International Journal of Innovative Research and Scientific Studies, 8(6) 2025, pages: 1333-1347

IJIRSS International Journal of Innovative Research & Scientific Studies

**ISSN:** 2617-6548

URL: www.ijirss.com

# Modeling railway section capacity under semi-automatic blocking: Implications for freight logistics and schedule reliability

Zhanibek Shukamanov[1*], Sergei Bessonenko[2], Gulfariza Suleimenova[3], Liliya Vakhitova[4]

[1]*Department of Automation and Telecommunications, Faculty of Technology, M.Kh. Dulaty Taraz University, Tole bi 60, 080000 Taraz, Kazakhstan.*
[2]*Department of Transportation Operations Management, Siberian Transport University, 191 Dusi Kovalchuk Street, 630000 Novosibirsk, Russia.*
[3,4]*Department of Automation and Control, Institute of Energy and Digital Technologies, M. Tynyshpayev University, Shevchenko 97, 050000 Almaty, Kazakhstan.*

Corresponding author: Zhanibek Shukamanov (*Email: zhanibek.shukamanov@gmail.com*)

## Abstract

This study aims to enhance the accuracy of railway section capacity estimation in freight logistics by accounting for random delays inherent in semi-automatic blocking (SAB) systems. These systems are still widely used in regions such as Kazakhstan, where manual dispatching introduces variability that affects scheduling and throughput reliability. A discrete-event simulation model was developed using AnyLogic software to assess the impact of dispatch delays on capacity. The delays were modeled as random variables following three statistical distributions: normal (for stable operation), exponential (for high uncertainty), and empirical (based on observed field data). The model calculates train intervals and daily throughput under each delay scenario. The results show that ignoring stochastic delays may lead to capacity overestimation by up to 30%. Depending on the delay profile, estimated throughput ranged from 64 to 77 trains per day. The empirical distribution yielded the most realistic result, approximately 69 trains per day, closely reflecting real operational conditions. Accounting for delay variability improves the realism and accuracy of capacity estimation, especially in manually operated systems where automation is limited. The proposed method offers a decision-support tool for logistics planners and infrastructure managers to improve scheduling, evaluate operational risks, and guide investments in dispatch system modernization.

**Keywords:** Capacity modeling, Rail logistics, Schedule reliability, Semi-automatic blocking, Stochastic delays.

**DOI:** 10.53894/ijirss.v8i6.9914
**Funding:** This study was funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant Number: AP22685501).
**History: Received:** 22 July 2025 / **Revised:** 26 August 2025 / **Accepted:** 28 August 2025 / **Published:** 18 September 2025

**Competing Interests:** The authors declare that they have no competing interests.
**Authors' Contributions:** Conceptualization, Zh.Sh., G.S and L.V.; methodology, Zh.Sh. and S.B.; software, Zh.Sh.; validation, Zh.Sh., G.S. and L.V.; formal analysis; investigation, Zh.Sh.; resources, G.S.; data curation, Zh.Sh.; writing – original draft preparation, Zh.Sh. and S.B.; writing – review and editing, L.V.; visualization, G.S. and S.B.; supervision, Zh.Sh.; project administration, Zh.Sh.; funding acquisition, Zh.Sh. All authors have read and agreed to the published version of the manuscript.
**Transparency:** The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
**Publisher:** Innovative Research Publishing

## 1. Introduction

In the context of limited opportunities to expand railway infrastructure through the construction of new tracks or large-scale modernization, optimizing the use of existing assets has become a strategic objective for sustainable transportation development. Increasing line capacity without significant capital investments is especially critical given the growing competition between transportation modes and the rising demand for reliable and efficient freight logistics. The capacity of a railway line directly influences not only the number of trains that can pass through a section per unit of time but also the stability of transport schedules, operational costs, profitability, and the overall efficiency of logistics chains [1, 2].

Today, the intensification of freight flows along key corridors highlights insufficient capacity as a major constraint not only to transportation network performance but also to broader economic growth. Empirical evidence shows that congested railway sections lead to extended waiting times, decreased average speeds, and frequent disruptions in freight and passenger schedules [3, 4]. Moreover, limited capacity increases operational risks and reduces resilience, particularly in high-density corridors where flexibility and maneuverability are already restricted.

Capacity management and development are regarded as priorities in strategic transportation programs. For instance, the European Commission's Smart and Sustainable Mobility Strategy highlights digitalization and the optimization of railway traffic management as critical elements for enhancing network efficiency [5]. Similar approaches have been observed in Asian countries, where comprehensive measures are being taken to increase the operational flexibility of existing lines due to high infrastructure pressure [6].

Among the key determinants of railway capacity is the type of interval control system in use. These systems regulate train spacing to ensure safe operation, and the length of the required interval governs how many trains can traverse a section in a given time frame [7]. Several factors influence interval duration, including rolling stock characteristics (e.g., braking distance), track layout, the degree of automation in dispatching, and the responsiveness of traffic control systems.

The intervals between trains depend on several factors, including the technical characteristics of the rolling stock (e.g., braking track), track profile, level of automation of the traffic control system, and reaction time to changes in the situation. In traditional systems, such as relay autoblocking, the line is divided into fixed block sections, each of which can only be occupied by one train. The next train is permitted to enter a block section only after the previous train has fully vacated it. While this approach provides a high level of safety, it also leads to significant limitations on the minimum possible interval, especially on sections with heavy traffic [8].

Modern interval control systems use flexible approaches that enable real-time control. Digital and radio communication technologies, such as radio interlocking (including ETCS Levels 2 and 3, CBTC, etc.), occupy a special place among these systems. They are based on constant data exchange between the train and the control center. These systems eliminate the need for rigidly defined block section lengths, forming dynamic intervals that adapt to the current situation [9, 10]. This significantly reduces the time gap between trains, increases schedule density, and enhances the section's capacity without requiring additional infrastructure.

Thus, the choice and application of a particular interval train control system have a decisive impact on the efficiency of railway infrastructure utilization. The transition from traditional relay systems to more modern digital solutions not only increases throughput capacity but also provides more flexible traffic control, which is especially important in conditions of high transportation demand dynamics and the need to respond quickly to schedule changes [11].

However, the implementation of such systems requires complex analysis and justification in terms of efficiency, feasibility, and possible limitations. In this regard, simulation modeling methods that allow reproducing the operation of a railway section under different conditions and traffic control scenarios are becoming relevant. With the help of simulation, it is possible to assess the impact of various factors on throughput capacity, identify bottlenecks in traffic organization, and compare the effectiveness of different regulation systems [12, 13].

While much of the existing literature on capacity modeling focuses on highly automated systems, large portions of the railway network particularly in post-Soviet countries continue to rely on semi-automatic blocking (SAB). In SAB systems, dispatch authorization is confirmed manually, introducing variability and delay into train operations. These delays are often stochastic in nature but are rarely captured in conventional capacity estimates, leading to overoptimistic planning and reduced schedule reliability.

This study addresses this gap by developing and validating a simulation-based model for estimating railway section capacity under SAB conditions, explicitly incorporating random dispatch delays. By comparing different statistical distributions normal, exponential, and empirical the study evaluates their impact on inter-train intervals and overall daily throughput, with the goal of improving schedule reliability and informing logistics planning under infrastructure and automation constraints.

## 2. Materials and Methods

### *2.1. Features of Semi-Automatic Locking*

Semi-automatic interlocking is one of the earliest and simplest forms of interval train control. Historically, it represents a transitional stage from manual to automated traffic control methods. First implemented in the late 19th century, it is still used today on low-volume railroad sections, particularly in countries with extensive infrastructure and limited automation. SAB operation is based on the presence of telegraphic or telephone communication between station duty officers bordering the crossing. Manual control is used instead of automatic train position control. Permission to occupy the crossing is given only after receiving confirmation of clearance from the previous train from the destination station. Thus, only one train can occupy a level crossing at a time. The next train receives permission to depart only after the crossing has been cleared. This significantly increases the technical inter-train interval and reduces the section's capacity [14, 15]. In practice, this scheme significantly reduces the sustainability of Kazakhstan's logistics chains and increases the risk of incidents due to personnel errors, confirmed in up to 35% of cases in dense traffic [16]. Table 1 summarizes the key features of interval train control systems.

**Table 1.**
Main characteristics of interval train control systems.

| Characteristics | SAB | AB | RB |
|---|---|---|---|
| Traffic control | Manual | Automatic | Digital |
| Confirmation time | 2-10 minutes | 0.5-1 minutes | < 0.2 minutes |
| Interval type | Fixed | Depends on the block party | Dynamic |
| Participation of the duty officer | Yes | No | No |
| Reliability | Middle | High | Very High |
| Throughput capacity | 50-65% of the possible | 80-90% | 90-100% |

Technically, the system consists of telephone or radio communication, operational negotiations between stations, and the recording of departures and arrivals in paper or electronic logs. However, there are no automatic sensors that record the passage of trains along the line. Because the entire system relies on trust in personnel and their actions, it is vulnerable to human error. Studies have shown that up to 35% of all incidents in heavy traffic are caused by personnel errors when using SABs [16]. Additionally, SAB does not allow for centralized monitoring of train movement, which limits the use of intelligent decision support systems and automatic schedule regulation. In the event of abnormal situations, such as an emergency stop, communication breakdown, or technical malfunction, the system's response depends entirely on humans. This makes ensuring the safety and sustainability of the transportation process difficult [17].

Nevertheless, semi-automatic interlocking is still used on local lines and sections with limited passenger traffic. Its main advantages are the low cost of implementation and operation and the relative ease of maintenance. In some regions, modernized forms of SABs include microprocessor-based control elements, also known as computer-based SABs. These systems retain the logic of manual coordination but use digital communication, maintain an electronic message transmission protocol, and implement automatic time and event recording [18, 19].

According to simulation studies, using SAB on sections with a traffic intensity of more than 30 pairs of trains per day results in significant time losses and inefficient infrastructure usage. Studies on real Kazakhstani routes demonstrate the significant potential of transitioning to digital systems. For instance, a model of container traffic along the Dostyk-Zhezkazgan-Iletsk corridor revealed that modernizing the infrastructure and transitioning to virtual interlockings could increase the section's throughput capacity and resource efficiency [20]. Similar results were obtained for the Trans-Caspian route corridor. Modeling revealed that traffic significantly accelerated and downtime was reduced due to the optimization of joint government and corporate logistics coordination [21].

One of Kazakhstan's defining features is its extensive regional lines with low traffic density. In these areas, the implementation of complex control systems may not be economically feasible. Here, SAB remains important, particularly as a temporary measure before transitioning to digital systems. However, studies show that SAB reduces throughput to 55-65% even at low traffic volumes (10-20 pairs of trains per day), which is much lower than similar indicators for autoblocking or radio-blocking (about 90%).

Additionally, implementing a comprehensive SAB in Kazakhstan's harsh climate is complicated by unreliable communication and limited infrastructure. Modeling these conditions shows that, in the event of communication failure or personnel error, the time between trains increases by 30-40%. This leads to disruptions in logistics schedules and reduces the competitiveness of the railway sector [20, 21].

Thus, SAB in Kazakhstan is a limited, temporary solution for low-intensity lines. However, its long-term use results in reduced transparency, reliability, and efficiency in transportation processes.

Unlike automatic blocking, where the inter-train interval is solely determined by the time a train passes a block partition and when it is released, SAB adds another component: the time spent waiting for confirmation and preparing for departure. Consequently, the minimum possible interval between trains increases, as does the capacity of the section.

Moreover, the waiting time for confirmation is not constant. It depends on staff workload, communication speed, reaction time, and many other external factors. This gives the waiting time a stochastic character. This factor is of scientific interest and requires modeling that takes random delays into account.

### *2.2. Factors Causing Delays*

In a semi-automatic interlocking system, the delay between the arrival of the previous train and the authorization of the next train is a key element that directly impacts throughput. Unlike automatic systems, where the sequence of actions is standardized and performed by technical means, the process of confirming arrival and granting permission in SAB depends on a variety of external and internal factors. These delays are not fixed and are subject to random deviations, giving them a stochastic character [22].

The main reasons for variation in resolution transfer time fall into the following categories:

1. Human factor: The timeliness of information transmission depends on the station duty officer, who is responsible for controlling and dispatching trains. Under conditions of high workload, fatigue, shift work, and lack of automation, there may be delayed reactions, errors in signal transmission, the need for multiple confirmations, and other subjective factors. This is especially true at sites where the duty officer performs multiple functions. The human factor likely contributes the most, since the work of a station duty officer requires high concentration, accuracy, stress resistance, and the ability to multitask. In conditions involving a large number of operations (receiving, dispatching, maneuvers, and communication), there is a probability of error.

- Delayed response to a request for dispatch;
- Errors in checking the condition of the crossing;
- Time losses for clarification of information;
- Transitions between tasks, or cognitive switches.

This is especially common in areas with high traffic density, poor lighting, or understaffing, particularly at night. Additionally, human performance is inconsistent the same duty officer may perform differently at the beginning and end of a shift, on weekdays versus weekends, and in normal versus stressful situations [23].

2. The peculiarities of the communication channel. In most railway sections, SAB still operates using wired analog channels, telephone lines, and radio stations. These are subject to:

- Temporary disruptions;
- Dial-up delay;
- Blurred hearing, noise, interference;
- Non-confirmation of transfer.

Even when people are on duty and ready to interact, delays may occur due to the technical limitations of the channel. Modern digital technologies, such as GSM-R, TETRA, and LTE-R, are still implemented in a fragmented manner. Their implementation requires large-scale modernization, which necessitates accounting for delays in current systems. In areas without modern digital facilities, such as GSM-R or TETRA, communication between stations is carried out via analog channels, wire lines, or radios. These channels are often subject to noise, temporary failures, and quality fluctuations. Transition areas between technologically sophisticated and less sophisticated areas are particularly sensitive to such failures. Delays can occur in both signal generation and acknowledgment, especially if the equipment requires manual operation [24].

3. External Conditions: Temporary disruptions in radio communications, impaired visibility of signals, and the need for additional checks can result from weather effects (e.g., rain, thunderstorms, high winds, and fog), temperature variations, and electromagnetic interference. These factors increase the total time required to transmit authorization [25].

4. Organizational and infrastructure failures: Sometimes delays are caused by unforeseen situations: temporary unavailability of the duty officer, track maintenance, delays of previous trains, failure of signaling devices, high station load during peak hours, etc. These events occur infrequently, but their impact can be critical, especially in dense traffic schedules [26].

The combined effect of these factors leads to the fact that the time of resolution transfer at SAB becomes a variable value, changing from train to train and from day to day. For this reason, in order to build a realistic throughput model, it is necessary to take into account the stochastic nature of delay by setting it as a random variable with a certain distribution law [27].

### *2.3. Formalization of the Throughput Estimation Model*

Given the above, the delay in transmitting train departure authorization under semi-automatic interlocking conditions can be represented as a random variable, $\tau_d$. It reflects the additional time that elapses between the arrival of the previous train and the station duty officer's authorization of the next train's departure [28].

Classical approaches to capacity estimation consider the period of the train schedule at the boundary crossing, denoted as $T_p$ . This period includes the total travel time of a pair of trains along the boundary crossing. It also takes into account their acceleration and deceleration, as well as the station time intervals at the split points adjacent to the boundary crossing (a and b, respectively), provided by the skip pattern.

$$T_p = \sum t_x + \tau_a + \tau_b \, , \qquad (1)$$

where $T_p$ – period of the train schedule at the limiting crossing; $\sum t_x$ – total time of movement of a pair of trains along the limiting crossing, taking into account their acceleration and deceleration provided for by the skip scheme; $\tau_a, \tau_b$ – station time intervals, stipulated by the train passing scheme, at the separate points a and b, respectively, adjacent to the boundary crossing.

However, under SAB conditions, the period of the train schedule at the restraining crossing must account for the delay: $\tau_d$

$$T_p = \left(\sum t_x + \tau_a + \tau_b\right) + \tau_d \, , \qquad (2)$$

where $\tau_d$ – random resolution transmission delay.

Thus, the throughput capacity of the single-track railroad section $N$ is calculated as the number of pairs of trains per day using the following formula:

$$N = \frac{(1440 - t_{tech}) \cdot a_r}{T_p} = \frac{(1440 - t_{tech}) \cdot a_r}{(\sum t_x + \tau_a + \tau_b) + \tau_d}, \qquad (3)$$

where 1440 minutes in a day; $t_{tech}$ exclusion for train traffic from the daily duration of the time (24 hours) required for maintenance of technical devices and scheduled repairs of devices (for single-track railroad sections, 75 minutes is accepted); $a_r$ – reliability coefficient (for single-track electrified railroad lines - 0.93, and for non-electrified lines - 0.92).

This formalization enables us to quantify the impact of stochastic delays on intervals and throughput. The following will model the $\tau_d$ using different probability distributions that reflect typical SAB operating scenarios [28].

### *2.4. Delay probability distributions*

In most traditional bandwidth estimation models, the transmission delay of the resolution is assumed to be zero or a constant, known, and repeatable value. This assumption is only acceptable for highly automated systems (e.g., automatic interlocking with digital signaling), where stable transmission of information is possible. However, under SAB conditions, this assumption significantly distorts the modeling results and overestimates the actual capacity of the section [29].

Classical estimates of SAB throughput often make the simplifying assumption that:

- or, there is no delay at all, and the resolution is instantaneous;
- or, it may have a fixed value (e.g., one minute).

Such assumptions are only appropriate in cases of full process automation, which is rarely observed in practice. Real delays fluctuate and cannot be accurately predicted, leading to calculation errors.

This is especially critical in tight schedules, where each extra interval leads to the loss of a train slot, and thus to reduced throughput, the formation of «tails», and even failures in neighboring sections [29].

To eliminate this drawback and increase the model's validity, this study treats the delay, $\tau_d$, as a random variable that obeys a known law of probability distribution. This formulation allows us to model the system's behavior more realistically, take into account fluctuations, instability, and the probabilistic nature of delays, and perform a comparative analysis under different operating conditions.

Introducing a stochastic element into the description of the resolution transfer process enables us to evaluate the likelihood of significant delays and their impact on train interval formation, the total number of possible departures during the calculation period, and consequently, the overall capacity of the railway section. This is particularly important when analyzing boundary scenarios where a line operates at capacity, as even small fluctuations can cause cascading effects and schedule shifts [30].

Thus, including the random variable $\tau_d$ in the calculation of the train interval not only refines the estimate but also elevates the analysis to a new level, presenting the system as dynamic and stochastic rather than static and simplified [30]

To achieve more realistic modeling, this study treats the parameter $\tau_d$ as a random variable that obeys one of the following probability distributions.

- Normal distribution occurs when personnel work is stable;
- Exponential distribution: Ideal for high fluctuations and failures;
- Empirical distribution is based on actual, observed data.

Each of them reflects certain operational features and scenarios of railroad sections.

#### *2.4.1. Normal distribution*

The normal distribution, also known as the Gaussian distribution, is one of the most commonly used distribution laws in applied statistics and probability theory. In the context of modeling railway processes, for example, it is used to describe natural fluctuations caused by the combined effect of many weak random factors [31].

The normal distribution is mathematically defined as follows:

$$f(\tau_d) = \frac{1}{\sigma\sqrt{2\pi}} \cdot exp\left(-\frac{(\tau_d - \mu)^2}{2\sigma^2}\right), \qquad (4)$$

where $f(\tau_d)$ – probability density; $\mu$ – mathematical expectation (mean value); $\sigma$ – standard deviation (characterizes the «spread» of values).

This paper models the delays in the normal distribution using the following parameters:

- $\mu = 3$ minutes – average delay;

- $\sigma = 1$ minute – standard deviation.

These values correspond to the average delays observed at sites with relatively stable connectivity and manual control.

To eliminate incorrect values (negative delays), trimming of the distribution to the left is applied:

$$\tau_{d-modeled} = \max(0, \tau_d)\,, \qquad (5)$$

This is acceptable, as the delay cannot possibly be less than zero.
The normal distribution is particularly useful for describing situations in which:

- Communication between stations is stable, but there is no automation.
- The operator (station duty officer) works at a predictable speed with permissible fluctuations;
- The influence of extraneous factors (e.g., weather, workload, and complexity of the station layout) is insignificant.

Examples of such sites include:

- Single-track crossings with good visibility and regular schedules;
- Lightly loaded stations with experienced personnel;
- Stations are equipped with analog communications and a high level of performance discipline.

The normal distribution is characterized by two parameters: mean $\mu = 3$ minutes and standard deviation $\sigma = 1$ minutes. This choice reflects moderate deviations and relatively stable personnel and equipment operation.

$$\tau_d \sim N(\mu = 3, \sigma^2 = 1)\,, \qquad (6)$$

Figure 1 shows the histogram of the normal delay distribution generated by $\mu = 3$ and $\sigma = 1$. Characteristic features:

- Bell-shaped: most of the delays are concentrated near the mean value;
- Approximately 68% of the values fall within $\mu \pm \sigma$ (2 to 4 minutes);
- Both delays that are too short (less than one minute) and those that are too long (more than five minutes) are unlikely.

The density curve (KDE) superimposed on the histogram confirms the theoretical shape.

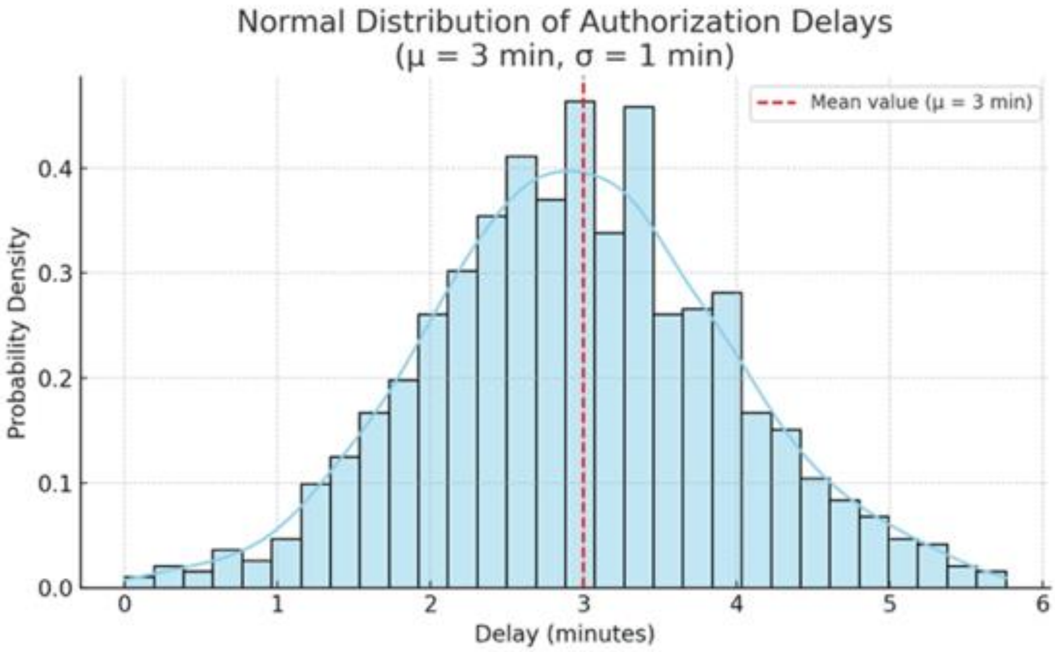


**Figure 1.**
Normal Distribution of Authorization Delays.

The normal distribution provides the best approximation of SAB system behavior under stable personnel and infrastructure conditions. It accurately reproduces fluctuations in the interval between train departures and estimates the average capacity of a section under typical conditions.

However, for sites with unstable connectivity, human error, or a high degree of uncertainty, exponential or empirical distributions are more appropriate.

*2.4.2. Exponential Distribution*

The exponential distribution is widely used in mass service theory, reliability modeling, and systems with rare events. It describes processes in which the time between events (e.g., delays, failures, or requests) follows a decreasing probability law: The longer the delay, the less likely the event is to occur; however, the probability is never zero [32].

The probability density function is mathematically represented by the following formula:

$$f(\tau_d) = \lambda e^{-\lambda \tau_d}, \qquad \tau_d \geq 0 \qquad (7)$$

where λ – intensity parameter (frequency of occurrence of events).

The average delay value is determined:

$$\mu = \frac{1}{\lambda} \tag{8}$$

The exponential distribution applies when delays are predominantly random and short intervals are likely, though long delays are still possible. It is described by one parameter: intensity ($\lambda$). The study used a value of $\lambda = 0.4$, which corresponds to an average value of $\mu = 2.5$. This brings the model closer to a scenario involving high uncertainty and unstable coupling.

$$\tau_d \sim Exp(\lambda = 0.4), \tag{9}$$

The exponential distribution best describes unstable, irregular systems where:

- The connection may be cut off;
- The operator may be distracted or take a long time to respond;
- The order in which requests are handled may be disrupted;
- Competition between trains may occur.

Characteristics:

- High density near zero: Many short delays;
- A long right tail indicates rare but possible long delays;
- Asymmetry: Unlike a normal distribution, this distribution is not symmetric.

Figure 2 illustrates an exponential distribution with a parameter of $\lambda = 0.4$.

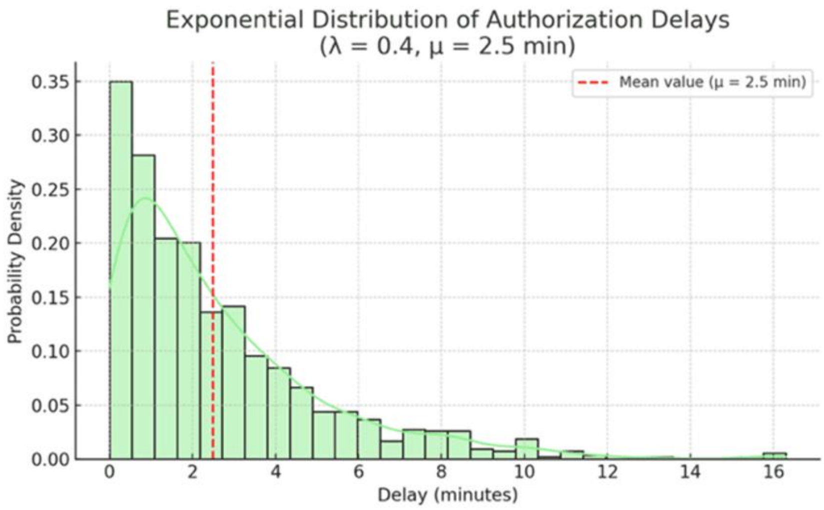


**Figure 2.**
Exponential of Authorization Delays.

In addition:

- More than 50% of all delays are less than two minutes;
- About 10% exceed five minutes;
- A small but non-zero probability is observed even for $\tau > 10$ min.

The exponential distribution allows us to model the unstable sections of the SAB, where delays are rare but significant. This distribution is particularly useful for analyzing the boundary states of the system, the stability of the schedule against failures, and the necessity of buffers between trains.

### *2.4.3. Empirical Distribution*

The empirical distribution is the most realistic way of modeling random variables, based not on theoretical assumptions but on real observed data. In the context of this paper, it is used to approximate the behavior of resolution transmission delays in SAB, taking into account all technical, human, organizational, and random factors operating in a real system [33].

The empirical distribution may be:

- Discrete (if fixed delay values are observed, e.g., 0.5, 1, 2, or 5 minutes);
- Continuous (if the data are smoothed or collected with high precision).

The empirical distribution is based on a discrete set of values that approximate the actual observed data. This approach allows for the consideration of peculiarities in the actual delay distribution, including asymmetry, heavy tails, and clusters. The following discrete set of values was used in the study:

$$\tau_d \in \{0.5, 1, 1.5, 2, 3, 4, 5, 6, 7, 10\}, \tag{10}$$

The values were selected based on an analysis of observations and expert judgment. They reflect the actual structure of delays on SAB sections. The probabilities were manually set, with a maximum delay frequency of 1-3 minutes and a probability of long (>7 minutes) delays of no more than 5%.

This paper applies a discrete approximation to reflect the realities of sites with a limited set of possible situations.

Figure 3 presents a histogram showing the empirical distribution of delays obtained through probabilistic modeling and approximated by observed data.

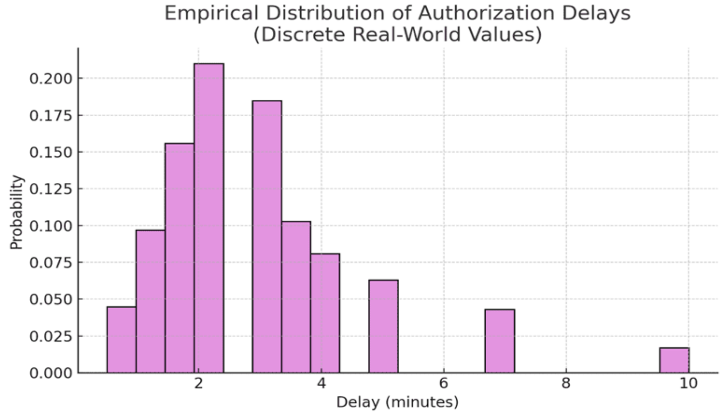


**Figure 3.**
Empirical Distribution of Authorization Delays.

Probability distribution:
- The maximum delay is approximately two to three minutes, which is the most frequently observed delay.
- A significant fraction is in the range of 1-4 minutes;
- A small but non-zero fraction constitutes rare, long delays of 7-10 minutes.

This type of distribution does not require symmetry or smoothness. It also does not require the existence of a mathematical formula. This type of distribution describes the system's actual behavior, which makes it particularly valuable for model validation.

The empirical distribution enables us to estimate both the average delay and its actual variation, the latter of which is particularly important:
- In schedule modeling;
- Timetable reliability assessment;
- Train buffer calculations.

Although empirical simulation often yields lower throughput than normal or exponential distributions, it is more accurate.

An empirical distribution is a key tool for improving the accuracy of a SAB simulation model, particularly when the system's behavior does not lend itself to a simple theoretical description. It reflects local characteristics, considers human and organizational factors, and establishes a foundation for realistic scenario analysis [34]. Table 2 presents a summary table of the distributions' parameters.

**Table 2.**
Summary table of distribution parameters.

| **Distribution type** | **Average value of** ? ? | **Parameters** | **Application characteristic** |
|---|---|---|---|
| Normal | 3.0 min | $\mu = 3, \sigma = 1$ | Steady-state conditions and mean deviations |
| Exponential | 2.5 min | $\lambda = 0.4$ | Unstable communication and long, infrequent delays |
| Empirical | ~ 2.9 min | Discrete set | Approximation to real observations |

These three distributions were chosen because a comparative analysis of models with different behaviors is necessary. These behaviors range from symmetric and narrow (normal) to asymmetric and exponentially damped (exponential), as well as those close to actual dispatch practices (empirical).

### *2.5. Simulation Model*

A simulation model was developed and implemented in the AnyLogic environment using a discrete event approach to evaluate the impact of stochastic delays on the throughput of a section with semi-automatic interlocking. The model's main task is to reproduce the process of a train moving along a single-track crossing. This includes taking into account transit time, waiting time for clearance, and stochastic delay $\tau_d$, which is modeled by one of three distributions: normal, exponential, or empirical (Figure 4).

The model simulates the following sequence of events:
- Release of the crossing by the previous train;
- The delay of the next train departure authorization is a random variable, $\tau_d$.
- Departure of the next train after authorization;
- A section with a fixed time $T_p$ is passed, after which the cycle repeats.

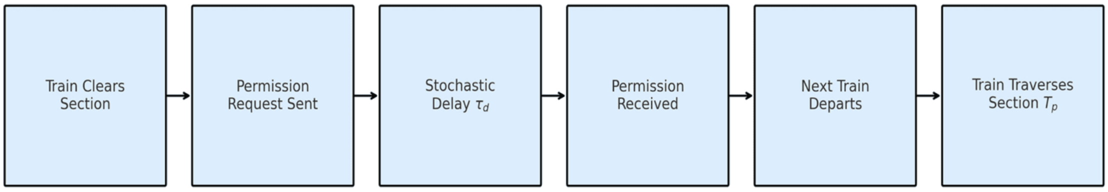


**Figure 4.**
Logic of the Simulation Model Considering Resolution Transfer Delay in SABs.

Thus, each cycle models one inter-train interval, which is calculated as follows: $T_p = (\sum t_x + \tau_a + \tau_b) + \tau_d$ (Figure 5).

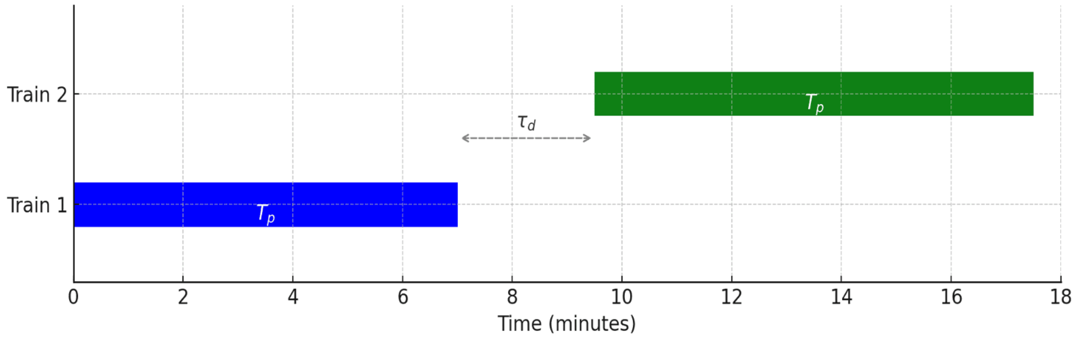


**Figure 5**.
Time diagram showing the delay $\tau_d$ in sending two trains.

The average throughput of the section is calculated based on these data using Equation 2.
The main modeling parameters are presented in Table 3.

**Table 3.**
Modeling parameters.

| Parameter | Significance |
|---|---|
| Passage time | 17 minutes (on average) |
| Lock type | Semi-automatic |
| Number of trains in the experiment | 2,000 trains |
| Time horizon | 30 days (1,440 minutes per day) |
| Repeats of the experiment | 10 runs per allocation |
| Type of modeling | Discrete event approach |

For comparison, simulations were run for each of the three delay distributions:

- Scenario 1: Normal distribution (stable operating environment);
- Scenario 2: Exponential distribution (unpredictable environment);
- Scenario 3: Empirical Distribution (Realistic Dispatching).

The model provides the following in the output:

- Array of train intervals;
- Statistics for each day (number of trains per day);
- Average throughput per scenario;
- Graphs of distributions and densities.

The simulation model was implemented in the AnyLogic software environment using a discrete event approach. Train movement was modeled as a sequence of events: release of the section, transfer of permission with a delay of $\tau_d$, departure of the next train, and passage of the train through the crossing in time $T_p$.

Delays were modeled as random variables with specified distributions (normal, exponential, and empirical). For each simulation, the total number of trains passing through the section in 1,440 model minutes (one day) was calculated and used as an estimate of capacity.

*2.6. Limitations and Assumptions of the Model*

The developed model has a simplified structure that corresponds to the purpose of comparative analysis. However, to expand its applicability, it is necessary to consider a number of limitations and assumptions:

1. Only one single-track crossing is modeled without considering the internal logic of stations, arrival systems, overtaking maneuvers, and technological windows;

2. All trains are assumed to be homogeneous in terms of speed, length, priority, and transit time. In real conditions, differences between passenger, freight, and special trains can significantly impact intervals;

3. The influence of the external environment is not explicitly modeled. Factors such as weather anomalies, communication failures, accidents, and repair work can significantly impact delays. However, they are summarized here as a stochastic component, $\tau_d$;

4. Although the technical time is average, it may vary in practice depending on the gradient, rail condition, braking characteristics of rolling stock, and other factors.

5. Delays $\tau_d$ are independent of previous events, which excludes the possibility of cascading effects typical of congested traffic schedules;

6. Peak and off-peak intervals during the day are not considered; the entire time period is modeled as uniformly congested. In reality, nighttime and daytime intervals may differ.

7. The model does not consider crossing intervals at split points, which are important on single-track sections. In reality, a train may wait for an oncoming train, increasing the actual inter-train interval.

8. The model does not consider intervals of non-simultaneous arrivals, which is especially relevant when there are restrictions on arrival and departure tracks.

These limitations define the bounds of the model's applicability but do not diminish its significance as an analytical tool for assessing the influence of random factors on interval traffic control during SAB. In the future, the model can be adapted for multi-track sections, hybrid interlocking systems, interaction with automated schedules, and consideration of external factors such as weather and accidents. Thus, the presented approach serves as a basis for further research under real operational constraints.

## 3. Results

The simulation model was implemented in the AnyLogic environment using a discrete event approach. The simulation considered a railroad section with semi-automatic interlocking, where trains follow a single-track crossing and receive permission to depart only after the previous train's arrival is confirmed.

The following input data were assumed in the modeling process:

- Technical time of passing the crossing: 17 minutes;
- Total modeling duration: 30 days;
- Modeling horizon: 1,440 minutes per day;
- Number of trains passing through the model: ~2,000 per scenario.
- The intervals between departures were generated based on the following:
- Normal distribution ($\mu = 3, \sigma = 1$);
- Exponential distribution ($\lambda = 0.4$, $\mu = 2.5$);
- Empirical distribution: Discrete set of 0.5-10 minutes based on observations.

The model was run 10 times for each scenario with different initial conditions (i.e., different random number generators).

Then, the average throughput value was calculated, and the statistical stability of the results was assessed.

Fluctuations in the final values did not exceed ±1.5 trains per day, confirming the stability of the model and the accuracy of the implemented approach.

Table 4 presents the calculated daily train throughput for different delay distributions. Additionally, the relationship between the average delay, $\tau_d$, and the final throughput of the section was plotted.

**Table 4.**
Estimated daily train throughput for different delay distributions.

| Delay distribution | Average delay | Standard deviation | Throughput capacity (trains/day) |
|---|---|---|---|
| Normal $\mu = 3, \sigma = 1$ | 3.0 | 1.0 | 64.0 |
| Exponential $\lambda = 0.4$ | 2.5 | 2.5 | 77.0 |
| Empirical (discrete) | 2.85 | 1.8 | 69.0 |

Figure 6 shows the expected decrease in throughput capacity with increasing average delay. This dependence is key to risk assessment when organizing traffic on sections with semi-automatic interlocking.

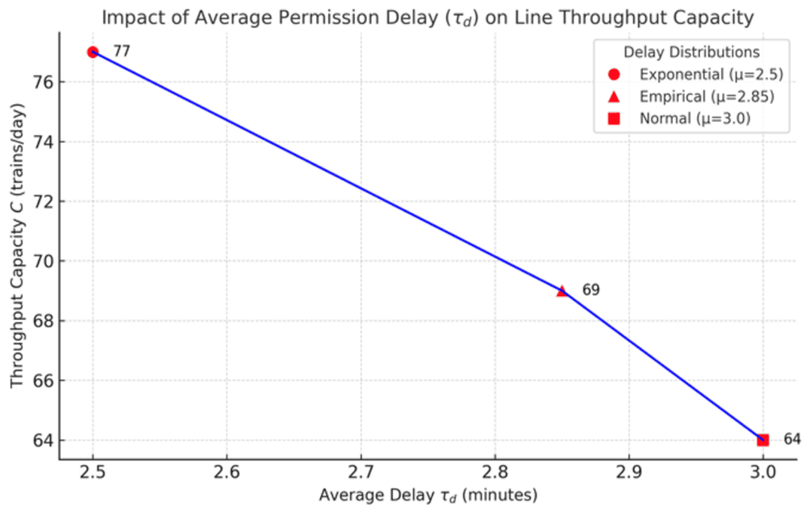


**Figure 6.**
The dependence of throughput on the average delay $\tau_d$ of resolution transmission.

Figure 7 shows a comparison of two probabilistic models of resolution transmission delays: the normal and exponential distributions.

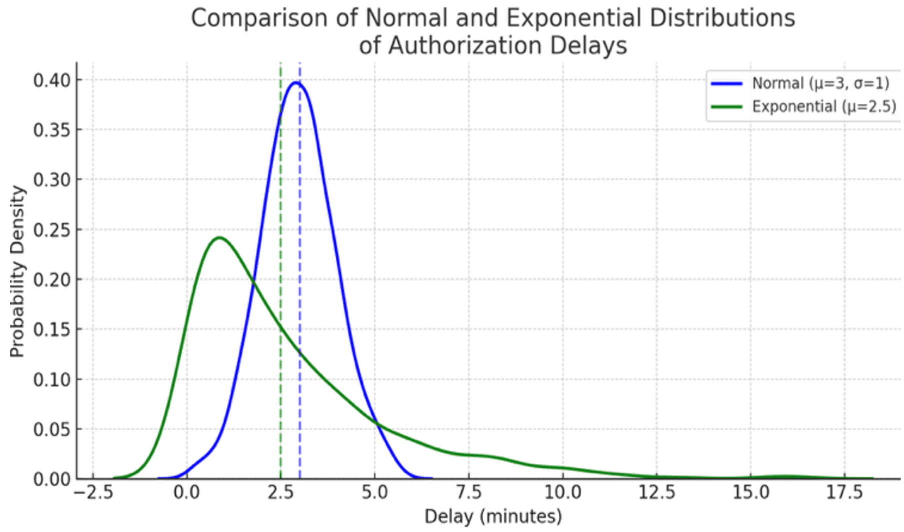


**Figure 7.**
A Comparison of Normal and Exponential Distributions of Authorization Delays

The purpose of the analysis is to demonstrate how differences in statistical properties affect the interpretation of reliability and capacity in sites with semi-automatic interlocking.

A normal distribution suggests that the most likely delay is three minutes, whereas an exponential distribution indicates that most delays are brief, although there is a possibility of rare, large deviations.

A normal distribution has a symmetric «tail» that decreases sharply outside the interval $\mu \pm 3\sigma$.

The exponential distribution has an unbounded, long right tail, which reflects the possibility of rare but significant delays.

If modeling areas where communication failures or operator errors lead to significant delays of 5-10 minutes, the normal model will underestimate them, whereas the exponential model will account for them.

A comparative analysis shows that the choice of delay distribution should correspond to the technological maturity and stability level of the site infrastructure. Using only one model can result in systematic errors when estimating throughput and schedule stability.

Therefore, this paper considers both distributions as alternative SAB operating scenarios, with the possibility of transitioning to an empirical model based on real data.

Figure 8 shows a comparative histogram of the probability densities for the three types of resolution transmission delay distributions.

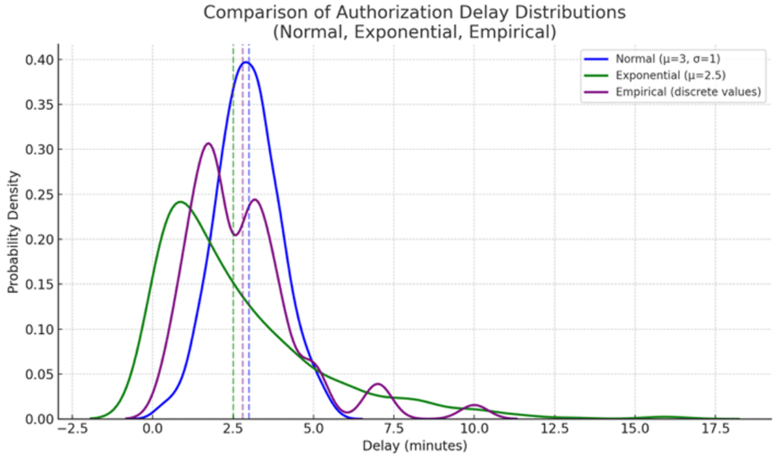


**Figure 8.**
Comparison of Authorization Delay Distributions (Normal, Exponential, Empirical).

In addition:
- The normal distribution is symmetric with a maximum at $\mu = 3$ min;
- Exponential: Asymmetric with a peak at zero and a long right tail;
- Empirical: Reflects real delay frequencies, including sharp jumps of two to three minutes and a small probability of long delays.

The analysis shows that the type of delay distribution has a significant effect on throughput estimation:
- The normal distribution provides the most conservative estimate. It limits outliers and assumes predictable staff performance.
- On the contrary, the exponential distribution reflects unstable operations and allows for rare but critically long delays. However, due to the high density of short delays, the overall throughput is higher.
- The empirical distribution provides an intermediate result that is closer to reality. It takes into account the impact of rare, long delays and the typical behavior of operational staff.

Therefore, relying solely on a theoretical distribution, especially a symmetric one, can result in overestimation or underestimation.

The proposed approach enables flexible customization of the model for specific railroad sections and operating conditions, such as climate, communication technologies, and human factors.

## 4. Discussion

The simulation results obtained in this study clearly demonstrate the significant influence of stochastic delays in authorization transmission on the throughput of railway sections operating under semi-automatic interlocking (SAB) systems. Unlike traditional modeling approaches, which treat such delays as constant or negligible, our model incorporates the randomness of dispatch decisions as a variable dependent on a variety of factors, including human behavior, technical readiness, procedural inconsistencies, and external conditions such as weather. This modeling perspective brings the simulation closer to real-world railway operations, where schedule deviations often arise unpredictably.

The analysis of different delay distributions revealed that the statistical nature of dispatch delays directly affects the minimum allowable interval between trains and, consequently, the achievable daily throughput. The normal distribution, being symmetrical and concentrated around the mean, is representative of stable operating conditions and strict schedule adherence. However, this distribution produced the most conservative estimate of capacity, approximately 64 trains per day, reflecting the cautious nature of planning under stable yet inflexible scenarios.

In contrast, the exponential distribution simulates environments characterized by occasional severe delays but a high probability of very short ones. This led to a maximum modeled capacity of 77 trains per day. Such a result, while optimistic, highlights the operational risk that accompanies uncertainty, as the system becomes vulnerable to delay clustering or service breakdown under stress. The empirical distribution, derived from field data, displayed irregularities such as outliers, skewed tails, and jumps in delay frequency. This model yielded a realistic intermediate capacity of around 69 trains per day, offering the most practical estimate for high-load corridors operating with limited automation and variable human input.

From a logistics perspective, the implications of these findings are substantial. Underestimating the stochastic nature of dispatch delays can result in unrealistic schedules, inefficient allocation of rolling stock, increased operational buffers, and ultimately, reduced reliability of delivery timelines. Conversely, overestimating capacity may lead to schedule congestion, missed connections, and cascading failures throughout the supply chain, particularly in multimodal hubs where precise timing is critical.

The developed modeling approach allows planners and dispatch operators to assess how delay variability influences throughput and to determine an appropriate safety margin in timetables. It supports decisions related to freight train slotting, resource distribution, and infrastructure load balancing, which are crucial for maintaining service levels in dynamic logistics networks. Moreover, the methodology provides a quantitative foundation for evaluating the potential benefits of digitalization and automation, especially in corridors where SAB systems still dominate operations.

Despite its practical value, the model does present several limitations. It does not yet account for geometric complexities such as gradients, curves, or station influence, nor does it simulate interactions with junctions, weather emergencies, or train heterogeneity (e.g., fast vs. slow freight). Additionally, the dispatch delay ($\tau_d$) is modeled as an independent stochastic variable, whereas in reality, it may correlate with cumulative network delays, technical failures, or operational constraints. These simplifications, however, do not detract from the model's utility as a robust starting point for more detailed simulations.

The flexibility of the approach also enables adaptation to other control systems. For example, with appropriate parameter adjustments, the model can be extended to simulate automatic blocking systems, where delays are less stochastic, or radio-based systems (ETCS, CBTC), where communication disruptions may be rare but critical. In the long term, the model may serve as a basis for hybrid simulations that track the gradual transition from semi-automatic to digital systems.

Future improvements could include the integration of real-time statistical dispatch data, incorporation of node and multi-track interactions, modeling of seasonally dependent disruptions, and creation of digital twins of key rail corridors. Furthermore, an optimization module could be developed to evaluate trade-offs between capacity, schedule resilience, failure sensitivity, and alternative dispatching policies. Such a tool would be of particular value to logistics managers, infrastructure planners, and policymakers aiming to modernize rail operations while minimizing risk and maximizing efficiency.

Several previous studies have addressed railway capacity modeling using deterministic or semi-deterministic approaches. For instance, Hansen and Pachl [35] proposed analytical methods for capacity assessment based on fixed block systems and standardized intervals, but did not account for dispatch variability or stochastic behavior of control systems [35]. Similarly, Artan and Şahin [36] examined delay propagation using simulation tools, yet their focus remained on high-speed and passenger-oriented systems with advanced automation, limiting applicability to freight-oriented semi-automatic networks [36].

By contrast, our study uniquely contributes to the literature by focusing on semi-automated lines with manual authorization systems, a configuration still prevalent in many post-Soviet and Asian countries. Moreover, the inclusion of empirical delay distributions derived from field observations allows for a more realistic assessment of operational uncertainty compared to models based solely on assumed distributions (e.g., normal or Poisson). While some recent works, such as those by Yin et al. [37] have used stochastic modeling for urban rail or metro systems [37] their methodologies are not readily transferable to freight corridors with long inter-train distances and different traffic dynamics [38, 39]. Therefore, our approach fills a critical methodological gap and offers practical value for freight logistics planning in partially modernized railway systems.

## 5. Conclusions

This study develops and implements a stochastic model for estimating the throughput capacity of a railroad section under semi-automatic blocking, accounting for delays in permission-to-depart transmission as a random variable. Unlike traditional deterministic approaches, the proposed methodology uses a probabilistic description of this key element of the train movement process. The delay was modeled using three distributions: normal, exponential, and empirical, which allowed for a comprehensive analysis of the impact of fluctuations on train intervals and, consequently, on the section's throughput capacity.

The results clearly demonstrate that overlooking stochastic elements in dispatch procedures may lead to substantial miscalculations in capacity planning, particularly in high-demand corridors where infrastructure is constrained. Depending on the assumed distribution, the modeled throughput ranged from 64 to 77 trains per day. The empirical distribution, derived from observed field data, yielded the most realistic estimate (approximately 69 trains), capturing the irregularities and asymmetries inherent in semi-automated operations. These findings underscore the need for more data-driven and uncertainty-aware approaches to timetable development, resource allocation, and dispatch strategy in freight logistics.

The proposed methodology offers practical value to railway operators, logistics planners, and infrastructure managers, especially in regions where legacy dispatch systems are still prevalent. By adjusting the distributional parameters or eliminating delay variability altogether, the model can be extended to simulate automatic and radio-based interlocking systems. Furthermore, the ability to model stochastic delays supports more accurate assessments of schedule resilience, operational risk, and infrastructure utilization under various disturbance scenarios.

Further research should aim to expand the model by integrating real statistical data, analyzing the combined effects of weather and technical factors, and introducing multi-track and hub scenarios. The development of digital twins for SAB-managed corridors, real-time virtual models capable of dynamic monitoring and forecasting, represents a promising direction for both research and practical deployment. Collectively, these efforts provide a scientific and methodological basis for improving the robustness, flexibility, and efficiency of railway logistics under constrained modernization conditions.

## References

[1] Y. Wei, L. Ma, and J. Yin, "Integrated capacity analysis model and optimization based on conflict resolution for high-speed railway," in *2022 China Automation Congress (CAC)*, 2022, pp. 2710-2715. https://doi.org/10.1109/cac57257.2022.10055147

[2] Y. Pu, S. Srikukenthiran, E. Morrow, A. Shalaby, and W. Klumpenhouwer, "Capacity analysis of a passenger rail hub using integrated railway and pedestrian simulation," *Urban Rail Transit*, vol. 8, no. 1, pp. 1-15, 2022. https://doi.org/10.1007/s40864-021-00162-7

[3] W. Zhou, S. Li, J. Kang, and Y. Huang, "Capacity-oriented train scheduling of high-speed railway considering the operation and maintenance of rolling stock," *Mathematics*, vol. 10, no. 10, p. 1639, 2022. https://doi.org/10.3390/math10101639

[4] V. Ľupták, M. Dedík, P. Morihladko, P. Šulko, and L. Pečený, "Assessment of the railway line capacity on the railway network using a new innovative method," *Sustainability*, vol. 17, no. 10, p. 4476, 2025. https://doi.org/10.3390/su17104476
[5] K. Tánczos, "Adapting the ITF's transport policy with a focus on the EU's strategy of sustainable mobility," *Periodica Polytechnica Transportation Engineering*, vol. 53, no. 2, pp. 222-225, 2025. https://doi.org/10.3311/pptr.39780
[6] R. Wang, L. Nie, and Y. Tan, "Evaluating line capacity with an analytical UIC code 406 compression method and blocking time stairway," *Energies*, vol. 13, no. 7, p. 1853, 2020. https://doi.org/10.3390/en13071853
[7] A. Baturin and A. Osipov, "Modern systems of interval control of train traffic of railways of the world," *Peace of Transport*, vol. 21, no. 2, pp. 18-26, 2023.
[8] R. Aliev, "Model coordinate system of interval regulation train traffic," presented at the International Conference on Computational Techniques and Applications, 2021.
[9] Q. Wu, X. Ge, Q.-L. Han, and Y. Liu, "Railway virtual coupling: A survey of emerging control techniques," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 5, pp. 3239-3255, 2023. https://doi.org/10.1109/tiv.2023.3260851
[10] D. Basile, M. H. ter Beek, A. Ferrari, and A. Legay, "Exploring the ERTMS/ETCS full moving block specification: An experience with formal methods," *International Journal on Software Tools for Technology Transfer*, vol. 24, no. 3, pp. 351-370, 2022. https://doi.org/10.1007/s10009-022-00653-3
[11] Y. He, J. Lv, and T. Tang, "Communication-based train control with dynamic headway based on trajectory prediction," *Actuators*, vol. 11, no. 8, p. 237, 2022. https://doi.org/10.3390/act11080237
[12] G. Wojnar and M. Irlik, "Multicriteria train running model and simulator for railway capacity assessment," *Transport Problems*, vol. 17, pp. 199–212, 2022. https://doi.org/10.20858/tp.2022.17.2.17
[13] P. Folęga, M. Irlik, A. Kochan, and G. Wojnar, "Railway line capacity relative to additional block division," *Promet-Traffic&Transportation*, vol. 36, no. 1, pp. 42-54, 2024. https://doi.org/10.7307/ptt.v36i1.360
[14] J. Široký, P. Nachtigall, E. Tischer, and J. Gašparík, "Simulation of railway lines with a simplified interlocking system," *Sustainability*, vol. 13, no. 3, p. 1394, 2021. https://doi.org/10.3390/su13031394
[15] V. Ranjbar, N. O. Olsson, and H. Sipilä, "Impact of signalling system on capacity–Comparing legacy ATC, ETCS level 2 and ETCS hybrid level 3 systems," *Journal of Rail Transport Planning & Management*, vol. 23, p. 100322, 2022. https://doi.org/10.1016/j.jrtpm.2022.100322
[16] W. Wójcik, M. Orunbekov, A. Toygozhinova, and A. Seitbekova, "Difficulties in TETRA operation with moving block in Kazakhstan," *Przeglad Elektrotechniczny*, vol. 96, no. 9, pp. 129-132, 2020. https://doi.org/10.15199/48.2020.09.27
[17] A. Saitov, J. Kurbanov, Z. Toshboyev, and S. Boltayev, "Improvement of control devices for road sections of railway automation and telemechanics," *E3S Web of Conferences*, vol. 264, p. 05031, 2021. https://doi.org/10.1051/e3sconf/202126405031
[18] Z. Łukasik, T. Ciszewski, J. Młyńczak, W. Nowakowski, and J. Wojciechowski, "Assessment of the safety of microprocessor-based semi-automatic block signalling system," in *Contemporary Challenges of Transport Systems and Traffic Engineering: 13th Scientific and Technical Conference" Transport Systems. Theory and Practice 2016" Katowice, Poland, September 19-21, 2016 Selected Papers*, 2016. https://doi.org/10.1007/978-3-319-43985-3_12
[19] M. Almammadova and N. Kamenieva, "Microprocessor system control stage: Railway tracks example," *E3S Web of Conferences*, vol. 508, p. 03001, 2024. https://doi.org/10.1051/e3sconf/202450803001
[20] M. Orunbekov, V. Shults, G. Suleimenova, S. Sarbayev, and A. Seitbekova, "Development of an imitation model of container transportation by railway area of dostyk-zhezkazghan-iletsk," *Journal of Applied Engineering Science*, vol. 19, no. 3, pp. 676-680, 2021. https://doi.org/10.5937/jaes0-29665
[21] A. Zhumanov, Z. Kegenbekov, and J. Tolujevs, "Trans-caspian international transport route infrastructure assessment using simulation modelling," *Transport and Telecommunication*, vol. 25, no. 1, pp. 11-19, 2024. https://doi.org/10.2478/ttj-2024-0002
[22] J. Wu *et al.*, "The bounds of improvements toward real-time forecast of multi-scenario train delays," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 3, pp. 2445-2456, 2021. https://doi.org/10.1109/tits.2021.3099031
[23] A. Janota, R. Pirník, J. Ždánsky, and P. Nagy, "Human factor analysis of the railway traffic operators," *Machines*, vol. 10, no. 9, p. 820, 2022. https://doi.org/10.3390/machines10090820
[24] A. Nikitin, A. Manakov, I. Kushpil, A. Kostrominov, and A. Osminin, "On the issue of using digital radio communications of the DMR standard to control the train traffic on Russian railways," in *2020 IEEE East-West Design & Test Symposium (EWDTS)*, 2020. https://doi.org/10.1109/ewdts50664.2020.9224707
[25] P. E. Afolabi, S. O. Ikubanni, B. O. Adebesin, S. J. Adebiyi, and E. F. Adeleke, "Effects of weather parameters on signal attenuation: A brief review," in *2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG)*, 2024. https://doi.org/10.1109/seb4sdg60871.2024.10629947
[26] G. Mukunzi and C.-W. Palmqvist, "The impact of railway incidents on train delays: A case of the swedish railway network," *Journal of Rail Transport Planning & Management*, vol. 30, p. 100445, 2024. https://doi.org/10.1016/j.jrtpm.2024.100445
[27] J. Lessan, L. Fu, C. Wen, P. Huang, and C. Jiang, "Stochastic model of train running time and arrival delay: A case study of Wuhan–Guangzhou high-speed rail," *Transportation Research Record*, vol. 2672, no. 10, pp. 215-223, 2018. https://doi.org/10.1177/0361198118780830
[28] M. M. Dekker, A. N. Medvedev, J. Rombouts, G. Siudem, and L. Tupikina, "Modelling railway delay propagation as diffusion-like spreading," *EPJ Data Science*, vol. 11, no. 1, p. 44, 2022. https://doi.org/10.1140/epjds/s13688-022-00359-1
[29] E. Tischer, P. Nachtigall, and J. Široký, "The use of simulation modelling for determining the capacity of railway lines in the Czech conditions," *Open Engineering*, vol. 10, no. 1, pp. 224-231, 2020. https://doi.org/10.1515/eng-2020-0026
[30] F. Corman, A. Trivella, and M. Keyvan-Ekbatani, "Stochastic process in railway traffic flow: Models, methods and implications," *Transportation Research Part C: Emerging Technologies*, vol. 128, p. 103167, 2021. https://doi.org/10.1016/j.trc.2021.103167
[31] S. Dong, Z. Tang, M. Wu, and J. Zhang, "Stochastic dynamic simulation of railway vehicles collision using data-driven modelling approach," *Railway Engineering Science*, vol. 30, no. 4, pp. 512-531, 2022. https://doi.org/10.1007/s40534-022-00273-2
[32] Y. Yang, P. Huang, Q. Peng, J. Li, and C. Wen, "Statistical delay distribution analysis on high-speed railway trains," *Journal of Modern Transportation*, vol. 27, no. 3, pp. 188-197, 2019. https://doi.org/10.1007/s40534-019-0188-z
[33] M. M. Dekker, "Geographic delay characterization of railway systems," *Scientific Reports*, vol. 11, no. 1, p. 20860, 2021. https://doi.org/10.1038/s41598-021-00361-z

[34] T. Spanninger, B. Büchel, and F. Corman, "Probabilistic predictions of train delay evolution," in *2021 7th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, 2021. https://doi.org/10.1109/mt-its49943.2021.9529333

[35] I. A. Hansen and J. Pachl, *Railway timetabling & operations: analysis, modelling, optimisation, simulation, performance evaluation, 2nd revised and extended edition*. Hamburg: DVV Media Group GmbH, Eurailpress, 2014.

[36] M. Ș. Artan and İ. Șahin, "A stochastic model for reliability analysis of periodic train timetables," *Transportmetrica B: Transport Dynamics,* vol. 11, no. 1, pp. 572-589, 2023. https://doi.org/10.1080/21680566.2022.2103051

[37] X. Yin, J. Chen, and Y. Li, "Simulation-based resilience evaluation for urban rail transit transfer stations," *Sustainability,* vol. 16, no. 9, p. 3790, 2024. https://doi.org/10.3390/su16093790

[38] A. D'Ariano, L. Meng, G. Centulio, and F. Corman, "Integrated stochastic optimization approaches for tactical scheduling of trains and railway infrastructure maintenance," *Computers & Industrial Engineering,* vol. 127, pp. 1315-1335, 2019. https://doi.org/10.1016/j.cie.2017.12.010

[39] H. Zhang, G. Lu, Y. Lei, G. Zhang, and I. Niyitanga, "A hybrid framework for synchronized passenger and train traffic simulation in an urban rail transit network," *International Journal of Rail Transportation*, vol. 11, no. 6, pp. 912-941, 2023. https://doi.org/10.1080/23248378.2022.2109522