

Research Article

Modelling of Freight Trains Classification Using Queueing System Subject to Breakdowns

Michal Dorda and Dušan Teichmann

Faculty of Mechanical Engineering, VŠB-Technical University of Ostrava, Institute of Transport, 17 listopadu 15, 708 33 Ostrava, Czech Republic

Correspondence should be addressed to Michal Dorda; michal.dorda@vsb.cz

Received 4 July 2013; Accepted 2 September 2013

Academic Editor: Massimo Scalia

Copyright © 2013 M. Dorda and D. Teichmann. 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 presents a mathematical model and a simulation model of the freight trains classification process. We model the process as a queueing system with a server which is represented by a hump at a marshalling yard. We distinguish two types of shunting over the hump; primary shunting represents the classification of inbound freight trains over the hump (it is the primary function of marshalling yards), and secondary shunting is, for example, represented by the classification of trains of wagons entering the yard via industrial sidings. Inbound freight trains are considered to be customers in the system, and all needs of secondary shunting are failures of the hump because performing secondary shunting occupies the hump, and therefore inbound freight trains cannot be sorted. All random variables of the model are considered to be exponentially distributed with the exception of customer service times which are Erlang distributed. The mathematical model was created using method of stages and can be solved numerically employing a suitable software tool. The simulation model was created using coloured Petri nets. Both models are tested in conditions of a marshalling yard.

1. Introduction

Marshalling yards represent important nodes on each railway net. They have to fulfil a lot of functions but one of the most important tasks is sorting (or classification) inbound freight trains and making up outbound freight trains. Besides, these tasks (marshalling yards) have also other functions, but they are without direct connection with the process, we would like to model in the paper.

In order to fulfil required tasks, marshalling yards are usually equipped with specialized sets of tracks. From the view of the process which will be modelled in the paper, the following sets of tracks are important. Marshalling yards are equipped with reception sidings, consisting of several arrival tracks on which trains, entering the marshalling yard, are being prepared for their classification. The consequential classification process is being carried out by means of the treatment of the gravitation over a specialized facility called a hump; trains of wagons are being humped into sorting

sidings consisting of so called classification tracks. The mutual arrangement of reception sidings, the hump, and sorting sidings can differ in individual yards. One of the possible arrangements of them is in series; see Figure 1. We will consider such type of arrangement in the paper.

Between the moment when an inbound freight train journey is ended on a selected arrival track and the moment when an outbound freight train journey can begin by its departure from a departure track, a lot of technological operations have to be done. One of them is sorting wagons over the hump. From the view of queueing theory, the hump can be considered to be a server; inbound freight trains then represent customers entering the queueing system in order to be serviced and to be classified according to desired destinations of individual wagons.

But in practice, it is often usual to distinguish two types of shunting—we can call them primary and secondary shuntings. Primary shunting is represented by the classification of inbound freight trains entering the yard. Beside

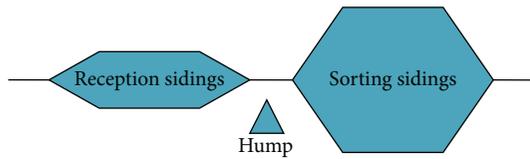


FIGURE 1: Considered arrangement of reception sidings, the hump, and sorting sidings.

primary shunting, it is often necessary to realize the so-called secondary shunting. Secondary shunting, for example, rises from the classification of trains of wagons entering the yard via industrial sidings which often lead into the yard.

Because the hump is equipped only with a single track, it is obvious that performing secondary shunting means that the hump is not able to carry out its main task—sorting inbound freight trains. All requests for secondary shunting, that we are not able to perform in other way or more effective than using the hump, are considered to be failures of the hump. Secondary shunting is also the failure of an arrival track due to the serial arrangement of the yard (train of wagons is shunted from an arrival track towards the hump).

The main aim of the paper is to demonstrate a novel approach to modelling operations on the hump. We model the hump as the server subject to failures. We apply two different approaches. The first approach is based on creating corresponding mathematical model of the problem and its numerical solving; the second approach is based on simulation of the problem using coloured Petri net. We compare the outcomes reached by both approaches in order to validate our models. The presented paper is a substantially enhanced version of our paper [1] written in Czech; we reused some facts and data, which are necessary for a case study and added a lot of new findings regarding the studied problem.

The paper is organized as follows. In Section 2, a short review of the available literature is presented. In Section 3 we make necessary assumptions, and the mathematical model is described. Section 4 is focused on a description of created simulation model. In Section 5, we apply the models in the conditions of a real marshalling yard in the Czech Republic. In Section 6, we make some conclusions.

2. State of the Art

To model queueing systems arising in practical applications, we can employ two basic approaches. The first one is based on making up a mathematical model of the modelled queueing system. The mathematical model can be solved numerically using suitable software tool. In some cases, closed-form formulas enabling the computation of requested system characteristics (probabilities of individual states, performance measures such as the mean number of the customers in the system, etc.) can be derived if it is possible. The second one is a simulation approach consisting of the creation of a simulation model, and requested system characteristics are obtained on the basis of simulation experiments with the simulation model.

Regarding mathematical models of single-server queueing systems, many authors have modelled queueing systems under different assumptions. Many of them use the exponential or general distribution to model random variables occurring in queueing models. Queueing models in which the Erlang distribution is used are not so common despite the fact that such single-server queueing systems can be relatively easily solved. We can mention some queueing models in which authors assumed the Erlang distribution in order to model some random variables of the model.

Wang and Kuo [2] modelled a single-server queueing system represented by a group of identical machines which are maintained by a single nonreliable server—it is a modification of the so-called machine-repairmen problem. The authors of the paper assumed that service times of the machines and repair times of the server are Erlang distributed; other random variables of the model are assumed to be exponentially distributed. Wang and Huang [3] studied a finite single-server queueing system. They assumed that customers come to the system according to the Poisson stream, and service times are considered to be Erlang distributed. The server is so-called removable that means the system turns on and turns off the server in dependency on the number of customers in the system.

Jain and Agrawal [4] also studied a single-server queueing system; they assumed the server takes vacation when it becomes idle and may break down when it is busy. The authors used the Erlang distribution for modelling service times. Yu et al. [5] modelled a single-server queueing system with so-called no damage service interruptions. The model was created in order to model a congestion problem in a Chinese hospital. Service times of customers are considered to be Erlang distributed; the service of a customer can be interrupted after finishing the first stage of service.

Pearn and Chang [6] solved a problem of optimal management of the N -policy in the case of $M/E_k/1$ queueing system with a removable service station. The Erlang distribution was used to model service times of customers. The authors applied outcomes of paper [7]; in the paper some closed-form expressions, for example, for the expected number of customers in the system when the server is turned off, were derived.

As regards using queueing theory for modelling transport processes, we can mention the following papers. In book [8] some mathematical models relating to the rail transport are shown. The authors, for example, presented using a finite queueing system $M/M/n/n$ for modelling reception sidings of marshalling yards. In the model, it is assumed that every arrival track constitutes a server; customers are represented by inbound freight trains. The authors also presented a mathematical model of the process of the preparation of inbound trains for sorting and their sorting over a hump as two queueing systems placed in series; the queueing model constitutes a two-dimensional Markov process. Another two-stage queueing model was presented in paper [9].

In paper [10], a simple queueing system is used in order to model a railway information system. Sharma and Sirohi [11] modelled a container unloader using a finite single-server queueing system with the server subject failures; the

authors further assumed two types of trailers which represent customers in the system. In paper [12], queueing theory is employed in order to estimate container dwell times at rail intermodal terminals and transit times through rail line-haul corridors. Authors of paper [13] presented a queueing network model enabling the computation of selected performance measures for railway networks; the authors derived closed-form formulas for mean delay. In paper [14], two multiserver queueing models were used in order to evaluate the performance of a river port.

In order to study the behaviour of queueing systems using simulation, a lot of simulation tools can be used. For example, the authors of paper [15] compared three discrete-event simulation tools ARENA, WITNESS, and SIMUL 8; the comparison was done for an example of M/G/1 manufacturing queueing system.

Another possible tool enabling simulation of queueing systems is represented by Petri nets. Coloured Petri nets represent a mathematical tool which enables modelling and simulation of discrete-event dynamic systems. More information about coloured Petri nets can be found, for example, in book [16]. Some examples of using Petri nets for queueing systems modelling and simulation can be found, for example, in paper [17] or [18].

3. Mathematical Model

In this section, we will present a mathematical model of the freight trains classification process employing a queueing model with a server, which is constituted by the hump, subject to breakdowns. Let us notice that the presented mathematical model and also simulation model, which will be described later in the text, are based on models presented in paper [19] which was accepted for publication in conference proceedings of 31st International conference on Mathematical Methods in Economics held in Jihlava.

On the basis of previous section, we know that there are usually two types of shunting. Individual requests for primary shunting-trains of wagons waiting on arrival tracks after finishing their preparation for the classification process will be called customers. Let the input process of customers be the Poisson process with the parameter λ . That means interarrival times are exponentially distributed with the mean interarrival time equal to $1/\lambda$. The process of classification over the hump will be considered to be service of customers. The easiest way to create the mathematical model would be in the case of exponentially distributed service times. But in practice there is a problem for using this assumption. From measuring executed by the authors, it resulted that it is more suitable to use a continuous random distribution which offers us greater variability of using than the exponential distribution. Such type of distribution could be the Erlang distribution; the distribution is defined by the shape parameter denoted by n , where we will consider that $n \geq 2$, and the scale parameter denoted by $n\mu$. The Erlang distribution with the parameters n and $n\mu$ arises from n mutually independent exponential distributions with the parameter $n\mu$. Therefore, Erlang service

times can be modelled as n exponential service times. The mean service time is equal to $n/n\mu = 1/\mu$.

By performing secondary shunting, we will further in the text call failures of the server because during secondary shunting the hump and also an arrival track are occupied. If there are several consecutive needs of secondary shunting, then we will unite them into a single failure; this enables us to make the following assumption. Let us assume that we can suppose that times between failures are exponentially distributed with the parameter η . It is clear that the mean time between failures (that means the time when there is no failure in the system) is equal to the reciprocal value of the parameter η . Times to repair are also exponentially distributed but with the parameter ζ . In the case that a failure arises when a customer is being serviced, the customer is served, and after finishing the service we can begin to repair the server.

Regarding the capacity of the queue of waiting customers, it is obvious that the maximal number of trains waiting for sorting should be equal to the number of tracks in the reception sidings denoted by m . But because it was mentioned that in the case of primary shunting and also secondary shunting an arrival track is occupied, we will consider the capacity of the queue to be equal to $m - 1$.

In order to describe possible states of the modelled system, it is necessary to establish the following discrete random variables.

- (i) The variable K takes its values from the set $\{0, 1, \dots, m\}$, and the value of the variable describes the number of the customers being in the system.
- (ii) The variable P takes its values from the set $\{0, 1, \dots, n - 1\}$ and describes how many phases of service the server has already finished.
- (iii) The last variable F takes its values from the set $\{0, 1, 2\}$ with the following meaning. If the variable F is equal to 0, there is no failure (no need of secondary shunting) in the system. If the variable F is equal to 1, the failure of the system is waiting for repair (there is a need of secondary shunting, but the realization of it has to wait because the classification of an inbound freight train is being carried out). And finally, if the variable F is equal to 2, the failure of the system is being repaired (secondary shunting is being carried out).

Using the above mentioned random variables, we can describe each possible state of the system by a triplet (k, p, f) . The state space of the system is formed by a union of the following states subsets:

$$\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3, \quad (1)$$

where

- (i) the states in the subset $\Omega_1 = \{(k, p, f) : k \in \{1, \dots, m\}, p \in \{0, \dots, n - 1\}, f = 0\}$ represent the states in which k customers are found in the system, and the p phases of customer service have already been finished and there is no failure in the system. To these states we have to add the state $(0, 0, 0)$ which corresponds to the state in which the system is empty;

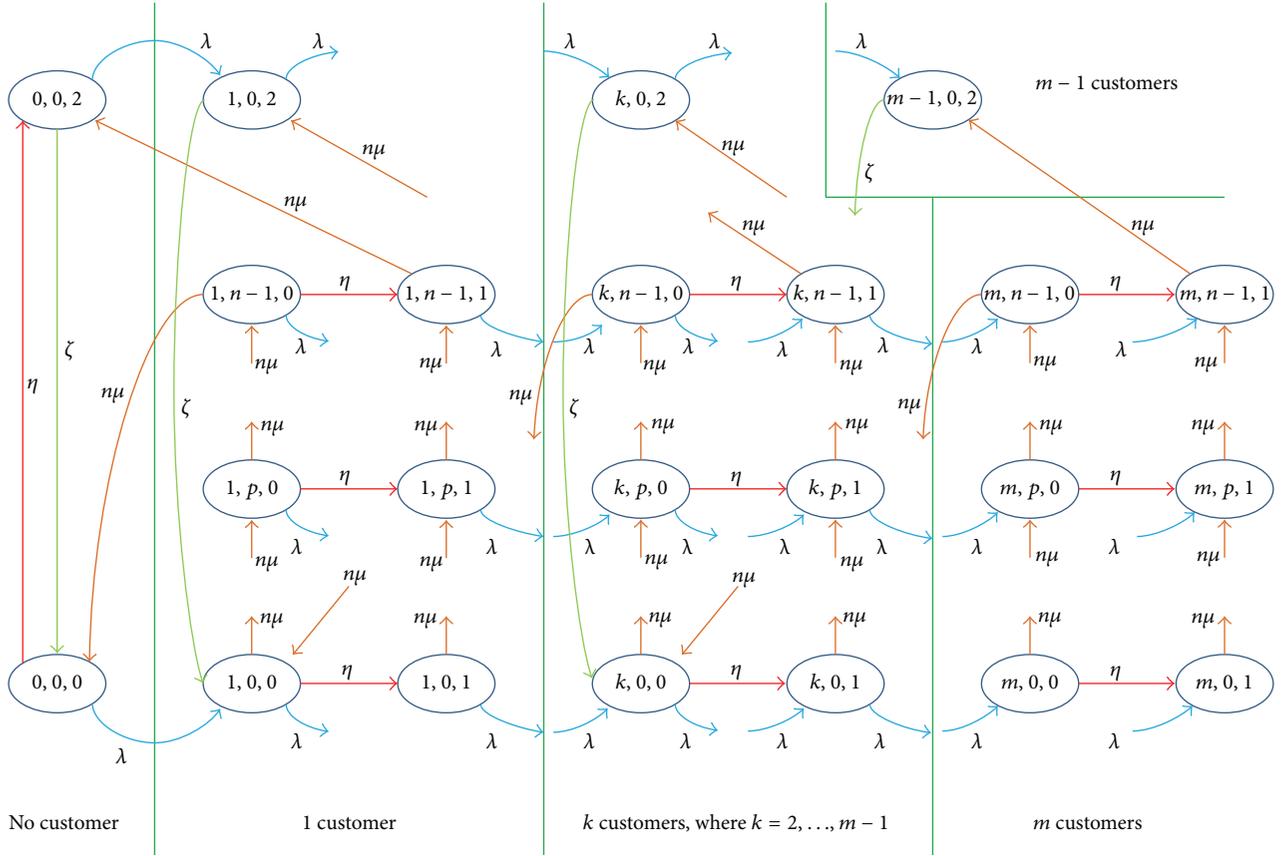


FIGURE 2: The state transition diagram of the system.

(ii) the states in the subset $\Omega_2 = \{(k, p, f), k \in \{1, \dots, m\}, p \in \{0, \dots, n-1\}, f = 1\}$ correspond to the states in which k customers are found in the system, and the p phases of customer service have already been finished and there is a failure of the server in the system waiting for repair;

(iii) the states in the subset $\Omega_3 = \{(k, p, f), k \in \{1, \dots, m-1\}, p = 0, f = 2\}$ are the states in which k customers are waiting in the queue and the server is being repaired.

Now, we can illustrate the studied queueing model graphically as a state transition diagram; see Figure 2. Vertices of the diagram represent the individual states of the system, and oriented edges correspond to the possible transitions in the system; weights of the edges represent the corresponding transition rates.

To simplify the creation of the state transition diagram, we applied the three-dimensional state description in the form of (k, p, f) . However, this state description has an essential lack for our purposes—it is not suitable for numerical computations in Matlab which works with matrices. Therefore, it is necessary to develop an alternative one-dimensional state description in the following form.

(i) We denote the states (k, p, f) for $k = 1, \dots, m, p = 0, \dots, n-1$, and $f = 0, 1$ using a single value $f \cdot m \cdot n + p \cdot m + k$.

(ii) The states $(k, 0, 2)$ for $k = 0, \dots, m-1$ can be denoted using a single value $2 \cdot m \cdot n + k + 1$.

(iii) We label the state $(0, 0, 0)$ as $2 \cdot m \cdot n + m + 1$.

Applying the alternative one-dimensional state description, we are able to write the equation system in the form we need for using Matlab. For the one-dimensional state description, we get the following linear equation system:

$$(\lambda + \eta) \cdot P_{2 \cdot m \cdot n + m \cdot k + 1} = n \cdot \mu \cdot P_{(n-1) \cdot m + 1} + \zeta \cdot P_{2 \cdot m \cdot n + 1}, \quad (2)$$

$$\begin{aligned} (\lambda + n \cdot \mu + \eta) \cdot P_1 \\ = \lambda \cdot P_{2 \cdot m \cdot n + m + 1} + n \cdot \mu \cdot P_{(n-1) \cdot m + 2} + \zeta \cdot P_{2 \cdot m \cdot n + 2}, \end{aligned} \quad (3)$$

for $k = 2, \dots, m-1$:

$$\begin{aligned} (\lambda + n \cdot \mu + \eta) \cdot P_k \\ = \lambda \cdot P_{k-1} + n \cdot \mu \cdot P_{(n-1) \cdot m + k + 1} + \zeta \cdot P_{2 \cdot m \cdot n + k + 1}, \end{aligned} \quad (4)$$

$$(n \cdot \mu + \eta) \cdot P_m = \lambda \cdot P_{m-1}, \quad (5)$$

$$(\lambda + n \cdot \mu) \cdot P_{m \cdot n + 1} = \eta \cdot P_1, \quad (6)$$

for $k = 2, \dots, m-1$:

$$(\lambda + n \cdot \mu) \cdot P_{m-n+k} = \eta \cdot P_k + \lambda \cdot P_{m-n+k-1}, \quad (7)$$

$$n \cdot \mu \cdot P_{m-n+m} = \eta \cdot P_m + \lambda \cdot P_{m-n+m-1}, \quad (8)$$

for $p = 1, \dots, n-1$:

$$(\lambda + n \cdot \mu + \eta) \cdot P_{p-m+1} = n \cdot \mu \cdot P_{(p-1) \cdot m+1}, \quad (9)$$

for $k = 2, \dots, m-1$ and $p = 1, \dots, n-1$:

$$\begin{aligned} & (\lambda + n \cdot \mu + \eta) \cdot P_{p-m+k} \\ & = \lambda \cdot P_{p-m+k-1} + n \cdot \mu \cdot P_{(p-1) \cdot m+k}, \end{aligned} \quad (10)$$

for $p = 1, \dots, n-1$:

$$\begin{aligned} & (n \cdot \mu + \eta) \cdot P_{p-m+m} \\ & = \lambda \cdot P_{p-m+m-1} + n \cdot \mu \cdot P_{(p-1) \cdot m+m}, \end{aligned} \quad (11)$$

for $p = 1, \dots, n-1$:

$$\begin{aligned} & (\lambda + n \cdot \mu) \cdot P_{m-n+p-m+1} \\ & = \eta \cdot P_{p-m+1} + n \cdot \mu \cdot P_{m-n+(p-1) \cdot m+1}, \end{aligned} \quad (12)$$

for $k = 2, \dots, m-1$ and $p = 1, \dots, n-1$:

$$\begin{aligned} & (\lambda + n \cdot \mu) \cdot P_{m-n+p-m+k} \\ & = \eta \cdot P_{p-m+k} + \lambda \cdot P_{m-n+p-m+k-1} + n \cdot \mu \cdot P_{m-n+(p-1) \cdot m+k}, \end{aligned} \quad (13)$$

for $p = 1, \dots, n-1$:

$$\begin{aligned} & n \cdot \mu \cdot P_{m-n+p-m+m} \\ & = \eta \cdot P_{m-p+m} + \lambda \cdot P_{m-n+p-m+m-1} + n \cdot \mu \cdot P_{m-n+(p-1) \cdot m+m}, \end{aligned} \quad (14)$$

$$\begin{aligned} & (\lambda + \zeta) \cdot P_{2 \cdot m-n+1} \\ & = \eta \cdot P_{2 \cdot m-n+m+1} + n \cdot \mu \cdot P_{m-n+(n-1) \cdot m+1}, \end{aligned} \quad (15)$$

for $k = 1, \dots, m-2$:

$$\begin{aligned} & (\lambda + \zeta) \cdot P_{2 \cdot m-n+k+1} \\ & = n \cdot \mu \cdot P_{m-n+(n-1) \cdot m+k+1} + \lambda \cdot P_{2 \cdot m-n+k}, \end{aligned} \quad (16)$$

$$\zeta \cdot P_{2 \cdot m-n+m} = n \cdot \mu \cdot P_{m-n+(n-1) \cdot m+m} + \lambda \cdot P_{2 \cdot m-n+m-1}. \quad (17)$$

Because equation set (2)–(17) is not linearly independent and one equation is therefore redundant and can be omitted, it is necessary to incorporate the normalization condition of probability in the following form:

$$\begin{aligned} & P_{2 \cdot m-n+m+1} + \sum_{k=1}^m \sum_{p=0}^{n-1} \sum_{f=0}^1 P_{f \cdot m-n+p-m+k} \\ & + \sum_{k=0}^{m-1} P_{2 \cdot m-n+k+1}. \end{aligned} \quad (18)$$

After omitting (2), we get the equation system of $2 \cdot m \cdot n + m + 1$ linear equations formed by (3) up to (18) with $2 \cdot m \cdot n + m + 1$ unknown stationary probabilities. The equation system can be numerically solved in Matlab. Let us mention that the mathematical model using the original three-dimensional state description is presented in our conference paper [19].

The mean number of customers found in the service (or in our case the utilization of the hump by primary shunting) ES can be computed according to the following formula:

$$ES = \sum_{k=1}^m \sum_{p=0}^{n-1} \sum_{f=0}^1 P_{f \cdot m-n+p-m+k}. \quad (19)$$

For the mean number of waiting customers EL , it has to hold,

$$\begin{aligned} EL &= \sum_{k=2}^m (k-1) \cdot \sum_{p=0}^{n-1} \sum_{f=0}^1 P_{f \cdot m-n+p-m+k} \\ &+ \sum_{k=1}^{m-1} k \cdot P_{2 \cdot m-n+k+1}. \end{aligned} \quad (20)$$

It is obvious that for the mean number of customers found in the system EK it holds

$$\begin{aligned} EK &= ES + EL = \sum_{k=1}^m k \cdot \sum_{p=0}^{n-1} \sum_{f=0}^1 P_{f \cdot m-n+p-m+k} \\ &+ \sum_{k=1}^{m-1} k \cdot P_{2 \cdot m-n+k+1}. \end{aligned} \quad (21)$$

For the mean number of the broken server (or the utilization of the hump by secondary shunting) EF , it can be written as

$$EF = \sum_{k=0}^{m-1} P_{2 \cdot m-n+k+1}. \quad (22)$$

And finally, for the overall utilization of the hump κ , we can write

$$\begin{aligned} \kappa &= ES + EF \\ &= \sum_{k=1}^m \sum_{p=0}^{n-1} \sum_{f=0}^1 P_{f \cdot m-n+p-m+k} + \sum_{k=0}^{m-1} P_{2 \cdot m-n+k+1} \\ &= 1 - P_{2 \cdot m-n+m+1}. \end{aligned} \quad (23)$$

4. Simulation Model

To validate the proposed mathematical model by comparing reached outcomes, a simulation model of the studied queueing system was created. There are a lot of possible simulation software tools which can be employed to create the simulation model of the studied queueing system (e.g., WITNESS or ARENA). However, we did not use any of them; we chose to develop the model based on coloured Petri nets. We employed the software CPN Tools in version 3.0.4 to create the model. In Figure 3, the created Petri net is depicted in the

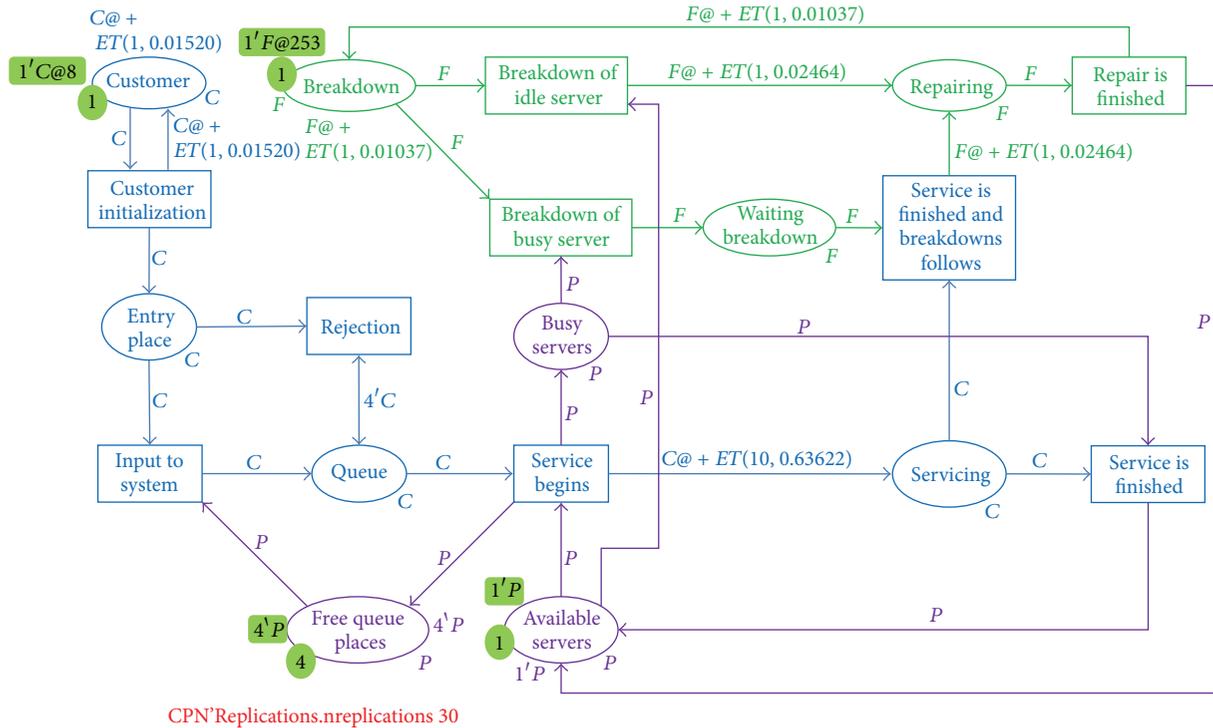


FIGURE 3: The simulation model created using CPN Tools—Version 3.0.4.

initial marking; the values of the model parameters, which are used in the model in Figure 3, correspond to the initial values from the case study—see Section 5.1. The inscription of the model is not depicted in Figure 3 but is described in the text.

The simulation model consists of 10 places and 9 transitions; places are depicted as ovals and transitions as rectangles. Places and transitions modelling the flow of customers through the system are highlighted by blue colour, elements of the Petri net relating to failures of the server are highlighted by green colour, and auxiliary elements of the model are highlighted by violet colour. Places and transitions are connected with arcs which model the relations amongst the individual elements of the model; the arcs with their arc expressions define the flows of tokens in the net.

The model works with following tokens.

- (i) Tokens “C” represent customers; these tokens are defined to be timed in order to model exponentially distributed interarrival times and Erlang distributed service times. Please notice that timed tokens are labelled by so-called timestamp which declares what time the corresponding token can be used at the earliest. Tokens “C” are assigned to places highlighted by blue colour. The declaration is “colset C = with C timed”.
- (ii) Token “F” models failure of the server and is also declared to be timed to model exponentially distributed times between failures and times to repair. Token “F” can be found only in places highlighted by green colour. The declaration is “colset F = with F timed”.

- (iii) Auxiliary tokens “P” model sources of the queueing system (e.g., free places in the queue or free places in the service). Tokens “P” are associated with places depicted by violet colour. The declaration is “colset P = with P”.

The place “Customer” and the transition “Customer initialization” model the Poisson input stream of customers coming to the modelled queueing system in order to be serviced. Exponentially distributed interarrival times are ensured by the arc expression “ $C@ + ET(1, 0.01520)$ ” which is assigned to the arc exiting the transition “Customer initialization” and entering the place “Customer”. Let us mention that the first customer comes to the system after reaching exponentially distributed time with the parameter $\lambda = 0.01520 \text{ h}^{-1}$ that is ensured by defining the initial marking of the place “Customer” in the form “ $C@ + ET(1, 0.01520)$ ”.

Tokens “C” being in the place “Entry place” model customers that have just been generated. Now, one of two disjunctive events can happen.

- (i) The customer can enter the queue—this event can happen only when a free place is in the queue. Free places in the queue are modelled via tokens “P” being in the auxiliary place “Free queue places.” In Figure 3, the initial marking of the place “Free queue places” is equal to “ $4P$ ”, which means that the capacity of the queue is equal to 4 places and the queue is empty in the initial marking (for the value of simulation time equal to 0). It is clear that in the case that the customer joins the queue it is necessary to take away a token “P” from the place “Free queue places”. It is ensured

by the arc exiting the place “Free queue places” and entering the transition “Input to system” that the arc expression is equal to “P”. The transition “Input to system” models the entry of customers into the queue.

- (ii) The customer has to be rejected because no free place is in the queue—that can happen only when the queue is full which means there are in general $m - 1$ customers in the queue. In the case of the model depicted in Figure 3 it means 4 customers have to wait in the queue. The rejection of customers is modelled by the transition “Rejection” which is connected with the place “Queue” by the test arc with the arc expression “4!C”; the arc expression assures that the customer is rejected only when the queue is full.

Tokens “C” in the place “Queue” represent customers waiting for the service. The first customer in the queue can enter the service only when the server is available. The fact that the server is available and ready to serve the customer is modelled by a token “P” in the auxiliary place “Available servers”. The initial marking of the place is defined by the expression “1!P” that means the server is assumed to be available in the initial marking of the model. The transition “Service begins” models the beginning of the service and its firing moves the token “C”, which represents the customer entering the service, from the place “Queue” to the place “Servicing”. Moreover, the transition “Service begins” can be fired only when a token “P” is in the place “Available servers” (the server has to be available). Firing the transition “Service begins” changes the state of the server from available into busy (the token “P” is taken away from the place “Available servers” and is sent to the place “Busy servers”).

A token “C” in the place “Servicing” models a customer which is being served right now. Erlang distributed service times are ensured by the arc expression “C@ + ET(10, 0.63622)” which is defined for the arc exiting the transition “Service begins” and entering the place “Servicing”. The arc expression updates the timestamp of the token “C” modelling the customer entering the service. If any failure has not occurred during the service of the customer, the service of the customer is finished by firing the transition “Service is finished”. Firing the transition changes the state of the server from busy into available by moving the token “P” from the place “Busy servers” into the place “Available servers”.

In the initial marking a token “F” is in the place “Breakdown”. The timestamp of the token corresponds to the value of simulation time when a new failure should occur in the system (in Figure 3 the value of the timestamp is equal to 253 minutes). However, we have to distinguish two possible events:

- (i) The server is being available when a failure occurs—this event is modelled by firing the transition “Breakdown of idle server”. In this case we can immediately begin to repair the server, repairing of the server is modelled by the place named “Repairing”. Moreover, the server is not available when the failure is being repaired, therefore it is necessary to take away the

token “P” from the place “Available servers”. This is done via the arc connecting the place “Available servers” with the transition “Breakdown of idle server”.

- (ii) The server is serving a customer when a failure occurs—this event is modelled by firing the transition “Breakdown of busy server”. Because the fact that failures of the server are considered to be non-preemptive, the failure has to wait in order to be repaired—the token “F” modelling the failure is sent to the place “Waiting breakdown.” After finishing the service of the customer we can begin to repair the failure; this event is modelled by firing the transition “Service is finished and breakdown follows.”

For both cases the arcs entering the place “Repairing” have the arc expression defined in the form “F@ + ET(1, 0.02464)”; the arc expression assures the update of the timestamp of the token “F”, modelling the failure of the server, in order to model exponentially distributed times to repair. When the actual value of simulation time reaches the value which corresponds to the value of the timestamp of the token “F”, which is found in the place “Repairing”, then the repair of the server can be finished by firing the transition “Repair is finished”. Firing the transition changes the state of the server into available and the token “F” is returned back to the place “Breakdown” with updated value of the timestamp.

To obtain requested simulation outcomes some monitoring functions had to be defined:

- (i) The monitoring functions named “ES,” “EL,” and “EF” enable estimation of the performance measures which were considered in Section 3.
- (ii) Because the queueing model is finite, it is obvious that it can happen that an arriving customer has not any free place in the system and therefore it has to be rejected. To estimate the probability that the customer is rejected upon its arrival we were obliged to define two monitoring functions “Number of Rejected Customers” and “Number of Arriving Customers”. These two functions count how many times the transition, to which the function is bound, was fired during each simulation run. The probability P_{rej} can be then estimated as ratio of these values. The problem with the rejection of arriving customers will be discussed later in Section 5.

To generate concrete values of the random variables used in model we defined a function “fun ET(k, mi) = round(erlang(k, mi/1.0));” where k is the shape parameter and mi is the scale parameter of the Erlang probability distribution; the scale parameter is expressed in $[\text{min}^{-1}]$. If we substitute $k = 1$, then the function generates exponentially distributed values. We applied a minute as the basic unit of time for the simulation experiments.

In order to break each simulation run after reaching a predefined value of simulation time we were obliged to define a breakpoint function. The defined function breaks each simulation run after reaching the simulation time equal to 525 600 minutes; that corresponds to 365 days of the hump operation. Moreover, due to the fact that the simulation model

is stochastic, it is also necessary to execute several simulation runs with the model in order to compute confidence intervals for monitored performance measures. To ensure it we added an auxiliary text “CPN’ Replications.nreplications 30” into the model. Evaluating the text 30 simulation runs are automatically performed and simulation outcomes are gained.

5. Case Study

As a demonstration example of using proposed models in the conditions of a real marshalling yard, we will model the classification process at the marshalling yard “Ostrava pravé nádraží”. The yard is a part of the railway station “Ostrava hlavní nádraží”. To the railway station “Ostrava hlavní nádraží” two marshalling yards belong, one of them is “Ostrava pravé nádraží” and the second one is “Ostrava levé nádraží”. For each marshalling yard it is defined for which relations the yard classifies inbound freight trains and makes up outbound freight trains.

The marshalling yard “Ostrava pravé nádraží” is equipped with reception sidings consisting of 5 arrival tracks. By primary shunting we mean the classification of inbound freight trains entering the marshalling yard to be sorted according to individual train wagons destinations.

However, there are also frequent needs for performing secondary shunting in the conditions of the yard. Several industrial tracks are connected to it, the industrial tracks lead into sorting sidings of the yard. Therefore, the yard has to also perform secondary shunting of trains of wagons entering the yard via the industrial tracks; such trains of wagons are shunted from the sorting sidings onto selected arrival track and then are sorted over the hump onto classification tracks. Another need for secondary shunting arises in the case when at “Ostrava levé nádraží” some wagons, which destination belongs to the destinations determined for “Ostrava pravé nádraží”, have been gathered. Then the wagons have to be shunted from “Ostrava levé nádraží” to “Ostrava pravé nádraží”, the wagons are shunted onto free arrival track of the yard “Ostrava pravé nádraží”. Consequently, the wagons have to be shunted over the hump to sort them according their desired destinations.

Due to the fact that the reception sidings, the hump and the sorting sidings of the yard are placed in series, an arrival track and the hump are being occupied when primary or secondary shunting is being performed. Therefore, we will consider that capacity of the queue, intended for waiting inbound freight trains which have been yet prepared for sorting, is equal to 4 (the yard has got 5 arrival tracks but 1 arrival track is being occupied by the shunting process). That means the total capacity m of the modelled system is equal to 5.

5.1. Input Data Analysis. Before we proceed to modelling the shunting process in the marshalling yard using the proposed models, it is necessary to analyze all random variables which have to be considered in the models. All data we needed were obtained by analyzing fulfilled diagrams of station work which were provided by employees of the yard (station dispatchers) and by performing own survey. We obtained 4

data sets containing customer interarrival times, customer service times, times between failures, and times to repair.

The data sets were processed using the following progress.

- (i) Exploratory data analysis—the computation of basic random variable characteristics (sample average, sample deviation, etc.).
- (ii) Construction of histogram of observed frequencies in order to reveal whether random distributions which we assumed in Section 3 can be used to model the corresponding random variable. To estimate the number of bins k we applied Sturges’ rule defined by the formula $k = 1 + 3.3 \log n$, where n is sample size of corresponding data set.
- (iii) On the basis of exploratory data analysis and constructed frequency histograms we can choose suitable theoretical probability distribution and consequently we are able to estimate unknown parameters of the considered distributions employing maximum likelihood method.
- (iv) Finally we have to test the concordance of individual empirical data sets with the chosen theoretical random distribution using Pearson’s χ^2 goodness-of-fit test.

On the basis of executed tests, we did not reject the following hypotheses.

- (i) Customer interarrival times are exponentially distributed with the parameter $\lambda_{\text{ent}} = 0.01520 \text{ min}^{-1}$.
- (ii) Customer service times are Erlang distributed with the shape parameter $n = 10$ and the scale parameter $n \cdot \mu = 10 \cdot \mu = 0.63622 \text{ min}^{-1}$.
- (iii) Times between failures are exponentially distributed with the parameter $\eta = 0.01037 \text{ min}^{-1}$.
- (iv) Service times are exponentially distributed with the parameter $\zeta = 0.02464 \text{ min}^{-1}$.

Now it is necessary to realize one important fact we have mentioned earlier in Section 4. We developed models of queueing system which is finite. For finite queueing systems, it is typical that such systems reject customers upon their arrival in the case that the system is full, and there is no additional free place in the system to accept new customer. For such queueing systems, we can write the following formula:

$$\lambda_{\text{ent}} = \lambda \cdot (1 - P_{\text{rej}}), \quad (24)$$

where λ represents the mean number of customers coming to the system per unit time, P_{rej} the probability that the customer is rejected upon its arrival, and λ_{ent} corresponds to the mean number of customers that truly enter the system per unit time. For the probability P_{rej} , it has to hold

$$P_{\text{rej}} = \sum_{p=0}^{n-1} \sum_{f=0}^1 P_{f \cdot m + p \cdot m + m} + P_{2 \cdot m + m}. \quad (25)$$

TABLE 1: Summary of numerical results.

Iteration	1	2	3	4	5	6	7	8
λ_i [min^{-1}]	0.01520	0.01536	0.01552	0.01544	0.01540	0.01538	0.01537	0.01536
$P_{\text{rej},i}$ [—]	0.01026	0.01058	0.01090	0.01074	0.01066	0.01062	0.01060	0.01058
$\lambda_{\text{ent},i}$ [min^{-1}]	0.01504	0.01520	0.01535	0.01527	0.01524	0.01522	0.01521	0.01520
ε_i [—]	0.000156	0.000002	-0.000151	-0.000074	-0.000036	-0.000017	-0.000007	0.000002

TABLE 2: Summary of reached outcomes.

	Mathematical model	Simulation model	
		Lower bound	Upper bound
<i>ES</i>	0.23887	0.23724	0.23944
<i>EL</i>	0.32530	0.32358	0.33270
<i>EK</i>	0.56417	0.56139	0.57156
<i>EF</i>	0.29025	0.28973	0.29303
κ	0.52912	0.52774	0.53169

In our case it is necessary to realize that customers in our queueing system are represented by inbound freight trains which have been yet prepared for sorting over the hump. The preparation of inbound freight trains is done on the arrival track on which the inbound freight train has stopped. When all necessary operations with the train have been finished, the customer represented by the prepared train of wagons enters our modelled queueing system. And because by analyzing diagrams of station work we obtained these moments, it is obvious that we do not know the value of the parameter λ but we know the value of the parameter λ_{ent} as listed above.

5.2. Algorithm for the Approximation of the Unknown Value λ . Therefore, to make the outcomes gained by proposed models more accurate, it is necessary to approximate the unknown value of the parameter λ . It can be done by simple iterative algorithm. The algorithm is shown by a flow chart divided into two parts—the first part is depicted in Figure 4 and the second part in Figure 5. Let us notice that the symbol ε denotes the predefined maximal error of the approximation of the value λ and λ_{ent} corresponds to the value of customer interarrival time given in Section 5.1.

The first stage of the algorithm is based on iterative increasing of the value of the parameter λ . In the case that we do not reach the requested accuracy of the approximation in the frame of the first stage then it is necessary to use the second stage of the algorithm which is based on bisection method.

5.3. Summary of Reached Outcomes. In this section we present reached outcomes. In the frame of the first phase of the computations it was necessary to obtain the unknown value of the parameter λ according to the algorithm described in previous section. Let the required value of the maximal error ε be equal to 0.00001.

The outcomes reached using the proposed algorithm are summarized in Table 1. We can see that after 2nd iteration we have reached the value of the parameter λ with demanded

accuracy. Let us mention that this value we obtained using the first stage of the algorithm (see Figure 4). To test the second stage of the algorithm we proceeded with the computations according to the flow chart depicted in Figure 5. We can see that 7th iteration gave us the outcome with requested accuracy, after 8th iteration we reached the same error as after 2nd iteration. The requested value of the parameter λ is equal to 0.01536 min^{-1} .

Now we can compute the performance measures we are interested in. The computed values of them are listed in the second column of Table 2. The values of the performance measures were also estimated using the simulation model presented in Section 4, for each performance measure we computed 95% both-sided confidence interval. We can see that for all performance measures the value computed by the proposed mathematical model lies in the confidence interval gained by the simulation experiments. Therefore we can state that both created models model the same queueing system and can be used individually in order to get the performance measures of the studied system.

In Table 2 we can see the occupation of the hump by primary shunting equals to 23.89% and the occupation of it by secondary shunting equals to 29.03%, the overall hump utilization is then approximately 53%.

6. Conclusions

In the paper, a novel approach to modelling the freight trains classification process performed over the hump in the marshalling yard was presented. Two types of shunting over the hump were considered—primary shunting representing the classification of inbound freight trains and secondary shunting which arises, for example, in the case of the classification of trains of wagons entering the yard via industrial sidings. When secondary shunting is being performed, primary shunting, which is considered to be the main task of marshalling yards, cannot be carried out. Due to that the fact, all needs for secondary shunting are considered to be failures of the service facility which is represented by the hump. Considering dividing the shunting processes enables, for example, to study the marshalling yard capacity for primary shunting under the assumption of certain extent of secondary shunting.

The process was modelled as the finite queueing system with the server subject to failures. We developed two models. The first one is the mathematical model and the second one is the simulation model. The mathematical model was created employing Markov processes; the simulation one was created using coloured Petri nets. Both models were tested in the

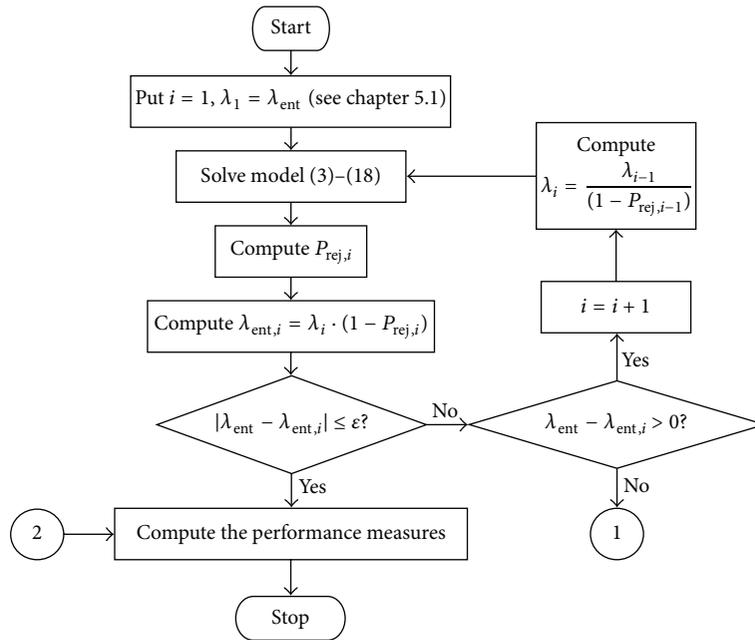


FIGURE 4: The first stage of the algorithm.

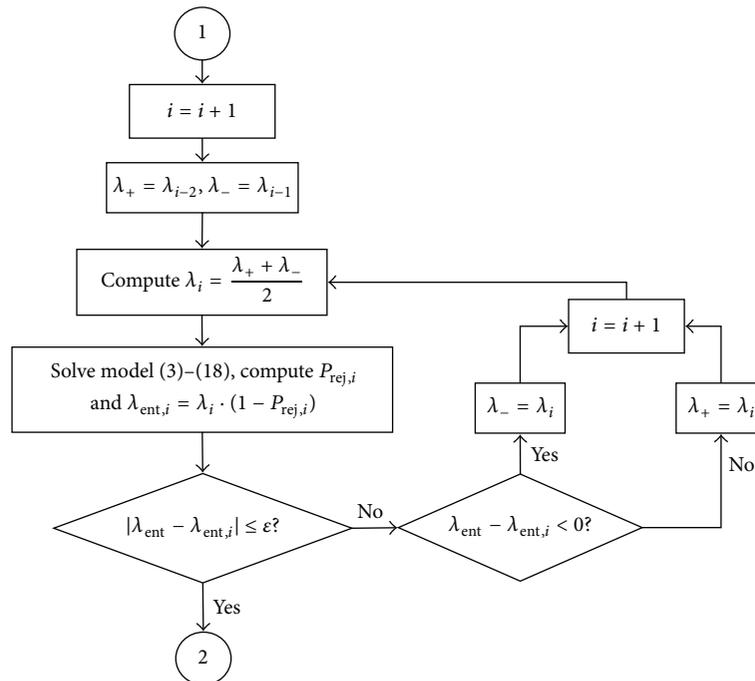


FIGURE 5: The second stage of the algorithm.

condition of the marshalling yard “*Ostrava pravé nádraží.*” From diagrams of station work and measures in the yard, we obtained necessary information about random processes involved in the model.

On the basis of the reached outcomes, we can state that presented models are successfully able to model the classification processes performed over the hump. Regarding our future research, we would like to generalize the proposed mathematical model to enable using the model also for the

cases when interarrival times, times between failures, or times to repair are not exponentially distributed.

Conflict of Interests

The authors of the paper state that they do not have any conflict of interests with any commercial identity which was mentioned in the presented paper.

References

- [1] M. Dorda and D. Teichmann, "Modelování provozu na na svázném pahrbku seřadovací stanice pomocí nespolehlivého systému hromadné obsluhy," *Perner's Contacts*, vol. 5, no. 1, pp. 38–46, 2010.
- [2] K.-H. Wang and M.-Y. Kuo, "Profit analysis of the M/Ek/1 machine repair problem with a non-reliable service station," *Computers and Industrial Engineering*, vol. 32, no. 3, pp. 587–594, 1997.
- [3] K.-H. Wang and H.-M. Huang, "Optimal control of a removable server in an M/Ek/1 queueing system with finite capacity," *Microelectronics Reliability*, vol. 35, no. 7, pp. 1023–1030, 1995.
- [4] M. Jain and P. K. Agrawal, "M/E_k/1 queueing system with working vacation," *Quality Technology & Quantitative Management*, vol. 4, no. 4, pp. 455–470, 2007.
- [5] M. Yu, Y. Tang, Y. Fu, and L. Pan, "An M/Ek/1 queueing system with no damage service interruptions," *Mathematical and Computer Modelling*, vol. 54, no. 5-6, pp. 1262–1272, 2011.
- [6] W. L. Pearn and Y. C. Chang, "Optimal management of the N-policy M/Ek/1 queueing system with a removable service station: a sensitivity investigation," *Computers and Operations Research*, vol. 31, no. 7, pp. 1001–1015, 2004.
- [7] K.-H. Wang, "Optimal control of an M/Ek/1 queueing system with removable service station subject to breakdowns," *Journal of the Operational Research Society*, vol. 48, no. 9, pp. 936–942, 1997.
- [8] F. Brandalík and P. Kluvánek, *Operační Analýza v Železniční Dopravě*, ALFA Bratislava, 1984.
- [9] O. Hein, "A two stage queue model for a marshalling yard," *Rail International*, vol. 3, no. 4, pp. 249–259, 1972.
- [10] A. Melikyan, "The railway information system as a queueing system," *Transport*, vol. 24, no. 4, pp. 339–344, 2009.
- [11] K. C. Sharma and A. Sirohi, "Cost analysis of the unloader queueing system with a single unloader subject to breakdown with two types of trailers," *Opsearch*, vol. 47, no. 1, pp. 93–103, 2010.
- [12] R. C. Leachman and P. Jula, "Estimating flow times for containerized imports from Asia to the United States through the Western rail network," *Transportation Research E*, vol. 48, no. 1, pp. 296–309, 2012.
- [13] T. Huisman, R. J. Boucherie, and N. M. Van Dijk, "A solvable queueing network model for railway networks and its validation and applications for the Netherlands," *European Journal of Operational Research*, vol. 142, no. 1, pp. 30–51, 2002.
- [14] B. Dragović, N.-K. Park, N. Zrnić, and R. Meštrović, "Mathematical models of multiserver queueing system for dynamic performance evaluation in port," *Mathematical Problems in Engineering*, vol. 2012, Article ID 710834, 19 pages, 2012.
- [15] P. Semanco and D. Marton, "Simulation tools evaluation using theoretical manufacturing model," *Acta Polytechnica Hungarica*, vol. 10, no. 2, pp. 193–204, 2013.
- [16] K. Jensen, *Coloured Petri Nets: Basic Concepts, Analysis Methods and Practical Use*, Springer, 1996.
- [17] M. Dorda and D. Teichmann, "Cost analysis of finite multi-server markov queueing system subject to breakdowns," in *Proceedings of the 29th International Conference Mathematical Methods in Economics*, pp. 134–139, Janská Dolina, Slovakia, September 2011.
- [18] M. Dorda and D. Teichmann, "About a modification of E_r/E_s/1/m queueing system subject to breakdowns," in *Proceedings of the 30th International Conference Mathematical Methods in Economics*, pp. 117–122, Karviná, Czech Republic, September 2012.
- [19] M. Dorda and D. Teichmann, "On finite single-server queue subject to non-preemptive breakdowns," in *Proceedings of the 31th International Conference Mathematical Methods in Economics*, Jihlava, Czech Republic, September 2013.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

