DOI: http://doi.org/10.25728/cs.2023.2.5

CREATING FEATURE SPACES AND AUTOREGRESSIVE MODELS TO FORECAST RAILWAY TRACK DEVIATIONS

A.Yu. Vladova

Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
Financial University under the Government of the Russian Federation, Moscow, Russia

✉ avladova@ipu.ru

Abstract. Diagnosis of railway tracks reveals the deviations of rail parameters in the plan and profile from their nominal values. If the deviations approach the limit values, the speeds of trains must be reduced. Therefore, forecasting changes in the deviations is a topical problem. Despite the significant amount of diagnostic data collected, railway operators underuse machine learning methods to improve the quality of forecasting. The proposed approach differs from known counterparts as follows. First, the dimensionality of the feature space is increased by calculating the variation of the amplitudes of deviations from the nominal values and two types of areas (the deviation length multiplied by the amplitude and the deviation length multiplied by the variation of the amplitude); subsequently, this space is represented in the 3D matrix form. Second, a set of control parameters is formed; it includes the time and space discretization step, the type of seasonal fluctuations, the number of trend change points, etc. Third, the deviations are forecasted in groups differing in type and position along the track. Forecasting is based on minimizing the empirical risk criterion. As a result, a family of autoregressive models is obtained for each discretization interval along the length of the railway track.

Keywords: time series, diagnosis, software package, discrete technological process.

INTRODUCTION

Railways operate in difficult conditions, i.e., under the effect of natural and climatic factors and loads from passing trains. Railway tracks are diagnosed to measure the track geometry and reveal the deviations of rail parameters in the plan and profile from their nominal values. The revealed deviations are divided as follows:

– track width (narrowing, widening),

– rail level position (misalignments, smooth level variations),

– sag on the right and left rails in the vertical plane (right and left sag),

– rail position in the plan (lining).

The revealed deviations are associated with discrete time instants and track distance and have a nonuniform distribution. If the deviations approach the limit values, the speeds of trains must be reduced. Therefore, forecasting the deviations is a topical problem.

Thus, the goal of this paper is to improve the accuracy of forecasting the deviations of railway tracks. To achieve the goal, we formulate and solve the following tasks: review modern forecasting methods for the sizes of defects (deviations) in extended objects; analyze statistically the deviation parameters as multidimensional data; form the sets of input and output parameters and the combinations of their values; forecast the deviations of railway tracks.

1. A LITERATURE SURVEY

1.1. General analytics

We studied the literature using *Dimensions*, the world's largest linked research database with 135 million publications, 153 million patents, and 7 million grants. The mosaic diagram in Fig. 1 shows the distribution of publications by different fields for the query "*railway degradation*" (in thousand pieces).

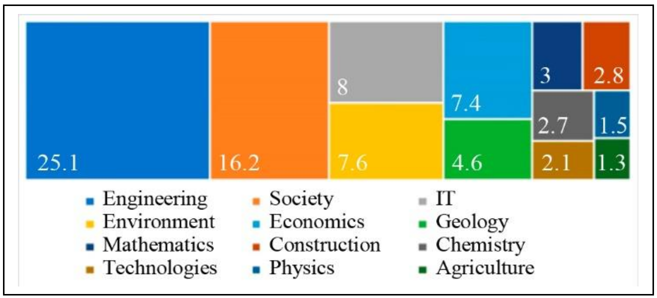

Fig. 1. The distribution of publications (in thousand pieces) by fields for 2012–2022.

With the permanently growing length and operation of railways, the issues of identification and forecasting of their condition are becoming acute. This process is reflected by the increased number of publications and their citations (Fig. 2) over the past 11 years. The characteristic rise of publications and citations is observed in two fields, "Engineering" and "IT"; it seems to be connected with the appearance of new sensors and big data processing methods.

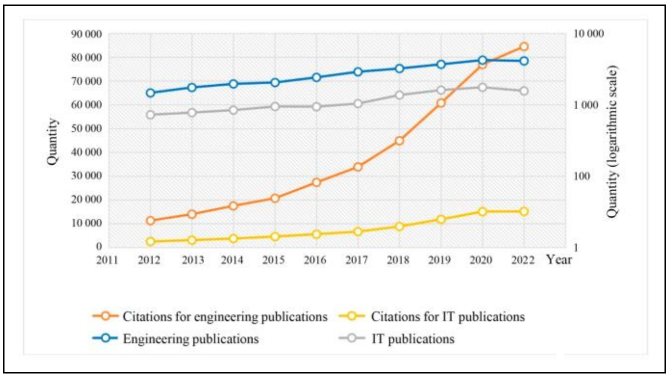

Fig. 2. The dynamics of publications and their citations for the query "*railway degradation*" in the fields "Engineering" and "IT."

1.2. Forecasting railway track deviations

The approaches to forecasting the condition of rail transport systems evolve due to different types of sensors and automatic processing methods for sensor data. The survey [1] was devoted to the condition monitoring of rail transport systems; the authors listed the main areas of development, including the creation of onboard systems based on different types of sensors and corresponding data processing methods. The publication [2] proposed a method to classify car-body vibration signs for the automatic detection of track defects. The causes of the fast degradation of a railway crossing were studied in [3]. The dynamic characteristics of the crossing were assessed using sensor equipment; according to the conclusions, the crossing characteristics degraded fast due to train hunting (the significant effect of wheel impact forces on the rail). The author [4] statistically analyzed diagnostic car data, constructed the probability densities of rail distributions by type, newness, purpose, and class, and clustered the results with transit tonnage recommendations. The paper [5] presented a model for estimating the intensity of railway track wear based on the slope angles of transient curves. According to the literature analysis, forecasting with the multidimensional time series characterizing the technical condition of railway tracks involves two basic types of models [6]:

- autoregressive models, which forecast the values of the time series $\check{z}$ using the past values $z(t), \ldots, z(t-p)$, e.g., by the additive model $\check{z}(t) = a + \varphi_1 z(t-1) + \ldots + \varphi_p z(t-p) + \epsilon(t)$, where φ_p are the weight coefficients and $\epsilon(t)$ is the error;
- regression models of the form $\check{z}(t) = f\left(x_1(t), \ldots, x_n(t)\right) + \epsilon(t)$, which forecast the time series values using a set of features $x_1(t), \ldots, x_n(t)$.

The main disadvantage of all these approaches is the limited use of big data and modern big data processing methods to improve the quality of forecasting the condition of railway tracks. However, in the literature devoted to the analysis of climatic, geotechnical [7, 8], and seismic data [9] as well as astronomical observations, we observe the wide application of classical multivariate parametric models and the Fourier transform (on the one hand) and the relatively recent wavelet transforms, canonical correlations, and trajectory matrices (on the other hand). Furthermore, these models can be adopted in optimization problems to find an optimal set of parameters [10].

1.3. Time series forecasting models implemented in machine learning libraries

We analyzed machine learning methods for time series forecasting on the example of Darts and GreyKite, the open source libraries. Darts contains the implementations of several regression models (ARIMA, Exponential Smoothing, Prophet, forecasting based on the Fast Fourier transform (FFT), etc. [11]). GreyKite provides time series forecasting based on Silverkite, Prophet, and ARIMA models, especially suitable for series with trend change points or seasonal fluctuations [12]. See Table 1 for a qualitative analysis of Silverkite, Prophet, and ARIMA, three modern models.

Table 1

Qualitative analysis of models

Criterion	Silverkite	Prophet	ARIMA
The speed of calculations	High	Low	Medium
The accuracy of forecasting	Very good	Good	Satisfactory
Interpretability	Good	Very good	Good
The ease of use	Satisfactory	Good	Very good
Visibility	Good	Very good	Satisfactory

After the analysis, we selected Prophet [13], an additive regression model with adjustable components. It has the form

$$\check{z}(t) = g(t) + s(t) + h(t) + \epsilon(t) \qquad (1)$$

with the following notations: t is time; $z(t)$ is the factual time series value; $\check{z}(t)$ is the forecasted time series value; $g(t)$ is the trend component described by a piecewise linear, piecewise logistic, or smooth function; $s(t)$ is the seasonal component (periodic changes), estimated using the partial Fourier sum; $h(t)$ is the component responsible for trend reversal instants (e.g., repair schedule); finally, $\epsilon(t)$ is the error containing the information neglected by this model.

The model allows managing different parameters (the seasonality of the time series, the number of trend change points, and the time step) and is stable to missed data and trend shifts. As a rule, it handles data outliers well.

Adjusting properly the components in equation (1) will ensure an acceptable quality of forecasting at the model output. This quality is estimated through the mean absolute error in percentage [14]:

$$\text{MAE}\,\% = \frac{\sum_{t=1}^{N} \left| z(t) - \check{z}(t) \right|}{\check{z}(t)}, \qquad (2)$$

where N denotes the number of time series points.

In addition, Prophet visualizes the confidence intervals [8] to establish the range of the output feature.

Thus, the problem under consideration has the following statement: on a training sample $\{ x_t : t = 1, ..., h\}$, find the control parameter vector U of the model $A(X, U)$ by minimizing the empirical risk (2). The vectors X and U take values from given subsets of the Euclidean spaces E^n and E^r, respectively: $X \subset E^n$, $U \subset E^r$. Constraints can be imposed on X and U: $g_i(X, U) \geq 0,\ i = 1, ..., m$.

2. FORECASTING THE PARAMETERS OF DEVIATIONS

2.1. Statistical analysis of initial data

The initial data occupy 8.9 Mb, are distributed in 10 columns (features), and contain 101 thousand rows. Among them, 91 rows have missed data; see Table 2 below.

Table 2

Initial data format

Assigned index	Feature	Range	The number of unique values
1D	Kilometer	[1, 650]	645
	Picket, m	[1, 1105]	1076
2D	Year	[2018, 2021]	4
	Month	[1, 12]	12
	Day	[1, 31]	29
3D	The code of deviation	[2065, 2161]	17
	Deviation	[Left sag, Wid]	18
Features	Amplitude, mm	[6, 1543]	144
	Nominal amplitude, mm	[10, 1535]	156
	Length, mm	[1, 308]	151
	The degree of danger	[1, 4]	4

According to the analysis results, the data are unbalanced:

- by year (Fig. 3*a*). There were as many deviations in 2020 as years 2018, 2019, and the first two months of 2021 combined;
- by the type of deviations (Fig. 3*b*). The distributions by year and type are presented in Fig. 3*d*. In 2020, the number of diagnosed deviations significantly increased; in 2021, the types of deviations increased, from 11 in 2018 to 18 in 2021. Moreover, the most frequent deviations changed from year to year: sag (Left sag, Right sag), misalignment (Mis), narrowing (Nar), widening (Wid), and smooth level variation (Lev);
- by the railway track length. Significantly more deviations were detected at the railway hubs compared to the track distance on average (Fig. 3*c*).

Due to the small number of some types of deviations, it is possible to combine them into a group or to generate additional data using bootstrap methods.

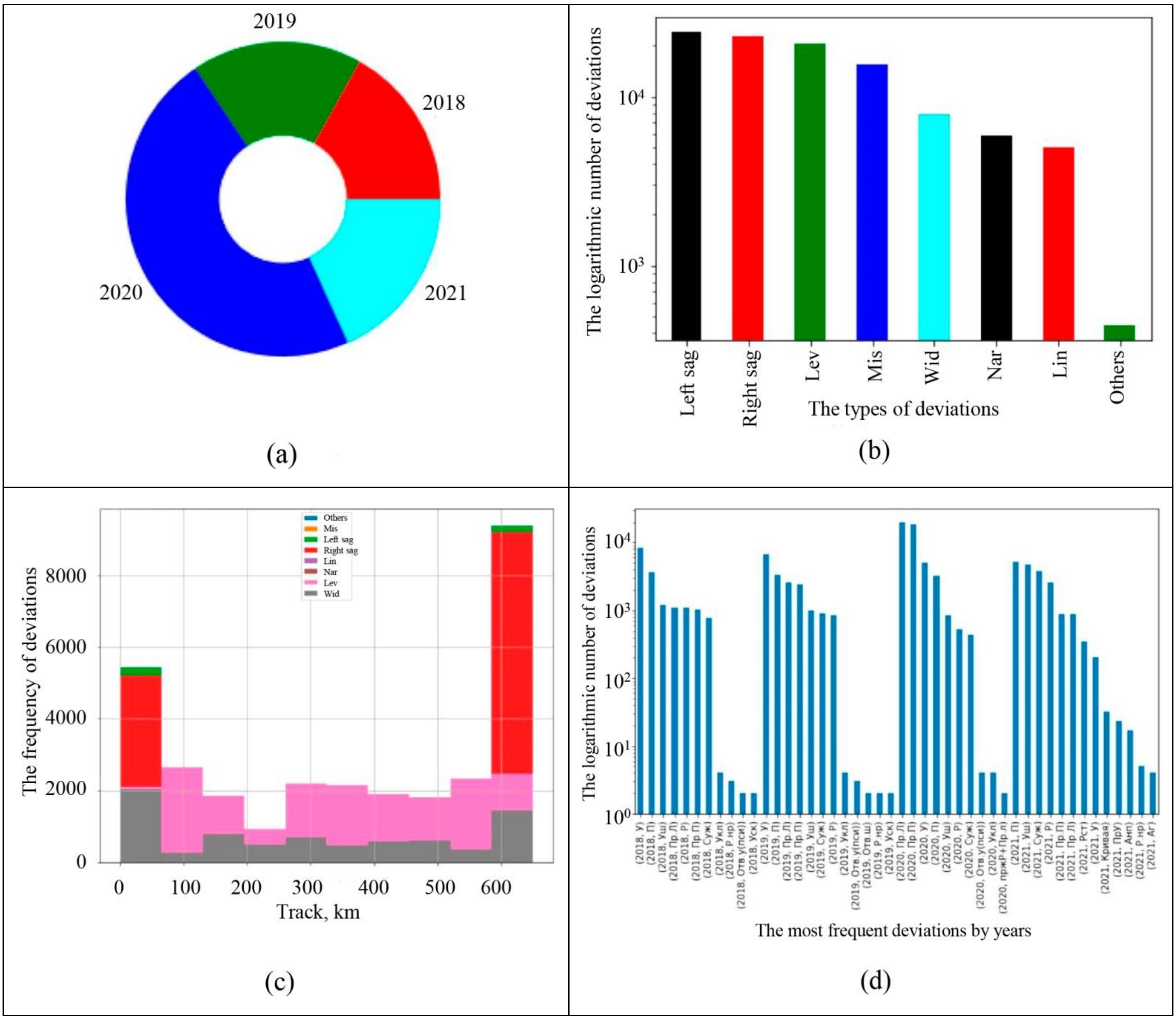

Fig. 3. Unbalanced data: (a) by year, (b) by the type of deviations, (c) by distance, and (d) by year and the type of deviations. In Figs. 3*b*–3*d*, Lin denotes lining.

2.1.1. Identify trend and seasonality

The presence of a trend and seasonality was checked using the upsampling and downsampling operations by varying the discretization step of the time grid:

• The trend for decreasing the amplitudes of deviations is traced by calculating their median values for each year. However, the lengths of the deviations first decreased and then sharply increased (Fig. 4*a*); the trend change point corresponds to year 2020. Note that "Danger," an aggregate parameter, is indifferent to changes in the median amplitude and the length of deviations.

• Downsampling by time (bimonthly data) and transition to the logarithmic scale (Fig. 4*d*) showed that the amplitudes of deviations have the highest variations in July. These changes suggest the presence of seasonality in the data. The locally optimal lengths of deviations, to some extent, follow those of the amplitudes with a time lag but have no pronounced seasonality. Also, no seasonal fluctuations were graphically identified for the degree of danger.

2.1.2. Increasing the dimensionality of the feature space

Since the time is represented by three integer features ("Year," "Month," and "Day"), the calculated feature "Time" was formed on their basis. The combination of two features, "Kilometer" and "Picket," characterizes the location of a deviation and forms the new feature “Distance.” To increase the stationarity of the time series, the differences between the features “Amplitude” and “Nominal Amplitude” were calculated and the attribute “Variation” was created. For a more complete description of the attribute “Variation” with a planar rather than point value, we introduced two aggregate features, “Area_A” and “Area_V” (the deviation length multiplied by the amplitude and the deviation length multiplied by the variation of the amplitude, respectively). In addition, 11 infrequent deviations were combined into the feature “Others” to increase the number of time series values. Fig. 5 demonstrates the peculiarities of the new features:

• the differences in the spread of deviations: only positive values for “Widening,” “Misalignment,” and “Sag”; only negative values for “Narrowing” with a significant degree of danger; a large dispersion for “Smooth level variations” (Fig. 5*a*);

• the clusters of the feature “Area_A” values by the types of deviations when passing to the logarithmic scale (Fig. 5*b*).

Hence, it is necessary to use different forecasting models [15, 16] for each type of deviations or to normalize their values.

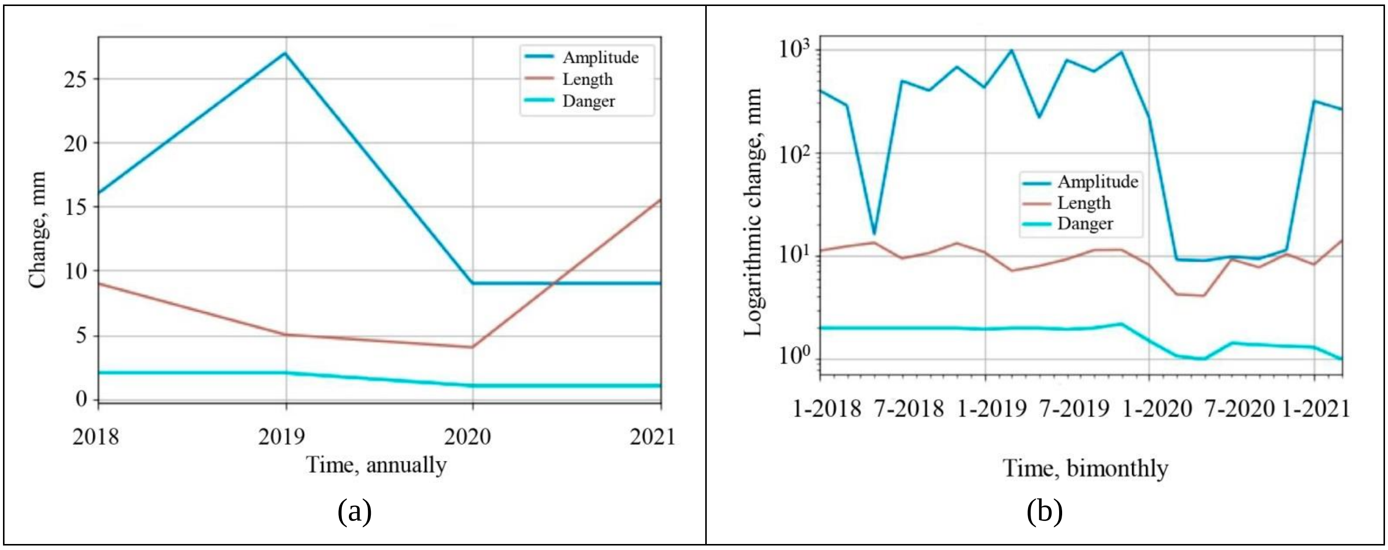

Fig. 4. The identification of (a) trend and (b) seasonality in the features “Amplitude,” “Deviation length,” and “The degree of danger.”

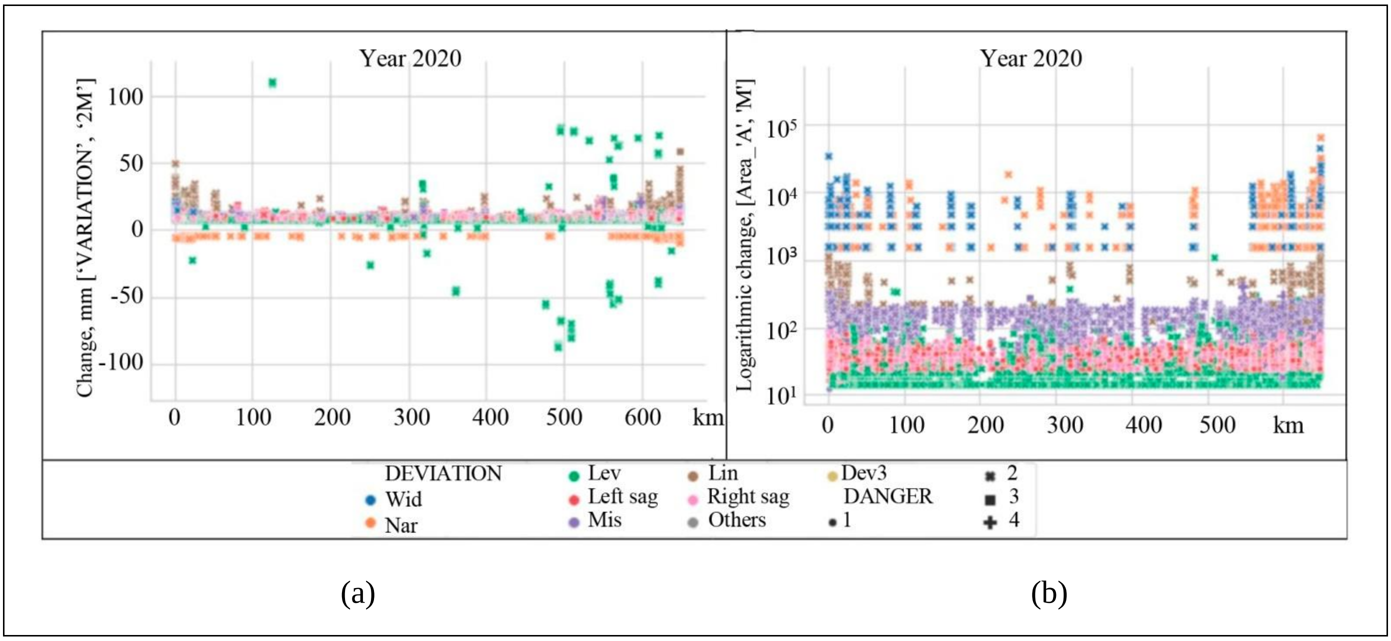

Fig. 5. Analysis of the dispersion by the type of deviation: (a) the spread of values and (b) value clusters.

2.1.3. Representing the data in the 3D matrix form

Each type of deviations may have specific dynamics (the law of variation), and each kilometer of railway tracks may have geotechnical peculiarities. Therefore, in the forecasting models, the length and amplitude of deviations, their nominal values as well as the aggregate features were written as a 3D matrix with the following axes:

– the location along the railway track,

– the time (the instant of detecting a deviation),

– the code of deviation.

A fragment of the transformed data is demonstrated in Fig. 6.

2.1.4. Probability densities

Probability densities were constructed on different time grids (with averaging from several weeks to a quarter) for the initial values of lengths, amplitudes, the variations of amplitudes, the areas of each type of deviations, and the degree of deviation. As it turned out, the values of the features “Variation” and “Area_V” are closest to the Gaussian distribution (Fig. 7).

The probability densities of the other features correspond to a two-tailed distribution. Therefore, their values were normalized into the interval [0, 1] by the formula

$$x_{i,\ \text{norm}} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, \qquad (3)$$

where $x_{i,\ \text{norm}}$ denotes the normalized value of a feature, $x_{\min}$ and $x_{\max}$ are its smallest and largest values, respectively, and x_i is the ith initial value of the feature.

Thus, the statistical analysis of the parameters of deviations as multidimensional data revealed the following properties: the data are unbalanced by year and the type of deviations, and trends and local minima appear when varying the discretization step of the data by year, month, and quarter. As a result, the least frequent types of deviations were grouped; the dimensionality of the feature space was increased by adding the position of deviations and the time instants of their detection and by calculating the variation of the deviation amplitudes from the nominal values and the corresponding areas of deviations (by the length, amplitude, and amplitude variation); the resulting feature space was represented in the 3D matrix form.

			Km	Amplitude	Normative	Type	Year	Danger	Length	Variation	Area_A	Area_V
Code	Track	Time										
2081	**1469**	**2018-01-29**	1	1529.0	1520	Wid	2018	2	4	9.0	6116.0	36.0
	1614	**2018-01-29**	1	1530.0	1520	Wid	2018	2	2	10.0	3060.0	20.0

Fig. 6. A fragment of the transformed data

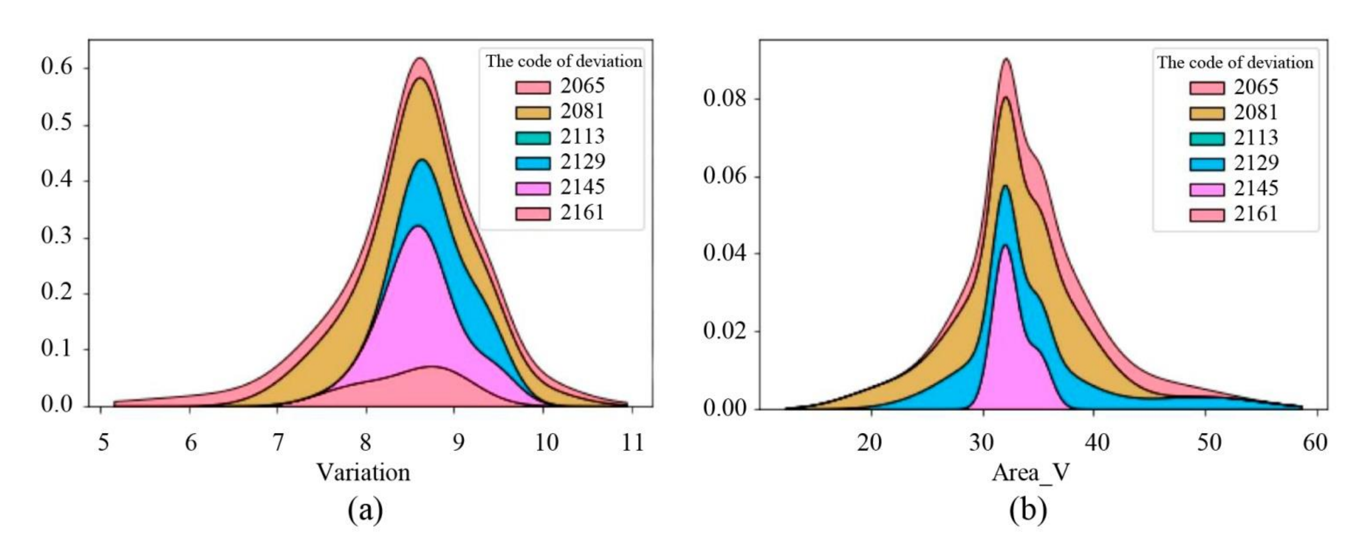

Fig. 7. The probability densities of feature values: (a) “Variation,” averaged on the six-week scale, and (b) “Area_V,” averaged on the seven-week scale.

2.2. The method for forecasting the parameters of railway track deviations

The proposed method for forecasting the parameters of railway track deviations includes three stages as follows. The first stage includes statistical analysis to group the rare types of deviations and form the model parameter vector *X* (the distance and time, the variation of amplitudes from the nominal values, and the areas of deviations). Then, the feature values are indexed along the three axes (distance, the code of deviation, and time) to form time series for each kilometer and the type of deviation.

In the second stage, the feature space is increased by adding the normalized and logarithmic lengths, amplitudes, the variations of amplitudes, areas, and the degrees of danger). Then, the control parameter vector *U* is constructed, including the time discretization step, seasonality, and the number of trend change points. After that, the constraints $g_i(U)$ are formed considering the following parameters:

– data depth and forecasting step: one year; three, two, or one quarter; two or one month; five, three, or two weeks;

– the type of seasonality: annual, quarterly, monthly, or weekly;

– the number of trend change points, according to the number of repairs;

– the type of seasonality model (additive or multiplicative);

– the rate of seasonal fluctuations (the number of terms in the partial Fourier sum).

In the third stage, under sufficiently many deviations in the group characterized by the type of deviation and the track kilometer, the model *A* (1) is trained in each group with estimating the empirical risk MAE% (2). As a result, the optimal values of the parameter vector *U* are determined for each group in terms of the minimum empirical risk criterion.

The practical significance of this method consists in the possibility to gain new knowledge and optimize railway transport.

2.3. Implementation

Table 3 shows the sets of model parameters corresponding to the best quality of forecasting.

The mean absolute error MAE% ranges from 2.6% to 8.4% for the recommended model parameters. According to the calculation results, we draw the following conclusions:

– Quarterly and annual grids significantly degrade the quality of forecasting.

– The best forecasts are for “Misalignment,” “Lining,” “Left sag,” “Widening,” and “Others” (the aggregate of 11 rarely encountered deviations).

– At most two trend change points should be selected.

– For the five-week grid, the best values of MAE% are achieved in the case of two trend change points.

– The lowest values of MAE% are achieved when considering the weekly, monthly, and quarterly seasonal fluctuations.

– Taking the logarithm of the feature values worsens the quality of forecasting.

– Normalizing the feature values into the interval [0, 1] improves the quality of forecasting.

Table 3

The results of forecasting

MAE%	Time step	Deviation	Variation	The number of trend change points
2.68	1 month	Misalignment	AREA_A	0
3.03	3 weeks	Lining	LENGTH	0
3.27	3 weeks	Widening	LENGTH	2
3.89	3 weeks	Left sag	VARIATION	0
4.36	1 month	Misalignment	AMPLITUDE	1
4.45	3 weeks	Others	LENGTH	2
4.60	5 weeks	Misalignment	VARIATION	2
5.34	1 month	Misalignment	AMPLITUDE	1
5.34	1 month	Right sag	AMPLITUDE	2
5.74	3 weeks	Others	AREA_V	0
5.74	3 weeks	Widening	AREA_V	0
6.59	1 month	Misalignment	DEGREE	2
7.90	5 weeks	Lining	LENGTH	2
8.39	3 weeks	Others	AREA_A	0
8.39	3 weeks	Widening	AREA_A	0

– Centering the values (subtracting the mean) and normalizing them (dividing by the dispersion) and using multiplicative regressors for seasonality worsens the quality of forecasting.

Figure 8*a* shows the one-month forecast for the feature "Misalignment"; Fig. 8*b*, the three-week forecast for the feature "Widening." The black dots are the factual values, the blue line corresponds to the model values, and the blue domain is the uncertainty corridor. Figures 8*c* and 8*d* present the identified trends; Figs. 8*d*–8*h*, the variations of the seasonal component, monthly and quarterly (Figs. 8*e* and 8*g*), and weekly

1-month time discretization Feature "Amplitude," value normalization into the interval [0, 1] 1 km, 1 trend change point Quarterly and monthly seasonality (Fourier order = 2) One-month forecast MAE% 4.36	3-weeks time discretization Feature "Length," value normalization into the interval [0, 1] 24 km, 2 trend change points Monthly and weekly seasonality (Fourier order = 3) One-month forecast MAE% 3.27

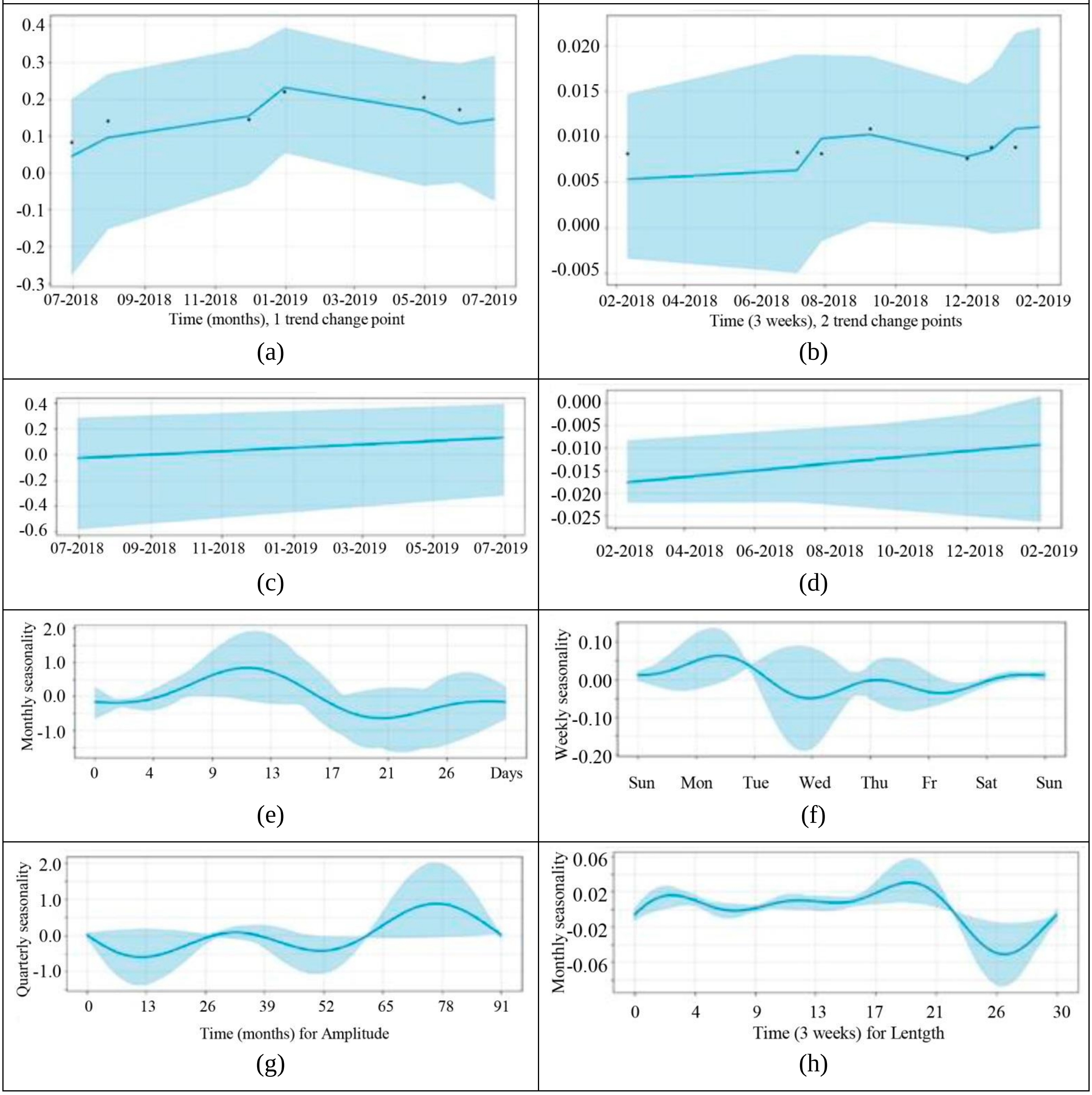

Fig. 8. The results of forecasting: (a), (b) model-based; (c), (d) trend; (e)–(h) seasonality.

and monthly (Figs. 8*f* and 8*h*). For both models, the uncertainty corridors for the seasonal and trend components are calculated as well.

According to the one-month forecast, the length and amplitude of deviations are increasing. The last point of the training sample is the local optimum point. The growth of amplitudes for the feature “Misalignment” falls on the last month in the quarter. This is probably due to repair work.

To implement the method, we developed a program in the Python language. The program forms a family of forecasting models for different groups of the types of deviations and the kilometers of roadway tracks on the features. It can be used to optimize railway transportation, which is of practical significance.

The Python language and the Google Colaboratory platform were applied to manipulate, visualize, and study the data. The statistical analysis results were visualized using Matplotlib, Pandas, and Seaborn libraries. The models were trained using SciPy, Sklearn, and Prophet libraries.

CONCLUSIONS

According to the literature survey, machine learning methods have proven their effectiveness in identifying and forecasting the condition of objects under a fairly long history of observations.

The statistical analysis carried out in this paper allows:

– revealing the number of deviations on any given time interval for any part of the railway track;

– comparing the sums of the variations of deviations from the nominal values for any railway track section at different time instants;

– normalizing the feature values into a certain interval;

– establishing data unbalancedness for different time intervals and types of deviations;

– synthesizing control parameters and features (time and distance, normalized and logarithmic lengths, amplitudes, the variations of amplitudes, areas, and the degrees of danger).

The parameters of railway track deviations have been statistically analyzed and the feature space has been supplemented by the variation of the amplitudes of deviations from the nominal values and the areas of deviations and variations.

The proposed method for forecasting the parameters of railway track deviations allows:

– defining the control parameter vector U;

– forming the constraints $g_i(U)$;

– obtaining a family of forecasting models by the types of deviations and railway track sections in terms of the minimum error (relative risk) criterion.

The control parameter vector of the forecasting model considers the type of seasonal fluctuations (from weekly to annual), the type of their description (additive or multiplicative), the rate of seasonality change (the number of terms in the partial Fourier sum), and the number of trend change points (in the general case, the number of time series points minus one).

To increase the accuracy of forecasting, including deviations with a shallow detection history, we have proposed to apply upsampling along the railway track (one and a half kilometers, two kilometers, etc.).

Data Availability. *A code fragment and a data fragment (three years of homogeneous measurements on two trace sections) that confirm the presented results can be found on the author's GitHub account: https://github.com/avladova/Railway-track-deviations or the author's website http://vladova.ru/About.*

REFERENCES

1. Kostrzewski, M. and Melnik, R., Condition Monitoring of Rail Transport Systems: A Bibliometric Performance Analysis and Systematic Literature Review, *Sensors*, 2021, vol. 21, no. 14, art. no. 4710.
2. Tsunashima, H., Condition Monitoring of Railway Tracks from Car-Body Vibration Using a Machine Learning Technique, *Applied Sciences*, 2019, vol. 9, art. no. 2734.
3. Liu, X. and Markine, V.L., Train Hunting Related Fast Degradation of a Railway Crossing-Condition Monitoring and Numerical Verification, *Sensors*, 2020, vol. 20, no. 8, art. no. 2278.
4. Vladova, A.Yu., Identification of the Railway Track Technical State, *Proceedings of 2021 14th International Conference “Management of Large-Scale System Development” (MLSD)*, Moscow, 2021, pp. 1–5.
5. Dubitskii, I.S., Enin, A.V., and Vladova, A.Yu., Analysis of Railway Track Wear Dynamics, *Trudy 14-oi mezhdunarodnoi konferentsii “Upravlenie razvitiem krupnomasshtabnykh sistem”* (Proceedings of the 14th International Conference “Management of Large-Scale System Development” (MLSD-2021)), Moscow, 2021, pp. 979–985. (In Russian.)
6. Lukashin, Yu.P., *Adaptivnye metody kratkosrochnogo prognozirovaniya vremennyh ryadov* (Adaptive Methods for Short-Term Time Series Forecasting), Moscow: Finansy i Statistika, 2003. (In Russian.)
7. Vladova, A.Yu., Harmonic Analysis of Temperatures of the Frozen Soil of Right-of-Way of the Linear Object, *Occupational Safety in Industry*, 2017, no. 7, pp. 25–30. (In Russian.)
8. Vladova, A.Yu., Remote Geotechnical Monitoring of a Buried Oil Pipeline, *Mathematics*, 2022, vol. 10, no. 11, art. no. 1813.
9. Lyubushin, A.A., *Analiz dannykh sistem geofizicheskogo i ekologicheskogo monitoringa* (Analysis of Geophysical and

Environmental Monitoring Systems Data), Moscow: Nauka, 2007. (In Russian.)
10. Valeev, S.G., *Regressionnoe modelirovanie pri obrabotke dannykh* (Regression Modeling in Data Processing), Kazan: FEN, Tatarstan Academy of Sciences, 2001. (In Russian.)
11. Herzen, J., Lässig, F., Piazetta, S.G., Neuer, T., et al., Darts: User-Friendly Modern Machine Learning for Time Series, *Journal of Machine Learning Research*, 2022, vol. 23, pp. 1–6.
12. Hosseini, R., Yang, K., Chen, A., and Patra, S., A Flexible Forecasting Model for Production Systems, *ArXiv: 2105.01098v1*, 2021, DOI: https://doi.org/10.48550/arXiv.2105.01098.
13. Taylor, S. and Letham, B., Forecasting at Scale, *The American Statistician*, 2018, vol. 72, no. 1, pp. 37–45.
14. Vandeput, N., Forecast KPI: RMSE, MAE, MAPE & Bias Towards Data Science, URL: https://towardsdatascience.com/forecast-kpi-rmse-mae-mape-bias-cdc5703d242d. (Accessed February 4, 2023.)
15. Hartomo, K. and Nataliani, Y., A New Model for Learning-Based Forecasting Procedure by Combining k-means Clustering and Time Series Forecasting Algorithms, *PeerJ Computer Science*, 2021, vol. 7, no. 2, art. no. e534. DOI: https://doi.org/10.7717/peerj-cs.534.
16. Vivas E., Allende-Cid H., and Salas R., A Systematic Review of Statistical and Machine Learning Methods for Electrical Power Forecasting with Reported MAPE Score, *Entropy*, 2020, vol. 22, no. 12, art. no. 1412.

This paper was recommended for publication by A.S. Mandel, a member of the Editorial Board.

Received November 30, 2021, and revised March 27, 2023. Accepted April 13, 2023.

Author information

Vladova, Alla Yur'evna. Dr. Sci. (Eng.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia; Financial University under the Government of the Russian Federation, Moscow, Russia
✉ avladova@ipu.ru

Cite this paper

Vladova, A.Yu., Creating Feature Spaces and Autoregressive Models to Forecast Railway Track Deviations. *Control Sciences* **2**, 46–55 (2023). http://doi.org/10.25728/cs.2023.2.5

Translated into English by *Alexander Yu. Mazurov*,
Cand. Sci. (Phys.–Math.),
Trapeznikov Institute of Control Sciences,
Russian Academy of Sciences, Moscow, Russia
✉ alexander.mazurov08@gmail.com