) \\
&\quad - c_{e1} Q(q_i \mid p) - c_{e0}.
\end{aligned}
\tag{3}
$$

The profit function of 3PL is

$$
\begin{aligned}
\Omega_p(q_i, p) &= (1-\phi_2) p (1-\alpha) Q(q_i \mid p) \\
&\quad - w_2(1-\alpha) Q(q_i \mid p) - c_{p1} Q(q_i \mid p) \\
&\quad - c_{p0} + t Q(q_i \mid p).
\end{aligned}
\tag{4}
$$

The profit function of supply chain is

$$
\begin{aligned}
\Omega(q_i, p) &= p Q(q_i \mid p) - (c_{s1} + c_{e1} + c_{p1}) Q(q_i \mid p) \\
&\quad - c_{s0i} - c_{e0} - c_{p0} + t Q(q_i \mid p) \\
&= (p+t) Q(q_i \mid p) - (c_{s1} + c_{e1} + c_{p1}) Q(q_i \mid p) \\
&\quad - c_{s0i} - c_{e0} - c_{p0}.
\end{aligned}
\tag{5}
$$

4. The Research on Profit and Coordination of Supply Chain under E-Commerce Environment

4.1. The Existence of Maximum Profit of Supply Chain. Finding the first and the second derivative of $\Omega(q_i, p)$ of p for a given q_i according to (5), we can obtain

$$
\begin{aligned}
\frac{\partial \Omega(q_i, p)}{\partial p} &= Q(q_i \mid p) + p \frac{\partial Q(q_i \mid p)}{\partial p} \\
&\quad + (t - c_{s1} - c_{e1} - c_{p1}) \frac{\partial Q(q_i \mid p)}{\partial p} \\
&= Q(q_i \mid p) - (p + t - c_{s1} - c_{e1} - c_{p1}) \\
&\quad \times \int_0^{q_i} \frac{\partial F(y \mid p)}{\partial p} dy,
\end{aligned}
\tag{6}
$$

$$
\begin{aligned}
\frac{\partial^2 \Omega(q_i, p)}{\partial p^2} &= 2 \frac{\partial Q(q_i \mid p)}{\partial p} + (p + t - c_{s1} - c_{e1} - c_{p1}) \\
&\quad \times \frac{\partial^2 Q(q_i \mid p)}{\partial p^2} \\
&= -2 \int_0^{q_i} \frac{\partial F(y \mid p)}{\partial p} dy - (p + t - c_{s1} - c_{e1} - c_{p1}) \\
&\quad \times \int_0^{q_i} \frac{\partial^2 F(y \mid p)}{\partial p^2} dy.
\end{aligned}
\tag{7}
$$

Analyzing (5) and (6), then we get that the maximal value exists when certain conditions are satisfied.

Theorem 1. *When the conditions $\partial^2 F(y \mid p)/\partial p^2 > 0$ and $p \geq c_{s1} + c_{e1} + c_{p1} - t$ are met, there exists only one p_i^* for maximum supply quantity of each product, which makes the profit of supply chain maximum.*

Proof. For $\partial F(y \mid p)/\partial p > 0$, $\partial^2 F(y \mid p)/\partial p^2 > 0$ and $p \geq c_{s1} + c_{e1} + c_{p1} - t$; then

$$
\begin{aligned}
\frac{\partial^2 \Omega(q_i, p)}{\partial p^2} &= -2 \int_0^{q_i} \frac{\partial F(y \mid p)}{\partial p} dy \\
&\quad - (p + t - c_{s1} - c_{e1} - c_{p1}) \\
&\quad \times \int_0^{q_i} \frac{\partial^2 F(y \mid p)}{\partial p^2} dy < 0.
\end{aligned}
\tag{8}
$$

That is, $\Omega(q_i, p)$ is a concave function on p.

When $p_1 = c_{s1} + c_{e1} + c_{p1} - t$, for $F(y \mid p_1) \le 1$, then $\int_0^{q_i} F(y \mid p_1) dy \le q_i$, and

$$\begin{aligned}\left.\frac{\partial\Omega(q_i, p)}{\partial p}\right|_{p=p_1} &= Q(q_i \mid p_1) - (p_1 + t - c_{s1} - c_{e1} - c_{p1}) \\ &\quad \cdot \int_0^{q_i} \left.\frac{\partial F(y \mid p)}{\partial p}\right|_{p=p_1} dy \\ &= q_i - \int_0^{q_i} F(y \mid p_1) dy \ge 0.\end{aligned} \tag{9}$$

Letting $\Delta = \int_0^{q_i} (\partial F(y \mid p)/\partial p)|_{p=p_1} dy$, with the hypothesis $\partial F(y \mid p)/\partial p > 0$, we get $\Delta > 0$.

Since $\partial^2 F(y \mid p)/\partial p^2 > 0$, then $\partial F(y \mid p)/\partial p$ is a increasing function of p.

So, when $p_2 > p_1$, then $(\partial F(y \mid p)/\partial p)|_{p=p_2} > (\partial F(y \mid p)/\partial p)|_{p=p_1}$ holds and $\int_0^{q_i} (\partial F(y \mid p)/\partial p)|_{p=p_2} dy > \int_0^{q_i} (\partial F(y \mid p)/\partial p)|_{p=p_1} dy = \Delta$ holds.

If we let $p_2 > c_{s1} + c_{e1} + c_{p1} - t + (q_i/\Delta)$, then

$$\begin{aligned}\left.\frac{\partial\Omega(q_i, p)}{\partial p}\right|_{p=p_2} &= q_i - \int_0^{q_i} \partial F(y \mid p)\Big|_{p=p_2} dy \\ &\quad - (p_2 - c_{s1} - c_{e1} - c_{p1} + t) \\ &\quad \times \int_0^{q_i} \left.\frac{\partial F(y \mid p)}{\partial p}\right|_{p=p_2} dy \\ &\le q_i - (p_2 - c_{s1} - c_{e1} - c_{p1} + t) \\ &\quad \times \int_0^{q_i} \left.\frac{\partial F(y \mid p)}{\partial p}\right|_{p=p_2} dy \\ &\le q_i - \frac{q_i}{\Delta}\Delta = 0.\end{aligned} \tag{10}$$

From (9) and (10), we know that there exist p_1 and p_2 for maximum supply quantity q_i of each product, which makes $(\partial\Omega(q_i, p)/\partial p)|_{p=p_1} \ge 0$ and $(\partial\Omega(q_i, p)/\partial p)|_{p=p_2} \le 0$ true. According to the Zero Theorem, we obtain that there must be only one p_i^* for each q_i, which makes $(\partial\Omega(q_i, p)/\partial p)|_{p=p_i^*} = 0$ true, and then the profit of the whole supply chain achieves the maximum value.

In the supply chain model under e-commerce environment, we suppose that the profit of supply chain achieves the maximum value $\Omega(q_{\text{opt}}, p^*_{\text{opt}})$ when the supply quantity of S is q_{opt} and the corresponding final price of product of EC and 3PL is p^*_{opt}. □

4.2. The Coordination in Behavior of Supply Chain Members. When the maximum supply quantity of product is q_{opt}, the final selling price is p^*_{opt}; that is, when the profit of supply chain achieves the maximum, the profit functions in the supply chain and each member in this model are as follows.

The profit function of supplier S is

$$\begin{aligned}&\Omega_s(q_{\text{opt}}, p^*_{\text{opt}}) \\ &\quad = w_1 \alpha Q(q_{\text{opt}} \mid p^*_{\text{opt}}) + \phi_1 p^*_{\text{opt}} \alpha Q(q_{\text{opt}} \mid p^*_{\text{opt}}) \\ &\qquad + w_2 (1-\alpha) Q(q_{\text{opt}} \mid p^*_{\text{opt}}) \\ &\qquad + \phi_2 p^*_{\text{opt}} (1-\alpha) Q(q_{\text{opt}} \mid p^*_{\text{opt}}) \\ &\qquad - c_{s1} Q(q_{\text{opt}} \mid p^*_{\text{opt}}) - c_{s0\text{opt}} \\ &\quad = [\alpha w_1 + (1-\alpha) w_2 + \alpha\phi_1 p^*_{\text{opt}} + (1-\alpha)\phi_2 p^*_{\text{opt}} - c_{s1}] \\ &\qquad \times Q(q_{\text{opt}} \mid p^*_{\text{opt}}) - c_{s0\text{opt}}.\end{aligned} \tag{11}$$

The profit function of e-commerce platform is

$$\begin{aligned}&\Omega_e(q_{\text{opt}}, p^*_{\text{opt}}) \\ &\quad = (1-\phi_1) p^*_{\text{opt}} \alpha Q(q_{\text{opt}} \mid p^*_{\text{opt}}) \\ &\qquad - w_1 \alpha Q(q_{\text{opt}} \mid p^*_{\text{opt}}) - c_{e1} Q(q_{\text{opt}} \mid p^*_{\text{opt}}) - c_{e0} \\ &\quad = [\alpha(1-\phi_1) p^*_{\text{opt}} - \alpha w_1 - c_{e1}] Q(q_{\text{opt}} \mid p^*_{\text{opt}}) - c_{e0}.\end{aligned} \tag{12}$$

The profit function of 3PL is

$$\begin{aligned}&\Omega_p(q_{\text{opt}}, p^*_{\text{opt}}) \\ &\quad = (1-\phi_2) p^*_{\text{opt}} (1-\alpha) Q(q_{\text{opt}} \mid p^*_{\text{opt}}) \\ &\qquad - w_2 (1-\alpha) Q(q_{\text{opt}} \mid p^*_{\text{opt}}) - c_{p1} Q(q_{\text{opt}} \mid p^*_{\text{opt}}) \\ &\qquad - c_{p0} + t Q(q_{\text{opt}} \mid p^*_{\text{opt}}) \\ &\quad = [(1-\alpha)(1-\phi_2) p^*_{\text{opt}} - (1-\alpha) w_2 - c_{p1} + t] \\ &\qquad \times Q(q_{\text{opt}} \mid p^*_{\text{opt}}) - c_{p0}.\end{aligned} \tag{13}$$

The profit function of supply chain is

$$\begin{aligned}&\Omega(q_{\text{opt}}, p^*_{\text{opt}}) \\ &\quad = p^*_{\text{opt}} Q(q_{\text{opt}} \mid p^*_{\text{opt}}) - (c_{s1} + c_{e1} + c_{p1}) Q(q_{\text{opt}} \mid p^*_{\text{opt}}) \\ &\qquad + t Q(q_{\text{opt}} \mid p^*_{\text{opt}}) - c_{s0\text{opt}} - c_{e0} - c_{p0} \\ &\quad = (p^*_{\text{opt}} + t - c_{s1} - c_{e1} - c_{p1}) Q(q_{\text{opt}} \mid p^*_{\text{opt}}) \\ &\qquad - (c_{s0\text{opt}} + c_{e0} + c_{p0}).\end{aligned} \tag{14}$$

Analysing (11), (12), (13), and (14), then we get that $\Omega_s(q_{\text{opt}}, p^*_{\text{opt}})$, $\Omega_e(q_{\text{opt}}, p^*_{\text{opt}})$, $\Omega_p(q_{\text{opt}}, p^*_{\text{opt}})$, and $\Omega(q_{\text{opt}}, p^*_{\text{opt}})$ are all unary functions of $Q(q_{\text{opt}} \mid p^*_{\text{opt}})$. Therefore, if there exist constants λ_1, λ_2, $0 \le \lambda_1 \le 1$ and $0 \le \lambda_2 \le 1$ can meet the following conditions

$$\begin{gathered}
\alpha w_1 + (1-\alpha) w_2 + \alpha\phi_1 p^*_{\text{opt}} + (1-\alpha)\phi_2 p^*_{\text{opt}} - c_{s1} \\
= \lambda_1 \left(p^*_{\text{opt}} + t - c_{s1} - c_{e1} - c_{p1}\right), \\
\alpha(1-\phi_1) p^*_{\text{opt}} - \alpha w_1 - c_{e1} = \lambda_2 \left(p^*_{\text{opt}} + t - c_{s1} - c_{e1} - c_{p1}\right), \\
(1-\alpha)(1-\phi_2) p^*_{\text{opt}} - (1-\alpha) w_2 - c_{p1} + t \\
= (1-\lambda_1-\lambda_2)\left(p^*_{\text{opt}} + t - c_{s1} - c_{e1} - c_{p1}\right), \\
c_{s0\text{opt}} = \lambda_1 \left(c_{s0\text{opt}} + c_{e0} + c_{p0}\right), \\
c_{e0} = \lambda_2 \left(c_{s0\text{opt}} + c_{e0} + c_{p0}\right), \\
c_{p0} = (1-\lambda_1-\lambda_2)\left(c_{s0\text{opt}} + c_{e0} + c_{p0}\right).
\end{gathered} \tag{15}$$

Then we obtain

$$\begin{aligned}
\Omega_s\left(q_{\text{opt}}, p^*_{\text{opt}}\right) &= \lambda_1 \Omega\left(q_{\text{opt}}, p^*_{\text{opt}}\right), \\
\Omega_e\left(q_{\text{opt}}, p^*_{\text{opt}}\right) &= \lambda_2 \Omega\left(q_{\text{opt}}, p^*_{\text{opt}}\right), \\
\Omega_p\left(q_{\text{opt}}, p^*_{\text{opt}}\right) &= (1-\lambda_1-\lambda_2)\,\Omega\left(q_{\text{opt}}, p^*_{\text{opt}}\right).
\end{aligned} \tag{16}$$

Since $\lambda_1, \lambda_2 \in [0,1]$ and $\Omega_s(q_{\text{opt}}, p^*_{\text{opt}})$, $\Omega_e(q_{\text{opt}}, p^*_{\text{opt}})$, and $\Omega_p(q_{\text{opt}}, p^*_{\text{opt}})$ can be seen as unary functions with coefficient greater than zero of $\Omega(q_{\text{opt}}, p^*_{\text{opt}})$, supplier S, e-commerce platform EC, and the third-party logistics 3PL achieve the maximum profit when the entire supply chain gets maximum profit. And they get λ_1, λ_2, and $(1-\lambda_1-\lambda_2)$ part of the maximum profit of entire supply chain, respectively. Then we know that the maximum profit of supply chain is consistent with the maximum profit of each member, and then the model realizes the coordination effect on supply chain members.

With the set of unary equations (15), we have

$$\begin{aligned}
\lambda_1 &= \frac{c_{s0\text{opt}}}{c_{s0\text{opt}} + c_{e0} + c_{p0}}, \\
\lambda_2 &= \frac{c_{e0}}{c_{s0\text{opt}} + c_{e0} + c_{p0}}.
\end{aligned} \tag{17}$$

Let

$$\begin{aligned}
\phi_1 &= 1 - \frac{\lambda_2}{\alpha} = \frac{\alpha\left(c_{s0\text{opt}} + c_{e0} + c_{p0}\right) - c_{e0}}{\alpha\left(c_{s0\text{opt}} + c_{e0} + c_{p0}\right)}, \\
\phi_2 &= \frac{\lambda_1 + \lambda_2 - \alpha}{1-\alpha} = \frac{c_{s0\text{opt}} + c_{e0} - \alpha\left(c_{s0\text{opt}} + c_{e0} + c_{p0}\right)}{(1-\alpha)\left(c_{s0\text{opt}} + c_{e0} + c_{p0}\right)}.
\end{aligned} \tag{18}$$

Then

$$\begin{aligned}
w_1 &= \frac{c_{e0}\left(c_{s1} + c_{p1} - t\right) - c_{e1}\left(c_{s0\text{opt}} + c_{p0}\right)}{\alpha\left(c_{s0\text{opt}} + c_{e0} + c_{p0}\right)}, \\
w_2 &= \frac{t\left(c_{s0\text{opt}} + c_{e0}\right) + c_{p0}\left(c_{s1} + c_{e1}\right) - c_{p1}\left(c_{s0\text{opt}} + c_{e0}\right)}{(1-\alpha)\left(c_{s0\text{opt}} + c_{e0} + c_{p0}\right)}.
\end{aligned} \tag{19}$$

According to the requirement of revenue sharing ratio, ϕ_1 and ϕ_2 should meet constraints $0 < \phi_1 < 1$ and $0 < \phi_2 < 1$, respectively. And since S sells the product to EC and 3PL at prices of w_1, w_2, respectively, then $w_1 > 0$, $w_2 > 0$.

Analysing (18), we get that if $0 < \phi_1 < 1$ and $0 < \phi_2 < 1$, we must have $c_{e0}/(c_{s0\text{opt}} + c_{p0}) < \alpha/(1-\alpha)$ and $(c_{s0\text{opt}} + c_{e0})/c_{p0} > \alpha/(1-\alpha)$. Analysing (19), we know that if $w_1 > 0$ and $w_2 > 0$, then $c_{e0}/(c_{s0\text{opt}} + c_{p0}) > c_{e1}/(c_{s1} + c_{p1} - t)$ and $c_{p0}/(c_{s0\text{opt}} + c_{e0}) > (c_{p1} - t)/(c_{s1} + c_{e1})$. Consequently, we can realize the coordination of the supply chain under e-commerce environment according to (18) and (19) to set the coordinate parameters ϕ_1, ϕ_2, w_1, and w_2 when relevant cost parameters of each member meet the above conditions.

Above all, the constructed model of profit distribution of supply chain in e-commerce environment can maximize the profit of supply chain when $\partial^2 F(y \mid p)/\partial p^2 > 0$, supply quantity of product is q_{opt}, and the corresponding selling price is $p^*_{\text{opt}} \ge c_{s1} + c_{e1} + c_{p1} - t$. Furthermore, coordinate parameters of supply chain ϕ_1, ϕ_2, w_1, and w_2 can be obtained from (18) and (19), when relevant cost parameters of each member satisfy the conditions $c_{e0}/(c_{s0\text{opt}} + c_{p0}) < \alpha/(1-\alpha)$, $(c_{s0\text{opt}} + c_{e0})/c_{p0} > \alpha/(1-\alpha)$, $c_{e0}/(c_{s0\text{opt}} + c_{p0}) > c_{e1}/(c_{s1} + c_{p1} - t)$, and $c_{p0}/(c_{s0\text{opt}} + c_{e0}) > (c_{p1} - t)/(c_{s1} + c_{e1})$. Then the entire supply chain achieves the maximum profit, and supply chain members S, EC, and 3PL earn the maximum profit at the same time. S, EC, and 3PL can get $\alpha\phi_1 + (1-\alpha)\phi_2$, $\alpha(1-\phi_1)$, and $(1-\alpha)(1-\phi_2)$ part of the profit of supply chain, respectively. Moreover, the shared ratio of profit only lies in the cost parameters of each member; that is, it has nothing to do with the demand.

5. Conclusion

In this paper, literature on profit distribution of supply chain is analysed first, and based on the analysis, we study a two-level and four-party supply chain which is composed of a supplier, an e-commerce platform, the third-party logistics, and demander; then a mathematical model of supply chain's profit is formulated. The profit distribution of supply chain under e-commerce is investigated from two aspects: the existence of maximum profit of supply chain and the model to coordinate the behaviors of members through the analysis of the profit functions of supply chain and each member.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (NSFC) (Grant no. 71201093), the Promotive Research Fund for Excellent Young and Middle-aged Scientists of Shandong Province (Grant no. BS2012SF012), and Independent Innovation Foundation of Shandong University (Grant no. IFYT14011).

References

[1] S. Chopra and P. Meindl, *Supply Chain Management: Strategy, Planning and Operation*, Social Sciences Academic Press, 2009.

[2] J. Hou and L. Zhao, "Research on two-part revenue-sharing contract under demand disruption," *Journal of Southeast University*, vol. 23, pp. 57–63, 2007.

[3] G. P. Cachon, "Supply chain coordination with contracts," in *Handbooks in Operations Research and Management Science: Supply Chain Management*, S. Graves and T. de Kok, Eds., North-Holland, Amsterdam, The Netherlands, 2002.

[4] C. X. Wang and M. Benaroch, "Supply chain coordination in buyer centric B2B electronic markets," *International Journal of Production Economics*, vol. 92, no. 2, pp. 113–124, 2004.

[5] G. P. Cachon and M. A. Lariviere, "Supply chain coordination with revenue-sharing contracts: strengths and limitations," *Management Science*, vol. 51, no. 1, pp. 30–44, 2005.

[6] M. Leng and M. Parlar, "Game-theoretic analyses of decentralized assembly supply chains: non-cooperative equilibria vs. coordination with cost-sharing contracts," *European Journal of Operational Research*, vol. 204, no. 1, pp. 96–104, 2010.

[7] S. Y. Lakhal, "An operational profit sharing and transfer pricing model for network-manufacturing companies," *European Journal of Operational Research*, vol. 175, no. 1, pp. 543–565, 2006.

[8] L. A. Guardiola, A. Meca, and J. Timmer, "Cooperation and profit allocation in distribution chains," *Decision Support Systems*, vol. 44, no. 1, pp. 17–27, 2007.

[9] R. W. Seifert, R. I. Zequeira, and S. Liao, "A three-echelon supply chain with price-only contracts and sub-supply chain coordination," *International Journal of Production Economics*, vol. 138, no. 2, pp. 345–353, 2012.

[10] I. Sila and D. Dobni, "Patterns of B2B e-commerce usage in SMEs," *Industrial Management and Data Systems*, vol. 112, no. 8, pp. 1255–1271, 2012.

[11] Y.-W. Zhou, J. Min, and S. K. Goyal, "Supply-chain coordination under an inventory-level-dependent demand rate," *International Journal of Production Economics*, vol. 113, no. 2, pp. 518–527, 2008.

[12] S. Li, Z. Zhu, and L. Huang, "Supply chain coordination and decision making under consignment contract with revenue sharing," *International Journal of Production Economics*, vol. 120, no. 1, pp. 88–99, 2009.

[13] E. Cao, C. Wan, and M. Lai, "Coordination of a supply chain with one manufacturer and multiple competing retailers under simultaneous demand and cost disruptions," *International Journal of Production Economics*, vol. 141, no. 1, pp. 425–433, 2013.

[14] J. Sun and L. Debo, "Sustaining long-term supply chain partnerships using price-only contracts," *European Journal of Operational Research*, vol. 233, no. 3, pp. 557–565, 2014.

[15] E. Feess and J.-H. Thun, "Surplus division and investment incentives in supply chains: a biform-game analysis," *European Journal of Operational Research*, vol. 234, no. 3, pp. 763–773, 2014.

[16] S. S. Chauhan and J.-M. Proth, "Analysis of a supply chain partnership with revenue sharing," *International Journal of Production Economics*, vol. 97, no. 1, pp. 44–51, 2005.

[17] W.-G. Zhang, J. Fu, H. Li, and W. Xu, "Coordination of supply chain with a revenue-sharing contract under demand disruptions when retailers compete," *International Journal of Production Economics*, vol. 138, no. 1, pp. 68–75, 2012.

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 789754, 14 pages
http://dx.doi.org/10.1155/2014/789754

Research Article

Routing Optimization of Intelligent Vehicle in Automated Warehouse

Yan-cong Zhou,[1] Yong-feng Dong,[2] Hong-mei Xia,[2] and Jun-hua Gu[2]

[1] *School of Information Engineering, Tianjin University of Commerce, Tianjin 300134, China*
[2] *School of Computer Science and Software, Hebei University of Technology, Tianjin 300401, China*

Correspondence should be addressed to Yan-cong Zhou; zycong78@126.com

Received 11 April 2014; Accepted 12 May 2014; Published 16 June 2014

Academic Editor: Xiang Li

Copyright © 2014 Yan-cong Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Routing optimization is a key technology in the intelligent warehouse logistics. In order to get an optimal route for warehouse intelligent vehicle, routing optimization in complex global dynamic environment is studied. A new evolutionary ant colony algorithm based on RFID and knowledge-refinement is proposed. The new algorithm gets environmental information timely through the RFID technology and updates the environment map at the same time. It adopts elite ant kept, fallback, and pheromones limitation adjustment strategy. The current optimal route in population space is optimized based on experiential knowledge. The experimental results show that the new algorithm has higher convergence speed and can jump out the U-type or V-type obstacle traps easily. It can also find the global optimal route or approximate optimal one with higher probability in the complex dynamic environment. The new algorithm is proved feasible and effective by simulation results.

1. Introduction

With the development of factory automation and logistics automation, traditional production and material handling are more automated and intelligent, and a lot of large-scale storage centers come into being. However, due to the complexity of goods classification, frequent turnover, large number of houses, complex road conditions, and so on, finding target shelves location in the large-scale storage center is a very difficult task. When working people always need the help of memories or carrying signs for tendency guidelines, the result is inefficient and error-prone, and human, material, and financial resources are wasted on a certain degree. Intelligent vehicle is a cargo automatically handling car in intelligent transport systems that can meet the requirements of warehouse and flexible manufacturing systems, and it is one of the most crucial parts of the whole logistics automation and production automation.

Generally intelligent vehicle also called automated guided vehicle is a mobile robot. Intelligent vehicle is an integrated system which involves computer systems, sensors, automatic control, mechanical, communications, and other technologies. It has been widely applied in many areas such as industry, agriculture, military, and other fields and is one of the hot research issues in the field of robotic applications. Intelligent vehicle system is controlled by computer, with the characters of autonomous navigation, automatically path planning and executing tasks, independently avoiding obstacle, and has the advantages of high degree of automation, easy scheduling and management, safety and reliability, and so forth. Its goal is to achieve the automatically handling of goods. The key technology of intelligent vehicle is similar to the mobile robot and the difficult and critical points include navigation, routing optimization, task scheduling, coordination of multi-AGV control, and information fusion technology. Routing optimization is the focus of this study.

The routing optimization of intelligent vehicle is searching one optimal or approximate optimal route with a specific performance (such as shortest distance, less time, etc.) from starting point to target point in theenvironment with obstacles. The route searching is the prerequisite for intelligent vehicle performing various complex tasks, so the research for effective route planning in a complex environment is necessary and significant in the field of intelligent warehouse. Depending on the degree that environmental information is known, route searching can be divided into two categories: (1) global route searching with environmental information known; (2) local route searching with environmental information unknown or partially unknown. Global route searching is to find the optimal or approximate optimal route to meet certain performance from the starting point to the target point according to the priori model. The key problem is the environment model building and the route searching strategy. Local route searching means in the unknown or partially unknown environment, the intelligent vehicle according to the sensor's information, including the location, size, and shape information of the obstacle, gives a satisfied route with no collision.

For path planning problem of mobile robot, many scholars have had extensive research and also gotten some achievement. The traditional robot path planning methods such as artificial potential [1], visibility graph [2], graph searching, and grid decoupling method each have the advantages and disadvantages. With the development of intelligent algorithms, many scholars introduce them into the field of robot path planning. The intelligent algorithms improve the performance of the robot path planning to some extent, but they also have their own defects. For example, genetic algorithm is easy to fall into local optimum and has premature, slower convergence and other issues. With the increasing of obstacles or encountering complex terrain especially, the complexity of the intelligent algorithms will greatly increase, and we even cannot find the optimal solution [3].

Tuncer and Yildirim [4] in order to solve the problem of genetic algorithm of premature and slow convergence proposed a dynamic path planning of mobile robot based on with improved genetic algorithm. Yang and Fu [5] in order to improve the searching efficiency of genetic algorithm proposed a new mobile robot path planning algorithm combined with grid method and chaos genetic algorithm. There are some problems in the neural network system, such as networks large-scale, common performance, and easy to make a robot into an infinite loop. Glasius et al. [6] proposed a neural network model based on Hopfield network with dynamically avoiding obstacles. The model can avoid local minimum point, but it is difficult to adapt to the dynamic and high speed environment. Ant colony algorithm is easy to fall into local optimal solution and the U-shaped or V-shaped trap. Cai et al. [7] proposed a new method combined with the ant colony algorithm and fuzzy control technology. Although this method can solve the robot path planning problem, it is more complex and difficult. In order to improve execution speed and search efficiency, Liu [8] proposed an improved algorithm based on ant colony algorithm and genetic algorithm. Liu and Cheng [9] presented vision detection colony algorithm with elitist strategy to improve the efficiency. Zhou and Hua [10] presented an improved ant colony algorithm using simulated annealing algorithm to improve the pheromone evaporation coefficient.

About the routing optimization in logistics and warehouse, scholars also do a lot of work, and some do the research based on intelligent algorithms.

Sun [11] researched the path planning of automated guided vehicle system. He built the map model with graph theory in the process of AGV path planning and searched the shortest path by Dijkstra algorithm. Wang and Feng [12] researched the picking path plan for carousel based on ant colony optimization algorithm. They gave a mathematical model for hierarchical leveled carousel system with single picking station of automatic stereoscopic warehouse and proposed an improved ant colony optimization (ACO) algorithm. Zeng and Zong [13] researched routing optimization of AS/RS based on simulated annealing genetic algorithm. The simulated annealing algorithm and genetic algorithm are combined to solve the AS/RS. Pang and Lu [14] researched the path picking optimization of automated warehouses based on the ant colony generic algorithm. They gave an initial population by the ant colony algorithm and then solved the model with the genetic algorithms. Liu et al. [15] researched dynamic material handling route planning based on real-time operation conditions. They considered the complex and changeable material demand environment of the production system and set up a dynamic material transporting routing optimization model, which considered several demanders and multifarious convey angles. References [16, 17] introduced the RFID technology into the path planning. Chen et al. [16] researched the indoor path planning for seeing robot eyes based on RFID. Guo [17] researched the intelligent navigation and scheduling of vehicles in warehouses. For the requirements of saving time and energy, he used bridging RFID module to take charge of the navigation function. Optimal route was generated by a combinative strategy of topological-index and A^* algorithm.

This paper will introduce RFID technology into large-scale warehousing center, and the use of RFID will make the environment map updated automatically and timely, so the routing optimization problem from starting point to target shelf will be solved more effectively. In the process of route searching, for the ant colony algorithm easy to fall into local optimum, and hard to jump out the U-type or V-type trap, the paper proposes an evolutionary ant colony algorithm based on the prior knowledge, to effectively find the optimal route.

The rest of this paper is organized as follows. Our work environment model is formulated in Section 2. In Section 3, we provide the goal of routing optimization problem and present the definitions of ant colony algorithm's parameters. The detailed steps and flow charts are given in Section 3. In Section 4 the simulation results and analysis of four different algorithms are provided. Finally, we conclude our paper in Section 5.

2. Modeling the Working Environment for Intelligent Vehicle

Modeling the working environment is the first step for the path optimization. A reasonable description of the working environment can decrease the searching steps in the process of searching the optimal path and reduce the complexity on time and space. In this section, we will model the working environment, which is mainly the presentations of obstacles, destination, and action space. Shelves and goods in large-scale storage center may not only be the action destination, but also obstacles, which make the working environment complicated. So to facilitate the following analysis, we use grid method to model the working environment. And some assumptions are given as follows.

(1) Assume the storage to be rectangular, and there are some static obstacles, such as shelves and goods. The reason for the static assumption is that, first, all shelves and immobile goods are static; second, although the goods can be moved, corresponding to the moving vehicle, the moving goods are also static. Therefore, all obstacles are considered to be static. However, the working environment can be changed dynamically after each transport.

(2) The intelligent vehicle is viewed as a particle without size.

(3) Expand each obstacle into a circumscribed rectangle. If the expansion cannot fill a complete grid, then it will be considered as a grid.

(4) Assume the starting point to be fixed, and the destination may be different as the changing of the goods. However, the environment will be updated after goods is moved to the destination, so the destination of the new path optimization problem is still static.

In the following, we will model the working space. Let C denote the whole storage space including all action space and finite obstacles. Define a rectangular coordinate system where the left upper corner is the origin of the coordinate and the upper border of C is x-axis, and the left border of C is y-axis. Let $x_{\max}$ denote the maximum value of the horizontal axis of the point in C, and $y_{\max}$ denote the maximum value of the longitudinal axis of the point in C, so the work space is

$$A = \{(x, y) \mid x \in [0, x_{\max}], y \in [0, y_{\max}]\}. \tag{1}$$

Let R_c denote the maximum length of each action step and divide the domain C into the grids with equality segmentation, and let both lengths of each step in x-axis and y-axis be R_c. The column number of the grid domain is denoted by N_x, and the row number is denoted by N_y. In this paper, we consider the rectangular C to be a square. Let $N_x = N_y$. So, the continuous domain A can give a discrete space A_d defined by

$$A_d = \{(i, j)\, i, j = 0, 1, 2, \ldots, N_x\}. \tag{2}$$

We set the sequence number for each grid, so A_d also can be denoted by

$$A_d = \left\{(x_k, y_k) \mid x_k = \operatorname{fix}\left(\frac{k}{N_x}\right) + 1,\ y_k = k \bmod N_x + 1,\right.$$
$$\left. k = 0, 1, \ldots, N_x^2 - 1\right\}. \tag{3}$$

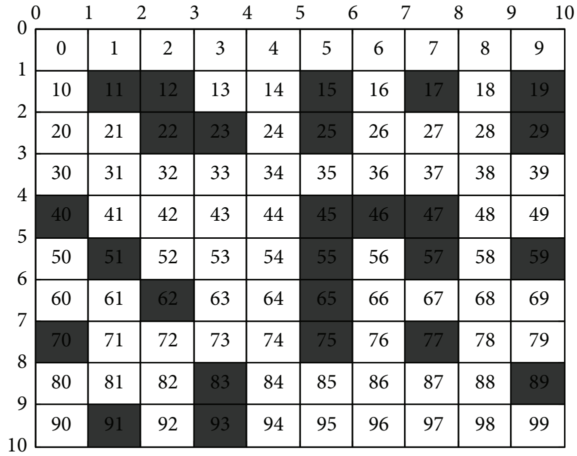


Figure 1: Environment map.

In Figure 1, we give a grid description of a working space with 100 grids. In the grid domain, a black grid denotes an obstacle and a white grid denotes an action space.

3. Modeling for Routing Optimization of Intelligence Vehicle

3.1. Objective Function. We consider a working environment in Figure 1, where the start point is (x_s, y_s), the destination is (x_d, y_d). The routing optimization is to find an optimal route from all feasible paths. The feasible path is a path from (x_s, y_s) to (x_d, y_d) and can avoid possible obstacles. Generally we set the start point to be (x_1, y_1).

We define the path length by Euclidean distance, so the length of an edge with the points $p_i(x_i, y_i)$ and $p_{i+1}(x_{i+1}, y_{i+1})$ is given as follows:

$$d = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}. \tag{4}$$

The path with L grids can be denoted as follows:

$$p_{i_1}(x_{i_1}, y_{i_1}) \longrightarrow \cdots \longrightarrow p_{i_k}(x_{i_k}, y_{i_k}) \longrightarrow \cdots \longrightarrow p_{i_L}(x_{i_L}, y_{i_L}),$$
$$i_1 = 1, \quad i_L = n. \tag{5}$$

Therefore, the total length $D(L)$ of the path is

$$D(L) = \sum_{k=1}^{L-1} \sqrt{(x_{i_{k+1}} - x_{i_k})^2 + (y_{i_{k+1}} - y_{i_k})^2}. \tag{6}$$

In practical operations, when making a turn the vehicle needs to judge the obstacle, calculate its size and space extent, and relocate the direction, which cost a lot of time and energy, so we should reduce the number of turns as possible as we can. Therefore, the routing optimization problem has two objectives: the least number of turns and the shortest path length. We consider a weighted sum objective function as follows:

$$F1 = \left(1 + \frac{1}{\sqrt{1+L}}\right) \times D \times w_d + \text{curve} \times w_c, \quad (7)$$

where L is the total number of all grids in a path, $D/\sqrt{1+L}$ is a correction item, and curve is the total number of turns. w_d and w_c are the weight, and $w_d + w_c = 1$.

To facilitate the observation and the show of experiment results, we use the following new objective function:

$$F = \frac{100}{\left(1 + \left(1/\sqrt{1+L}\right)\right) \times D \times w_d + \text{curve} \times w_c}. \quad (8)$$

In summary, the path optimization problem can be summarized as the following model:

$$\max\{G(L), L = 1, 2, \ldots, I\}$$

$$\text{s.t.} \begin{cases} G(L) = \max_P F \\ \text{s.t.} \begin{cases} \left(\theta x_{i_k} + (1-\theta) x_{i_{k+1}}, \theta y_{i_k} + (1-\theta) y_{i_{k+1}}\right) \notin A_O, \\ \qquad \text{for any } \theta \in [0,1], \ i_k, i_{k+1} \in S_a, \\ p = (p_1, p_2, \ldots, p_m), p_k = \left(x_{i_k}, y_{i_k}\right), \\ \left(x_{i_1}, y_{i_1}\right) = (x_1, y_1), \\ \left(x_{i_L}, y_{i_L}\right) = (x_d, y_d), \end{cases} \end{cases} \quad (9)$$

where I denotes the total number of feasible grids in the working environment; that is, $I = N_x^2 - s$. Here we set the upper-bound of L to be I, which is just the reason that the path including withdraw steps is not optimal.

3.2. Parameters of Ant Colony Algorithm. When looking for food, ants release special secretions called pheromones on their paths, which will evaporate with time. The later ants will select one path with the probability that is proportional to the intensity of pheromones on the path. When more ants pass through one path, there will be more pheromones released on the path, and then this path will be selected by ants with higher probability. Thus, a kind of positive feedback mechanism is formed by which ants can eventually find the optimal route. The parameters and strategies of basic model of ant colony algorithm are as follows.

Let m be the number of ant colony algorithm.

Definition 1. Tabu table is an array of two dimensions, which records the traversed nodes of each ant. tabu_k records currently traversed nodes of ant k in each generation. When ant k reaches the target point, the route of ant k is just given by the tabu table tabu_k.

Definition 2. $\tau_{ij}(t)$ is the pheromones factor. t represents time and (i, j) represents the path from node i to node j, $(i, j = 1, 2, \ldots, n)$. $\tau_{ij}(t)$ represents the retained pheromones of edge (i, j) at time t.

Definition 3. Heuristic factor η_{ij} denotes expectation degree that the ants move from node i to node j and usually is defined as follows:

$$\eta_{ij} = \frac{1}{d_{ij}}, \quad (10)$$

where d_{ij} is the distance between node i and node j.

Definition 4. Let P_{ij}^k be the transition probability with which ant transfers from node i to node j. The definition is as follows:

$$P_{ij}^k = \begin{cases} \dfrac{\left[\tau_{ij}(t)\right]^{\alpha}\left[\eta_{ij}(t)\right]^{\beta}}{\sum_{s \in \text{allowed}_k}\left[\tau_{ij}(t)\right]^{\alpha}\left[\eta_{ij}(t)\right]^{\beta}}, & \text{if } j \in \text{allowed}_k, \\ 0, & \text{else}, \end{cases} \quad (11)$$

where allowed_k (where $k = \{1, 2, \ldots, n\}$) represents the set of allowed next nodes which can be selected by ant k on the current environment. α and β denote the degree of importance of pheromones on the path and heuristic factor η_{ij}.

Definition 5. $\Delta\tau_{ij}(t)$ is the pheromone increment. It represents the pheromone increment on edge (i, j) after time Δt. The definition is as formula (12). Generally when initialized, $\Delta\tau_{ij}(0)$ is always set to zero:

$$\Delta\tau_{ij}(t) = \sum_{k=1}^{m} \Delta\tau_{ij}(t), \quad (12)$$

where $\Delta\tau_{ij}^k(t)$ represents the pheromones increment of ant k on edge (i, j) after time Δt. The definition is as follows:

$$\Delta\tau_{ij}^k(t) = \begin{cases} \dfrac{Q}{L_k}, & \text{ant } k \text{ pass edge } (i, j), \\ 0, & \text{others}, \end{cases} \quad (13)$$

where Q is the enhancing coefficient of pheromones, which affects the speed of convergence to a certain extent; L_k represents the distance of the route which is created by ant k in the current iteration.

Therefore, after the time Δt, the pheromones on each edges can be updated by the following:

$$\tau_{ij}(t + \Delta t) = (1 - \rho)\,\tau_{ij}(t) + \Delta\tau_{ij}(t), \quad (14)$$

where ρ is the evaporating coefficient of pheromones, $1 - \rho$ represents the retain factor. In order to prevent the unlimited accumulation of pheromones, set $\rho \in [0, 1]$.

For the first goal of shortest distance, the basic idea of ant colony algorithm is to place m ants at the starting point at the same time, and each ant selects one feasible node with a certain probability and meanwhile updates the local pheromones. The ants select the next available node with the same strategy until they reach the target point. Thus, the path passed by each ant is a feasible solution and then in accordance with their contribution to the problem they update global pheromones. If the conditions of termination are met, the current optimal solution is output, otherwise the next iteration continues.

About the second goal of reducing the turns of route, the modeling and related strategies will be described in detail in Section 4.

4. Routing Optimization for Intelligent Vehicle Based on Evolutionary Ant Colony Algorithm

4.1. Steps of Ant Colony Algorithm. The general steps for solving the routing optimization problem by ant colony algorithm are as follows.

Step 1 (*initialization*). The ants are placed at the starting point S, and S is added to the tabu list tabu_k. Let the initial pheromones of each side be a constant; $\tau_{ij}(0) = \tau_0/d$ (τ_0 is a constant, d is the distance to the next grid, $d = 1$ or $\sqrt{2}$). Here, we redefine the initial pheromones with the distance factor considered, so it is different from the traditional ant colony algorithm $\tau_{ij}(0) = \tau$. The new definition helps to improve the convergence. Set the current experiment iterations NG = 1; the maximum iterations are NGMAX.

Step 2 (*select the next available node*). In the algorithm we select the next available node j with roulette strategy. At any time t, transfer probability P_{ij}^{k} from node i to j is as shown in formula (11).

Step 3 (*pheromones update*). After time Δt, the pheromones are updated according to formula (14), and the pheromones evaporating coefficient is adaptive. The pheromones evaporating coefficient is very important when the environment map is complex. If ρ is too small, it is very easy to fall into local optimum solution. If ρ is too large, it will reduce the convergence speed of the algorithm. So in this paper the pheromones evaporating coefficient is dynamically adjusted according to the situation of path length. As formula (15), if the path length of the path set has distinct difference, it will slow convergence speed or, otherwise, accelerate convergence speed:

$$\rho(t) = \frac{D_{\text{ave}} - D_{\min}}{D_{\max} - D_{\min}}, \tag{15}$$

where $D_{\max}$ is the length of the longest path, D_{ave} represents the average length of all paths, and $D_{\min}$ is the length of the shortest path.

Here we also set a limitation for the pheromones evaporating coefficient. The limitation is just to prevent the algorithm falling into local optimum due to the coefficient being too big or too small. Here the maximum and minimum values are given.

Step 4. Set iteration NG = NG + 1. If NG > NGMAX, then go to Step 5; otherwise, adopt elite ant strategy. The ant with best fitness value in this iteration is chosen as the elite ant, which is automatically selected into the next iteration, and thus can increase the impact of the optimal route of the previous iteration and improve the convergence of the algorithm. Go to Step 2.

Step 5. Output the optimal route and the algorithm end.

In order to prevent the ants falling into a U-shaped or V-shaped trap, fallback strategy is adopted in the algorithm. When ants fall into the trap, if there is no good method to deal with the situation, the ants will be in "dead" state that the current feasible node set is empty, so the entire algorithm will be influenced. In the paper when ant k falls into U-shaped or V-shaped trap, we let it back to the previous node, and then the previous node is added to tabu table tabu_k. If now the feasible nodes set is still empty, do backing until the feasible set of ant is not empty.

The whole algorithm flow chart is as shown in Figure 2.

4.2. Strategy of Routing Optimization. As referred to in Section 3.1, routing optimization of intelligent vehicle has two goals: one is the shorter distance, and the second is the fewer number of turns. Inspired by the cultural algorithm, we propose the reducing turns strategy based on the prior knowledge.

The main idea of the cultural algorithm is that in the population space individuals have individual experience during the evolutionary process, and the individual experience will be passed to the belief space through the function Accept(). Individual experience received in belief space will be compared and optimized according to certain rules, thus forming groups experience, and then update the group experience with update() function according to the existing group and individual experience. In belief space, after the formation of the group experience the behavior of individuals in the population space will be modified by Influence() function, in order to enable individuals to achieve higher evolutionary efficiency. The basic framework of cultural algorithm [18] is as shown in Figure 3.

Considering the second goal, we should minimize the turns of the route. This goal can be achieved by the optimization operation based on the a priori knowledge on the route which is obtained during iterations of ant colony algorithm. The optimization operation is the experience update in belief space of cultural colony algorithm. In the belief space the two optimizing operations are based on the a priori knowledge. One operation is abandoning the roundabout and the other is reducing turns by parallelogram strategy.

(1) Strategy of Abandoning the Roundabout. Roundabout will emerge when ants are looking for food, and it will influence the pheromones of the path, which is not conducive for routing optimization, and thus will mislead other ants, so the roundabout makes the algorithm have poor convergence. In order to improve the convergence of the algorithm, the operation of abandoning roundabout must be applied to the current route. Set the grid number of one path to be Ln; calculate the allowed nodes set for the current node. If the next nodes (except the first next node of the current node)

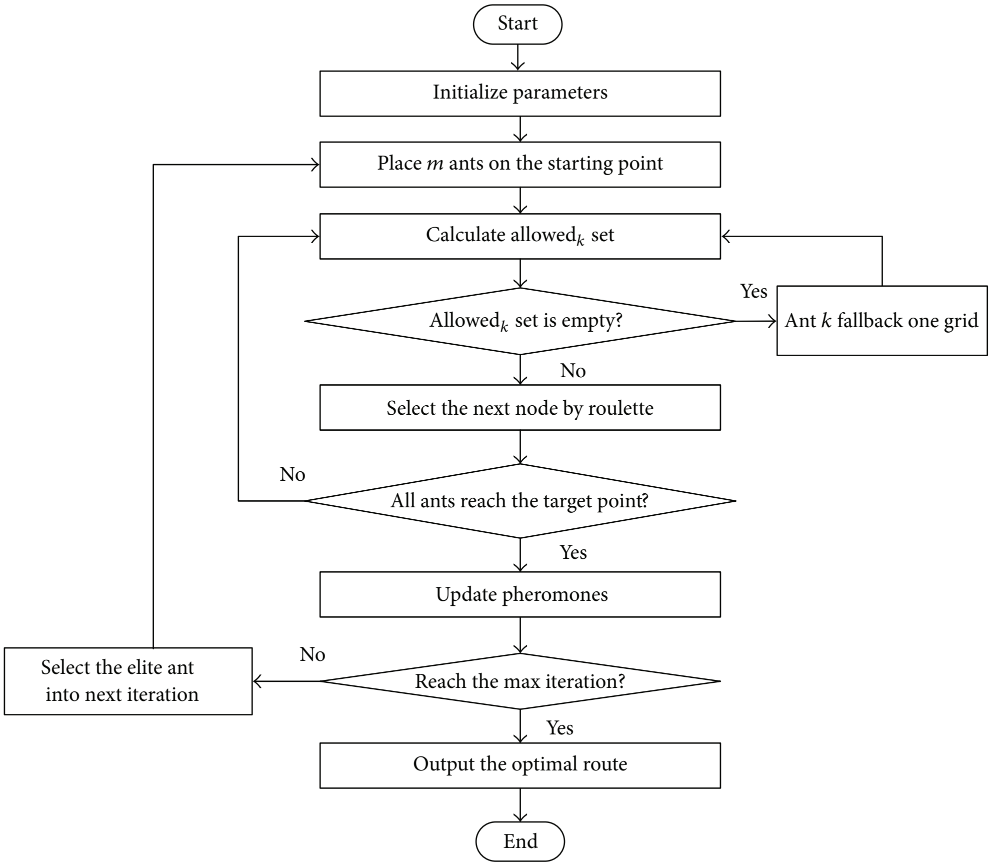


FIGURE 2: Flow chart of ant colony algorithm.

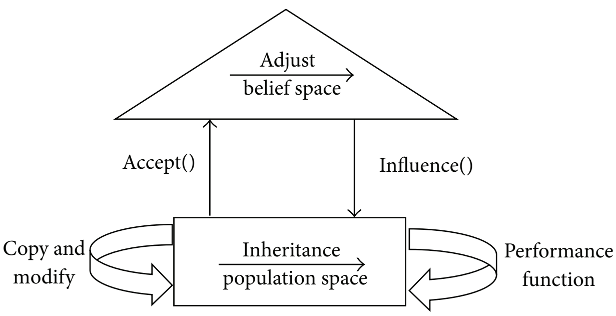


FIGURE 3: Basic framework of cultural algorithm.

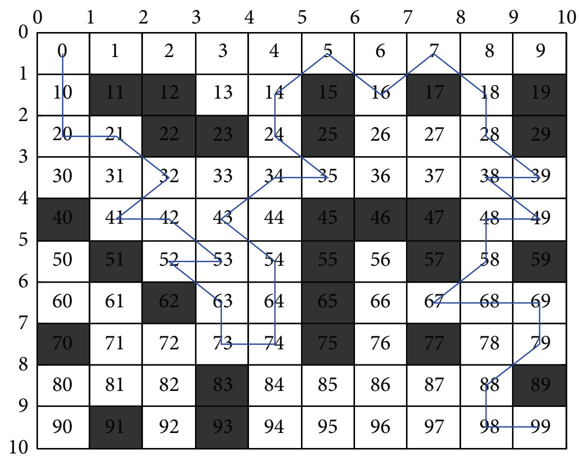


FIGURE 4: Route before abandoning roundabout.

after current node are in the allowed nodes set, you can delete the nodes between the current node and the next node directly from the circuitous path, and then the quality of the path will be improved. Figure 4 shows the original circuitous path. For node 32, its allowed node set is (31, 33, 41, 42, 43). Except the next node 41, we find node 43 in the next nodes set and also in the allowed nodes set, so the path between node 43 and node 32 can be deleted. The new route after the optimization is shown in Figure 5. From Figure 4 and Figure 5 we can see the quality of the route is significantly improved after the operation of abandoning roundabout and the number of turns is effectively reduced also. The flow chart of optimization operation for abandoning the roundabout is shown in Figure 6.

(2) Strategy of Reducing Turns by Parallelogram. The strategy is to reduce the number of vehicle turns and thus can reduce

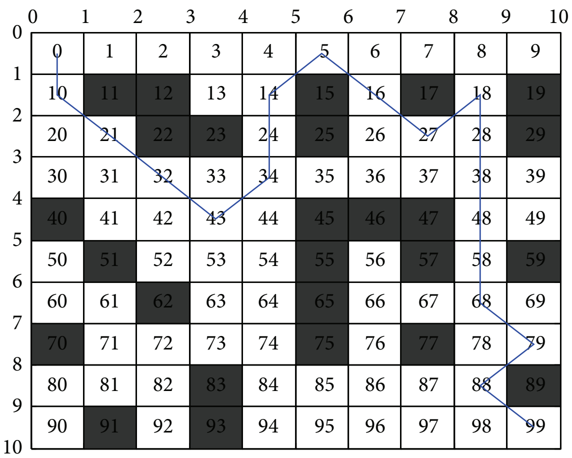


Figure 5: Route after operation of abandoning roundabout.

the vehicle's walking energy consumption and improve the fitness value of route. Select three consecutive nodes: the connection between the first node and the previous node on the route forming segment 1 and the connection between second and third node forming segment 2. When segment 1 and segment 2 are parallel, draw a parallelogram with the three nodes as the three vertices of the parallelogram, and then we will get two new segments which are not on the original route. If the grids passing by the new segments are all free grid, then the two original segments on the route will be deleted, the two new segments replace the original, and a new route then comes up. Obviously, the new route has less turns than the old one. The map and the original route are shown in Figure 7. We explain the strategy in detail taking nodes 10, 43, and 63, for example. Segment 1 between nodes 0 and 10 parallels segment 2 between nodes 43 and 63, so draw up a parallelogram with nodes 10, 43, and 63 as the three vertices of the parallelogram. Then, two new segments come up. One is segment between nodes 10 and 30, the other is the one between nodes 30 and 63. The grids passing by the two new segments are all free grid, so the two old segments are substituted by the two new ones. Then the new route comes up and is as shown in Figure 8. From Figure 7 and Figure 8 we can see after the optimization of reducing turns by parallelogram that the new route has less turns and higher fitness value than the old one. The flow chart of optimization operation for reducing turns by parallelogram is shown in Figure 9.

4.3. Steps of Evolutionary Ant Colony Algorithm. The integration of Sections 4.1 and 4.2 is the whole evolutionary ant colony algorithm. The algorithm takes the ant colony as the population space and the optimization for the route with operation of abandoning roundabout and reducing the turns by parallelogram based on the a priori knowledge as the group experience updating of belief space. The main steps of the evolutionary ant colony algorithm for routing optimization are as follows.

(1) In the population space the route set is generated by ant colony algorithm, and each route with individual experience is delivered to the belief space.

(2) In the belief space all the routes are optimized by strategies of abandoning roundabout and reducing turns by parallelogram based on the a priori knowledge and thus form the groups experience.

(3) The optimized routes are delivered back to the population space to update the pheromones. Repeat steps (1)–(3) until the end condition of the algorithm is satisfied.

The flow chart of evolutionary ant colony algorithm is as shown in Figure 10. Specific operation is as follows.

(1) Population Space. Detailed operation steps in population space are as follows.

(1) Build model for working environment of intelligent vehicle by grid method. Obtain information of the current environment through RFID technology and then build the current environment map. Determine the starting point and the destination point.

(2) Parameter initialization for ant colony algorithm. Set the initial value of iteration, initial time, initial pheromones, tabu list, and so forth.

(3) According to the ant colony algorithm mentioned in Section 4.1, select nodes and update pheromones, and generate corresponding route set for m ants.

(4) Deliver the route set to the belief space by Accept() function.

(5) Receive the updated route set from the belief space by Influence() function.

(6) Global pheromones are updated.

(7) Judging the termination conditions. If the condition is not satisfied, go to step (8). If satisfied, output the optimal route and the intelligent vehicle advances according to the route through the assistance of RFID. During the process of returning starting point from target point the environment map is updated through RFID technology timely and intelligent vehicle preparing for the next goods handling.

(8) The number of iterations $t = t + 1$; go to step (3).

(2) Belief Space. Detailed operation steps in belief space are as follows.

4.4. Evolutionary Genetic Algorithm. For the problem of routing optimization, coupled with the a priori knowledge is a very effective method. In order to verify this conclusion, apply the a priori knowledge into genetic algorithm. Also based on the framework of cultural algorithm, genetic populations are the population space, and in the belief space group experience is updated based on the a priori knowledge. The algorithm is called evolutionary genetic algorithm. The routing optimization for intelligent vehicle is mainly involved in the following

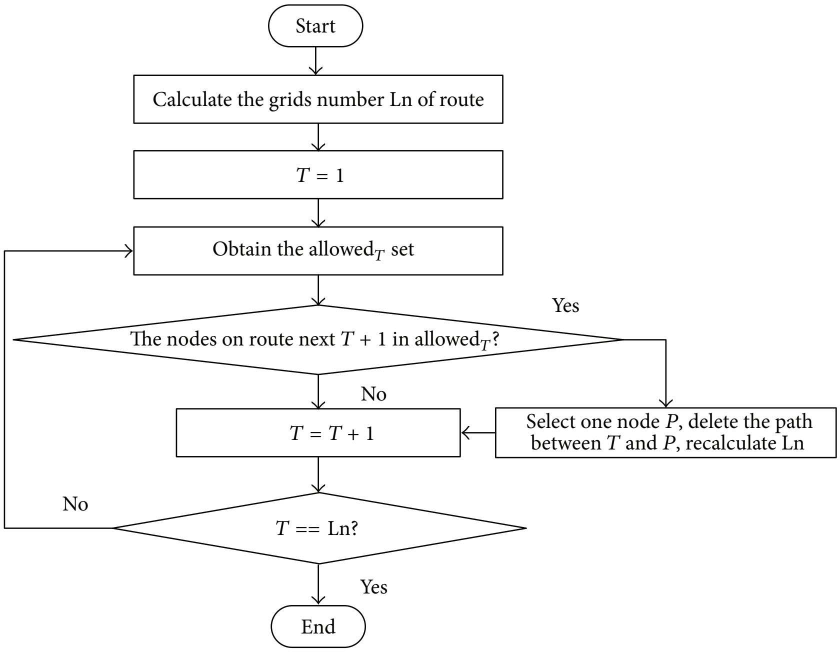


Figure 6: Flow chart of abandoning the roundabout.

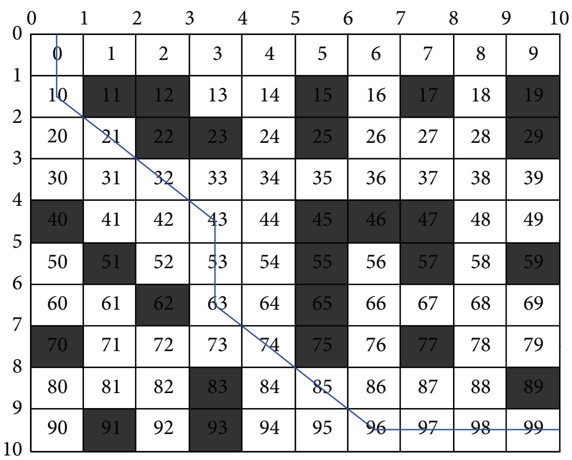


Figure 7: Original route before reducing turns.

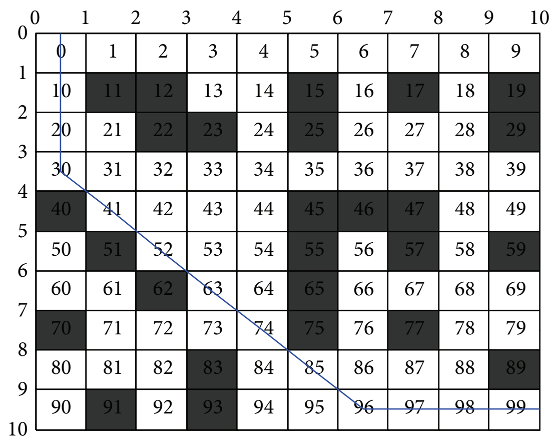


Figure 8: Route after reducing turns by parallelogram.

points: (1) initializing the population, giving the original route set in the feasible region, (2) giving appropriate fitness function combined with the actual working environment of the intelligent vehicle, (3) according to different situations of population adaptive genetic algorithm selecting appropriate crossover and mutation operator, and (4) keeping the diversity of the population in the belief space. Flow chart of evolutionary algorithm is shown in Figure 11. Specific steps of algorithm are as follows.

(1) Population Space

(1) Initializing the genetic algorithm, setting the iteration number $T = 1$.

(2) Modeling for working environment of intelligent vehicle by grid method, obtaining the feasible route set and set it as the initial population of genetic algorithm.

(3) According to the fitness function to calculate each individual's fitness value, do the selection operation for the population with roulette selection, adopt elitist kept strategy, and generate new populations.

(4) Do the single point crossover operation on the adjacent chromosomes to generate new individuals.

(5) Do mutation operation on part of individuals with mutation probability to produce new individuals.

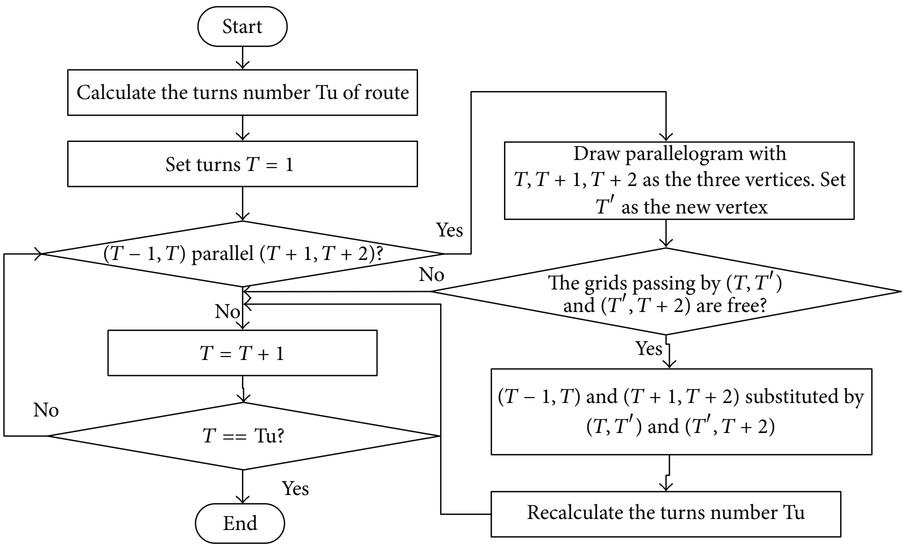


Figure 9: Flow chart of reducing turns by parallelogram.

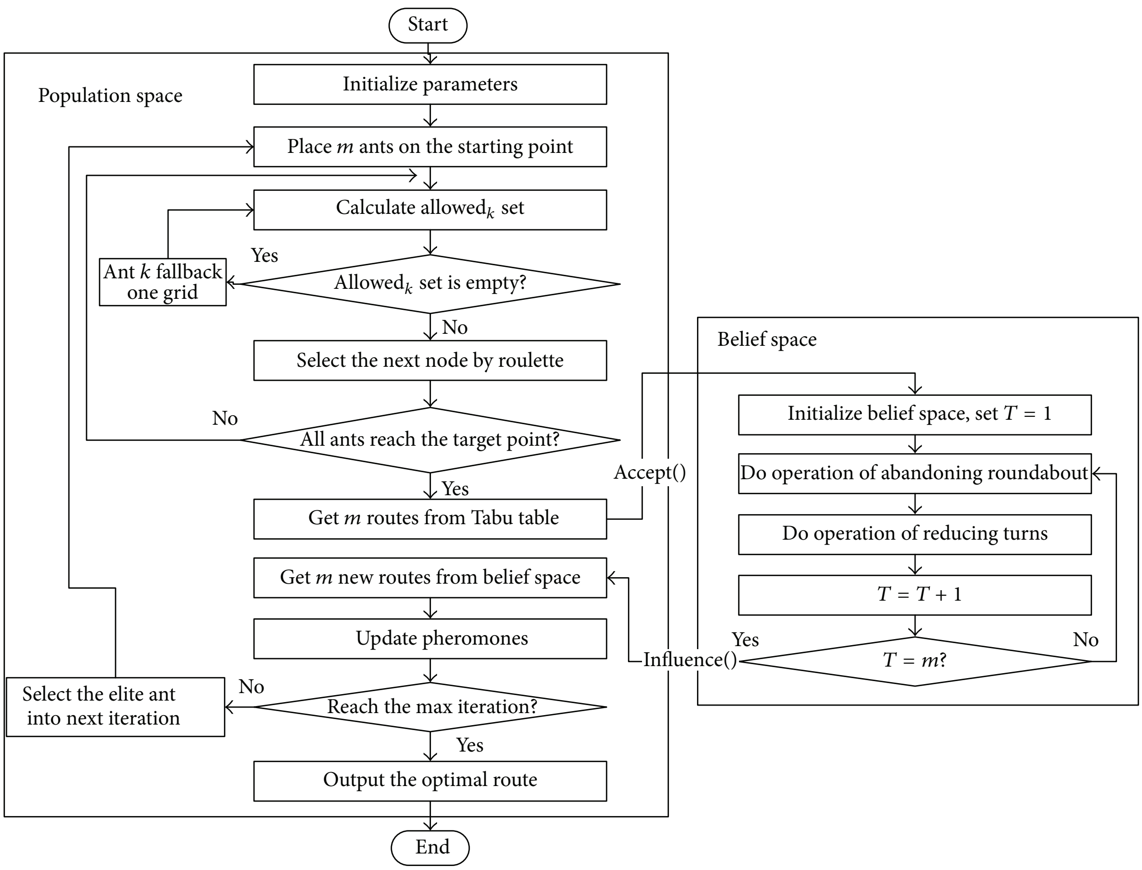


Figure 10: Flow chart of evolutionary ant colony algorithm.

(6) The new population is delivered to the belief space by Accept() function.

(7) Receive the new generation of population by Influence() function of belief space.

(8) Judging the conditions for termination $T < t\text{max}$, if the termination condition is satisfied, output the best individual.

(9) $T = T + 1$, go to step (3).

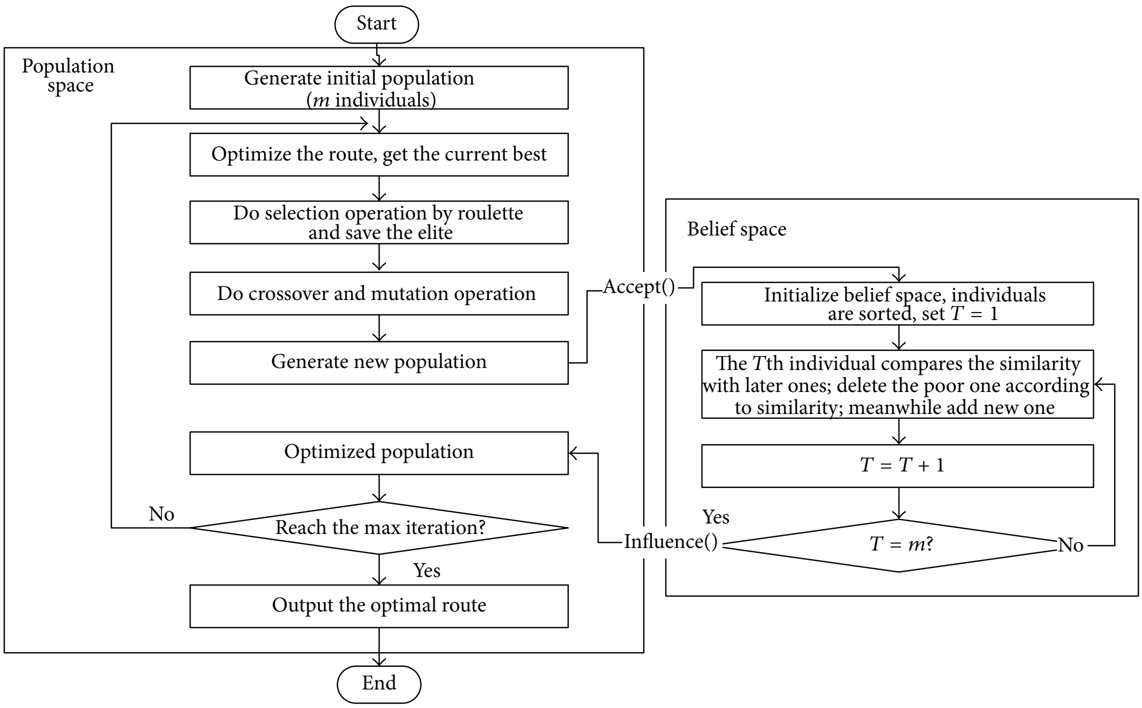


Figure 11: Flow chart of evolutionary genetic algorithm.

(2) Belief Space. The steps in belief space are as follows: the population after selection, crossover, and mutation of population space is delivered to belief space; the individuals with fitness value less than a certain threshold in the population do similarity comparison with each other; if the similarity is greater than a certain threshold, then the individuals with lower fitness value are deleted and randomly generate a new individual to join the populations. When comparison finishes the new populations return to population space.

5. Simulation

5.1. Simulation Environment and Parameter Settings. In order to study the efficiency of evolutionary ant colony algorithm for intelligent vehicle searching optimal route, a lot of simulation experiments are done. The hardware environment of simulation: Processor Core (TM) i3-2120, CUP 3.30 GHz, RAM 6 GB, 64-bit operating system, and hard disk 500 GB. Operating system is Windows 8, and Matlab 7.10.0 is the programming tool. Environment map can be changed as the actual environment change. Here we select three maps with $20 * 20$ grids, $50 * 50$ grids, and $100 * 100$ grids. All parameters are as shown in Table 1, where m represents the number of ants, α and β represent, respectively, the importance of the pheromones and heuristic factor, ρ represents the global pheromones evaporation coefficient, τ_0 represents the initial value of pheromones, and Q represents intensity factor of pheromones.

5.2. Simulation Results. The environment mapI is as shown in Figure 12. The start point is S, and the destination point is G. We make the comparison with four algorithms: the evolutionary ant colony algorithm (EAC), the ant colony genetic algorithm (AC-GA) in [8], the improved ant colony algorithm (SA-AC) in [10], and the evolutionary genetic algorithm (EGA). There are two algorithms (EAC and EGA) based on experiential knowledge. The four algorithms all run 50 times; randomly select one result; the optimal fitness convergence results are as shown in Figure 13. We can see

Table 1: Parameters of evolutionary ant colony algorithm.

m	α	β	ρ	τ_0	Q
20	0.6	0.8	0.3	10	100

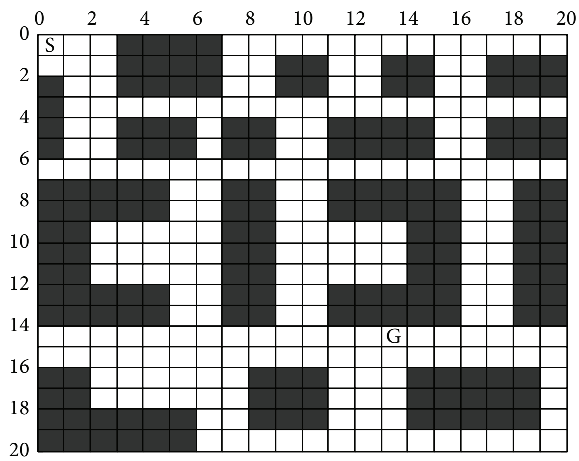


Figure 12: Environment map I.

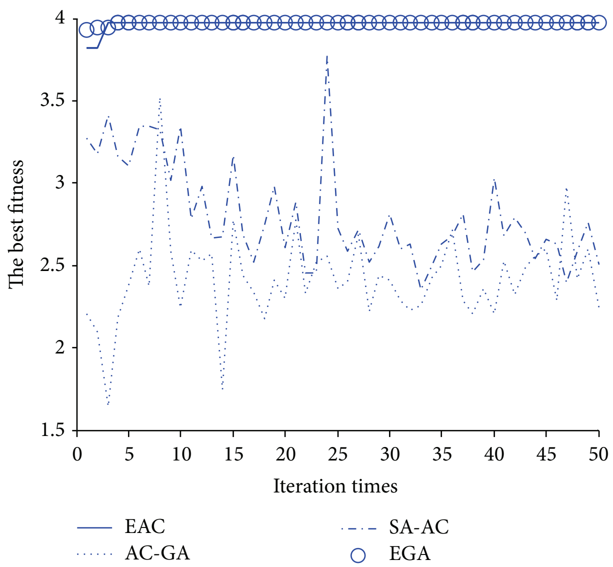


Figure 13: Convergence curves about the optimal fitness on map I.

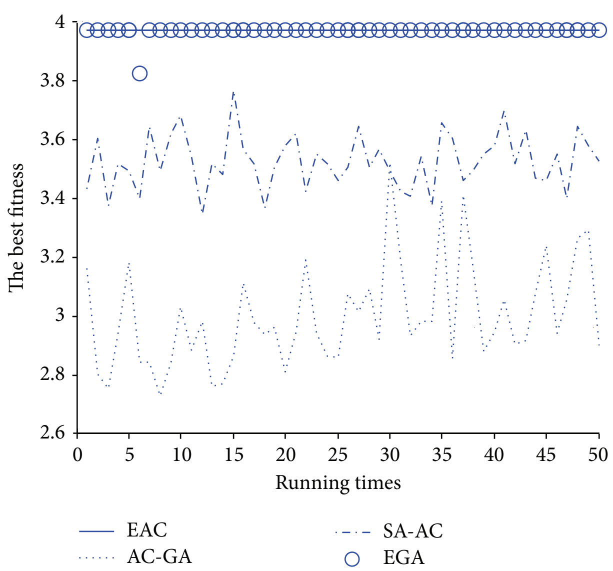


Figure 14: Convergence curves about the optimal fitness running 50 times on map I.

from Figure 13 EAC algorithm has the highest efficiency and the best convergence. Although EGA also can find the optimal route, its convergence speed is slower than EAC algorithm. The AC-GA algorithm and the improved SA-AC algorithm can hardly find the optimal route and have poorer convergence than the evolutionary algorithms.

The comparison of four algorithms on fitness value of 50 times is as shown in Figure 14. We can see that the EAC algorithm can find the optimal route every time, and the EGA can find the optimal route with high probability. The AC-GA algorithm and the SA-AC algorithm cannot find the optimal route yet, and their fluctuations are relatively large, but they all can find the approximate optimal route. We also

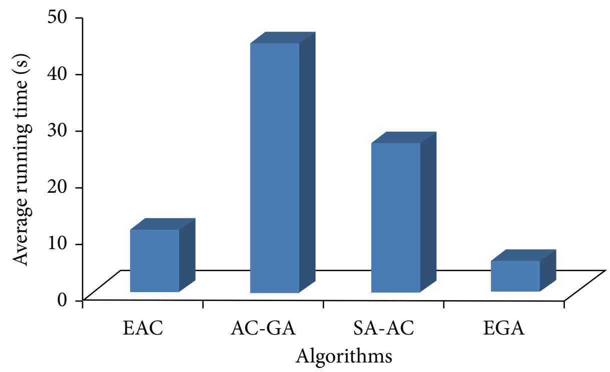


Figure 15: Comparison of average running time on map I.

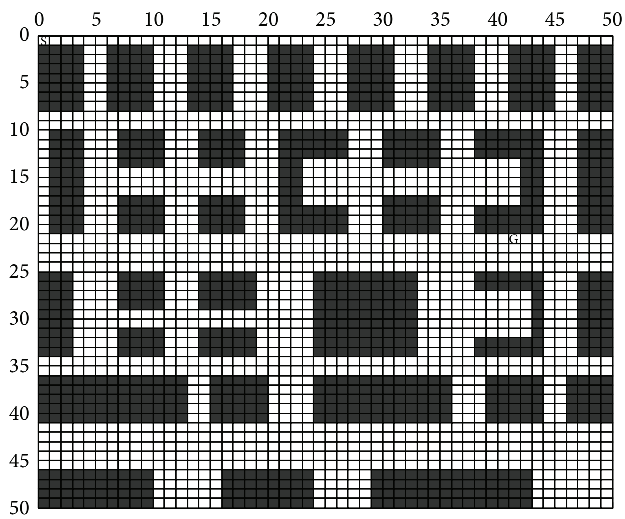


Figure 16: The working environment map II.

compare the average running time of four algorithms, and the result is as shown in Figure 15. From Figure 15 we can see that EAC and EGA run faster than AC-GA and SA-AC. EGA is the fastest and AC-GA is the lowest. In summary, the EAC algorithm is the most effective one.

We do a lot of simulation experiments with 30 different environment maps and have the same conclusion. Randomly select another two maps as shown in Figure 16 is $50 * 50$ map, and Figure 17 is $100 * 100$ map. The experimental comparison results of map II are as shown in Figures 18, 19, and 20, and the comparison results of map III are as shown in Figures 21, 22, and 23. The experimental results prove that the EAC algorithm has the highest optimal searching efficiency and best convergence. The EGA has the shortest running time. The more complex the working environment maps are, the more superior the EAC is. In summary, EAC algorithm is a feasible and an effective algorithm for routing optimization of automated vehicle.

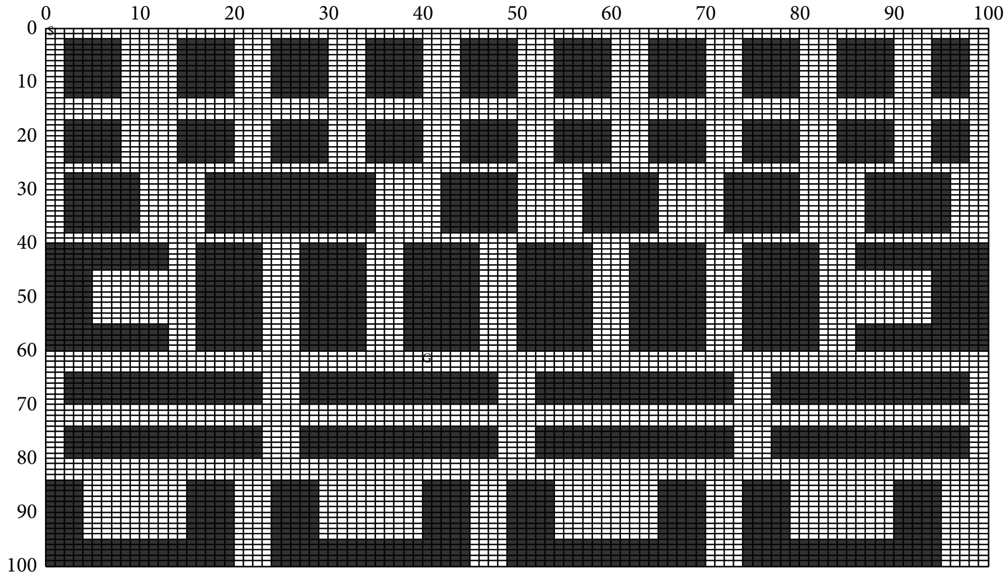


Figure 17: The working environment map III.

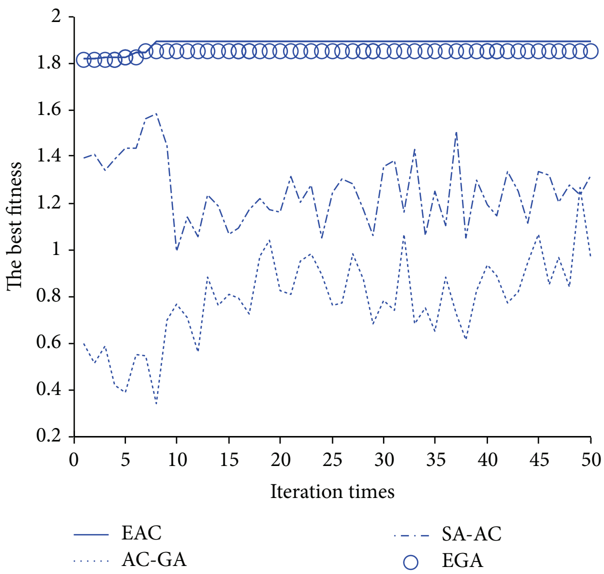


Figure 18: Convergence curves about the optimal fitness on map II.

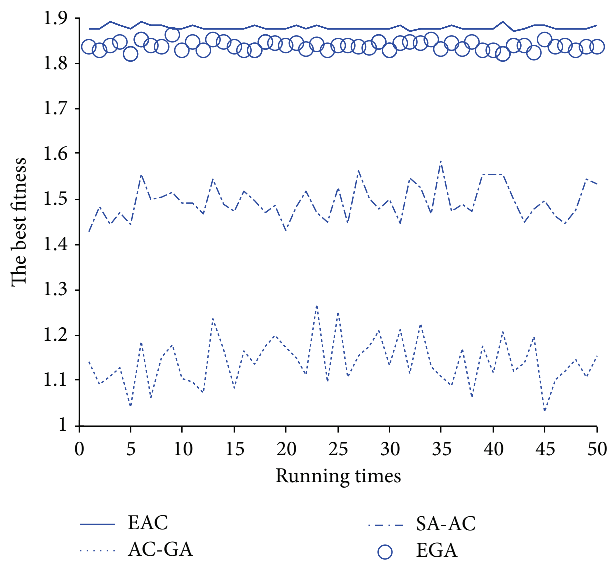


Figure 19: Convergence curves about the optimal fitness running 50 times on map II.

6. Conclusion

Intelligent vehicle is one of the most crucial parts of the whole logistics automation and production automation, in which the routing optimization is one key technology. In the paper we study the route searching problem based on evolutionary ant colony algorithm with RFID technology. We first build the environment map and give the target goal. When searching the optimal route in order to overcome the defect of traditional ant colony algorithm, such as easy falling into local optimum and slow convergence, based on the experiential knowledge, we propose an evolutionary ant colony algorithm.

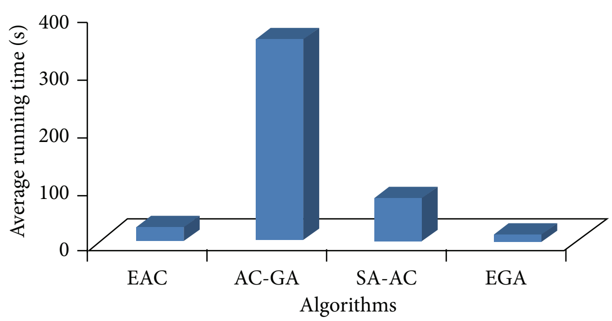


Figure 20: Comparison of average running time on map II.

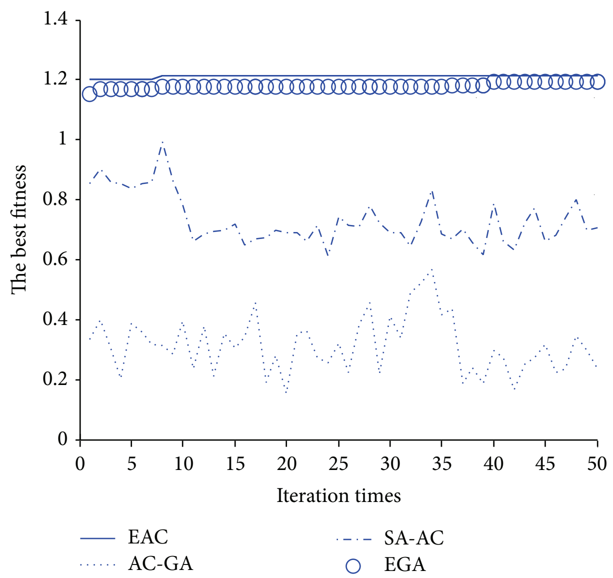


Figure 21: Convergence curves about the optimal fitness on map III.

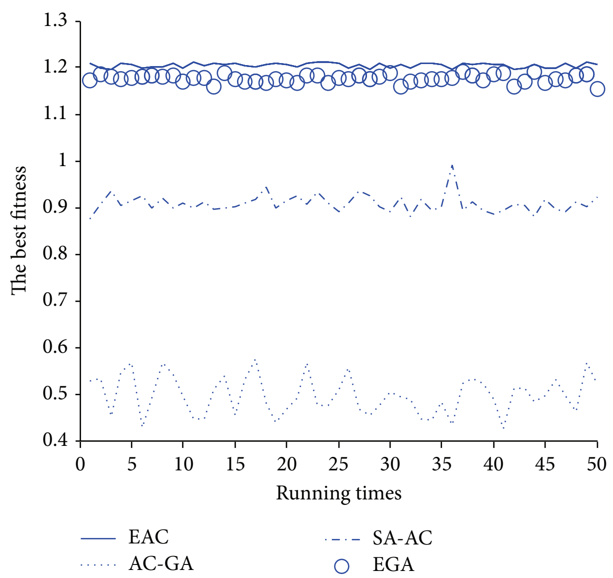


Figure 22: Convergence curves about the optimal fitness running 50 times on map III.

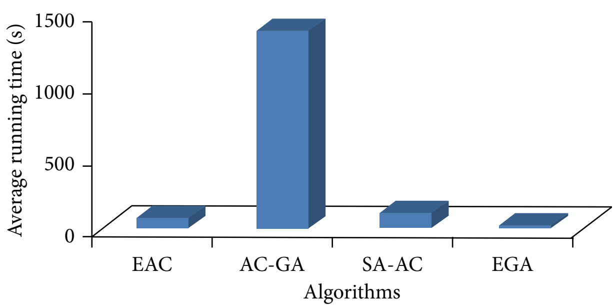


Figure 23: Comparison of average running time on map III.

The new algorithm adopts elite ant strategy, ant fallback strategy, and pheromones evaporation coefficient adaptive adjustment strategy which is proved feasible and effective. When the groups experience is updated, the optimizing operations of abandoning roundabout and reducing turns by parallelogram based on experiential knowledge are done. A lot of experimental results show that the new algorithm is practical and efficient. It is also proved that the algorithm has a high convergence speed and can find the optimal route with higher probability. Due to the fact that the actual working is more complex, how to use the advanced technology to help the vehicle obtain more information timely and dynamically avoiding obstacles is still worth researching.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research is supported by the Natural Science Foundation of Tianjin, research Fund no. 14JCYBJC15900, Tianjin University Innovation Fund, Fund no. 2014XRG-0100, and Humanities and Social Sciences Project of Shandong Provincial Education Department, Fund no. J13WG19.

References

[1] Z.-Z. Yu, J.-H. Yan, J. Zhao, Z.-F. Chen, and Y.-H. Zhu, "Mobile robot path planning based on improved artificial potential field method," *Journal of Harbin Institute of Technology*, vol. 43, no. 1, pp. 50–55, 2011.

[2] X. Sijun and C. Qiying, "A visibility graph based path planning algorithm for mobile robot," *Computer Application and Software*, vol. 3, no. 28, pp. 220–236, 2011.

[3] F. Kuang and Y.-N. Wang, "Robot path planning based on hybrid artificial potential field/genetic algorithm," *Journal of System Simulation*, vol. 18, no. 3, pp. 774–777, 2006.

[4] A. Tuncer and M. Yildirim, "Dynamic path planning of mobile robots with improved genetic algorithm," *Computers and Electrical Engineering*, vol. 38, no. 6, pp. 1564–1572, 2012.

[5] X.-F. Yang and J.-H. Fu, "Mobile robot path planning based on grid algorithm and CGA," *Computer Simulation*, vol. 29, no. 7, pp. 223–226, 2012.

[6] R. Glasius, A. Komoda, and S. C. A. M. Gielen, "A biologically inspired neural net for trajectory formation and obstacle avoidance," *Biological Cybernetics*, vol. 74, no. 6, pp. 511–520, 1996.

[7] W. Cai, Q. Zhu, and J. Hu, "Path planning based on biphasic ant colony algorithm and fuzzy control in dynamic environment," in *Proceedings of the 2nd International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC '10)*, vol. 1, pp. 333–336, August 2010.

[8] C. L. Liu, "Improved ant colony genetic optimization algorithm and its application," *Journal of Computer Applications*, vol. 33, no. 11, pp. 3111–3113, 2013.

[9] T.-F. Liu and R.-Y. Cheng, "Ant colony algorithm with elitist strategy and vision detection for mobile robot path planning," *Computer Applications*, vol. 1, no. 1, pp. 92–93, 2008.

[10] Z.-P. Zhou and L. Hua, "Improved colony algorithm for path planning under complex environment," *Computer Engineering and Design*, vol. 32, no. 5, pp. 1773–1776, 2011.

[11] Q. Sun, *The Research on Path Planning of Automated Guided Vehicle System*, Zhejiang University, Hangzhou, China, 2012.

[12] G. Wang and Y.-J. Feng, "Picking path plan for carousel based on ant colony optimization algorithm," *Computer Engineering*, vol. 36, no. 3, pp. 221–223, 2010.

[13] M. R. Zeng and H. L. Zong, "Route optimization of AS/RS based on simulated annealing genetic algorithm," *Manufacture Automation*, vol. 31, no. 4, pp. 17–19, 2009.

[14] L. Pang and J. G. Lu, "Order picking optimization of automated warehouses based on the ant colony generic algorithm," *Computer Engineering & Science*, vol. 34, no. 3, pp. 148–151, 2012.

[15] M. Z. Liu, J. Lu, M. G. Ge, Z. Q. Jiang, and J. Hu, "Dynamic material handling route planning based on real-time operation conditions," *Journal of Hefei University of Technology*, vol. 31, no. 12, pp. 1948–1952, 2008.

[16] C. Chen, J. Tang, and Z. G. Jin, "Indoor path planning for seeing eyes robot based on RFID," *Journal of Jiangsu University of Science and Technology*, vol. 27, no. 1, pp. 60–63, 2013.

[17] X. X. Guo, *Intelligent Navigation and Scheduling of Vehicles in Warehouses Based on RFID*, Zhejiang University, Hangzhou, China, 2011.

[18] J.-H. Gu, P.-P. Fan, Q.-Z. Song, and E.-H. Liu, "Improved culture ant colony optimization method for solving TSP problem," *Computer Engineering and Applications*, vol. 46, no. 26, pp. 49–52, 2010.

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 576213, 7 pages
http://dx.doi.org/10.1155/2014/576213

Research Article

Optimization of the Actuarial Model of Defined Contribution Pension Plan

Yan Li,[1] Yuchen Huang,[1] and Yancong Zhou[2]

[1] *School of Insurance and Economics, University of International Business and Economics, Beijing 100029, China*
[2] *School of Information Engineering, Tianjin University of Commerce, Tianjin 300134, China*

Correspondence should be addressed to Yancong Zhou; zycong78@126.com

Received 17 March 2014; Accepted 20 May 2014; Published 12 June 2014

Academic Editor: Xiaochen Sun

Copyright © 2014 Yan Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The paper focuses on the actuarial models of defined contribution pension plan. Through assumptions and calculations, the expected replacement ratios of three different defined contribution pension plans are compared. Specially, more significant considerable factors are put forward in the further cost and risk analyses. In order to get an assessment of current status, the paper finds a relationship between the replacement ratio and the pension investment rate using econometrics method. Based on an appropriate investment rate of 6%, an expected replacement ratio of 20% is reached.

1. Introduction

China's current social security system was established July 1, 1997, which was updated in 2005. When it comes to supplemental pension plan, it is used as a way to provide more benefits over the social security system which is discussed in [1]. The supplemental pension plan functions in the following two fields. The first one is to complement the insufficient social pension insurance treatment. The second one is to stabilize employees and keep enterprises employees focus on their work. Retirement plans of supplemental pension are divided into defined benefit and defined contribution plans. In a defined benefit plan, the funding is intended to ensure that sufficient pension funds will be available to pay promised benefits. In a defined contribution plan in [2], the benefit received equals the contributions to the plan, adjusted for possible investment.

Though China supplemental pension plan has been carried out commonly, there are some large troubles. In 2011, the replacement ratio of whole pension plan was only 42.9%. When reaching 80%, we will judge that the protection of retirement employees is adequate. Some researchers expect that the average replacement ratio of developed supplemental pension system will reach 20%, which is appropriate.

Foreign researches on supplemental pension plan are wide. Leime and Seling regarded supplemental pension plan as a deferred survival annuity. Within the prescribed period of time, the insured is to pay a fee every year to the insurer, which will be maintained until the insured retires. If the insured is still survival at retirement, the insured can achieve a certain amount from the insurer each year until death. The actuarial present value of the insured's contribution equals the actuarial present value of the insurer's future payments.

In this paper, application is given priority. We try to compare different types of defined contribution pension plans and test their effectiveness under the consideration of costs and risks. All of these try to give some explanations of the actual application of the pension plan and provide some suggestions for China pension market.

2. Data and Assumption

The sample data used in our analyses is contained in Table 1, simply the basic information of employees, which comes from a certain governmental institution. The institution has 15 executive management level employees.

The China life insurance industry experience life table (2000–2003) is also used in the spreadsheet model.

As discussed by Bills and Lyon, the following assumptions are also used in the model in [3].

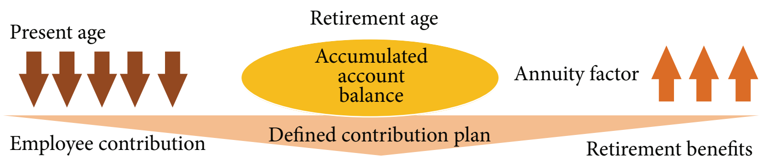


FIGURE 1: Defined Contribution Plan Mechanisms. Resource: Model Calculation Analysis.

TABLE 1: Sample employees information.

ID	Name	Gender	Present age	Entry age	Present salary per month
1	A	F	54	46	29190
2	B	M	47	45	39591
3	C	M	46	32	65100
4	D	M	44	32	29190
5	E	F	42	34	26625
6	F	F	46	38	23000
7	G	M	40	31	39591
8	H	F	36	26	15954
9	I	M	40	28	27800
10	J	F	50	36	5023
11	K	M	44	37	27957
12	L	F	50	32	4556
13	M	F	37	30	27957
14	N	M	49	39	26625
15	O	F	47	30	16751

Resource: government institution sample data.

TABLE 2: Assumption of salary increase rate.

Future service years	Salary increase rate
$0 < \text{Year} \leq 5$	10%
$5 < \text{Year} \leq 10$	8%
$10 < \text{Year} \leq 15$	6%
$\text{Year} > 15$	5%

Resource: spreadsheet model assumption.

(i) In this paper, all of the pension plans are based on mechanism of defined contribution model.

(ii) The present time is 7/1/2013.

(iii) The data of annuity factor in life table is used for calculations.

(iv) The retirement age for male is 60, for female 55.

(v) The retirement benefits are paid until death, with a discount rate of 80%. This assumption presents the possibility of deducting benefits and potential risks.

(vi) The fixed rate of monthly salary contribution is 10%, which is a general level.

(vii) The future salary increase rate per year is related to future services year as Table 2 shows.

(viii) The interest rate is 5% per year, 0.41% per month.

The risk-free rate varies from 5% to 6%. The minimum one, 5%, is first chosen in the model analysis. This assumption is performed in the following sensitivity analysis.

(ix) The Investment return rate is 5%.

The investment of China supplemental pension individual accounts has many limits. The percentage of fixed interest products (mainly the Treasury bond) must be greater than 50%. The percentage of equity products must be less than 30%. In particular, stocks must be less than 20%. In practice, the expected invested return is 5.5%. We assume that it is 5%.

3. Model Description (The Spreadsheet Model Is Contained in Table 3)

The work is completed using a basic "defined contribution plan" spreadsheet model. The relative symbols are contained in Table 4. Three different defined contribution plans are tested. The analysis is performed using a spreadsheet model, entering the established assumptions as parameters into the model. Using the Input fields, various parameters related to the employees' information are entered. The model is used to calculate replacement ratios at retirement age for an individual employee under the conditions of different plans.

Calculation Original Mechanism. See Figure 1.

Symbol Description. See Table 4.

Calculation Basic Formula.

(i) Original defined contribution plan: $\text{EC} * \text{Sn} = \text{RB} * \text{AF}$; $\text{RR} = \text{RB}/S$.

(ii) Modified defined contribution plan: $\text{IA} + \text{EC} * \text{Sn} = \text{RB} * \text{AF}$; $\text{RR} = \text{RB}/S$.

(iii) Hybrid plan: $\text{EC} * \text{Sn} = \text{RB} * \text{AF}$; $\text{RR} = \text{RB}/S$ $(\text{RB} \geq \text{PB})$; $\text{RR} = \text{PB}/S$ $(\text{RB} < \text{PB})$.

4. Analysis

4.1. Sample Human Demographic Analysis. Human demographic information is an important external factor for a pension plan. We do some analyses on the data of 15 executive management level employees and show some findings and graphics below.

Table 3: Summary.

ID	Gender	Present age	Entry age	Present salary per month	Original DC	Modified DC	Hybrid plan	Modified cost	Hybrid cost
1	F	54	46	29,190	0.79%	5.72%	24.92%	342,531.00	1,409,256.00
2	M	47	45	39,591	8.02%	9.16%	9.47%	99,636.00	192,012.00
3	M	46	32	65,100	8.57%	19.40%	8.57%	1,569,581.00	0.00
4	M	44	32	29,190	10.39%	20.43%	12.56%	571,364.00	215,232.00
5	F	42	34	26,625	6.87%	11.43%	14.09%	312,432.00	746,028.00
6	F	46	38	23,000	4.56%	8.56%	17.40%	269,894.00	1,074,372.00
7	M	40	31	39,591	12.91%	19.86%	12.91%	536,557.00	0.00
8	F	36	26	15,954	10.52%	17.32%	19.84%	246,680.00	683,952.00
9	M	40	28	27,800	12.91%	22.95%	12.91%	544,157.00	0.00
10	F	50	36	5,023	2.68%	11.08%	98.89%	121,106.00	1,416,276.00
11	M	44	37	27,957	10.39%	15.52%	13.11%	279,672.00	258,828.00
12	F	50	32	4,556	2.68%	14.74%	109.03%	157,799.00	1,419,936.00
13	F	37	30	27,957	9.98%	14.38%	11.89%	279,672.00	233,820.00
14	M	49	39	26,625	6.89%	14.04%	15.83%	411,674.00	704,244.00
15	F	47	30	16,751	4.13%	15.30%	25.80%	532,804.00	1,222,500.00

Table 4: Symbol determination.

Employer contribution per month	EC
Annuity factor (life table)	AF
Years before retirement	N
Replacement ratio	RR
Initial account amounts	IA
Accumulated account balance	AA
Retirement benefits per month	RB
Accumulated value of annuity	Sn
Salary per month	S
Promised benefits	PB

Resource: model symbol determination.

(i) The average age of employees is high, 44.80. For men, the average age is 44.28; for women, the average age is 45.25. As the women will get retired 5 years earlier than men, the pressure of pension plan for women is greater.

(ii) All employees are under the same social security system, which I don't consider in my analyses. The defined contribution pension plan is used for supplemental pension plan.

(iii) Some employees have high age with shorter future service years. 5 employees will retire in 10 years. Their salary is on the low side, compared to the younger. The monthly salary of two employees aged 50 is less than 5,500. The future accumulated years for individual accounts may not be enough.

(iv) The salary of the employees who have a longer past services year (more than 10 years) is more different, 65,100 as the maximum and 4,556 as the minimum. The salary for shorter service years' employees is more stable.

4.2. *Defined Contribution Pension Plan*

4.2.1. Original Defined Contribution Pension Plan. A defined contribution plan will provide many details to determine the annual contribution made by employer on behalf of each employee as well as any relevant regulations. In analysis, contribution rate is assumed. The formula determines how much an employee receives from each plan.

Based on the data, assumptions, and model above, the replacement ratio of all employees are calculated under original defined contribution plan. First, we calculate the expected accumulated individual account at retirement age for each employee. Then, with the annuity factor provided by life table, we estimate the future retirement benefits. Finally, the replacement ratio is obtained.

The results are summarized in Table 5. The results vary from 0.79% (the oldest one) to 12.91% (the youngest one). We can almost see that "the younger, the better". The average replacement ratio is 7.49%, far less than the average replacement ratio of China, 20%. In fact, no employees will have a reasonable replacement ratio, over 20%.

The accumulated period (future service years) of individual accounts is adequate, certainly including the sample company. The pure pension plan, defined contribution, can't provide appropriate retirement benefits coverage.

4.2.2. Modified Defined Contribution Pension Plan. According to the results in Table 5, we can conclude that the supplemental retirement compensation, simply the defined contribution pension plan, is not enough for employees and more compensation expenses are necessary. Some solutions are proposed.

Table 5: Replacement ratio for original defined contribution pension plan.

ID	Gender	Present age	Entry age	Present salary per month	Original replacement ratio
1	F	54	46	29,190.00	0.79%
2	M	47	45	39,591.00	8.02%
3	M	46	32	65,100.00	8.57%
4	M	44	32	29,190.00	10.39%
5	F	42	34	26,625.00	6.87%
6	F	46	38	23,000.00	4.56%
7	M	40	31	39,591.00	12.91%
8	F	36	26	15,954.00	10.52%
9	M	40	28	27,800.00	12.91%
10	F	50	36	5,023.00	2.68%
11	M	44	37	27,957.00	10.39%
12	F	50	32	4,556.00	2.68%
13	F	37	30	27,957.00	9.98%
14	M	49	39	26,625.00	6.89%
15	F	47	30	16,751.00	4.13%
		Average			7.49%

Resource: model calculation results.

Considering the past services years for each employee, there should be an initial accumulated individual account. This amount will also be invested until retirement. We assume that the past salary is the present salary provided and the contribution rate is still fixed 10%. Therefore, the accumulated individual account at retirement age, the expected replacement ratio, will increase. The results are summarized in Table 6.

From the results, we can see that the compensation expense is a huge expenditure for employer. Some employees have served for long years. The modified defined contribution plan may be easy to accept. The initial account is a big amount.

However, the replacement ratio is increased but it is still unexpected, almost still less than 20%, possibly because of the future increased salary and the inadequate investment return. The average replacement ratio of this plan is 14.66%, with an average increase of 7.17%. Employees 4, 7, and 9 may be satisfied, as their length of service is long and their salary is not too high. With a relatively low salary, the pressure of pension benefit will be much smaller.

4.2.3. Hybrid Pension Plan. The hybrid pension plan will guarantee the profits more effectively. In this plan, the benefit formula looks like a defined contribution plan, but the plan is really a defined benefit plan. The plan administrator keeps track of the monthly contributions into the account and guarantees a crediting pension level. If the employees earn less, the plan must fund the shortfall.

Our sample company is a governmental one, reasonably with better pension care. As some employees' estimated monthly pension is not enough, the following hybrid pension plan is more direct. We assume that every employee should have a credit monthly pension of 8,000. Therefore, it is necessary to replenish future pension for those with less than certain 8,000. Thus the expected replacement ratio for certain employees will increase. The results of hybrid pension plan are summarized in Table 7.

The hybrid pension is better than the modified pension. The 8,000 guarantee per month is directly added to the future

Table 6: Replacement ratio for modified defined contribution pension plan.

ID	Gender	Present age	Entry age	Present salary per month	Replacement ratio
1	F	54	46	29,190.00	5.72%
2	M	47	45	39,591.00	9.16%
3	M	46	32	65,100.00	19.40%
4	M	44	32	29,190.00	20.43%
5	F	42	34	26,625.00	11.43%
6	F	46	38	23,000.00	8.56%
7	M	40	31	39,591.00	19.86%
8	F	36	26	15,954.00	17.32%
9	M	40	28	27,800.00	22.95%
10	F	50	36	5,023.00	11.08%
11	M	44	37	27,957.00	15.52%
12	F	50	32	4,556.00	14.74%
13	F	37	30	27,957.00	14.38%
14	M	49	39	26,625.00	14.04%
15	F	47	30	16,751.00	15.30%
		Average			14.66%

Resource: model calculation results.

Table 7: Replacement ratio for hybrid pension plan.

ID	Gender	Present age	Entry age	Present salary per month	Replacement ratio
1	F	54	46	29,190.00	24.92%
2	M	47	45	39,591.00	9.47%
3	M	46	32	65,100.00	8.57%
4	M	44	32	29,190.00	12.56%
5	F	42	34	26,625.00	14.09%
6	F	46	38	23,000.00	17.40%
7	M	40	31	39,591.00	12.91%
8	F	36	26	15,954.00	19.84%
9	M	40	28	27,800.00	12.91%
10	F	50	36	5,023.00	98.89%
11	M	44	37	27,957.00	13.11%
12	F	50	32	4,556.00	109.03%
13	F	37	30	27,957.00	11.89%
14	M	49	39	26,625.00	15.83%
15	F	47	30	16,751.00	25.80%
		Average			27.15%

Resource: model calculation results.

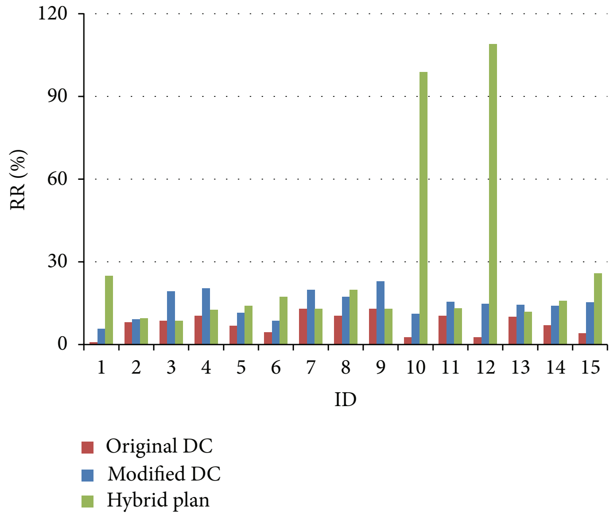


FIGURE 2: The replacement ratios of three defined contribution pension plans. Resource: spreadsheet model calculation results.

retirement benefits, more effectively. The average replacement ratio of hybrid pension plan is 27.15%, with an average increase of 19.66%.

The minimum replacement ratio is 8.57%, with no change, due to his maximum salary. The replacement ratio of employees 10 and 12 even exceed 95% because of their relatively low salary. Just due to the low salary, their initial account in modified defined contribution plan is small, ineffectively.

Up to now, three defined contribution pension plans have been analyzed above. The comparison graph is shown in Figure 2. It is obvious that, the replacement ratio is improved under the two improved plans, especially the hybrid plan.

4.3. Cost Analysis. The modified defined contribution pension plan and the hybrid pension plan can be regarded as improved defined contribution plan. There is an estimated compensation expense for this improvement.

In modified defined contribution pension plan, the sample employer must supplement the initial individual account value at once. Employees with longer past service years require more expenditure. Certainly, there is a time value for this compensation. In this paper, we just measure the present value of the required costs.

In hybrid pension plan, the compensation expenses only appear for the employees who need to be assisted. There is a measurable amount for employers. With this added amount, the monthly 8,000 pension benefits will be reached. In order to compare these two plans' compensation expenses, they are both measured by the use of present value and completed in the model.

The detailed results are summarized in Table 8.

According to the results, there is no doubt that more expenses make more benefits. As the average replacement ratio of hybrid plan increases more, the estimated compensation expenses for employer are much larger.

TABLE 8: Present value of compensation expenses.

ID	Gender	Present age	Entry age	Cost for modified defined contribution plan	Cost for hybrid plan
1	F	54	46	342,531.00	1,409,256.00
2	M	47	45	99,636.00	192,012.00
3	M	46	32	1,569,581.00	0.00
4	M	44	32	571,364.00	215,232.00
5	F	42	34	312,432.00	746,028.00
6	F	46	38	269,894.00	1,074,372.00
7	M	40	31	536,557.00	0.00
8	F	36	26	246,680.00	683,952.00
9	M	40	28	544,157.00	0.00
10	F	50	36	121,106.00	1,416,276.00
11	M	44	37	279,672.00	258,828.00
12	F	50	32	157,799.00	1,419,936.00
13	F	37	30	279,672.00	233,820.00
14	M	49	39	411,674.00	704,244.00
15	F	47	30	532,804.00	1,222,500.00
		Total cost		6,275,559.00	9,576,456.00

Resource: model calculation results.

4.4. Risk Analysis. Every retirement pension plan presents several risks and the design of each plan determines how the risks are divided between employers and employees. Employers have the main responsibility on defined contribution plans.

One obvious risk is investment return. If the asset earns more than expected, the employee under this risk is happier. If the asset earns less, there may be a loss of benefits. In a word, the ultimate accumulated accounts depend on investment earnings to some degree. Investment returns help provide for promised benefits. Everyone expects to have an account which is overperformed. This risk is predominant in the actual operation.

Another risk is inflation rate. In this paper, we regard inflation as one component of investment return. The investment return will provide some protection. The inflation depreciates the value of future accumulated accounts.

The next risk is longevity. In this paper, we assume that the future pension benefits will be paid until death. Regardless of the possibility of ignoring this risk, it is traditionally considered that a defined benefit plan puts this risk on the employers and a defined contribution plan puts this risk on the employees. In our opinion, this conclusion is reasonable as a defined benefit plan/hybrid plan guarantees a promised benefit.

Finally, the salary increase risk is considered. The salary at retirement age is a predominant factor which needs to be monitored. If the salary is relatively higher than

Table 9: Results of sensitivity analysis (replacement ratio).

Tested parameter	Changes	Original defined contribution	Modified defined contribution	Hybrid plan
Investment return	(+1%)	0.55%	1.45%	0.21%
	(−1%)	−0.50%	−1.29%	−0.19%
Salary increase	(+1%)	−0.50%	−1.27%	−1.89%
	(−1%)	0.55%	1.46%	2.12%

Resource: model calculation results.

expected, there is more pressure and more contributions for the certain pension plan. The final influence is uncertain. In the assumption, the salary increase rate is predicted based on the future service years.

4.5. Sensitivity Analysis. In order to measure the potential risks for several retirement pension plans, the following sensitivity analysis is completed. As we assume that the pension benefits will be paid until death, longevity risk is not considered in this part. The results are summarized in Table 9.

Based on the results in Table 9, it is evident that the employees are under some risks when the investment return rate declines by 1% annually. If the inflation rate increases by 1% annually, it will have a similar impact on the replacement ratios. The investment return risk of hybrid plan is not as severe as the other two pension plans. The modified defined contribution pension plan will be influenced most. This result is easy to understand, in a modified defined contribution pension plan, and the time value of the initial individual account at retirement age is the core of this plan.

The salary at retirement age will have a direct influence on the replacement ratio. The hybrid pension plan is most sensitive to this risk. Most employees have a fixed monthly pension benefit of 8,000 in this plan; the change of the denominator of replacement ratio is much more important. The modified defined contribution pension plan is still influenced by this risk. When the salary increases faster, the accumulated individual account, and thus estimated pension benefit, will be greater.

In conclusion, though the two improved pension plans will get a better result of replacement ratios, the potential risks will increase to some degree. The modified defined contribution pension plan is sensitive to both of the risk parameters. The hybrid pension plan is more sensitive to salary increase rate. When it comes to our traditional defined contribution pension plan, it is the steadiest one, with less fluctuation. As a result, both of the risks need to be monitored.

5. Forecast

In order to forecast the pension replacement ratio, we simplify defined contribution pension plan. The important parameters are kept retained, including investment income rate/interest rate (i), pension benefit (B), contribution rate (c), contribution years (n), pension paid years (m), present salary, (S) and salary increase rate (g), as in [4].

Based on the parameters listed above, we can get the accumulated individual account. We assume that the contribution is made at the end of years and investment income rate is not equal to salary increase rate.

$$\begin{aligned}\text{Account} &= cS(1+g)(1+i)^{n-1} \\ &\quad + cS(1+g)^2(1+i)^{n-2} + \cdots + cS(1+g)^n \\ &= \frac{cS(1+g)\left[(1+i)^n - (1+g)^n\right]}{i-g}, \\ \text{Account} &= B\frac{1-1/(1+i)^m}{i}.\end{aligned} \tag{1}$$

According to the formula (1), the formula of B is obtained.

So,

$$B = \frac{ciS(1+g)(1+i)^m\left[(1+i)^n - (1+g)^n\right]}{(i-g)\left[(1+i)^m - 1\right]}. \tag{2}$$

Finally, the replacement ratio

$$\text{RR} = \frac{B}{S(1+g)^n} = \frac{ci(1+g)(1+i)^m\left[(1+i)^n/(1+g)^n - 1\right]}{(i-g)\left[(1+i)^m - 1\right]}. \tag{3}$$

According to the sample data, employee 7 is chosen for example. The future service year is 20 years, $n = 20$. The fixed contribution rate is 10%, $c = 10\%$. The salary increase rate is 5%, $g = 5\%$. According to the life table and China pension market, we assume that m equals 18.77 years. Up to now, we find a basic formula relationship between investment return rate (i) and pension replacement ratio. As the investment return rate changes, the replacement ratio will change.

In order to forecast the replacement ratio, Table 10 is shown.

It is obvious that the replacement ratio increases as the interest rate and certainly investment return increase. With the method of econometrics, linear regression, the relation between them is presented by "Eviews 5.0". To minimize the expected error as much as possible, we vary the form of interest rate, using quadratic and cube interest rate as explaining variables. The biquadrate interest rate is not significant, which is discarded.

Regression Formula:

$$\begin{gathered}\text{RR} = 0.050925 + 2.388865i - 11.97265i^2 + 216.2676i^3 + \varepsilon \\ \text{Std.Error:} \\ (0.006067) \quad (0.277030) \quad (3.748962) \quad (15.51312) \\ t\text{-Statistic:} \\ (8.393567) \quad (8.623115) \quad (-3.19259) \quad (13.94095) \\ \text{Prob:} \\ (0.0002) \quad (0.0001) \quad (0.0188) \quad (0.0000) \\ R\text{-Squared:} \quad 0.999971.\end{gathered} \tag{4}$$

Table 10: Forecast relationship (partial).

Interest rate (i)	3%	4%	6%	7%	8%	9%	10%	11%	12%	13%
Replacement ratio (RR)	11.81%	14.04%	19.77%	23.43%	27.73%	32.45%	38.71%	45.69%	53.86%	63.44%

Resource: model calculation results.

Table 11: Investment returns rate.

Year	2007	2008	2009	2010	2011	2012
Asset	8.86B	35.572B	132.795B	227.158B	326.579B	401.98B
Investment rate	7.66%	2.6%	4.36%	2.85%	2.53%	5.68%

Resource: the national enterprise annuity fund business data in 2012, Human resources social security fund supervision, 2013.4.

It is clear that the regression formula fits this relationship well. Under the condition that the other factors are unchanged, with the improvement of investment return, the growth of enterprise annuity replacement rate is more significantly higher than the growth of investment return.

Certainly, the improvement of defined contribution pension plan design will increase the replacement ratio. If we devote ourselves to increasing the investment return rate, the replacement ratio will increase too. According to the corresponding linear regression relationship analysis, if the enterprise annuity replacement ratio level is about 25%, the annual investment yields need to exceed 7%.

The investment structure in China is not well developed now. If we want to have a better investment return, some limits must be cancelled. The investment income rate is around 6%, which will result in a replacement ratio of 20%.

In Table 11 there are the historical investment returns in China. The investment return is not satisfying during recent years. In a word, it is significant for us to think how to increase investment benefits as well as guarantee the security.

In the analysis, we just assume that 5% is the investment return rate. The actuarial assumption is used in the calculations, but it will never exactly reflect the plan's experience this year. The difference between actual and assumed experience creates actuarial gains and losses. A gain means the actual experience was more favorable, and less costly, than assumed. A loss means the opposite. If we want to forecast the future replacement ratio, it is important to get reasonable interest rate estimation which is discussed in [5].

6. Conclusions and Recommendations

The traditional defined contribution plan cannot satisfy the employees' demand, with shorter time to accumulate pension accounts. The initial amounts are not enough. According to the results of two improved plans proposed, modified defined contribution plan and the hybrid plan, both the replacement ratios become improved. The modified defined contribution pension plan considers the past service years of each employee while the hybrid pension plan presents the feature of defined benefit plan and guarantees a fixed pension benefit. The effectiveness of the latter one is better. The two improved defined contribution plans both mean a huge expenditure to employer. In particular, the cost of hybrid pension plan is much larger. In addition, the potential risks for retirement plans are analyzed, investment return, inflation rate, longevity risk rate, and salary increase rate. The relationship between investment return and replacement ratio is clear. If the investment return can be raised under a secure method, the pension market will have a great improvement as discussed in [6]. Much more consideration about the pension accumulated accounts investment strategy is necessary.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors thank the referee and the associate editor for their very useful suggestions and detailed comments. This paper is supported by the Fundamental Research Funds for the Central University in UIBE (13QNLG01).

References

[1] G. E. Rejda, *Social Insurance and Economic Security*, Prentice Hall, Upper Saddle River, NJ, USA, 1999.

[2] Ippolito and A. Richard, "Pension terminations for reversion," *Journal of Pension Planning and Compliance*, vol. 12, no. 3, pp. 221–422, 1986.

[3] C. Bellis and R. Lyon, *Understanding Actuarial Management: The Actuarial Control Cycle*, Society of Actuary, 2003.

[4] Y. Weili and N. Xiao, "The Relationship between Enterprise Annuity Fund Investment Rate and Replacement Ratio," Taiping Pension Company, LTD. Liaoning Branch.

[5] D. Dufresne, "Stability of pension systems when rates of return are random," *Insurance Mathematics and Economics*, vol. 8, no. 1, pp. 71–76, 1989.

[6] H. Yaodong, "The choice of the mode of the supplemental pension plans in China," *Journal of Chongqing Business School*, no. 3, pp. 50–52, 2002.

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 701938, 7 pages
http://dx.doi.org/10.1155/2014/701938

Review Article

Operations Management of Logistics and Supply Chain: Issues and Directions

Xiang Li

College of Economic and Social Development, Nankai University, Tianjin 300071, China

Correspondence should be addressed to Xiang Li; xiangli@nankai.edu.cn

Received 11 April 2014; Accepted 13 May 2014; Published 10 June 2014

Academic Editor: Xiaochen Sun

Copyright © 2014 Xiang Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

There has been consensus that logistics as well as supply chain management is a vital research field, yet with few literature reviews on this topic. This paper sets out to propose some hot issues in the current research, through a review of related literature from the perspective of operations management. In addition, we generate some insights and future research directions in this field.

1. Introduction

Organizations adopt numerous business improvement methodologies to improve business performance. Logistics as well as supply chain management has been regarded to be the crucial factor for the companies to obtain competitive edge. In fact, logistics as well as supply chain management has received attention since the early 1980s, yet conceptually the management of supply chains is not particularly well understood, and many authors have highlighted the necessity of clear definitional constructs and conceptual frameworks on supply chain management. In this paper, we provide a tutorial on the current research of operations management of logistics and supply chain. We first clarify the conception of logistics and supply chain management in this paper, which defines the scope of our related research papers. The core of this paper is that we provide several hot issues in this field with examples to show how these researches contribute from different research angles. Finally, we conclude the paper with the insights obtained from our analysis and future study directions in this field.

The paper is organized as follows. In the next section, we specify the definitions of the terms of logistics and supply chain used in our paper, with a comparison between these two popular conceptions. In Section 3, which is the core section of this paper, we provide several hot topics in current research with detailed examples. In Section 4, we provide insights and further research directions.

2. Conception and Scope

2.1. Logistics. Logistics is the management of the flow of goods between the point of origin and the point of consumption in order to meet some requirements, for example, of customers or corporations. The resources managed in logistics can include physical items, such as food, materials, animals, equipment, and liquids, as well as abstract items, such as time, information, particles, and energy. The logistics of physical items usually involves the integration of information flow, material handling, production, packaging, inventory, transportation, warehousing, and often security. The complexity of logistics can be modeled, analyzed, visualized, and optimized by dedicated simulation software. The minimization of the use of resources is a common motivation in logistics for import and export.

Note that the above definition of logistics is not unified, although it might be indeed, in current environment, a commonly acknowledged one. For example, Council of Logistics Management (now renamed as Council of Supply Chain Management Professionals) referred to logistics as "*the process of planning, implementing, and controlling the efficient, effective flow and storage of goods, services, and related information from point of origin to point of consumption for the purpose of conforming to customer requirements*," which includes inbound, outbound, internal, and external movements and return of materials for environmental purposes.

As we can see, the concept of logistics focuses on the product flow, which is the meaning by which this word has been translated in Chinese. It also puts emphasis on the activities of handling product, which include the storage, transportation, distribution, and packaging and processing. Although business logistics involves many activities, the traditional research of operations management on logistics mainly relates to the fields of logistics facility, transportation, and inventory planning.

2.2. Supply Chain. Compared to "logistics," there appears to be even less consensus on the definition of the term "supply chain management." Kathawala and Abdou [1] point out that SCM "has been poorly defined and there is a high degree of variability in people's minds about what is meant." Nevertheless, we present a rather widely adopted definition, which is given by Mentzer et al. [2] which is rather broad, not confined to any specific discipline area, and adequately reflecting the breadth of issues that are usually covered under this term: "*Supply chain management is defined as the systemic, strategic coordination of the traditional business functions and the tactics across these business functions within a particular company and across businesses within the supply chain, for the purposes of improving the long-term performance of the individual companies and the supply chain as a whole.*"

The terms of "logistics" and "supply chain" are usually comparative in academy and industry, since both of them are closely relevant to the product circulation during its whole life cycle, and both have been regarded as the central unit of competitive analysis of model management science. Generally speaking, supply chain is a more broadened conception with a wider range which can involve other similar subjects, such as network sourcing, supply pipeline management, value chain management, and value stream management [3–5].

In addition, we can see that the conception of logistics has no relationship with organization, which is the opposite of supply chain, since supply chain is made up of multiple organizations, usually companies. An important issue in supply chain management is that companies will not seek to achieve cost reductions or profit improvement at the expense of their supply chain partners but rather seek to make the supply chain as a whole more competitive. Hence, the contention that it is supply chains, and not a single company, that compete is a central tenet in the field of supply chain management [6]. A central research methodology for supply chain management is game theory (and also incentive theory for the scenario of incomplete information).

3. Hot Issues

Due to the extensive research ranges in operations management of logistics and supply chain management, we cannot possibly make a comprehensive review in one paper. In this section, we point out several of the most important issues and hot topics in recent research, which draws great attention from both academy and industry.

3.1. Inventory and Transportation Management on Specific Fields. As has been pointed out in the previous section, the operations research on logistics management still mainly focuses on the traditional domain, that is, the inventory (including production planning) and transportation management. However, a noticeable phenomenon is that most papers are putting emphasis on specific fields with remarkable features captured into their models and thus making new contributions to the literature.

For example, the inventory management of perishable products (also referred to as deteriorating product) is a rather old and mature field in logistics and supply chain management, with replenishment policies for inventory being the main focus of study. Whitin [7] investigated such a problem, where fashion goods deteriorating at the end of certain storage periods were considered. Since then, considerable attention has been paid to this line of research. Nahmias [8] provides a comprehensive survey of research published before the 1980s. Studies in recent years on the deteriorating inventory models can be found in Raafat [9] and Goyal and Giri's [10] papers, in which relevant literature published in the 1980s and 1990s is reviewed, respectively. A more updated review is given in Blackburn and Scudder's [11] paper. However, new models can still be developed to capture the current management feature and obtain new managerial insights. Generally, two types of perishable loss, quantity loss and quality loss, may take place for a perishable product. The majority of the literature has dealt mainly with only one type of loss. In this regard, Cai et al. [12] adopt a stochastic model to study a supply chain in which a distributor procures from a producer a quantity of a fresh product. During the transportation process, the distributor has to make an appropriate effort to preserve the freshness of the product, and his success in this respect impacts both the quality and quantity of the product delivered to the market. Cai et al. [13] further extend the model into a 3-stage supply chain with outsourcing transportation involved.

Another important field is transportation. It is generally known that the research on VRP (vehicle routing problem) and its various extensions has been extensive. However, other new domains on transportation can still be interesting topics. For example, the remarkable growth in intermodal transportation over the past decade has not been matched by a comparable level of academic activity, and, hence, the research on intermodal transportation appears to have a great potential. Chang [14] explores one of the intermodal operational issues: how to select best routes for shipments through the international intermodal network. The problem is formulated as a multiobjective multimodal multicommodity flow problem with time windows and concave costs, and an efficient heuristic is proposed. Vermaa and Verter [15] present a first attempt for the development of an analytical framework for planning rail-truck intermodal transportation of hazardous materials by developing a biobjective optimization model to plan and manage intermodal shipments to represent the current practice; the routing decisions in the model are driven by the delivery times specified by the customers. Bruns and Knust [16] study the problem of load planning for trains in intermodal container terminals. The objective is to assign load units to wagons of a train such that the utilization of the train is maximized and setup and transportation costs in the terminal are minimized. Bruns et al. [17] further study the problem of robust load planning for trains in intermodal container terminals. The goal of load planning is to choose wagon settings and assign load units to wagons of a train such that the utilization of the train is maximized and setup and transportation costs in the terminal are minimized. García et al. [18] adopt a new hybrid approach by combining OR techniques with AI search methods in order to obtain good quality solutions for complex intermodal transport problems, by exploiting the benefits of both kinds of techniques. The solution has been applied to a real-world problem from one of the largest Spanish companies using intermodal transportation.

3.2. Sourcing and Marketing in Supply Chain. Sourcing is the first step in a supply chain. The research on sourcing has been extensive in recent years. This leaves open room for a supplier to improve efficiency over time by further optimizing the production processes. In general, OEMs' shifting of more development and engineering work, which require complex tasks and customized products, to their suppliers implies a significant potential for a supplier to accumulate knowledge and experience from learning, thus reducing costs over time [19–21]. This dynamic change of supply costs affects the negotiation of sourcing contracts.

A noticeable issue is the utilization of auctioning in the sourcing strategy. One of the first researches in this regard might be Chen's [22], which studies a procurement problem with one buyer and multiple potential suppliers who hold private information about their own production costs. An optimal procurement strategy is considered for the buyer who first specifies a payment for each possible purchase quantity and then invites the suppliers to bid for this contract. The auction can be conducted in many formats such as the English auction, the Dutch auction, the first-priced auction, sealed-bid auction, and the Vickrey auction. Chen and Vulcano [23] study a supply chain where an upstream supplier auctions his inventory or capacity as a bundle, which formulates the problem as a two-stage supply chain comprising a single supplier and two resellers. Huh and Janakiraman [24] study periodic-review inventory replenishment problems with auctions and other sales channels and show that the optimality of (s, S) inventory replenishment policies extends well beyond the traditional sales environments studied so far in the inventory literature. Chen et al. [25, 26] study a supply chain in which a single buyer wishes to procure a package of products or services from various competing suppliers that possess private cost information and show how the buyer can optimize his/her profit and at the same time coordinate the channel by using a contract scheme involving auctions, audits, and profit sharing.

For a supplier that provides critical and customized components, the demand closely depends on, and hence is susceptible to, the variation of the final product demand. In the automotive industry, unstable and uncertain domestic volume of individual models is cited as one of the biggest challenges faced by manufacturers due to increased consumer choices [27]. The consumer electronics industry is notorious for risk stemming from short product life cycles and high demand uncertainty [28]. Furthermore, there is typically more uncertainty about the future demand than about the current demand. This demand uncertainty adds another source of future uncertainty, besides possible supplier switching (in a short-term relationship), that influences the decision of initial capacity investment.

Marketing is another end in supply chain. The collaboration with marketing science massively extends the domain of supply chain management. Pricing, promotion, and channel management are the three most important areas in this regard. Pricing and promotion are the central issues in marketing management, let alone under consideration of the supply chain environment. Li and Graves [29] explore the pricing decisions during intergenerational product transition, by formulating the dynamic pricing problem and deriving the optimal prices for both the old and new products. The optimal initial inventory for each product is also determined, and a heuristic method is discussed. Li and Zhang [30] study the preorder strategy that a seller may use to sell a perishable product in an uncertain market with heterogeneous consumers. They find that accurate demand information may improve the availability of the product, which undermines the seller's ability to charge a high preorder price. As a result, advance demand information may hurt the seller's profit due to its negative impact on the preorder season. Sainathan [31] considers pricing and ordering decisions faced by a retailer selling a perishable product with a two-period shelf life over an infinite horizon. Sinitsyn [32] investigates the outcome of a price competition between two firms, each producing two complementary products. It is found that each firm predominantly promotes its complementary products together, which is correlationally supported by data in the shampoo and conditioner and in the cake mix and cake frosting categories. Liu et al. [33] examine the efficacy of cost sharing in a model of two competing manufacturer-retailer supply chains who sell partially substitutable products that may differ in market size. Some counterintuitive findings suggest that the firms performing the advertising would rather bear the costs entirely if this protects their unit profit margin. Gao et al. [34] show that the weather-conditional rebate program can increase sales by price discriminating among a customer's postpurchase states. Taking advantage of the early sales, it can also reduce the inventory holding cost and ordering cost and hence can increase the retailer's expected profits.

In addition, channel management is also an important interface between marketing and supply chain. Chen et al. [25, 26] study a manufacturer's problem of managing his direct online sales channel together with an independently owned bricks-and-mortar retail channel, when the channels compete in service. They identify optimal dual channel strategies that depend on the channel environment described by factors such as the cost of managing a direct channel, retailer inconvenience, and some product characteristics. Brynjolfsson et al. [35] investigate local market structures for traditional retailers and then match these data to a dataset on consumer demand via two direct channels: Internet and catalog. Their analyses show that Internet retailers face significant competition from brick-and-mortar retailers when selling mainstream products but are virtually immune from competition when selling niche products. Guo [36] investigates optimal disclosure strategies/formats in a channel setting with bilateral monopolies and shows that retail disclosure leads to more equilibrium information revelation. Chiang [37] extends the single-period vertical price interaction in a manufacturer-retailer dyad to a multiperiod setting, in which a manufacturer distributes a durable product through an exclusive retailer to an exhaustible population of consumers with heterogeneous reservation prices. The open-loop, feedback, and myopic equilibria for this dynamic pricing game are explored and compared to the centralized solution.

3.3. Green Logistics and Supply Chain. Green logistics refers to a logistics form which plans and implements green transport, green storage, green packaging, green circulation processing, green recovery, and other activities via advanced logistics technology. It aims to reduce environmental pollution and resource consumption arising from logistics activity so as to realize a "win-win" consequence in logistics development and eco-environmental conservation. As an important avenue for realizing the sustainable development strategy, greater attention has been given to green logistics which will play an important role in industrial upgrading, transformation of economic structure, promotion of logistics development level, and other relevant aspects. Green supply chain is the supply chain management with similar objectives and core implications. Green logistics as well as supply chain management is also usually referred to "sustainable" management.

A typical field in green logistics and supply chain management is reverse logistics, sometimes called closed-loop supply chains, in which there are reverse flows of used products (postconsumer) back to manufacturers. There has been substantial research into production planning and inventory management in remanufacturing systems. Simpson [38] first studies a periodic review inventory system with stochastic and mutually dependent demands and returns and provides the optimality of a three-parameter inventory policy. Kelle and Silver [39] consider a different model with independent demand and return processes, where all returned products should be remanufactured. Inderfurth [40] shows that the optimal policy derived by Simpson [38] is still optimal in the case of fixed cost when lead times for remanufacturing and manufacturing are identical. Van der Laan et al. [41] analyze a push control strategy and a pull control strategy in a hybrid system and compare them with the traditional systems without remanufacturing. Teunter et al. [42] explore the superior inventory strategies for hybrid manufacturing/remanufacturing systems with a long lead time for manufacturing and a short lead time for remanufacturing. Wang et al. [43] analyze the impacts of the amount of products manufactured and the proportion of the remanufactured part to the returned products on the total cost of the hybrid system, showing that the cost could be reduced significantly if these two critical values are optimally set. Other related works include Kiesmüller [44], Tang and Grubbström [45], Aras et al. [46]. For a comprehensive review, I refer the reader to Fleischmann et al. [47], Dekker et al. [48], and Ilgin and Gupta [49].

A typical feature in reverse logistics and closed-loop supply chains is the quality uncertainty of acquired used product, which is usually expressed by a random remanufacturing yield and has been studied in some recent papers. Inderfurth [50] shows that the uncertainty in returns and demand can be an obstacle to an environmental-benign recovery strategy within a reverse logistics system. Inderfurth and Langella [51] develop heuristics for the problem of obtaining parts for remanufacturing by disassembling used products or procuring new ones, under the consideration of random disassembly yields. Galbreth and Blackburn [52] explore acquisition and sorting/remanufacturing policies in the case of a continuum of quality levels for cores with fixed quality distribution. The main premise is that remanufacturing costs will go down if only the returned products with better quality are remanufactured. Ketzenberg et al. [53] explore the value of information in the context of a firm that faces uncertainty with respect to demand, product return, and product remanufacturing yield by first analyzing a simple single-period model and then proving that the results carry over multiperiod setting. Çorbacioğlu and van der Laan [54] analyze a two-product system with end-product stock containing both manufactured and remanufactured products while the remanufacturable stock may contain products of different quality. Zikopoulos and Tagaras [55] investigate the production problem in a reverse supply chain consisting of two collection sites and a refurbishing site and examine how the profitability of reuse activities is affected by uncertainty regarding the quality of returned products. Denizel et al. [56] propose a stochastic programming formulation to solve the remanufacturing production planning problem when inputs of the remanufacturing system have different and uncertain quality levels and capacity constraints.

Although the research on remanufacturing systems is vast, there are only a few papers that consider a market-driven acquisition channel for used products. Guide and Jayaraman [57] and Guide and van Wassenhove [58] are the first to investigate this field, pointing out the importance of used product acquisition management to deal with the uncertainty in timing, quantity, and quality of the returned products. Guide et al. [59] develop a quantitative model to determine the optimal acquisition prices of used products and the selling price of remanufactured products, assuming that the quantity of return items can be fully controlled by the acquisition price. Bakal and Akcali [60] extend the model of Guide et al. [59] into the case of random remanufacturing yield and analyze the impact of yield on the remanufacturing profitability. Karakayali et al. [61] study the problem of determining the optimal acquisition price of the end-of-life products and the selling price of the remanufactured parts under centralized as well as decentralized remanufacturer-driven and collector-driven decentralized channels.

3.4. Behavior Operations. The decisions under the consumers' behavior are important for the firms to gain competitive edge and obtain more profit. The customer's behavior can be loss averse, risk averse, regretful, and strategic, and the papers incorporating such factors are regarded as increasingly important. Kök and Xu [62] study assortment planning and pricing for a product category with heterogeneous product types from two brands by modeling consumer choice using the nested multinomial logit framework with two different hierarchical structures: a brand-primary model in which consumers choose a brand first and then a product type in the chosen brand and a type-primary model in which consumers choose a product type first and then a brand within that product type. Nasiry and Popescu [63] study the dynamic pricing implications of a new, behaviorally motivated reference price mechanism based on the peak-end memory mode, which suggests that consumers anchor on a reference price that is a weighted average of the lowest and most recent prices. They find that a range of constant pricing policies is optimal for the corresponding dynamic pricing problem. Nasiry and Popescu [64] further characterize the effect of anticipated regret on consumer decisions and on firm profits and policies in an advance selling context where buyers have uncertain valuations. Tereyağoğlu and Veeraraghavan [65] propose a model that addresses pricing and production decisions for a firm, using the rational expectations framework. They show that firms may offer high availability of goods despite the presence of conspicuous consumption and scarcity strategies are harder to adopt as demand variability increases. Parlaktürk [66] considers a firm that sells two vertically (quality) differentiated products to strategically forward-looking consumers over two periods, setting the prices dynamically in each period. It is found that the loss due to strategic customer behavior can be less with two product variants compared to the single-product benchmark, which indicates that product variety can serve as a lever when dealing with strategic customers. Cachon and Swinney [67] consider a retailer that sells a product with uncertain demand over a finite selling season, with three types of consumers: myopic, bargain-hunting, and strategic consumers. They find that the retailer stocks less, takes smaller price discounts, and earns lower profit if strategic consumers are present than if there are no strategic consumers, and a retailer should generally avoid committing to a price path over the season.

Another stream of research focuses on the risk attitude of the firms in the supply chain. Lau's [68] might be the first piece of work that studies the newsvendor boy problem under mean-variance framework, which takes the variance of system profit or cost into the utility function. Other recent works employing similar methodology to investigate supply chain problem include H. S. Lau and A. H. L. Lau [69] on supply chain model with return policy, Buzacott et al. [70] on the commitment-option contracts, Choi et al. [71] on channel coordination, and Wei and Choi [72] on wholesale pricing and profit sharing scheme.

4. Insights and Future Directions

From the above analysis, we can absorb the following insights and future directions in the area of operations research of logistics and supply chain management.

First, the logistics issue regarding the people's livelihood becomes a hot spot. The traditional research in this regard is related to perishable product, fashion product, and electronic product, which have short life cycle. Nowadays, such topics might include city logistics, emergency logistics, and agriculture supply chain.

Second, new directions on logistics and supply chain management can be brought about by the development of economy and technology. A typical example is the information technology which leads to the research on e-business and related distribution channel choice. Nowadays, the common usage of RFID, cloud technique, and big data can be important research directions for future study.

Third, the environmental related research will continue to be big issue. With the steady increase in global population and economic scale, resource crisis, ecological damage, environmental pollution, and other issues have drawn universal concern. It has been the consensus of the international community to attain socioeconomic sustainable development through a greener economic pattern and lifestyle. Many countries create a new outlook in industrial and technical competition by increasing investment in the green logistics and supply chain field, formulating and implementing various bills, plans, and strategies, and strengthening the implementation of green economic development strategy. In the future, the range of this topic will not only be just remanufacturing, reverse logistics, and closed-loop supply chain. Low-carbon issues can be an important research direction.

Finally, multimethodology is an important direction for future study. Traditionally, major research methodologies in operations management can be classified into several categories, such as theoretical modeling, computation and simulations, surveys, cases, event studies, and behavioral experiments. In recent years, there is an emerging trend towards combining multiple research methodologies to explore research problems in logistics and supply chain management. For example, in addressing the issues of supply chain coordination, some papers establish the respective models and verify the findings by real-world cases and some papers conduct behavioral experiments with the goal of exploring the real-world relevance of some theoretical models. Moreover, the number of the papers with new applications of the existing methodology, such as cooperative game and behavior operations, is expected to grow continuously.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

The author gratefully acknowledges the support by the Fundamental Research Funds for the Central Universities, no. NKZXB1228.

References

[1] Y. K. Kathawala and K. Abdou, "Supply chain evaluation in the service industry: a framework development compared to

manufacturing," *Managerial Auditing Journal*, vol. 18, no. 2, pp. 140–149, 2003.

[2] J. T. Mentzer, W. DeWitt, J. S. Keebler et al., "Defining supply chain management," *Journal of Business Logistics*, vol. 22, no. 2, pp. 1–25, 2001.

[3] R. Lamming and J. Hampson, "The environment as a supply chain management issue," *The British Journal of Management*, vol. 7, supplement 1, pp. S45–S62, 1996.

[4] M. J. Saunders, *Strategic Purchasing and Supply Chain Management*, Pitman, London, UK, 1997.

[5] M. J. Saunders, "Making strategic decisions and actions in purchasing and supply chain management," in *Proceedings of the 6th International IPSERA Conference*, pp. T1/6 1–T1/6 9, Naples, Italy, 1997.

[6] S. Croom, P. Romano, and M. Giannakis, "Supply chain management: an analytical framework for critical literature review," *European Journal of Purchasing and Supply Management*, vol. 6, no. 1, pp. 67–83, 2000.

[7] T. M. Whitin, *Theory of Inventory Management*, Princeton University Press, Princeton, NJ, USA, 1957.

[8] S. Nahmias, "Perishable inventory theory: a review," *Operations Research*, vol. 30, no. 4, pp. 680–708, 1982.

[9] F. Raafat, "Survey of literature on continuously deteriorating inventory models," *Journal of the Operational Research Society*, vol. 42, no. 1, pp. 27–37, 1991.

[10] S. K. Goyal and B. C. Giri, "Recent trends in modeling of deteriorating inventory," *European Journal of Operational Research*, vol. 134, no. 1, pp. 1–16, 2001.

[11] J. Blackburn and G. Scudder, "Supply chain strategies for perishable products: the case of fresh produce," *Production and Operations Management*, vol. 18, no. 2, pp. 129–137, 2009.

[12] X. Q. Cai, J. Chen, Y. B. Xiao, and X. L. Xu, "Optimization and coordination of fresh product supply chains with freshness keeping effort," *Production and Operations Management*, vol. 19, no. 3, pp. 261–278, 2010.

[13] X. Q. Cai, J. Chen, Y. B. Xiao, X. L. Xu, and G. Yu, "Fresh-product supply chain management with logistics outsourcing," *Omega*, vol. 41, no. 4, pp. 752–765, 2013.

[14] T. S. Chang, "Best routes selection in international intermodal networks," *Computers and Operations Research*, vol. 35, no. 9, pp. 2877–2891, 2008.

[15] M. Verma and V. Verter, "A lead-time based approach for planning rail-truck intermodal transportation of dangerous goods," *European Journal of Operational Research*, vol. 202, no. 3, pp. 696–706, 2010.

[16] F. Bruns and S. Knust, "Optimized load planning of trains in intermodal transportation," *OR Spectrum*, vol. 34, no. 3, pp. 511–533, 2012.

[17] F. Bruns, M. Goerigk, S. Knust, and A. Schöbel, "Robust load planning of trains in intermodal transportation," *OR Spectrum*, vol. 1, pp. 1–38, 2013.

[18] J. García, J. E. Florez, Á. Torralba et al., "Combining linear programming and automated planning to solve intermodal transportation problems," *European Journal of Operational Research*, vol. 227, no. 1, pp. 216–226, 2013.

[19] R. Monczka, R. Trent, and R. Handfield, *Purchasing and Supply Chain Management*, Thomson South-Western, Mason, Mich, USA, 2005.

[20] J. Gray, B. Tomlin, and A. Roth, "The effect of learning and strategic behavior on manufacturing outsourcing decisions," Working Paper, Kenan-Flagler Business School, University of North Carolina, 2005.

[21] T. R. Lewis and H. Yildirim, "Managing dynamic competition," *The American Economic Review*, vol. 92, no. 4, pp. 779–797, 2002.

[22] F. Chen, "Auctioning supply contracts," *Management Science*, vol. 53, no. 10, pp. 1562–1576, 2007.

[23] Y. J. Chen and G. Vulcano, "Effects of information disclosure under first- and second-price auctions in a supply chain setting," *Manufacturing and Service Operations Management*, vol. 11, no. 2, pp. 299–316, 2009.

[24] W. T. Huh and G. Janakiraman, "Inventory management with auctions and other sales channels: optimality of (s, S) policies," *Management Science*, vol. 54, no. 1, pp. 139–150, 2008.

[25] K.-Y. Chen, M. Kaya, and Ö. Özer, "Dual sales channelmanagement with service competition," *Manufacturing and Service Operations Management*, vol. 10, no. 4, pp. 654–675, 2008.

[26] Y. J. Chen, S. Seshadri, and E. Zemel, "Sourcing through auctions and audits," *Production and Operations Management*, vol. 17, no. 2, pp. 121–138, 2008.

[27] B. Anderson, "GM's global supply footprint," in *Proceedings of the Supplier Industry in Transition: The New Geography of Auto Production*, 2006.

[28] J. Carbone, "What buyers look for in contract manufacturers," *Purchasing*, vol. 128, no. 4, pp. 32–38, 2000.

[29] H. Li and S. C. Graves, "Pricing decisions during intergenerational product transition," *Production and Operations Management*, vol. 21, no. 1, pp. 14–28, 2012.

[30] C. Li and F. Zhang, "Advance demand information, price discrimination, and preorder strategies," *Manufacturing and Service Operations Management*, vol. 15, no. 1, pp. 57–71, 2013.

[31] A. Sainathan, "Pricing and replenishment of competing perishable product variants under dynamic demand substitution," *Production and Operations Management*, vol. 22, no. 5, pp. 1157–1181, 2013.

[32] M. Sinitsyn, "Coordination of price promotions in complementary categories," *Management Science*, vol. 58, no. 11, pp. 2076–2094, 2012.

[33] B. Liu, G. Cai, and A. A. Tsay, "Advertising in asymmetric competing supply chains," *Production and Operations Management*.

[34] F. Gao, O. C. Demirag, and F. Y. Chen, "Early sales of seasonal products with weather-conditional rebates," *Production and Operations Management*, vol. 21, no. 4, pp. 778–794, 2012.

[35] E. Brynjolfsson, Y. Hu, and M. S. Rahman, "Battle of the retail channels: how product selection and geography drive cross-channel competition," *Management Science*, vol. 55, no. 11, pp. 1755–1765, 2009.

[36] L. Guo, "Quality disclosure formats in a distribution channel," *Management Science*, vol. 55, no. 9, pp. 1513–1526, 2009.

[37] W.-Y. K. Chiang, "Supply chain dynamics and channel efficiency in durable product pricing and distribution," *Manufacturing and Service Operations Management*, vol. 14, no. 2, pp. 327–343, 2012.

[38] V. P. Simpson, "Optimum solution structure for a repairable inventory problem," *Operations Research*, vol. 26, no. 2, pp. 270–281, 1978.

[39] P. Kelle and E. A. Silver, "Forecasting the returns of reusable containers," *Journal of Operations Management*, vol. 8, no. 1, pp. 17–35, 1989.

[40] K. Inderfurth, "Simple optimal replenishment and disposal policies for a product recovery system with leadtimes," *OR Spectrum*, vol. 19, no. 2, pp. 111–122, 1997.

[41] E. van der Laan, M. Salomon, R. Dekker, and L. van Wassenhove, "Inventory control in hybrid systems with remanufacturing," *Management Science*, vol. 45, no. 5, pp. 733–747, 1999.

[42] R. Teunter, E. van der Laan, and D. Vlachos, "Inventory strategies for systems with fast remanufacturing," *Journal of the Operational Research Society*, vol. 55, no. 5, pp. 475–484, 2004.

[43] J. Wang, J. Zhao, and X. Wang, "Optimum policy in hybrid manufacturing/remanufacturing system," *Computers and Industrial Engineering*, vol. 60, no. 3, pp. 411–419, 2011.

[44] G. P. Kiesmüller, "A new approach for controlling a hybrid stochastic manufacturing/remanufacturing system with inventories and different leadtimes," *European Journal of Operational Research*, vol. 147, no. 1, pp. 62–71, 2003.

[45] O. Tang and R. W. Grubbström, "Considering stochastic lead times in a manufacturing/remanufacturing system with deterministic demands and returns," *International Journal of Production Economics*, vol. 93-94, no. 1, pp. 285–300, 2005.

[46] N. Aras, V. Verter, and T. Boyaci, "Coordination and priority decisions in hybrid manufacturing/remanufacturing systems," *Production and Operations Management*, vol. 15, no. 4, pp. 528–543, 2006.

[47] M. Fleischmann, J. M. Bloemhof-Ruwaard, R. Dekker, E. van der Laan, J. A. E. E. van Nunen, and L. N. van Wassenhove, "Quantitative models for reverse logistics: a review," *European Journal of Operational Research*, vol. 103, no. 1, pp. 1–17, 1997.

[48] R. Dekker, M. Fleischmann, K. Inderfurth, and L. N. van Wassenhove, *Reverse Logistics: Quantitative Models for Closed-Loop Supply Chains*, Springer, New York, NY, USA, 2004.

[49] M. A. Ilgin and S. M. Gupta, "Environmentally conscious manufacturing and product recovery (ECMPRO): a review of the state of the art," *Journal of Environmental Management*, vol. 91, no. 3, pp. 563–591, 2010.

[50] K. Inderfurth, "Impact of uncertainties on recovery behavior in a remanufacturing environment: a numerical analysis," *International Journal of Physical Distribution and Logistics Management*, vol. 35, no. 5, pp. 318–336, 2005.

[51] K. Inderfurth and I. M. Langella, "Heuristics for solving disassemble-to-order problems with stochastic yields," *OR Spectrum*, vol. 28, no. 1, pp. 73–99, 2006.

[52] M. R. Galbreth and J. D. Blackburn, "Optimal acquisition and sorting policies for remanufacturing," *Production and Operations Management*, vol. 15, no. 3, pp. 384–392, 2006.

[53] M. E. Ketzenberg, E. A. van der Laan, and R. H. Teunter, "Value of information in closed loop supply chains," *Production and Operations Management*, vol. 15, no. 3, pp. 393–406, 2006.

[54] U. Çorbacioğlu and E. A. van der Laan, "Setting the holding cost rates in a two-product system with remanufacturing," *International Journal of Production Economics*, vol. 109, no. 1-2, pp. 185–194, 2007.

[55] C. Zikopoulos and G. Tagaras, "Impact of uncertainty in the quality of returns on the profitability of a single-period refurbishing operation," *European Journal of Operational Research*, vol. 182, no. 1, pp. 205–225, 2007.

[56] M. Denizel, M. Ferguson, and G. C. Souza, "Multiperiod remanufacturing planning with uncertain quality of inputs," *IEEE Transactions on Engineering Management*, vol. 57, no. 3, pp. 394–404, 2010.

[57] V. D. R. Guide Jr. and V. Jayaraman, "Product acquisition management: current industry practice and a proposed framework," *International Journal of Production Research*, vol. 38, no. 16, pp. 3779–3800, 2000.

[58] V. D. R. Guide Jr. and L. N. van Wassenhove, "Managing product returns for remanufacturing," *Production and Operations Management*, vol. 10, no. 2, pp. 142–155, 2001.

[59] V. D. R. Guide Jr., R. H. Teunter, and L. N. van Wassenhove, "Matching supply and demand to maximize profits from remanufacturing," *Manufacturing and Service Operations Management*, vol. 5, no. 4, pp. 303–316, 2003.

[60] I. S. Bakal and E. Akcali, "Effects of random yield in remanufacturing with price-sensitive supply and demand," *Production and Operations Management*, vol. 15, no. 3, pp. 407–420, 2006.

[61] I. Karakayali, H. Emir-Farinas, and E. Akcali, "An analysis of decentralized collection and processing of end-of-life products," *Journal of Operations Management*, vol. 25, no. 6, pp. 1161–1183, 2007.

[62] A. G. Kök and Y. Xu, "Optimal and competitive assortments with endogenous pricing under hierarchical consumer choice models," *Management Science*, vol. 57, no. 9, pp. 1546–1563, 2011.

[63] J. Nasiry and I. Popescu, "Dynamic pricing with loss-averse consumers and peak-end anchoring," *Operations Research*, vol. 59, no. 6, pp. 1361–1368, 2011.

[64] J. Nasiry and I. Popescu, "Advance selling when consumers regret," *Management Science*, vol. 58, no. 6, pp. 1160–1177, 2012.

[65] N. Tereyağoğlu and S. Veeraraghavan, "Selling to conspicuous consumers: pricing, production, and sourcing decisions," *Management Science*, vol. 58, no. 12, pp. 2168–2189, 2012.

[66] A. K. Parlaktürk, "The value of product variety when selling to strategic consumers," *Manufacturing and Service Operations Management*, vol. 14, no. 3, pp. 371–385, 2012.

[67] G. P. Cachon and R. Swinney, "Purchasing, pricing, and quick response in the presence of strategic consumers," *Management Science*, vol. 55, no. 3, pp. 497–511, 2009.

[68] H. S. Lau, "The newsboy problem under alternative optimization objectives," *Journal of the Operational Research Society*, vol. 31, no. 6, pp. 525–535, 1980.

[69] H. S. Lau and A. H. L. Lau, "Manufacturer's pricing strategy and return policy for a single-period commodity," *European Journal of Operational Research*, vol. 116, no. 2, pp. 291–304, 1999.

[70] J. Buzacott, H. Yan, and H. Zhang, "Risk analysis of commitment-option contracts with forecast updates," Working Paper, York University, Toronto, Canada, 2003.

[71] T. M. Choi, D. Li, H. Yan, and C.-H. Chiu, "Channel coordination in supply chains with agents having mean-variance objectives," *Omega*, vol. 36, no. 4, pp. 565–576, 2008.

[72] Y. Wei and T. M. Choi, "Mean-variance analysis of supply chains under wholesale pricing and profit sharing schemes," *European Journal of Operational Research*, vol. 204, no. 2, pp. 255–262, 2010.

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 323764, 10 pages
http://dx.doi.org/10.1155/2014/323764

Research Article

A Large-Scale Network Data Analysis via Sparse and Low Rank Reconstruction

Liang Fu Lu,[1] Zheng-Hai Huang,[1,2] Mohammed A. Ambusaidi,[3] and Kui-Xiang Gou[1]

[1] *Department of Mathematics, School of Science, Tianjin University, Tianjin 300072, China*
[2] *Center for Applied Mathematics of Tianjin University, Tianjin 300072, China*
[3] *Faculty of Engineering and IT, University of Technology, Sydney, NSW 2007, Australia*

Correspondence should be addressed to Kui-Xiang Gou; gkxiang@tju.edu.cn

Received 11 April 2014; Accepted 28 April 2014; Published 26 May 2014

Academic Editor: Xiang Li

Copyright © 2014 Liang Fu Lu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the rapid growth of data communications in size and complexity, the threat of malicious activities and computer crimes has increased accordingly as well. Thus, investigating efficient data processing techniques for network operation and management over large-scale network traffic is highly required. Some mathematical approaches on flow-level traffic data have been proposed due to the importance of analyzing the structure and situation of the network. Different from the state-of-the-art studies, we first propose a new decomposition model based on accelerated proximal gradient method for packet-level traffic data. In addition, we present the iterative scheme of the algorithm for network anomaly detection problem, which is termed as NAD-APG. Based on the approach, we carry out the intrusion detection for packet-level network traffic data no matter whether it is polluted by noise or not. Finally, we design a prototype system for network anomalies detection such as Probe and R2L attacks. The experiments have shown that our approach is effective in revealing the patterns of network traffic data and detecting attacks from large-scale network traffic. Moreover, the experiments have demonstrated the robustness of the algorithm as well even when the network traffic is polluted by the large volume anomalies and noise.

1. Introduction

The rapid growth of data communication through the Internet and World Wide Web has led to vast amounts of information available online. In addition, business and government organizations create large amounts of both structured and unstructured information which need to be processed, analyzed, and linked. The large-scale network data plays a popular and important role in network operation and management. Consequently, high-dimensional data and multivariate data are becoming commonplace as the number of applications increases, such as statistical and demographic computation and digital libraries. Though it can provide flexible and cost-saving IT solutions for the end users, it is much easier in causing a great deal of problems such as network and system security issues due to its sharing and centralizing computing resources.

In general, network managers consider that the packet-level and flow-level data constitute the traditional network traffic data. On one hand, the packet-level data analysis performs successfully in maintaining simple, scalable, highly available, and robust networks [1]. On the other hand, flow-level data analysis has become popular in recent studies because the data can describe the network-level status and behavior of communication networks from origin nodes to destination nodes (OD) effectively [2]. However, there is no absolute way to secure the data and data transformations in large-scale networking systems. The existing techniques and tools of securing a network system still rely heavily on human experiences. Most of them require human involvement in analyzing and detecting anomalies and intrusions. Moreover, the existing networked-data analysis techniques are mainly based on the complete data, which limits the application of them. Unveiling the anomalies is a crucial task, especially nowadays, as big data acquisition and storage become increasingly difficult with the increasing amount of data due to the sampling bandwidth and storage space constraints [3]. Some approaches have been achieved in flow-level network data [2, 4, 5]. They reconstructed all origin-destination flows via compressive sensing methods by leveraging the low intrinsic-dimensionality of OD flows and the sparse nature of anomalies. Meanwhile, due to the difficulties of collecting and processing the large-scale packet-level data, to the best of our knowledge, few researchers pay attention to analyzing the incomplete packet-level data for managing and controlling the whole network.

Intrusion detection systems are security managements systems developed to find inconsistency with expected patterns in network traffic data, which is termed as well in literature [3] as novelty detection, anomaly mining, and noising mining. They play an important role in detecting different types of network attacks including Denial of Service (DOS), surveillance and other probing (Probe), unauthorized access to local super user (root) privilege (U2R), and unauthorized access from a remote machine (R2L) attacks. Intrusion detection approaches can be categorized into two main categories: signature-based and anomaly-based detection. Signature-based or misuse-based detection systems detect on-going anomalies by looking for a match with any predefined attack signature [6]. Anomaly-based detection, on the other hand, makes an assumption that intruders' behaviors are different from that of normal network traffic. Therefore, any deviation from the normal flow can be considered as an attack [7].

To enhance the human perception and understanding of different types of network intrusions and attacks, and inspired by the literature [5], approaches on network traffic data analysis and network anomalies detection based on compressed sensing in big data are put forth in this paper. As pointed out in literature [2, 3], on one hand, the number of normal data instances is much more than the number of anomaly data ones, which exactly meets the sparsity requirements of compressed sensing theory. On the other hand, traffic matrices usually have low effective dimensions because they can be well approximated by a few principal components that correspond to the largest singular values of the matrices, which are introduced by Lakhina et al. [8] by using of Principal Component Analysis (PCA) method to traffic matrix analysis.

Therefore, at the first stage, we propose a new decomposition model for packet-level traffic matrix. Then, we present the iterative algorithm based on accelerated proximal gradient method for network anomaly detection problem, which is termed as NAD-APG. Based on the approach, we carry out the intrusion detection for network traffic data no matter whether it is polluted by noise or not. Finally, we design a prototype system for network anomalies detection such as Probe and R2L attacks and so on. The experiments have shown that our approach is effective in revealing the patterns of network traffic data and detecting attacks from large-scale network traffic. In addition, the experiments have demonstrated the robustness of the algorithm as well even when the network traffic is polluted by the large volume anomalies and noise.

The rest of the paper is organized as follows. Section 2 gives an overview of existing methods on structural analysis of network traffic via compressed sensing techniques. Section 3 presents our approach on anomaly detection in network traffic based on accelerated proximal gradient line research method (APGL). The experimental evaluation of our new approaches is explored in Section 4. Finally, conclusions and future work are presented in Section 5.

2. Related Work

It has become popular in recent studies that considering the traffic matrix analysis as the main flow-level data because the traffic matrix can describe the network-level status and behavior of communication networks from origin nodes to destination nodes (OD) effectively and it is a combination of different classes of network traffic to represent how much data is transmitted during different time intervals [8].

As one of the most widely used methods to analyze traffic matrix, PCA was put forth in [9] by Lakhina et al. They calculated the principal component that corresponds to the largest singular value of the matrix and utilized these principal components to approximate the original traffic matrix. Moreover, they improved this method and proposed volume anomaly detection approach based on PCA-subspace [8]. In the following approaches, researchers improved the classical PCA method and proposed distributed PCA [10], network anomography [11], and traffic matrix evaluation from adaptivity and bias perspectives [12]. However, as mentioned by the literatures [5, 8, 13], there are some limitations when we utilize the PCA method to deal with the traffic matrix in order to analyze and manage the whole network, such as the fluctuation of estimation error with the volume change of anomalies, the sensitivity to the choice of parameters, and failure on exploiting the sparsity of anomalies.

Therefore, due to the increasing complexity and amounts of internet applications, the acquisition and storage of big data becomes more and more difficult. Moreover, to overcome the limitations of PCA, researchers have obtained some approaches for analyzing end-to-end network traffic in recent years. To solve the problem that PCA performs poorly in polluted traffic matrix by large volume anomalies, Lakhina et al. [8] proposed structural analysis by decomposing the network traffic matrix into deterministic traffic, anomaly traffic, and the noise traffic matrix. They analyzed that the decomposition problem is equivalent to the relaxed principal component pursuit method. A distributed estimation method to unveil the anomalies presented in OD flows using proximal gradient method was proposed by Mardani et al. [5]. A centralized solver and the in-network processing of link-load measurements were analyzed in their work as well. While Nie et al. found that the size of OD flows obeys the power laws [4]. By using this characteristic and restricted isometric property in compressed sensing theory, they reconstructed all OD flows with the help of partial observed samples from backbone network traffic data.

However, all the current researches pay too much attention on the network traffic in flow-level network. Therefore, we propose a novel approach based on the latest method from compressed sensing to reveal the abnormal patterns by dealing with the packet-level network data. Firstly, we propose to apply the most popular accelerated proximal gradient line search method (APGL) [14] to recover the low rank matrices with network traffic data. Moreover, to get a more accurate and robust approximation to reconstruct traffic matrix, motivated partly by the literature [15], we propose a traffic matrix decomposition method based on the APGL algorithm. Finally, the simulation results and analysis describe the effectiveness and robustness of our approaches in network traffic data.

3. Overall Approach

3.1. Principal Component Analysis Method. Principal component analysis (PCA), as a widely used method in high dimensional data analysis, can be viewed as a coordinate transformation process which transforms the redundant data points to a low dimensional system. As pointed out in literature [2, 9], each row vector x_i of the traffic matrix $X \in \mathbb{R}^{m\times n}$ is considered as a data point. PCA is performed by calculating the principal component vectors of X, which are represented as $\{v_i, i = 1, 2, \ldots, n\}$. The first principal component vector enjoys the property of the maximum variance of the original matrix X. Similarly, the tth principal component vector v_t, $t = 2, 3, \ldots, m$ captures the maximum variance of the residual traffic matrix as follows:

$$v_t = \arg\max_{\|v\|=1} \left\| \left(X - \sum_{i=1}^{t-1} X v_i {v_i}^T \right) v \right\|. \tag{1}$$

Corresponding to v_i, we denote another unit vector u_i as

$$u_i = \frac{X v_i}{\|X v_i\|}, \quad i = 1, 2, 3, \ldots, m. \tag{2}$$

It is noted that vectors u_i and v_i ($i = 1, 2, 3, \ldots, m$) form the orthogonal basis of $\mathbb{R}^n$, respectively. Therefore, the traffic matrix can be decomposed by the following formula:

$$X = \sum_{i=1}^{m} \|X v_i\| u_i {v_i}^T. \tag{3}$$

If we denote $\sigma_i := \|X v_i\|$, the traditional PCA method can be recited by the famous singular value decomposition (SVD) method in the research field of matrix computation as follows:

$$X = \sum_{i=1}^{m} \sigma_i u_i {v_i}^T, \tag{4}$$

where v_i can be achieved by calculating the eigenvectors of the matrix $X^T X$, while $\sigma_i = \sqrt{\lambda_i} \cdot \lambda_i$ is the corresponding eigenvalue of the matrix $X^T X$. That is to say,

$$X^T X v_i = \lambda_i v_i := {\sigma_i}^2 v_i. \tag{5}$$

SVD plays an important role for its revealing, interesting, and attractive algebraic properties and conveys important geometrical and theoretical in-sights about transformations. The entries of each matrix obtained by the SVD algorithm have their special physical significances. According to the rationale of Eckart-Young theorem, $\sum_{i=1}^{r} \sigma_i u_i {v_i}^T$ $(1 \le r \le m)$ is considered to be the best rank-r approximation of X, that is, $\sum_{i=1}^{r} \sigma_i u_i {v_i}^T = \arg\min_{\mathrm{rank}(Y)\le r} \|X - Y\|_F$, where $\|\cdot\|_F$ denotes the Frobenius norm.

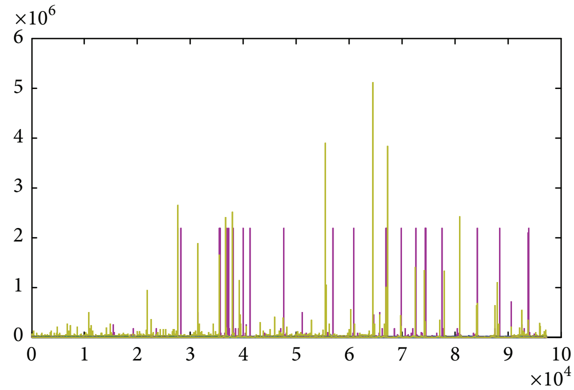


Figure 1: Visualization for original normal traffic with 97278 data items.

To the matrix expression for the PCA, it seeks an optimal estimate of A via the following constrained optimization:

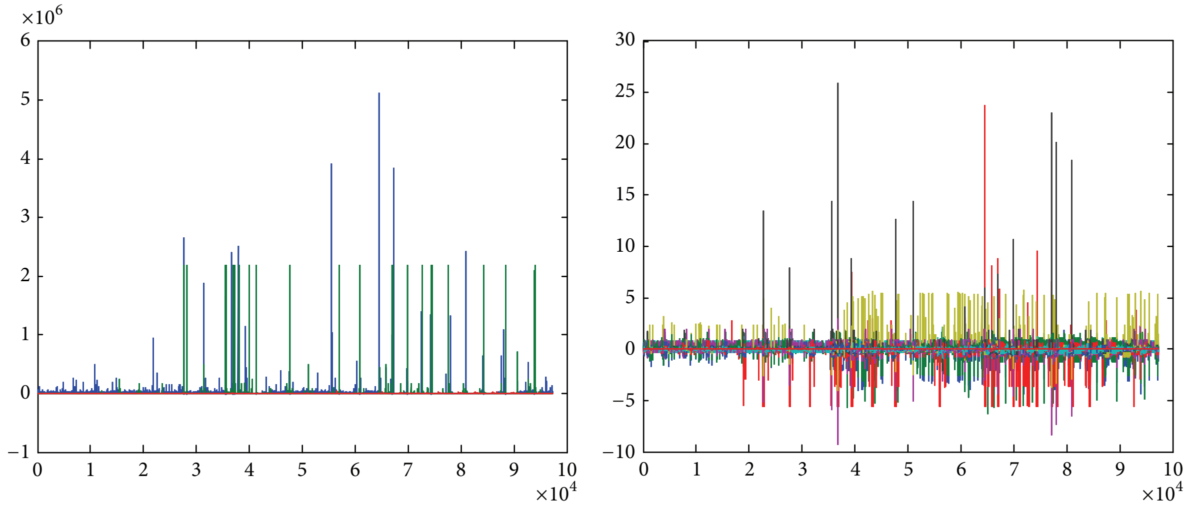
$$\begin{aligned} \min_{A,E} \quad & \|E\|_F, \\ \text{subject to} \quad & \mathrm{rank}(A) \le r, \\ & X = A + E, \end{aligned} \tag{6}$$

where $A, E \in \mathbb{R}^{m\times n}$, $r \ll \min(m, n)$. In fact, the optimal estimate of A is the projection of the columns of X onto the subspace spanned by the r principal left singular vectors of X [16].

3.2. Network Anomaly Detection Algorithm Based on APG. Though classical PCA method processes the data with the corruption of small Gaussian noise effectively, it always breaks down under large corruption [16]. Therefore, to recover a low-rank matrix A from a corrupted data matrix $X = A + E$, where some of the matrix E may be of arbitrarily large magnitude, Wright et al. [17] proposed a method termed as Robust PCA (RPCA) which can exactly recover the low-rank matrix in the presence of gross errors. Based on their analysis for the above optimization problem, the Lagrangian reformulation of it is

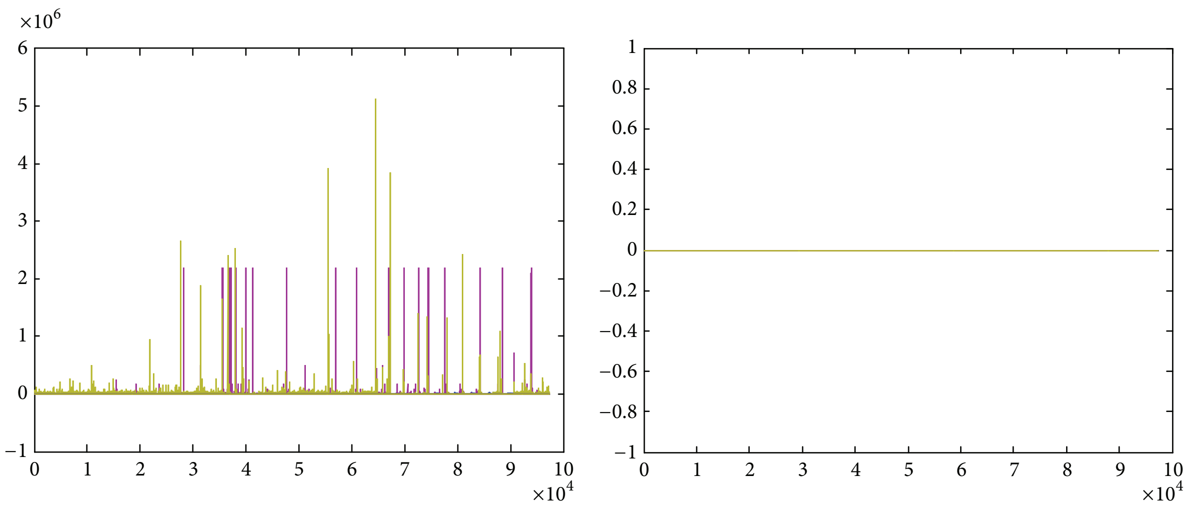
$$\begin{aligned} \min_{A,E} \quad & \mathrm{rank}(A) + \lambda \|E\|_0, \\ \text{subject to} \quad & X = A + E, \end{aligned} \tag{7}$$

where λ is a positive parameter that balances the two terms. Unfortunately, the above optimization problem is NP-hard in general due to the nonconvexity and discontinuous nature of the rank function. Moreover, the nuclear norm $\|\cdot\|_*$ (the sum of singular values of a matrix) is well known as a convex surrogate of the nonconvex matrix rank function. Therefore, in the literature [17], researchers proposed to solve



(a) (b)

FIGURE 2: (a) Principal components of original data in PCA. (b) Residual matrix visualization in PCA.



(a) (b)

FIGURE 3: (a) Normal traffic in NAD-APG method. (b) Visualization for abnormal traffic using NAD-APG.

the following convex optimization problem by replacing the l^0-norm with l^1-norm and $\text{rank}(A)$ with $\|A\|_*$:

$$\begin{aligned} \min_{A,E} \quad & \|A\|_* + \lambda\|E\|_1, \\ \text{subject to} \quad & X = A + E. \end{aligned} \tag{8}$$

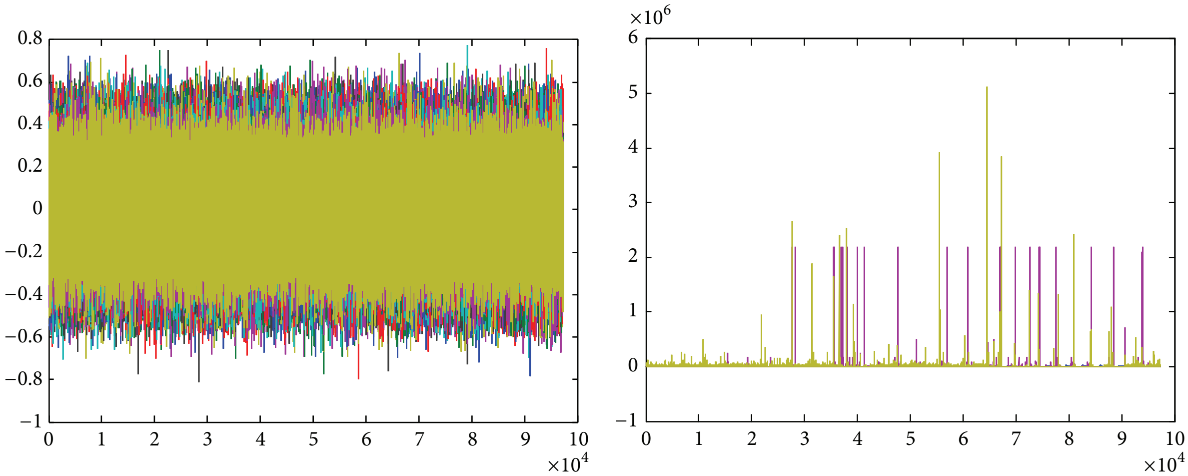
To develop faster and more scalable algorithms associated with the robust PCA, one popular method among state-of-art research approaches [16, 18–21] dubbed accelerated proximal gradient algorithm (APG) is widely exploited to seek an optimal solution of a soft constrained version of the convex problem (8). In this paper, we adopt the algorithm partially in the literature [16]. The main model is the following unstrained minimization optimization problem:

$$\min_{A,E} \mu\|A\|_* + \lambda\mu\|E\|_1 + \frac{1}{2}\|X - A - E\|_F^2. \tag{9}$$

Moreover, they summarized the convergence of the algorithm theoretically as follows.

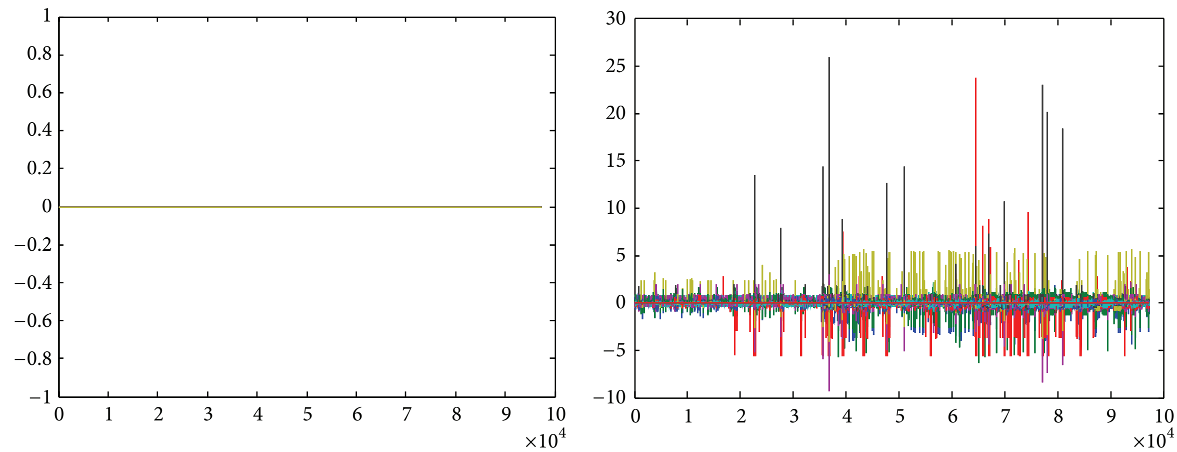
Theorem 1 (see [16]). *Suppose that* $F(A,E) = \mu\|A\|_* + \lambda\mu\|E\|_1 + (1/2)\|X - A - E\|_F^2$. *For all* $k > \log(\mu_0/\mu)/\log(1/\eta)$, *any solution* X^* *of the problem* (9), *we have* $F(X) - F(X^*) \leq 4\|X_{k_0} - X^*\|_F^2/(k - k_0 + 1)^2$.

The APG algorithm solves the optimization problem (9) by iteratively updating A, E, and other parameters. At last, in



(a) (b)

FIGURE 4: (a) Gaussian white noise matrix. (b) Anomaly-free traffic matrix with Gaussian white noise.



(a) (b)

FIGURE 5: (a) Abnormal traffic using NAD-APG. (b) Residual matrix with Gaussian white noise in PCA.

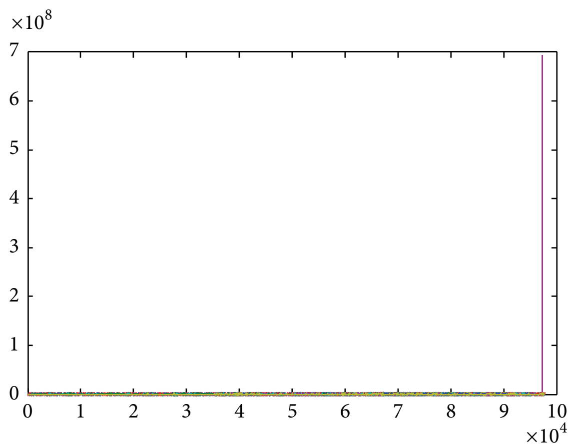
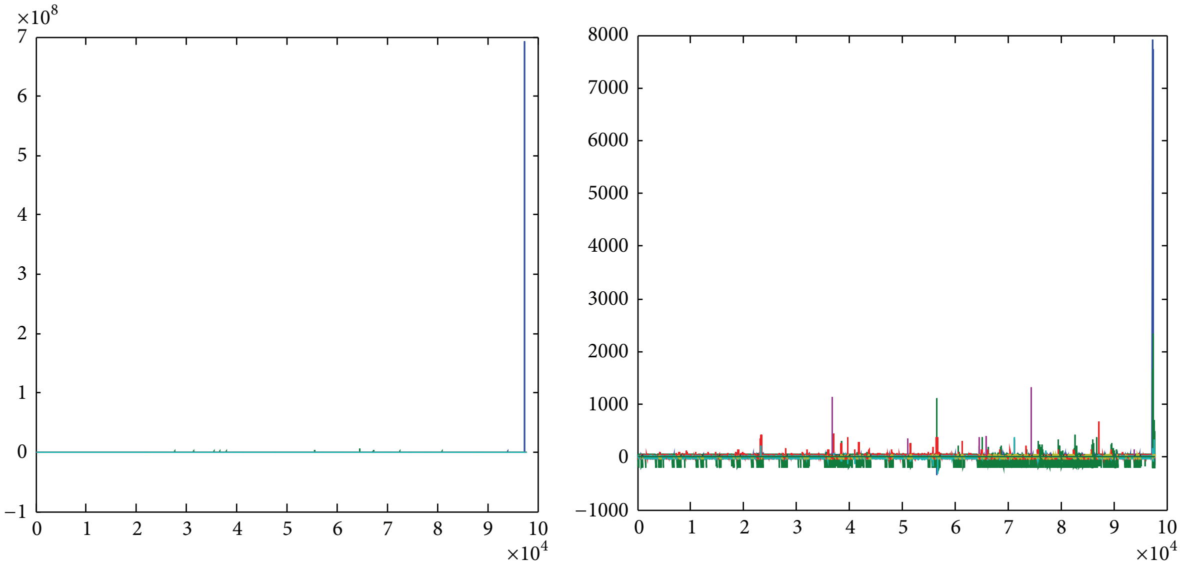


FIGURE 6: Original normal traffic mixed with 500 Probe attack data.

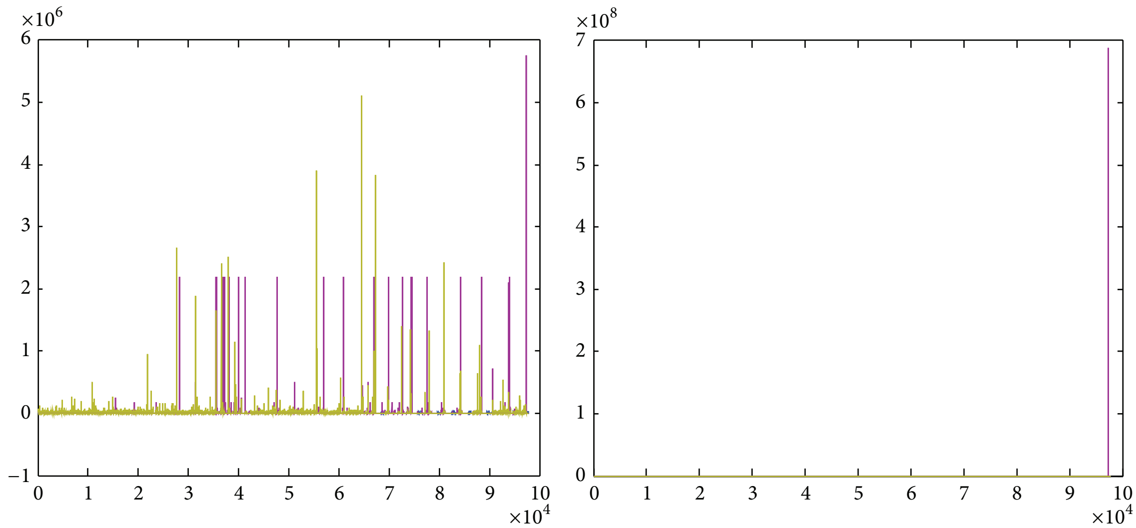
the kth iteration, we update A_{k+1} and E_{k+1} as the following iterative scheme:

$$A_{k+1} = \mathbb{S}_{\mu/2}\left(Y_k^A + \frac{X - Y_k^A - Y_k^E}{2}\right),$$
$$E_{k+1} = \mathbb{S}_{\mu/2}\left(Y_k^E + \frac{X - Y_k^A - Y_k^E}{2}\right),$$
$$\text{where } \mathbb{S}_{\mu/2}(x) = \begin{cases} x - \dfrac{\mu}{2}, & \text{if } x > \dfrac{\mu}{2}; \\ x + \dfrac{\mu}{2}, & \text{if } x < \dfrac{\mu}{2}; \\ 0, & \text{otherwise,} \end{cases} \tag{10}$$



(a) (b)

Figure 7: (a) Principal components of hybrid data in PCA. (b) Residual matrix visualization in PCA.



(a) (b)

Figure 8: (a) Normal traffic in NAD-APG method. (b) Abnormal patterns shown in NAD-APG.

Y^A_{k+1}, Y^E_{k+1} and t_{k+1} are updated in the same way as [16] as follows:

$$Y^A_{k+1} = A_{k+1} + \frac{t_k - 1}{t_{k+1}} \left(A_{k+1} - A_k\right),$$
$$Y^E_{k+1} = E_{k+1} + \frac{t_k - 1}{t_{k+1}} \left(E_{k+1} - E_k\right), \quad (11)$$
$$t_{k+1} = \frac{1 + \sqrt{1 + 4t_k^2}}{2}.$$

Here we summarize the main procedure for solving our network anomaly detection problem by APG algorithm, which is called NAD-APG (see Algorithm 1). As pointed out in [19], the algorithm has a convergence rate of $O(1/k^2)$.

4. Experiments and Results

In this section, we conduct several experiments on different attack types to show the effectiveness of our proposed approach.

4.1. The Data Set. Currently, there are only few public datasets for intrusion detection evaluation. According to the literature review by Tsai et al. [22], the majority of the IDS experiments

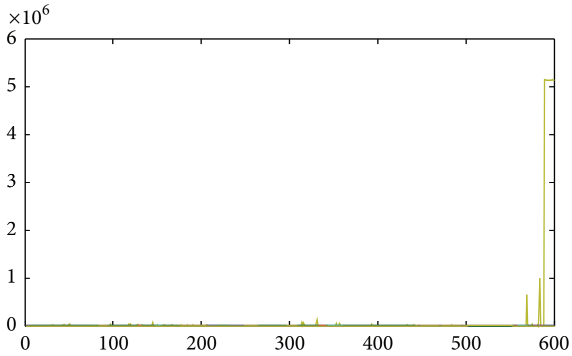


FIGURE 9: Visualization for 500 normal and 100 R2L network traffic.

are performed on the KDD Cup 99 datasets. It is the most comprehensive dataset that is still widely applied to compare and measure the performance of IDSs. Therefore, in order to facilitate fair and rational comparisons with other state-of-the-art detection approaches, we select the KDD Cup 99 dataset to evaluate the performance of our approach for detection. This dataset was derived from the DARPA 1998 datset. It contains training data with approximately five million connection records and test data with about two million connection records. Each record in this dataset is unique with 41 features. KDD Cup 99 dataset includes normal traffic and our different types of attacks, namely, Probe, Denial of Service (DOS), User o Root (U2R), and Remote to User (R2U). More details about these attacks are given by [23].

During our experiments, we use 10% KDD Cup 99 for training and testing. Literature review shows that a significant number of state-of-the-art IDSs, such as [23, 24], were evaluated using 10% KDD Cup 99 data. Therefore, training and testing our system on the 10% KDD Cup 99 data can help to provide a fair comparison with those approaches. The 10% KDD Cup 99 consists of 494,021 TCP/IP connection records simulated in a military network environment, US Air Force LAN. Each record is labeled as either normal or an attack, and it has 41 different quantitative and qualitative features. These features are generally categorized into three main groups. The first group is the basic features (i.e., attributes 1 to 9) that can be extracted from a TCP/IP connection. The second group refers to features 10 to 22 that are named as content-based features presenting the information derived from network packet payloads. The third group corresponds to the traffic-based features, which are carried by the features 23 to 41 of each record. A complete list of the set of features and the detailed description is available in [25].

4.2. Experimental Results. To demonstrate our method for managing the internet network, especially detecting the abnormal behaviors from the normal network traffic, we conduct PCA and NAD-APG methods for normal traffic mixed with some attacks in our experiments.

Firstly, we randomly choose some normal traffic data which consists of 97278 data items (see Figure 1). In this paper, we consider these data as the normally-free traffic to test the performance of our method. After using PCA, we can find that the principal components almost enjoy the whole property of the original normal traffic. Here, the sum of the variances of the first ten principal components is near 100% of the total variance of the original data. Figure 2 shows us the details of two matrices after decomposing the original matrix using PCA, where Figure 2(a) shows the visualization for matrix which is composed of top ten principal components and Figure 2(b) shows the details of residual matrix. However, it is very difficult for us to find any pattern from them. Moreover, there are still some data characteristics left in residual matrix after we apply PCA to the original normal data, which confuse the network analysts greatly. If we use our proposed method to the above normal traffic, two matrices with low-rank and sparse properties can be obtained to show the normal and abnormal traffic, respectively. Figure 3 visualizes the details of the two matrices, where Figure 3(a) represents the low-rank matrix with main properties of the original normal traffic. In our paper, we term this matrix as normal traffic matrix. While Figure 3(b) displays the matrix with all zero elements, which is termed as abnormal traffic matrix. Therefore, if we decompose the unique normal traffic into two matrices, we can obtain normal matrix uniquely. To sum up, Figure 3 shows the correctness and effectiveness of our proposed NAD-APG. Furthermore, if the traffic was polluted by noise (in this paper, we refer the noise to be Gaussian white noise with 0 mean and 1 variance), NAD-APG can identify the anomaly-free traffic accurately.

Figures 4 and 5 show the robust property of our proposed NAD-APG method, where (a) is the visualization of Gaussian white noise added to the normal traffic and (b) is the recovery of anomaly-free traffic. In fact, we obtained the all zero matrix as well in this decomposition, which means the traffic is anomaly-free. Figure 5 compares the effectiveness of NAD-APG method with PCA. In the visualization of residual matrix of PCA, it is obvious that we cannot find any pattern of attacks.

To evaluate the effectiveness of our method for detecting attacks in the whole internet, we add 500 "Probe" data items to the normal traffic, which means that the amount of the total data items is 97778 (see Figure 6). As we all know, Probe attacks refer to attackers that typically probe the victim's network or host by searching through the network or host for open ports before they launch an attack on a given host [26]. Therefore, there may be a large volume of traffic in a short time interval. Figure 6 displays the details of hybrid traffic. Firstly, we try to use PCA to detect the Probe attack. Figure 7(a) shows us the principal components of the hybrid traffic data which occupy the 99% contribution to the whole data. Though only 1% of energy of the whole data is left in the residual matrix, we find that there are still some intrinsic properties of the original data set which does not show any valuable pattern for attacks. However, if we test the hybrid data using NAD-APG method, the sparse traffic matrix obtained from the algorithm shows the attack pattern apparently. This enables the network manager to identify the anomalous packets from the normal traffic and can improve the accuracy of attacks detection. Figures 8(a) and 8(b) show the patterns of normal and abnormal traffic decomposed by the NAD-APG scheme.

Input: Network traffic matrix $X \in \mathbb{R}^{m\times n}$, λ and tolerance ε.
Initialize: $A_0 \leftarrow 0; E_0 \leftarrow 0; t_0 \leftarrow 1;$
Repeat
Step 1. Update A_{k+1}, E_{k+1} as $A_{k+1} = \mathbb{S}_{\mu/2}(Y_k^A + ((X - Y_k^A - Y_k^E)/2))$ and $E_{k+1} = \mathbb{S}_{\mu/2}(Y_k^E + ((X - Y_k^A - Y_k^E)/2))$.
Step 2. Let $t_{k+1} = (1 + \sqrt{1 + 4t_k^2})/2$.
Step 3. Update Y_{k+1}^A and Y_{k+1}^E: $Y_{k+1}^A = A_{k+1} + ((t_k - 1)/(t_{k+1}))(A_{k+1} - A_k)$, $Y_{k+1}^E = E_{k+1} + ((t_k - 1)/(t_{k+1}))(E_{k+1} - E_k)$
Until $\|A_{k+1} - A_k\| \le \varepsilon; \|E_{k+1} - E_k\| \le \varepsilon; A \leftarrow A_{k+1}, E \leftarrow E_{k+1}$.
Analyze: normal traffic pattern matrix A; abnormal traffic matrix E.
Output: Is there any abnormal activity or not in the whole traffic?

Algorithm 1: Anomaly detection using APG.

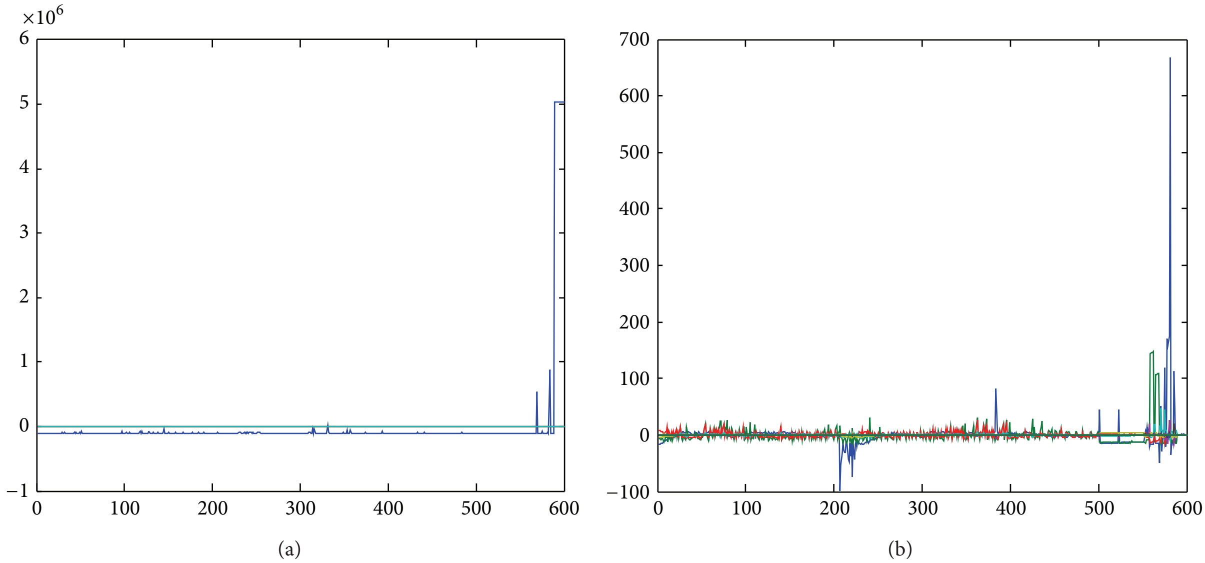


Figure 10: (a) Principal components of NR data in PCA. (b) Residual matrix of NR data in PCA.

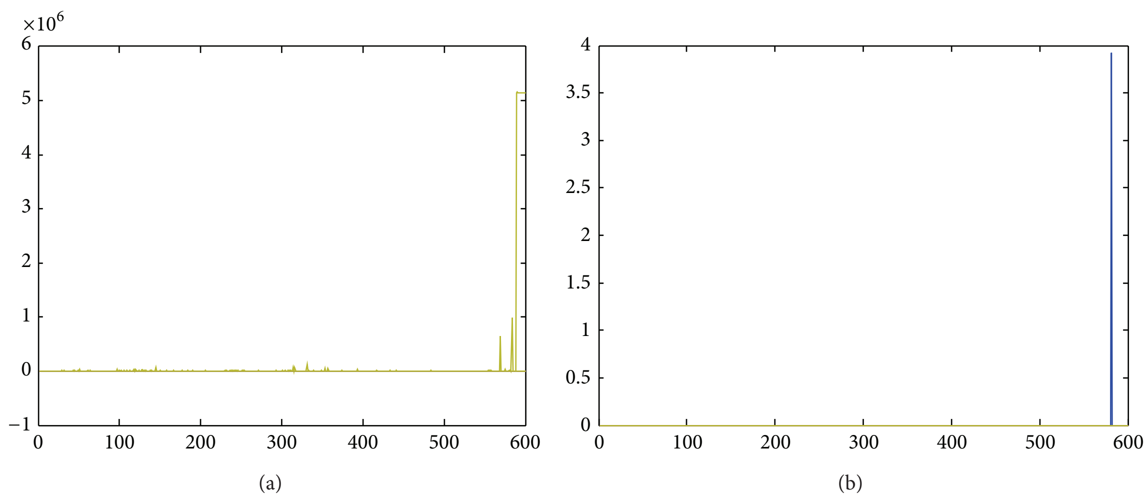


Figure 11: (a) Normal traffic of NR in NAD-APG method. (b) Abnormal patterns of NR in NAD-APG.

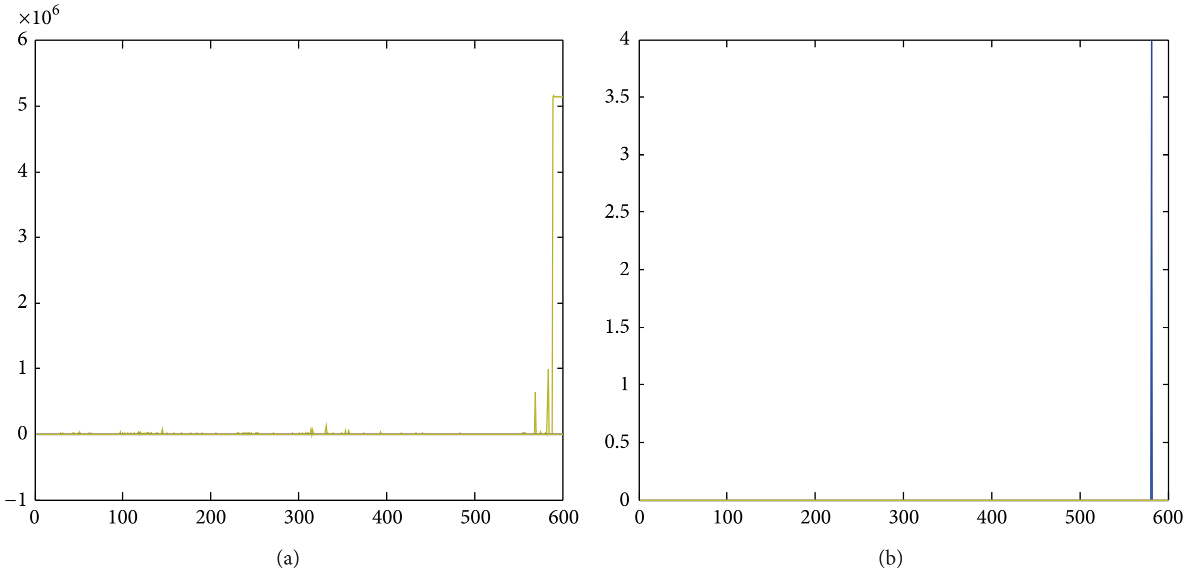


Figure 12: (a) Normal traffic of NRN in NAD-APG method. (b) Abnormal patterns of NRN in NAD-APG.

Table 1: Rank of low-rank matrix in the decomposition of NAD-APG method for attacks detection.

Different sampling data	Data items	Rank (sampling data)	Rank (low-rank matrix)
Normal traffic	97378	38	10
Normal traffic + Probe	97778	38	4
500 Normal traffic + 100 R2L traffic	600	36	4
Normal traffic + 100 R2L traffic	97478	38	4

To further demonstrate the advantages of our proposed scheme in small sampling data, we randomly choose 500 normal data items and 100 R2L attacks traffic as our test data set (here we term this data set as NR). R2L attacks always reveal the unauthorized local access from a remote machine. Moreover, the data values in R2L attacks are always much larger than the common data traffic. Therefore, we can find that the fluctuation in the scale of the data values occurs from Figure 9. Figures 10 and 11 represent the details of the normal and abnormal traffic matrices obtained from the two different decomposition algorithms, where the residual matrix in PCA is still confusing. Therefore, it is very difficult for network analysts to find the R2L attacks pattern from it. On the contrary, the attack pattern can apparently be found from the sparse matrix as shown in Figure 11(b). Even sometimes the real network traffic data are polluted by the noise (here we term this data set as NRN), especially the Gaussian white noise; the NAD-APG method can separate the normal and R2L attack patterns apparently. Figures 12(a) and 12(b) reveal the different patterns hidden in the network traffic, respectively.

To sum up, the experiments implemented above show that the low-rank matrix represents the normal traffic and the sparse matrix can always reveal the patterns of different attacks when we use our proposed NAD-APG scheme to detect anomalies. Table 1 describes the different ranks of the normal traffic matrix in detecting different attacks. There is no doubt that the low-rank matrices in processing different sampling data have much lower ranks than the original ones. However, the rank of low-rank matrix is ten when we deal with the pure normal network traffic, which may be caused by the pure type of data.

5. Conclusion

This paper introduced a new decomposition model for packet-level network traffic data no matter whether it was polluted by large-scale anomalies and noise or not. We presented the iterative algorithm based on accelerated proximal gradient method, which was termed as NAD-APG. Moreover, we designed a prototype system for network anomalies detection such as Probe and R2L attacks and so on. The experiments have shown that our approach is effective in revealing the patterns of network traffic data and detecting attacks from a variety of networking patterns. In addition, the experiments have demonstrated the robustness of the algorithm as well when the network traffic is polluted by the large volume anomalies and noise.

Though it is effective in detecting the attacks from the large-volume network traffic, it is difficult to classify the abnormal activities. Therefore, leveraging some feature selection and classification methods to our approaches to enhance the efficiency of intrusion detection is considered as our near future work. On the other hand, we will do more researches on APG algorithm itself and make our method more powerful and practical.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (no. 11171252 and no. 61202379) and the Doctoral Fund of Ministry of Education of China (no. 20120032120041).

References

[1] C. Fraleigh, S. Moon, B. Lyles et al., "Packet-level traffic measurements from the sprint IP backbone," *IEEE Network*, vol. 17, no. 6, pp. 6–16, 2003.

[2] Z. Wang, K. Hu, K. Xu, B. Yin, and X. Dong, "Structural analysis of network traffic matrix via relaxed principal component pursuit," *Computer Networks*, vol. 56, no. 7, pp. 2049–2067, 2012.

[3] W. Wang, D. Lu, X. Zhou, B. Zhang, and J. Mu, "Statistical wavelet-based anomaly detection in big data with compressive sensing," *EURASIP Journal on Wireless Communications & Networking*, vol. 2013, no. 1, article 269, pp. 1–6, 2013.

[4] L. Nie, D. Jiang, and L. Guo, "A power laws-based reconstruction approach to end-to-end network traffic," *Journal of Network and Computer Applications*, vol. 36, no. 2, pp. 898–907, 2012.

[5] M. Mardani, G. Mateos, and G. B. Giannakis, "Unveiling anomalies in large-scale networks via sparsity and low rank," in *Proceedings of the Conference Record of the 45th Asilomar Conference on Signals, Systems and Computers (ASILOMAR '11)*, pp. 403–407, November 2011.

[6] G. Vigna and R. A. Kemmerer, "NetSTAT: a network-based intrusion detection approach," in *Proceedings of the 14th Annual Computer Security Applications Conference*, pp. 25–34, 1998.

[7] A. Hassanzadeh and B. Sadeghian, "Intrusion detection with data correlation relation graph," in *Proceedings of the 3rd International Conference on Availability, Security, and Reliability (ARES '08)*, pp. 982–989, March 2008.

[8] A. Lakhina, K. Papagiannaki, M. Crovella, C. Diot, E. D. Kolaczyk, and N. Taft, "Structural analysis of network traffic flows," *ACM SIGMETRICS Performance Evaluation Review*, vol. 32, no. 1, pp. 61–72, 2004.

[9] A. Lakhina, M. Crovella, and C. Diot, "Diagnosing network-wide traffic anomalies," *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 4, pp. 219–230, 2004.

[10] L. Huang, M. I. Jordan, A. Joseph, M. Garofalakis, and N. Taft, "In-network PCA and anomaly detection," in *Advances in Neural Information Processing Systems*, pp. 617–624, 2006.

[11] Y. Zhang, Z. Ge, A. Greenberg, and M. Roughan, "Network anomography," in *Proceedings of the 5th ACM SIGCOMM Conference on Internet Measurement (IMC '05)*, p. 30, 2005.

[12] A. Soule, A. Lakhina, N. Taft et al., "Traffic matrices: balancing measurements, inference and modeling," *ACM SIGMETRICS Performance Evaluation Review*, vol. 33, no. 1, pp. 362–373, 2005.

[13] H. Ringberg, A. Soule, J. Rexford, and C. Diot, "Sensitivity of PCA for traffic anomaly detection," *ACM SIGMETRICS Performance Evaluation Review*, vol. 35, no. 1, pp. 109–120, 2007.

[14] K.-C. Toh and S. Yun, "An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems," *Pacific Journal of Optimization*, vol. 6, no. 3, pp. 615–640, 2010.

[15] H. Yao, D. Zhang, J. Ye, X. Li, and X. He, "Fast and accurate matrix completion via truncated nuclear norm regularization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 9, pp. 2117–2130, 2013.

[16] A. Ganesh, Z. Lin, J. Wright, L. Wu, M. Chen, and Y. Ma, "Fast algorithms for recovering a corrupted low-rank matrix," in *Proceedings of the 2009 3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP '09)*, pp. 213–216, December 2009.

[17] J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma, "Robust principal component analysis: exact recovery of corrupted low-rank matrices by convex optimization," in *Proceedings of the Neural Information Processing Systems*, 2009.

[18] R. He, W. S. Zheng, T. Tan, and Z. Sun, "Half-quadratic based iterative minimization for robust sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 2, pp. 261–275, 2014.

[19] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[20] Y.-F. Li, Y.-J. Zhang, and Z.-H. Huang, "A reweighted nuclear norm minimization algorithm for low rank matrix recovery," *Journal of Computational and Applied Mathematics*, vol. 263, pp. 338–350, 2014.

[21] M. Zhang, Z.-H. Huang, and Y. Zhang, "Restricted p-isometry properties of nonconvex matrix recovery," *IEEE Transactions on Information Theory*, vol. 59, no. 7, pp. 4316–4323, 2013.

[22] C.-F. Tsai, Y.-F. Hsu, C.-Y. Lin, and W.-Y. Lin, "Intrusion detection by machine learning: a review," *Expert Systems with Applications*, vol. 36, no. 10, pp. 11994–12000, 2009.

[23] S. Mukkamala, A. H. Sung, and A. Abraham, "Intrusion detection using an ensemble of intelligent paradigms," *Journal of Network and Computer Applications*, vol. 28, no. 2, pp. 167–182, 2005.

[24] A. Mitrokotsa, M. Tsagkaris, and C. Douligeris, "Intrusion detection in mobile Ad Hoc Networks using classification algorithms," *IFIP International Federation for Information Processing*, vol. 265, pp. 133–144, 2008.

[25] F. Amiri, M. Rezaei Yousefi, C. Lucas, A. Shakery, and N. Yazdani, "Mutual information-based feature selection for intrusion detection systems," *Journal of Network and Computer Applications*, vol. 34, no. 4, pp. 1184–1199, 2011.

[26] K. Labib and V. R. Vemuri, "Detecting and visualizing denial-of-service and network probe attacks using principal component analysis," in *Proceedings of 3rd Conference on Security and Network Architectures (SAR '04)*, La Londe, France, June 2004.

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 816980, 7 pages
http://dx.doi.org/10.1155/2014/816980

Research Article

An Approach for Simple Linear Profile Gauge R&R Studies

Liangxing Shi,[1] Wei Chen,[1] and Liang Fu Lu[2]

[1] *Department of Industrial Engineering, Tianjin University, Tianjin 300072, China*
[2] *Department of Mathematics, School of Science, Tianjin University, Tianjin 300072, China*

Correspondence should be addressed to Liang Fu Lu; liangfulv@gmail.com

Received 11 April 2014; Accepted 12 May 2014; Published 26 May 2014

Academic Editor: Xiang Li

Copyright © 2014 Liangxing Shi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Gauge repeatability and reproducibility studies are significant to quality improvement and quality control. The approaches are always applied to determine the capability of the measurement system. Much of the literature in this field mainly focuses on univariate and multivariate measurement systems. However, the state-of-the-art methods are not appropriate when the quality of a product is characterized by a profile. Therefore, this paper proposes a method for the measured values which can be characterized by a simple linear profile. In addition, the slopes and intercepts of these profiles often vary due to measurement error. Thus, the simple linear profile gauge studies can be considered as a two-response (slope and intercept) problem. X-values transformation is used to make the slope and intercept of each profile independent. ANOVA is utilized to estimate the variance component of measurement error and other sources of variation. Then, the criteria precision to tolerance ratio and percent R&R are introduced to assess the simple linear profile measurement system capability. Finally, the proposed approach is applied to the spring length and elasticity measurement which demonstrates how to implement the method.

1. Introduction

Gauge repeatability and reproducibility (GR&R) studies are usually adopted for determining the capability of the measurement systems. Burdick et al. [1] and Montgomery [2] proposed that one of the objectives of measurement systems capability analysis (MSCA) is to determine whether the gauge is capable. Evaluating the capability of measurement system is necessary for other researches, such as process capability analysis, control charts, and design of experiment. In practice, a reliable measurement system is significant to quality improvement and quality control.

Burdick et al. [1] reviewed the measurement system capability analysis. In their paper, the criteria, typical model, and GR&R experiment were all introduced. Attribute data are widely applied in industry; Lyu and Chen [3] evaluated the R&R of a measurement system for attribute data based on the generalized linear models (GLMs). Some papers considered confidence intervals for gauge capability studies. Borror et al. [4] presented two methods to construct confidence intervals for variance components: one is the restricted maximum likelihood (REML) method and the other is the modified large sample (MLS) method. Gong et al. [5] proposed an approach by applying unweighted sums of squares method when confidence intervals for unbalanced two-factor gauge study were considered.

In practice, many products need several critical characteristics to describe their quality. For these cases, the univariate MSCA method may not be appropriate especially when the measured characteristics have some correlations. Thus, numerous studies in this field considered multivariate MSCA. An approach to analyze the two-dimensional GR&R of imbalance measurement was introduced by Sweeney [6]. He concluded that the variation can be underestimated via treating such data as one-response. Majeske [7] proposed the method and the criteria for multivariate MSCA by regarding the measured values as a vector and assuming that the measured values follow a multivariate normal distribution. MANOA is applied to estimate the variance-component matrices.

Gauge repeatability and reproducibility studies are widely used in practice. Li and Al-Refaie [8] utilized the define-measure-analyze-improve-control (DMAIC) procedure to improve the quality system involving measurements.

Erdmann et al. [9] gave an example of GR&R in a hospital, and the experiment is about an ear thermometer for temperature measurements.

The existing research proposed many methods for univariate and multivariate MSCA. However, the measured values in some situations can be characterized by profiles. The measurement for the elasticity and length of springs is a good example. In this case, the elasticity of spring is a simple linear profile of an independent variable: the length of spring. In this paper, an approach for the situation where the measured values can be expressed as a simple linear profile is proposed.

The remainder of this paper is organized as follows. A brief review of gauge R&R is presented in Section 2. Section 3 proposes the approach and criteria for simple linear profile gauge R&R. A case study is introduced in Section 4. The last section contains our concluding remarks.

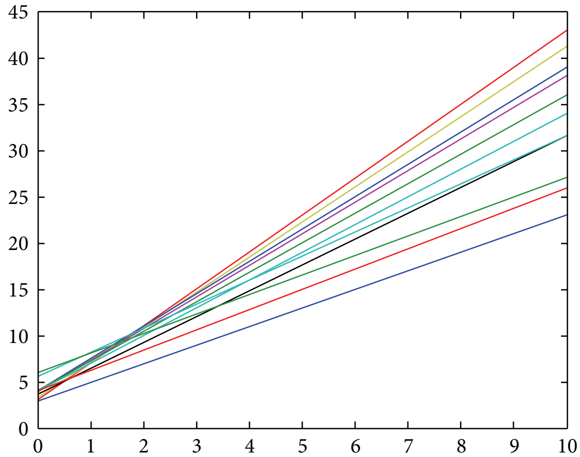


Figure 1: The differences of the profiles.

2. The Model and Criteria for GR&R

Typically, gauge studies are always designed with two factors, i parts and j operators, and each operator measures each part k times. The measured values are represented as M_{ijk}. These values are always treated as true values by the operators because they usually contain errors. The typical model with a two-factor design for variance analysis of M_{ijk} is as follows, see Burdick et al. [1] and Montgomery [2]:

$$M_{ijk} = \mu + p_i + o_j + (op)_{ij} + \varepsilon_{ijk}, \tag{1}$$

where $i = 1, 2, \ldots, p$, $j = 1, 2, \ldots, o$, and $k = 1, 2, \ldots, r$. μ is the true value. p_i, o_j, $(op)_{ij}$, and ε_{ijk} are all independent random variables which represent the effect of part, operator, operator-part interaction, and random error. Assume that they are all normally distributed, $p_i \sim N(0, \sigma_p^2)$, $o_j \sim N(0, \sigma_o^2)$, $(op)_{ij} \sim N(0, \sigma_{(op)}^2)$, and $\varepsilon_{ijk} \sim N(0, \sigma_e^2)$. The variance of M_{ijk} is given as follows:

$$\begin{aligned} \sigma_{M_{ijk}}^2 &= \sigma_p^2 + \sigma_o^2 + \sigma_{(op)}^2 + \sigma_e^2, \\ \sigma_G^2 &= \sigma_o^2 + \sigma_{(op)}^2 + \sigma_e^2, \end{aligned} \tag{2}$$

where σ_G^2 is the variance component which represents the effect of gauge.

Two criteria which are always utilized to assess the gauge capability are introduced in AIAG [10]. One is the ratio between the measurement precision estimate and the tolerance of the parts, $P/T\%$. The other is the ratio of the measurement precision estimate and the process width, R&R%. The criteria are formulated as follows:

$$\begin{aligned} \frac{P}{T}\% &= \frac{K\sigma_G}{\text{USL} - \text{LSL}} \times 100\%, \\ \text{R\&R}\% &= \frac{\sigma_G}{\sqrt{\sigma_G^2 + \sigma_p^2}} \times 100\%, \end{aligned} \tag{3}$$

where USL and LSL are upper specification limit and lower specification limit of the part. K is either 5.15 or 6. According to AIAG [10], the measurement system is unacceptable if the value exceeds 30%, and it is acceptable if the value is lower than 10%. If the value is between 10% and 30%, the measurement system needs to be improved. References considering these criteria include Wheeler and Lyday [11] and Woodall and Borror [12].

3. The GR&R Method for Simple Linear Profile

In practice, the quality of some products can be described by simple linear profiles when they have a collection of critical characteristics to be measured. In simple linear profile gauge study, a two-factor (i parts and j operators) design is considered. Each operator measures all of the parts k times. The measured values are characterized by a simple linear profile:

$$Y_{ijk} = A_{0(ijk)} + A_{1(ijk)} X_{ijk} + \varepsilon_{ijk}, \tag{4}$$

where $i = 1, 2, \ldots, p$, $j = 1, 2, \ldots, o$, and $k = 1, 2, \ldots, r$. $A_{0(ijk)}$ is the intercept, and $A_{1(ijk)}$ is the slope. The random variable ε_{ijk} is independent and normally distributed, $\varepsilon_{ijk} \sim (0, \sigma^2)$. Due to measurement error, the profiles always vary, and the variation is reflected by the slopes and the intercepts. As is shown in Figure 1, where one profile represents one time measurement, the slopes and intercepts of these profiles are different.

Thus, when considering measurement error, (4) is given as follows:

$$\begin{aligned} Y_{ijk} = {} & \left[A_0 + p_{iA_0} + o_{jA_0} + (op)_{ijA_0} + \varepsilon_{ijkA_0}\right] \\ & + \left[A_1 + p_{iA_1} + o_{jA_1} + (op)_{ijA_1} + \varepsilon_{ijkA_1}\right] X_{ijk} + \varepsilon_{ijk}, \end{aligned} \tag{5}$$

where A_0 and A_1 are constants. p_{iA_0}, o_{jA_0}, $(op)_{ijA_0}$, and ε_{ijkA_0} are independent random variables which represent the effect of part, operator, operator-part interaction, and the random error on $A_{0(ijk)}$, respectively. p_{iA_1}, o_{jA_1}, $(op)_{ijA_1}$, and ε_{ijkA_1} are independent random variables which represent the effect of part, operator, operator-part interaction, and the random error on $A_{1(ijk)}$, respectively.

3.1. The Transforming Model. For simple linear profile gauge study, $A_{0(ijk)}$ and $A_{1(ijk)}$ can be obtained easily when the measured results are profiles directly. However, it is necessary to estimate the two parameters in some cases. The least-square method, see Johnson and Wichern [13], is used to estimate parameters and fit profiles. The fitted profile is the following:

$$\widehat{Y}_{ijk} = \widehat{A}_{0(ijk)} + \widehat{A}_{1(ijk)} X_{ijk}, \quad (6)$$

where $\widehat{Y}_{ijk}$ is the fitted value. $\widehat{A}_{0(ijk)}$ and $\widehat{A}_{1(ijk)}$ are the least-square estimates of $A_{0(ijk)}$ and $A_{1(ijk)}$ which are normally distributed. $\widehat{A}_{0(ijk)}$ and $\widehat{A}_{1(ijk)}$ are correlative, and the covariance between them is σ_{01}^2.

Kim et al. [14] introduced a method which can make the intercept and the slope of each profile independent via transforming the X-values, so that the model can be simplified. An alternative form of the model in (6) is the following:

$$Y_{ijk} = B_{0(ijk)} + B_{1(ijk)} X_{ijk}^* + \varepsilon_{ijk}, \quad (7)$$

where $X_{ijk}^* = X_{ijk} - \overline{X}$, $B_{0(ijk)} = A_{0(ijk)} + A_{1(ijk)}\overline{X}$, and $B_{1(ijk)} = A_{1(ijk)}$. In this situation, the least-squares estimator of $B_{0(ijk)}$ is $\widehat{B}_{0(ijk)} = \widehat{A}_{0(ijk)} + \widehat{A}_{1(ijk)}\overline{X}$, and the least-squares estimator of $B_{1(ijk)}$ is $\widehat{B}_{1(ijk)} = \widehat{A}_{1(ijk)}$. Both of $\widehat{B}_{0(ijk)}$ and $\widehat{B}_{1(ijk)}$ are normally distributed, and the covariance between them is zero, so $\widehat{B}_{0(ijk)}$ and $\widehat{B}_{1(ijk)}$ for each profile are independent. A separate gauge R&R analysis can be applied to $\widehat{B}_{0(ijk)}$ and $\widehat{B}_{1(ijk)}$, respectively.

3.2. The Model for Variance Analysis. In the alternative form, the measurement error is reflected by $\widehat{B}_{0(ijk)}$ and $\widehat{B}_{1(ijk)}$. When a two-factor design experiment is considered, the model for variance analysis of measurement error on $\widehat{B}_{0(ijk)}$ and $\widehat{B}_{1(ijk)}$ is the following:

$$\begin{aligned} \widehat{B}_{0(ijk)} &= \mu_0 + p_{iB_0} + o_{jB_0} + (op)_{ijB_0} + \varepsilon_{ijkB_0}, \\ \widehat{B}_{1(ijk)} &= \mu_1 + p_{iB_1} + o_{jB_1} + (op)_{ijB_1} + \varepsilon_{ijkB_1}, \end{aligned} \quad (8)$$

where μ_0 and μ_1 are constants. p_{iB_0} and p_{iB_1} are independent random variables which represent the part effect on $\widehat{B}_{0(ijk)}$ and $\widehat{B}_{1(ijk)}$, respectively. o_{jB_0} and o_{jB_1} are independent random variables which represent the operator effect on $\widehat{B}_{0(ijk)}$ and $\widehat{B}_{1(ijk)}$, respectively. $(op)_{ijB_0}$ and $(op)_{ijB_1}$ are independent random variables which represent the operator-part interaction effect on $\widehat{B}_{0(ijk)}$ and $\widehat{B}_{1(ijk)}$, respectively. ε_{ijkB_0} and ε_{ijkB_1} are independent random variables which represent the random error effect on $\widehat{B}_{0(ijk)}$ and $\widehat{B}_{1(ijk)}$, respectively. Assume that p_{iB_0}, o_{jB_0}, $(op)_{ijB_0}$, ε_{ijkB_0}, p_{iB_1}, o_{jB_1}, $(op)_{ijB_1}$, and ε_{ijkB_1} are all normally distributed, where $p_{iB_0} \sim N(0, \sigma_{p0}^2)$, $o_{jB_0} \sim N(0, \sigma_{o0}^2)$, $(op)_{ijB_0} \sim N(0, \sigma_{op0}^2)$, $\varepsilon_{ijkB_0} \sim N(0, \sigma_0^2)$, $p_{iB_1} \sim N(0, \sigma_{p1}^2)$, $o_{jB_1} \sim N(0, \sigma_{o1}^2)$, $(op)_{ijB_1} \sim N(0, \sigma_{op1}^2)$, and $\varepsilon_{ijkB_1} \sim N(0, \sigma_1^2)$.

The variance of $\widehat{B}_{0(ijk)}$ considering the measurement error is

$$V\left[\widehat{B}_{0(ijk)}\right] = \sigma_{p0}^2 + \sigma_{o0}^2 + \sigma_{op0}^2 + \sigma_0^2. \quad (9)$$

The variance component of $\widehat{B}_{0(ijk)}$ for the gauge is

$$\sigma_{G0}^2 = \sigma_{o0}^2 + \sigma_{op0}^2 + \sigma_0^2. \quad (10)$$

The variance component of $\widehat{B}_{0(ijk)}$ for the measured values is

$$\sigma_{M0}^2 = \sigma_{p0}^2 + \sigma_{G0}^2. \quad (11)$$

The variance of $\widehat{B}_{1(ijk)}$ considering the measurement error is

$$V\left[\widehat{B}_{1(ijk)}\right] = \sigma_{p1}^2 + \sigma_{o1}^2 + \sigma_{op1}^2 + \sigma_1^2. \quad (12)$$

The variance component of $\widehat{B}_{1(ijk)}$ for the gauge is

$$\sigma_{G1}^2 = \sigma_{o1}^2 + \sigma_{op1}^2 + \sigma_1^2. \quad (13)$$

The variance component of $\widehat{B}_{1(ijk)}$ for the measured values is

$$\sigma_{M1}^2 = \sigma_{p1}^2 + \sigma_{G1}^2. \quad (14)$$

A two-factor ANOVA with an interaction term is applied to estimate these variance components.

3.3. The Assessment Criteria. Two responses ($\widehat{B}_{0(ijk)}$ and $\widehat{B}_{1(ijk)}$) are obtained in the gauge study for simple linear profile. The univariate criteria $P/T\%$ and R&R% can be calculated for $\widehat{B}_{0(ijk)}$ and $\widehat{B}_{1(ijk)}$, respectively. Four equations are utilized to assess the gauge capability for simple linear profile:

$$\begin{aligned} \frac{P}{T_0}\% &= \frac{K\sigma_{G0}}{\mathrm{USL}_0 - \mathrm{LSL}_0} \times 100\%, \\ \mathrm{R\&R}_0\% &= \frac{\sigma_{G0}}{\sqrt{\sigma_{G0}^2 + \sigma_{p0}^2}} \times 100\%, \\ \frac{P}{T_1}\% &= \frac{K\sigma_{G1}}{\mathrm{USL}_1 - \mathrm{LSL}_1} \times 100\%, \\ \mathrm{R\&R}_1\% &= \frac{\sigma_{G1}}{\sqrt{\sigma_{G1}^2 + \sigma_{p1}^2}} \times 100\%. \end{aligned} \quad (15)$$

According to AIAG [10], the measurement system is unacceptable if one of the above assessed values exceeds 30%. It is acceptable if the values are all lower than 10%. If the values are all between 10% and 30%, the measurement system needs to be improved.

Table 1: Profile variance component of $\widehat{B}_{0(ijk)}$.

σ_{p0}^2	0.007894
σ_{o0}^2	0.000009
σ_{op0}^2	0
σ_0^2	0.000149

Table 2: Profile variance component of $\widehat{B}_{1(ijk)}$.

σ_{p1}^2	0.000554
σ_{o1}^2	0.000001
σ_{op1}^2	0.000002
σ_1^2	0.000010

4. The Introductory Case

An example of gauge study about spring measurement is introduced in this section. In the quality improvement, it is necessary to measure the elasticity and the length of spring. According to Hooke's law, when the spring has reached a state of equilibrium, its elasticity is a simple linear profile of the amount by which the free end of the spring is displaced from its relaxed position (when it is not stretched). In this case, three operators made three measurements on each of ten springs, X_{ijk} is the length of spring after compression or elongation, and Y_{ijk} is the different elasticity when the spring is of different length.

The least-square method is used to fit profile and estimate the parameters $\widehat{A}_{0(ijk)}$ and $\widehat{A}_{1(ijk)}$ after the measurement. Then, $\widehat{B}_{0(ijk)}$ and $\widehat{B}_{1(ijk)}$ in (7) can be obtained via transforming X-values. A two-factor ANOVA with an interaction term is utilized to estimate the variance components of $\widehat{B}_{0(ijk)}$ and $\widehat{B}_{1(ijk)}$, respectively. The variance components of $\widehat{B}_{0(ijk)}$ are shown in Table 1. The variance components of $\widehat{B}_{1(ijk)}$ are shown in Table 2.

According to (10), (11), (13), and (14), the variance components for the gauge and the measured values of $\widehat{B}_{0(ijk)}$ and $\widehat{B}_{1(ijk)}$ are the following:

$$\begin{aligned} \sigma_{G0}^2 &= 0.000158, \\ \sigma_{M0}^2 &= 0.000090, \\ \sigma_{G1}^2 &= 0.000013, \\ \sigma_{M1}^2 &= 0.000567. \end{aligned} \tag{16}$$

The upper and lower specifications on $\widehat{B}_{0(ijk)}$ are $\mathrm{USL}_0 = 1.3283$ and $\mathrm{LSL}_0 = 0.9717$. Then, the criteria for $\widehat{B}_{0(ijk)}$ are the following:

$$\begin{aligned} \frac{P}{T_0}\% &= 21.18\%, \\ \mathrm{R\&R}_0\% &= 14.03\%. \end{aligned} \tag{17}$$

The upper and lower specifications on $\widehat{B}_{1(ijk)}$ are $\mathrm{USL}_1 = -0.2464$ and $\mathrm{LSL}_1 = -0.3223$. Then, the criteria for $\widehat{B}_{1(ijk)}$ are estimated as follows:

$$\begin{aligned} \frac{P}{T_1}\% &= 28.30\%, \\ \mathrm{R\&R}_1\% &= 15.16\%. \end{aligned} \tag{18}$$

In this case, all of the criteria are between 10% and 30%. According to AIAG [10], the measurement system needs to be improved. The improvement will be made based on the gauge R&R experiment. In this case, the approach used to improve the capability of measurement system provides a clearer instruction. The instruction includes the precise location where the spring should be placed and the measurement operations that should be conducted. This improvement can decrease the differences among the operators.

5. Conclusion

Gauge repeatability and reproducibility studies are important to guarantee the validity of data, which is essential to other researches. Numerous existing studies in this area are about univariate and multivariate measurement systems, but these methods may not be suitable when the quality of a product should be characterized by a profile. This paper proposes an approach to assess the gauge capability when a simple linear profile is used to reflect product quality. Our proposed method can simplify the measurement problems effectively, especially for those with the multi-dimensional measured values. The example of spring measurement is presented in this paper which shows how to implement the proposed method. Further, more and more methods and criteria should be proposed to assess the capability of the measurement system.

Appendix

In this appendix, the original measured values of the introductory case are provided as Table 3.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Table 3: The measured data of spring case.

Part	Operator	X_{ijk} 17	16	15	13.5	12.5	11
		Y_{ijk}					
1	1	0.26	0.59	0.86	1.37	1.68	2.19
1	1	0.24	0.57	0.86	1.37	1.66	2.17
1	1	0.24	0.55	0.86	1.35	1.66	2.17
1	2	0.26	0.57	0.86	1.35	1.66	2.17
1	2	0.24	0.55	0.86	1.37	1.66	2.17
1	2	0.26	0.57	0.86	1.37	1.66	2.17
1	3	0.26	0.59	0.86	1.35	1.66	2.17
1	3	0.24	0.55	0.86	1.37	1.68	2.17
1	3	0.24	0.57	0.86	1.35	1.66	2.17
2	1	0.46	0.75	1.06	1.53	1.84	2.33
2	1	0.39	0.73	0.99	1.48	1.77	2.26
2	1	0.39	0.73	0.99	1.48	1.79	2.24
2	2	0.44	0.73	0.99	1.48	1.79	2.24
2	2	0.44	0.75	0.99	1.46	1.79	2.26
2	2	0.39	0.75	1.04	1.51	1.79	2.26
2	3	0.46	0.75	1.06	1.48	1.79	2.26
2	3	0.44	0.73	1.04	1.48	1.79	2.26
2	3	0.39	0.75	1.02	1.48	1.79	2.24
3	1	0.22	0.51	0.79	1.26	1.59	2.06
3	1	0.22	0.48	0.77	1.24	1.57	2.04
3	1	0.19	0.48	0.77	1.24	1.57	2.06
3	2	0.22	0.46	0.75	1.24	1.55	2.04
3	2	0.22	0.46	0.75	1.24	1.55	2.04
3	2	0.22	0.46	0.77	1.24	1.55	2.06
3	3	0.22	0.48	0.79	1.26	1.59	2.06
3	3	0.22	0.48	0.77	1.22	1.55	2.04
3	3	0.19	0.46	0.77	1.26	1.57	2.06
4	1	0.26	0.55	0.82	1.24	1.53	1.97
4	1	0.26	0.53	0.82	1.24	1.53	1.95
4	1	0.26	0.53	0.79	1.22	1.53	1.95
4	2	0.26	0.53	0.79	1.22	1.51	1.93
4	2	0.26	0.53	0.79	1.22	1.48	1.88
4	2	0.26	0.53	0.79	1.22	1.51	1.93
4	3	0.26	0.55	0.82	1.22	1.48	1.93
4	3	0.26	0.53	0.79	1.22	1.53	1.95
4	3	0.26	0.53	0.79	1.22	1.48	1.93
5	1	0.19	0.48	0.79	1.26	1.59	2.06
5	1	0.19	0.48	0.79	1.26	1.55	2.04
5	1	0.19	0.48	0.79	1.26	1.55	2.02
5	2	0.19	0.51	0.77	1.24	1.55	2.04
5	2	0.19	0.48	0.79	1.24	1.55	2.02
5	2	0.19	0.42	0.73	1.19	1.48	2.04
5	3	0.19	0.48	0.79	1.26	1.57	2.06
5	3	0.19	0.46	0.79	1.24	1.55	2.02
5	3	0.19	0.46	0.77	1.24	1.55	2.02
6	1	0.28	0.57	0.82	1.22	1.46	1.86
6	1	0.28	0.55	0.79	1.19	1.42	1.79
6	1	0.28	0.55	0.79	1.19	1.46	1.84
6	2	0.28	0.55	0.82	1.22	1.46	1.84

Table 3: Continued.

Part	Operator	X_{ijk} 17	16	15	13.5	12.5	11
		Y_{ijk}					
6	2	0.28	0.55	0.79	1.22	1.46	1.84
6	2	0.28	0.57	0.82	1.22	1.46	1.84
6	3	0.28	0.55	0.79	1.22	1.46	1.84
6	3	0.28	0.55	0.79	1.19	1.44	1.82
6	3	0.26	0.53	0.79	1.19	1.42	1.79
7	1	0.19	0.48	0.77	1.24	1.53	1.99
7	1	0.19	0.46	0.77	1.22	1.53	1.97
7	1	0.17	0.46	0.75	1.22	1.53	1.97
7	2	0.19	0.46	0.77	1.22	1.53	1.97
7	2	0.17	0.46	0.77	1.22	1.53	1.95
7	2	0.19	0.46	0.77	1.22	1.53	1.99
7	3	0.19	0.46	0.75	1.22	1.53	1.97
7	3	0.19	0.46	0.77	1.22	1.55	1.97
7	3	0.17	0.46	0.75	1.22	1.51	1.97
8	1	0.26	0.55	0.84	1.33	1.64	2.13
8	1	0.24	0.53	0.82	1.31	1.62	2.08
8	1	0.24	0.53	0.82	1.33	1.64	2.08
8	2	0.26	0.55	0.82	1.28	1.62	2.08
8	2	0.24	0.53	0.82	1.28	1.59	2.06
8	2	0.26	0.57	0.82	1.33	1.62	2.08
8	3	0.26	0.53	0.82	1.33	1.62	2.08
8	3	0.26	0.53	0.82	1.33	1.64	2.08
8	3	0.24	0.53	0.82	1.28	1.62	2.06
9	1	0.15	0.44	0.75	1.22	1.53	1.99
9	1	0.15	0.39	0.75	1.22	1.53	1.99
9	1	0.13	0.39	0.73	1.22	1.53	1.99
9	2	0.15	0.39	0.73	1.22	1.53	2.02
9	2	0.17	0.42	0.73	1.22	1.53	2.02
9	2	0.15	0.42	0.75	1.22	1.53	1.99
9	3	0.15	0.39	0.73	1.22	1.53	2.02
9	3	0.15	0.39	0.75	1.22	1.53	1.99
9	3	0.15	0.39	0.73	1.19	1.51	1.97
10	1	0.26	0.55	0.79	1.19	1.42	1.82
10	1	0.26	0.53	0.79	1.19	1.42	1.82
10	1	0.26	0.53	0.77	1.17	1.39	1.75
10	2	0.28	0.55	0.79	1.19	1.42	1.82
10	2	0.26	0.53	0.77	1.15	1.37	1.75
10	2	0.28	0.57	0.79	1.19	1.44	1.82
10	3	0.26	0.53	0.79	1.15	1.39	1.77
10	3	0.26	0.53	0.79	1.15	1.39	1.75
10	3	0.26	0.53	0.77	1.13	1.37	1.77

Acknowledgments

The authors wish to thank Dr. Jinyu Yang for introducing the elementary idea of the paper. This research is supported by the National Natural Science Foundation of China (71102140) and the National Research Foundation for the Doctoral Program of Higher Education of China (20100032120030).

References

[1] R. K. Burdick, C. M. Borror, and D. C. Montgomery, "A review of methods for measurement systems capability analysis," *Journal of Quality Technology*, vol. 35, no. 4, pp. 342–354, 2003.

[2] D. C. Montgomery, *Introduction to Statistical Quality Control*, John Wiley and Sons, 2007.

[3] J. Lyu and M.-N. Chen, "Gauge capability studies for attribute data," *Quality and Reliability Engineering International*, vol. 24, no. 1, pp. 71–82, 2008.

[4] C. M. Borror, D. C. Montgomery, and G. C. Runger, "Confidence intervals for variance components from gauge capability studies," *Quality and Reliability Engineering International*, vol. 13, no. 6, pp. 361–369, 1997.

[5] L. Gong, R. K. Burdick, and J. Quiroz, "Confidence intervals for unbalanced two-factor gauge R&R studies," *Quality and Reliability Engineering International*, vol. 21, no. 8, pp. 727–741, 2005.

[6] S. Sweeney, "Analysis of two-dimensional gage repeatability and reproducibility," *Quality Engineering*, vol. 19, no. 1, pp. 29–37, 2007.

[7] K. D. Majeske, "Approval criteria for multivariate measurement systems," *Journal of Quality Technology*, vol. 40, no. 2, pp. 140–153, 2008.

[8] M.-H. C. Li and A. Al-Refaie, "Improving wooden parts' quality by adopting DMAIC procedure," *Quality and Reliability Engineering International*, vol. 24, no. 3, pp. 351–360, 2008.

[9] T. P. Erdmann, R. J. M. M. Does, and S. Bisgaard, "Quality quandaries: a gage R&R study in a hospital," *Quality Engineering*, vol. 22, no. 1, pp. 46–53, 2010.

[10] Automotive Industry Action Group, *Measurement Systems Analysis: Reference Manual*, Automotive Industry Action Group, 2010.

[11] D. J. Wheeler and R. W. Lyday, *Evaluating the Measurement Process*, Statistical Process Controls, 1984.

[12] W. H. Woodall and C. M. Borror, "Some relationships between gage R&R criteria," *Quality and Reliability Engineering International*, vol. 24, no. 1, pp. 99–106, 2008.

[13] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, Prentice hall, Upper Saddle River, NJ, USA, 2002.

[14] K. Kim, M. A. Mahmoud, and W. H. Woodall, "On the monitoring of linear profiles," *Journal of Quality Technology*, vol. 35, no. 3, pp. 317–328, 2003.

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 578280, 6 pages
http://dx.doi.org/10.1155/2014/578280

Research Article

Optimal Order Strategy in Uncertain Demands with Free Shipping Option

Qing-Chun Meng,[1] Teng Zhang,[2] Ming Li,[1] and Xiao-Xia Rong[2]

[1] *School of Management, Shandong University, Jinan, Shandong 250100, China*
[2] *School of Mathematics, Shandong University, Jinan, Shandong 250100, China*

Correspondence should be addressed to Xiao-Xia Rong; rongxiaoxia@sdu.edu.cn

Received 10 March 2014; Accepted 8 May 2014; Published 25 May 2014

Academic Editor: Xiang Li

Copyright © 2014 Qing-Chun Meng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Free shipping with conditions has become one of the most effective marketing tools; more and more companies especially e-business companies prefer to offer free shipping to buyers whenever their orders exceed the minimum quantity specified by them. But in practice, the demands of buyers are uncertain, which are affected by weather, season, and many other factors. Firstly, we model the centralization ordering problem of retailers who face stochastic demands when suppliers offer free shipping, in which limited distributional information such as known mean, support, and some deviation measures of the random data is needed only. Then, based on the linear decision rule mainly for stochastic programming, we analyze the optimal order strategies of retailers and discuss the approximate solution. Further, we present the core allocation between all retailers via dual and cooperative game theory. The existence of core shows that each retailer is pleased to cooperate with others in the centralization problem. Finally, a numerical example is implemented to discuss how uncertain data and parameters affect the optimal solution.

1. Introduction

With the rapid development of e-commerce and logistics industry, free shipping offered by e-business companies has become an effective means of attracting and keeping customers. Many business-to-consumer and business-to-business (B2B) companies now offer free shipping to buyers who spend more than a specific amount. More and more companies and businesses begin to take free shipping strategy, such as Amazon online bookstore; when the cost reaches $25, the buyer can get free shipping service. Taobao and Dangdang also offer free shipping if the certain amount is satisfied. The growth and evolution of the e-commerce sector have highlighted the importance of shipping and handling (S&H) fees for business models. The supplier can effectively reduce order processing costs and implementation costs, if they can reduce the frequent small shipments. Otherwise, if many retailers are joined together to order, the cost may be saved for the satisfaction of free shipping. Therefore, free shipping schedules have become an interest study for both the supplier and demander. Survey evidence indicates that shipping fees are the main complaint of more than 50 percent of online shoppers and that more than 60 percent of shoppers have abandoned an order when shipping fees are added. Academic work has further confirmed that fulfillment issues are a key driver of customer satisfaction.

In this paper, we study the centralization ordering problem with uncertain demands considering free shipping. Considering a supply chain includes a supplier and a number of retailers whose demands for commodities are uncertain, and retailers focus on how free shipping schedules impact their ordering strategies. Lewis et al. [1] used an ordered probability model to account for the effects of nonlinear and discontinuous free shipping on purchasing decisions. It shows that the retailers are very sensitive to shipping charges, and promotions such as free shipping and free shipping for orders that exceed some size threshold are very effective in generating additional sales. Leng et al. [2, 3] considered the free shipping strategies of business-to-consumer and business-to-business environments companies by modeling the problem as a leader-follower game under complete information where the leader is the seller and the follower is the buyer. Yang et al. [4] analyzed the price-threshold relationship which was inverted-U shaped and explored how prices and the free shipping threshold interact to affect the optimal policy. The initial price level dictates the price dispersion for homogenous goods increases when the threshold is lowered. Zhou et al. [5] gave the management of stochastic inventory systems with free shipping option. Abad and Aggarwal [6] studied the pricing decisions with random demand in order to reduce transport costs, which is free shipping with condition. Hua et al. [7] studied the optimal order strategy of a retailer who faces deterministic or stochastic demand when suppliers offer free shipping. It analyzes the impacts of the transportation cost on the retailer's optimal order strategy based on EOQ model and newsvendor model.

One issue of the above study is the assumption of full distributional knowledge of the uncertain data. Because such information may rarely be available in practice, it has rekindled recent interests in robust optimization as an alternative perspective of data uncertainty. In robust optimization, compared with the full distributional knowledge which is hardly got, limited distributional information such as known mean, support, and some deviation measures of the random data is required only. So in this paper, we consider the uncertain demand in general and study the optimal ordering model with free shipping.

The paper is organized as follows. In Section 2, we give the stochastic programming model of optimal order strategy about the retailers, in which the demands are uncertain with free shipping option. In Section 3, based on the linear decision rules, we analyze the robust counterpart of the stochastic programming model and formulate a new equivalent determined model. In Section 4 we use cooperative game theory to get the core of all retailers. In Section 5, a numerical experiment confirms that order incidence is affected by free shipping option and varying interval of demand. Finally, Section 6 concludes this paper.

2. Problem Description

Assuming there are a supplier and a number of retailers, who just trade only a type of goods. The supplier offers the goods to retailers whose demands for commodities are uncertain. All retailers order goods uniformly and order price is constant. Only when the total ordering amount reaches a certain threshold, the supplier can offer free shipping for retailers. Here we consider how to maximize the benefits of all retailers by selecting their optimal order quantity. In this problem, we take the following assumptions. The retailers are all rational and their inventories are inadequate to meet the real demands, so they are willing to participate in group to order goods according to their actual situation. In addition, there is no competition among retailers and they are willing to participate in the group to pay the total minimum fee. So the problem aims to minimize the total cost of all retailers with constraint that their demands are met.

At first, we denote the following notions. The notions m and c are the retail price and order price of the goods separately and q is the known threshold of free shipping. The notion n is the number of retailers, and the random demand of the ith retailer is $d_i(\tilde{z})$ in which $\tilde{z}$ means random environment, and all $d_i(\tilde{z})$ are independent. L_i is the current inventory of the ith retailer. We denote the order quantity x_i by the decision variable. Only when the condition $\sum_{i=1}^{n} x_i \geq q$ holds, the supplier can offer free shipping for retailers; otherwise the retailers should cost the $f(\sum_{i=1}^{n} x_i)$. The symbol y is 0-1 variable, in which 1 means payoff for the transport and 0 is free shipping. The symbol $w_i(\tilde{z})$ is the amount of the shortage of goods of the ith retailer, which is caused by $d_i(\tilde{z})$ in uncertain environment. Then the model of optimal order strategy with free shipping option is given as follows (1):

$$\min \quad c \times \left(\sum_{i=1}^{n} x_i\right) + f\left(\sum_{i=1}^{n} x_i\right) \times y + m \times E\left(\sum_{i=1}^{n} w_i(\tilde{z})\right),$$
$$\text{s.t.} \quad \begin{cases} x_i + w_i(\tilde{z}) + L_i \geq d_i(\tilde{z}), & i = 1, 2, \ldots n, \\ x_i, w_i(\tilde{z}) \geq 0, & i = 1, 2, \ldots n, \\ y = 0, & \text{if} \sum_{i=1}^{n} x_i \geq q, \\ y = 1, & \text{if} \sum_{i=1}^{n} x_i < q, \end{cases} \tag{1}$$

where the objective function contains the ordering cost, transportation cost, and penalty cost incurred at the retailers if the demands are not satisfied. For the penalty cost is relevant to uncertain realization of $\tilde{z}$, the expectation is taken here. According to the objective function, we require that the order quantity of the goods be not too much to add inventory and at the same time the shortage be not too much to increase cost. The first constraint means that, for the ith retailer, the sum of ordering quantity, shortage, and inventory quantity be not less than their demand. The second constraint shows that the decision variables of ordering quantity and shortage are nonnegative. The third constraint ensures that when the ordering quantity is not less than the given threshold q, the supplier can offer free shipping for retailers; otherwise the retailers should pay for the transportation cost.

The slack variable $v_i(\tilde{z})$ is added in the first constraint, so model (1) can be rewritten to the model

$$\min \quad c \times \left(\sum_{i=1}^{n} x_i\right) + f\left(\sum_{i=1}^{n} x_i\right) \times y + m \times E\left(\sum_{i=1}^{n} w_i(\tilde{z})\right),$$
$$\text{s.t.} \quad \begin{cases} x_i + w_i(\tilde{z}) + L_i - v_i(\tilde{z}) = d_i(\tilde{z}), & i = 1, 2, \ldots n, \\ x_i, w_i(\tilde{z}), \ v_i(\tilde{z}) \geq 0, & i = 1, 2, \ldots n, \\ y = 0, & \text{if} \sum_{i=1}^{n} x_i \geq q, \\ y = 1, & \text{if} \sum_{i=1}^{n} x_i < q. \end{cases} \tag{2}$$

In this paper, we assume the random demand $d_i(\tilde{z})$ is generalized variable, whose only limited distributional information is known, but distributional function and other full knowledge are unknown. Then, we discuss the solvability approximately of the stochastic programming (2).

3. Approximation via Decision Rule

It is difficult to solve the general stochastic programming; besides this, the full distributional knowledge of the uncertain data is needed, which may rarely be available in practice. But assuming only limited distribution information is known such as mean, support, and some deviation measures of the random data, linear decision rule is the key enabling method that permits scalability to multistage models [8]. Interesting applications include designing supplier-retailer contracts, network design under uncertainty, and crashing projects with uncertain activity times. Even though linear decision rule allows us to derive tractable formulations in a variety of applications, it may lead to infeasible instances [9]. This fact motivates people to refine linear decision rule and improve it to a general linear decision rule, which improves the objective value. Chen et al. [10] gave the conclusion that when complete recourse exists, the general linear decision rule is equal to the linear decision rule. Because the recourse matrix in this paper is special and the support is bounded closed set, it is feasible to analyze the solvability of model (2) using the linear decision rule.

3.1. A Two-Stage Stochastic Linear Programming Model. The linear decision rule is used mainly to solve the multistage stochastic programming. For the stochastic linear programming (3), decision x has to be made before the actual value of $\tilde{z}$ is realized which consists the first stage. After applying the decision and after the uncertainty is realized, the subjects of (3) may be not satisfied and the optimal second-stage decisions or recourse decisions are carried out, in which $w(\tilde{z})$ is recourse variable and W is recourse matrix. So the subjects of (3) are satisfied and, at the same time, the cost that is aroused by the recourse in (4) is minimized. The problems (3) and (4) can be rewritten in model (5) equivalently:

$$\begin{aligned} &\min \quad c^T x + E\left(Q\left(x,\tilde{z}\right)\right), \\ &\text{s.t.} \quad \begin{cases} T\left(\tilde{z}\right) x = h\left(\tilde{z}\right), \\ \qquad x \geq 0, \end{cases} \end{aligned} \tag{3}$$

$$\begin{aligned} &Q\left(x,\tilde{z}\right) = \min \quad mw\left(\tilde{z}\right), \\ &\text{s.t. } T\left(\tilde{z}\right) x + Ww\left(\tilde{z}\right) = h\left(\tilde{z}\right), \end{aligned} \tag{4}$$

$$\begin{aligned} &\min \quad c^T x + mE\left(w\left(\tilde{z}\right)\right), \\ &\text{s.t.} \quad \begin{cases} T\left(\tilde{z}\right) x + Ww\left(\tilde{z}\right) = h\left(\tilde{z}\right), \\ \qquad\qquad x \geq 0. \end{cases} \end{aligned} \tag{5}$$

Comparing models (2) and (5), we know easily that (2) is also a two-stage stochastic linear program with the special recourse matrix.

3.2. Problem Analysis Based on Linear Decision Rule. According to the literature [10], assuming the stochastic programming with the relatively complete recourse, the second-stage problem is surely feasible for any choice of feasible first-stage decision vector x. The complete recourse is defined on the matrix W such that, for any t, there exists $w \geq 0$, satisfying $Ww = t$. Hence, the definition of complete recourse depends only on the structure of the matrix W, which makes the problem easier to solve. For the model (2), if we let $W_i = [-1, 1]$ for each retailer, there exists a simple case of complete recourse where the special matrix is

$$\mathbf{W} = \begin{bmatrix} -1 & 1 & & & & & \\ & & -1 & 1 & & & \\ & & & & \cdots & & \\ & & & & & -1 & 1 \end{bmatrix}. \tag{6}$$

So it is feasible to analyze the solvability of the problem (2) via the linear decision rule. Based on the linear decision rule given in [9], we assume both the recourse variable $w_i(\tilde{z})$ and the slack variable $v_i(\tilde{z})$ are the affine functions. For convenience, we describe them by using vector form below. Let $w(\tilde{z}) = w^{\mathbf{0}} + \sum_{k=1}^{N} w^k \tilde{z}_k$, where the coefficients w^k are unknown. At the same time, the demands of all retailers are in the same linear form $d(\tilde{z}) = d^0 + \sum_{k=1}^{N} d^k \tilde{z}_k$. Generally for the uncertain variable $\tilde{z}$, we may assume its mean is 0 and the support is $W_{\tilde{z}} = [-\underline{z}, \overline{z}]$, $\underline{z} > 0$, $\overline{z} > 0$, which is also the value set of $\tilde{z}$. So the model can be adjusted to the following problem:

$$\begin{aligned} &Z_{\text{STOC}} = \min \quad c^T x + f\left(x\right) y + me^T w^0, \\ &\text{s.t.} \quad \begin{cases} x + L + w^0 - v^0 = d^0, & \\ w^k - v^k = d^k, & k = 1, 2, \ldots, N, \\ x, w, v \geq 0, & \\ y = 0, & \text{if} \displaystyle\sum_{i=1}^{n} x_i \geq q, \\ y = 1, & \text{if} \displaystyle\sum_{i=1}^{n} x_i < q. \end{cases} \end{aligned} \tag{7}$$

Theorem 1. *For model* (7)*, the third subject $w(\tilde{z}) \geq 0$, for all $\tilde{z} \in W_{\tilde{z}} = [-\underline{z}, \overline{z}]$, holds if and only if there is $w^0 \geq \overline{z} \sum_{k=1}^{N} t^{1k} + \underline{z} \sum_{k=1}^{N} s^{1k}$.*

Proof. Let $w^k = s^{1k} - t^{1k}$; $s^{1k}, t^{1k} \in R^n$; $s^{1k}, t^{1k} \geq 0$, then the following equivalence relations are easy to get:

$$\begin{aligned} &w\left(\tilde{z}\right) = w^0 + \sum_{k=1}^{N} w^k \tilde{z}_k \geq 0 \\ &\Longleftrightarrow w^0 \geq -\sum_{k=1}^{N} w^k \tilde{z}_k = -\sum_{k=1}^{N} \left(s^{1k} - t^{1k}\right) \tilde{z}_k \\ &\qquad = \sum_{k=1}^{N} \left(t^{1k} \tilde{z}_k - s^{1k} \tilde{z}_k\right) \\ &\Longleftrightarrow w^0 \geq \max \sum_{k=1}^{N} \left(t^{1k} \tilde{z}_k - s^{1k} \tilde{z}_k\right) = \overline{z} \sum_{k=1}^{N} t^{1k} + \underline{z} \sum_{k=1}^{N} s^{1k}. \end{aligned} \tag{8}$$

□

Similar analysis is considered for vector v in model (7), and based on the assumption that the transportation cost is linear function, the uncertain model (7) is rewritten to the determinate model

$$Z_{\mathrm{LDR}} = \min \quad c^T x + re^T xy + me^T w^0,$$
$$\text{s.t.} \quad \begin{cases} x + L + w^0 - v^0 = d^0, & \\ w^k - v^k = d^k, & k = 1, 2, \dots N, \\ w^0 \geq \overline{z}\sum_{k=1}^{N} t^{1k} + \underline{z}\sum_{k=1}^{N} s^{1k}, & \\ v^0 \geq \overline{z}\sum_{k=1}^{N} t^{2k} + \underline{z}\sum_{k=1}^{N} s^{2k}, & \\ x, s^1, t^1, s^2, t^2 \geq 0, & \\ y = 0, & \text{if} \sum_{i=1}^{n} x_i \geq q, \\ y = 1, & \text{if} \sum_{i=1}^{n} x_i < q. \end{cases} \tag{9}$$

Because model (9) is the linear case of the model (7), the following conclusion is obtained.

Theorem 2 ($Z_{\mathrm{STOC}} \leq Z_{\mathrm{LDR}}$). *Hence, by introducing the linear decision rule for the primal model* (2)*, we get the robust model* (7)*, whose optimal value is approximate to the value of the determinate model* (9)*. Regarding the issue of the bound on the objective function, the literature [10] made detailed discussions.*

4. Cooperative Game of the Ordering Centralization Problem

In the ordering centralization problem considering free shipping, we only focus on the allocation of the expected cost. Whether it is possible to derive a stable allocation of the actual cost for each demand realization, the cost of each retailer is reduced. In this section, based on the dual theory of stochastic programming, we analyze the centralization problem with free shipping.

First we briefly introduce the concepts of cooperative game theory that will be used. Let $N = \{1 \cdots n\}$ be the set collection of players. A collection of players $S \subseteq N$ is called a coalition. A characteristic cost function is defined for each coalition $S \subseteq N$. A cooperative game is defined by the pair (N, C). Given a cooperative game, there are many ways to divide the cost allocation of the game among the players. The cost allocation has been extensively studied in the literature. We focus on the core allocation, which is defined below.

A vector $l = (l_1, l_2, \dots, l_n)$ is called an imputation of the game (N, C) if $\sum_{j \in N} l_j = C(N)$ and $l_j \leq C(\{j\})$ for every $j \in N$. One can interpret an imputation as a division of $C(N)$ that charges every player at most as much as that they will play by themselves. When this idea is generalized to every coalition of players, the notion of core is given.

Definition 3. An allocation $l = (l_1, l_2, \dots, l_n)$ is in the core of the game of (N, C), if $\sum_{j \in N} l_j = C(N)$ and $\sum_{j \in S} l_j \leq C(S)$ for every $S \subseteq N$.

Then, we consider the following two-stage stochastic linear programming:

$$\begin{aligned} \min \quad & v_1^T x_1 + E\left[v_2^T x_2\right], \\ \text{s.t.} \quad & A_{11} x_1 = b_1, \\ & A_{12} x_1 + A_{22} x_2 = b_2(\tilde{z}), \\ & x_1, x_2 \geq 0, \end{aligned} \tag{10}$$

where x_1 is the decision variable of the first stage, x_2 is the recourse variable of the second stage in which $A_{12}x_1 = b_2(\tilde{z})$ is not satisfied for the realization of $\tilde{z}$, and $v_2^T x_2$ is the recourse cost.

The dual problem of problem (10) is

$$\begin{aligned} \max \quad & E\left[b_1^T \pi_1 + b_2(\tilde{z})^T \pi_2(\tilde{z})\right], \\ \text{s.t.} \quad & A_{11}^T \pi_1 + E\left[A_{12}^T \pi_2(\tilde{z})\right] \leq v_1, \\ & A_{22}^T \pi_2(\tilde{z}) \leq v_2. \end{aligned} \tag{11}$$

For the realization of $\tilde{z} \in W$, if W is a finite set, then model (10) is a linear programming and the strict dual condition is satisfied. Otherwise, the strict dual condition may not be satisfied [11]. In the literature [12], there is a conclusion that the optimal value of (10) is equal to the optimal value of model (11) when model (10) is feasible and has relatively complete recourse matrix. So we can study the dual problem of the primal problem (1) to find the relationship of them.

Here, assuming that the condition of free-transportation cost holds, the problem (1) is equal to the problem (12). In the model (12), for every collaboration $S \subseteq N$, let the set S replace the set N; then the collaboration model (13) is given as follows:

$$C(S) = \min \quad c\left(\sum_{i=1}^{n} x_i\right) + mE\left(\sum_{i=1}^{n} w_i(\tilde{z})\right),$$
$$\text{s.t.} \quad \begin{cases} \sum_{i=1}^{n} x_i \geq q, & \\ x_i + w_i(\tilde{z}) \geq d_i(\tilde{z}) - L_i, & i = 1, 2, \dots n, \\ x_i, w_i(\tilde{z}) \geq 0, & i = 1, 2, \dots n, \end{cases} \tag{12}$$

$$\min \quad c\left(\sum_{i \in S} x_i\right) + mE\left(\sum_{i \in S} w_i(\tilde{z})\right),$$
$$\text{s.t.} \quad \begin{cases} \sum_{i \in S} x_i \geq q, & \\ x_i + w_i(\tilde{z}) \geq d_i(\tilde{z}) - L_i, & i \in S, \\ x_i, w_i(\tilde{z}) \geq 0, & i \in S. \end{cases} \tag{13}$$

It is easy to see that the dual problem of model (13) is

$$D(S) = \max q\alpha + \sum_{i=1}^{|S|} E\left(\left(d_i(\tilde{z}) - L_i\right)\alpha_i(\tilde{z})\right),$$
$$\text{s.t.} \begin{cases} E\left(\alpha_i(\tilde{z})\right) + \alpha \le c, & i \in S, \\ \alpha_i(\tilde{z}) \le m, & i \in S, \\ \alpha, \alpha_i(\tilde{z}) \ge 0, & i \in S. \end{cases} \tag{14}$$

In Section 3, we have given the example that the relatively complete recourse matrix exists in the inventory centralization problem considering free shipping. Hence we propose the following results.

Theorem 4. *For any collection of retailers $S \subseteq N$, the optimal value of* (13) *is equal to the optimal value of* (14).

Theorem 5. *The allocation in the core of retailers exists in model* (1) *of cooperative ordering game with free shipping option.*

Proof. For $S = N$, we denote by $(\alpha_j^*(\tilde{z}), \alpha^*)$ the optimal solution of problem Dual(N) and let

$$l_j = \varepsilon_j + E\left(\left(d_j(\tilde{z}) - L_j\right)\alpha_j^*(\tilde{z})\right), \tag{15}$$

where $\sum_{j=1}^{N} \varepsilon_j = q\alpha,\ \varepsilon_j \ge 0$. From Theorem 4, we know that

$$\sum_{j\in N} l_j = \sum_{j\in N}\left(\varepsilon_j + E\left(\left(d_j(\tilde{z}) - L_j\right)\alpha_j^*(\tilde{z})\right)\right) = C(N). \tag{16}$$

On the other hand, for every $S \subseteq N$, because $(\alpha_j^*(\tilde{z}), \alpha^*)$ is a feasible solution to Dual(S), we have the following inequality:

$$\sum_{j\in S} l_j = \sum_{j\in S}\left(\varepsilon_j + E\left(\left(d_j(\tilde{z}) - L_j\right)\alpha_j^*(\tilde{z})\right)\right) \le q\alpha + \sum_{j\in S} E\left(\left(d_j(\tilde{z}) - L_j\right)\alpha_j^*(\tilde{z})\right) \le C(S). \tag{17}$$

By Definition 3, the vector $(l_1 \cdots l_N)$ defined by (15) is an allocation in the core of the cooperative inventory game (N, C). □

In general, a core is the set of cost allocations under which no coalitions should be charged more than they would pay if they were to separate and follow an optimal strategy for themselves. That is, no coalition will be better off by deviation from the grand coalition. Since the core is nonempty, there is at least one allocation of the cost that is considered advantageous by all players. So for each retailer of the inventory centralization problem, he is pleased to cooperate with others based on the cost-vector $(l_1 \cdots l_N)$ defined by (15).

5. Numerical Experiment

Assuming that there are three retailers and they order goods from the same supplier. The retail price of the unit goods is $m = 70$ and the order price is $c = 40$. The transportation function is proportional to order quantity; that is, $r = 4$. The given threshold is $q = 60$. The current inventory of retailers is 5, 8, and 10, respectively, and their demands are all in the same form $d_i(\tilde{z}) = 30 + \tilde{z}$, where the mean of $\tilde{z}$ is zero and the support is $\tilde{z} \in [-1, 2]$.

Solving model (9) with the above data by MATALB solver, we get the optimal ordering strategy of three retailers being 24, 21, and 19, respectively and the optimal cost value is 2770, in which the free shipping is satisfied for the total ordering number which is 64. On the other hand, if the free shipping is not satisfied, the optimal ordering strategy is 20, 17, and 15, respectively, and the total ordering number is 52, and the optimal cost value is 3338. This numerical experiment shows that if all retailers participate in group to order goods, the total cost may reduce for more ordering quantity. So all rational retailers are willing to order jointly when their inventories are inadequate to meet the real demands with the consideration of free shipping option.

Then we consider how the support of $\tilde{z}$ affects the optimal value. For the normal demand is 30 with a little uncertain changes, we test other 5 kinds of support with the maximal interval being 5. There are global optimal solutions in the cases of $[-1, 2]$, $[-1, 1]$, and $[-2, 2]$, but in other cases MATALB solver shows the current solution may be nonoptimal and the local optimal solutions can be found after more than one hundred iterations. The experiment shows the support has obvious effect on the optimal value of model (9) which is given in Table 1.

Table 1: The relationship between support and the optimal value.

Support of $\tilde{z}$	$[-1, 2]$	$[-1, 1]$	$[-2, 2]$	$[-3, 3]$	$[-4, 4]$	$[-5, 5]$
The optimal value Z_{LDR}	2770	2796.6	2871.6	4017.5	3676.7	3728.7

Further, we consider how the threshold q influences the optimal value. At the beginning we take the value 5 times according to the current value that is 60, and the optimal function values are given in Table 2 (the current value may be local optimal according to the solver for the cases that the threshold is more than 70). It shows that the optimal cost value increases from $q = 45$ to $q = 55$; then the minimal value reaches the bottom from $q = 55$ to $q = 60$; afterwards it increases from $q = 60$ to $q = 80$. In order to find the relationship better, get 35 groups of the threshold and optimal value ($q = 45, 46, \ldots, 80$ resp.), which are displayed in Figure 1. The experiment shows that the optimal value obeys piecewise function with obvious character.

Table 2: The relationship between threshold and the optimal value.

Threshold of q	45	50	55	60	65	70	75	80
The optimal value Z_{LDR}	3280	3159	2770	2770	2803	3026	3026	3026

At last we try to compute the core of the primal problem above based on the discussion in Section 4. For the condition of free-transportation cost is satisfied, the total minimal cost is 2770. In (14), the total maximal payment is equal to 2770

Figure 1: The detailed relationship between threshold and the optimal value.

also. According to dual theory, we have $\alpha^* = 0$, and the costs that three retailers should take are $l_1 = 25E(\alpha_1(\tilde{z}))$, $l_2 = 22E(\alpha_2(\tilde{z}))$, and $l_3 = 20E(\alpha_3(\tilde{z}))$ with constraint $E(\alpha_i(\tilde{z})) \leq 40$ and for $\alpha_i(\tilde{z}) \leq 70$ all i, where $\alpha_i(\tilde{z})$ is random decision variable. The allocations satisfying restrictions above are feasible.

6. Conclusions

It is a key research content of supply chain management. This research will provide effective effort for scientific decision making and e-business activity. This paper studied the optimal order problem under the uncertainty of demand and proposed the stochastic programming model, in which the objective function is to minimize the total cost. Considering the limited information of the uncertain variable in the model, we used the linear decision rule; one of the robust optimization methods to analyze this model and get the approximate model which is tractable. Afterwards, evidence attained from the numerical experiment strongly suggested its effectiveness and efficiency. Furthermore, this paper analyzed the core allocation between all retailers via dual and cooperative game theory. The results showed that the ordering centralization can save the cost of all retailers and it is economic for the society. We will study the application of robust optimization to supply management with other uncertainties or in the view of supplier.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This paper is supported by the National Nature Science Foundation of China (NSFC) (Grant no. 71250002) and Independent Innovation Foundation of Shandong University (Grant no. IFW12107).

References

[1] M. Lewis, V. Singh, and S. Fay, "An empirical study of the impact of nonlinear shipping and handling fees on purchase incidence and expenditure decisions," *Marketing Science*, vol. 25, no. 1, pp. 51–64, 2006.

[2] M. Leng and M. Parlar, "Free shipping and purchasing decisions in B2B transactions: a game-theoretic analysis," *IIE Transactions*, vol. 37, no. 12, pp. 1119–1128, 2005.

[3] M. Leng and R. Becerril-Arreola, "Joint pricing and contingent free-shipping decisions in B2C transactions," *Production and Operations Management*, vol. 19, no. 4, pp. 390–405, 2010.

[4] Y. Yang, S. Essegaier, and D. R. Bell, "Free shipping and repeat buying on the internet: theory and evidence," Manuscript, Wharton School, University of Pennsylvania, 2005.

[5] B. Zhou, M. N. Katehakis, and Y. Zhao, "Managing stochastic inventory systems with free shipping option," *European Journal of Operational Research*, vol. 196, no. 1, pp. 186–197, 2009.

[6] P. L. Abad and V. Aggarwal, "Incorporating transport cost in the lot size and pricing decisions with downward sloping demand," *International Journal of Production Economics*, vol. 95, no. 3, pp. 297–305, 2005.

[7] G.-W. Hua, J. Li, X. Tian, and F. M. Yang, "Optimal order strategy with free shipping option," *System Engineering Theory and Practice*, vol. 30, no. 3, pp. 506–512, 2010 (Chinese).

[8] A. Shapiro and A. Nemirovski, "On complexity of stochastic programming problems," in *Continuous Optimization*, pp. 111–146, Springer, New York, NY, USA, 2005.

[9] X. Chen, M. Sim, and P. Sun, "A robust optimization perspective on stochastic programming," *Operations Research*, vol. 55, no. 6, pp. 1058–1071, 2007.

[10] X. Chen, M. Sim, P. Sun, and J. Zhang, "A linear decision-based approximation approach to stochastic programming," *Operations Research*, vol. 56, no. 2, pp. 344–357, 2008.

[11] X. Chen and J. Zhang, "A stochastic programming duality approach to inventory centralization games," *Operations Research*, vol. 57, no. 4, pp. 840–851, 2009.

[12] R. T. Rockafellar and R. J.-B. Wets, "Stochastic convex programming: relatively complete recourse and induced feasibility," *SIAM Journal on Control and Optimization*, vol. 14, no. 3, pp. 574–589, 1976.

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 289239, 10 pages
http://dx.doi.org/10.1155/2014/289239

Research Article

Oil Well Characterization and Artificial Gas Lift Optimization Using Neural Networks Combined with Genetic Algorithm

Chukwuka G. Monyei,[1] Aderemi O. Adewumi,[2] and Michael O. Obolo[3]

[1] *Department of Electrical and Electronic Engineering, University of Ibadan, Nigeria*
[2] *School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, South Africa*
[3] *Department of Petroleum Engineering, University of Ibadan, Nigeria*

Correspondence should be addressed to Aderemi O. Adewumi; laremtj@gmail.com

Received 24 January 2014; Revised 23 March 2014; Accepted 8 April 2014; Published 22 May 2014

Academic Editor: Weihua Liu

Copyright © 2014 Chukwuka G. Monyei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper examines the characterization of six oil wells and the allocation of gas considering limited and unlimited case scenario. Artificial gas lift involves injecting high-pressured gas from the surface into the producing fluid column through one or more subsurface valves set at predetermined depths. This improves recovery by reducing the bottom-hole pressure at which wells become uneconomical and are thus abandoned. This paper presents a successive application of modified artificial neural network (MANN) combined with a mild intrusive genetic algorithm (MIGA) to the oil well characteristics with promising results. This method helps to prevent the overallocation of gas to wells for recovery purposes while also maximizing oil production by ensuring that computed allocation configuration ensures maximum economic accrual. Results obtained show marked improvements in the allocation especially in terms of economic returns.

1. Introduction

Petroleum, a limited natural resource, is a nonrenewable form of energy on which humans largely depend. This leads to pressing market demands, accessibility issues, and competitive market environment that force oil companies to seek technologies and procedures that can give competitive advantage and meet environmental restrictions while streamlining production processes and cutting costs [1].

Artificial gas lift (AGL) is a recovery process that involves the use of gases, produced (from oil) or purchased, which are pumped into the well bore to maintain formation pressure, that is, the pressure at which the fluid flows to the surface. There are two types of AGL, namely, intermittent gas lift and continuous gas lift [1]. However, this paper is not concerned with the types of AGL but rather its distribution or allocation.

The gas lift process involves the injection of high pressure gas at the bottom of the production tubing of an oil well [1–3]. In other words, AGL involves injecting high-pressured gas from the surface into the producing fluid column through one or more subsurface valves set at predetermined depths [2, 3]. This helps to improve recovery by reducing the bottom-hole pressure at which wells become uneconomic, resulting in being abandoned. The gas, mixed with the oil, diminishes the weight of the fluid column thereby reducing the downhole pressure. A low downhole pressure induces a flux of fluids from the reservoir to the well. The produced fluid is composed of oil, gas, and water. The water must be treated before being discharged which incurs costs while the gas can be either reused in the process or sent to customers and other facilities [3]. In large oil fields, several separators are used to divide the three phases. This gives rise to the problem of maximizing production by allocating lift-gas to the wells while defining the routing from wells to separators and observing separator capacities [2, 3].

In particular, the gas lift operation of oil fields is one of many production processes whose performances can be improved. As the internal pressure in high-depth or depleted reservoirs can force the flow of only a fraction of their oil to the surface, the use of artificial means becomes imperative to lift the oil, especially for deep reservoirs that are found off-shore. Two examples of artificial lifting are submerged pumps and continuous injection of gas [3]. Although the former can, in principle, recover most of the oil, its operating costs are excessively high for today's oil prices, not to mention the potential of an unfavourable energy trade and other technical hindrances. The gas lift technique, on the other hand, harnesses the reservoir's gas by injecting natural gas into the production tubing so as to reduce the weight of the oil column, thereby elevating the mix of oil, gas, and water to the surface.

A motivation for this work therefore arises from the need to reduce wastage in gas allocation especially in the unlimited scenario thus freeing up gas thus freeing up gas for other uses such as domestic, transportation and electricity generation purposes. With global outcry to the insidious effect of green-house gases on our environment and the need for prudent management of scarce resources, this paper aims at providing a cost-effective method for solving the problem of gas allocation for recovery purposes in the oil and gas industry.

The rest of this paper is organized as follows: Section 2 presents a brief overview of related works in literature while Section 3 further describes the problem as well as modelling approach. Section 4 gives an overview of the methodology adopted to solve the oil well characterization and AGL problem while results obtained are discussed in Section 5. The final section presents some useful conclusions as well as direction for further works.

2. Literature Review

Gas injection has been used to maintain reservoir pressure at some selected levels or to supplement natural reservoir energy to a lesser degree by reinjection of a portion of the produced gas. Complete or partial pressure maintenance operations can result in increased hydrocarbon recovery and improved reservoir production characteristics. A general position opines that daily oil production increases concomitantly with gas up to a certain level where further gas injection yields a decrease in oil production with increased gas cost [2]. Ray and Sarker [2] developed a multiobjective constrained algorithm to optimize gas lift allocation within the constraint of limited available gas. The proposed solution was applied to six and fifty-six well problems with single and multiobjective problem formulations. De Souza et al. [4] described a case study that involved modelling and optimizing gas allocation for deep water offshore petroleum production with interest in determining the rate of injected gas flow that guarantees maximum oil production, profit, and optimal design of gas lift system considering capital cost of compressors, turbine, and gas pipeline constraints. The problem was modelled as a nonlinear optimization problem and solved as a two-phase network flow model. Codas and Camponogara [1] addressed the problem of gas lift allocation with separator routing constraint using a mixed integer linear model solved using the CPLEX software. Mahmudi and Sadeghi [5] used a hybrid computational model consisting of genetic algorithm (GA) and Marquardt algorithm to optimize gas allocation under various constraints including effects of tubing diameter, rates of gas injection, and separator pressure on the economic return of the well over a long period.

From the foregone and other research trends, the problem of well characterization and gas allocation is mainly addressed independently. Moreover, specific interests of earlier works did not focus on configuration selection but rather on routing mechanisms such as separator routing constraint, effects of tubing diameter, rates of gas injection and separator pressure on the economic return of the well over long period, and capital cost of compressors, turbine, and gas pipeline amongst others. Most previous works therefore focused on the contribution of the artificial gas lift layout and material selection in ensuring optimum gas yield. This paper seeks to complement ongoing research by proposing a combined modified artificial neural network (MANN) and mild intrusive genetic algorithm (MIGA) intelligent technique for optimum well characterization and gas lift allocation to achieve maximum economic yield. We seek to characterize the oil wells and get maximum produced oil using limited available gas which indirectly results in increased economic returns. Further details and description of the gas lift problem can be found in [2–4, 6–8].

3. Problem Definition

3.1. Statement of the Problem. An oil field consisting of six wells is considered in this paper. AGL is employed in improving the recovery of the respective wells. Available gases at times are not sufficient to guarantee maximum oil production from the wells while classical techniques also do not guarantee optimal allocation of these available gases; hence an efficient algorithm that can optimally allocate them by selecting the best configuration that can ensure optimum economic accrual is sought. This paper proposes the MANN-MIGA approach for this problem.

3.2. Modelling. This paper addresses a case scenario involving six wells and limited gas supply that is currently not able to guarantee maximum production. We employ modified versions of artificial neural networks (ANN) and GA which has been successfully used in literature (e.g., see [9, 10]). These algorithms seek to optimize the allocation of available gas quantity among the six wells under consideration by allocating quantities that guarantee maximum economic accrual. First, MANN is used to characterise the oil production in terms of B/D (B/D refers to unit of measuring oil production output in terms of barrels per day) and the capacity of each well with respect to the gas injected (in MMscf/D), MMscf/D refers to unit of measuring gases in terms of million standard cubic feet per day. Next, MIGA is used to select the best values for each oil well while observing the following.

(i) Total gas allocated for the six wells does not exceed the available gas quantity for optimization.

(ii) Economic accrual is more important than allocating entire gas and must therefore be maximized.

(iii) Where there are two feasible values for gas allocation with the same oil production, the lowest value is chosen.

The work seeks to address the issue of efficient usage of scarce gas thus preventing wastage and enhancing oil production. In the event of other uses of gas emanating, optimization of the available gas for optimum oil production, maximum economic accrual, and other emanating needs is possible.

The main objective is to develop a strategy for the optimal allocation of limited (and unlimited) gas supply in an oil field involving six wells. Gas allocation is to be minimized while profit from both the sale of oil and remaining unallocated gas is to be maximized by maximizing oil production. We assume full dispatch of gas thus neglecting interaction between gas, conveying media, and other intermeddling media.

If i represents the index of the wells and n is the number of wells, the objective function is formulated as follows

$$\begin{aligned} \text{Minimize gas allocation} &= \sum_{i=1}^{n} \left[G_u(\cdot)\right], \\ \text{Maximize oil production} &= \sum_{i=1}^{n} \left[O_u(\cdot)\right] \\ &= \sum_{i=1}^{n} \Big[a + bG_u(\cdot) \\ &\qquad + cG_u(\cdot)^2 + dG_u(\cdot)^3\Big], \\ \text{Maximize profit from sales} &= P_g \sum_{i=1}^{n} \left[O_u(\cdot)\right] + P_o O_m, \end{aligned} \tag{1}$$

where $G_u(\cdot)$ is gas unit allocated for well I; $O_u(\cdot)$ is respective oil produced for respective gas allocated to well i; a, b, c, and d are constants for derived polynomial function and P_g is unit price of gas ($/$B$); P_o is unit price of oil ($/MMscf); O_m is unallocated gas $= T_a - \sum_{i=1}^{n}[G_u(\cdot)]$; T_a is (un)limited gas available for allocation.

Subject to

(1)

$$\sum_{i=1}^{n} \left[G_u(\cdot)\right] \leq T_a \gamma i; \tag{2}$$

(2) i, if G_i is a or b and $G_i(a) = G_i(b) = O_{i,}$, where O_i is oil produced for $G_i(\cdot)$, if $a < b$ gas allocated $= G_i(a)$.

Also, the equation governing gas allocation (MMscf/D) and oil produced (B/D) is given as

$$Q_{Li} = aQ_{Qi}^2 + bQ_{Qi} + c, \tag{3}$$

where i is well number, Q_{Li} is oil produced per well (in B/D), Q_{Qi} is gas allocated (in MMscf/D), and a, b, and c are constants

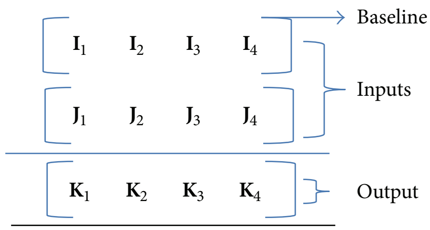


Figure 1: Received inputs and output during training.

Table 1

Well	Value a	Value b	Value c
1	1138.71	799.94	−284.00
2	841.34	893.50	−277.69
3	131.28	61.54	0
4	135.92	39.08	0
5	125.38	49.05	0
6	156.35	89.46	0

The value of the constants a, b, and c for the six wells is given in Table 1.

The values used in training the MANN (Section 4) were obtained from [8].

4. Methodology

As stated earlier, this work adopts combined algorithms of MANN-MIGA which are essential modified forms of ANN and GA. The choice lies in the facts that the underlying techniques have been successfully used in other similar problems as they offer needed efficiency, speed, and flexibility [6, 7]. Moreover, the combination of MANN and MIGA has been efficiently used to solve other problems in literature with promising results [9, 10]. We present below a detailed description of the methods as applied to the current problem.

4.1. Modified Artificial Neural Network. In characterizing the gas injection/oil production of the wells, MANN was used in generating a model. In designing the algorithm for the neural network, a simple regression formula was used. The algorithm receives the inputs, sorts them out, adjusts its parameters, and computes the expected result. A step by step overview of the algorithm is presented as follows.

(1) The inputs are received in the proper (same) dimension (Figure 1).

(2) Interpolation is then carried out on each input matrix and each output matrix using their respective minimum and maximum values in generating an equivalent value of the contribution of each input to the expected answer or result.

(3) Received interpolated input(s) and output(s) are sorted (ranked) concurrently in order to obtain a minimum and maximum value with the first matrix serving as a baseline (Figure 2).

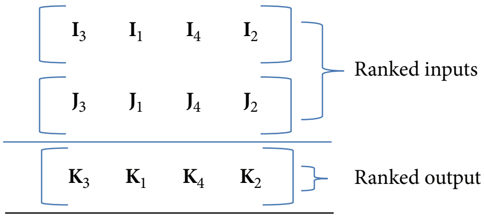


Figure 2: Ranked received inputs and output.

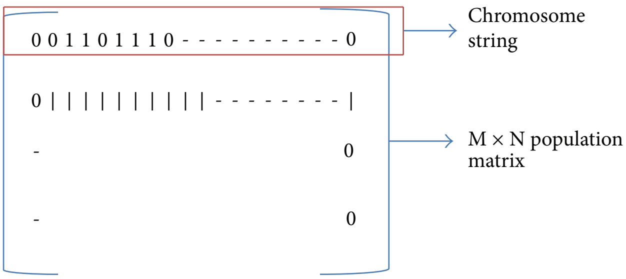


Figure 3: Sample space for MIGA.

(4) Since MATLAB makes use of matrices, received inputs are thus stored in matrices. The output matrix is checked for crossovers (transition from a high-low-high or low-high-low number) using the first input matrix as baseline.

(5) The transitions obtained from the output matrix are used in generating the order of the polynomial function or curve in which our data is to be fitted into (5)–(10). As can be surmised from (5)–(10), there are two inputs and one output (see Figure 1). The modelling of the first input is calculated using (5)–(7) which are used in generating the respective values of a, b, and c. Our first output is thus given as $\mathbf{K}(\mathbf{I}) = a\mathbf{I}^2 + b\mathbf{I} + c$. The second input is modeled similarly using (8)–(10) which are used in generating the respective values of $a1$, $b1$, and $c1$. The second output is thus given as $\mathbf{K}(\mathbf{J}) = a1\mathbf{J}^2 + b1\mathbf{J} + c1$. Our final output is thus the average of these two values given as $\mathbf{K} = (\mathbf{K}(\mathbf{I}) + \mathbf{K}(\mathbf{J}))/2$.

(6) Generated value(s) is/are then recalculated in order to obtain actual values which are then fed forward to the output.

(7) Before the generated values are displayed, they are adjusted for errors using appropriate weights. For the purpose of this analysis, the weights have been assumed to be unity and there is no back propagation network provided to assist in adjusting the displayed values. The inputs therefore must be reliable and fairly accurate.

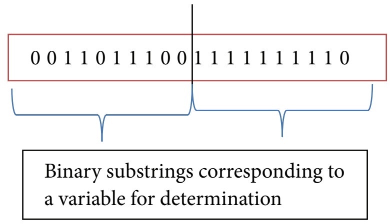


Figure 4

If

$$\begin{aligned} \mathbf{K}_1 &> \mathbf{K}_3 \\ \mathbf{K}_1 &> \mathbf{K}_4 \\ \mathbf{K}_2 &< \mathbf{K}_4, \end{aligned} \tag{4}$$

where $\mathbf{I}_i$, $\mathbf{J}_i$, and $\mathbf{K}_i$ $(i = 1, 2, 3, 4)$ are valid inputs and output, respectively, then the number of crossovers could be determined from the ranked output. From (1) therefore, the ranked output of Figure 2 has 1 crossover.

One crossover therefore signifies that our data could be modeled using a polynomial of order 2 (i.e., a quadratic equation). In modelling therefore the following equations are used:

$$\sum \mathbf{K}_i = a * n + b\sum_{i=1}^{n}\mathbf{I}_i + c\sum_{i=1}^{n}\mathbf{I}_i^{2}, \tag{5}$$

$$\sum \mathbf{K}_i\mathbf{I}_i = a\sum_{i=1}^{n}\mathbf{I}_i + b\sum_{i=1}^{n}\mathbf{I}_i^{2} + c\sum_{i=1}^{n}\mathbf{I}_i^{3}, \tag{6}$$

$$\sum \mathbf{K}_i\mathbf{I}_i^{2} = a\sum_{i=1}^{n}\mathbf{I}_i^{2} + b\sum_{i=1}^{n}\mathbf{I}_i^{3} + c\sum_{i=1}^{n}\mathbf{I}_i^{4}, \tag{7}$$

$$\sum \mathbf{K}_i = a1 * n + b1\sum_{i=1}^{n}\mathbf{J}_i + c1\sum_{i=1}^{n}\mathbf{J}_i^{2}, \tag{8}$$

$$\sum \mathbf{K}_i\mathbf{J}_i = a1\sum_{i=1}^{n}\mathbf{J}_i + b1\sum_{i=1}^{n}\mathbf{J}_i^{2} + c1\sum_{i=1}^{n}\mathbf{J}_i^{3}, \tag{9}$$

$$\sum \mathbf{K}_i\mathbf{J}_i^{2} = a1\sum_{i=1}^{n}\mathbf{J}_i^{2} + b1\sum_{i=1}^{n}\mathbf{J}_i^{3} + c1\sum_{i=1}^{n}\mathbf{J}_i^{4}. \tag{10}$$

4.2. Mild Intrusive Genetic Algorithm. MIGA is used in combination with the MANN in arriving at optimum solutions for gas allocation under limited and unlimited conditions. GA is an evolutionary algorithm that mimics the principle of natural selection, reproduction, and survival of the fittest in solving complex optimization problems [11]. It has been widely and successfully used in literature howbeit in modified or improved form from the standard GA [11, 12] when it comes to some complex problems [13–15]. GA has also been applied to difficult problem involving the control of gas pipeline transmission [6, 7].

MIGA is designed as a modified population-based technique. Figure 3 provides a sample space that illustrates the

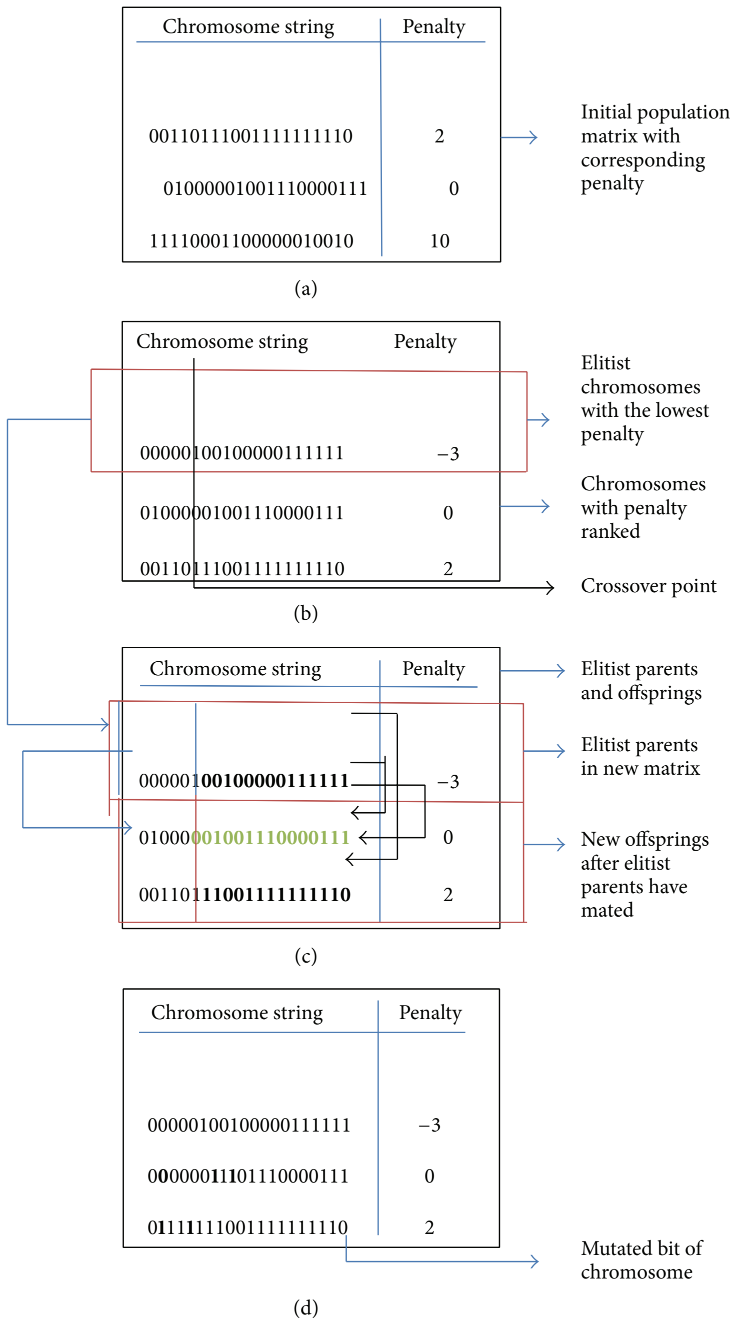


Figure 5: Illustration of MIGA steps.

environment for the activities of MIGA. GA allows for various ways of defining the population structure (chromosome) [14, 16].

In this study, MIGA uses the binary representation. It then runs through the usual steps of selection, crossover, elitism, and mutation as illustrated in Figure 5. Each chromosome string (Figure 4) corresponds to a solution whose fitness is tested for optimality (fitness value of zero). Figure 5(b) shows an instance of a poor solution with fitness value farther from zero.

5. Simulation Experiment, Results, and Discussions

5.1. Simulations Environments. Simulation experiment for this work was done with MATLAB R2009a. Results were interpreted in form of graphs (as shown in Figures 6–13) and tables (as depicted in Tables 2 and 3) generated within this environment. The figures are grouped sometimes into three groups based on the well allocation number as presented and briefly discussed below. Experiment was done on a system

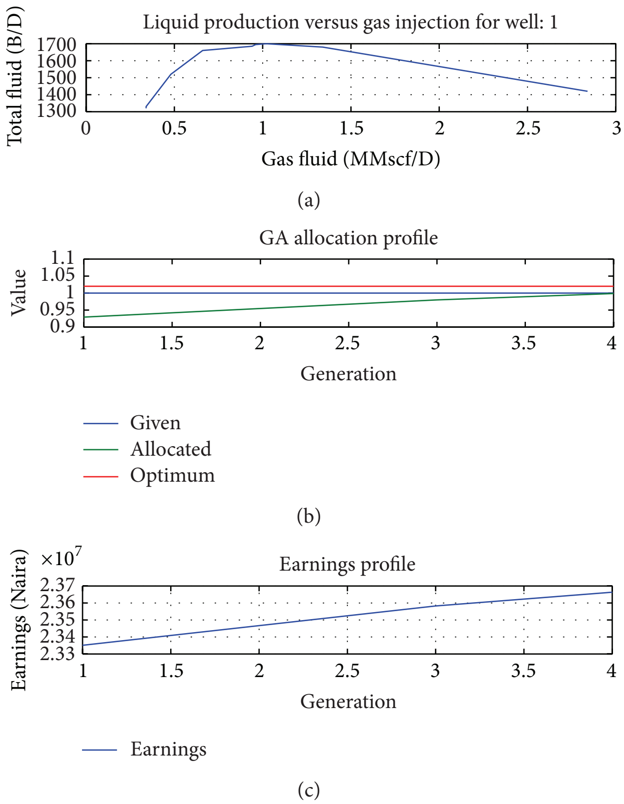


Figure 6: Well 1's allocation profile, GA allocation profile, and the earnings (Naira) profile for a limited scenario.

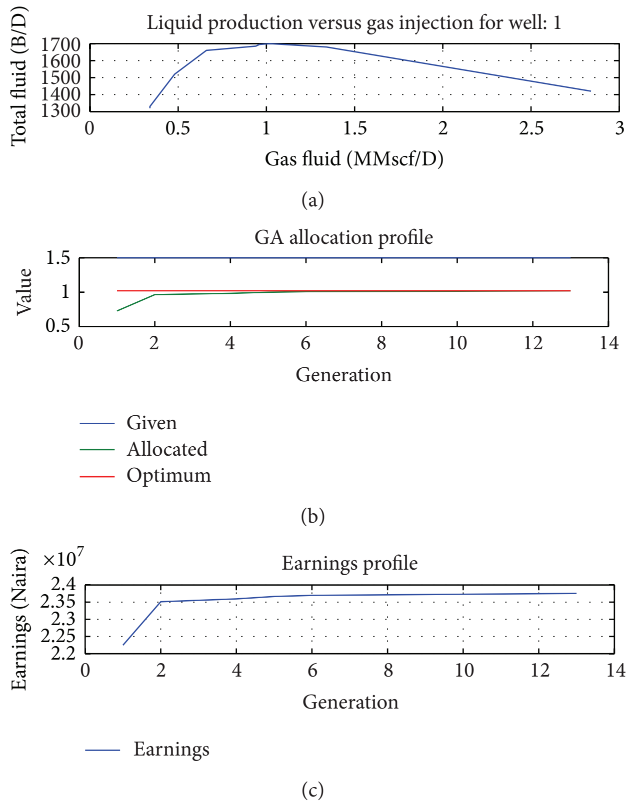


Figure 7: Well 1's allocation profile, GA allocation profile, and earnings (Naira) profile in an unlimited scenario.

with 4 GB DDR3 Memory, 500 GB HDD, and an Intel Core^{TM} i3-380 M Processor.

5.2. Results and Discussion. Based on the application of the proposed methodology in Section 4, results obtained for gas allocation to the six wells under the limited and unlimited scenario are presented mainly in graphical forms as given in Figures 6–13. The behaviour of the proposed MANN-MIGA in effectively characterizing and allocating gas is discussed subsequently. Also discussed is the economic accrual obtained by methodology for different allocation configuration within defined generation run. Tables 2 and 3 show the best allocation values for the limited and unlimited scenario and also the average allocation.

Figure 6(a) displays the curve establishing the relationship between the allocated gas (in MMscf/D) and the oil produced (in B/D). It will be noted in this figure that the allocation of gases does not exceed the maximum point on the curve as already defined in given constraints otherwise it would have led to wastage since increasing allocation beyond the maximum point yields reduced oil production. The curve displayed in Figure 6(a) also echoes with the curves displayed in Figures 7(a), 8(a), 9(a), 10(a), 11(a), 12(a), and 13(a). The curves establishing the gas allocated versus oil produced for wells 1, 2, 3, and 6 are displayed in Figures 6(a) and 7(a), Figures 8(a) and 9(a), Figures 10(a) and 11(a), and Figures 11(a) and 12(a), respectively. It is observed that results illustrated in these graphs satisfy the expected constraints; that is, the allocation of gases does not exceed the value corresponding to the maximum oil that could be produced from that well.

In an attempt to allocate gas under the limited scenario in order to achieve maximum economic yield, Figures 6(b), 8(b), 10(b), and 12(b) show how best gas was distributed among wells 1, 2, 3, and 6, respectively. The available gas in each of the considered scenarios was less than the sum of the gas value(s) corresponding to peak oil production for the number of wells under consideration. The MANN-MIGA results shown in Figures 6(b), 8(b), 10(b), and 12(b) in allocating gas present the obtained best configuration that ensures maximum economic yield. The economic yield being referred to includes the combined accrual from sales of the respective oil produced and the unallocated gas using current market values.

Similarly, the combined MANN-MIGA was also used in allocating gas during an unlimited scenario (where the available gas for allocation exceeds the combined sum of the gas value(s) corresponding to peak oil production for the number of wells under consideration). Figures 7(b), 9(b), 11(b), and 13(b) show the allocation profiles for 1, 2, 3, and 6 wells during an unlimited scenario. Figures 7(b), 9(b), 11(b), and 13(b) display the earnings based on the MANN-MIGA values and current market values for gas and oil.

Table 2: Limited gas optimization values for different well combinations.

Well number	1	1-2	1–3	1–6
Optimum value (MMscf/D)	1.02	2.16	3.29	7.71
Given value (MMscf/D)	1	2	3	5
Allocated (MMscf/D)	Well 1: 0.999	Well 1: 0.9666 Well 2: 1.0324	Well 1: 0.96933 Well 2: 1.0285 Well 3: 1.0018	Well 1: 0.75867 Well 2: 0.78643 Well 3: 0.75565 Well 4: 0.50765 Well 5: 1.2298 Well 6: 0.96
Unallocated (MMscf/D)	0.001	0.001	0.00037	0.0018
% allocation	99.9	99.9	99.99	99.96

Table 3: Unlimited gas optimization values for different well combinations.

Well number	1	1-2	1–3	1–6
Optimum value (MMscf/D)	1.02	2.16	3.29	7.71
Given value (MMscf/D)	1.5	3	5	9
Allocated (MMscf/D)	Well 1: 1.02	Well 1: 1.0173 Well 2: 1.14	Well 1: 0.996 Well 2: 1.14 Well 3: 1.13	Well 1: 0.81467 Well 2: 1.0593 Well 3: 1.0592 Well 4: 1.99 Well 5: 1.4112 Well 6: 0.99554
Unallocated (MMscf/D)	nil	0.0027	0.024	1.38
% allocation	100	99.88	99.27	82.1

In summary, four scenarios were considered for two conditions in each scenario, that is, limited and unlimited gas availability. The first scenario involved well 1 only, second scenario involved wells 1 and 2 only, and third scenario involved 3 wells, 1, 2, and 3 only, while fourth scenario involved the six wells. In each scenario, MANN-MIGA was tested for both limited and unlimited gas (in MMscf/D) availability. Figures 6(b), 8(b), 10(b), and 12(b) show the allocation profile of MANN-MIGA in the limited scenario. As can be observed, gas quantity available for allocation, that is, given quantity (blue line), is less than the optimum value (red line) needed for maximum oil production. The allocation profile (green line) traces the given profile (blue line) in maximizing oil production. It should be noted that, in maximizing oil production and earnings (as shown in Figures 6, 8, 10, and 12), the criticality of the allocation configuration cannot be overemphasized. In essence, the configuration with the highest earnings is chosen.

Similarly, MANN-MIGA was also tested for the unlimited condition. As can be observed, Figures 7(b), 9(b), 11(b), and 13(b) show the allocation profile under this condition. A further observation of the aforelisted figures shows that the given quantity (blue line) is greater than the optimum value (red line) needed for maximum oil production. Gas allocation (green line) is thus expected to follow the optimum value (red line) in generating maximum oil production values. The earning profiles for the different considered scenario under the unlimited condition are shown in Figures 7(c), 9(c), 11(c), and 13(c). In generating allocation configuration values and earnings, an oil price of ₦14, 000/B (Naira (₦) is the currency for Nigeria where the work was carried out and based. The displayed value is in terms of Naira per barrel) and gas price of ₦400/MMscf were used.

The subsequent tables show the new values obtained in both the limited and unlimited scenarios. A critical observation of Figures 6(b), 8(b), 10(b), and 12(b) for the limited scenario and Figures 7(b), 9(b), 11(b), and 13(b) for the unlimited scenario reveals some deviations. As presented in Table 2, the MANN-MIGA under the limited scenario has a better allocation (about 99.94 on average) compared to the MANN-MIGA allocation values presented in Table 3 during the unlimited scenario (about 95.31 on average). As can be observed from Tables 2 and 3, three allocation terms are presented as further shown in the earlier considered figures. They are the optimum, given, and allocated values. The optimum value shows the maximum value that can be allocated to the well(s) under consideration guaranteeing the maximum oil that can be produced. The given value denotes the scenario under consideration, limited or unlimited, while the allocated value describes the ability of the MANN-MIGA in optimally distributing the given value for optimal economic/monetary yield. Our algorithm is therefore able to characterize the wells using a least square approximation method embedded into MANN and also allocate gas to the well combinations under consideration during limited (over 99% on average) and unlimited scenario (over 95% on

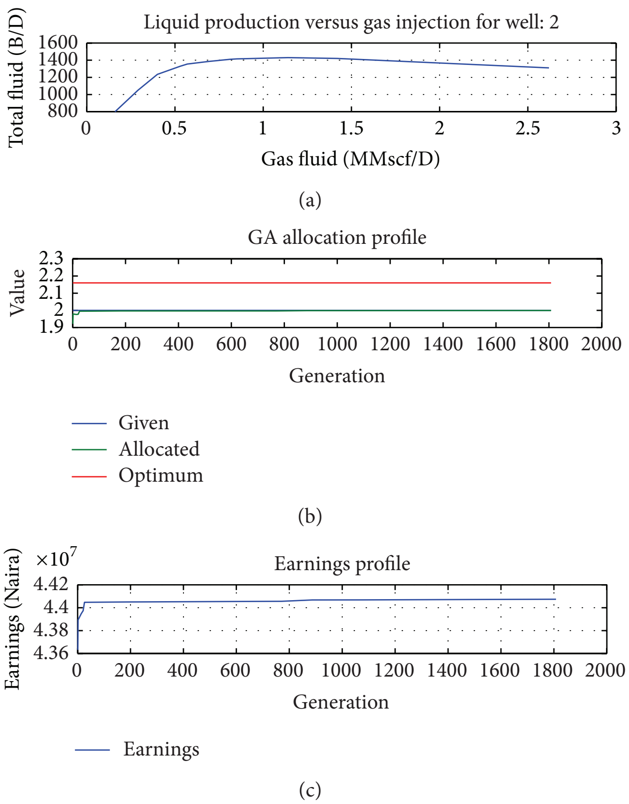


FIGURE 8: Well 2's allocation profile, wells 1-2 GA allocation profile, and wells 1-2 earnings (Naira) for a limited scenario.

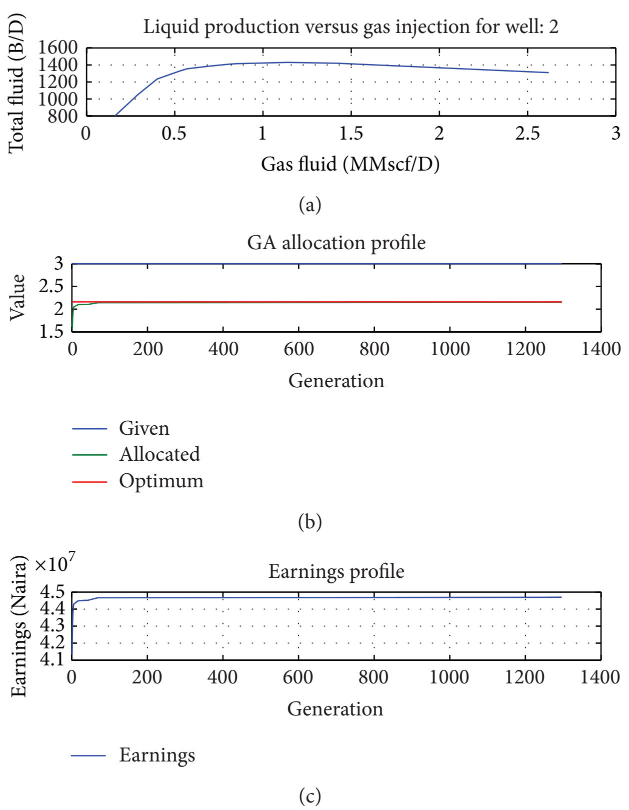


FIGURE 9: Well 2's allocation profile, wells 1-2 GA allocation profile, and wells 1-2 earnings (Naira) in an unlimited scenario.

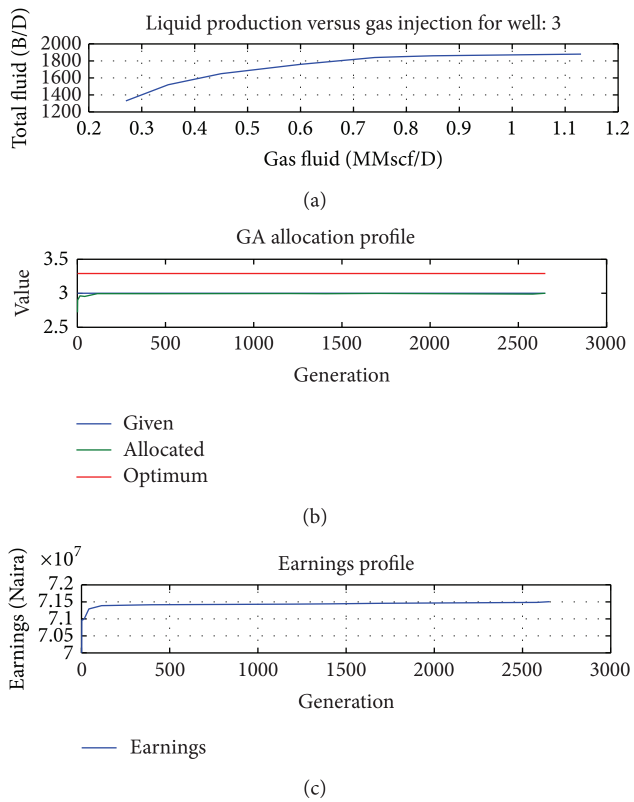


FIGURE 10: Well 3's allocation profile, wells 1–3 GA allocation profile, and earnings (Naira) in a limited scenario.

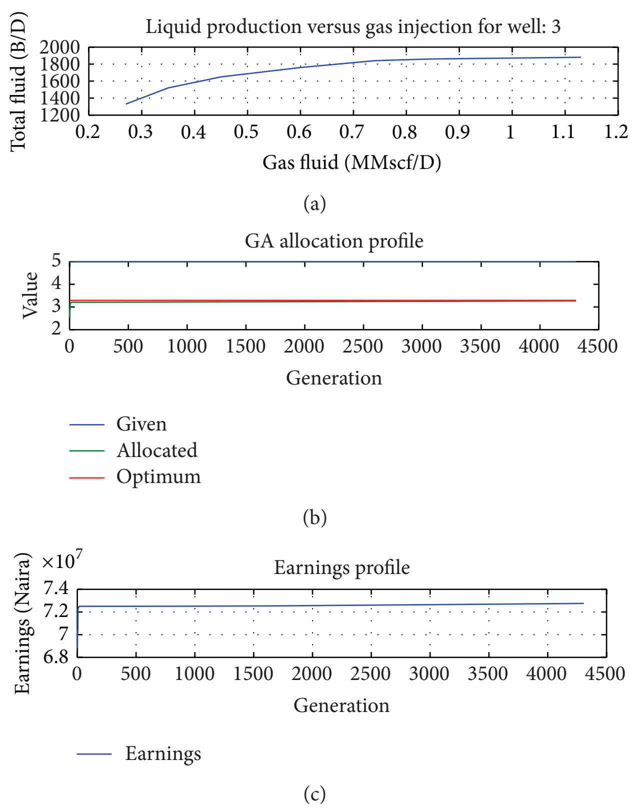


FIGURE 11: Well 3's allocation profile, wells 1–3 GA allocation profile, and earnings (Naira) in an unlimited scenario.

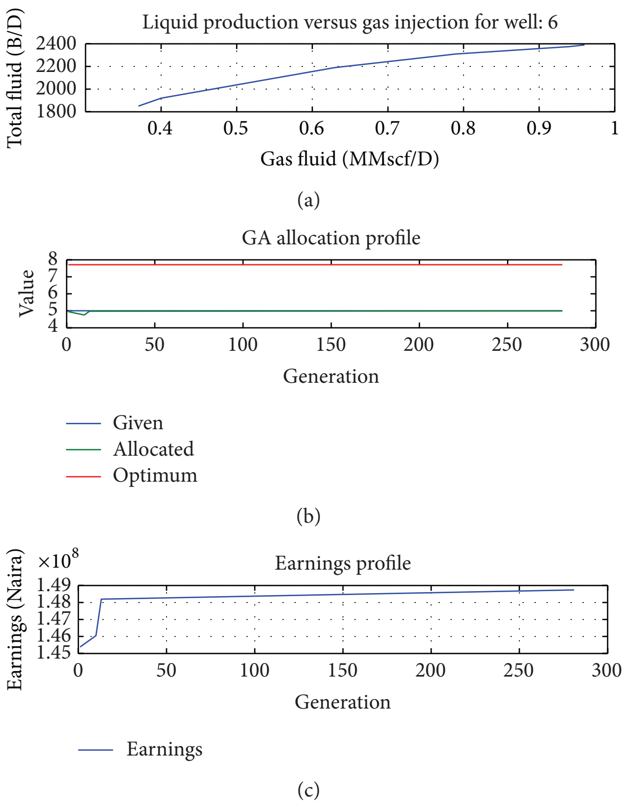


FIGURE 12: Well 6's allocation profile, wells 1–6 GA allocation profile, and earnings (Naira) for a limited scenario.

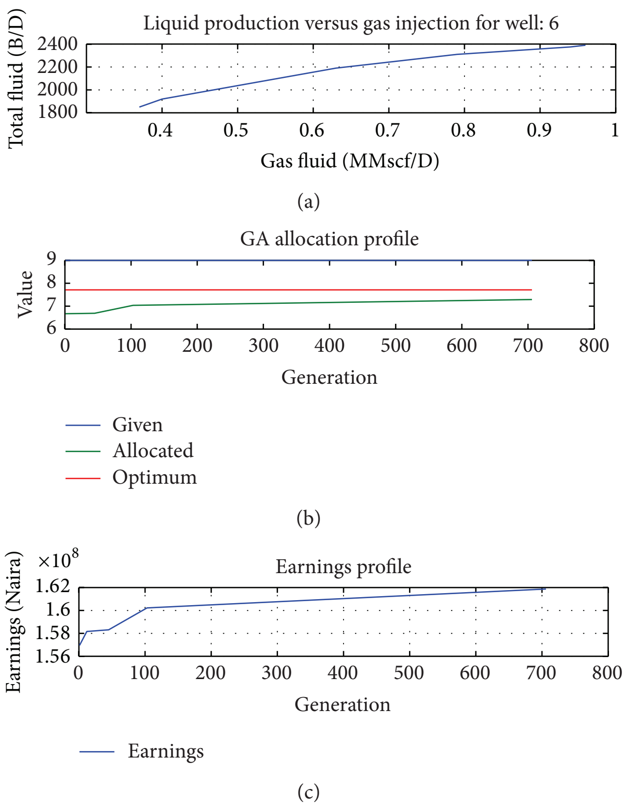


FIGURE 13: Well 6's allocation profile, wells 1–6 GA allocation profile, and earnings (Naira) for an unlimited scenario.

average) using the MIGA as a stochastic optimization tool. The algorithm provided consistent results on multiple runs and at a fast run time.

6. Conclusion and Future Work

This paper presents a characterization of oil wells and gas allocation in both the limited and unlimited scenarios such as presenting a cost-effective means for gas allocation for the oil and gas industry. While earlier works considered the characterization and allocation separately, this work has been able to both characterize and use generated values from the characterization in allocating gas. A mathematical model is obtained for characterizing the gas allocated and oil produced for the wells under consideration. Wells 1 and 2 were characterized using a quadratic equation while the rest were subsequently characterized using linear curve fitting techniques. A combined MANN-MIGA was adapted and applied extensively in evolving a relationship between the gas allocated and oil produced for the six wells considered. These values have been further used in computing the optimum gas allocation per well. The values generated show a remarkable allocation by our approach. The mild intrusive property of our GA arises from its ability to allocate the least possible gas value for maximum oil production. The approach neglects gas values that exceeded the maximum for optimal oil production. The MIGA was then used in generating optimized values which met given conditions and yielded improved economic returns. The MANN-MIGA thus proves useful to the oil and gas industry as it not only provides characterized equations, but also allocates gas based on preset conditions.

Future work might consider further fine-tuning of the algorithm for improved performance especially in optimizing gas allocation under the unlimited scenario. Also, other natured-inspired techniques like variants particle swarm optimization [17] and other recent stochastic algorithms [18] can be investigated for comparative analysis.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] A. Codas and E. Camponogara, "Mixed-integer linear optimization for optimal lift-gas allocation with well-separator routing," *European Journal of Operational Research*, vol. 217, no. 1, pp. 222–231, 2012.

[2] T. Ray and R. Sarker, "Genetic algorithm for solving a gas lift optimization problem," *Journal of Petroleum Science and Engineering*, vol. 59, no. 1-2, pp. 84–96, 2007.

[3] E. Camponogara and P. H. R. Nakashima, "Solving a gas-lift optimization problem by dynamic programming," *European Journal of Operational Research*, vol. 174, no. 2, pp. 1220–1246, 2006.

[4] J. N. M. de Souza, J. L. de Medeiros, A. L. H. Costa, and G. C. Nunes, "Modeling, simulation and optimization of continuous

gas lift systems for deepwater offshore petroleum production," *Journal of Petroleum Science and Engineering*, vol. 72, no. 3-4, pp. 277–289, 2010.

[5] M. Mahmudi and M. T. Sadeghi, "The optimization of continuous gas lift process using an integrated compositional model," *Journal of Petroleum Science and Engineering*, vol. 108, pp. 321–327, 2013.

[6] R. Sharma, K. Fjalestad, and B. Glemmestad, "Optimization of lift gas allocation in a gas lifted oil field as non-linear optimization problem," *Modeling, Identification and Control*, vol. 33, no. 1, pp. 13–25, 2012.

[7] O. G. Santos, S. N. Bordalo, and F. J. S. Alhanati, "Study of the dynamics, optimization and selection of intermittent gas-lift methods—a comprehensive model," *Journal of Petroleum Science and Engineering*, vol. 32, no. 2–4, pp. 231–248, 2001.

[8] E. P. Kanu, J. Mach, and K. E. Brown, "Economic approach to oil production and gas allocation in continuous gas lift," *Journal of Petroleum Technology*, vol. 33, no. 10, pp. 1887–1892, 1981.

[9] A. Ibitola and C. G. Monyei, "Genetic algorithm for students allocation to halls of residence using energy consumption as discriminant adaptive computing," *Information Systems, Development Informatics & Allied Research Journal*, vol. 3, no. 4, pp. 31–36.

[10] C. G. Monyei and O. A. Fakolujo, "Optimized enhanced control system for the Unibadan's virtual power plant project using genetic algorithm," *African Journal of Computing & ICT*, vol. 6, no. 4, pp. 16–20, 2013.

[11] D. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, Mass, USA, 1989.

[12] A. O. Adewumi, B. A. Sawyerr, and M. M. Ali, "A heuristic solution to the university timetabling problem," *Engineering Computations*, vol. 26, no. 8, pp. 972–984, 2009.

[13] A. R. Yildiz and F. Ozturk, "Hybrid enhanced genetic algorithm to select optimal machining parameters in turning operation," *Proceedings of the Institution of Mechanical Engineers B: Journal of Engineering Manufacture*, vol. 220, no. 12, pp. 2041–2053, 2006.

[14] T. C. Nwaoha, Z. Yang, J. Wang, and S. Bonsall, "Application of genetic algorithm to risk-based maintenance operations of liquefied natural gas carrier systems," *Proceedings of the Institution of Mechanical Engineers E: Journal of Process Mechanical Engineering*, vol. 225, no. 1, pp. 40–52, 2011.

[15] P. Guo, X. Wang, and Y. Han, "The enhanced genetic algorithms for the optimization design," in *Proceedings of the 3rd International Conference on BioMedical Engineering and Informatics (BMEI '10)*, pp. 2990–2994, October 2010.

[16] A. O. Adewumi and M. M. Ali, "A multi-level genetic algorithm for a multi-stage space allocation problem," *Mathematical and Computer Modelling*, vol. 51, no. 1-2, pp. 109–126, 2010.

[17] M. A. Arasomwan and A. O. Adewumi, "An adaptive velocity particle swarm optimization for high-dimensional function optimization," in *Proceedings of the IEEE Congress on Evolutionary Computation (CEC '13)*, pp. 2352–2359, 2013.

[18] S. Chetty and A. O. Adewumi, "Three new stochastic local search algorithms for continuous optimization problems," *Computational Optimization and Applications*, vol. 56, no. 3, pp. 675–721, 2013.

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 972487, 13 pages
http://dx.doi.org/10.1155/2014/972487

Research Article

Precommitted Investment Strategy versus Time-Consistent Investment Strategy for a Dual Risk Model

Lidong Zhang,[1,2] Ximin Rong,[3] and Ziping Du[4]

[1] *School of Management, Tianjin University, Tianjin 300072, China*
[2] *School of Science, Tianjin University of Science & Technology, Tianjin 300457, China*
[3] *School of Science, Tianjin University, Tianjin 300072, China*
[4] *College of Economics & Management, Tianjin University of Science Technology, Tianjin 300222, China*

Correspondence should be addressed to Ximin Rong; rongximin@tju.edu.cn

Received 23 February 2014; Accepted 30 April 2014; Published 19 May 2014

Academic Editor: Xiang Li

Copyright © 2014 Lidong Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We are concerned with optimal investment strategy for a dual risk model. We assume that the company can invest into a risk-free asset and a risky asset. Short-selling and borrowing money are allowed. Due to lack of iterated-expectation property, the Bellman Optimization Principle does not hold. Thus we investigate the precommitted strategy and time-consistent strategy, respectively. We take three steps to derive the precommitted investment strategy. Furthermore, the time-consistent investment strategy is also obtained by solving the extended Hamilton-Jacobi-Bellman equations. We compare the precommitted strategy with time-consistent strategy and find that these different strategies have different advantages: the former can make value function maximized at the original time $t = 0$ and the latter strategy is time-consistent for the whole time horizon. Finally, numerical analysis is presented for our results.

1. Introduction

We start from a dual risk model; namely, the surplus process of a company is given by

$$R(t) = x + \sum_{j=1}^{N(t)} Z_j - ct, \quad (1)$$

where x is the initial capital and c is the rate of expenses; the positive incomes or profits Z_j $(j = 1, 2, 3, \ldots)$ arrive as a Poisson process $\{N(t)\}_{t\geq 0}$ with intensity λ; Z_j $(j = 1, 2, 3, \ldots)$ are independent and identically distributed (i.i.d) nonnegative random variables with the first and second moments μ_z and σ_z^2. The expected increase of the surplus per unit time satisfies the positive loading condition: $p = \lambda\mu_z - c > 0$. The model is called dual as opposed to the Cramer-Lundberg risk model with applications to insurance.

The dual risk model can be described as the amount of capital of a company engaging in research and development. The company pays expenses and gets occasional revenues. Revenues are interpreted as the values of future gains from an invention or discovery, while the decrease of surplus can represent costs of production, payments to employees, maintenance of equipment, and so forth. This is also the case for a portfolio of life annuities, where the risk consists of survival and the event death leads to gains. Furthermore, many scholars investigated the dual risk model under different criterions. Stochastic control theory and the related methodologies are the main tools for finding the related minimum probability of ruin or the maximum value for the expected utility of terminal wealth or the maximum value for the expected present value of the dividends minus capital injections. For related works see, for example, Avanzi et al. [1, 2], Zhu and Yang [3], Dai et al. [4], and Yao et al. [5] and references therein.

In this paper, we are concerned with optimal investment strategy for the dual risk model under mean-variance criterion. Our objective is to find the optimal investment strategy such that the expected terminal wealth is maximized and the variance of the terminal wealth is minimized. According to the biobjective optimization theory, the alternative objective is to find a strategy which maximizes the expected terminal wealth minus the variance of the terminal wealth; namely,

$$\text{(MV)} \begin{cases} E_{0,x}(X(T)) - \dfrac{\gamma}{2}\operatorname{Var}_{0,x}(X(T)) \longrightarrow \max \\ l \in U, \end{cases} \quad (2)$$

where γ is a prespecified risk aversion coefficient, $E_{0,x}[\cdot] = E[\cdot \mid X^l(0) = x]$, and $\operatorname{Var}_{0,x}[\cdot] = \operatorname{Var}[\cdot \mid X^l(0) = x]$.

It is well known that this criterion lacks the iterated-expectation property, so problem (MV) is time-inconsistent which means that Bellman Optimality Principle does not hold. The lack of time consistency leads to conceptual as well as computational problems. Recently there are two approaches to the conceptual problem.

(i) Fix one initial point $(0, x_0)$ and then try to find a strategy l^* which maximizes $E_{0,x_0}(X(T)) - \gamma \operatorname{Var}_{0,x_0}(X(T))$. We then simply disregard the fact that at later points in time the strategy will not be optimal. Such a strategy is known as a precommitted strategy. Zhou and Li [6] and Li and Ng [7] made big breakthrough works and proposed an embedding method to derive the precommitted strategy for problem (MV).

(ii) We take the time-inconsistency seriously and formulate the problem within a game theoretic framework. Problem (MV) is viewed as a game, where the players are the future incarnations of our own preferences and Nash equilibrium points can be derived in order to address the general time-inconsistency. See more references in Björk and Murgoci [8], Ekeland and Lazrak [9], and Kryger and Steffensen [10].

Thus, problem (MV) can be reduced to a resolvable problem by virtue of some techniques including the embedding technique [6, 7] and the game theoretical technique [8]. Our contributions are as follows. Firstly, we take three steps to deal with the precommitted investment problem. This method is different from the embedding technique in Zhou and Li [6] and Li and Ng [7]. Secondly, the time-consistent strategy from the game theoretical perspective is also derived and we also present economics implications and sensitivity analysis for our results.

This paper is organized as follows. Section 2 formulates problem (MV) and gives the main results. Section 3 provides numerical analysis for our results.

2. Model and Main Results

This section starts with a filtered complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\le t\le T}, \mathcal{P})$, where T represents the fixed time horizon and $\mathcal{F}_t$ stands for the information available at time t. We recall that the dual risk model and the surplus process of the company are subject to the following stochastic process:

$$R(t) = x + \sum_{j=1}^{N(t)} Z_j - ct. \quad (3)$$

The company invests its wealth into a Black-Scholes market consisting of a risk-free asset and a risky asset. Denote the amount of money invested in the risky asset at time t by $l(t)$, and the allowance of short-selling and borrowing money implies $l(t) \in \mathbb{R}$. The price of the risk-free asset $S_0(t)$ is modeled by

$$dS_0(t) = r_0(t) S_0(t)\,dt, \quad S_0(0) = s_0 \quad (4)$$

and the price of the risky asset $S_1(t)$ evolves according to the dynamics

$$dS_1(t) = S_1(t)\left(r_1(t)\,dt + \sigma(t)\,dW(t)\right), \quad S_1(0) = s_1, \quad (5)$$

where $W(t)$ is a standard Brownian motion which is independent of $\sum_{j=1}^{N(t)} Z_j$, the appreciation rate $r_1(t)$ and risk-free rate $r_0(t)$ satisfy $r_1(t) > r_0(t)$, $\sigma(t)$ is the volatility coefficient, and they are all bounded continuous functions. Denote the resulting surplus process after incorporating strategy l into (3) by $X^l(t)$, and $X^l(t)$ is subject to the following stochastic equation:

$$\begin{aligned} dX^l(t) &= \left(r_0(t) X^l(t) + r(t) l(t) - c\right) dt \\ &\quad + \sigma(t) l(t)\,dW(t) + d\sum_{j=1}^{N(t)} Z_j, \end{aligned} \quad (6)$$

where $r(t) = r_1(t) - r_0(t)$. Denote that $C^{1,2}(Q) = \{\phi(t,x) \mid \phi(t,\cdot)$ is once continuously differentiable on $[0,T]$ and $\phi(\cdot,x)$ is twice continuously differentiable on $\mathbb{R}\}$.

According to the theory of stochastic calculus, the infinitesimal operator of the surplus process $X^l(t)$ is

$$\begin{aligned} \mathcal{A}\phi(t,x) &= \phi_t(t,x) + \phi_x(t,x)\left(r_0(t)x - c + r(t)l(t)\right) \\ &\quad + \frac{1}{2}\phi_{xx}(t,x)\, l^2(t)\,\sigma^2(t) \\ &\quad + \lambda\left[E\phi(t,x+Z) - \phi(t,x)\right], \\ &\quad \text{for } \phi(t,x) \in C^{1,2}(Q). \end{aligned} \quad (7)$$

In order to solve problem (MV), we firstly give the definition of admissible strategy for the dual risk process $X^l(t)$.

Definition 1. A strategy $l = \{l(t), t \ge 0\}$ is said to be admissible if

(1) $l : [0,\infty] \times \Omega \rightarrow \mathbb{R}$ is an $\mathcal{F}$-adapted process;

(2) $l(t)$ satisfies the integrability condition: $E\int_0^t l^2(t)ds < \infty$ almost surely, for all $t \ge 0$;

(3) SDE (6) has a unique solution corresponding to l.

Denote the set of all admissible strategies over the time horizon $[0,T]$ by U. Our aim is to find a strategy which maximizes the expected terminal wealth minus the variance of the terminal wealth; in other words, we aim to solve problem (MV). In the next subsection we will give the optimal precommitted investment strategy and the time-consistent strategy.

2.1. The Optimal Precommitted Investment Strategy for Problem (MV). This subsection puts forward an idea to solve problem (MV).

Firstly, problem (MV) can be transformed into a constrained problem. For $d \in \mathbb{R}$, define the following problem with a constrained expectation of terminal wealth:

$$(\mathrm{CMV}) \begin{cases} E_{0,x}\left(X^{l}(T)\right) - \frac{\gamma}{2}\mathrm{Var}_{0,x}\left(X^{l}(T)\right) \longrightarrow \max \\ E_{0,x}\left(X^{l}(T)\right) = d, \\ l \in U. \end{cases} \tag{8}$$

Considering the terminal condition $E_{0,x}(X^{l}(T)) = d$ sufficiently, problem (CMV) equals the following problem:

$$\left(\overline{\mathrm{CMV}}\right) \begin{cases} -\frac{\gamma}{2}E_{0,x}\left(X^{l}(T)\right)^{2} + d + \frac{\gamma}{2}d^{2} \longrightarrow \max \\ E_{0,x}\left(X^{l}(T)\right) = d, \\ l \in U. \end{cases} \tag{9}$$

$V(x,d)$ and $V(x)$ represent the value functions for problem $(\overline{\mathrm{CMV}})$ and problem (MV), respectively, and they have the following relationship $V(x) = \sup_{d\in R} V(x,d)$.

Secondly, problem $(\overline{\mathrm{CMV}})$ is solved by Lagrange technique. By introducing a Lagrange multiplier $y := y(x,d) \in \mathbb{R}$, define a quadratic problem without the constrained expectation of the terminal wealth:

$$(\mathrm{BM}) \begin{cases} E_{0,x}\left[yX^{l}(T) - \frac{\gamma}{2}\left(X^{l}(T)\right)^{2}\right] + d \\ \quad + \frac{\gamma}{2}d^{2} - dy \longrightarrow \max \\ l \in U; \end{cases} \tag{10}$$

$V(x,d,y)$ represents the value function for problem (BM). The duality theory implies that the value function for problem $(\overline{\mathrm{CMV}})$ satisfies that $V(x,d) = \inf_{y\in\mathbb{R}} V(x,d,y)$; then we have

$$V(x) \sup_{d\in R} \inf_{y\in\mathbb{R}} V(x,d,y). \tag{11}$$

We will take three steps to derive the precommitted investment strategy for problem (MV). Based on the discussion in Appendix A, we can derive the precommitted investment strategy and the value function.

Theorem 2. *For problem (MV), the optimal precommitted investment strategy is given by*

$$l^{*}(t) = \frac{r(t)}{\sigma^{2}(t)} \left\{ e^{-\int_{t}^{T} r_0(s)ds} \left(p\int_{0}^{t} e^{\int_{s}^{T} r_0(v)dv} ds + xe^{\int_{0}^{T} r_0(s)ds} + \frac{e^{\int_{0}^{T}\xi(s)ds}}{\gamma} \right) - x(t) \right\}, \tag{12}$$

and the value function is given by

$$V(x) = xe^{\int_{0}^{T} r_0(s)ds} + p\int_{0}^{T} e^{\int_{s}^{T} r_0(u)du} ds - \frac{\gamma\lambda\sigma_z^{2}}{2}\int_{0}^{T} e^{\int_{s}^{T}(2r_0(u)-\xi(u))du} ds + \frac{\left(e^{\int_{0}^{T}\xi(s)ds} - 1\right)}{2\gamma}. \tag{13}$$

Furthermore, the efficient frontier of problem (MV) at initial state $(0,x)$ is given by

$$E_{0,x}\left(X^{l^{*}}(T)\right) = p\int_{0}^{T} e^{\int_{s}^{T} r_0(v)dv} ds + xe^{\int_{0}^{T} r_0(s)ds} + \sqrt{\left[\mathrm{Var}_{0,x}\left(X^{l^{*}}(T)\right) - \lambda\sigma_z^{2}\int_{0}^{T} e^{\int_{s}^{T}(2r_0(u)-\xi(u))du} ds\right]\left(e^{\int_{0}^{T}\xi(s)ds} - 1\right)}. \tag{14}$$

Remark 3. Our method is different from the embedding technique proposed by Zhou and Li [6]. Problem (MV) can be embedded into a class of auxiliary stochastic linear-quadratic (LQ) problems, and the precommitted strategy for problem (MV) was derived by solving the LQ problems. Correspondingly, we take three steps to derive the precommitted strategy. Compared with the embedding method, our method is simple but easy to implement for solving problem (MV).

2.2. The Time-Consistent Investment Strategy for Problem (MV). This subsection provides the time-consistent investment strategy for problem (MV). Firstly, define a dynamic problem

$$\left(\overline{\mathrm{MV}}\right) \begin{cases} J(t,x,l) = E_{t,x}\left[X^{l}(T)\right] \\ \quad -\frac{\gamma}{2}\mathrm{Var}_{t,x}\left[X^{l}(T)\right] \longrightarrow \max \\ l \in U. \end{cases} \tag{15}$$

According to the theory stated in Björk and Murgoci [8], we can convert this time-inconsistent problem into a time-consistent problem. The equilibrium strategy is defined by the same way in Björk and Murgoci [8] or Zeng et al. [11].

Definition 4 (equilibrium strategy). For any fixed chosen initial state $(t, x) \in Q := [0, T] \times R$, consider an admissible strategy $l_*(t, x)$. Choose two fixed real numbers $\widetilde{l} > 0$ and $\varepsilon > 0$ and define the following strategy:

$$l^{\varepsilon}(s, x) = \begin{cases} \widetilde{l}, & \text{for } (s, x) \in [t, t+\varepsilon] \times \mathbb{R} \\ l_*(s, x), & \text{for } (s, x) \in [t+\varepsilon, T] \times \mathbb{R}. \end{cases} \tag{16}$$

If

$$\liminf_{\epsilon \to 0} \frac{J(t, x, l_*) - J(t, x, l^{\varepsilon})}{\varepsilon} \geq 0, \quad \forall \widetilde{l} \in \mathbb{R}^+, \ (t, \ x) \in Q, \tag{17}$$

then $l_*(t, x)$ is called an equilibrium strategy and the corresponding equilibrium value function $\psi(t, x)$ for problem $(\overline{\text{MV}})$ is given by

$$\psi(t, x) = J(t, x, l_*) = E_{t,x}\left[X^{l_*}(T)\right] - \frac{\gamma}{2}\text{Var}_{t,x}\left[X^{l_*}(T)\right]. \tag{18}$$

By Definition 4, we know that the equilibrium strategy is time-consistent. So the equilibrium strategy $l_*(t)$ and the equilibrium value function $\psi(t, x)$ are called optimal time-consistent strategy and optimal equilibrium value function for problem $(\overline{\text{MV}})$, respectively. It is easy to see that the equilibrium value function $V^*(x)$ for problem (MV) satisfies $V^*(x) = \psi(0, x)$. Based on the discussion in Appendix B, the time-consistent investment strategy and the equilibrium value function for problem (MV) are given by the following theorem.

Theorem 5. *For the dual model, the optimal time-consistent strategy $l_*(t)$ is given by*

$$l_*(t) = \frac{r(t)}{\gamma \sigma^2(t)} e^{-\int_t^T r_0(s) ds} \tag{19}$$

and the optimal equilibrium value function is given by

$$\begin{aligned} V^*(x) &= e^{\int_0^T r_0(s) ds} x + p \int_0^T e^{\int_s^T r_0(u) du} ds \\ &\quad - \frac{\gamma \lambda \sigma_z^2}{2} \int_0^T e^{2\int_s^T r_0(u) du} ds + \frac{1}{2\gamma} \int_0^T \xi(s)\, ds. \end{aligned} \tag{20}$$

Furthermore, the efficient frontier under time-consistent strategy is given by the following equation:

$$\begin{aligned} E_{0,x}\left(X^{l_*}(T)\right) &= e^{\int_0^T r_0(s) ds} x + p \int_0^T e^{\int_s^T r_0(u) du} ds \\ &\quad + \sqrt{\left[\text{Var}_{0,x}\left(X^{l_*}(T)\right) - \lambda \sigma_z^2 \int_0^T e^{2\int_s^T r_0(u) du} ds\right] \int_0^T \xi(s)\, ds}. \end{aligned} \tag{21}$$

Remark 6. The time-consistent strategy is not affected by the initial information and the current wealth which is different from the precommitted strategy. While the time-consistent strategy is time deterministic, the precommitted strategy is stochastically dependent on the current wealth. We exploit Monte Carlo methods to simulate the precommitted investment strategy. We compare the average of 1000 tracks of the precommmitted investment strategy with the time-consistent strategy. Figure 1(a) shows that the time-consistent investment strategy is smaller than the average of the precommitted investment strategy; that is to say, the company should invest more money into the risky asset under the precommitted investment strategy.

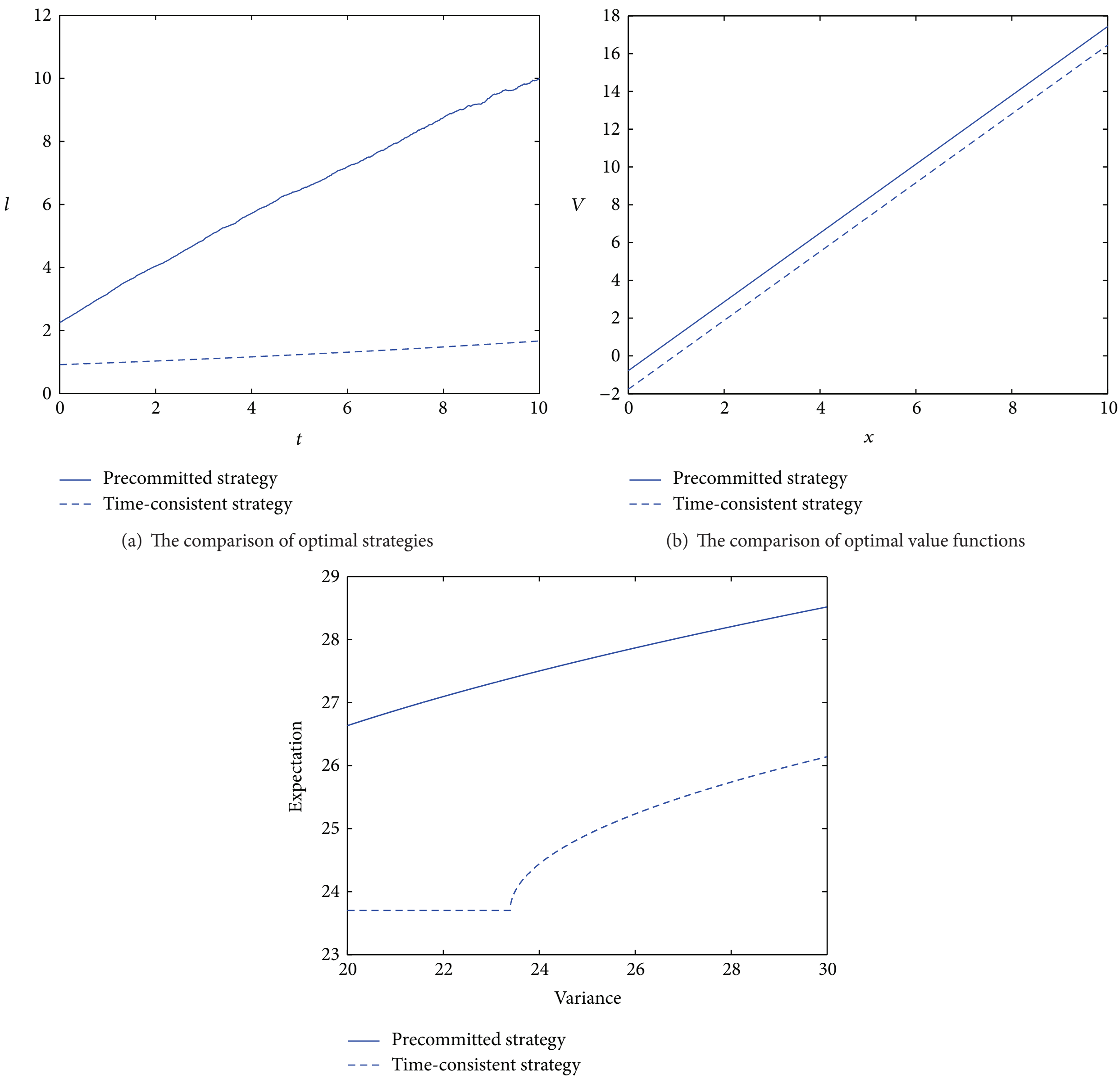
Remark 7. The optimal value function $V(x)$ and the equilibrium value function $V^*(x)$ are not the same for problem (MV). From (13) and (20), it is not hard to calculate that $V(x) > V^*(x)$ which is illustrated by Figure 1(b). It follows from (14) and (21) that the efficient frontiers for problem (MV) are not straight lines but hyperbolas in mean-standard variance plane and that the expectation of terminal wealth under the time-consistent strategy is never bigger than the one under the precommitted strategy for the same variance of terminal wealth which is illustrated by Figure 1(c).

Remark 8. From Remark 7, we can see that the value function $V(x)$ is larger than the equilibrium value function $V^*(x)$ and that the efficient frontier under the time-consistent equilibrium strategy is never above the efficient frontier under the precommitted strategy. But the conclusion that the precommitted strategy is prior to the time-consistent equilibrium strategy is not right, for the precommitted strategy is a global optimal strategy only at $t = 0$ and the time-consistent strategy is a suboptimal strategy for all $t \geq 0$.

3. Numerical Analysis

This section provides some numerical examples to illustrate the effect of parameters on the optimal strategies and the corresponding value functions. For convenience but without loss of generality, all the parameters involved are constants. For the following numerical illustrations, unless otherwise stated, the basic parameters are given by $r_0 = 0.06$, $r_1 = 0.15$, $c = 0.6$, $x = 10$, $\gamma = 0.6$, $\lambda = 1$, $\mu_z = 1$, $\sigma_z = 1.1$, $\sigma = 0.3$, and $T = 10$.

3.1. Analysis of the Precommitted Strategy and the Corresponding Value Function. This subsection will provide numerical



(a) The comparison of optimal strategies

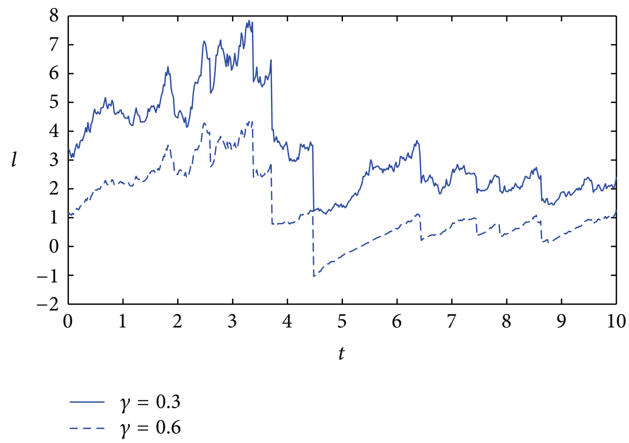
(b) The comparison of optimal value functions

(c) The comparison of different efficient frontiers

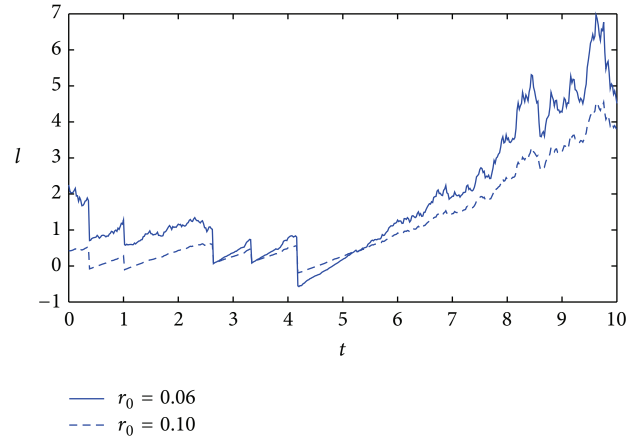
Figure 1: The comparisons between precommitted strategy and time-consistent strategy with parameters $r_0 = 0.06$, $r_1 = 0.15$, $c = 0.6$, $x = 10$, $\gamma = 0.6$, $\lambda = 1$, $\mu_z = 1$, $\sigma_z = 1.1$, $\sigma = 0.3$, and $T = 10$.

examples to show how the parameters effect the precommitted strategy and the value function.

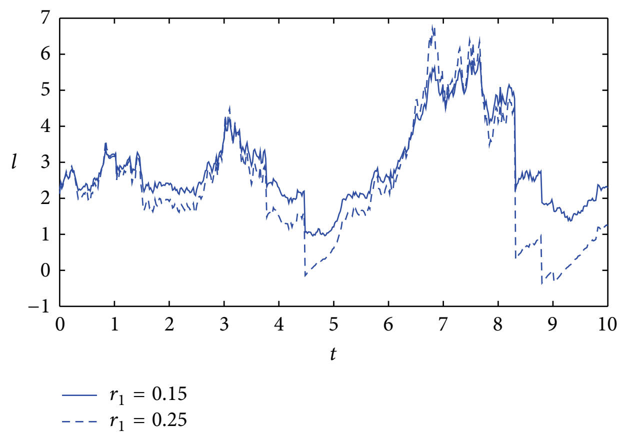
Because the precommitted investment strategy depends on the current wealth, we will investigate the effect of all the parameters for the precommitted strategy in the same sample trajectory by stochastic simulations. In order to model the trajectory, we assume that $\{N(t)\}_{t\geq 0}$ is a Poisson process with intensity $\lambda = 1$ and the profit or the income Z_j ($j = 1, 2, 3, \ldots$) is exponentially distributed with mean 1. From (12), we can see that the precommitted investment strategy increases when the current wealth decreases; namely, if the current wealth is smaller, the company should invest more money in the risky asset. Figure 1 shows how the coefficients involved impact on the optimal precommitted investment strategy. From Figure 2, we can conclude the following: (1) the precommitted investment strategy is decreasing with respect to γ and σ which shows that the more the company dislikes risk or the larger the market's risk is, the less amount the company invests in the risky asset; (2) the precommitted investment strategy is also decreasing with respect to c which shows that the more the company's expenditure is, the less amount the company invests in the risky asset; (3) the precommitted investment strategy has a more complex relation with r_1 and r_0 because the increase of r_1 or r_0



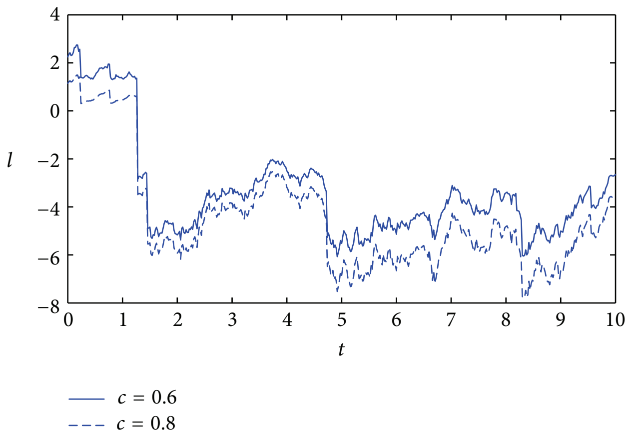
(a) The effect of γ on the optimal precommitted strategy



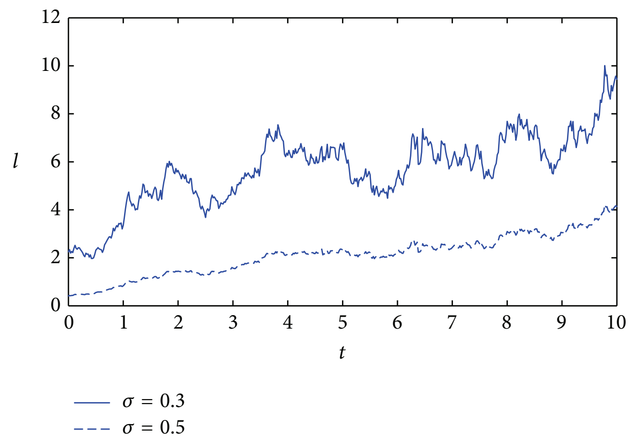
(b) The effect of r_0 on the optimal precommitted strategy



(c) The effect of r_1 on the optimal precommitted strategy



(d) The effect of c on the optimal precommitted strategy

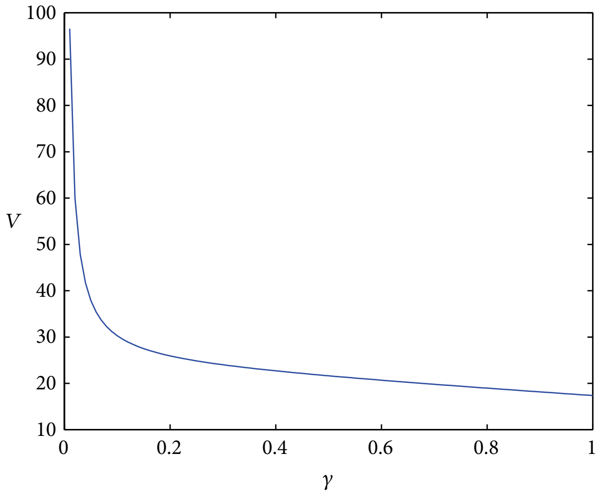


(e) The effect of σ on the optimal precommitted strategy

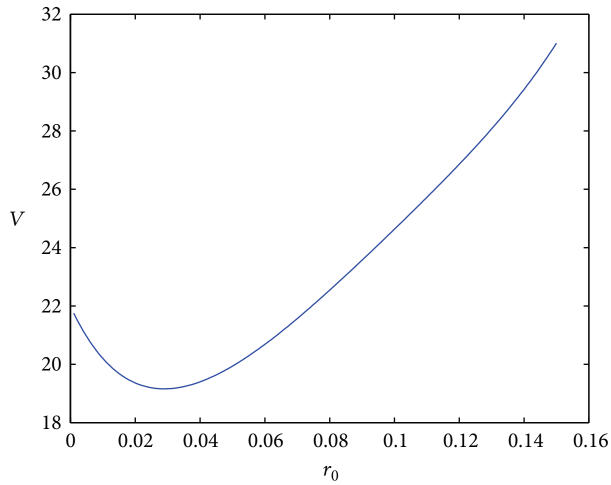
Figure 2: The effect of parameters on the optimal precommitted strategy.

can increase the deterministic part and the stochastic part (current wealth) of the precommited strategy which results in the uncertainty of their difference.

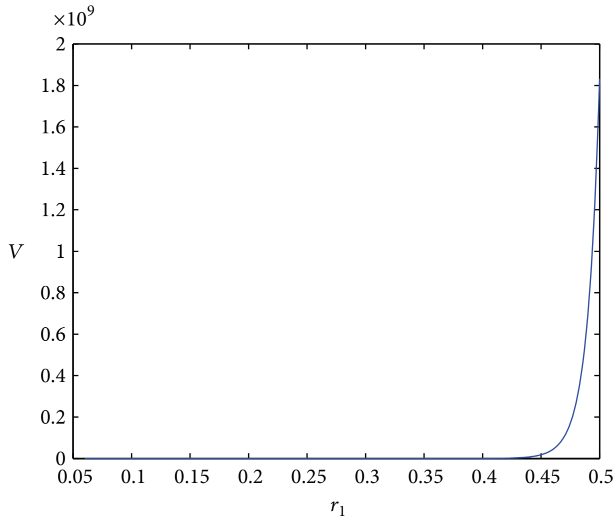
Secondly, Figure 3 shows how the coefficients involved impact on the value function. Figure 3(a) illustrates that the value function is decreasing with respect to γ, namely, the larger risk aversion the company has, the smaller the value function is; Figure 3(b) reveals when the risk-free rate is small enough, the value function decreases and when the risk-free rate is close to the the appreciation rate, the value function increases; Figure 3(c) illustrates that the value function is increasing with respect to r_1, namely, the bigger the appreciation rate is, the bigger the value function is; Figure 3(d) reveals that the value function is decreasing



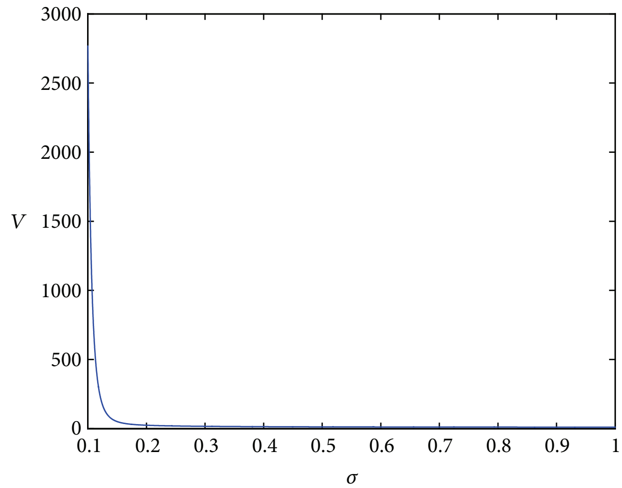
(a) The effect of γ on the optimal value function



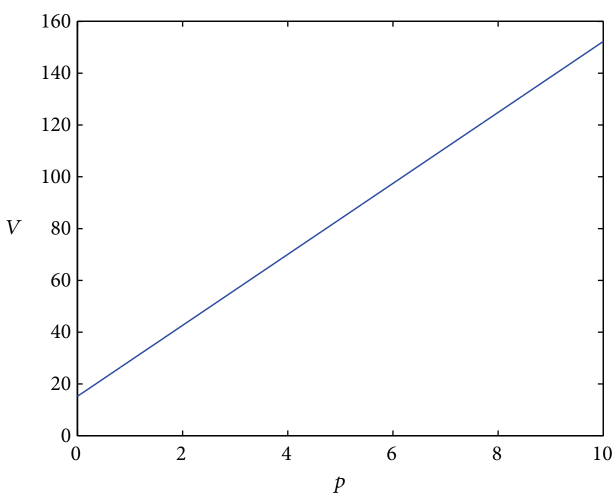
(b) The effect of r_0 on the optimal value function



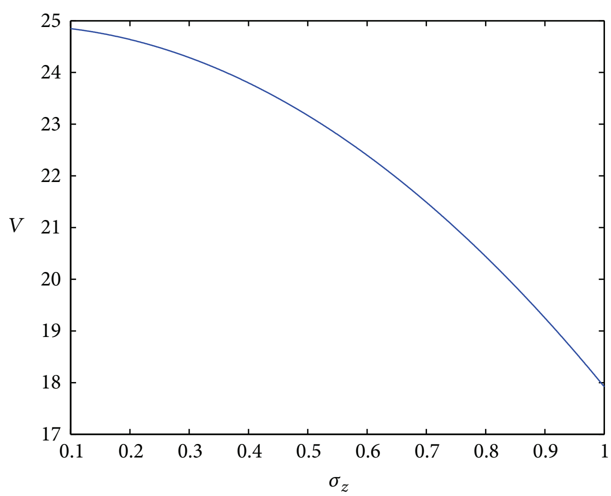
(c) The effect of r_1 on the optimal value function



(d) The effect of σ on the optimal value function

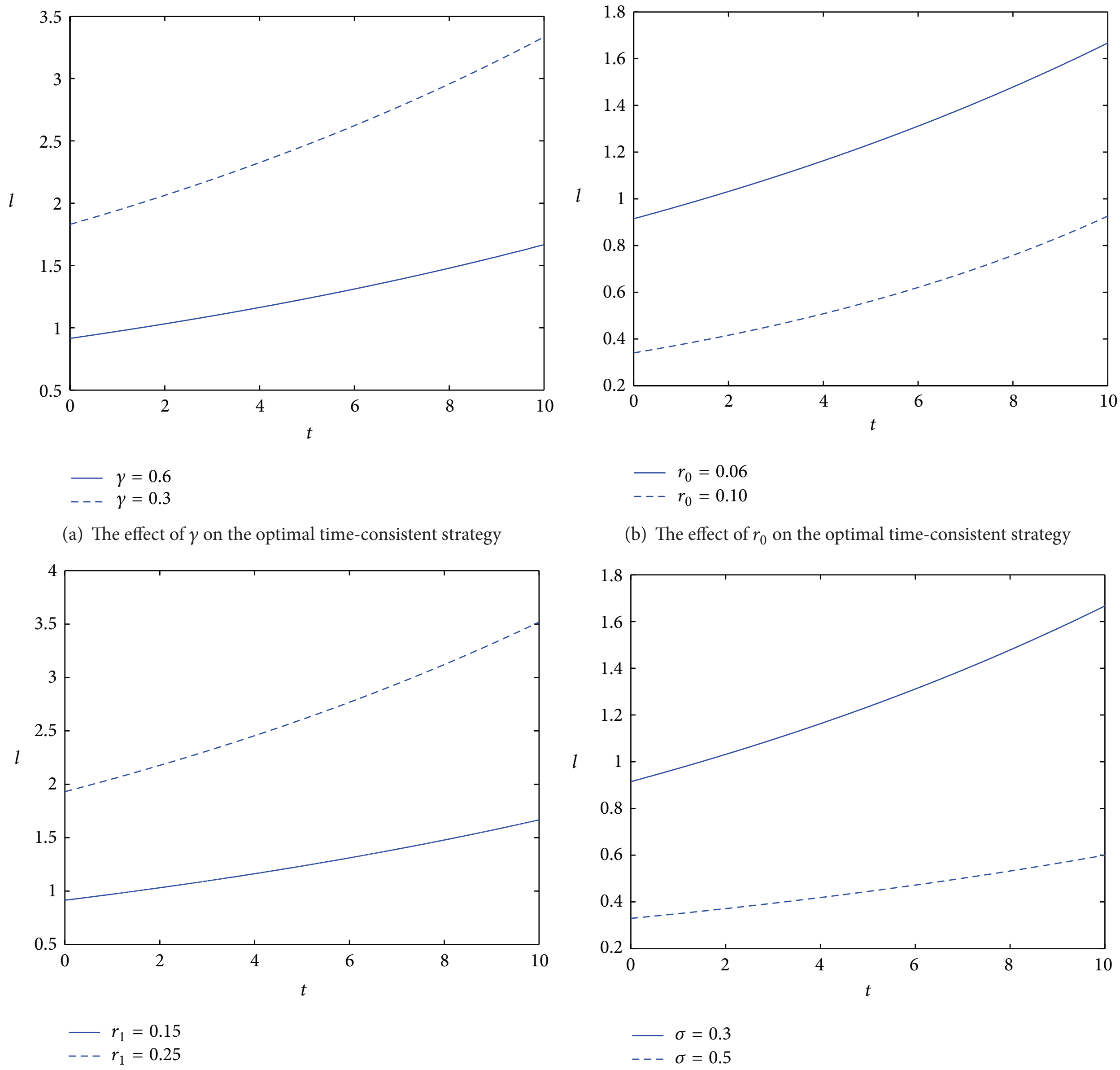


(e) The effect of p on the optimal value function



(f) The effect of σ_z on the optimal value function

Figure 3: The effect of parameters on the optimal value function.



(a) The effect of γ on the optimal time-consistent strategy

(b) The effect of r_0 on the optimal time-consistent strategy

(c) The effect of r_1 on the optimal time-consistent strategy

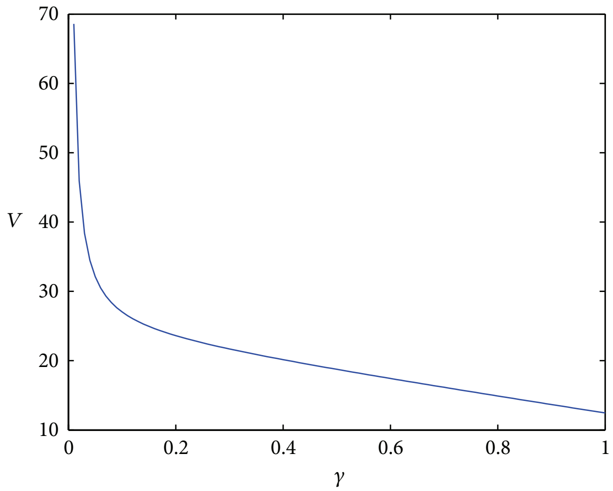
(d) The effect of σ on the optimal time-consistent strategy

FIGURE 4: The effect of parameters on the optimal time-consistent strategy.

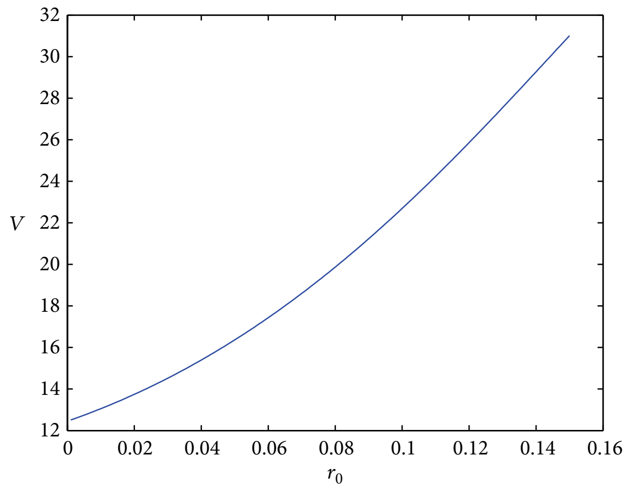
with respect to σ, namely, the bigger the volatility of the market's risky asset is, the smaller the value function is. Recalling $p = \lambda\mu - c$, Figure 3(e) illustrates that the value function is increasing with respect to p, namely, the larger the expectation of the positive income is or the smaller the expense rate of the company is or the bigger the intensity of the jumps of the profit is, the bigger the value function becomes; Figure 3(f) shows the value function is increasing with respect to σ_z, namely, the smaller the second moment of the positive income is, the bigger the value function becomes.

3.2. Analysis of the Time-Consistent Strategy and the Equilibrium Value Function. This subsection will work on numerical analysis of the time-consistent strategy and the equilibrium value function.

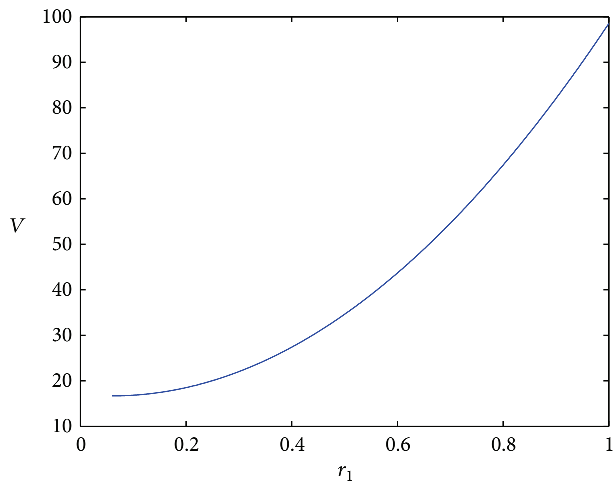
Firstly, we will show how the coefficients involved impact on the time-consistent investment strategy. From (19) we can see that the time-consistent investment strategy is independent of the current wealth. Figure 4(a) illustrates that the time-consistent investment strategy is decreasing with respect to γ, namely, the more the company dislikes risk, the less amount the company invests in the risky asset; Figure 4(b) reveals that the time-consistent investment strategy is decreasing with respect to r_0, namely, the smaller the risk-free rate is, the more amount the company invests in the risky asset; Figure 4(c) reveals that the time-consistent investment strategy is increasing with respect to r_1, namely,



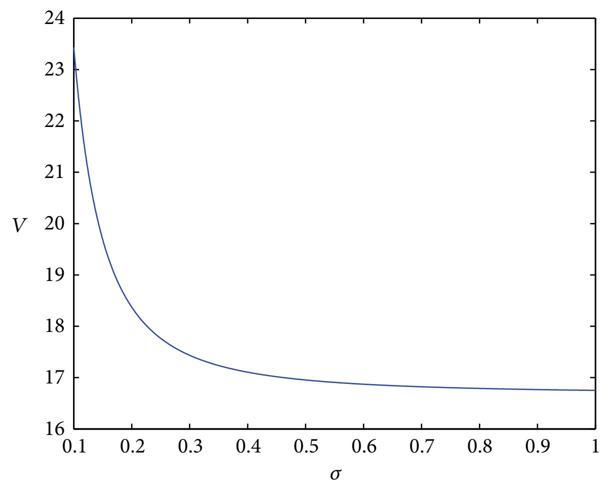
(a) The effect of γ on the optimal equilibrium value function



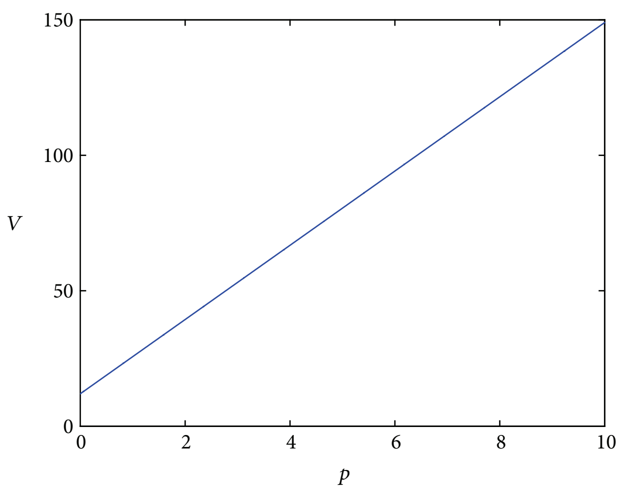
(b) The effect of r_0 on the optimal equilibrium value function



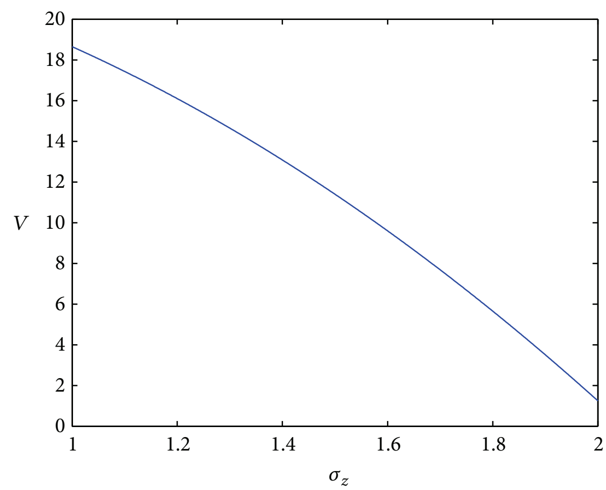
(c) The effect of r_1 on the optimal equilibrium value function



(d) The effect of σ on the optimal equilibrium value function



(e) The effect of p on the optimal equilibrium value function



(f) The effect of σ_z on the optimal equilibrium value function

FIGURE 5: The effect of parameters on the optimal equilibrium value function.

when the appreciation rate r_1 increases, the company should invest more money in the risky asset; Figure 4(d) tells that the time-consistent investment strategy is decreasing with respect to σ, namely, when volatility of the risky asset increases, the company should invest more money in the risk-free asset.

Secondly, Figure 5 shows how the coefficients involved impact on the equilibrium value function. The parameters γ, r_1, σ, p, and σ_z have the similar effect on the equilibrium value function as their effect on the value function with precommitment discussed in Section 3.1. Furthermore, the equilibrium value function is increasing with respect to the risk-free rate.

4. Conclusion

In this paper, optimal investment strategies for a dual risk model are explored under mean-variance criterion. We assume that a company can invest into a finance market which consists of a risk-free asset and a risky asset. We have derived the optimal precommitted strategy and the time-consistent equilibrium strategy for problem (MV). In the end, numerical analysis is given for optimal investment strategies and the value functions. From the comparisons on the value functions and efficient frontiers, it seems to be that the precommitted strategy is better than the time-consistent strategy. Unfortunately, the precommitted strategy is not time-consistent. Thus these investment strategies have different advantages. These strategies are all important for the company.

Appendices

A.

In this appendix, we will take three steps to derive the precommitted investment strategy for problem (MV).

Step 1. The closed expressions for the optimal investment strategy and value function are derived by solving the related Hamilton-Jacobi-Bellman (HJB) equation for problem (BM). Based on similar arguments in Fleming and Soner [12], we obtain the HJB equation which the value function $V(x, d, y)$ satisfies.

Lemma A.1 (verification theorem). *If a function $W(t,x) \in C^{1,2}(Q)$ and a function $l^*(t) \in U$ satisfy the following HJB equation:*

$$\sup_{l\in U}\left\{\mathscr{A}^l W(t,x)\right\} = 0, \tag{A.1}$$

$$W(T,x) = -\frac{\gamma}{2}x^2 + yx + \frac{\gamma}{2}d^2 + d - yd, \tag{A.2}$$

$$l^*(t) = \arg\sup\left\{\mathscr{A}^l W(t,x)\right\}, \tag{A.3}$$

then $V(x,d,y) = W(0,x)$ and $l^*(t)$ is the optimal investment strategy.

Next, we will construct the solution to problem (BM). Assume that there exists a real function $W(t,x)$ which satisfies the condition in Lemma A.1. Inserting (7) into (A.1), we have

$$\begin{aligned}\sup_{l\in U}\Big\{&W_t(t,x) + W_x(t,x)\left(r_0(t)x - c + r(t)l(t)\right)\\ &+\frac{1}{2}W_{xx}(t,x)l^2(t)\sigma^2(t)\\ &+\lambda E\left[W(t,x+Z) - W(t,x)\right]\Big\} = 0.\end{aligned} \tag{A.4}$$

Because the structure of (A.4) and the boundary condition of $W(t,x)$ given by (A.2) are quadratic in x, it is natural to conjecture that

$$W(t,x) = A(t)x^2 + B(t)x + C(t), \tag{A.5}$$

$$A(T) = -\frac{\gamma}{2}, \qquad B(T) = y, \qquad C(T) = \frac{\gamma}{2}d^2 + d - yd. \tag{A.6}$$

The corresponding partial derivatives are given by the following equations:

$$\begin{aligned}W_t(t,x) &= \dot{A}(t)x^2 + \dot{B}(t)x + \dot{C}(t),\\ W_x(t,x) &= 2A(t)x + B(t),\\ W_{xx}(t,x) &= 2A(t).\end{aligned} \tag{A.7}$$

Inserting (A.5)–(A.7) into (A.4) yields

$$\begin{aligned}\sup_{l\in U}\Big\{&\dot{A}(t)x^2 + \dot{B}(t)x + \dot{C}(t) + (2A(t)x + B(t))\\ &\times\left(r_0(t)x - c + r(t)l(t) + \lambda\mu_z\right)\\ &+A(t)\left(l^2(t)\sigma^2(t) + \lambda\sigma_z^2\right)\Big\} = 0.\end{aligned} \tag{A.8}$$

Differentiating the function in the left bracket of (A.8) with respect to l and setting the derivative to zero, we get

$$l(t) = -\frac{(2A(t)x + B(t))r(t)}{2A(t)\sigma^2(t)}. \tag{A.9}$$

Inserting (A.9) into (A.8), we have the following equation:

$$\begin{aligned}&\left(\dot{A}(t) + A(t)\left(2r_0(t) - \xi(t)\right)\right)x^2\\ &+\left(\dot{B}(t) + B(t)\left(r_0(t) - \xi(t)\right) + 2Ap\right)x\\ &+\left(\dot{C}(t) + \lambda\sigma_z^2 A(t) + pB(t) - \frac{B^2(t)}{4A(t)}\xi(t)\right) = 0,\end{aligned} \tag{A.10}$$

where

$$\xi(t) = \frac{r^2(t)}{\sigma^2(t)}, \qquad p = \lambda\mu_z - c > 0. \tag{A.11}$$

To ensure the above equation holds, we require that

$$\begin{gathered}\dot{A}(t)+A(t)\left(2r_0(t)-\xi(t)\right)=0,\quad A(T)=-\frac{\gamma}{2},\\ \dot{B}(t)+B(t)\left(r_0(t)-\xi(t)\right)+2Ap=0,\quad B(T)=y,\\ \dot{C}(t)+\lambda\sigma_z^2A(t)+pB(t)-\frac{B^2(t)}{4A(t)}\xi(t)=0,\\ C(T)=\frac{\gamma}{2}d^2+d-yd.\end{gathered}\tag{A.12}$$

By solving the system of equations, we have

$$\begin{aligned}A(t)&=-\frac{\gamma}{2}e^{\int_t^T(2r_0(s)-\xi(s))ds},\\ B(t)&=e^{\int_t^T(r_0(s)-\xi(s))ds}\left(-p\gamma\int_t^T e^{\int_s^T r_0(u)du}ds+y\right),\\ C(t)&=-\frac{\gamma\lambda\sigma_z^2}{2}\int_t^T e^{\int_s^T(2r_0(u)-\xi(u))du}ds\\ &\quad-\frac{1}{2\gamma}e^{-\int_t^T\xi(u)du}\left[-p\gamma\int_t^T e^{\int_s^T r_0(v)dv}ds+y\right]^2\\ &\quad+\frac{\gamma}{2}d^2+d-dy+\frac{y^2}{2\gamma}.\end{aligned}\tag{A.13}$$

Substituting (A.13) into (A.9), we have the precommitted investment strategy for problem (MV):

$$l(t)=\frac{r(t)}{\sigma^2(t)}\left\{e^{-\int_t^T r_0(s)ds}\left(-p\int_t^T e^{\int_s^T r_0(u)du}ds+\frac{y}{\gamma}\right)-x(t)\right\}.\tag{A.14}$$

Correspondingly, the value function for problem (BM) is given by

$$\begin{aligned}V(x,d,y)&=-\frac{\gamma\lambda\sigma_z^2}{2}\int_0^T e^{\int_s^T(2r_0(u)-\xi(u))du}ds\\ &\quad-\frac{\gamma}{2}e^{-\int_0^T\xi(s)ds}\left\{xe^{\int_0^T r_0(s)ds}\right.\\ &\qquad\left.+\left[p\int_0^T e^{\int_s^T r_0(u)du}ds-\frac{y}{\gamma}\right]\right\}^2\\ &\quad+\frac{\gamma}{2}d^2+d-dy+\frac{y^2}{2\gamma}.\end{aligned}\tag{A.15}$$

Step 2. Problem ($\overline{\text{CMV}}$) is solved by virtue of the relationship of the value functions for problem (BM) and problem ($\overline{\text{CMV}}$). Differentiating $V(x,d,y)$ with respect to y yields

$$\begin{aligned}\frac{\partial V}{\partial y}&=e^{-\int_0^T\xi(s)ds}\left\{xe^{\int_0^T r_0(s)ds}+\left[p\int_0^T e^{\int_s^T r_0(u)du}ds-\frac{y}{\gamma}\right]\right\}\\ &\quad-d+\frac{y}{\gamma},\end{aligned}\tag{A.16}$$

$$\frac{\partial^2 V}{\partial y^2}=\frac{1-e^{-\int_0^T\xi(s)ds}}{\gamma}>0.\tag{A.17}$$

Setting the derivative of $V(x,d,y)$ to zero yields

$$y^*=\frac{\left[p\int_0^T e^{\int_s^T r_0(v)dv}ds+xe^{\int_0^T r_0(s)ds}-de^{\int_0^T\xi(s)ds}\right]\gamma}{1-e^{\int_0^T\xi(s)ds}}.\tag{A.18}$$

By virtue of the condition of a stationary point becoming an extreme point, we conclude that y^* is the point which minimizes $V(x,d,y)$. The optimal investment strategy $l^*(t)$ and the value function $V(x,d)$ for problem ($\overline{\text{CMV}}$) can be derived by inserting (A.17) into (A.14)-(A.15) as follows:

$$\begin{aligned}l^*(t)&=\frac{r(t)}{\sigma^2(t)}\left\{e^{-\int_t^T r_0(s)ds}\right.\\ &\quad\times\left(-p\int_t^T e^{\int_s^T r_0(u)du}ds\right.\\ &\quad+\left(p\int_0^T e^{\int_s^T r_0(v)dv}ds+xe^{\int_0^T r_0(s)ds}\right.\\ &\quad\left.-de^{\int_0^T\xi(s)ds}\right)\\ &\quad\left.\left.\times\left(1-e^{\int_0^T\xi(s)ds}\right)^{-1}\right)-x(t)\right\},\\ V(x,d)&=-\frac{\gamma\lambda\sigma_z^2}{2}\int_0^T e^{\int_s^T(2r_0(u)-\xi(u))du}ds\\ &\quad+\frac{\gamma}{2\left(1-e^{\int_0^T\xi(s)ds}\right)}\\ &\quad\times\left\{xe^{\int_0^T r_0(s)ds}+p\int_0^T e^{\int_s^T r_0(u)du}ds-d\right\}^2+d.\end{aligned}\tag{A.19}$$

Step 3. Finally problem (MV) is solved by virtue of $V(x)=\sup_{d\in R}V(x,d)$. Differentiating $V(x,d)$ at d, we can conclude that

$$\begin{aligned}\frac{\partial V(x,d)}{\partial d}&=\frac{\gamma}{1-e^{\int_0^T\xi(s)ds}}\\ &\quad\times\left\{d-xe^{\int_0^T r_0(s)ds}-p\int_0^T e^{\int_s^T r_0(u)du}ds\right\}+1,\\ \frac{\partial^2 V(x,d)}{\partial d^2}&=\frac{\gamma}{1-e^{\int_0^T\xi(s)ds}}<0.\end{aligned}\tag{A.20}$$

From the extreme value theory, the optimal expected terminal wealth d^* does exist. d^* can be derived by $\partial V(x,d)/\partial d=0$ and is given by the following equation

$$d^*=p\int_0^T e^{\int_s^T r_0(v)dv}ds+xe^{\int_0^T r_0(s)ds}+\frac{\left(e^{\int_0^T\xi(s)ds}-1\right)}{\gamma}.\tag{A.21}$$

By inserting (A.21) into (A.19), Theorem 2 is proved.

B.

This section will prove Theorem 5. Due to the relationship of the equilibrium value function for problem (MV) and problem ($\overline{\text{MV}}$), we firstly derive the time-consistent investment strategy for problem ($\overline{\text{MV}}$). From standard arguments as in Björk and Murgoci [8], we obtain the extended HJB equations for problem ($\overline{\text{MV}}$).

Lemma B.1 (verification theorem). *If the two functions $W(t,x)$ and $h(t,x) \in C^{1,2}(Q)$ satisfy the following extended HJB equations:*

$$\sup_{l\in U}\left\{\mathscr{A}^l W(t,x) - \mathscr{A}^l\left(\frac{\gamma}{2}h^2(t,x)\right) + \gamma h(t,x)\mathscr{A}^l h(t,x)\right\} = 0, \tag{B.1}$$

$$W(T,x) = x, \tag{B.2}$$

$$\mathscr{A}^{l_*} h(t,x) = 0, \tag{B.3}$$

$$h(T,x) = x, \tag{B.4}$$

where

$$\begin{aligned} l_*(t) = \arg\sup &\left\{\mathscr{A}W(t,x) - \mathscr{A}\left(\frac{\gamma}{2}h^2(t,x)\right)\right. \\ &\left. + \gamma h(t,x)\mathscr{A}h(t,x)\right\}, \end{aligned} \tag{B.5}$$

then $\psi(t,x) = W(t,x)$, $E_{t,x}(X^{l^*}(t)) = h(t,x)$, and $l_*(t)$ is the optimal time-consistent strategy.

Next, we will construct the solution to problem ($\overline{\text{MV}}$). Assume that there exist two real functions $W(t,x)$ and $h(t,x)$ satisfying the conditions in Lemma B.1. By virtue of (7), (B.1) can be rewritten as

$$\begin{aligned} \sup_{l\in U}&\left\{W_t(t,x) + W_x(t,x)\left(r_0(t)x - c + r(t)l(t)\right)\right. \\ &+ \frac{1}{2}\left(W_{xx}(t,x) - \gamma h_x^2(t,x)\right)l^2(t)\sigma^2(t) \\ &+ \lambda E\left[W(t,x+Z) - \frac{\gamma}{2}h(t,x+Z)\right. \\ &\qquad \left.\times\left(h(t,x+Z) - 2h(t,x)\right)\right] \\ &\left. -\lambda\left[W(t,x) + \frac{\gamma}{2}h^2(t,x)\right]\right\} = 0. \end{aligned} \tag{B.6}$$

Equation (B.3) can be also rewritten as

$$\begin{aligned} &h_t(t,x) + h_x(t,x)\left(r_0(t)x - c + r(t)l_*(t)\right) \\ &\quad + \frac{h_{xx}(t,x)}{2}l_*^2(t)\sigma^2(t) + \lambda E\left[h(t,x+Z) - h(t,x)\right] = 0, \end{aligned} \tag{B.7}$$

where l_* is determined below.

Since the linear structure of (B.6)-(B.7) and the boundary conditions of $W(t,x)$ and $h(t,x)$ given by (B.2) and (B.4) are linear in x, we can conjecture that $W(t,x)$ and $h(t,x)$ have the following expressions:

$$W(t,x) = M(t)x + N(t), \quad M(T) = 1,\ N(T) = 0, \tag{B.8}$$

$$h(t,x) = m(t)x + n(t), \quad m(T) = 1,\ n(T) = 0. \tag{B.9}$$

Thus, the corresponding partial derivatives are calculated as follows:

$$\begin{gathered} W_t(t,x) = \dot{M}(t)x + \dot{N}(t), \qquad W_x(t,x) = M(t), \\ W_{xx}(t,x) = 0, \end{gathered} \tag{B.10}$$

$$\begin{gathered} h_t(t,x) = \dot{m}(t)x + \dot{n}(t), \qquad h_x(t,x) = m(t), \\ h_{xx}(t,x) = 0. \end{gathered} \tag{B.11}$$

By substituting (B.8)–(B.11) into (B.6), it yields that

$$\begin{aligned} \sup_{l\in U}&\left\{\dot{M}(t)x + \dot{N}(t) + M\left(r_0(t)x - c + r(t)l(t) + \lambda\mu_z\right)\right. \\ &\left. -\frac{\gamma}{2}m^2(t)\left[\lambda\sigma_z^2 + l^2(t)\sigma^2(t)\right]\right\} = 0. \end{aligned} \tag{B.12}$$

Differentiating the function in the left bracket of (B.12) with respect to l and setting the derivative to zero, we get

$$l_*(t) = \frac{M(t)r(t)}{\gamma m^2(t)\sigma^2(t)}. \tag{B.13}$$

Inserting (B.13) into (B.12) and (B.7), we have

$$\begin{aligned} &\left(\dot{M}(t) + r_0(t)M(t)\right)x + \dot{N}(t) - cM(t) + \lambda\mu_z M(t) \\ &\quad - \frac{\gamma}{2}m^2(t)\lambda\sigma_z^2 + \frac{M^2(t)\xi(t)}{2\gamma m^2(t)} = 0, \\ &\left(\dot{m}(t) + r_0(t)m(t)\right)x + \dot{n}(t) - cm(t) \\ &\quad + \lambda\mu_z m(t) + \frac{M(t)\xi(t)}{\gamma m(t)} = 0. \end{aligned} \tag{B.14}$$

In order to ensure the above equations hold, we require

$$\begin{gathered} \dot{M}(t) + r_0(t)M(t) = 0, \quad M(T) = 1, \\ \dot{N}(t) + \left(\lambda\mu_z - c\right)M - \frac{\gamma}{2}m^2(t)\lambda\sigma_z^2 + \frac{M^2(t)\xi(t)}{2\gamma m^2(t)} = 0, \\ N(T) = 0, \\ \dot{m}(t) + r_0(t)m(t) = 0, \quad m(T) = 1, \\ \dot{n}(t) + \left(\lambda\mu_z - c\right)m(t) + \frac{M(t)\xi(t)}{\gamma m(t)} = 0, \quad n(T) = 0. \end{gathered} \tag{B.15}$$

By simple calculation, the solutions to (B.15) are given by the following equations:

$$M(t) = e^{\int_t^T r_0(s)ds},$$
$$N(t) = (\lambda\mu_z - c)\int_t^T e^{\int_s^T r_0(u)du} ds - \frac{\gamma}{2}\lambda\sigma_z^2 \int_t^T e^{2\int_s^T r_0(u)du} ds + \frac{1}{2\gamma}\int_t^T \xi(s)\, ds, \quad \text{(B.16)}$$
$$m(t) = e^{\int_t^T r_0(s)ds},$$
$$n(t) = (\lambda\mu_z - c)\int_t^T e^{\int_s^T r_0(u)du} ds + \frac{1}{\gamma}\int_t^T \xi(s)\, ds.$$

Substituting (B.16) into (B.13) and (B.8), we derive the explicit expressions for $W(t,x)$ and $h(t,x)$ for problem ($\overline{\text{MV}}$):

$$l_*(t) = \frac{r(t)}{\gamma\sigma^2(t)} e^{-\int_t^T r_0(s)ds},$$
$$W(t,x) = e^{\int_t^T r_0(s)ds} x + p\int_t^T e^{\int_s^T r_0(u)du} ds - \frac{\gamma\lambda\sigma_z^2}{2}\int_t^T e^{2\int_s^T r_0(u)du} ds + \frac{1}{2\gamma}\int_t^T \xi(s)\, ds. \quad \text{(B.17)}$$

When the original time t is equal to 0, the time-consistent equilibrium strategy and the equilibrium value function for problem (MV) are shown in Theorem 5.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research is supported by the National Natural Science Foundation of China (Grant nos. 11201335, 11301376, and 71071111) and the Research Project of the Social Science and Humanity on Young Fund of the Ministry of Education (Grant no. 11YJC910007).

References

[1] B. Avanzi and H. U. Gerber, "Optimal dividends in the dual model with diffusion," *Astin Bulletin*, vol. 38, no. 2, pp. 653–667, 2008.

[2] B. Avanzi, H. U. Gerber, and E. S. W. Shiu, "Optimal dividends in the dual model," *Insurance: Mathematics & Economics*, vol. 41, no. 1, pp. 111–123, 2007.

[3] J. Zhu and H. Yang, "Ruin probabilities of a dual Markov-modulated risk model," *Communications in Statistics*, vol. 37, no. 18–20, pp. 3298–3307, 2008.

[4] H. Dai, Z. Liu, and N. Luan, "Optimal dividend strategies in a dual model with capital injections," *Mathematical Methods of Operations Research*, vol. 72, no. 1, pp. 129–143, 2010.

[5] D. Yao, H. Yang, and R. Wang, "Optimal dividend and capital injection problem in the dual model with proportional and fixed transaction costs," *European Journal of Operational Research*, vol. 211, no. 3, pp. 568–576, 2011.

[6] X. Y. Zhou and D. Li, "Continuous-time mean-variance portfolio selection: a stochastic LQ framework," *Applied Mathematics and Optimization*, vol. 42, no. 1, pp. 19–33, 2000.

[7] D. Li and W.-L. Ng, "Optimal dynamic portfolio selection: multiperiod mean-variance formulation," *Mathematical Finance*, vol. 10, no. 3, pp. 387–406, 2000.

[8] T. Björk and A. Murgoci, "A general theory of Markovian time inconsistent stochastic control problems," *Social Science Research Network*, 2010.

[9] I. Ekeland and A. Lazrak, *Being Serious about Non-Commitment: Subgame Perfect Equilibrium in Continuous Time*, University of British Columbia, 2006.

[10] E. M. Kryger and M. Steffensen, "Some solvable portfolio problems with quadratic and collective objectives," *Social Science Research Network*, 2010.

[11] Y. Zeng, Z. Li, and Y. Lai, "Time-consistent investment and reinsurance strategies for mean-variance insurers with jumps," *Insurance: Mathematics & Economics*, vol. 52, no. 3, pp. 498–507, 2013.

[12] W. H. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions*, vol. 25 of *Stochastic Modelling and Applied Probability*, Springer, New York, NY, USA, 2nd edition, 2006.

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 527598, 7 pages
http://dx.doi.org/10.1155/2014/527598

Research Article

Research on Overconfidence in Decision-Making for the Capacity Recovery of Damaged Power Systems

Xing Bao[1,2]

[1] *Department of Business Administration, Zhejiang Institute of Administration, 1000 Wenyixilu, Hangzhou 311121, China*
[2] *Zhejiang Scientific Development Research Center, 1000 Wenyixilu, Hangzhou 311121, China*

Correspondence should be addressed to Xing Bao; goldbxing@gmail.com

Received 12 March 2014; Accepted 10 April 2014; Published 8 May 2014

Academic Editor: Weihua Liu

Copyright © 2014 Xing Bao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper studies the influence of two types of overconfident behavior, overestimation and overprecision, on decision of capacity recovery when power system's critical capacity is seriously damaged. A newsvendor model is used to prove that increasing regulatory punishment for electricity shortage and providing subsidy for capacity recovery are conducive measures to calibrate insufficient service level caused by an overconfident manager. The research also finds that the manager's overprecision behavior both negatively and positively influences the decision of capacity recovery, and a calibration method could motivate manager to recover more capacity by tuning up the ratio of punishment and subsidy. However, the effectiveness of the calibration mentioned above is inevitably weakened due to the less capacity recovery given by an overestimated manager. This research also indicates that the manager should pay close attention to the random disturbance whose distribution peak is left skewed, and correspondingly more capacity recovery should be given to improve the service level of power system during the disruption.

1. Introduction

Empirical researches in the field of knowledge management have found that a manager's confidence in the accuracy of his/her priori knowledge contributes to his/her decision-making when facing a complicated situation with incomplete information or highly fluctuated environment. A manager's intention and confidence to apply a priori knowledge may be strengthened when his/her observation of the situation is compatible with his/her knowledge (cf. [1]). A manager could strengthen his/her intention and confidence to apply a priori knowledge during decision by referring to past successful experiences, but this is not always the case. In an unexpected situation with extremely low probability, an overconfident and experienced manager may lose the capability to calibrate the accuracy of his/her a priori knowledge during the decision-making process (cf. [2]).

Some researchers have indicated that the more experience a manager has, the more likely he/she is to become overconfident in making decision (cf. [3]). In October 2012, the author of this paper conducted a survey of 50 senior managers of a district branch of the State Grid Corporation of China (SGCC) who suffered great disruption in the winter of 2008, nearly 75% of its critical transmission network was damaged by unexpected frozen rain, and it took over 2 weeks of recovery, which exceeded nearly one week than expected. After taking an in-depth interview with one of its senior managers in 2012, an interesting phenomenon was found: a large portion of mistakes were caused by middle-level managers who were making decision at the first line of disruption management. A successive interview with 50 middle-level managers was taken, and the results showed that 80% of the middle-level managers interviewed confessed that mistakes frequently happened in making decision during disruption management. However, it is interesting that most of the managers attributed the mistakes to their overconfidence in applying a priori knowledge. Similarly, the executives of Tokyo Electric Power Company (TEPCO) declined international assistance after the occurrence of the Fukushima Nuclear Accident in 2011. This was partly due to the fact that, as the world's leading nuclear power operator, TEPCO's executives overestimated their abilities of disruption management, and as a result, there is a nuclear leakage crisis continuing to this day. For public infrastructures, such as electric power systems, rigorous regulations are enacted by the Government of China to ensure that managers of these systems are taking effective disruption management, for the purpose of securing and preserving the public interest as great as possible. However, in recent years, government officials and researchers have frequently realized the drawbacks of the existing regulations after the investigations of the 2003 blackout in the United States and Canada. According to the behavioral operation management (BOM) theory, the effectiveness of government regulations might be weakened, or even counteracted, due to the decision bias given by a bounded rational manager, particularly affected by an overconfident manager (cf. [4]). However, few studies have been conducted on the topic of how an overconfident behavior influences the manager's decision bias during the disruption. According to the knowledge of the author, this paper is the first disruption management study on the decision bias of an overconfident manager in the scenario that the critical capacity of a power system is seriously damaged by unexpected events. This paper also investigates the calibration methods, specifically regulatory punishment and subsidy, for the purpose of reducing the decision bias given by an overconfident manager during capacity recovery process.

This paper is organized into 6 sections. After the Introduction, relevant literatures on disruption management and overconfidence in operation management are reviewed as detailed as possible in Section 2. In Section 3, a newsvendor model corresponding to the disruption cost is firstly presented as a basic model that would be compared with the subsequent model of an overconfident manager. In Section 4, a modification to the newsvendor model mentioned in Section 3 is presented by two kinds of overconfident behaviors, overestimation and overprecision. And, corre-spondingly, the decision bias on capacity recovery is demonstrated. In Section 5, numerical simulations of capacity recovery decision biases are presented to illustrate the effects of random disturbances with the distribution functions being symmetrically and asymmetrically distributed. Finally, some interesting managerial insights and future researches are presented in Section 6.

2. Literature Review

In the past decade, operational systems and supply chains have frequently been crippled by unexpected catastrophes, such as natural disasters, human-made hazards, and terrorist attacks. Knowledge of traditional risk management is facing the challenges, because some unexpected events with extremely low probability have significant negative impacts on the whole supply chain from the disrupted node; thus, the huge negative consequences escalate quickly throughout the supply chain in a "snowball effect" (cf. [5]). Therefore, disruption management is focused on to reduce risks and expenses by researchers in recent years in the field of operation management.

Professor Sheffi may have been the first to research disruption management by conducting his studies on the security problems for international supply chains under the consideration of international terrorism (cf. [6]). Later, some researchers (e.g., Norrman, Hendricks, Oke, etc.) reviewed the mitigation strategies of lean supply chains through empirical and case studies in the context of disruption (cf. [7–9]). Some other researchers (e.g., Chopra, Kleindorfer, Zsidisin, etc.) presented disruption management frameworks of supply chains from the point of business continuity planning and flexible operation strategies (cf. [10–12]). Other researchers conducted studies through mathematical models of operation management, which follow the framework proposed by Tang (cf. [13]), based on the assumption that nodes of a supply chain might suffer disruptions. For example, in order to mitigate the disruption of main supplier, some researchers present selection models of supplier bases; the questions of how many suppliers are the best and when to start up the backup suppliers under different objective functions are investigated in detail in the consideration of cutting down the disruption cost while maintaining a certain service level (cf. [14–17]). To mitigate the cost in case of production capacity being disrupted, rescheduling algorithm is presented in order to get production plans that both operational cost and computation speed are satisfied in redispatching the residual capacity (cf. [18]), and to mitigate the risk in case of production capacity being disrupted, the decision of *ex-ante* preventive capacity investment or the short-term capacity trading strategy with partners is presented by multiobjective programming models (cf. [19]). Most of the mathematical models focus on the inventory strategies in case of disruption, for example, preventive inventory control strategies for managing supply chain disruption risk (cf. [20]), or performances of different inventory models and supply chain structures when facing transportation disruptions (cf. [21]), or optimal inventory policies with advanced warning of disruptions (cf. [22]). Other researchers have adopted game theory and contract theory to investigate incentive mechanisms or coordination policies that mitigate or prevent disruptions, for example, optimal subsidies for suppliers under a competitive manufacturing supply chain facing the disruptions of supplier bankruptcies (cf. [23]), or how a firm (buyer) can use incentive mechanisms to motivate a supplier's investment in capacity restoration, thus generating the restoration enhancement (RE) strategy or supplier diversification (SD) strategy when incentives are given *ex-ante* or *ex post* the supplier's disruption (cf. [24]), or how a disrupted supplier chose either to pay a penalty or to use backup production to manufacturer when supplier was privileged with private information (cf. [25]).

The mitigation strategies mentioned above generally assume that the "*decision- maker is completely rational*" throughout the entire cycle of disruption management, which does not account for the influence of bounded rationality on managers. If the bounded rationality of individuals in the decision-making process is neglected, a systematic bias will inevitably happen when compared to that of a completely rational model (cf. [26, 27]). When an experienced manager is making a decision, overconfidence is the most

Table 1: Notation descriptions in the proposed model.

D	Electricity demand of the public during disruption period
d	Residual capacity that can still be in operation after disruption and $d < D$
μ	Manager's decision of capacity recovery
$\overline{X}$	Random disturbance to capacity recovery process with its probability density and distribution functions of $f(x)$ and $F(x)$, respectively, with the mean and variance being φ and σ^2, respectively
k	Electricity shortage cost per capacity, which is also the regulatory punishment according to Electricity Regulatory Ordinance of China, given by regulation number 599
c_1	Recovery cost per damaged capacity
c_0	Operation cost per normal capacity
α	Factor of overestimation behavior of the manager
β	Factor of overprecision behavior of the manager

significant display of bounded rationality (cf. [28]), which is also validated by some behavioral experiments under the scenario of newsvendor decision (cf. [29–31]). However, mathematical models are not as common as empirical and experimental studies in the operation management field (cf. [32]). As far as the knowledge of the author, Croson et al. were pioneers who studied the decision bias by a newsvendor model of incorporating the overestimation and overprecision behaviors (cf. [33]). Inspired by the research of Croson (cf. [33]), this paper intends to study the capacity recovery decision through a newsvendor model when power system's capacity is damaged, of which the decision biases are investigated by considering the manager being overestimated and overprecise during disruption management. In addition, the calibration capability of regulatory punishment and subsidy in reducing the decision bias of capacity recovery is analyzed in detail, and managerial insights are presented by theoretical proofs and numerical simulations.

3. Description of Problems

It is assumed that the critical operation capacity of a power system is seriously damaged due to unexpected events, such as freezing rain, man-made sabotage, and operational mistakes, and as a result the service level is lowered and cannot satisfy the electricity demand of the public. In this case, the manager starts the recovery program at the first time in order to improve the service level of power system and to minimize operation and recovery costs during the disruption. Notations used in this paper are listed in Table 1.

Given that the recovery process can be fully controlled, the quantity of capacity recovery decision R is rather simple, and $R = D - d$. However, the fact is that random disturbance could change the capacity recovery decision. Therefore, $R = D - d + \overline{X}$, and $\overline{X}$ is the random disturbance with the mean and the variance being φ and σ^2, respectively, in which φ can be either positive or negative. Further, $R = D - d + \varphi + \sigma X = \mu + \sigma X$, where $X \sim (0, 1)$ is the standardized random disturbance with probability density and distribution function being $f(x)$ and $F(x)$, respectively. The mean and the variance of R are μ and σ^2, respectively. In this case, μ is the decision variable, which is assumed to be studied in the following sections.

It is assumed that the electricity demand of the public is fully satisfied by the operational capacity of the power system. The capacity per unit can meet the demand per unit, and excess capacity is unable to be lagged to the next period; thus, there is no electricity inventory cost. Therefore, a newsvendor model can be used to describe the operation cost $c(R)$, which the manager should balance during disruption:

$$c(R) = c_0 d + c_1 R + k[D - d - R]^+. \tag{1}$$

The substitution of $R = \mu + \sigma X$ into (1) gives the disruption cost function, with respect to the mean μ, as

$$C(\mu) = c_0 d + c_1 \mu + k \int_{-\infty}^{(D-d-\mu)/\sigma} (D - d - \mu - \sigma x) f(x)\, dx. \tag{2}$$

Hence, if a manager is completely rational, the optimal capacity recovery μ^* can be obtained by minimizing $C(\mu)$.

Proposition 1. *In the context of a completely rational manager, $C(\mu)$ is a convex function, and there is a unique $\mu^* = \max\{0, D - d - \sigma\lambda\}$ that makes $C(\mu^*) = \min C(\mu)$, where $\lambda = F^{-1}(c_1/k)$.*

Corollary 2. *μ^* is positively correlated with k, and is negatively correlated with c_1; that is, $\partial\mu^*/\partial k > 0, \partial\mu^*/\partial c_1 < 0$.*

Corollary 2 shows that, with the assumption of a rational manager, increasing k (e.g., increasing regulatory punishment for electricity shortage) and decreasing c_1 (e.g., providing subsidy for the recovery of damaged capacity) may strengthen the intention of manager to increase the decision of capacity recovery.

Proposition 1 also demonstrates that any bias away from μ^* will increase the disruption cost. In addition, according to Proposition 1 and Corollary 2, it can be seen that $\lambda = F^{-1}(c_1/k) = D - d - \mu^*/\sigma$ is virtually the upper quartile of the distribution function of X at the probability of c_1/k. And it is clear that (1) $d + \mu^* \geq D$ when $\lambda < 0$, which means that electricity demand can be fully satisfied during disruption; (2) $d + \mu^* < D$ when $\lambda > 0$, which means that there is electricity shortage during disruption. Corollary 2 indicates that increasing regulatory punishment (e.g., increasing k)

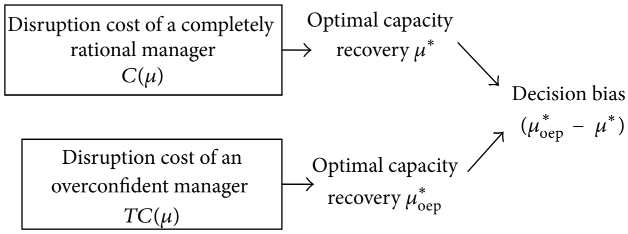


Figure 1: Decision bias model due to overconfidence.

and provision of subsidy for capacity recovery (e.g., reducing c_1) could help to improve power system's service level by increasing the final capacity (e.g., $d + \mu^*$) during disruption. Otherwise, an insufficient service level will be ensured by a low-capacity recovery decision.

4. The Decision Bias of an Overconfident Manager

According to the survey and interview conducted on 50 managers of a district branch of SGCC, very few managers could follow the decision provided by the $C(\mu)$ in a completely rational decision mode, which is described in Section 3. During disruption, an overconfident manager tends to make his/her decisions based on experiences of past successful practice. The more successful past experiences he/she has had, the more optimistically he/she will regard the random disturbance in the decision-making process. Therefore, an overconfident manager makes the decision of capacity recovery (e.g., μ^*_{oep}) by the cost function of $TC(\mu)$ that is given by Model II in Figure 1. Obviously, there is a decision bias of capacity recovery (e.g., $\mu^*_{\text{oep}} - \mu^*$) between Model I and Model II, and, thus, the disruption cost is increased in Model II.

Croson et al. proposed two types of overconfidence: overestimation and overprecision (cf. [33]), which are also followed by this research. First, overestimation will cause manager to make an overoptimistic estimation over the mean φ of external random disturbance; that is, $(D - d + \varphi)/\mu = \alpha \ge 1$, where α is the overestimation factor and $E(R) = \alpha\mu$. The greater the value of α is, the more overestimated the manager behavior is. Second, overprecision behavior would cause the manager to regard the random disturbance variance as $\text{var}(R) = [(1-\beta)\sigma]^2$, where β is the behavioral factor of overprecision. $\beta \to 0$ represents a manager with no overprecision behavior, and $\beta \to 1$ represents a completely overprecise manager.

It is difficult to distinguish between these two types of overconfidence that a manager may engage in decision-making during disruption. Therefore, this paper assumes that both of these two ways of overconfident behavior are taking into effects, and the impact on the decision bias of capacity recovery would be presented by modifying the model of Section 3. Furthermore, the calibration of reducing decision bias will be reinvestigated under an overconfident manager.

Accordingly, the capacity recovery decision made by an overestimated and overprecise manager is R_{oep}, and $R_{\text{oep}} = \alpha\mu_{\text{oep}} + (1-\beta)\sigma X$, with its mean and variance being $E(R_{\text{oep}}) = \alpha\mu_{\text{oep}}$ and $\text{var}(R_{\text{oep}}) = [(1-\beta)\sigma]^2$, respectively, where $\alpha \ge 1$ and $0 \le \beta \le 1$. Correspondingly, the disruption cost function takes the form of $TC_{\text{oep}}(\alpha\mu_{\text{oep}}, \alpha, \beta)$ according to Model II, as (3), where $U = (D - d - \alpha\mu_{\text{oep}})/(1-\beta)\sigma$:

$$\begin{aligned} TC_{\text{oep}}\left(\alpha\mu_{\text{oep}}, \alpha, \beta\right) &= c_0 d + c_1 \alpha\mu_{\text{oep}} \\ &\quad + k\int_{-\infty}^{U}\left(D - d - \alpha\mu_{\text{oep}} - (1-\beta)\,\sigma x\right) f(x)\,dx. \end{aligned} \tag{3}$$

Proposition 3. *The disruption cost $TC_{oep}(\alpha\mu_{oep}, \alpha, \beta)$ given by Model II is convex in μ_{oep}, and there is a unique $\mu^*_{oep} = (\mu^*/\alpha) + (\beta\sigma\lambda/\alpha)$ that yields $TC_{oep}(\mu^*_{oep}, \alpha, \beta) = \min TC_{oep}$.*

The proof of Proposition 3 is similar to that of Proposition 1 and is omitted herein.

Corollary 4. *The decision bias $(\mu^*_{oep} - \mu^*)$ is negatively correlated with α; that is, $\partial(\mu^*_{oep} - \mu^*)/\partial\alpha \le 0$ and $\partial^2(\mu^*_{oep} - \mu^*)/\partial\alpha^2 \ge 0$, and the absolute decision bias has a positive correlation with β; that is, $\partial|\mu^*_{oep} - \mu^*|/\partial\beta = |\sigma\lambda/\alpha| \ge 0$.*

Corollary 5. *μ^*_{oep} is positively correlated with k while being negatively correlated with c_1; that is, $\partial\mu^*_{oep}/\partial k \ge 0$ and $\partial\mu^*_{oep}/\partial c_1 \le 0$, respectively. The decision bias $(\mu^*_{oep} - \mu^*)$ is negatively correlated with k and positively correlated with c_1; that is, $\partial(\mu^*_{oep} - \mu^*)/\partial k \le 0$ and $\partial(\mu^*_{oep} - \mu^*)/\partial c_1 \ge 0$. The cost bias $TC_{oep}(\mu^*_{oep}) - C(\mu^*_{oep})$ is linearly independent of (α, β) and is convex in α and β, respectively.*

Hereto, this paper has addressed the mathematical model modified by overestimation and overprecision behavior and proofs of decision bias with respect to parameters of punishment and subsidy. Furthermore, three inferences can be obtained through Corollaries 2 to 5.

Inference 1. When $\alpha > 1$ and $\beta = 0$, that is, managers only present the behavior of overestimation, α will cause decision bias of capacity recovery through the "multiplicative effect;" that is, $\mu^*_{\text{oe}} = \mu^*/\alpha$. α will inevitably lead to the decrease of the decision of capacity recovery, that is, $(\mu^*_{\text{oe}} - \mu^*) = ((1-\alpha)/\alpha)\mu^* < 0$, which may cause the power system to have a low service level during disruption. Furthermore, $\partial(\mu^*_{\text{oe}} - \mu^*)/\partial\alpha = -\mu^*/\alpha^2 < 0$, which means that even less capacity recovery is decided when α increases.

Inference 2. When $\alpha = 1$ and $\beta \ne 0$, that is, managers only present the behavior of overprecision, β will cause decision bias of capacity recovery through the "additive effect;" that is, $\mu^*_{\text{op}} = \mu^* + \beta\sigma\lambda$. Considering that $\lambda = F^{-1}(c_1/k)$, the decision bias, that is, $(\mu^*_{\text{op}} - \mu^*)$, is dependent on both the distribution of X and the ratio c_1/k. When $\lambda > 0$, factor of β would cause the increasing of the capacity recovery and relieve the negative bias caused by α. In contrast, when $\lambda < 0$, factor of β may further decrease the capacity recovery and deepen the

Table 2: Decision bias ($\mu^*_{\text{oep,lft}} - \mu^*_{\text{oep,nml}}$) for left-skewed disturbance.

	$\beta = 0.0$	$\beta = 0.2$	$\beta = 0.4$	$\beta = 0.6$	$\beta = 0.8$
$\alpha = 1.0$	0.0341	0.0509	0.1019	0.1528	0.2038
$\alpha = 1.2$	0.0404	0.0849	0.1274	0.1699	0.2123
$\alpha = 1.4$	0.0736	0.1091	0.1456	0.1819	0.2183
$\alpha = 1.6$	0.0984	0.1272	0.1592	0.1911	0.2824
$\alpha = 1.8$	0.2078	0.1800	0.1698	0.1983	0.2794

Table 3: Decision bias of ($\mu^*_{\text{oep,rgt}} - \mu^*_{\text{oep,nml}}$) for right-skewed disturbance.

	$\beta = 0.0$	$\beta = 0.2$	$\beta = 0.4$	$\beta = 0.6$	$\beta = 0.8$
$\alpha = 1.0$	−0.0673	−0.09665	−0.1933	−0.29	−0.3866
$\alpha = 1.2$	−0.0806	−0.1611	−0.2416	−0.3221	−0.4027
$\alpha = 1.4$	−0.1384	−0.2071	−0.2757	−0.3452	−0.4142
$\alpha = 1.6$	−0.1816	−0.2418	−0.3019	−0.3621	−0.4233
$\alpha = 1.8$	−0.2152	−0.23	−0.3222	−0.3757	−0.4292

negative bias caused by α, thus causing a lower service level when compared to that when $\alpha = 1$.

Inference 3. Given $\partial(\mu^*_{\text{oep}} - \mu^*)/\partial k \leq 0$ and $\partial(\mu^*_{\text{oep}} - \mu^*)/\partial c_1 \geq 0$, the regulators can calibrate the decision bias aroused by manager's overconfident behavior through increasing regulatory punishment for capacity shortage (i.e., increasing k) and providing subsidy for capacity recovery (i.e., decreasing c_1). However, close attention should be paid to the distribution function of random disturbance and the sign of λ. When the random disturbance is normally distributed and $c_1/k > 0.5$, an overconfident manager will generally increase the decision of capacity recovery. Otherwise, capacity recovery will be decreased. In addition, it should be noted that $\mu^*_{\text{op}} = D - d$ is independent of k and c_1 when $\beta = 1$, which means that punishment and subsidy from regulators will become completely invalid.

5. Numerical Analysis

Decision given by an overconfident manager is investigated by a modified newsvendor model, which could be easily evaluated by first- and second-order moments model of $\mu^*_{\text{oep,nml}} = D - d + z\sigma$, wherein z is the z-value given by the value of c_1/k by assuming that the random disturbance is normally distributed. However, this is not always the case, since many random disturbances present to be asymmetrically distributed with their peaks of distribution functions being left- or right-skewed from the mean value, which means that the third- and fourth-order moments of disturbance would take in function in evaluating μ^*_{oep}. Thus, neglecting the skewness of the random disturbance X will cause additional bias away from μ^*_{oep} (cf. [34]).

In the numerical analysis below, $\mu^*_{\text{oep,nml}}$ is supposed to be the recovery decision given by an overconfident manager, which is also served as the baseline when actual disturbance is not symmetrically distributed. $\mu^*_{\text{oep, lft}}$ and $\mu^*_{\text{oep,rgt}}$ are the capacity recoveries when the peaks of random disturbance distribution are left- and right-skewed, respectively. The skewed random disturbance is given by a partial student distribution expressed as

$$f(z_t \mid \xi, \nu) = \begin{cases} \dfrac{2}{\xi + 1/\xi} sg\left[\xi(sz_t + m) \mid \nu\right], & z_t < -\dfrac{m}{s}, \\ \dfrac{2}{\xi + 1/\xi} sg\left[\dfrac{(sz_t + m)}{\xi} \mid \nu\right], & z_t > -\dfrac{m}{s}, \end{cases} \tag{4}$$

where $s = \xi^2 + (1/\xi^2) - 1 - m^2$, $m = (\Gamma[(\nu-1)/2]\sqrt{\nu-2}/\sqrt{\pi} \cdot \Gamma(\nu/2)) \cdot (\xi - 1/\xi)$, $g[\cdot \mid \nu]$ is a symmetric student distribution, and ν is the parameter of fat-tail. When $0 < \xi < 1$, $f(z_t \mid \xi, \nu)$ is left-skewed, while $\xi > 1$, $f(z_t \mid \xi, \nu)$ is right-skewed. $\nu = 5$, $\xi = 0.5$, and $\xi = 2$ are set with the skewness of 2.1 to the left and to the right correspondingly. Other parameters used in numerical analysis are taken as $d = 1$, $D = 5$, $\sigma = 2$, $c_0 = 2$, $k = 10$, $c_1 = 6$. Tables 2 and 3 show the numerical results for the biases of ($\mu^*_{\text{oep, lft}} - \mu^*_{\text{oep,nml}}$) and ($\mu^*_{\text{oep,rgt}} - \mu^*_{\text{oep,nml}}$), respectively.

From Table 2, we can see that capacity recovery is less for a left-skewed random disturbance than that for a normal distributed disturbance; that is, $\mu^*_{\text{oep,nml}} < \mu^*_{\text{oep, lft}}$. However, $\mu^*_{\text{oep,nml}} > \mu^*_{\text{oep,rgt}}$ according to numerical results given by Table 3. It is largely the result that the negative effect of a left-skewed random disturbance is relatively underestimated by manager who uses the evaluation of $\mu^*_{\text{oep,nml}}$ instead, and it should be offset by a positive calibration value on $\mu^*_{\text{oep,nml}}$ in order to improve the service level during disruption, while the negative effect of a right-skewed random disturbance is relatively overestimated by manager when using $\mu^*_{\text{oep,nml}}$ instead, and more capacity recovery decision is foreseen, which would obviously increase the service level of power system during disruption. As a consequence, replacing $\mu^*_{\text{oep,rgt}}$ with $\mu^*_{\text{oep,nml}}$ could be regarded as having no harm to the interest and welfare of the public.

6. Managerial Insights and Future Research

In this paper, decision bias of capacity recovery is investigated under the scenario of power system's capacity being damaged by unexpected events. Modified newsvendor models are presented to show the decision bias given by an overconfident manager and a completely rational manager, respectively. And several interesting managerial insights as well as calibration methods for improving performance of disruption management are obtained through theoretical proofs and numerical simulations, which are given as below.

Firstly, for the sake of improving the power system's service level during disruption, regulators should increase regulatory punishment for electricity shortage and provide subsidy for capacity recovery at the same time, in order to calibrate the decision biases in reducing the capacity recovery by an overconfident manager. In other words, "*carrot-and-stick*" regulatory mechanism could push managers to recover more damaged capacity. Moreover, fortunately, "*carrot-and-stick*" regulatory mechanism can be always in function no matter the manager is overconfident or not.

Secondly, a manager's overestimation behavior will inevitably lower the decision of capacity recovery; thus, calibration capability of "*carrot-and-stick*" regulatory mechanism could be inevitably weakened, and a lower service level would be foreseen during the disruption of power system. For the interest and welfare of the public, higher punishment for electricity shortage should be prescribed by regulation to improve electricity service level in case of the manager being overconfident.

Thirdly, manager's overprecision behavior may have both negative and positive influences on the capacity recovery decision. More capacity recovery would be decided when the value of the upper quartile $\lambda = F^{-1}(c_1/k)$ of random disturbance is positive; otherwise, less capacity recovery would be decided. Furthermore, for improving the power system's service level during disruption, regulator should pay close attention to ensuring that the ratio c_1/k is less than 50%, which indicates that capacity recovery cost after subsidy must be less 50% the value of punishment for electricity shortage.

Finally, negative impact on capacity recovery would be underestimated under a left-skewed disturbance, and lower electricity service level would be caused due to the decision of less capacity recovery. In contrast, higher electricity service level would be foreseen due to the decision of more capacity under right-skewed disturbance and normal distributed disturbance, respectively. As a result of requiring higher electricity service level by regulator, the situation of right-skewed disturbance should be closely paid attention to by a manger during disruption management.

However, two assumptions made in this paper are expected to be relaxed in future modeling research. The first assumption is that this paper only studies exogenous punishment of electricity shortage, and it would be desirable to investigate an endogenous punishment given by a game model between regulator and manager. The second assumption is that the capacity recovery decision of Model II only considers a single period; however, the overconfident behavior might be changed during the disruption, and thus the correlations of decision between neighboring periods during disruption should be investigated in future research despite the difficulties in creating mathematical model.

Appendix

Proof of Proposition 1. Taking the first and the second derivatives of $C(\mu)$ with respect to μ, we can obtain $\partial C(\mu)/\partial\mu = c_1 - kF((D-d-\mu)/\sigma)$, $\partial^2 C(\mu)/\partial\mu^2 = (k/\sigma)f((D-d-\mu)/\sigma)$. Since $\partial^2 C(\mu)/\partial\mu^2$ is nonnegative, $C(\mu)$ is a convex function, and the optimal solution of μ^* can be derived from $\partial C(\mu)/\partial\mu = 0$; that is, $\mu^* = D-d-\sigma F^{-1}(c_1/k)$. Since μ^* is nonnegative, then $\mu^* = \max\{0, D-d-\sigma\lambda\}$. □

Proof of Corollary 2. Based on Proposition 1, taking the first derivative of μ^* with respect to k and c_1 yields $\partial\mu^*/\partial k = \sigma F(\lambda)/kf(\lambda) > 0$, $\partial\mu^*/\partial c_1 = -\sigma/kf(\lambda) < 0$. □

Proof of Corollary 4. According to Proposition 3, $(\mu^*_{\text{oep}} - \mu^*) = \mu^*/\alpha + \beta\sigma\lambda/\alpha - \mu^*$ when $\mu^* + \beta\sigma\lambda > 0$. Its first-order derivative with α is $\partial(\mu^*_{\text{oep}} - \mu^*)/\partial\alpha = (-1/\alpha^2)(\mu^* + \beta\sigma\lambda) \le 0$. Similarly, the second-order derivative with α is $\partial^2(\mu^*_{\text{oep}} - \mu^*)/\partial\alpha^2 = (1/\alpha^3)(\mu^* + \beta\sigma\lambda)^2 \ge 0$. Similarly, $\partial(\mu^*_{\text{oep}} - \mu^*)/\partial\beta = \sigma\lambda/\alpha$, and it can be easily obtained that the sign of $(\mu^*_{\text{oep}} - \mu^*)$ depends on λ. The absolute decision bias is nonnegative; that is, $\partial|\mu^*_{\text{oep}} - \mu^*|/\partial\beta = |\sigma\lambda/\alpha| \ge 0$. □

Proof of Corollary 5. According to Proposition 3, it can be obtained that $\partial\mu^*_{\text{oep}}/\partial k = (1/\alpha)((\partial\mu^*/\partial k) + \beta\sigma(\partial\lambda/\partial k)) = \sigma(1-\beta)F(\lambda)/\alpha kf(\lambda) \ge 0$, $\partial\mu^*_{\text{oep}}/\partial c_1 = \sigma(\beta-1)/\alpha kf(\lambda) \le 0$. It can also be obtained that $(\partial\mu^*_{\text{oep}}/\partial k) - (\partial\mu^*/\partial k) = -\sigma(\alpha + \beta - 1)F(\lambda)/\alpha kf(\lambda) \le 0$, $(\partial\mu^*_{\text{oep}}/\partial c_1) - (\partial\mu^*/\partial c_1) = \sigma(\alpha + \beta - 1)/\alpha kF(\lambda) \ge 0$. Similarly, it could be obtained that $\partial(TC_{\text{oep}} - C)/\partial\alpha = \mu^*_{\text{oep}}[c_1 - kF(U)]$, $\partial(TC_{\text{oep}} - C)/\partial\beta = \sigma k\int_{-\infty}^{U} xf(x)dx$, and then $\partial^2(TC_{\text{oep}} - C)/\partial\alpha^2 = (\mu^2_{\text{oep}}/(1-\beta)\sigma)kf(U) \ge 0$, $\partial^2(TC_{\text{oep}} - C)/\partial\beta^2 = (\sigma/(1-\beta))kU^2 f(U) \ge 0$, $\partial^2(TC_{\text{oep}} - C)/\partial\alpha\partial\beta = -(\mu^*_{\text{oep}}/(1-\beta))kUf(U)$. It can be derived that the Hessian matrix is 0. Therefore, the cost bias $TC_{\text{oep}}(\mu^*_{\text{oep}}) - C(\mu^*_{\text{oep}})$ is linearly independent of (α, β), and fortunately $TC_{\text{oep}}(\mu^*_{\text{oep}}) - C(\mu^*_{\text{oep}})$ is convex in α and β, respectively. It means that the influences of α and β on the disruption cost bias are independent of each other. The calibration methods on α would have no influence on those on β. □

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The author thanks the support of the National Natural Science Foundation of China (Grants nos. 71302033 and 71301146), the Research Fund for the Doctoral Program of Higher Education of China (Grant no. 2011332612002), the Zhejiang Provincial Natural Science Foundation of China (Grants nos. LQ12G01004 and LY13G020011), and the Zhejiang Provincial Social Science Foundation of China (Grant no. 13ZJQN048YB).

References

[1] R. Germain, C. Dröge, and W. Christensen, "The mediating role of operations knowledge in the relationship of context with performance," *Journal of Operations Management*, vol. 19, no. 4, pp. 453–469, 2001.

[2] K. P. Gopalakrishna and S. Min, "A firm's capability to calibrate supply chain knowledge-Antecedents and consequences," *Industrial Marketing Management*, vol. 39, no. 8, pp. 1365–1375, 2010.

[3] F. Gino and G. P. Pisano, "Why leaders don't learn from success," *Harvard Business Review*, vol. 89, no. 4, pp. 68–74, 2011.

[4] F. Gino and G. Pisano, "Toward a theory of behavioral operations," *Manufacturing and Service Operations Management*, vol. 10, no. 4, pp. 676–691, 2008.

[5] A. Swierczek, "The impact of supply chain integration on the "snowball effect" in the transmission of disruptions: an empirical evaluation of the model," *International Journal of Production Economics*, 2013.

[6] Y. Sheffi, "Supply chain management under the threat of international terrorism," *International Journal of Logistics Management*, vol. 12, no. 2, pp. 1–11, 2001.

[7] A. Norrman and U. Jansson, "Ericsson's proactive risk management approach after a serious sub-supplier accident," *International Journal of Physical Distribution and Logistics Management*, vol. 34, no. 5, pp. 434–456, 2004.

[8] K. B. Hendricks and V. R. Singhal, "An empirical analysis of the effect of supply chain disruptions on long-run stock price performance and equity risk of the firm," *Production and Operations Management*, vol. 14, no. 1, pp. 35–52, 2005.

[9] A. Oke and M. Gopalakrishnan, "Managing disruptions in supply chains: a case study of a retail supply chain," *International Journal of Production Economics*, vol. 118, no. 1, pp. 168–174, 2009.

[10] S. Chopra and M. S. Sohi, "Managing risk to avoid supply chain breakdown," *MIT Sloan Management Review*, vol. 46, no. 1, pp. 53–62, 2004.

[11] P. R. Kleindorfer and G. H. Saad, "Managing disruption risks in supply chains," *Production and Operations Management*, vol. 14, no. 1, pp. 53–68, 2005.

[12] G. A. Zsidisin, S. A. Melnyk, and G. L. Ragatz, "An institutional theory perspective of business continuity planning for purchasing and supply management," *International Journal of Production Research*, vol. 43, no. 16, pp. 3401–3420, 2005.

[13] C. S. Tang, "Review: perspectives in supply chain risk management," *International Journal of Production Economics*, vol. 103, pp. 451–488, 2006.

[14] P. D. Berger, A. Gerstenfeld, and A. Z. Zeng, "How many suppliers are best? A decision-analysis approach," *Omega*, vol. 32, no. 1, pp. 9–15, 2004.

[15] P. L. Meena, S. P. Sarmah, and A. Sarkar, "Sourcing decisions under risks of catastrophic event disruptions," *Transportation Research E: Logistics and Transportation Review*, vol. 47, no. 6, pp. 1058–1074, 2011.

[16] A. Sarkar and P. K. J. Mohapatra, "Determining the optimal size of supply base with the consideration of risks of supply disruptions," *International Journal of Production Economics*, vol. 119, no. 1, pp. 122–135, 2009.

[17] A. J. Schmitt, "Strategies for customer service level protection under multi-echelon supply chain disruption risk," *Transportation Research B: Methodological*, vol. 45, no. 8, pp. 1266–1283, 2011.

[18] X. Qi, J. F. Bard, and G. Yu, "Disruption management for machine scheduling: the case of SPT schedules," *International Journal of Production Economics*, vol. 103, no. 1, pp. 166–184, 2006.

[19] M. Wu and W. Chang, "A short-term capacity trading method for semiconductor fabs with partnership," *Expert Systems with Applications*, vol. 33, no. 2, pp. 476–483, 2007.

[20] B. Tomlin, "On the value of mitigation and contingency strategies for managing supply chain disruption risks," *Management Science*, vol. 52, no. 5, pp. 639–657, 2006.

[21] M. C. Wilson, "The impact of transportation disruptions on supply chain performance," *Transportation Research E: Logistics and Transportation Review*, vol. 43, no. 4, pp. 295–320, 2007.

[22] L. V. Synder and B. T. Tomlin, "Inventory management with advanced warning of disruptions," Working Paper, P. C. Rossin College of Engineering and Applied Sciences, Lehigh University, Bethlehm, Pa, USA, 2008.

[23] A. A. Wadecki, V. Babich, and O. Q. Wu, "Manufacturer's competition and subsidies to suppliers," Working Paper, 2010, http://ssrn.com/abstract=1616949.

[24] X. X. Hu, H. Gurnani, and L. Wang, "Managing risk of supply disruptions: Incentives for capacity restoration," *Production and Operations Management*, vol. 22, no. 1, pp. 137–150, 2013.

[25] Z. Yang, G. Aydin, V. Babich, and D. R. Beil, "Supply disruptions, asymmetric information, and a backup production option," *Management Science*, vol. 55, no. 2, pp. 192–209, 2009.

[26] E. Bendoly, R. Croson, P. Goncalves, and K. Schultz, "Bodies of knowledge for research in behavioral operations," *Production and Operations Management*, vol. 19, no. 4, pp. 434–452, 2010.

[27] D. Y. Wu and K. Y. Chen, "Supply chain contract design: impact of bounded rationality and individual heterogeneity," *Production and Operation Management*, vol. 23, no. 2, pp. 253–268, 2014.

[28] U. Malmendier and G. Tate, "CEO overconfidence and corporate investment," *The Journal of Finance*, vol. 60, no. 6, pp. 2661–2700, 2005.

[29] S. Rafaeli and G. Ravid, "Information sharing as enabler for the virtual team: an experimental approach to assessing the role of electronic mail in disintermediation," *Information Systems Journal*, vol. 13, no. 2, pp. 191–206, 2003.

[30] F. de Vericourt, K. Jain, J. N. Bearden et al., "Sex, risk and the newsvendor," *Journal of Operations Management*, vol. 31, pp. 86–92, 2013.

[31] Y. F. Ren and R. Croson, "Overconfidence in newsvendor orders: an experimental study," *Management Science*, vol. 59, no. 11, pp. 2502–2517, 2013.

[32] L. V. Synder, Z. Atan, P. Peng, Y. Rong, and A. J. Schmitt, "OR/MS Models for supply chain disruptions: a review," Working Paper, 2012, http://ssrn.com/abstract=1689882.

[33] R. Croson, D. Croson, and Y. Ren, "How to manage an overconfident newsvendor," Working Paper, University of Texas at Dallas, 2008.

[34] O. Tang and R. W. Grubbström, "On using higher-order moments for stochastic inventory systems," *International Journal of Production Economics*, vol. 104, no. 2, pp. 454–461, 2006.

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 193143, 7 pages
http://dx.doi.org/10.1155/2014/193143

Research Article

Bifurcation in a Discrete Competition System

Li Xu, Lianjun Zou, Zhongxiang Chang, Shanshan Lou, Xiangwei Peng, and Guang Zhang

School of Science, Tianjin University of Commerce, Tianjin 300134, China

Correspondence should be addressed to Li Xu; beifang_xl@163.com

Received 9 February 2014; Accepted 9 April 2014; Published 28 April 2014

Academic Editor: Xiaochen Sun

Copyright © 2014 Li Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A new difference system is induced from a differential competition system by different discrete methods. We give theoretical analysis for local bifurcation of the fixed points and derive the conditions under which the local bifurcations such as flip occur at the fixed points. Furthermore, one- and two-dimensional diffusion systems are given when diffusion terms are added. We provide the Turing instability conditions by linearization method and inner product technique for the diffusion system with periodic boundary conditions. A series of numerical simulations are performed that not only verify the theoretical analysis, but also display some interesting dynamics.

1. Introduction

Interactions of different species may take many forms such as competition, predation, parasitism, and mutualism. One of the most important interactions is the competition relationship. The dynamic relationship between the two competition species is one of the dominant subjects in mathematical ecology due to its universal existence and importance. Lotka-Volterra competition systems are ecological models that describe the interaction among various competing species and have been extensively investigated in recent years (see [1–3] and the references therein). In the earlier literature, the two-competing species competition models are often formulated in the form of ordinary differential systems as follows:

$$\begin{aligned} u'(t) &= u(t)\left(r_1 - a_{11}u(t) - a_{12}v(t)\right), \\ v'(t) &= v(t)\left(r_2 - a_{21}v(t) - a_{22}v(t)\right), \end{aligned} \tag{1}$$

for $t \in [0, +\infty)$ $a_{ij} \geq 0, i, j = 1, 2$, where $u(t)$ and $v(t)$ are the quantities of the two species at time t, $r_1 > 0$ and $r_2 > 0$ are growth rates of the respective species, a_{11} and a_{22} represent the strength of the intraspecific competition, and a_{12} and a_{21} represent the strength of the interspecific competition.

The discrete time models governed by difference equation are more realistic than the continuous ones when the populations have nonoverlapping generations or the population statistics are compiled from given time intervals and not continuously. Moreover, since the discrete time models can also provide efficient computational models of continuous models for numerical simulations, it is reasonable to study discrete time models governed by difference equations.

Applying forward Euler scheme to the first equation of system (1) and obtaining a discrete analog of the second equation by considering a variation with piecewise constant arguments for certain terms on the right side (exponential discrete form) [4], we obtain the following equation:

$$\begin{aligned} u_{t+1} &= u_t\left(r_1 - a_{11}u_t - a_{12}v_t\right), \\ v_{t+1} &= v_t \exp\left(r_2 - a_{21}u_t - a_{22}v_t\right). \end{aligned} \tag{2}$$

For the sake of simplicity, let

$$\begin{aligned} u_{t+1} &= u_t\left(r_1 + 1 - a_{11}u_t - a_{12}v_t\right), \\ v_{t+1} &= v_t \exp\left(r_2 - a_{21}u_t - a_{22}v_t\right). \end{aligned} \tag{3}$$

By setting $U_t = a_{11}u_t$ and $V_t = a_{22}v_t$ and $r_1 = r_2, a_{11} = a_{22}$, and $a_{12} = a_{21}$, we have the following form:

$$\begin{aligned} u_{t+1} &= u_t\left(r + 1 - u_t - av_t\right) = f\left(u_t, v_t\right), \\ v_{t+1} &= v_t \exp\left(r - au_t - v_t\right) = g\left(u_t, v_t\right). \end{aligned} \tag{4}$$

Although numerical variations of system (1) have been extensively studied (see, e.g., the work in [5–8]), some discrete analogs may be found in [9–15], regarding attractivity, persistence, global stabilities of equilibrium, and other dynamics. Up to now, to the best of our knowledge, the discrete system (4) has not been investigated.

Since the pioneering theoretical works of Skellam [15] and Turing [16], many works have focused on the effect of spatial factors which play a crucial role in the stability of populations [17–19]. Many important epidemiological and ecological phenomena are strongly influenced by spatial heterogeneities because of the localized nature of transmission or other forms of interaction. Thus, spatial models are more suitable for describing the process of population development. Impact of spatial component on system has been widely investigated (e.g., see [20–22]). It may be a case in reality that the motion of individuals is random and isotropic; that is, without any preferred direction, the individuals are also absolute ones in microscopic sense, and each isolated individual exchanges materials by diffusion with its neighbors [19, 23]. Thus, it is reasonable to consider a 1D or 2D spatially discrete reaction diffusion system to explain the population system. Corresponding to the above analysis, we can obtain the following one-dimensional diffusion systems:

$$\begin{aligned} u_i^{t+1} &= u_i^t\left(r+1-u_i^t-av_i^t\right)+d_1\nabla^2 u_i^t,\\ v_i^{t+1} &= v_i^t\exp\left(r-au_i^t-v_i^t\right)+d_2\nabla^2 v_i^t, \end{aligned} \tag{5}$$

for $i \in \{1,2,\ldots,m\} = [1,m]$, $t \in Z^+$ and $\nabla^2 u_i^t = u_{i+1}^t - 2u_i^t + u_{i-1}^t$, $\nabla^2 v_i^t = v_{i+1}^t - 2v_i^t + v_{i-1}^t$, and two-dimensional diffusion systems:

$$\begin{aligned} u_{ij}^{t+1} &= u_{ij}^t\left(r+1-u_{ij}^t-av_{ij}^t\right)+d_1\nabla^2 u_{ij}^t,\\ v_{ij}^{t+1} &= v_{ij}^t\exp\left(r-au_{ij}^t-v_{ij}^t\right)+d_2\nabla^2 v_{ij}^t, \end{aligned} \tag{6}$$

for $i,j \in \{1,2,\ldots,m\} = [1,m]$, $t \in R^+ = [0,\infty)$ and $\nabla^2 u_{ij}^t = u_{i+1,j}^t + u_{i,j+1}^t + u_{i-1,j}^t + u_{i,j-1}^t - 4u_{ij}^t$, $\nabla^2 v_{ij}^t = v_{i+1,j}^t + v_{i,j+1}^t + v_{i-1,j}^t + v_{i,j-1}^t - 4v_{ij}^t$.

In this paper, we will study the dynamical behaviors of models (4), (5), and (6). By using the theory of difference equation, the theory of bifurcation, and the center manifold theorem we will establish the series of criteria on the existence and local stability of equilibria, flip bifurcation for the system (4). For the one- or two-dimensional diffusion systems, with periodic boundary conditions, the Turing instability (or Turing bifurcation) theory analysis will be given. Turing instability conditions can then be deduced combining linearization method and inner product technique. Furthermore, by means of the numerical simulations method, we will indicate the correctness and rationality of our results.

The paper is organized as follows. In Section 2, we study the existence and stability of equilibria points and the conditions of existence for flip bifurcation are verified for system (4). Turing instability conditions will be illustrated by linearization method and inner product technique for the system (5) and (6) with periodic boundary conditions in Section 3. A series of numerical simulations are performed that not only verify the theoretical analysis, but also display some interesting dynamics. For the system (4), the bifurcation diagrams are given. The impact of the system parameters and diffusion coefficients on patterns can also be observed visually for the given diffusion systems. Finally, some conclusions are given.

2. Analysis of Equilibria and Flip Bifurcation

Clearly, the system (4) has four possible steady states; that is, $E_0 = (0,0)$, exclusion points $E_1 = (r,0)$, $E_2 = (0,r)$, and nontrivial coexistence point $E_3 = (u^*, v^*)$, where

$$u^* = v^* = \frac{r}{a+1}. \tag{7}$$

The linearized form of (4) is then

$$\begin{aligned} u_{t+1} &= f_u u_t + f_v v_t,\\ v_{t+1} &= g_u u_t + g_v v_t, \end{aligned} \tag{8}$$

which has the Jacobian matrix

$$J_{E_i} = \begin{bmatrix} f_u & f_v \\ g_u & g_v \end{bmatrix}_{P_i} = \begin{bmatrix} 1-u & -au \\ -av & 1-v \end{bmatrix}_{P_i}, \quad i = 0,1,2,3. \tag{9}$$

The characteristic equation of the Jacobian matrix J can be written as

$$\lambda^2 + p\lambda + q = 0, \tag{10}$$

where $p = -(f_u + g_v)$ and $q = f_u g_v - f_v g_u$.

In order to discuss the stability of the fixed points of (4), we also need the following definitions [20]:

(1) if $|\lambda_1| < 1$ and $|\lambda_2| < 1$, then steady state E is called a sink and E is locally asymptotically stable;

(2) if $|\lambda_1| > 1$ and $|\lambda_2| > 1$, then E is called a source and E is unstable;

(3) if $|\lambda_1| > 1$ and $|\lambda_2| < 1$ (or $|\lambda_1| < 1$ and $|\lambda_2| > 1$), then E is called a saddle;

(4) if either $|\lambda_1| = 1$ and $|\lambda_2| \neq 1$ or $|\lambda_2| = 1$ and $|\lambda_1| \neq 1$, then E is called nonhyperbolic.

Case 1 (the fixed point $E_0 = (0,0)$). The linearization of (4) about E_0 has the Jacobian matrix

$$J_{E_0} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{11}$$

which has two eigenvalues

$$\lambda_1 = \lambda_2 = 1. \tag{12}$$

The fact means that the system is resonance at the fixed point P_0.

Case 2 (the fixed point $E_1 = (0,r)$). At the fixed point, the Jacobian matrix has the form

$$J_{E_1} = \begin{bmatrix} 1 & 0 \\ -ar & 1-r \end{bmatrix}, \tag{13}$$

and the corresponding eigenvalues of (13) are

$$\lambda_1 = 1, \qquad \lambda_2 = 1 - r. \tag{14}$$

r is a bifurcation parameter. And $r \neq 2$ implies $\lambda_2 \neq -1$, and the fixed point E_1 is nonhyperbolic.

Case 3 (the third fixed point $E_2 = (r, 0)$). The linearization of (4) about E_2 has the Jacobian matrix

$$J_{E_1} = \begin{bmatrix} 1-r & -ar \\ 0 & 1 \end{bmatrix}, \tag{15}$$

and the eigenvalues of (15) are

$$\lambda_1 = 1 - r, \qquad \lambda_2 = 1. \tag{16}$$

r is a bifurcation parameter. And $r \neq 2$ implies $\lambda_1 \neq -1$, and the fixed point E_2 is nonhyperbolic.

Case 4 (the fixed point $E_3 = (r/(a+1), r/(a+1))$). The linearization of (4) about E_3 has the Jacobian matrix

$$J_{E_3} = \begin{bmatrix} 1 - \dfrac{r}{a+1} & -\dfrac{ar}{a+1} \\ -\dfrac{ar}{a+1} & 1 - \dfrac{r}{a+1} \end{bmatrix}, \tag{17}$$

and the eigenvalues of (17) are

$$\lambda_1 = 1 - r, \qquad \lambda_2 = 1 + \frac{(a-1)r}{a+1}; \tag{18}$$

then, we have the following results:

(1) $|\lambda_1| < 1, |\lambda_2| < 1$ if and only if $0 < r < 2, 0 < a < 1$;

(2) $\lambda_1 = -1, |\lambda_2| \neq 1$ if and only if $r = 2, a \neq 1$;

(3) $\lambda_2 = 1, |\lambda_1| \neq 1$ if and only if $a = 2, r \neq 2$;

(4) $\lambda_2 = -1, |\lambda_1| \neq 1$ if and only if $a = (r-2)/(r+2), r \neq 2$;

(5) $|\lambda_1| < 1, |\lambda_2| > 1$ if and only if $0 < r < 2, a > 1$;

(6) $|\lambda_1| > 1, |\lambda_2| < 1$ if and only if $2 < r < (a+1)/(a-1)$;

(7) $|\lambda_1| > 1, |\lambda_2| > 1$ if and only if $r > 2, a > 1$.

The following theorem is the case that the fixed point E_3 is a flip bifurcation point.

Theorem 1. *The positive fixed point E_3 undergoes a flip bifurcation at the threshold $r_F = 2$.*

Proof. Let $\zeta_n = u_n - u^*$, $\eta_n = v_n - v^*$, and $\mu_n = r - 2$, and parameter μ_n is a new and dependent variable; the system (4) becomes

$$\begin{pmatrix} \zeta_{n+1} \\ \eta_{n+1} \\ \mu_{n+1} \end{pmatrix} = \begin{pmatrix} \left(\zeta_n + \dfrac{\mu_n + 2}{a+1}\right)(1 - \zeta_n - a\eta_n) - \dfrac{\mu_n + 2}{a+1} \\ \left(\eta_n + \dfrac{\mu_n + 2}{a+1}\right)\exp(-a\zeta_n - \eta_n) - \dfrac{\mu_n + 2}{a+1} \\ \mu_n \end{pmatrix}. \tag{19}$$

Let

$$T = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \tag{20}$$

then

$$T^{-1} = \begin{bmatrix} \dfrac{1}{2} & \dfrac{1}{2} & 0 \\ -\dfrac{1}{2} & \dfrac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{21}$$

By the following transformation:

$$\begin{pmatrix} \zeta_n \\ \eta_n \\ \mu_n \end{pmatrix} = T \begin{pmatrix} x_n \\ y_n \\ \delta_n \end{pmatrix}, \tag{22}$$

the system (19) can be changed into

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \\ \delta_{n+1} \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \dfrac{3a-1}{a+1} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_n \\ y_n \\ \delta_n \end{pmatrix} + \begin{pmatrix} f(x_n, y_n, \delta_n) \\ g(x_n, y_n, \delta_n) \\ 0 \end{pmatrix}, \tag{23}$$

where

$$\begin{aligned} f(x_n, y_n, \delta_n) = & -(a+1)x_n^2 - 2(a+1)x_n y_n - x_n \delta_n \\ & + \left(a - 1 + \frac{(a-1)^2}{a+1}\right) y_n^2 \\ & + o\left(\left(|x_n| + |y_n| + |\delta_n|\right)^3\right), \\ g(x_n, y_n, \delta_n) = & (a+1)x_n^2 - 2(a+1)x_n y_n \\ & - \frac{2(a-1)}{a+1} y_n \delta_n + \frac{(a-1)^2}{a+1} y_n^2 \\ & + o\left(\left(|x_n| + |y_n| + |\delta_n|\right)^3\right). \end{aligned} \tag{24}$$

Then, we can consider

$$y_n = h(x_n, \delta_n) = a_1 x_n^2 + a_2 x_n \delta_n + a_3 \delta_n^2 + o\left(\left(|x_n| + |\delta_n|\right)^3\right), \tag{25}$$

which must satisfy

$$\begin{aligned} & h(-x_n + f(x_n, y_n, \delta_n), \delta_{n+1}) \\ & = \frac{3a-1}{a+1} h(x_n, \delta_n) + (a+1)x_n^2 \\ & \quad - 2(a+1)x_n h(x_n, \delta_n) \\ & \quad - \frac{2(a-1)}{a+1} h(x_n, \delta_n)\delta_n \\ & \quad + \frac{(a-1)^2}{a+1} h^2(x_n, \delta_n) \\ & \quad + o\left(\left(|x_n| + |y_n| + |\delta_n|\right)^3\right). \end{aligned} \tag{26}$$

By calculating, we can get that

$$a_1 = \frac{(a+1)^2}{2(1-a)}, \qquad a_2 = 0, \qquad a_3 = 0, \tag{27}$$

and the system (19) is restricted to the center manifold, which is given by

$$\begin{aligned} f: x_{n+1} = & -x_n + \frac{a^3+4a^2+3a}{a-1}x_n^2 + (a+1)x_n\delta_n \\ & + \frac{1}{4}(a+1)^3 x_n^4 + o\left(|x_n|^4\right). \end{aligned} \tag{28}$$

Since

$$\begin{aligned} \left(\frac{\partial f}{\partial \delta}\frac{\partial^2 f}{\partial x^2} + 2\frac{\partial^2 f}{\partial x \partial \delta}\right)\Bigg|_{(0,0)} &= 2(a+1) \neq 0, \\ \left(\frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\right)^2 + \frac{1}{3}\frac{\partial^3 f}{\partial x^3}\right)\Bigg|_{(0,0)} &= 2\left(\frac{a^3+4a^2+3a}{a-1}\right)^2 > 0, \end{aligned} \tag{29}$$

system (4) undergoes a flip bifurcation at E_3. The proof is completed. □

3. Turing Bifurcation

In this section, we discuss the Turing bifurcation. Turing's theory shows that diffusion could destabilize an otherwise stable equilibrium of the reaction-diffusion system and lead to nonuniform spatial patterns. This kind of instability is usually called Turing instability or diffusion-driven instability [16].

3.1. One-Dimensional Case. We consider the following diffusion system:

$$\begin{aligned} u_i^{t+1} &= u_i^t\left(r+1-u_i^t - av_i^t\right) + d_1\nabla^2 u_i^t, \\ v_i^{t+1} &= v_i^t \exp\left(r - au_i^t - v_i^t\right) + d_2\nabla^2 v_i^t, \end{aligned} \tag{30}$$

with the periodic boundary conditions

$$\begin{aligned} u_0^t &= u_m^t, \qquad u_1^t = u_{m+1}^t, \\ v_0^t &= v_m^t, \qquad v_1^t = v_{m+1}^t, \end{aligned} \tag{31}$$

for $i \in \{1, 2, \ldots, m\} = [1, m]$ and $t \in Z^+$, where m is a positive integer,

$$\begin{aligned} \nabla^2 u_i^t &= u_{i+1}^t - 2u_i^t + u_{i-1}^t, \\ \nabla^2 v_i^t &= v_{i+1}^t - 2v_i^t + v_{i-1}^t. \end{aligned} \tag{32}$$

In order to study Turing instability of (30) and (31), we firstly consider eigenvalues of the following equation:

$$\nabla^2 X^i + \lambda X^i = 0 \tag{33}$$

with the periodic boundary conditions

$$X^0 = X^m, \qquad X^1 = X^{m+1}. \tag{34}$$

By calculating, the eigenvalue problem (33)-(34) has the eigenvalues

$$\lambda_s = 4\sin^2\frac{(s-1)\pi}{m} \quad \text{for } s \in [1, m]. \tag{35}$$

We linearise at the steady state, to get

$$w_i^{t+1} = Jw_i^t + D\nabla^2 w_i^t, \quad D = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \tag{36}$$

with the periodic boundary conditions

$$w_0^t = w_m^t, \qquad w_1^t = w_{m+1}^t, \tag{37}$$

where

$$w_i^t = \begin{pmatrix} u_i^t - u^* \\ v_i^t - v^* \end{pmatrix} = \begin{pmatrix} x_i^t \\ y_i^t \end{pmatrix}. \tag{38}$$

Then, respectively, taking the inner product of (36) with the corresponding eigenfunction X_s^i of the eigenvalue λ_s, we see that

$$\begin{aligned} \sum_{i=1}^m X_s^i x_i^{t+1} &= f_u \sum_{i=1}^m X_s^i x_i^t + f_v \sum_{i=1}^m X_s^i y_i^t + d_1 \sum_{i=1}^m X_s^i \nabla^2 x_i^t, \\ \sum_{i=1}^m X_s^i y_i^{t+1} &= g_u \sum_{i=1}^m X_s^i x_i^t + g_v \sum_{i=1}^m X_s^i y_i^t + d_2 \sum_{i=1}^m X_s^i \nabla^2 y_i^t. \end{aligned} \tag{39}$$

Let $U^t = \sum_{i=1}^m X_s^i x_i^t$ and $V^t = \sum_{i=1}^m X_s^i y_i^t$ and use the periodic boundary conditions (34) and (37); then we have

$$\begin{aligned} U^{t+1} &= f_u U^t + f_v V^t - d_1\lambda_s U^t, \\ V^{t+1} &= g_u U^t + g_v V^t - d_2\lambda_s V^t \end{aligned} \tag{40}$$

or

$$\begin{aligned} U^{t+1} &= \left(f_u - d_1\lambda_s\right)U^t + f_v V^t, \\ V^{t+1} &= g_u U^t + \left(g_v - d_2\lambda_s\right)V^t. \end{aligned} \tag{41}$$

Thus, the following fact can be obtained.

Proposition 2. *If* (u_i^t, v_i^t) *is a solution of the problem of* (30) *and* (31)*, then*

$$\left(U^t = \sum_{i=1}^m X_s^i x_i^t, V^t = \sum_{i=1}^m X_s^i y_i^t\right) \tag{42}$$

is a solution of (41)*, where* λ_s *is some eigenvalue of* (33)-(34) *and* X_s^i *is the corresponding eigenfunction. For some eigenvalue* λ_s *of* (33)-(34)*, if* (U^t, V^t) *is a solution of the system* (41)*, then*

$$\left(u_i^t = U^t X_s^i, \ v_i^t = V^t X_s^i\right) \tag{43}$$

is a solution of (30) *with the periodic boundary conditions* (31).

Proposition 3. *If there exist positive numbers d_1, d_2 and the eigenvalue λ_s of the problem* (33)-(34) *such that one of the conditions*

$$\begin{aligned} h(\lambda_s) &< (\lambda_s(d_1+d_2)-(f_{u^*}+g_{v^*}))-1, \\ h(\lambda_s) &< -(\lambda_s(d_1+d_2)-(f_{u^*}+g_{v^*}))-1 \end{aligned} \tag{44}$$

or

$$h(\lambda_s) > 1 \tag{45}$$

holds, then the problem (30) *and* (31) *at the fixed point* (u^*, v^*) *is unstable, where*

$$h(\lambda_s) = d_1 d_2 \lambda_s^2 - (d_1 g_{v^*} + d_2 f_{u^*})\lambda_s + (f_{u^*} g_{v^*} - f_{v^*} g_{u^*}). \tag{46}$$

For the system (5), we have the following results about instability of the positive equilibrium of system.

Theorem 4. $0 < r < 2$, $0 < a < 1$ *and Proposition 3 mean or show that the problem* (30) *and* (31) *is diffusion-driven unstable or Turing unstable.*

3.2. Two-Dimensional Case. In this subsection, we will pay our attention to the Turing instability analysis for the following two-dimensional system:

$$\begin{aligned} u_{ij}^{t+1} &= u_{ij}^t\left(r+1-u_{ij}^t-av_{ij}^t\right)+d_1\nabla^2 u_{ij}^t, \\ v_{ij}^{t+1} &= v_{ij}^t\exp\left(r-au_{ij}^t-v_{ij}^t\right)+d_2\nabla^2 v_{ij}^t, \end{aligned} \tag{47}$$

with the periodic boundary conditions

$$\begin{aligned} u_{i,0}^t &= u_{i,m}^t, \qquad u_{i,1}^t = u_{i,m+1}^t, \\ u_{0,j}^t &= u_{m,j}^t, \qquad u_{1,j}^t = u_{m+1,j}^t, \\ v_{i,0}^t &= v_{i,m}^t, \qquad v_{i,1}^t = v_{i,m+1}^t, \\ v_{0,j}^t &= v_{m,j}^t, \qquad v_{1,j}^t = v_{m+1,j}^t, \end{aligned} \tag{48}$$

for $i, j \in \{1, 2, \ldots, m\} = [1, m]$ and $t \in Z^+$, where m is a positive integer,

$$\begin{aligned} \nabla^2 u_{ij}^t &= u_{i+1,j}^t + u_{i,j+1}^t + u_{i-1,j}^t + u_{i,j-1}^t - 4u_{ij}^t, \\ \nabla^2 v_{ij}^t &= v_{i+1,j}^t + v_{i,j+1}^t + v_{i-1,j}^t + v_{i,j-1}^t - 4v_{ij}^t. \end{aligned} \tag{49}$$

The following theorem will show that the system (47) also undergoes Turing instability. Since the analysis is very similar to the one-dimensional case, the proof is omitted.

Theorem 5. *If there exist positive numbers d_1, d_2 and the eigenvalue k_{ls}^2 of the corresponding characteristic equation such that one of the conditions*

$$\begin{aligned} h\left(k_{ls}^2\right) &< \left(k_{ls}^2(d_1+d_2)-(f_{u^*}+g_{v^*})\right)-1, \\ h\left(k_{ls}^2\right) &< -\left(k_{ls}^2(d_1+d_2)-(f_{u^*}+g_{v^*})\right)-1 \end{aligned} \tag{50}$$

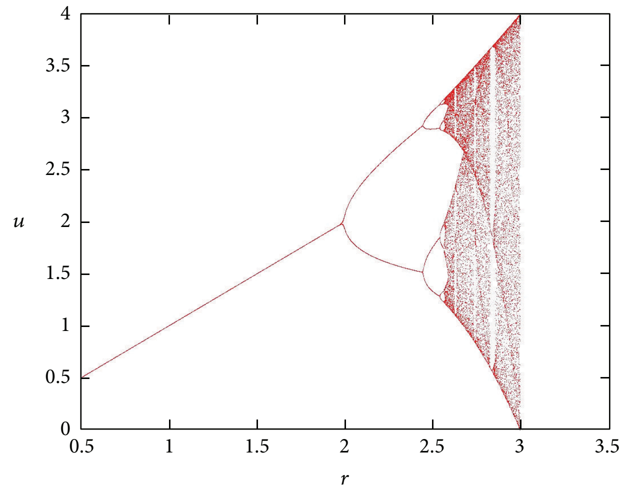


Figure 1: Bifurcation diagram for $r - u_t$ with $a = 0.511$.

or

$$h\left(k_{ls}^2\right) > 1 \tag{51}$$

holds and $0 < r < 2$, $0 < a < 1$ *then the problem* (47)-(48) *at the fixed point* (u^*, v^*) *is diffusion-driven unstable or Turing unstable, where*

$$\begin{aligned} h\left(k_{ls}^2\right) &= d_1 d_2 k_{ls}^4 - (d_1 g_{v^*} + d_2 f_{u^*}) k_{ls}^2 + (f_{u^*} g_{v^*} - f_{v^*} g_{u^*}), \\ k_{ls}^2 &= \lambda_{l,s} = 4\left(\sin^2\left(\frac{(l-1)\pi}{m}\right) + \sin^2\left(\frac{(s-1)\pi}{m}\right)\right) \\ &\qquad \textit{for } l, s \in [1, m]. \end{aligned} \tag{52}$$

4. Numerical Simulation

As is known to all, the bifurcation diagram provides a general view of the evolution process of the dynamical behaviors by plotting a state variable with the abscissa being one parameter. As a parameter varies, the dynamics of the system we concerned change through a local or global bifurcation which leads to the change of stability at the same time.

Now, r is considered as a parameter with the range 0.5–3.5 for the system (4). Since the bifurcation diagrams of $r - u_t$ are similar to the bifurcation diagrams of $r - v_t$, we will only show the former which can be seen from Figure 1.

Next, we performed a series of simulations for the reaction-diffusion systems, and, in each, the initial condition was always a small amplitude random perturbation 1% around the steady state. As a numerical example, we consider the bifurcation of the two-dimensional system (47)-(48). It is well known that Turing instability (bifurcation) is diffusion-driven instability; thus the diffusion rate is vital to the pattern formation. To investigate the effect of diffusion coefficients on patterns, by keeping all the other parameters of the system fixed ($a = 0.55$, $r = 0.71$, and $d_1 = 0.22$), we change a diffusion coefficient in the Turing instability

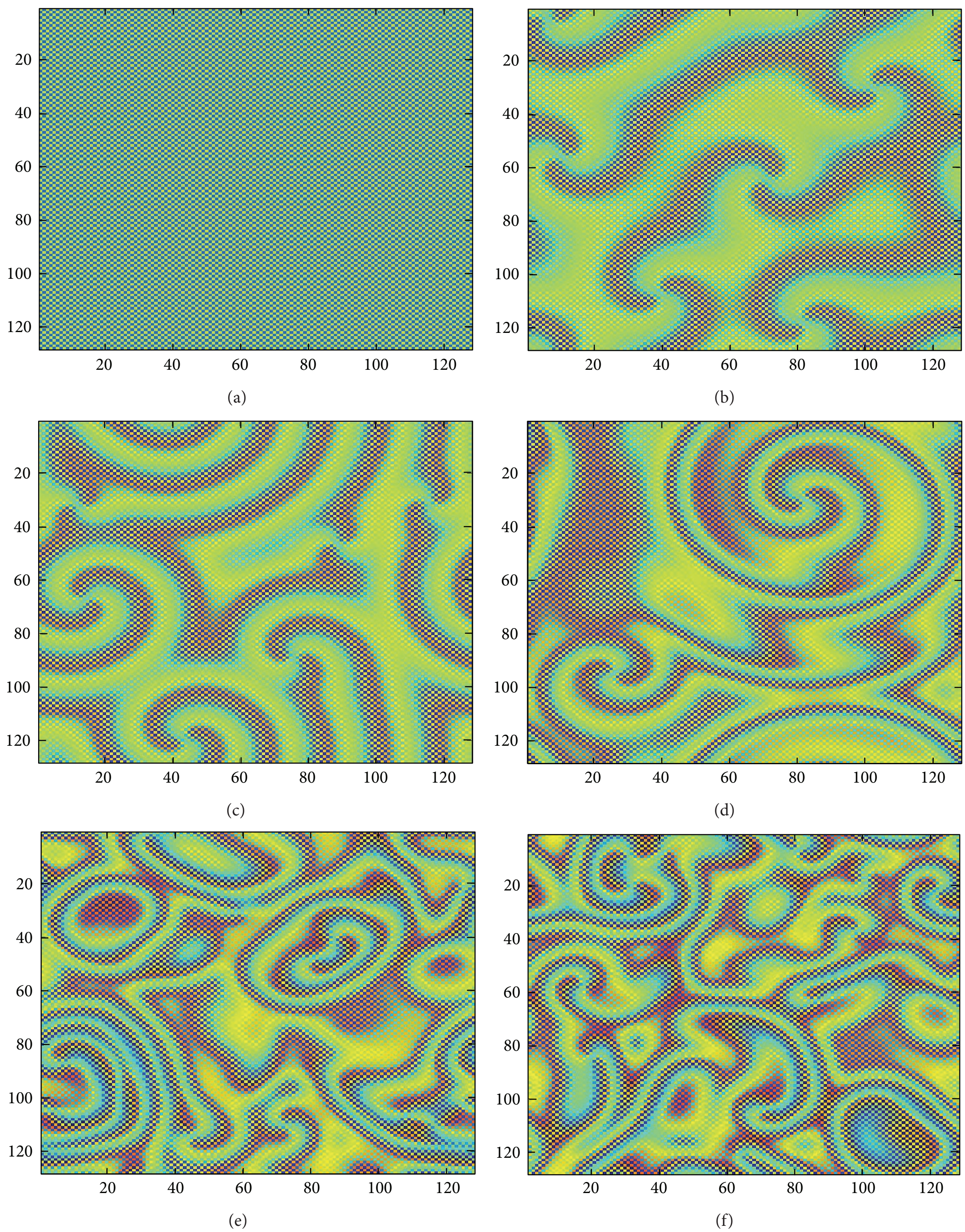


Figure 2: Pattern selection with the increase of d_2 in the Turing instability region when $a = 0.55$, $r = 0.71$, and $d_1 = 0.22$. (a) $d_2 = 0.195$. (b) $d_2 = 0.20$. (c) $d_2 = 0.205$. (d) $d_2 = 0.21$. (e) $d_2 = 0.215$. (f) $d_2 = 0.217$.

region (parameter space which satisfies Turing instability). Figures 2(a)–2(f) exhibit in detail the different distribution of patterns with varying values of d. If we let $d_2 = 0.19$, a stable pattern of square shapes, namely, stationary wave, is observed in Figure 2(a). With the increase of d_2, some tips appear, which attempt to form self-centered spiral waves (Figure 2(b)). Then some tips vanish and some tend to evolve to regular spiral waves, shown in Figure 2(c). Then, with the parameter evolution proceeding, the size of spiral waves rises, but the density of them decreases (see Figure 2(d)). If d_2 is further increased, we observe that, in parts of patterns, spiral waves begin to break up. Hardly can spiral waves be seen, disorder and chaotic structure are depicted in Figures 2(e) and 2(f).

5. Discussion and Conclusion

In this paper, we have applied different discrete schemes to convert the continuous Lotka-Volterra competition model to a new discrete model and studied the dynamical characteristic of the discrete model. Our theoretical analysis and numerical simulations have demonstrated that the discrete competition model undergoes flip bifurcation. Furthermore, when the effects of spatial factors are considered, we discuss the Turing instability conditions combining linearization method and inner product technique. The impact of the diffusion coefficients on patterns can also be observed visually, and some interesting situations can be observed. Indeed, the new discrete model can result in a rich set of patterns and we expect that it is more effective in practice.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors thank Dr. Y. D. Jiao for the valuable suggestions and thank the referees for their helpful comments. This work was financially supported by Tianjin University of Commerce with the Grant no. X0803, the National Natural Science Foundation of China under Grant nos. 11001072, 11371277, and Cultivation Program for Excellent Youth Teacher in University, Tianjin (507-125RCPY0314).

References

[1] Y. Takeuchi, *Global Dynamical Properties of Lotka-Volterra Systems*, World Scientific Publishing, Singapore, 1996.

[2] P. Waltman, *Competition Models in Populationbiology*, vol. 45 of *CBMS-NSF Regional Conference Series in Applied Mathematics*, SIAM, Philadelphia, Pa, USA, 1985.

[3] M. E. Fisher, *Analysis of difference equation models in population dynamics [Ph.D. thesis]*, University of Western Australia, 1982.

[4] P. Liu and X. Cui, "A discrete model of competition," *Mathematics and Computers in Simulation*, vol. 49, no. 1-2, pp. 1–12, 1999.

[5] F. B. Hildebrand, *Finite-Difference Equations and Simulations*, Prentice Hall, Englewood Cliffs, NJ, USA, 1968.

[6] J. K. Hale and A. S. Somolinos, "Competition for fluctuating nutrient," *Journal of Mathematical Biology*, vol. 18, no. 3, pp. 255–280, 1983.

[7] K. P. Hadeler and I. Gerstmann, "The discrete Rosenzweig model," *Mathematical Biosciences*, vol. 98, no. 1, pp. 49–72, 1990.

[8] Y. Kuang, *Delay Differential Equations with Applications in Population Dynamics*, vol. 191, Academic Press, New York, NY, USA, 1993.

[9] S. N. Elaydi, *An Introduction to Difference Equations*, Springer, Berlin, Germany, 1996.

[10] J. Maynard Smith, *Models in Ecology*, Cambridge University Press, 1974.

[11] P. Z. Liu and S. N. Elaydi, "Discrete competitive and cooperative models of Lotka-Volterra Type," *Journal of Computational Analysis and Applications*, vol. 3, pp. 53–73, 2001.

[12] Y. S. Wang and H. Wu, "Dynamics of competitive Lotka-Volterra systems that can be projected to a line," *Computers & Mathematics with Applications*, vol. 47, pp. 1263–1271, 2004.

[13] X. Xiong and Z. Zhang, "Periodic solutions of a discrete two-species competitive model with stage structure," *Mathematical and Computer Modelling*, vol. 48, no. 3-4, pp. 333–343, 2008.

[14] C. Niu and X. Chen, "Almost periodic sequence solutions of a discrete Lotka-Volterra competitive system with feedback control," *Nonlinear Analysis. Real World Applications*, vol. 10, no. 5, pp. 3152–3161, 2009.

[15] J. G. Skellam, "Random dispersal in theoretical populations," *Biometrika*, vol. 38, pp. 196–218, 1951.

[16] A. M. Turing, "The chemical basis of morphogenesis," *Philosophical Transactions of the Royal Society B*, vol. 237, pp. 37–72, 1953.

[17] J. Bascompte and R. V. Sole, "Spatially induced bifurcations in single-species population dynamics," *Journal of Animal Ecology*, vol. 63, pp. 256–264, 1994.

[18] L. A. D. Rodrigues, D. C. Mistro, and S. Petrovskii, "Pattern formation, long-term transients, and the Turing-Hopf bifurcation in a space- and time-discrete predator-prey system," *Bulletin of Mathematical Biology*, vol. 73, no. 8, pp. 1812–1840, 2011.

[19] Y. T. Han, B. Han, L. Zhang, L. Xu, M. F. Li, and G. Zhang, "Turing instability and labyrinthine patterns for a symmetric discrete comptitive Lotka-Volterra system," *WSEAS Transactions on Mathematics*, vol. 10, pp. 181–189, 2011.

[20] G. Q. Sun, Z. Jin, Q. X. Liu, and L. Li, "Dynamical complexity of a spatial predator-prey model with migration," *Ecological Modelling*, vol. 219, pp. 248–255, 2008.

[21] G. Q. Sun, Z. Jin, Q. X. Liu, and L. Li, "Spatial pattern of an epidemic model with cross-diffusion," *Chinese Physics B*, vol. 17, pp. 3936–3941, 2008.

[22] Y.-X. Wang and W.-T. Li, "Effect of cross-diffusion on the stationary problem of a diffusive competition model with a protection zone," *Nonlinear Analysis. Real World Applications*, vol. 14, no. 1, pp. 224–245, 2013.

[23] M. Li, B. Han, L. Xu, and G. Zhang, "Spiral patterns near Turing instability in a discrete reaction diffusion system," *Chaos, Solitons & Fractals*, vol. 49, pp. 1–6, 2013.