Predicting for train movement based on historical statistical data

A Domojirova, R Upyr, M Eremenko

Irkutsk State Transport University, 664074, RF, Irkutsk, Chernyshevskogo str., bld. 15
E-mail: domojirova.a@gmail.com

Abstract. Prediction of train movement is required for traffic control. The readiness of systems to process incoming train flow depends on the quality of planning the situation on railroad. But predicting is influenced by various factors, and that is why the output parameters of train flows do not correspond to the planned parameters. The article considers the reliability factor of technical means of infrastructure, and compares its effect on train flows with various technical equipment of railway sections: double-track mountain-pass, single-track, double-track. As a result, the most unstable to failures of technical means section (in comparison with others) was determined and the reliability of its technical equipment was studied in more detail.

1. Introduction

Prediction of train movement is the most important part of traffic control. Running trains on time, trains departure and delivery of goods to the consignee mainly depends on the quality of planning [1].

The planning scope of train traffic can vary, starting from the assistant station master’s planning for several trains, and ending with the train dispatcher’s planning of train crossings for the whole section. However, the planned train pattern is not always reflected in the executed train schedule. Various factors can interfere: human (incorrect decisions of operational and dispatching personnel, non-compliance with safety culture), natural and technogenic factors, technical factors (reliability of technical means).

In this article, the authors consider the reliability factor with different technical equipment of the infrastructure and its impact on predicting the train traffic movement.

2. Materials and methods

To determine the section with the most predicted train position, two types of analysis were used: regression and correlation.

Regression analysis allows you to determine the nature of the relationship between two variables: the number of failures and the number of delayed trains. The relationship between these variables is characterized by a mathematical model called the regression equation. This regression method was chosen because observations of uncontrolled events were performed [2].

Correlation analysis determines the degree of relationship between two or more random variables. Correlation analysis consists in calculating the correlation coefficients; in the study, for this purpose, the Pearson correlation coefficient formula was used (1):

$$r_{xy} = \frac{\sum (x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum (x_i-\bar{x})^2 \sum (y_i-\bar{y})^2}}$$

As a material for the study, we used statistical data on the number of failures of technical means and delayed trains at three sections with different technical equipment: double-track mountain-pass, single-track, double-track.
3. Results and discussion

3.1. Predicting on sections with different technical equipment

An analysis was conducted, considering the impact of failures on the number of delayed trains with different technical equipment: double-track mountain-pass, single-track and double-track sections.

According to the analysis, the mountain-pass section has the greatest dependence of the delayed trains on the number of failures, and, consequently, the dependence of train traffic on the reliability of technical means. This is proved by the highest correlation coefficient – 0.9859 (Figure 1).

\[ y = 20.662x - 73,252 \]

**Fig. 1.** Prediction of the train traffic on the mountain-pass section

This can be explained by the complexity in planning due to the presence of banking locomotives among freight and passenger trains, which ensure the traffic management in conditions of a complex profile. In such conditions, significant consequences arise due to an error or failure of a planned action. For example, a malfunction of the track that the banking locomotive was supposed to follow makes the timely departure of the train waiting to be pushed [3].

A single-track section has a lesser, but also high dependence on the technical equipment reliability with a correlation coefficient - 0.9339. The main problem of predicting on a single-track section is in the planning of crossing and the failure of crossing. These failures make it necessary to change the planning in accordance with the prevailing situations.
Fig. 2. Prediction of the train traffic on the single-track section.

Predicting on a low-intensity double-track section is the most stable compared to the above-mentioned sections. The correlation index – 0.7727. Failure on this section causes fewer changes for the train traffic, compared with mountain-pass and single-track sections.

Fig. 3. Prediction of the train traffic on the double-track section.

The scatter of data in the latter case on the graph (Figure 3) can be more strictly described by nonlinear regression. For this, the data for 5 years were processed and the dependence graph was rebuilt (Figure 4).
Thus, the polynomial function of the sixth order (2) most accurately describes the dependence on the section under consideration:

\[ y = 0.0001x^6 - 0.0316x^5 + 2.8399x^4 - 133.25x^3 + 3441.9x^2 - 46421x + 255669 \] \hspace{1cm} (2)

The correlation index was also recalculated and amounted to 0.8322.

3.2. Reliability of technical means on the mountain-pass section

The reliability of the technical equipment on the mountain-pass section was studied in more detail, since this section was determined as the most unstable to failures. A railway section Bolshoi Lug – Slyudyanka is one of the main sections of the Eastern polygon that limit traffic capacity. The maximum ups and downs of the section (18 %), as well as the minimum radius within curves (less than 300 meters) make it impossible to handle trains of the required dimensions. Currently, Bolshoi Lug – Slyudyanka-I section handles up to 96 pairs of freight trains. The requirements for banking locomotives make 35 units. By the year 2023, it is planned to increase the traffic capacity on the section to 137 pairs of trains per day, including 107 pairs of freight trains per day and 19 down banking locomotives straight joints. The carrying capacity of the section is going to increase to 116.7 million tons per year [4].

A huge variety of organizational, technical and technological measures are proposed to increase the capacity of the section. However, the condition of the existing technical equipment and its operability should be estimated before taking the most effective measures to increase the traffic capacity. In this regard, it is necessary to analyze the reliability of the technical means operation, since the traffic capacity and the safety of train traffic largely depends on the operation of technical equipment.

3.3. Failures of technical means on the mountain pass section

The problem of the failure occurrence is still relevant, since the duration of the operable state of technological means nowadays does not decrease in comparison with previous years. This is shown in Figure 4.

**Fig. 4.** Approximation of the functional dependence of the number of delayed trains on the number of failures on a single-track section
According to [5], reliability is a capability of the object to keep in time within the established limits the values of all parameters characterizing the ability to perform the required functions in the specified conditions of use, maintenance, repairs, storage and transportation. However, the object can be prevented from preserving these initial properties by a transition to an inoperative state or a limiting state – a failure. Failure in the technical means operation is a violation of the operational state of an object.

Failures are divided:
1) Category 1 – delaying passenger, suburban or freight trains for 1 hour or more, or leading to events;
2) Category 2 – delaying passenger trains from 6 minutes, freight trains from 15 minutes;
3) Category 3 – taking into account the involved services, due to which the failure and delays of passenger trains up to 6 minutes occurred, freight trains up to 15 minutes (the KASANT system – a comprehensive automated accounting system, monitoring the elimination of failures and analyzing their reliability – does not take into account).

An analysis was conducted of the failure rate of categories 1 and 2 for 12 months: from November’18 to October’19. March, July, August, September 2019 were marked with the largest number of failures (Table 1).

| Month               | Failures | Delayed trains |
|---------------------|----------|----------------|
| November, 2018      | 13       | 158            |
| December, 2018      | 11       | 125            |
| January, 2019       | 11       | 118            |
| February, 2019      | 11       | 142            |
| March, 2019         | 18       | 171            |
| April, 2019         | 13       | 162            |
| May, 2019           | 7        | 184            |
| June, 2019          | 10       | 217            |
| July, 2019          | 18       | 885            |
| August, 2019        | 16       | 403            |
| September, 2019     | 16       | 415            |
| October, 2019       | 9        | 135            |
| **Total**           | 153      | **3115**       |

**Table 1. Failure rate**
The most common faults are malfunctions of insulating joints, flexible lock of the contact system; the arc-extinguishing chamber of the vacuum circuit breaker, pantograph collector frame, the armature of the DC traction motor, the power contacts of the electromagnetic contactor of the power circuit; brake conduit of freight car, freight car reservoirs, axlebox mount of freight car, air distributor of freight car; lens traffic lights, relay of the automatic block system.

A large number of failures occur on stages that have complex profiles and are directly included in the Bolshoi Lug – Slyudyanka mountain-pass section.

**Fig. 6.** Distribution of failures by stages of the mountain-pass section over 5 years (2014-2015).

4. Conclusion

Reliability of technical means in the transportation process organization has an extremely high importance. The reliability of equipping railway sections and stations directly affects the service speed, traffic capacity, delivery time of goods and the safety of train traffic [6]. The influence of reliability on the quality of the transportation process increases significantly in the most difficult from the operation perspective sections. This is the mountain-pass section Bolshoi Lug – Slyudyanka and the single-track section – Ust-Ilimsk – Korshunikh-Angarskaya. The main factor limiting the quality and stability of predicting is the section profile for the mountain-pass section and one main track for the single-track section. Therefore, failure on these sections leads to more undesirable consequences in comparison with the sections having more favorable profiles and the second and third main tracks.

References

[1] Upyr R, Domojirova A 2020 *About railway sustainability on the train diagram* Available at SSRN: https://ssrn.com/abstract=3556382.
[2] Morgunov A, Revina I 2014 *Planning and analysis of experimental results* (Omsk: Publishing house of OmSTU)
[3] Domojirova A, Goncharova N, Upyr R 2020 The research of the reliability of the operational work of the railway line in the mountain pass section *TRANSPORT: science, equipment, management* (Scientific information collection) 6 24-28
[4] Olentsevich V, Vlasova N, Astashkov N 2019 Organizational and technical measures aimed at increasing the traffic capacity of the railway section, based on the analysis of patterns and factors of the integrated development of its material and technical base Scientific Cooperation Center "Interactive plus" 5
[5] Sapožnikov VL, Sapožnikov V, Shamanov V 2003 Reliability of railway automation, telemechanics and communication systems (Moscow: Marshrut)
[6] Nefedov A 1997 Reserves of resource-saving rail metal and train traffic safety Dissertation Abstract of Ph. D thesis. (Moscow)