1. Introduction

Shunting operations are an important component of the transportation process on railway transport. About 70% of the time of its circulation, a railroad carriage spends at the stations. Over the time of circulation, a railroad carriage several times undergoes processing. Shunting operations make up to 25% of all conveyance costs [1].

Shunting operations consist of separate operations. They include transposition maneuvers, which imply fulfillment of the following operations, include: the transposition of trains and railroad carriages from a park into another park, the transposition of railroad carriages from one track to another one, delivery (clearing) of railroad carriages to the cargo points and driveways, etc. [2].

The rational organization of shunting operations to a large degree determines the performance efficiency of station and use of shunting means. In connection with this, the improvement of the methods of planning and rate setting of shunting operations with the help of mathematical modeling is a relevant task in the area of organization and control of the transportation process on rail transport.

2. Literature review and problem statement

Setting the standards of duration of shunting operations consists in determining the time period necessary for fulfilling separate shunting operations. This is necessary for the correct organization of shunting operations, planning, evaluation and remuneration for the work of shunting brigades. The base of setting the standards is formed by separating the shunting operations into the simplest elements – semi-flights and operations. Thus, the task comes down to determining the duration of these elements [3].

The relevancy of this issue is not any less at present. It is due to a wide acceptance in the area of studies of rail transport of the method of imitation simulation and modern software for its realization in practice, in particular Anylogic [4] and GPSS [5], which also requires time measurements of the elements of shunting operations. Thus, time measurement is necessary for determining the optimum quantity of locomotives [6]. The operating costs increase with the low usage coefficient of traction rolling stock; the amount of pay for the use of railroad carriages may substantially grow with their use to full capacity because of their additional idle
time when expecting the cargo transfer. The discreteness of tabular norms decreases the accuracy of calculation of the planned indices of idle time of railroad carriages in the stations, which requires improvement in the procedure of their calculation. When determining the capacity of railway stations, planning of the execution of shunting operations is also a main process [7]. As a result of the use of approximate calculations, the problem of the non-fulfillment of technological schedules of servicing the cargo points may occur, as well as of disruptions of technological processes at industrial stations. Of special complexity is the task of determining the planned indicators of work, in particular, of the time period that railroad carriages spend at the stations [8].

For standardization of the time of fulfillment of shunting operations in the station, it is efficient to create special software complexes [9], which provide information support of dispatcher in the process of decision making about operational control of the transportation process.

For the purpose of determining the norms of time when fulfilling shunting operations with trains and railroad carriages on the main and industrial rail transport, different analytical and empirical models are used [10, 11].

Specific time norms for different ranges included in these models of variables are identified in the tabular form. This representation of the time norms causes certain inconveniences in the practical solution of operational problems. For the big systems, the use of tabular norms may lead to the accumulation of errors in the calculations, which will lead to the mismatch between the calculated and actual indicators. Therefore, it is more expedient to obtain, based on those developed in [10] standards of time for separate shunting operations, a polynomial model of duration of the given shunting operation, which explicitly reflects the effect on the magnitude of technological time not only by separate variables (factors), but also their interaction.

3. The aim and tasks of the study

The purpose of this work is the generalization of discrete data on the duration of fulfillment of shunting operations at the railway stations.

To achieve the goal, the following tasks were set:

- analysis of existing approaches to setting the standards of time for semi-flights;
- conducting a full factor experiment and simulation of duration of shunting operations;
- performing an analysis of the obtained models and testing their adequacy.

4. Full factor experiment and simulation of duration of fulfillment of the shunting operations

For obtaining the new (modified) models, we used two well-known methods of setting the standards for semi-flights based on the time measurement observation.

Method 1. Setting the standard by the simplified linear dependency [1]:

\[ t_{pr} = a + bm_{vmst} \]  

(1)

where \( a \) is the part of time of a semi-flight, connected with the movement of locomotive to assigned distance; \( b \) is the part of time of a semi-flight, connected with moving the railroad carriages up to assigned distance; \( m_{vmst} \) is the number of railroad carriages in the shunting train.

The norms of time \( a \) and \( b \), depending on the length of a semi-flight, are determined according to special tables [10].

The modified model of a semi-flight is built based on the superposition of the parameters \( a \) and \( b \). Using tabulated data [10], with the aid of OLS we obtained single-factor regression dependencies of the specified parameters on the length of a semi-flight:

\[ a=0.595+0.0015L_{pr}, \]  

(2)

- when the brakes are on

\[ b=0.0015L_{pr}^{0.475}, \]  

(3)

where \( L_{pr} \) is the length of a semi-flight, m;

- when the brakes are off

\[ b=0.0017L_{pr}^{0.541}. \]  

(4)

By substituting \( a \) and \( b \) in (1), we obtain:

- when the brakes are on

\[ t_{pr} = 0.595 + 0.0015L_{pr} + 0.0015L_{pr}^{0.475}m_{vmst}. \]  

(5)

- when the brakes are off

\[ t_{pr} = 0.595 + 0.0015L_{pr} + 0.0017L_{pr}^{0.541}m_{vmst}. \]  

(6)

Method 2. Three-parameter formula is used for setting the standard [2]

\[ t_{pr} = \frac{(\alpha_m + \beta_m m_{vmst})v_{ms} + 0.06L_{pr}}{v_{ms}}, \]  

(7)

where \( \alpha_m \) is the coefficient that considers the time connected with the change in the speed of motion of a locomotive by 1 km/h at acceleration and braking, \( \beta_m = 0.0407 \text{ min/km/h} \); \( \beta_m \) is the coefficient that considers additional time for changing the speed of motion of each railway carriage in the shunting train set by 1 km/h at acceleration and braking \( (\beta_m = 0.0017 \text{ min/km/h}) \); \( v_{ms} \) is the permissible speed of motion at shunting, km/h.

The problem is to obtain a mathematical model in the form of functional dependency

\[ t_{pr} = f(L_{pr}, m_{vmst}, v_{ms}), \]  

(8)

where \( t_{pr} \) is the technological time of a semi-flight, min.

For establishing this dependency explicitly, that is, to obtain equation of regression, we use a full factor experiment (FFE) of the 2^3 type with effects of the binary interaction [12, 13].

The levels and intervals of variation of variables are given in Table 1.

The coded factors are connected with the natural following dependencies:

\[ \bar{X}_i = \frac{L_{pr} - 5005}{4995}; \]  

(9)
In the experimental data are obtained based on the results of three parallel experiments:

\[
\bar{x}_a = \frac{m_{\text{min}} - 41}{40},
\]

\[
\bar{x}_b = \frac{v_{\text{min}} - 27.5}{12.5}.
\]

(10)

– we determine the calculated value of the Cochran’s criterion:

\[
G_r = \frac{S_{\text{supr}}^2}{\sum_{i=1}^{N} S_i^2}.
\]

(13)

\[
G_r = \frac{1,868}{3,764} = 0.496;
\]

– we determine the numbers of degrees of freedom: \(f_1=m-1=3-1=2; f_2=N-8;\)

– we select the level of significance \(\alpha=0.05;\)

– according to special table [9, 10] by the set \(\alpha, f_1\) and \(f_2\) we find critical value \(G_\alpha=0.5757;\)

– since \(G_r=0.496<G_\alpha=0.5757\), then the dispersions are uniform, and the experiments are reproducible.

**Table 1**

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>Levels of variation</th>
<th>Intervals of variation (\Delta)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of a semi-flight (l_{\text{pr}},) m</td>
<td>Low –1</td>
<td>Basic 0</td>
</tr>
<tr>
<td>Number of carriages in the shunting train set (n_{\text{trans}},) units</td>
<td>1</td>
<td>41</td>
</tr>
<tr>
<td>Speed of motion of the shunting train set (v_{\text{trans}},) km/h</td>
<td>15</td>
<td>27.5</td>
</tr>
</tbody>
</table>

Based on the obtained dependencies, we carry out the development of the mathematical model for the calculation of technological time of a semi-flight.

**5. Formalized description and development of a mathematical model with coded variables for the calculation of technological time of a semi-flight of the shunting operation**

For the formalized description of technological time of a semi-flight, we used the mathematical model with coded variables in the form of equation

\[
y = \hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3.
\]

(12)

The mean values of technological time (function of response) \(\bar{y}\) in the experimental data are obtained based on tabular norms [10] according to the results of three \((m=3)\) parallel experiments. The matrix of planning with the results of experiment and intermediate calculations are represented in Tables 2, 3.

**Table 2**

<table>
<thead>
<tr>
<th>No. of entry</th>
<th>(x_0)</th>
<th>(x_1)</th>
<th>(x_2)</th>
<th>(x_3)</th>
<th>(x_1x_3)</th>
<th>(x_2x_3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>5</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>7</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>8</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Construction of mathematical model according to the data from Table 2 and Table 3 is performed in the following order.

1. Test of reproducibility of experiments according to the results of three parallel experiments:

– we find the maximum value \(S_r^2 = 1,868\) in Table 2;

2. Determining the dispersion of reproducibility of the experiments:

\[
S_{\text{supr}}^2 = S_r^2 \bar{y} = \frac{1}{N} \sum_{i=1}^{N} S_i^2,
\]

(14)

\[
S_{\text{supr}}^2 = \frac{1}{8} \cdot 3.764 = 0.47.
\]

3. Calculation of coefficients of model (12) according to formula:

\[
b_i = \frac{1}{Nm} \sum_{i=1}^{N} x_i \bar{y},
\]

(15)

We obtain \(b_0=15.59; b_1=14.09; b_2=-1.78; b_3=-5.19;\) \(b_{12}=-0.68; b_{13}=-6.45; b_{23}=-0.013;\)

Testing the significance of regression coefficients:

– we find the variance of error of determining the coefficient

\[
S_{b_i}^2 = \frac{1}{N m} S_{\text{supr}}^2,
\]

(16)

\[
S_{b_i}^2 = \frac{1}{8 \cdot 3} \cdot 0.47 = 0.019;
\]

– we find the number of degrees of freedom of the dispersion of reproducibility

\[
f_{\text{res}} = N(m-1) = 8 \cdot (3-1) = 16;
\]

(17)

**Table 3**

<table>
<thead>
<tr>
<th>No. of entry</th>
<th>Output parameter (technological time), min</th>
<th>(y)</th>
<th>(y^2)</th>
<th>((y - \bar{y})^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.24</td>
<td>0.22</td>
<td>0.22</td>
<td>0.0002</td>
</tr>
<tr>
<td>2</td>
<td>39.50</td>
<td>39.00</td>
<td>39.89</td>
<td>0.1160</td>
</tr>
<tr>
<td>3</td>
<td>1.42</td>
<td>1.40</td>
<td>1.40</td>
<td>0.0002</td>
</tr>
<tr>
<td>4</td>
<td>14.88</td>
<td>14.89</td>
<td>14.89</td>
<td>1.4100</td>
</tr>
<tr>
<td>5</td>
<td>0.57</td>
<td>0.53</td>
<td>0.53</td>
<td>0.0012</td>
</tr>
<tr>
<td>6</td>
<td>15.01</td>
<td>15.47</td>
<td>15.47</td>
<td>0.3630</td>
</tr>
<tr>
<td>7</td>
<td>3.71</td>
<td>3.65</td>
<td>3.83</td>
<td>0.0050</td>
</tr>
<tr>
<td>8</td>
<td>18.15</td>
<td>19.36</td>
<td>19.33</td>
<td>1.8680</td>
</tr>
</tbody>
</table>
Control processes

– according to special table [12, 13] for the assigned $f_{res}$ and the level of significance $\alpha=0,05$, we find critical value $t_\alpha=2,12$;

– we determine confidence interval for the regression coefficients

$$\Delta b_i = t_\alpha \sqrt{s_{b_i}^2}, \quad (18)$$

$$\Delta b_i = 2,12 \sqrt{0,019} = 0,33;$$

– a comparative analysis demonstrates that the coefficient $b_{23}$ is insignificant, since the confidence interval $\Delta b_i=0,33>|b_{23}|=0,013$ for it.

As a result of experiment, we obtained a mathematical model in the form of the polynomial:

$$t_{per} = 15,59 + 14,09x_1 + 1,78X_2 - 5,79X_3 + 0,68X_3X_2 - 6,43X_3X_1,$$

$$t_{per} = 15,59 + 14,09X_1 + 1,78X_2 - 5,79X_3 + 0,68X_3X_2 - 6,43X_3X_1.$$  \hspace{1cm} (19)

This model describes technological time of moving of shunting train sets.

6. Discussion of results of calculation of technological time of a semi-flight of the shunting operation

Testing the adequacy of the mathematical model:

– from the obtained equation we calculate theoretical values of the function of response $\hat{y}$;

– we find the dispersion of adequacy

$$S^2_{ad} = \frac{m}{N-p} \sum_{i=1}^{N} (\hat{y} - \bar{y})^2,$$  \hspace{1cm} (20)

where $p$ is the number of significant factors.

$$S^2_{ad} = \frac{3}{8-6}1,935 = 0,967.$$

We find calculated value of the Fisher F-criterion:

$$F_i = \frac{S^2_{ad}}{S^2_{res}}, \quad (21)$$

$$F_i = \frac{0,967}{0,47} = 2,06.$$  

Determine the numbers of degrees of freedom:

$$f_1 = N - p;$$

$$f_1 = 8 - 6 = 2;$$

$$f_2 = N(m - 1);$$

$$f_2 = 8 \cdot (3 - 1) = 16.$$  \hspace{1cm} (23)

According to special table [12] for $f_1, f_2$ and $\alpha=0,05$ we find critical value $F_{0,05}=3,6$.

Since $F_i=2,06<F_{0,05}=3,6$, then the mathematical model is considered to be adequate.

The obtained equation of regression in the form (19) makes it possible to draw a number of fundamental conclusions. In particular, it follows from it that the least important, from the point of view of the effect on technological period of a semi-flight of shunting operations, is the factor of the quantity of railroad carriages in the shunting train set, as well as the combined influence of factors of the length of a semi-flight and the quantity of railroad carriages in the shunting train set. This is demonstrated by low absolute values of estimations of the coefficients at factors $x_1$ and $x_1x_2$. The largest influence on the output variable, from the point of view of its maximization, is exerted by the length of a semi-flight. An increase in the length of a semi-flight and in the quantity of railroad carriages in the shunting train set lead to the increase in the technological time of a semi-flight.

The opposite influence is rendered by the speed of motion of a shunting train set, the increase in which leads to reducing the technological time of a semi-flight of the shunting operation. Therefore, when examining selected input variables as the controlling factors of shunting operations at the railway stations, it is necessary to keep in mind that for the faster maximization of this process, it is necessary to increase the values of the length of a semi-flight and the quantity of railroad carriages in the shunting train set. At the same time, for its minimization, it is advisable to increase the speed of motion of the shunting train set, decreasing, in this case, the length of a semi-flight and the quantity of railroad carriages.

The response surface that describes combined effect of the considered input variables on technological time of a semi-flight of the shunting operation is displayed in Fig. 1. The values of input variables are represented in standardized form; moreover, the value of the factor “quantity of railroad carriages in the shunting train set” is accepted to be equal to zero, which corresponds to the fixation of the value of this variable at the basic level $m_{min}=-41$. It is obvious that by changing the values of this variable in the examined range of the area of planning $[-1; +1]$, it is possible to obtain different response surfaces, displaced relative to each other.

![Fig. 1. Response surface that describes combined effect of the length of a semi-flight $L_{sh}$ ($x_1$) and the speed of motion of the shunting train set $v_{sh}$ ($x_2$) on technological time of a semi-flight of the shunting operation $t_{per}$](image)
of equal level, which corresponds to the selected value of output variable, on the plane \((x_1, x_2), (x_1, x_3), (x_2, x_3)\) and subsequent construction of nomograms [14]. However, of the largest interest is the study of the response surface to detect local extrema that occur on the boundaries, described by the curves, obtained in the cross-section of response surface by cylinders of the radius \(r\). The latter are the analytical description of limitations in the form of the crests that allow determining the suboptimal values of output variables and the values of input variables corresponding to them [15]. In this case, the parametric description of corresponding equations takes the form:

\[
x^*(\lambda) = (\lambda I - A)^{-1} a,
\]

\[
r(\lambda) = \sqrt{x^*},
\]

\[
y^*(\lambda) = a_0 + 2a_1 x^* + x^* A x^*,
\]

(24)

where

\[
x^*_i = \frac{a_i}{2\lambda}
\]

are the suboptimal values of input variables,

\[
r = \sqrt{r^2}, \quad r^2 = \sum_{i=1}^{n} \left(\frac{a_i}{2\lambda}\right)^2
\]

is the radius of cylinder, which describes limitations to the input variables,

\[
y^*_k = 2a_1 x^* + x^* A x^*,
\]

are the suboptimal values of output variable, \(\lambda\) is the Lagrange multiplier, which acts as a parameter, \(a_0, A\) are the estimations of coefficients in description (19).

The application of this approach appears rather interesting particularly for those technological systems, in which a mathematical model of the process is described by the polynomial of \(n\) degree, and the limitations to input variables are linked to impossibility to exceed, under real practical conditions, the established technological regimes [16].

Fig. 3 demonstrates results of solution to the problem in the form of two graphs – dependency of radius on values of the Lagrange multiplier \(r=r(\lambda)\) (Fig. 3, \(a\)) and dependency of output variable on radius: \(y=y(r)\) (Fig. 3, \(b\)).

It follows from Fig. 3 that the theoretical minimum of the value of technological time of a semi-flight is reached at radius \(r=1,1\), which corresponds to the standardized values of input variables \(x_1^*=-1,1, x_2^*=0, x_3^*=-0,196\), obtained based on parametric description (24). This minimum is reached on the ridge line I. The maximum of the value of technological time of a semi-flight with regard to the limitations, superimposed on the input variables, for example \(r=2,14\), is reached at the appropriate radius and is \(t_{per}=46\) min (ridge line IV, Fig. 3, \(b\)). It is obvious that the ridge lines II and III are absent due to the uniqueness of the pole \(\lambda=0\).

It is necessary to note that this very approach to determining the best, in a sense, solutions – suboptimal values of output variable – makes it possible to avoid the excess expenditures, which invariably occur when solving the problems of experimental optimization. It is clear that conducting additional experiments, performed on the acting objects, in the direction towards the optimum is associated with significant costs and organizational difficulties. However, the use of ridge analysis provides the possibility of efficient study of the response surface and detecting the optimum values of input variables, understood in the sense of their maximum or minimum and reached at the boundaries of the examined area. The latter in particular are dictated by technological regimes, standardized in the process of operation of real objects. In this case, such limitations may be imposed by the following values of input variables: \(L_{vms}=10-10000\) m, \(m_{vms}=1-81\) units, \(v_{dm}=15-40\) km/h.

Let us proceed from the model with coded variables to the model with natural variables:

\[
\begin{align*}
pr vms &= 15,59 + 14,09 \left(\frac{L_{vms} - 5005}{4995}\right) + 1,78 \left(\frac{m_{vms} - 41}{40}\right) \\
n_{vms} &= -5,79 \left(\frac{L_{vms} - 5005}{4995}\right) + 0,68 \left(\frac{m_{vms} - 41}{40}\right) \\
n_{dm} &= -6,45 \left(\frac{L_{vms} - 5005}{4995}\right) + 0,68 \left(\frac{m_{vms} - 41}{40}\right)
\end{align*}
\]

(25)
After transformations, we finally obtain:
\[ t_{pr} = L_{pr} (0,00553 + 0,0000034m_{vms} - 0,000103v_{in}) + 0,0275m_{vms} + 0,052v_{in} - 1,06. \] \[ (26) \]

For different speeds of shunting operations, there were developed regression models of technological time for the semi-flights of the rounds of shunting locomotives for the train set and their transpositions from track to another track, from a park to a park (Table 4).

Mathematical models of the semi-flights of the rounds of locomotives and transposition of railroad carriages and train sets

<table>
<thead>
<tr>
<th>Motion speed, km/h</th>
<th>Type of formula</th>
<th>Range of change in parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Round of locomotive</td>
<td>Transposition</td>
<td>( t_{pr}, m )</td>
</tr>
<tr>
<td>From track to another track, from a carriage park to a carriage park:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>( t_{pr} = 0,103 + 0,00367L_{pr} )</td>
<td>( t_{pr} = 0,137 + 0,0030L_{pr} + 0,0217m_{vms} )</td>
</tr>
<tr>
<td>25</td>
<td>( t_{pr} = 0,254 + 0,00233L_{pr} )</td>
<td>( t_{pr} = 0,19 + 0,00238L_{pr} + 0,0217m_{vms} )</td>
</tr>
<tr>
<td>40</td>
<td>( t_{pr} = 0,367 + 0,00147L_{pr} )</td>
<td>( t_{pr} = 0,42 + 0,00148L_{pr} + 0,0347m_{vms} )</td>
</tr>
<tr>
<td>60</td>
<td>( t_{pr} = 1,06 + 0,0011L_{pr} )</td>
<td>( t_{pr} = 0,88 + 0,0011L_{pr} + 0,0452m_{vms} )</td>
</tr>
</tbody>
</table>

Mathematical models of transposition of railroad carriages from one cargo front to another:

\[-\text{without checking the overall dimensions:} \]
\[ t_{pr} = 0,18L_{pr} (0,0024 + 0,000057m_{vms} + 0,0092m_{vms}); \]
\[ (27) \]

\[-\text{with checking the overall dimensions:} \]
\[ t_{pr} = 0,053 + L_{pr} (0,00947 + 0,000065m_{vms}) + 0,136m_{vms}. \]
\[ (28) \]

The speeds of fulfilling shunting operations are determined by the Rules of technical operation.

Obtained polynomial models give the opportunity to calculate separate shunting operations, namely the transposition of railroad carriages, for the specified time norms without referring to tabular norms, which considerably simplifies solving the operational tasks.

Based on the obtained regression models, we plan to develop simulation models in the Anylogic software, which, in addition to the adequate standardization of duration of the shunting operations, will allow us to clearly represent shunting processes and to perform their analysis in detail.

7. Conclusions

1. Based on the analysis of the tabular norms of duration of the execution of shunting operations, we propose to use polynomial models as more convenient in the course of practical application and reflecting influence on the magnitude of technological time not only from separate variables (factors), but also their interaction.

2. To obtain regression dependency we applied a full factor experiment with effects of pair interaction. As a result of the experiment, we obtained a model of technological time of the transposition of shunting train sets, which adequately reflects the elements of transportation process. The dispersion of estimation of determining regression coefficient \( S^2_{adj} \) amounted to 0,0019, the dispersion of adequacy \( S^2_{adj} \) reached 0,967. The calculated values of the Fischer criterion \( F(2,06) \) are lower than the tabular 3,6, which indicates the adequacy of the model. The obtained dependencies are recommended to use when constructing simulation models of operation of the railway stations.

References