

Research Article

Resilience Analysis of Urban Road Networks Based on Adaptive Signal Controls: Day-to-Day Traffic Dynamics with Deep Reinforcement Learning

Wen-Long Shang ^{1,2}, Yanyan Chen,¹ Xingang Li ², and Washington Y. Ochieng³

¹Beijing Key Laboratory of Traffic Engineering, College of Metropolitan Transportation, Beijing University of Technology, Beijing 100124, China

²School of Traffic and Transportation, Beijing Jiaotong University, Beijing SW7 2AZ, China

³Centre for Transport Studies, Imperial College London, London SW7 2AZ, UK

Correspondence should be addressed to Wen-Long Shang; shangwl_imperial@bjut.edu.cn

Received 7 August 2020; Revised 12 October 2020; Accepted 11 November 2020; Published 21 November 2020

Academic Editor: Zhile Yang

Copyright © 2020 Wen-Long Shang 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.

Improving the resilience of urban road networks suffering from various disruptions has been a central focus for urban emergence management. However, to date the effective methods which may mitigate the negative impacts caused by the disruptions, such as road accidents and natural disasters, on urban road networks is highly insufficient. This study proposes a novel adaptive signal control strategy based on a doubly dynamic learning framework, which consists of deep reinforcement learning and day-to-day traffic dynamic learning, to improve the network performance by adjusting red/green time split. In this study, red time split is regarded as extra traffic flow to discourage drivers to use affected roads, so as to reduce congestion and improve the resilience when urban road networks are subject to different levels of disruptions. In addition, we utilize the convolution neural network as Q-network to approximate Q values, link flow distribution and link capacity are regarded as the state space, and actions are denoted as red/green time split. A small network is utilized as a numerical example, and a fixed time signal control and other two adaptive signal controls are employed for the comparisons with the proposed one. The results show that the proposed adaptive signal control based on deep reinforcement learning can achieve better resilience in most of the cases, particularly in the scenarios of moderate and severe disruptions. This study may shed light on the advantages of the proposed adaptive signal control dealing with major emergencies compared to others.

1. Introduction

It is widely accepted that urban road networks (URNs) underpin the prosperity of our society and economy, while URNs are easily exposed to various internal or external disruptions [1, 2]. If appropriate and timely measures cannot be implemented when URNs suffer from these disruptions, a large amount of loss of life and economic loss would be incurred. For example, on 1st August 2007, 140,000 daily vehicle trips were disturbed by the abrupt collapse of the I-35W bridge over the Mississippi River in Minneapolis. According to the Minnesota Department of Transportation [3], the loss was estimated up to \$400,000 per day [4]; on 21st

July 2012, the rainstorm in Beijing caused traffic paralysis within some areas of the city, and this disaster also led to the deaths of 37 people, with 1.9 million people affected approximately, and economic loss estimated in tens of billions RMB [5]. In addition, other types of disruptions, such as terrorist attacks, traffic accidents, and riots, are all able to cause huge damage to URNs. Therefore, given the significant costs caused by disruptions, analysing and improving resilience of URNs are economically and practically important.

Although resilience is very important to urban systems, there is no universal definition concerning resilience of urban road networks due to the diversity of networked

systems. Resilience has been studied in different fields, such as social-ecological systems [6], economics [7], and urban infrastructure [8]. Particularly, in the field of infrastructure, resilience can be described with the following four key characteristics [9, 10]:

- (1) **Robustness:** the inherent strength of or resistance in a system to withstand external demands without degradation or loss of functionality
- (2) **Redundancy:** system properties that allow for alternate options, choices, and substitutions under stress
- (3) **Resourcefulness:** the capacity to mobilize needed resources and services in an emergency
- (4) **Rapidity:** the speed with which disruption can be overcome and safety, services, and functionality stability can be restored

Currently, the above definition for resilience has been widely used in the field of transportation since such definition includes several characteristics of resilience. However, in reality, it is very difficult to combine all characteristics together to assess the resilience within one model framework, for example, Lhomme et al. [11] and Wang et al. [12] utilized the redundancy and rapidity to evaluate the resilience of traffic networks, respectively. In this study, we define resilience of URNs as the ability to recover to a stable state after various disruptions. Two characteristics of resilience are used to assess the resilience of URNs, that is, robustness and rapidity; in other words, our modelling and quantitative index for resilience require to take into account these as key performance indicators (KPIs) of resilience.

It is also expected that resilience of transport networks is measured in many ways. The minority of existing research explores resilience from the perspective of qualitative analysis. For example, Hughes and Healy [13] proposed a qualitative method to assess resilience of transport networks in New Zealand based on principles such as redundancy and adaptation. The method consists of many measures which are scored from 1 (very low) to 4 (very high). Meanwhile, the majority of research concerning resilience of transport networks focuses on quantitative analysis; to be more specific, these quantitative studies are conducted from two perspectives: topology-based model and mathematical programming-based model [14, 15]. In general, topology-based models mainly utilize some indices based on complex network theory to assess resilience, such as average path length [11], betweenness centrality [16], and giant connected component [17]. However, these topology-based measures tend to ignore the realistic characteristics of transport networks such as travel demand, road capacity, and traveller's behaviours, although such measures can be easily understood and efficiently computed. By contrast, research based on mathematical programming models is able to cover these realistic characteristics. Omer et al. [18] proposed a Networked Infrastructure Resiliency Assessment (NIRA) framework to assess resilience of networked infrastructure, and they assessed resilience based on the calculations before and after disruptions. Following this, Bhavathrathan and Patil [19]

proposed an indicator to access the resilience of the road network suffering from recurring capacity disruptions, which is calculated based on the ratio between the minimum possible expected system travel time (ESTT) at the state without disruptions and the ESTT at the critical stage. The critical-state ESTT is obtained from a minimax program. Nogal et al. [20] utilized the normalized area over the exhaustion curve, which is obtained from a dynamic restricted equilibrium model, to investigate the resilience of a traffic network experiencing extreme weather. In addition, a bilevel, three-stage stochastic mathematical program with partial user equilibrium constraints is proposed to investigate the travel time resilience of the network under different disaster scenarios [21]. Wang et al. [12] developed a day-to-day toll scheme to analyse and optimize the resilience of traffic systems after a disruption, and rapidity is used as an indicator of resilience.

The number of measures and indices for the resilience of networked systems is therefore considerable. Amongst research related to resilience analysis, how to utilize intelligent control measures such as adaptive signal control (ASC) to enhance the resilience of URNs suffering from disruptions has been a central focus. As a main tool to manage traffic, adaptive signal control is able to improve the performance of URNs and mitigate congestion and delays by adjusting the red/green time split, namely, the signal plan, according to traffic flow detected in real time.

Webster [22] was one of the first to explore how to model the signal settings and how such settings affect the traffic flow on a single junction. Following this, Robertson [23] developed a model (TRANSYT) which optimizes the whole network of traffic signals. However, these are all based on a very unrealistic fundamental assumption, that is, the chosen signal setting does not affect the route choices of drivers. In order to seek mutual interactions of traffic assignment and signal controls, many studies related to these have emerged. Allsop [24], Gartner [25], Smith [26, 27], and Dickson [28] are regarded as the earliest scholars who explored combination models of route choice and signal controls. Allsop [24] explored the relationship between the signal control and the route choice, and Gartner [25] realized that signal control may influence the demand pattern and thus improve the system-oriented total cost, although the assumption is that the route choice is insensitive to signal control, namely, route choice is fixed in his study. Dickson [28] investigated how signal settings at signalized junctions influence the route flow at the equilibrium state by taking into account the optimization of total travel cost. Meneguzzo [29] combined signal controls with stochastic user equilibrium (SUE) to study a combined traffic assignment and control (CTAC) problem. The purpose is to develop a methodological framework to evaluate the effectiveness of different adaptive signal controls with different levels of user information. Maher et al. [30] proposed a bilevel program including traffic signal optimization on congested networks and stochastic user equilibrium assignment.

Nowadays, two main adaptive signal control strategies exist. The first is an equisaturation policy [22], which is one of the oldest signal controls used to adjust traffic, and the second one is a P0 policy [26, 27, 31], which is less conventional and more recent. These two signal controls are

introduced in Section 2.2 in detail. The rationale for this policy is to reduce the green time split for the approaches with congested traffic and encourage more drivers to use routes with less traffic flow. In this way, the capacity of the network can be efficiently used. The exploration of the combination of traffic assignment and signal controls has been a central focus in the field of traffic control. Adaptive signal control (ASC) may facilitate the mitigation of congestion and traffic delay.

This study proposes a novel adaptive signal control method based on a doubly dynamic learning framework to improve the resilience of URNs when suffering from disruptions, and this learning framework consists of day-to-day traffic dynamic model and deep reinforcement learning. Day-to-day (DTD) dynamics can be used to describe and predict the daily evolution of traffic flows and drivers' learning processes on route costs, and the DTD dynamic model is adopted in this study because such an assignment method is so flexible that a wide range of behaviour rules, levels of aggregation, control measures, and various traffic models can be integrated within the same modelling framework [32, 33]. In this study, the resilience of URNs is observed from the perspective of the evolution of day-to-day traffic and quantified with the RAI index; various signal controls and distinct learning process are also incorporated into the model in order to demonstrate how adaptive signal controls (ASCs) improve the resilience by adjusting red/green time split when suffering from disruptions. For the detailed discussions on day-to-day traffic dynamic models, refer to Smith [34], Friesz et al. [35], Cantarella and Cascetta [36], Watling [37], Zhang and Nagurney [38], and Peeta and Yang [39]. In addition, deep reinforcement learning (DRL) is a relatively new concept and is the combination of deep learning and reinforcement learning. It has been used to solve many complex decision-making problems that were out of the scope for a machine [40]. El-Tantawy et al. [41] summarized the work using reinforcement learning to control traffic signals from 1997 to 2010. They mentioned that reinforcement learning is limited to tabular Q-learning, discrete state space is only used for small-scale systems, and the complex nature of traffic at intersections cannot be described well. In this study, DRL is used to combine with the DTD dynamic model to output optimal red/green time split, so as to improve the network performance of URNs against disruptions. DRL is able to handle complicated tasks with lower prior knowledge, even in high-dimensional space [40]; therefore, it has many potential applications in real world, such as robotics [42], autonomous driving [43], and economics [44]. For the detailed discussions on DRL, refer to Vincent et al. [40] and El-Tantawy et al. [41].

The aim of this study is to propose a novel adaptive signal control (ASC) strategy based on the DTD dynamic model and deep reinforcement learning (DRL) so as to improve the resilience of URNs suffering from different levels of disruptions, which captures the day-to-day learning behaviours of drivers and complex nature of traffic flow evolution and signal setting at intersections. Here, several signal control strategies are combined with the DTD traffic model by transforming the red time split into extra flow to illustrate

the mechanism how ASC guides traffic flow to less affected routes when disruptions occur. The proposed ASC is compared with P0, equisaturation, and fixed time signal control strategies to present its efficiency in improving the resilience of the URN after disruptions.

The remaining of this study is organized as follows. Section 2 introduces the methodology used in this study, which includes the proposed ASC based on a doubly dynamic learning framework, three components of the DTD dynamic model: route perception updating model, route choice model, and network loading model, and ASC based on deep reinforcement learning; other existing ASC methods and a relative area index (RAI) for quantifying resilience are introduced in detail. In Section 3, a numerical case study is presented to show traffic evolutions of the network under different levels of disruptions with different ASC strategies and to examine the effects of these signal controls in improving the resilience after the disruptions. Finally, conclusions are presented in Section 4.

2. Methodology

This section mainly introduces the methodology used in this study. A novel adaptive signal control strategy based on a doubly dynamic learning framework, which consists of deep reinforcement learning and traffic day-to-day dynamic learning, is introduced in detail. In addition, two existing adaptive signal controls and a relative area index (RAI) used for quantifying resilience of URNs are introduced briefly.

2.1. Adaptive Signal Control Strategy Based on a Doubly Dynamic Learning Framework

2.1.1. Traffic Model with Day-to-Day Dynamic Learning. In this study, the traffic model with day-to-day dynamic learning refers to the day-to-day (DTD) traffic dynamic model, which is able to be used to capture daily dynamic evolution of traffic flows via drivers' learning process on route perceptions. One of the advantages of the DTD dynamic model is that it is most appropriate for analysing traffic equilibration processes due to their flexibility in accommodating a wide range of behaviour rules, levels of aggregation, and various traffic models to be integrated within the same modelling framework [32]. The DTD model mainly consists of three components: route perception updating model, route choice model, and network loading model.

Assume there is a general urban road network, $G(N, A)$, where N is a set of nodes and A is a set of links. Here, W is used to denote a set of origin-destination (OD) pairs, and an OD pair is denoted as $(i, j) \in W$. Following this, we use T_{ij} to denote a fixed travel demand. In addition, h_t^r and c_t^r denote the flow and unit travel cost on route $r \in R$ on day t , respectively, and R represents the set of routes for OD pair $(i, j) \in W$.

(1) *Route Perception Updating Model.* For the DTD traffic model, two types of route perception updating model exist [45]; drivers' perceptions on the routes for the first type of model depend on the measured costs of a finite number of

previous days, while the second type updates drivers' perceptions on routes only based on the perceived cost and actual cost of the previous day [37]. Given that the first type uses weight to represent the influence level of previous days' cost on the route cost, the complexity of the model increases. Hence, we employ the second one developed by Walting [46], as shown in the following equation:

$$x_t^r = x_{t-1}^r + \alpha(c_{t-1}^r - x_{t-1}^r), \quad \alpha \in (0, 1), \quad (1)$$

where x_t^r and c_{t-1}^r are perceived and actual route cost on route r on day t and day $t-1$, respectively, and α represents the sensitivity of the route cost on the current day to the route cost on the previous day. α is a constant, and a smaller value for α indicates a stronger habitual tendency of drivers. In this study, this equation reflects drivers' day-to-day learning process.

(2) *Route Choice Model*. In this study, drivers' route choice behaviours are modelled with the logit model [47], and the proportion of drivers choosing route r on day t , P_t^r , is shown as follows:

$$P_t^r = \frac{1}{1 + \sum_{l \in R_{i,j}} e^{\theta(x_l^r - x_t^r)}}, \quad \forall r \in R_{i,j}, \quad (2)$$

where $R_{i,j}$ is the set of routes connecting OD pair $(i, j) \in W$ and θ represents the sensitivity of perceived route cost differences to the proportion of drivers using the route. θ is also termed the dispersion parameter by Prashker and Bekhor [48], which can be related to the quality of information.

Based on equation (2), flow assignment via the route choice for aggregate flow can be shown as follows:

$$h_t = T_{i,j} P^r(x_t^r), \quad \forall r \in R_{i,j}, i, j \in W. \quad (3)$$

According to Daganzo and Sheffi [49], route flow h can follow a stochastic user equilibrium when the model reaches convergence.

(3) *Network Loading Model*. In this study, the link flow u_t^a on link $a \in A$ on day t is able to be derived from the path flow h_t^r and is shown as follows:

$$u_t^a = \sum_{r \in R} \delta_{a,r} \cdot h_t^r, \quad \forall a \in A. \quad (4)$$

$\delta_{a,r}$ is a link-route incident matrix:

$$\delta_{a,r} = \begin{cases} 1 & \text{if } a \text{ belongs to } r, \\ 0 & \text{if } a \text{ does not belong to } r. \end{cases} \quad (5)$$

Afterwards, link cost can be obtained by the following function:

$$C_a = C_a(u_t^a), \quad \forall a \in A, \quad (6)$$

where C_a is the link travel cost function, which is assumed to be continuous and monotonically increasing. Bureau of Public Roads (BPR) [50] is used as the link cost function:

$$C_a(u_t^a) = A_a \left[1 + B \left(\frac{u_t^a}{K_a} \right)^\varphi \right], \quad \forall a \in A, \quad (7)$$

where A_a and K_a are free-flow time and link capacity, respectively, and B and φ are parameters. Following this, path cost is obtained with the following equation:

$$c_r = \sum_{a \in A} C_a \delta_{a,r}, \quad \forall a \in A. \quad (8)$$

The above process is known as the network loading problem. Mathematically, the network loading can be summarized as

$$\begin{aligned} c_t &= (c_t^r)_{r \in R}, \\ u_t &= (u_t^a)_{a \in A}, \\ c_t^r &= \sum_{a \in A} \delta_{a,r} \cdot C_a(u_t), \quad \text{for all } r \in R. \end{aligned} \quad (9)$$

This network loading process is used here to reflect the relationship between route cost and link flow.

(4) *Integrating Signal Control Strategies into the Traffic Model*. In the study, we assume that the signal light located at the intersections has two phases: red and green, and the sum of red time and green time split for a given link i is equal to 1, namely, $r_{l_i} = 1 - g_{l_i}$. Here, red/green time split represents the proportion of red/green time split for the link from the macroperspective.

In order to integrate different ASC strategies into the DTD traffic model, red time split can be regarded as extra traffic flow on roads. In the study, red time split based on different ASC strategies can be added into the BPR function as extra flow, which is mathematically presented as follows:

$$C_a(u_t^a) = A_a \left[1 + B \left(\frac{u_t^a + \beta r_{l_i}^a}{K_a} \right)^\varphi \right], \quad \forall a \in A, \quad (10)$$

where $r_{l_i}^a$ is the red time split on link a on day t and β is a parameter used for conversion from the red time split to link flow.

2.1.2. Adaptive Signal Control (ASC) Based on Deep Reinforcement Learning. This section introduces how deep reinforcement learning (DRL) is employed to self-adaptive signal controls. DRL may give better intelligence to adaptive signal control by more accurately capturing action characteristics. Here, we use deep Q-learning network (DQN) to learn how red/green time split of signal controls impacts traffic flow in the network.

(1) *Deep Q-Learning Model*. In general, the daily state characteristics of a road network are provided as the input of the DQN, and the DQN takes actions based on the highest score among all actions and inputs the actions into the environment to obtain the state on the next day. In order to solve the problem for unstable training, a deep Q-learning model with experience replay and target networks is designed to learn how traffic signals respond to variations of link flow. The deep Q-learning model used in the study is shown in Figure 1.

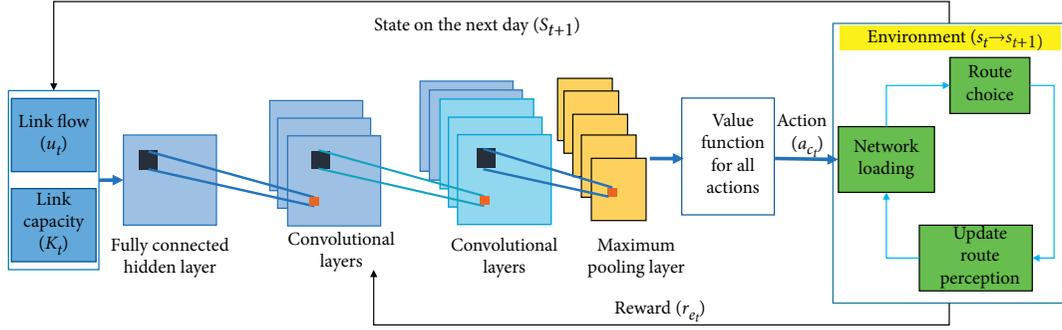


FIGURE 1: The deep Q-learning model used in the study.

(2) *Definitions of Elements for the DQN Applied to Traffic Signal Control.* In the study, elements of the DQN are defined as follows:

Agent (A_g): adaptive traffic signal in the road network

Environment (E): an urban road network (URN)

Policy (p_o): choosing specific red/green time split $((r_{t+1}, g_{t+1}) \in (0, 1))$ for the next day when agents are at state s

State (s): the capacity and traffic flow of each link of the URN are defined as the state, which can be represented as a tensor set $s_t = [K_t, u_t]$, where K_t and u_t are the capacity and flow of links, respectively

Action (a_c): red/green time split on the next day

Reward (r_e): relative area index (RAI) used for quantifying the resilience of URNs and the number of time steps (days) that the road network recovers to a new equilibria state

In the model, the adaptive traffic signal lights (agents) are used to output the red time split, which is calculated based on the link capacity and flow in the road network (environment). The DQN utilizes Q value to replace the rewards in the Markov decision process (MDP). According to Bellman's equations, Q value can be calculated by the following equation:

$$Q(s_t, a_{c_t}) = R(s_t, a_{c_t}) + \gamma \max_{a_{c_{t+1}}} Q(s_{t+1}, a_{c_{t+1}}), \quad \gamma \in (0, 1), \quad (11)$$

where γ is a discount factor, which represents the tradeoff between future and current rewards. The DQN chooses the red time split for each link at intersections for the next time step based on Q -scores. In the model, a convolutional neural network (CNN) is used to approximate the Q -function, which is presented as follows:

$$Q(s, a_c; w) \approx Q(s, a_c), \quad (12)$$

where w is the parameter of the CNN. w starts with random initialization. At the beginning of each training step, the CNN collects the state information as the input of the neural network; afterwards, it uses the forward process to generate Q value at the output layer of the CNN.

(3) *Red Time Split.* After the CNN generates the Q value, the traffic signal lights (agents) take current action based on a certain strategy. This study utilizes ϵ -greedy policy to take actions. Firstly, ϵ is assigned to a smaller value; then, a number $\delta \in (0, 1)$ is randomly generated. If δ satisfies equation (13), adaptive traffic signal lights will select the red time split with the highest Q value; otherwise, they randomly select a red time split.

$$\delta < 1 - \epsilon. \quad (13)$$

Following this, the chosen red time split is input into the DTD traffic dynamic model to obtain the state on the next day, and then agents obtain the state. Estimated return π_i can be calculated with the following equation:

$$\pi_i = R_{e_i} + \gamma \max_{a_c} q(s_i, a_c^*; w). \quad (14)$$

π_i consists of two parts: the first part is the accurate reward value, which is derived from the DTD dynamic model; the second part is the estimated maximum value based on the CNN. Here, a^* is the action to maximize the current Q value.

(4) *Updating Parameters.* The estimation of the Q value for the red time split requires to achieve better accuracy by updating the parameters of the CNN. At each time step, based on the interactions of actions, we update the parameters of the CNN by reducing the value of the following equation:

$$\frac{1}{G} \sum_{i \in G} [u_i - Q(a_c, s; w)]^2, \quad (15)$$

where G is the batch number. The CNN conducts gradient descent every time with multiple data batches. Here, updating parameters is to make estimated return close to the Q value predicted by the CNN.

(5) *Experience Replay and Target Network.* In order to increase the stability and expedite the convergence speed of the algorithm, the deep Q-network (DQN) algorithm with experience replay and the target network is utilized.

Experience replay is a technique that stabilizes the probability distribution of the experience, which may

improve the stability of training. Experience replay mainly consists of two key steps: storage and sampling replay.

Storage: store the track in the form of $\langle S_t, A_{e_t}, R_{e_{t+1}}, S_{t+1} \rangle$

Sampling replay: use random sampling to take one or more pieces of experiences from the storage

The target network is a network with exactly the same structure with the original neural network (ONN) but out of the ONN. The ONN is named as the evaluation network (EN). In the process of updating weight, only the weight of the EN is updated, but the weight of the target network is not updated. Since, during a period of time, when the target network does not change, the estimated return is relatively fixed; hence, the target network increases the stability for learning.

2.1.3. The Simulation of the ASC Strategy Based on a Doubly Dynamic Learning Framework. The components described in the previous two sections are formed into a complete doubly dynamic learning framework. This novel ASC with the doubly learning framework is based on the DQN, and the training process is presented in Table 1.

As can be seen from Table 1, we start with random initialization of the parameters w of the agent evaluation network, and the target network and its parameters are set up. In this study, the training of the DQN is performed by Keras, which is an open-source library to provide python interface for artificial neural networks. In the study, the epoch is set to be 5000 times, which means the times that the model runs to equilibrium. At each epoch loop, we first initialize the perceived route costs with free-flow time and assign the average link flow as the initial state. If the DTD model is not converged, we conduct the procedures from Step 3 to Step 10. To begin with, red time split is generated based on the Q value in equation (12) and the ϵ -greedy policy in (13). Following this, DTD dynamic learning starts running, network loading model in (9) is used to gain actual route costs, route perception updating model (1) is used to update the perceived route costs on the next day, and route flow is determined by the route choice model shown in (2) and (3). Afterwards, the new state and the actual route costs on the next day are achieved with equation (4) and the network loading model, respectively. The reward for the network takes -1 if the DTD model is not converged. Steps 6 to 9 are introduced in detail in Section 2.1.2. Based on the experience dataset, return can be calculated with (14), and parameters are updated according to (15). If the model is converged, one epoch is completed, reward takes 10000 minus RAI value, and the parameters of the target network are updated. At the last step, if the epoch completes the number of times set, the parameters are stored, and the training of the DQN ends; otherwise, return to Step 2.

When the DQN finishes training, such a learning framework including the DTD dynamic model and deep reinforcement learning can be used to enhance the resilience of the URN when suffering from disruptions. The detailed procedures are presented in Table 2.

As can be observed from Table 2, in the study, we assume that the disruptions take place on a given link when the URN achieves equilibrium, which implies the stable state of the network system. Then, based on equation (1), perceived route travel cost x_{t+1} is obtained, and route flow can be derived based on equations (2) and (3). Following this, use the trained DQN to generate the red time split corresponding to the link flow and road capacity; according to equation (10), the red time split is incorporated into the DTD model, and then based on network loading model in (9), actual route costs on the next day can be gained. If the model achieves equilibrium, the simulation ends; otherwise, return to Step 2. Table 2 exhibits how the proposed ASC strategy based on the DTD dynamic model with deep reinforcement learning adjusts the red time split to enhance the resilience of the URN suffering from disruptions.

2.2. Two Existing Adaptive Signal Controls. To date, there are two main existing adaptive signal control (ASC) strategies. The first is an equisaturation policy [22] which is one of the oldest signal controls used to adjust traffic, and the second one is a P0 policy [26, 27, 31], which is less conventional and more recent. The core of both signal controls is to calculate the red/green time based on the traffic delays in the assignment models.

2.2.1. Equisaturation Signal Policy. Equisaturation policy is regarded as one of the most widely adopted conventional signal-setting methods used in traffic engineering to handle combined traffic assignment and control (CTAC) problems [51]. Webster [22] originally proposed the equisaturation signal control policy, which stipulates

$$\text{link red - time cost} = \frac{x_i}{g_i s_{a_i}} = \frac{x_i}{(1 - r_i) s_{a_i}} = \frac{x_i}{s_{a_i} - r_i s_{a_i}}, \quad (16)$$

where s_{a_i} is the saturation flow on link i , g_i and r_i are the green time split and red time split on link i , respectively (they are dimensionless), and x_i is the flow on link i .

Red time split based on equisaturation can be written as follows:

$$r_i = 1 - z_i \tau, \quad i \in (1, \dots, n), \quad r_i \in [0, 1], \quad (17)$$

where $z_i = (x_i/s_{a_i})$ and $\tau = (1/(z_1 + z_2 + \dots + z_n))$; n denotes the number of incoming links for a general road junction controlled by signals. For the detailed mathematical derivation process, refer to Shang [1].

Then, the link flow is updated by adding extra flow in terms of red time split [52]:

$$x_i \leftarrow x_i + \vartheta r_i x_i, \quad i \in (1, \dots, n), \quad (18)$$

where ϑ is a constant multiplier, and in this study, ϑ takes 1.

2.2.2. P0 Signal Policy. Compared to the equisaturation policy, the P0 policy, introduced by Smith [27, 31], is more recent and specifically designed for use in CTAC modelling

TABLE 1: The training of the DQN.

Step 1	Initialize the parameters w of the agent evaluation network (EN) $Q(s, a_c; w)$; assign w to the parameters w^* of the target network $Q^*(s, a_c; w^*)$; set the training times $epoch$ and the batch number G ; set $epoch = 0$
Step 2	Set $epoch = epoch + 1$. Assign the free-flow costs of all routes as the initial perceived route costs of drivers; initialize the state s_t of the DTD dynamic model, and set $t = 0$
Step 3	<i>While</i> (DTD model does not reach convergence)
Step 4	Based on equations (12) and (13), generate actions A_c
Step 5	According to equations (7) and (9), obtain actual route costs c_t on day t based on A_{c_t} and s_t (link flow); according to route perception updating process (1), update the perceived route travel cost x_{t+1} on day $t + 1$ based on perceived route cost x_t and actual route costs c_t on day t ; following this, determine the route flows h_{t+1} on day $t + 1$ based on route choice probability formula (2) and route assignment equation (3); according to network loading model (9) and equation (4), the link flows u_{t+1} on day $t + 1$ (new state s_{t+1}) and the actual route costs c_{t+1} on day $t + 1$ are achieved; since the DTD model is not converged, reward $r_{e_{t+1}} = -1$
Step 6	Store the experience $\langle S_t, A_c, R_{e_{t+1}}, S_{t+1} \rangle$ into the experience database D
Step 7	Select a batch of experience $\langle s_i, A_c, r_{e_{t+1}}, s_{t+1} \rangle, i \in G$, from D
Step 8	Based on equation (14), calculate the return
Step 9	According to (15), update the parameters
Step 10	Update the state, actual route cost, perceived route cost, and route flow: $s_{t+1} \rightarrow s_t, c_t \rightarrow c_{t+1}, h_t \rightarrow h_{t+1}$, and $x_t \rightarrow x_{t+1}$
Step 11	When the equilibrium is reached, reward r_e takes 10000-RAI, and the RAI is derived from (23); update the parameters of the target network, $w \rightarrow w^*$. One $epoch$ ends
Step 12	If $epoch$ does not reach the set times, return to Step 2; otherwise, store the parameters w^* , and the training of the DQN ends

TABLE 2: Procedures for the proposed doubly dynamic learning framework.

Step 1	On day t , start with the equilibrium state of the URN under DTD dynamics, and add different levels of capacity reduction on links
Step 2	Based on the perceived route cost x_t and actual route cost c_t on day t , update the perceived route travel cost x_{t+1} on day $t + 1$ by performing DTD dynamic learning process (1)
Step 3	Based on route choice probability formula (2) and route assignment equation (3), determine the flow on all routes h_{t+1} on day $t + 1$
Step 4	Obtain link flow u_{t+1} on day $t + 1$ by using formula (4); based on the current state (link flow and capacity), utilize the trained DQN to output the action (red time split); following this, integrate the action into the DTD model by using equation (10), and then achieve the actual route cost c_{t+1} on day $t + 1$ by performing network loading (9)
Step 5	If convergence condition (22) is satisfied, stop; otherwise, return to Step 2

[51]. The P0 policy also assumes that red times may cause additional link delays that may be captured through some extra flow units on the relevant links. Let the augmented link flow be $x_i + s_i r_i$; the link red-time cost can be shown as follows:

$$\text{link red - time cost} = s_{a_i} b_i (x_i + s_{a_i} r_i), \quad (19)$$

where b_i is the nondecreasing link delay function (such as the BPR function).

According to the P0 policy, the red time split can be deduced in the following way:

$$r_i = \frac{K_i \sqrt{\left((\mu - s_{a_i} A_i) / (s_{a_i} A_i B_i) \right)} - x_i}{s_{a_i}}, \quad (20)$$

$$\forall i = 1, 2, 3, \dots, n, r_i \in [0, 1],$$

where K_i is the capacity of link i and A_i, B_i , and φ are positive parameters of the BPR link performance function. Here, μ can be obtained from the following algebraic equation:

$$\sum_{i=1}^n \frac{K_i \sqrt{\left((\mu - s_{a_i} A_i) / (s_{a_i} A_i B_i) \right)} - x_i}{s_{a_i}} = n - 1. \quad (21)$$

Here, for detailed mathematical derivation process, refer to Shang's work [1].

In addition to these two existing adaptive signal controls, we also utilize a fixed time signal control for comparisons with other ASC strategies. Apparently, the fixed time signal control implies that the red and green time split are equal within one signal phase.

2.3. Relative Area Index (RAI). In this study, resilience is measured based on two KPIs: robustness and rapidity of recovery, as described in Section 1. Here, the rapidity of recovery relates to the speed at which the network reaches a new equilibrium that is not necessarily the same as the previous one if the disruption is not removed and/or the network capacity is not restored. Therefore, rapidity of recovery can be quantified as the time between the day of the disruption and the time when the new equilibrium state is

reached. In our study, we use the following equation to determine whether the equilibrium is reached:

$$\frac{\|h_t - h_{t-1}\|}{\|h_{t-1}\|} \leq \rho, \quad (22)$$

where h_t is the set of route flow on day t and ρ is an extreme small value and takes 0.001 here. In the study, all experiments follow this way to determine the equilibrium.

The system evolution is illustrated in Figure 2, where a hypothetical disruption occurs on day t_1 . Before the disruption, the network traffic is at equilibrium with a constant total travel cost across different days. Immediately after the disruption, the network-wide travel cost is likely to increase followed by a slight decrease before traffic reaches a new equilibrium on day t_3 , as can be seen from Figure 2. In the simple case of Figure 2, the rapidity can be presented as the duration of the period between t_1 and t_3 . Meanwhile, robustness can relate to the resistance in a system to maintain its functionality after disturbance occurs. In this study, the network functionality is equated to the system-wide total travel cost, and robustness can be shown as the total travel cost on each day.

When the traffic network suffers from disruptions, the total cost of the network will evolve accordingly. As indicated in Figure 2, right after the disruption, the cost increases but is then likely to reduce as a result of adaptive routing of travellers and improved system efficiency thereafter.

Based on the above descriptions, resilience can be presented by the shadow area in Figure 2, which captures both features of robustness and rapidity. In order to quantify these two characteristics of resilience when URNs are subject to different levels of disruptions, a relative area index (RAI) is utilized here, which was first proposed by Shang et al. [53], as shown in the following:

$$\text{RAI} = \int_{t_1}^{t_3} w_e(t) \cdot \frac{[\text{TC}(t) - \text{TC}(t_1)]}{\text{TC}(t_1)} dt, \quad (23)$$

where $w_e(t)$ is the weight representing the effects of the disruption on each day. In some studies [53], the weight ranges from 10 to 1 based on the consideration of the cascading effects caused by consecutive capacity reductions. This study mainly focuses on the local capacity degradation occurring on a certain day, which is assumed to be consistent during the period of capacity degradation. Here, therefore, all weights w_e take 1.

In this study, RAI represents the running cost during recovery. This index assesses the cumulative loss of efficiency as a result of the disruption and is used in this study to measure the network's resilience. A larger RAI implies a less resilient URN.

3. Numerical Study

In this section, a small network is considered for the numerical study (as shown in Figure 3). The network consists of 9 nodes and 12 links, and there are 1000 travellers on the network. Only one origin-destination (OD) pair (from node a to node i) is taken into account,

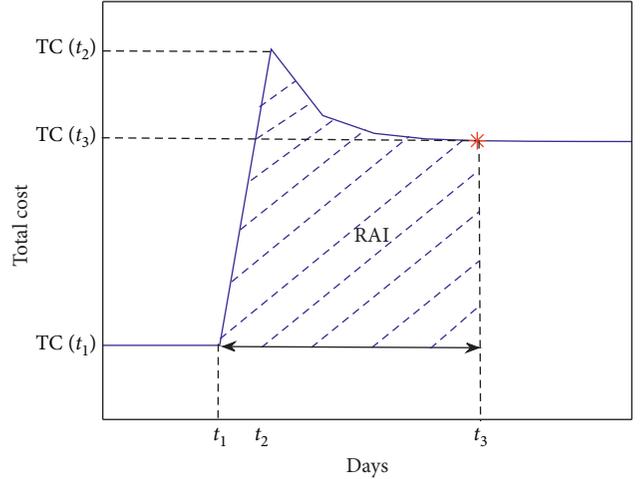


FIGURE 2: Evolution of the total travel cost before and after the disruption.

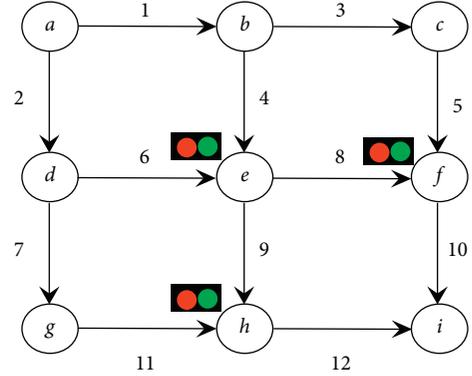


FIGURE 3: Signal control used in the intersection of the network.

TABLE 3: Link parameters of the small network.

Link ID	A	B	φ	K_a
1	25	0.15	4	1000
2	25	0.15	4	1000
3	25	0.15	4	1000
4	25	0.15	4	1000
5	25	0.15	4	1000
6	25	0.15	4	1000
7	25	0.15	4	1000
8	25	0.15	4	1000
9	25	0.15	4	1000
10	25	0.15	4	1000
11	25	0.15	4	1000
12	25	0.15	4	1000

and six routes exist in the network: $a \rightarrow b \rightarrow c \rightarrow f \rightarrow i$ (route 1), $a \rightarrow d \rightarrow e \rightarrow f \rightarrow i$ (route 2), $a \rightarrow b \rightarrow e \rightarrow f \rightarrow i$ (route 3), $a \rightarrow d \rightarrow g \rightarrow h \rightarrow i$ (route 4), $a \rightarrow d \rightarrow e \rightarrow h \rightarrow i$ (route 5), and $a \rightarrow b \rightarrow e \rightarrow h \rightarrow i$ (route 6). The other modelling parameters used for the BPR functions are summarized in Table 3.

Here, link 9 is assumed to be subject to different levels of disruptions, and we use 25%, 50%, and 75% capacity degradation to represent the mild, moderate, and severe disruptions. In this numerical study, the network is assumed to experience two stages: predisruption and postdisruption. Before the disruption, the network system reaches at equilibrium, which represents the normal state of the network. Once disruptions occur, four types of traffic signal control strategies are implemented: the proposed ASC, equisaturation, P0, and fixed time. After the disruption, the network will reach a new equilibrium with signal controls. The resulting network performance, as well as the resilience, is analysed in detail.

As can be seen from Figure 3, the network is controlled by traffic lights. Traffic signal control strategies are employed to react to the daily variation of network traffic flows. This is done by adjusting the green/red split for different approaches at relevant junctions. The rationale for this mechanism is that the time splits can respond to the delays caused by the disruptions, thereby reducing the flow fluctuation as well as the network-wide delay. Given that a signal control is needed only when there are conflicting approaches at a junction, in the case of the small network, signals are only considered at nodes e , f , and h (see Figure 3). Regarding the adaptive signal control (ASC) strategies, the equisaturation control policy, the P0 control policy, and the proposed ASC based on DRL are employed, and we may recap these from Section 2.

3.1. Traffic Evolution with Different Signal Control Strategies.

This numerical study mainly utilizes ASC as a main tool to induce drivers choose alternative routes, so as to mitigate the congestion caused by different levels of disruptions, namely, to improve the resilience of the network against disruptions. In the numerical example, the drivers' perceptions on different routes are initially set to be equal to their free-flow time. We use the methodology presented in Section 2 to carry out the simulation. In the study, we assume that different levels of disruptions take place when the initial equilibrium is obtained; thereafter, the ASC adjusts the red/green time split to respond to such unexpected disruptions. Here, we need to emphasize that the difference of the days for adding disruptions does not affect the results related to the resilience, and the shadow area under the curve between t_1 (the day when the initial equilibrium is obtained) and t_3 (the day when the new equilibrium is obtained), as shown in Figure 2, is our main focus.

Based on the simulations, the resulting route costs, route flows, and network-wide total cost over time are presented in Figures 4–7, and each one corresponds to a signal control (e.g., the proposed ASC, equisaturation, P0, and fixed time).

3.1.1. Equisaturation Control Strategy. Figure 4 presents how the route costs, route flows, and network-wide total costs evolve when equisaturation is used to adjust red/green split so as to control the traffic after different levels of disruptions. In this case, drivers adjust their route

perceptions and choices based on actual link costs where red time split is added as extra flow, and the red time split of the equisaturation signal control is derived from equation (17). As can be seen from Figure 4, the network starts with an arbitrary configuration of route flows and reaches an equilibrium state. The red vertical line represents the time when the disruptions occur and initial equilibrium is broken, and the black vertical line denotes the time of the attainment of a new equilibrium. With equisaturation signal control, the route costs and route flow do not react significantly to the minor disruption (25% capacity reduction), and the network takes approximately 40 days to converge to a new equilibrium. For the moderate and severe disruptions, the fluctuations of both route costs and route flow are more significant than those under mild disruption, and the network also takes longer days to reach a new equilibrium. The third column shows that the more severe the disruption, the more significant the increase of total cost, and the total cost at the new equilibrium is further deviated from the original equilibrium as the disruption is more severe.

3.1.2. P0 Control Strategy. As can be seen from Figure 5, the route costs, route flows, and network-wide total cost evolve over time when the P0 signal control is utilized to mitigate the congestion and delay caused by different levels of disruptions. In this case, red time split added into the link cost as extra delay is derived from equation (20). As we can see, when the network is subject to 25% capacity reduction (mild), the route costs, route flows, and total cost are less affected, and the network takes shorter days (19 days) to reach a new equilibrium under the influence of the P0 signal control, which is shorter than that under the equisaturation signal control. For moderate and severe disruptions, the route costs significantly increase and then reach stability, and the fluctuations on route flows are more significant. Compared to the case where the equisaturation signal control is employed, the network takes shorter days to reach equilibrium in these scenarios. Through the observation and analysis from Figures 4 and 5, it is concluded that the P0 signal control is superior to equisaturation in enhancing the resilience of the network experiencing all levels of disruptions.

3.1.3. DRL Control Strategy. When the proposed ASC is employed, in order to more clearly exhibit the evolution of route costs, route flow, and total cost under disruptions, we mainly present the curves after the equilibrium is broken. As can be seen from Figure 6, for 50% and 75% capacity reduction, route costs, route flows, and total cost are less fluctuated when the disruptions occur compared to the previous traditional ASC, and it also seems that the network takes shorter days to attain new equilibrium than the equisaturation signal control and similar days compared to the P0 signal control. For minor disruption, we can see that route costs, route flows, and total cost are less fluctuated than those under moderate and severe disruptions, and the speed of converging to new equilibrium is similar with the case

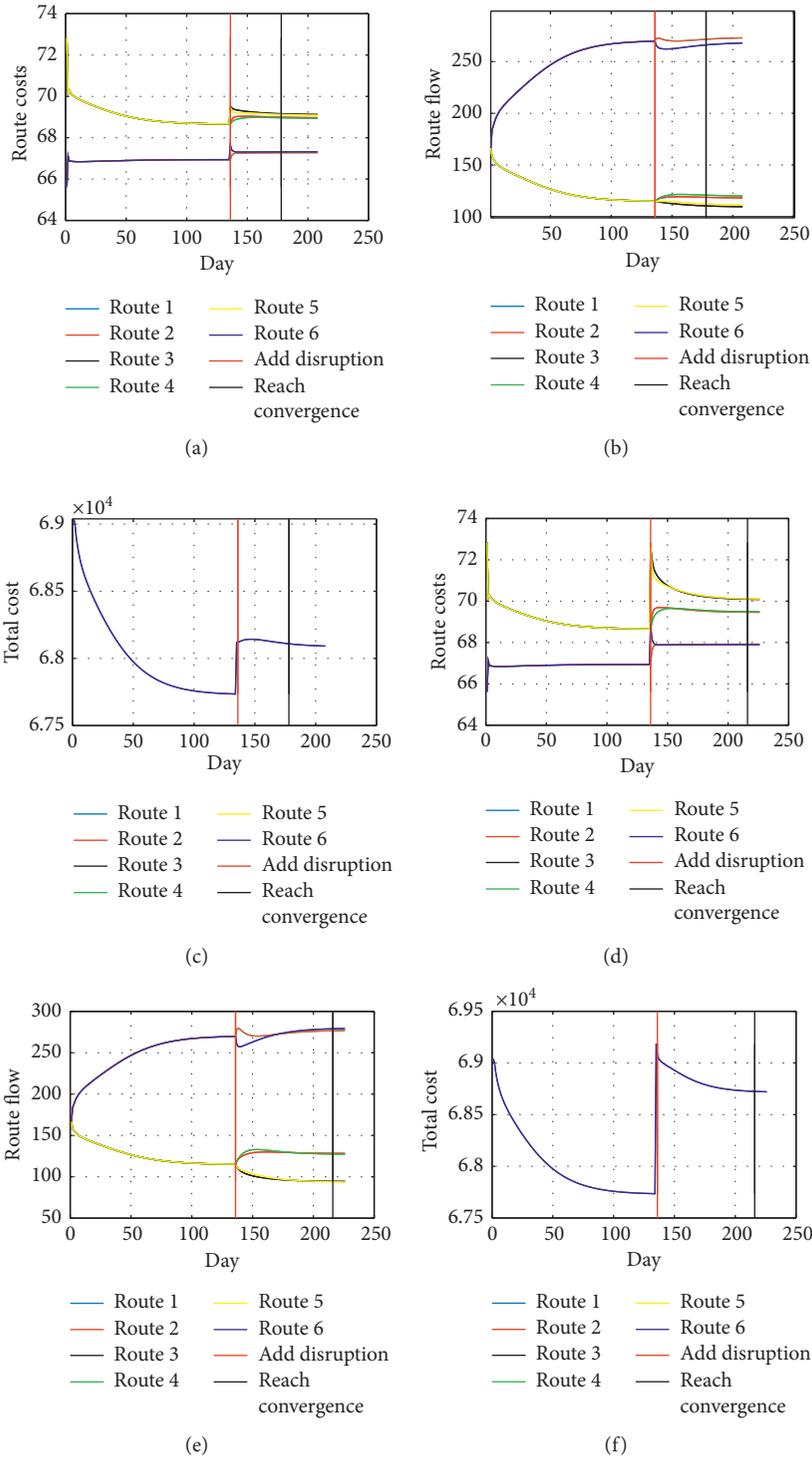


FIGURE 4: Continued.

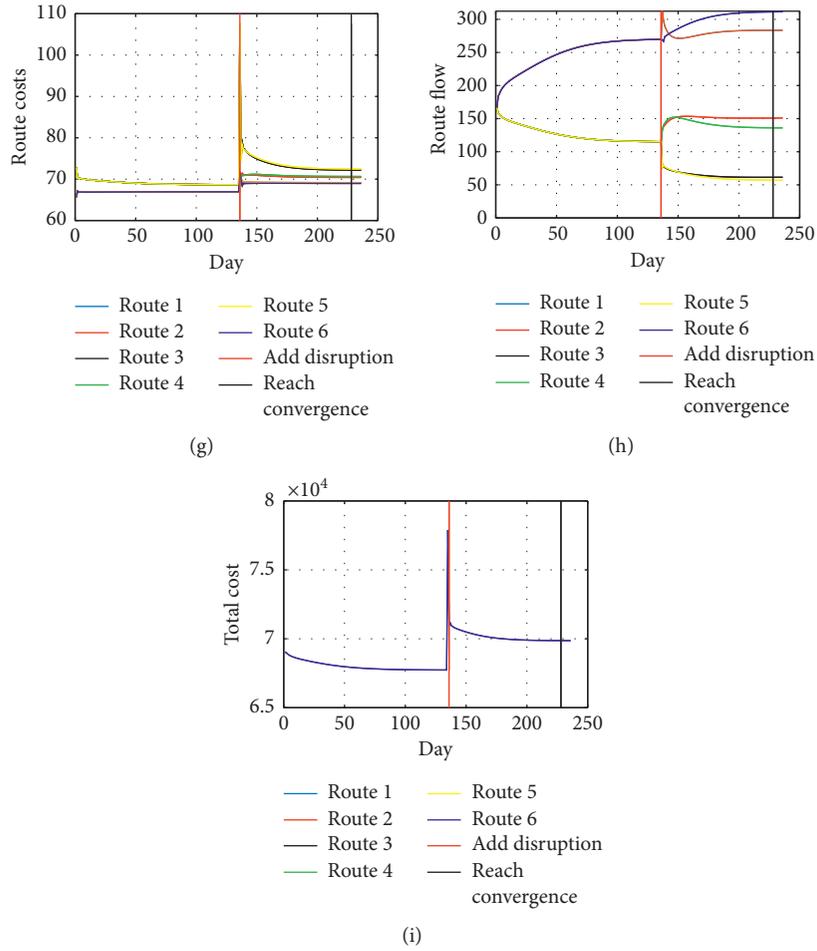


FIGURE 4: Evolution of route costs, route flow, and total cost under different levels of disruptions when the equisaturation ASC is employed. First row (from (a) to (c)): the evolution of route costs, route flow, and total cost when the capacity of link 9 reduces by 25%; second row (from (d) to (f)): the evolution of route costs, route flow, and total cost when the capacity of link 9 reduces by 50%; third row (from (g) to (i)): the evolution of route costs, route flow, and total cost when the capacity of link 9 reduces by 75%.

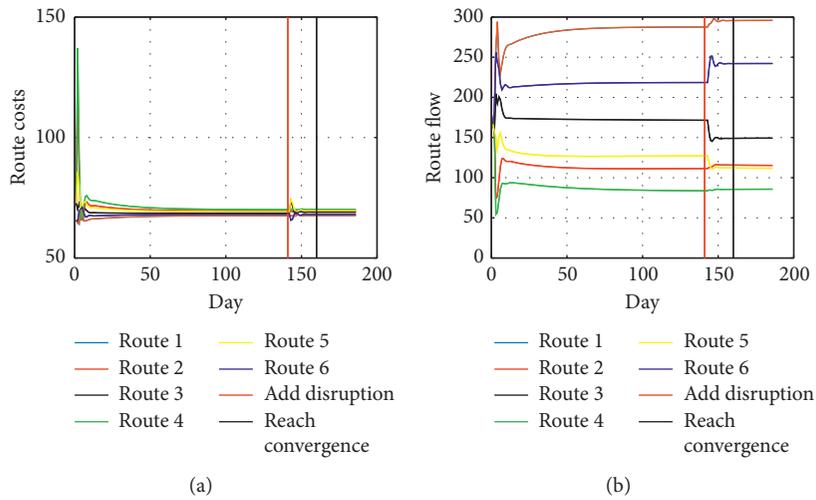


FIGURE 5: Continued.

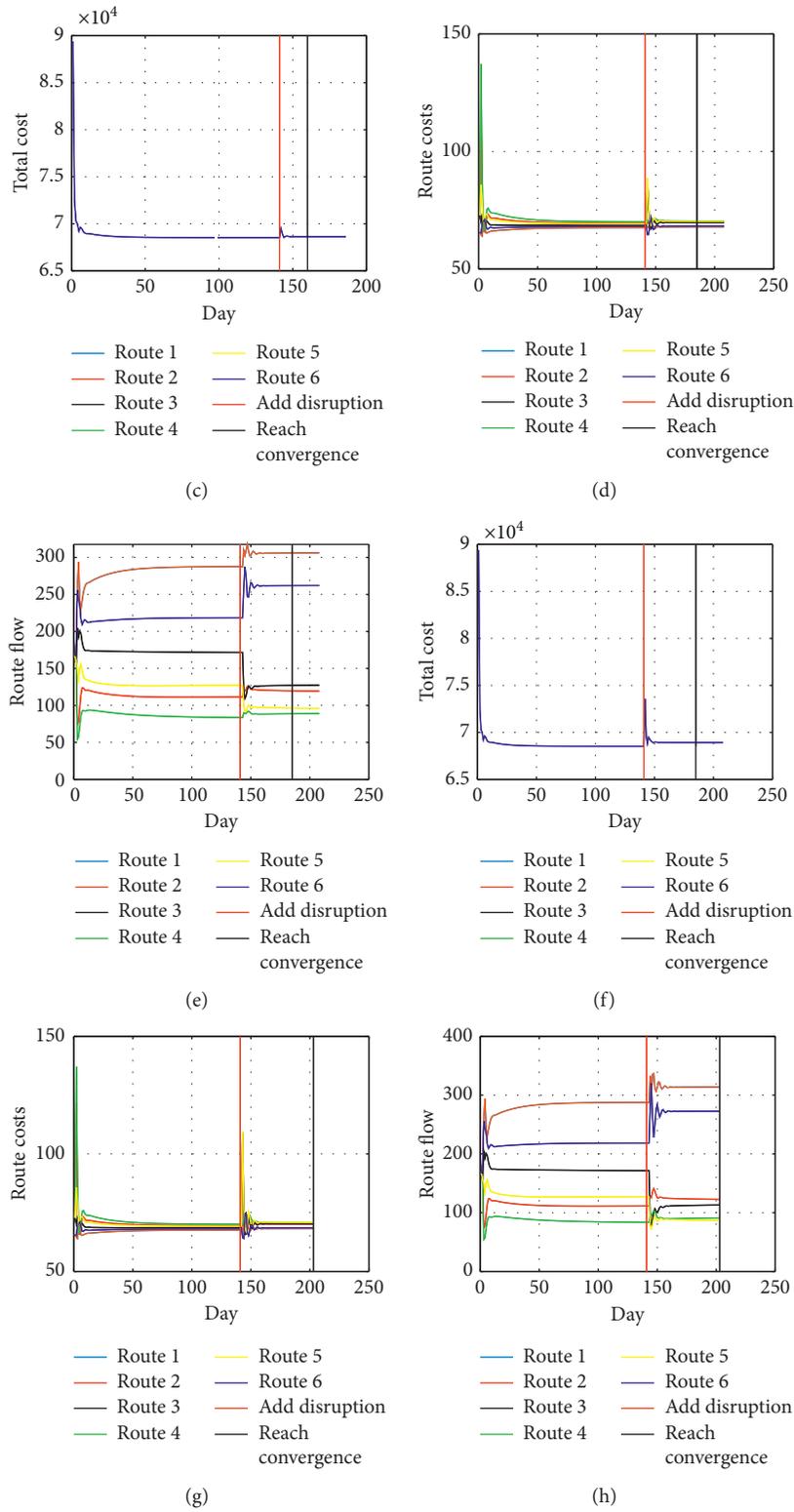
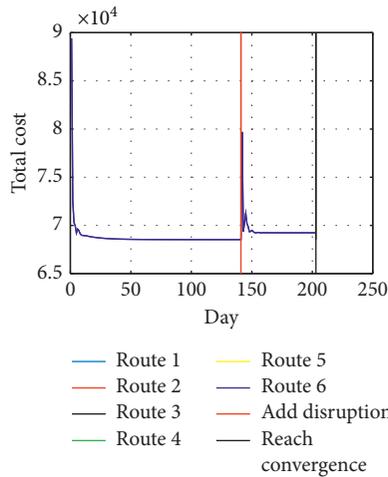


FIGURE 5: Continued.



(i)

FIGURE 5: Evolution of route costs, route flow, and total cost evolution under different levels of disruptions when the P0 ASC is employed. First row (from (a) to (c)): the evolution of route costs, route flow, and total cost when the capacity of link 9 reduces by 25%; second row (from (d) to (f)): the evolution of route costs, route flow, and total cost when the capacity of link 9 reduces by 50%; third row (from (g) to (i)): the evolution of route costs, route flow, and total cost when the capacity of link 9 reduces by 75%.

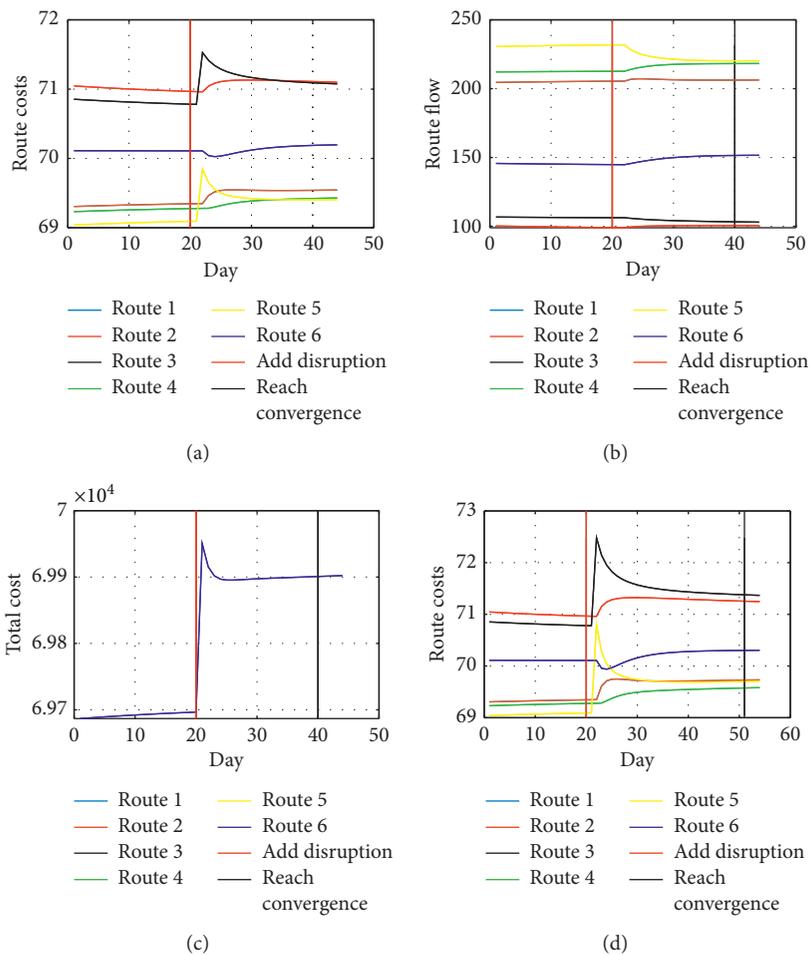


FIGURE 6: Continued.

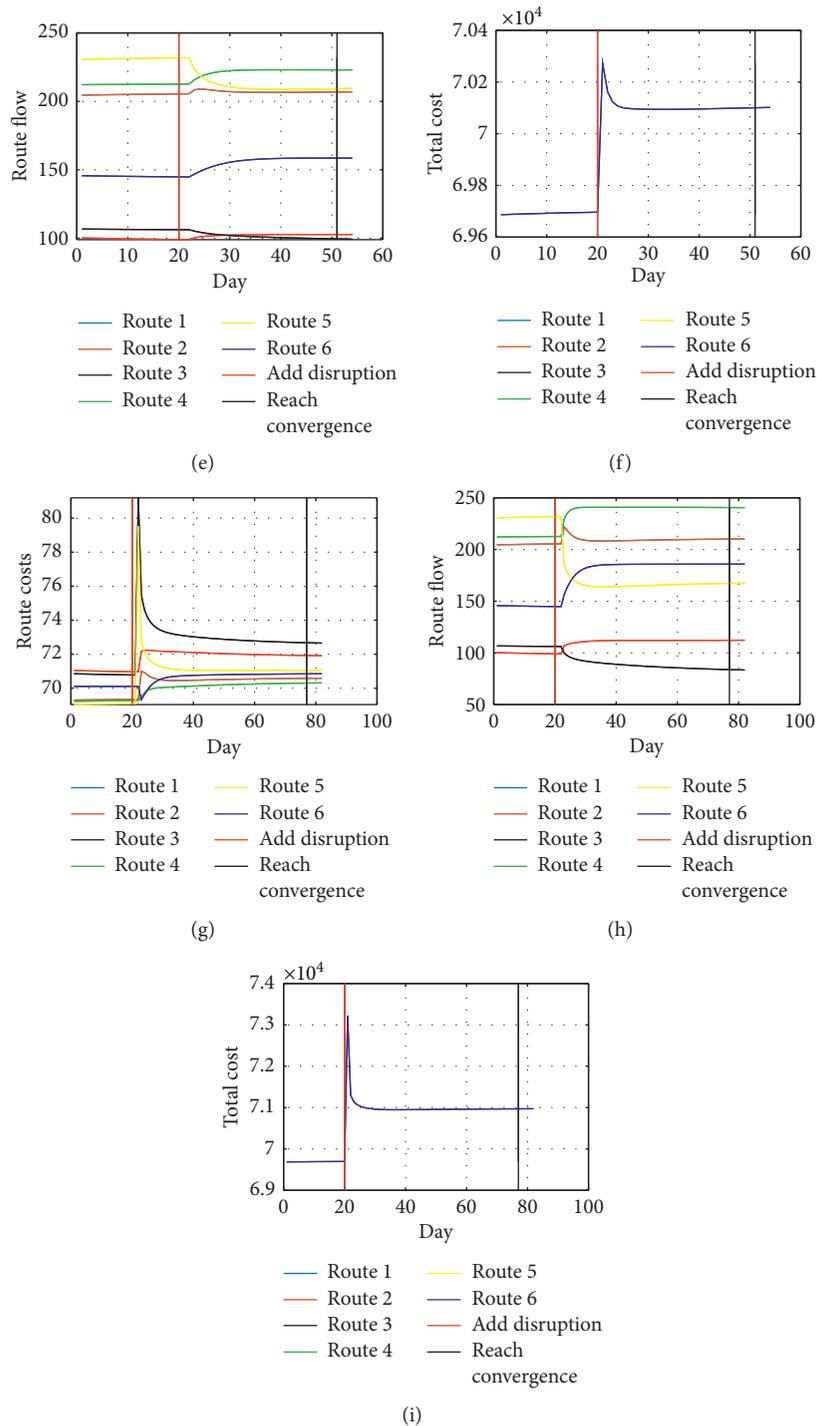


FIGURE 6: Evolution of route costs, route flow, and total cost evolution under different levels of disruptions when the ASC based on DRL is employed. First row (from (a) to (c)): the evolution of route costs, route flow, and total cost when the capacity of link 9 reduces by 25%; second row (from (d) to (f)): the evolution of route costs, route flow, and total cost when the capacity of link 9 reduces by 50%; third row (from (g) to (i)): the evolution of route costs, route flow, and total cost when the capacity of link 9 reduces by 75%.

using the P0 signal control. In addition, we can see from Figure 6 that the increase in the total cost is more significant when more severe disruption takes places, and the total cost at the new approximation equilibrium is further deviated from the original approximation equilibrium.

3.1.4. Fixed Time Control Strategy. In order to compare these adaptive signal controls with the traditional signal strategy, a fixed time signal control is utilized. As can be seen from Figure 7, for all levels of disruptions, the utilization of the fixed time signal control apparently gives rise to significant

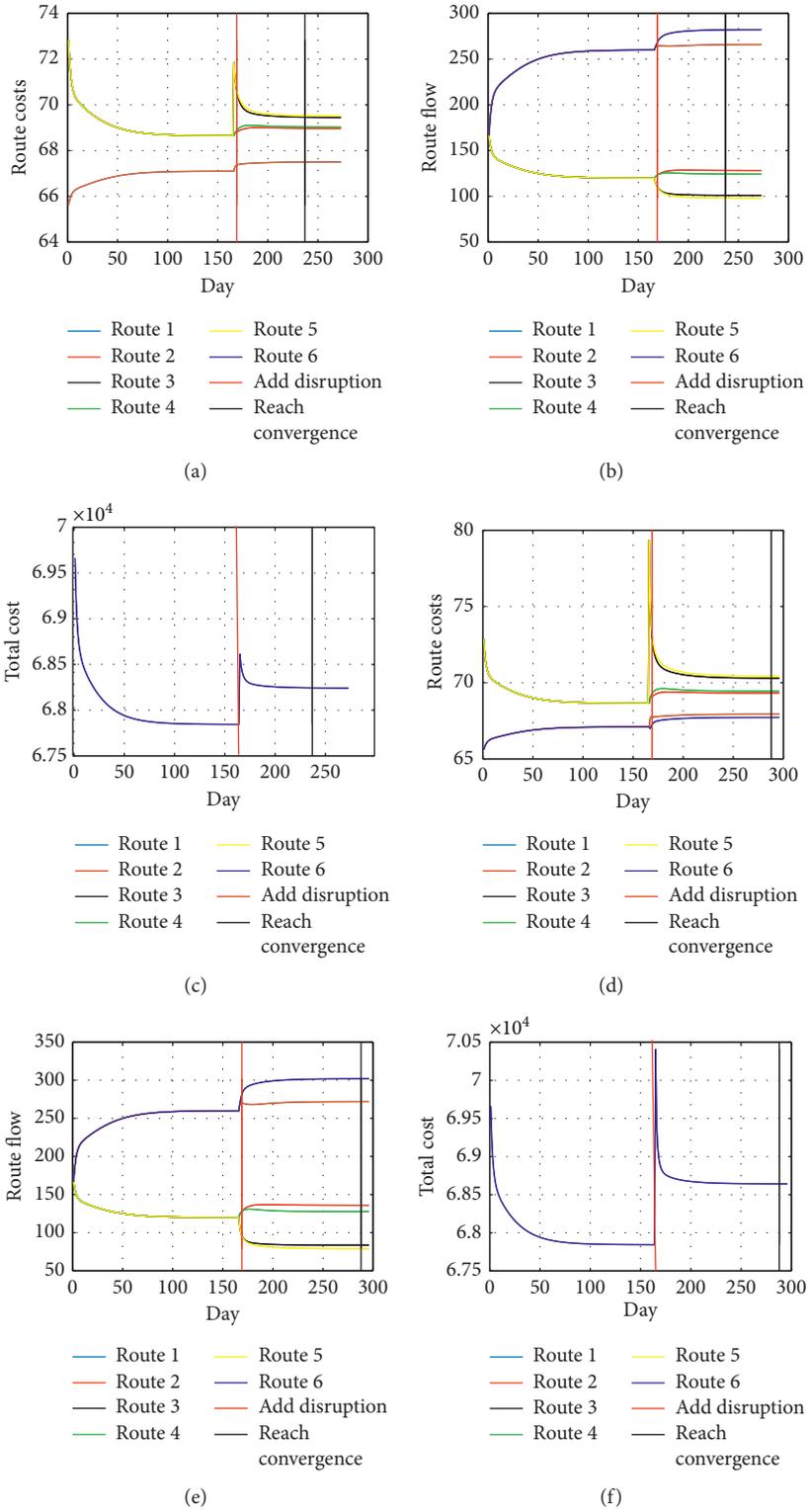


FIGURE 7: Continued.

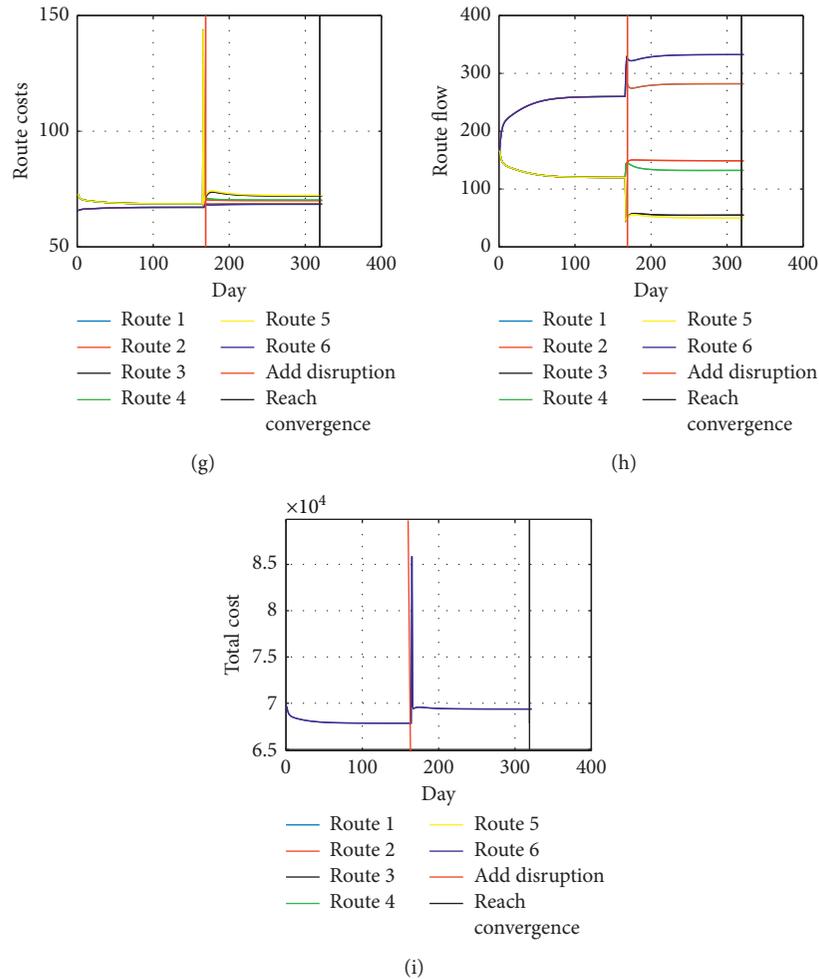


FIGURE 7: Evolution of route costs, route flow, and total cost evolution under different levels of disruptions when the fixed time signal control is employed. First row (from (a) to (c)): the evolution of route costs, route flow, and total cost when the capacity of link 9 reduces by 25%; second row (from (d) to (f)): the evolution of route costs, route flow, and total cost when the capacity of link 9 reduces by 50%; third row (from (g) to (i)): the evolution of route costs, route flow, and total cost when the capacity of link 9 reduces by 75%.

fluctuations of route costs, route flows, and total cost and also causes longer days to reach new equilibrium compared to equisaturation, P0, and DRL. The explanation for this is that the fixed time signal control cannot dynamically adjust red/green time split based on the flow distribution in the network, while P0, DRL, and equisaturation can automatically adjust signal timings based on certain learning rules.

Through the observations from Figures 4–7, we can see that the route costs, route flows, and network-wide total costs evolve over time as different types of signal controls are employed, respectively. For moderate and severe disruptions, ASC based on DRL shows the most efficient learning mechanism on improving resilience of the road network among four types of signal controls, although the case where P0 is employed for mild disruption presents better results. The quantitative results for all signal controls are presented in Table 4.

3.2. Resilience Analysis with Different Signal Control Strategies. In this study, rapidity of recovery and robustness can be regarded as key performance indicators (KPIs) of

TABLE 4: RAIs under different levels of disruptions with different signal controls.

Capacity reduction	Signal control			
	Equisaturation	P0	DRL	Fixed time
25%	0.0136	0.0037	0.0143	0.0731
50%	0.1481	0.0794	0.0536	0.8847
75%	1.8156	0.4087	0.2691	4.6031

resilience. In order to quantify these two KPIs with one comprehensive index, we utilize RAI to access resilience of the network against disruptions, as introduced in Section 2.3.

Adaptive signal control is widely used to improve the performance of URNs and thus to mitigate congestion and delays by adjusting the red/green time split based on information on route costs [54]. In the previous section, the brief discussions regarding how route costs, route flow, and total cost of the network evolve under different ASC strategies when the network is subject to different levels of

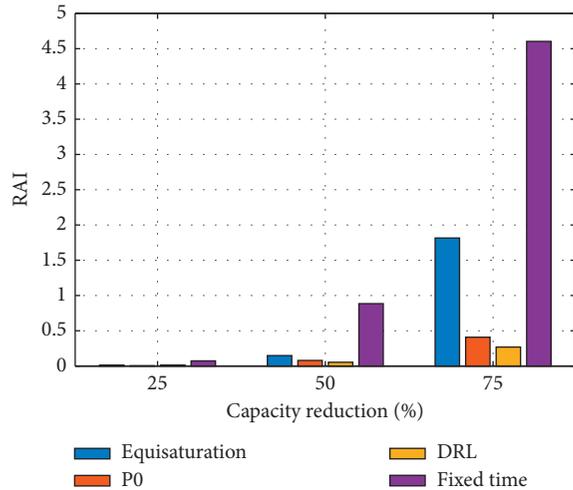


FIGURE 8: Comparisons of resilience for the small network under different levels of disruptions with different signal controls.

disruptions are presented, and a quantitative evaluation and analysis of resilience under different ASCs is conducted here based on the RAI. The results are summarized in Table 4. In addition, Figure 8 is also presented to visually observe the resilience of the network under different ASC strategies.

As can be seen from Table 4, in the case of 25% capacity reduction caused by a mild disruption, P0 signal control has the minimum RAI, which means that the P0 signal strategy is the most efficient in adjusting traffic flow so as to achieve the best resilience when the network suffers from the mild disruption, while the fixed time signal control achieves the worst resilience, and the RAI values for the fixed time signal control are much larger than those for ASC strategies. Apparently, since the fixed time signal control lacks dynamical mechanisms to adjust red/green time split based on variations of link flow, it shows the worst ability to improve the performance of URNs when suffering from different levels of disruptions.

In the scenario of the network suffering from moderate (50%) and severe (75%) disruptions, we can see that the ASC based on DRL expedites the network to achieve fastest recovery to normal (equilibrium) state after disruptions. It seems that the proposed novel doubly dynamic learning framework is very efficient in managing the traffic when more serious disruptions occur. The comparisons of different signal controls on improving resilience can also be visually presented in Figure 8.

To summarize, compared to the fixed time signal control, ASC strategies always perform better in improving the resilience of the road network when suffering from different levels of disruptions. This result is expected since ASC strategies tend to adjust red/green time split dynamically based on the traffic flow on the roads rather than equally assigning red/green time split in a static way. In addition, the proposed ASC outperforms other traditional ones in scenarios of moderate and severe disruptions, although a little worse than the P0 signal control when experiencing mild disruption. This suggests that the doubly dynamic learning framework consisting of DTD dynamic learning and deep reinforcement learning is able to efficiently adjust red/green time split globally in response to the disruptions which may

cause greater destructions of URNs, and our results shed light on the advantages of the proposed adaptive signal control dealing with unexpected major emergencies compared to others.

4. Conclusions and Future Work

Given increasing number of natural disasters and emergencies, resilience of URNs has received increasing attention. In order to improve the resilience of URNs when experiencing different levels of disruptions, this study proposes a novel adaptive signal control (ASC) strategy based on a doubly dynamic learning framework, which combines the DTD traffic dynamic model with deep reinforcement learning (DRL). This novel signal control takes into account the drivers' day-to-day learning process on route perceptions and ASC's learning mechanism on the flow distributions. In the study, red time split is regarded as extra flow, which can be incorporated into the network loading process. Through this way, the signal control strategies can be incorporated into the DTD dynamic model. We also utilize two existing adaptive signal controls: P0 and equisaturation and fixed time signal controls to compare with the proposed ASC. In the study, a small URN is used as a numerical example. The results show that three adaptive signal controls perform much better than the fixed time signal control in terms of improving the resilience when the network is subject to mild, moderate, and severe disruptions, and particularly, the proposed ASC suggests an apparent advantage of the combination of DTD dynamic learning and DRL in improving resilience of the network suffering from moderate and severe disruptions, which may provide valuable insights on the traffic management of URNs in response to major emergencies.

In the future, this research may be extended from several ways. Firstly, this study is limited by the capabilities of computational architecture and algorithms. In particular, this work utilizes a small network as a numerical study, which is appropriate for training and computing based on the proposed doubly dynamic learning framework. In the future, the use of high-performance computing including parallel/distributed computing and GPU could be considered as means to expedite the training and computational procedures. More computationally efficient models and algorithms can be considered as well, such as link-based traffic assignment models, as opposed to the path-based models that require path enumeration and that do not scale well when the network size increases. In addition, the DTD traffic dynamic model used in this study only considers the route flow and cost evolution from a macrotemporal granularity (days), which means that it does not explicitly incorporate the microtemporal dimension in network flow propagation. One of the consequences is the lack of accountability for real-time information provision, which plays a critical role in network and congestion management under external stress. In the future, one important extension of this work is the dynamic modelling of traffic networks, which considers the within-day fluctuation of network conditions such as traffic flow, congestion, and controls. In

addition to these, the network's performance before, during, and after the disruptions is considered in this study, but the recovery phase of network capacity is completely ignored. However, in reality, once the disruptions occur, the restoration of network capacity becomes an immediate concern, which may involve many topics including network stability, resource allocation, and infrastructure management. Therefore, in the future, exploring the resilience of URNs during the recovery phase with adaptive signal controls will be an interesting area.

Data Availability

All data generated or analysed during this research are included within this article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the Key Special Project of Beijing City (Grant no. Z181100003918011).

References

- [1] W. L. Shang, *Robustness and resilience analysis of Urban Road Networks*, PhD Thesis, Imperial College London, London, England, 2017.
- [2] E. E. Koks, R. Julie, Z. Conrad, and T. Mersedeh, "A global multi-hazard risk analysis of road and railway infrastructure assets," *Nature Communications*, pp. 1–11, 2019.
- [3] Minnesota Department of Transportation, *Road-user Cost Due to Unavailability of Interstate 35W Mississippi River Crossing at Minneapolis, Minnesota*, Minnesota Department of Transportation, Minneapolis, MN, USA, 2007, <http://www.dot.state.mn.us/i35wbridge/rebuild/pdfs/I-35WMississippiRiverCrossingRoad-UserCost.pdf>.
- [4] S. Zhu and D. Levinson, "Disruptions to transportation networks: a review," in *Network Reliability in Practice, Transportation Research, Economics and Policy*, et al. Springer, Berlin, Germany, 2012, http://link.springer.com/chapter/10.1007%2F978-1-4614-0947-2_2 Accessed.
- [5] China News, <http://www.chinanews.com/gn/2012/07-22/4050026.shtml>, 2012.
- [6] S. Carpenter, B. Walker, J. M. Anderies, and N. Abel, "From metaphor to measurement: resilience of what to what?" *Ecosystems*, vol. 4, no. 8, pp. 765–781, 2001.
- [7] E. Hoffman, *Building a Resilient Business*, Raptor Networks Technology Inc, Santa Ana, CA, USA, 2007.
- [8] N. O. Attoh-Okine, A. T. Cooper, and S. A. Mensah, "Formulation of resilience index of urban infrastructure using belief functions," *IEEE Systems Journal*, vol. 3, no. 2, pp. 147–153, 2009.
- [9] M. Bruneau and A. Reinhorn, "Exploring the concept of seismic resilience for acute care facilities," *Earthquake Spectra*, vol. 23, no. 1, pp. 41–62, 2007.
- [10] T. D. O'Rourke, "Critical infrastructure, interdependencies, and resilience," *Bridge, Washington-National Academy Of Engineering*, vol. 37, no. 1, p. 22, 2007.
- [11] S. Lhomme, D. Serre, Y. Diab, and R. Laganier, "Analyzing resilience of urban networks: a preliminary step towards more flood resilient cities," *Natural Hazards and Earth System Sciences*, vol. 13, no. 2, pp. 221–230, 2013.
- [12] Y. Wang, H. Liu, K. Han, T. L. Friesz, and T. Yao, "Day-to-day congestion pricing and network resilience," *Transportmetrica A: Transport Science*, vol. 11, no. 9, pp. 873–895, 2015.
- [13] J. Hughes and K. Healy, "Measuring the resilience of transport infrastructure," *NZ Transport Agency Research Report*, vol. 546, pp. 1–82, 2014.
- [14] W.-L. Shang, Y. Chen, and W. Y. Ochieng, "Resilience analysis of transport networks by combining variable message signs with agent-based day-to-day dynamic learning," *IEEE ACCESS*, vol. 8, pp. 104458–104468, 2020.
- [15] W.-L. Shang, C. Yanyan, S. Chengcheng, and Y. O. Washington, "Robustness analysis of urban road networks from topological and operational perspectives," *Mathematical Problems in Engineering*, vol. 2020, Article ID 5875803, 12 pages, 2020.
- [16] L. Schintler, S. Gorman, R. Kulkarni, and R. Stough, "Moving from protection to resiliency: a path to securing critical infrastructure," in *Critical Infrastructure Reliability and Vulnerability*, A. T. Murray and T. Grubestic, Eds., pp. 291–307, Springer, Berlin, Germany, 2007.
- [17] N. Y. Aydin, H. S. Duzgun, F. Wenzel, and H. R. Heinimann, "Integration of stress testing with graph theory to assess the resilience of urban road networks under seismic hazards," *Natural Hazards*, vol. 91, no. 1, pp. 37–68, 2018.
- [18] M. Omer, A. Mostashari, and R. Nilchiani, "Assessing resilience in a regional road-based transportation network," *International Journal of Industrial and Systems Engineering*, vol. 13, no. 4, pp. 389–408, 2013.
- [19] B. K. Bhavathrathan and G. R. Patil, "Quantifying resilience using a unique critical cost on road networks subject to recurring capacity disruptions," *Transportmetrica A: Transport Science*, vol. 11, no. 9, pp. 836–855, 2015.
- [20] M. Nogal, B. Martinez-Pastor, A. O'Connor, and C. Brian, "Dynamic restricted equilibrium model to determine statistically the resilience of a traffic network to extreme weather events," in *Proceedings of the 12th International Conference on Application of Statistics and Probability in Civil Engineering*, Vancouver, Canada, July 2015.
- [21] R. Faturechi and E. Miller-Hooks, "Travel time resilience of roadway networks under disaster," *Transportation Research Part B: Methodological*, vol. 70, pp. 47–64, 2014.
- [22] F. V. Webster, *Traffic Signal Settings*, Department of Transport, London, UK, Road Research Technical Paper N0.39, 1958.
- [23] Robertson, "TRANSYT method for area traffic control," *Traffic Engineering and Control*, vol. 11, pp. 276–281, 1969.
- [24] R. J. E. Allsop, "Some possibilities for using traffic control to influence trip distribution and route choice," in *Proceedings of the 6th International Symposium on Transportation and Traffic Theory*, pp. 345–374, Sydney, Australia, April 1974.
- [25] N. Gartner, "Area traffic control and network equilibrium," *Lecture Notes in Economics and Mathematical Systems*, vol. 118, pp. 274–297, 1976.
- [26] M. J. Smith, "The existence, uniqueness and stability of traffic equilibria," *Transportation Research Part B: Methodological*, vol. 13, no. 4, pp. 295–304, 1979.
- [27] M. J. Smith, "Traffic control and route-choice; a simple example," *Transportation Research Part B: Methodological*, vol. 13, no. 4, pp. 289–294, 1979.

- [28] T. J. Dickson, "A note on traffic assignment and signal timings in a signal-controlled road network," *Transport Research Part B: Methodological*, vol. 15, no. 4, pp. 264–271, 1981.
- [29] C. Meneguzzer, *Stochastic User Equilibrium Assignment with Traffic-responsive Signal Control*, European Regional Science Association, Belgium, Europe, 1998, <http://www-sre.wu-wien.ac.at/ersa/ersaconfs/ersa98/papers/337.pdf>.
- [30] M. J. Maher, X. Zhang, and D. V. Vliet, "A bi-level programming approach for trip matrix estimation and traffic control problems with stochastic user equilibrium link flows," *Transportation Research Part B: Methodological*, vol. 35, no. 1, pp. 23–40, 2001.
- [31] M. J. Smith, "A local traffic control policy which automatically maximises the overall travel capacity of an urban road network," in *Proceedings of the International Symposium on Traffic Control Systems*, pp. 11–32, Berkeley, CA, USA, December 1979.
- [32] D. Watling and M. L. Hazelton, "The dynamics and equilibria of day-to-day assignment models," *Networks and Spatial Economics*, vol. 3, no. 3, pp. 349–370, 2003.
- [33] W. Shang, K. Han, W. Y. Ochieng, and P. Angeloudis, "An agent-based day-to-day traffic evolution model with information percolation," *Transportmetrica A: Transport Science*, 2017.
- [34] M. J. Smith, "The stability of a dynamic model of traffic assignment—an application of a method of lyapunov," *Transportation Science*, vol. 18, no. 3, pp. 245–252, 1984.
- [35] T. L. Friesz, D. Bernstein, N. J. Mehta, R. L. Tobin, and S. Ganjalizadeh, "Day-to-day dynamic network disequilibria and idealized traveler information systems," *Operations Research*, vol. 42, no. 6, pp. 1120–1136, 1994.
- [36] G. E. Cantarella and E. Cascetta, "Dynamic processes and equilibrium in transportation networks: towards a unifying theory," *Transportation Science*, vol. 29, no. 4, pp. 305–329, 1995.
- [37] D. Watling, "Stability of the stochastic equilibrium assignment problem: a dynamical systems approach," *Transportation Research Part B: Methodological*, vol. 33, no. 4, pp. 281–312, 1999.
- [38] D. Zhang and A. Nagurney, "On the local and global stability of a travel route choice adjustment process," *Transportation Research Part B: Methodological*, vol. 30, no. 4, pp. 245–262, 1996.
- [39] S. Peeta and T.-H. Yang, "Stability issues for dynamic traffic assignment," *Automatica*, vol. 39, no. 1, pp. 21–34, 2003.
- [40] F.-L. Vincent, H. Peter, I. Riashat, G. B. Marc, and P. Joelle, "An introduction to deep reinforcement learning," *Foundations and Trends in Machine Learning*, vol. 11, pp. 219–354, 2018.
- [41] S. El-Tantawy, B. Abdulhai, and H. Abdelgawad, "Design of reinforcement learning parameters for seamless application of adaptive traffic signal control," *Journal of Intelligent Transportation Systems*, 2014.
- [42] L. Pinto, A. Marcin, W. Peter, Z. Wojciech, and A. Pieter, "Asymmetric actor critic for image-based robot learning," arXiv preprint arXiv:1710.06542, 2017.
- [43] X. Pan, Y. You, Z. Wang, and C. Lu, "Virtual to real reinforcement learning for autonomous driving," in *Proceedings of the British Machine Vision Conference (BMVC)*, T. K. Kim, S. Zafeiriou, G. Brostow, and K. Mikolajczyk, Eds., pp. 11.1–11.13, BMVA Press, London, UK, 2017.
- [44] Y. Deng, F. Bao, Y. Kong, Z. Ren, and Q. Dai, "Deep direct reinforcement learning for financial signal representation and trading," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 3, pp. 653–664, 2017.
- [45] E. Cascetta and G. E. Cantarella, "Modelling dynamics in transportation networks: state of the art and future developments," *Simulation Practice and Theory*, vol. 1, no. 2, pp. 65–91, 1993.
- [46] D. Watling, "Stability of the stochastic equilibrium assignment problem- a dynamical systems approach," *Transportation Research Part B*, vol. 33, pp. 281–312, 1999.
- [47] J. Bie and H. K. Lo, "Stability and attraction domains of traffic equilibria in a day-to-day dynamical system formulation," *Transportation Research Part B: Methodological*, vol. 44, no. 1, pp. 90–107, 2010.
- [48] J. N. Prashker and S. Bekhor, "Some observations on stochastic user equilibrium and system optimum of traffic assignment," *Transportation Research Part B: Methodological*, vol. 34, no. 4, pp. 277–291, 2000.
- [49] C. F. Daganzo and Y. Sheffi, "On stochastic models of traffic assignment," *Transportation Science*, vol. 11, no. 3, pp. 253–274, 1977.
- [50] Bureau of Public Roads, *Traffic Assignment Manual*, U.S. Department of Commerce, Urban Planning Division, Washington, DC, USA, 1964.
- [51] C. Meneguzzer, "Dynamic process models of combined traffic assignment and control with different signal updating strategies," *Journal of Advanced Transportation*, vol. 46, no. 4, pp. 351–365, 2012.
- [52] R. Liu and M. Smith, "Route choice and traffic signal control: a study of the stability and instability of a new dynamical model of route choice and traffic signal control," *Transportation Research Part B: Methodological*, vol. 77, pp. 123–145, 2015.
- [53] W. Shang, K. C. Pien, K. Pan, and W. Y. Ochieng, "Robustness and topology analysis of european air traffic network using complex network theory," in *Proceedings of the 94th Transportation Research Board annual meeting*, Washington, DC, USA, 2014.
- [54] H. Liu, K. Han, V. V. Gayah, T. L. Friesz, and T. Yao, "Data-driven linear decision rule approach for distributionally robust optimization of on-line signal control," *Transportation Research Part C: Emerging Technologies*, vol. 59, pp. 260–277, 2015.