The main task of railway transport is timely meeting the needs of country’s economy in transportation. In spite of the fact that recent years have seen a sharp decrease in the volume of transportation work, railway transport has retained its leading position in the overall transportation balance of freight traffic in Ukraine. Under current unstable and hard-to-predict economic, political and social conditions, railway transportation is the kind of transport that is in constant operation and is attractive to most cargo owners [1].

However, increased competition between the types of transport as well as wear and tear of fixed assets of railway transport may lead in the future to the loss by railways of part of the transportation market. This situation leads to the deterioration of financial and economic situation of the railway transport system. Determining prospective volumes of freight transportation by railroads will make it possible to choose the optimal model of behavior of transportation system. In this case, it is necessary to take into account the impact of the totality of economic, political, technological, natural factors, as well as the conditions of domestic transport market.

As the economy in Ukraine largely depends on the operation of railway, then hasty management decision will have a negative impact on other industries. Without the ability to predict the future, it is impossible to organize normal func-
The problem of development of the system of predicting the volumes of freight traffic under modern conditions has become especially relevant. Analysis and forecasting of the volumes of cargo transportation is the most essential tool for developing effective management decisions regarding the formation of a tariff strategy. In addition, it will contribute to choosing the optimal strategy for the development of this sector of industry, to determining necessary technical equipment of railways, to planning the needs for material, labor and financial resources. These steps should have a positive impact on attracting clients to railway transport.

Predicting the volumes of transportation by railway transport is not sufficiently tackled yet. Modern conditions of functioning of transport require increasing the area of prediction and improving the procedures and methods for developing forecasts.

A model of prediction should provide flexible adaptation of indicators on operational activity to fluctuations in demand. This, in turn, would reduce economic losses of railway transport and increase profits. Especially important is the accuracy of results of the prediction under conditions of the implemented structural reform of the sector.

2. Literature review and problem statement

Theoretical and methodological bases for the present research are the scientific works of scientists in the field of planning and forecasting of different spheres of activity. Article [2] highlighted the need for predicting future volumes of transportation by rail when forming the level of tariff for freight traffic. Paper [3] argues that the task of predicting future values of the time series is the basis for financial planning in the economy. Based on historical values of the time series, planning of the work of an enterprise is executed, as well as management and optimization of production volumes and warehouse control. Empirical results [4] confirm that the artificial neural network is an effective tool for forecasting currency exchange rate. This procedure provides evidence that there is a possibility to derive hidden information for reliable forecast of the future.

Forecasting the transportation of raw materials and finished goods is required to improve the accuracy of assessment of the throughput of rail networks [5]. The authors employ simulation modeling of the primary and secondary delays in the train traffic over particular section. Choosing the variants of forecasting using a non-linear function of the trend to determine the projected operational turnover for the next period is proposed in [6]. This will make it possible to plan in advance optimal routes of delivery of cargoes and help managers of transport enterprises to improve management system of the transportation process. Authors of article [7] propose a simple approach to the prediction of one-dimensional time series, which combines advice from several generalized regression neural networks. The proposed algorithm looks more powerful than the existing ones.

A decision-making process when planning the required amount of rolling stock is considered to be one of the most problematic issues for transport industry. Paper [8] developed a new approach to determining the size of the fleet in air transportation using the Monte Carlo modeling method. By using this method, an airline gains maximal profits. Such decision-making procedure could be applied for railway transportation to detect the required number of wagons for freight traffic, which is predicted.

To plan the work of freight transportation, [9] proposes an algorithm of fuzzy regression with accurate prediction and adaptability of the system to actual situation. The authors also developed a neural network with the inverse error spread in order to predict the cost of freight transportation. The authors argue that the algorithm of fuzzy regression is more effective than the neural network at large-scale planning and forecasting. In this case, article [10] states that to achieve accurate long-term forecast, it is better to employ a neural network model. Such model yields more accurate results of forecasting than the approaches based on the theory of chaos.

Prediction of flows is an essential component of transportation systems. Making unjustified decisions when transporting the goods leads to the growth of financial costs of the railway industry. Management of transports must provide making and implementing only optimal decisions, which can comprehensively evaluate performance indicators of transportation vehicles for different variants of transportation [11]. Article [12] indicated that genetic algorithms demonstrate not only high accuracy of calculations but rather provide the possibility of taking into account the constraints on throughput and processing capacity of facilities in the railway infrastructure. Applying this method will make it possible to solve the task of calculation of train formation plans for the entire polygon of railways in Ukraine based on previous data on the volumes of cargo transportations.

The experiment performed in [13] demonstrated high accuracy of prediction at short-term forecasting of passenger traffic. The result of prediction can be used for planning cargo operations and for managing revenues of the industry. To model freight transportation, paper [14] describes an approach using artificial neural networks and a neuro-fuzzy output system. Such approach combines the ability to train artificial neural networks and a transparent character of fuzzy logic. The model turns out to be highly adaptive and effective when exploring nonlinear interconnections between different variables. By using information about cargo flows, this approach may be applied on the Ukrainian railways.

For the reliable prediction of the volumes of freight traffic on the railways of Ukraine and the possibility of deriving hidden information when planning railway wagon flows, the present work considered the aforementioned existing experience in forecasting both on transport and in other areas.

3. The aim and objectives of the study

The goal of the present study is to improve a management technology over rolling stock based on the predictive values of volumes of cargo transportations by railway transport.

To accomplish the set goal, the following tasks had to be solved:
- to run a statistical analysis of time series of the volumes of freight traffic by railways of Ukraine;
- to develop a model for predicting the volumes of cargo transportation and verify results of the prediction by accuracy;
to propose a mathematical model capable of solving the optimization problem on the organization of railway wagon flows.

4.1. Time series analysis

Identification of the structure of time series is needed in order to construct a mathematical model for predicting the volumes of cargo transportation by railway transport. Prediction of future values of the time series is used to make effective decisions. A time series analysis will help to determine the nature of the series and to predict future values of the time series.

An analysis of time series and the realization of prediction were performed using the STATISTICA software [15, 16]. Manufacturer of the program is the firm StatSoft Inc. (USA).

Before starting the implementation of the prediction model, we performed a statistical analysis of time series. Prediction was conducted based on statistical data [17] on the volumes of transportation of grain and the products of flour mills from 2002 through 2016 (Fig. 1).

Seasonality of transportation is observed not for all types of cargo. Thus, when examining the volumes of transportation of building materials, irregularity factor amounted to 1.1.

Seasonal component affects the loading of rolling stock and transportation infrastructure. Over the periods of mass loading of seasonal cargoes, there may occur a deficit in wagons (grain wagons) and in the throughput of railway stations and sections. This situation leads to an increase in the operational work of railways and, thereby, may lead to financial losses in the industry. Therefore, a necessary component for normal existence of railway transport is the task of reliable prediction of the volume of freight traffic. Of special attention is a seasonal cargo.

4.2. Examining a method of prediction to determine the volumes of cargo transportation

The problem of prediction was realized by employing the most popular structural model, in particular, a model based on artificial neural networks (ANN). This method of forecasting possesses high qualities of adaptation to variable input data through self-learning. The main advantage of neural network models is the non-linearity, that is, the ability to establish non-linear dependencies between future and actual values of processes. Other important advantages are adaptability, scalability (a parallel structure of ANN speeds up computation) and the uniformity of analysis and projecting. However, ANN disadvantages include a lack of transparency of modeling, complexity of the choice of architecture, high requirements to reliability of the training sample, difficult choice of the learning algorithm and a resource-intensive learning process [18].

The main types of the functions of activation are given in article [18].

Mathematically, the structure of a neuron can be described by a pair of equations:

\[
y_i = \sum_{k=1}^{K} (w_{ik} \cos \omega_{ik} t + v_{ik} \sin \omega_{ik} t),
\]

where \( w_{ik}, v_{ik} \) are the non-correlated random magnitudes with zero mathematical expectation and equal variances, \( \omega_{ik} \) is the wavelength of sines or cosines, that is, frequency.

Non-stationarity of the time series is obvious. Trend character amounted to 1.7. In the structure of time series we observe a clear seasonal component, which is why one can make an assumption that the seasonal lag effect is 12 months. There is also a certain cyclicity spanning 3–4 years, which is visible in the chart (Fig. 1). However, the given cyclicity exerts a much weaker effect than that of the seasonal component.

Result of periodogram is shown in Fig. 2.

Fig. 2. Spectral analysis of time series

Fig. 1. Volumes of transportation of grain and the products of flour mills
where $x_i$, $w_i$ are, accordingly, the forces at inputs to neurons and the magnitudes of synaptic bonds; $w_0$ is the coefficient of displacement of a neuron (weight of additional input); $U$ is the induced local field; $f(u)$ is the transfer function or a function of activation.

Additional input and its appropriate weight is used to initialize the neuron. Initialization implies a shift of the activation function of neuron along the horizontal axis, that is, the formation of a neuron sensitivity threshold. In addition, sometimes the output of a neuron is specifically assigned with some random value called a shift. The shift can be considered to be a signal on the additional, always loaded, synapse.

The type of neuron connection determines the architecture of a neural network. According to article [19], depending on the type of neuron connection, networks are divided into:

– networks of direct distribution;
– recurrent neural networks;
– radial basis functions;
– self-organising maps or Kohonen networks.

The architecture of multilayer neural networks is used now probably most often. Usually, a network consists of multiple of sensor elements (input nodes) that form the input layer, one or more hidden layers of computing neurons and one output layer of neurons.

A multi-layer neural network can model function of almost any degree of complexity, and the number of layers and the quantity of elements in each layer determine complexity of a function. Determining the number of intermediate layers and the quantity of elements in them is an important issue during construction.

Neural networks are especially efficient for systems that possess a high degree of non-linearity. In order to evaluate the ability of a neural network to effectively predict the volumes of cargo transportation by railway transport, first of all, it is necessary to analyze previous statistical data on freight traffic.

The prediction was realized using the Rumelhart fully connected multilayer perceptron (MLP) with direct distribution, architecture 12-96-1. The signals are distributed in one direction, from the input layer of neurons, through the hidden layers to the output layer, and at the output neurons we receive a signal processing result.

The activation function of the input layer is identical, that is, the input signals are not converted but are passed on to the next layer. The given network, in contrast to its predecessor, Rosenblatt perceptron, employs nonlinear functions of activation, namely sigmoid. It allows strengthening weak signals and prevents oversaturation with powerful signals. Such a nonlinear activation function solves the dilemma of noise saturation. In the hidden layer, logistic function is applied (4), and at the output – hyperbolic tangent (5).

$$f(U) = \frac{1}{1 + e^{-U}},$$

$$f(U) = \frac{e^U - e^{-U}}{e^U + e^{-U}}. \tag{5}$$

At the output of a neuron we receive predictive values of the volumes of freight traffic by railway.

5. Results of modeling and assessment of the accuracy of prediction

We used statistical observations [17] of the volumes of transportation of grain and the products of flour mills from 2002 through 2016 as the input data. The training sample composed of 90% of the values, the control – 10%. We employed actual data on freight traffic over 8 months of 2016 as a test sample for cross-checking. Results of prediction are shown in Fig. 3.

![Fig. 3. Comparison of prediction range to actual values](image)

A diagram of distribution of residuals (Fig. 4) can be used to judge the adequacy of the model; in the ideal case, a graph resembles a normal distribution law and is bell-shaped.

![Fig. 4. Diagram of distribution of residuals](image)

Dependence between input and output data is shown in Fig. 5.

Judging by the figure, one can conclude that the network captures the overall trend of the series.

Results of the cross-check are shown in Fig. 6.

The magnitude of accuracy of predicting the transportation is determined through the value of the mean absolute percent error (MAPE). The magnitude of accuracy of prediction is the notion directly opposite to a prediction error. If the error of prediction is large, then the accuracy is low and, vice versa, if the prediction error is small, then the accuracy is large.
Comparison of the actual and predicted data of the examined freight transportation period is given in Table 1.

<table>
<thead>
<tr>
<th>Period</th>
<th>Actual data</th>
<th>Predicted data</th>
<th>Residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 2016</td>
<td>1.8</td>
<td>1.786232</td>
<td>0.007649</td>
</tr>
<tr>
<td>February 2016</td>
<td>2.8</td>
<td>2.713318</td>
<td>0.030958</td>
</tr>
<tr>
<td>March 2016</td>
<td>2.7</td>
<td>2.588141</td>
<td>0.041429</td>
</tr>
<tr>
<td>April 2016</td>
<td>2.5</td>
<td>2.277128</td>
<td>0.089149</td>
</tr>
<tr>
<td>May 2016</td>
<td>1.7</td>
<td>1.618872</td>
<td>0.047722</td>
</tr>
<tr>
<td>June 2016</td>
<td>1.3</td>
<td>1.396673</td>
<td>0.074364</td>
</tr>
<tr>
<td>July 2016</td>
<td>2.1</td>
<td>2.266609</td>
<td>0.079338</td>
</tr>
<tr>
<td>August 2016</td>
<td>3.3</td>
<td>3.055346</td>
<td>0.074107</td>
</tr>
<tr>
<td>MAPE</td>
<td></td>
<td></td>
<td>5.56 %</td>
</tr>
</tbody>
</table>

Speaking of high accuracy, we always talk about lower error in prediction. Prediction error characterizes to some extent the accuracy of prediction procedure that was applied. The experiment revealed that the mean absolute percentage error of the volumes of transportation of grain and the products of flour mills over 8 months of 2016 was 5.56%. Given that railway transport is a fairly inert system, MAPE indicator of 5.56% is sufficient for making management decisions.

Due to the prediction of the volumes of transportation over the next period, it is possible to find the optimal technology of freight traffic. The above predictive values of the indicators of cargo transportation were taken into account in subsequent calculations.

5.2. Formalization of the technology of management of railway wagons

Technology of the organization of cargo transportation should be based on the rational use of the rolling stock for the execution of the planned volumes of traffic. Making up a plan of allocation of railway wagons depends mainly on the number of wagons required for transportation. Another important parameter in planning is the location of railway wagons in a particular period of time.

Optimal technology of the organization of railway wagon flows implies minimization of operational costs for the transportation of cargo. When calculating a plan for the formation of trains and an operational plan for managing railway wagons, the basic criterion in the choice of railway wagon flows management is the minimal operating cost. However, the same approach to calculating the transportation plan for different kinds of cargo is not optimal. Therefore, in order to find the best variant to move the railway wagons, it is proposed to consider a parameter of transportation irregularity, or seasonality. The above studies have shown that the irregularity factor for different cargoes is disproportionate. Selection of the rational plan of distribution of wagons along a railway network is represented in the form of objective function:

\[
C(x_0) = \sum_{i} \sum_{j} x_{ij} k \cdot \Delta y + \Delta y \sum_{i} x_{ij} k \cdot \Delta y + \frac{\Delta D}{\Delta x_0} \rightarrow \min, \tag{6}
\]
where $C$ is the cost of cargo transportation, UAH; $x_i$ is the variable number of railway wagons that move from station $i$ to station $j$; $k$ is the irregularity factor of cargo transportation; $d_i$, $d_e$ are the shared load of wagons and empty wagons, respectively, that travel from station $i$ to station $j$; $d_i^m$, $d_e^m$ are the expenses related to the motion of loaded and empty wagons, respectively; $AD$ is the specific cost due to the railway wagon flow that exceeds the mean value, UAH; $\Delta k$ is the specific increase in the number of wagons transported from station $i$ to station $j$.

We consider the following factors to be constraints:

- constraints associated with the throughput of stations and sections

\[
\sum_{e \in e} N_{st/sect}^{-(e)} - \sum_{e \in e} N_{st/sect}^{+(e)} \geq \Pi; \tag{7}
\]

- constraints on the length of a freight train wagons

\[
\sum_{e \in e} x_{ij} \leq L_{st}; \tag{8}
\]

- a constraint that takes into account the term of cargo delivery

\[
t_{ij} + n_{it} + t_{se} \leq T. \tag{9}
\]

Under condition of arrival of the required number of wagons to load points

\[
1 \leq x_{ij} \leq \frac{a_i}{q_{ai}}, \tag{10}
\]

where $\sum N_{st/sect}^{-(e)}$ is the existing throughput of stations and sections, respectively, relative to trains; $\sum N_{st/sect}^{+(e)}$ is the required throughput of stations and sections, respectively, which is required for the execution of transportation plan, in trains; $\Pi$ is the number of trains that can be engaged in the distribution of railway wagons by destinations in a railway network. $\mu(e)$ is the intensity of railway wagon flow on the $e$-th railway station; $S$ is the number of stations that a train calls; $L_{st}$ is the length of dispatching tracks on the $e$-th station; $t_{ij}$ is the time needed to load $x_i$ railway wagons; $h$; $t_{ij}$ is the time needed to form a train; $h$; $t_{it}$ is the time needed to travel the distance to destination, $h$; $T$ is the time of cargo delivery, $h$; $a_i$ is the volume of cargo in the $i$-th point of loading, $t$; $q_{ai}$ is the average static railway wagon load, t/wagons.

The application on the network of railways of the result of solution to the proposed model enables a dispatcher – the one who handles wagons – to make rational management decisions. That is, to plan and carry out on time the distribution of railway wagons of a particular type depending on demand for freight wagons. This model takes into account both annual and monthly fluctuations in traffic. As well as loading of transportation vehicles and railway infrastructure. This technology makes it possible to take both long-term and operational decisions directly in the system of the organization of railway wagon flows.

In order to automate management decision-making by operational personnel on railway transport, we simulated organization of railway wagon flows using the software package “Graphic information system of a transportation network of railways TMkarta”. The simulation was carried out on a virtual railway polygon. The problem on modeling was to build an optimal plan for the distribution of wagons on the calculation railway polygon, taking into account a predicted plan of irregularity in the transportation of appropriate cargoes. Graphic interpretation of the ordered set of variants to manage a railway wagon flow on the calculation railway polygon is shown in Fig. 7.

![Fig. 7. Graphic interpretation of the ordered set of variants to manage a railway wagon flow](image)

The system of displacement of a railway wagon flow may be transferred from the initial state to the next by using transfer function $S^k$ (sequence of optimal management at each $k$-th step). Transfer function is determined by the assessment of technical, technological, operational and economic (cost) indicators.

A problem on choosing a technology to manage a wagon flow for a given railway network polygon is determined by a sequence of optimal management of the system

\[
S^1 \left\{ (\varepsilon(i))^{(1)}, (\varepsilon(i + 1))^{(1)}, \ldots, (\varepsilon(i + m))^{(1)} \right\}. \tag{11}
\]

In this case, a condition must be satisfied to maximize objective function $W(\varepsilon)$ – the saving obtained as a result of the implementation of optimal plan of displacement of a wagon flow during operational planning, which can be represented as

\[
W(\varepsilon) = \max_{k=1}^{n} \sum_{i=1}^{m} \varepsilon(i)^{k-1} S^k, \tag{12}
\]

where $n$ is the number of steps of the system; $\varepsilon$ is the set of stations along the path of a wagon displacement.

Method of optimal displacement of a railway wagon flow in the railway polygon implies different variations of the formation of wagon flows (route dispatch, transverse, sectional, assembled trains) depending on the technological limitations of the system. The proposed system makes it possible to perform calculations for the polygons of any size and at arbitrary time of planning.

6. Discussion of results of simulation

Formation of phenomena under examination is always affected by a multitude of different factors. It should be noted that accuracy of a one-time prediction does not yield the same results during subsequent calculations. Full match or significant difference between the prediction and its implementation can be a consequence of particularly favorable or unfavorable circumstances. A one-time “good” forecast can be obtained based on a “bad model” and vice versa. However, the benefits of a neural network include the capability to learn. Consequently, the quality of predictions can be judged only when there is a set of comparisons between predictions and their implementation.
Studies have shown that the devised predictive model using a neural network belongs to the high precision class. The experiment demonstrated that the mean absolute percentage error in transportation volume was 5.56%. Such mathematical apparatus could be used for finding predictive data on railway transport.

Using the prediction of the volumes of transportation over a subsequent period, we proposed optimal technology of cargo transportation. The optimal technology of the organization of railway wagon flows implies minimization of operational costs for the transportation of cargo. In order to find the best variant for the displacement of railway wagons, we propose to consider a traffic irregularity factor, or seasonality.

To automate management decision-making by operational personnel of railway transport, we simulated organization of wagon flows using the software. The simulation was performed on the virtual railway polygon. Based on the aforementioned mathematical apparatus, we built optimal plan for the distribution of wagons on the estimated railway polygon, considering the projected plan of irregularity in the transportation of appropriate cargo.

In order to calculate wagon flows of the entire polygon of railways, a fundamentally new automated system to manage and distribute wagons is required.

7. Conclusions

1. We performed a time series analysis, which helped define the nature of series for different types of cargo. A non-stationarity of the system was established. The structure of time series of the volumes of transportation of grain and the products of flour mills clearly manifests a seasonal component. Irregularity factor amounted to 1.7. For building cargoes, the irregularity is almost absent. Irregularity factor for this case was 1.1. That is why, when calculating a cargo transportation plan, as well as during operational management of wagon flows, the irregularity in the transportation of each cargo should be taken into account.

2. With regard to the features of input information, we formed a neural network model for the prediction of volumes of transportation. This method of forecasting possesses considerable properties of adaptation to variable input data through self-learning. The magnitude of accuracy in predicting the transportation is determined using the value of mean absolute percent error that equaled 5.56%. Given that railway transport is a fairly inert system, such indicator is sufficient for management decision-making.

3. We formalized technology of organization of railway wagon flows. To find the rational variant of displacing the wagons, we proposed to take into account irregularity factor, or seasonality. Optimality is accepted to be minimization of operational costs for the transportation of cargo. Based on the obtained data on transportation volumes, and by using the optimization model, we received a solution to the problem on optimal plan of the distribution of wagons on an estimated railway polygon. The procedure of wagon distribution is quite universal and makes it possible to adjust the plan of train formation in real time. The proposed model could be integrated into the automated working places of dispatcher – the one who handles wagons, and a train dispatcher, for making rational management decisions.

References

1. Introduction

Integration of Ukraine into the global economic system puts forward new requirements to the development and implementation of methods for the protection of economic interests of the state. Absolutely relevant is the need for achieving economic security when carrying out foreign trade operations.

An increase in the competitiveness and quality of transportation services for the economy should be promoted by the creation of sustainable, cost-effective system of transport and logistics.

Effective functioning of transportation sector at present is impossible without implementation of information and communication technologies, development of intelligent transport systems (ITS). ITS provide systemic changes