ISSN: 2755-9181

Review Article

Journal of Journalism and Media Management

Forecasting of the Daily Peak Electrical Load of a Machine-Building Enterprise

Vishniakou UA[1*] and Paloska KI[2]

[1]*Department of ICT. Research interests: information management and security, electronic business, intelligent control systems, IoT network, blockchain*

[2]*Department of Economic Informatics at BSUIR, postgraduate student*

[*]Corresponding author

Vishniakou UA Department of ICT. Research interests: information management and security, electronic business, intelligent control systems, IoT network, blockchain.

Received: October 02, 2026; **Accepted:** October 06, 2026; **Published:** October 16, 2026

ABSTRACT

The factors of the daily peak electrical load of a machine-building enterprise are investigated and an approach to its forecasting is proposed. Data of the Minsk Automobile Plant for 2024–2025 (731 days), together with meteorological and monthly production indicators, were used. The calculations were performed in the authors' adaptive system for comprehensive statistical analysis and forecasting (ASCSAF). Seven production-calendar mismatches and 12 violations of the relationship between electricity consumption and peak load were detected and eliminated; without their elimination the forecast error increases by about 25 %. Factor analysis identified three factors explaining 62.6 % of the variance. A combined scheme was adopted: ridge regression for working days, random forest for non-working days, and adaptive correction. The MAPE on working days was 4.75 % one day ahead and 7.14 % one month ahead.

Keywords: Peak Load, Data Preprocessing, Factor Analysis, Ridge Regression, Electricity Consumption Forecasting, Declared Capacity

Introduction

Under a two-part tariff, a significant share of an enterprise's electricity costs is determined by the declared capacity; therefore, even a small reduction in the monthly peak load yields an economic effect without cutting output. This requires identifying the factors of the peak and forecasting it – one day ahead for operational management and one month ahead for the capacity declaration.

Unlike a power system with stable daily and weekly dynamics the load of an individual plant is determined by the production schedule: the alternation of working days and days off, the shift pattern, and the melting schedule. Therefore, methods that are effective for a power system cannot always be transferred to the enterprise level and a high coefficient of determination here largely reflects only the difference between working and non-working days [1,2]. The problem is solved in the authors' adaptive system for comprehensive statistical analysis and forecasting, ASCSAF, which combines data quality control and preprocessing, correlation and factor analysis, a library of forecasting models and adaptive forecast correction with logging of all transformations [3,4]. This ensures reproducibility: when new data arrive, the analysis is repeated using a single algorithm.

Object of Study and Initial Data

The object of study is the Minsk Automobile Plant, a consumer with a connected capacity of 750 kW and above. The following data were used: daily data for 01.01.2024–31.12.2025 (peak load, electricity consumption E, working-day flag), the power profile by time-of-day zones for 181 days, monthly indicators of output and foundry production, and weather data for Minsk. The loading of the datasets and the initial control log are shown in Figure 1.

Data Quality Control and Preprocessing

The completeness of the series, value ranges, duplicates and logical relationships between the indicators were checked. The daily totals of electricity consumption and the monthly peaks matched the reporting for all 24 months; there are no missing values. The detected defects are listed in Table 1.

Calendar Flag

A «bridge» day is a calendar working day between non-working days; the shops are not started on such days, and the peak is only 22–25 % of the average working-day level, which corresponds to the base load. The rule relies only on the calendar, so such days are known in advance; over 731 days it triggered exactly three times with no false positives. Four Saturdays without a working mark had a peak of 88–107 % of the working-day average, with seven similar Saturdays marked as working, so the flag was restored as «working». The number of working days changed

Citation: Vishniakou UA, Paloska KI. Forecasting of the Daily Peak Electrical Load of a Machine-Building Enterprise. J Journalism Media Manag. 2026. 2(4): 1-7. DOI: doi.org/10.61440/JJMM.2026.v2.53

from 505 to 506. The rule describes typical rather than mandatory behaviour: if the enterprise decides to work on a bridge day, the forecast will be underestimated; therefore, the actual work schedule remains the preferred source of the flag, and the calendar rule is used when it is unavailable.

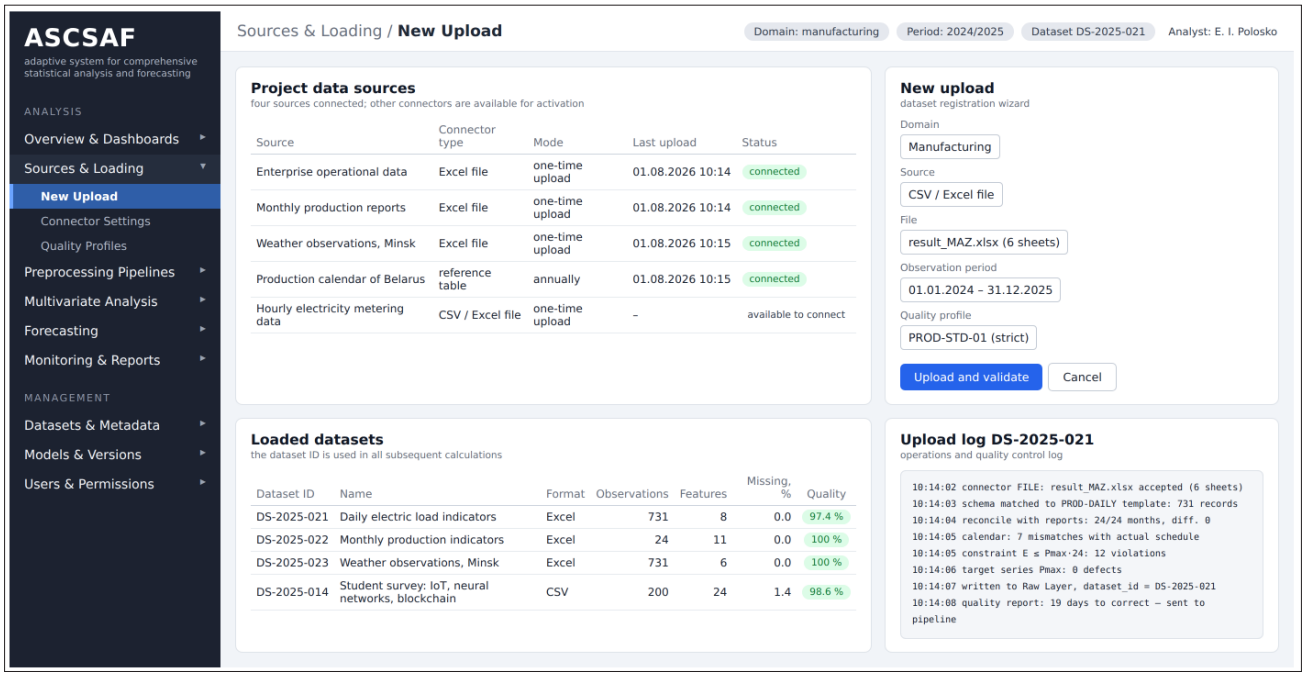

Figure 1: ASCSAF Screen: «Sources & Loading» Section

Table 1: Defects of the Initial Data and the Decisions Taken

Defect type	Days	Decision
Working day between non-working days («bridge day»)	3	Treated as non-working
Non-working day with full load	4	Flag corrected to «working»
Violation of the constraint E ≤ Pmax ×24	12	E restored by regression
Change in recording precision (from 25.08.2025)	129	Taken into account in interpretation

Electricity Consumption Anomalies

Physically, daily consumption is limited by the condition

E ≤ × 24,

where E is consumption, mWh; Pmax is the peak load, mW. Twelve violations were detected; for example, on 10.08.2025 the recorded consumption exceeded the physically possible value by a factor of 9.3, and on 31.10.2025 it was twice the monthly median with a normal peak. The target series Pmax contains no errors. The values of E were restored using regressions fitted on non-anomalous days:

E = 0.00888×Pmax + 60.05 − for non-working days,

E = 0.01172×Pmax − 26.33 − for working days.

The change in recording precision from 25.08.2025 (rounding to 1 kW and 1 mW) has practically no effect on the statistics and required no correction.

Formation of Signs

Weather and production indicators were joined to the daily data by date; monthly features were introduced with a time lag so that the forecast uses only the information available. A total of 31 features were formed for 671 observations in five blocks: calendar, seasonal (annual harmonics), meteorological (including heating degree-days), inertial (lags and moving averages of the load) and production. For factor analysis the data were standardised, and the minimum and maximum temperatures and the daily temperature range were excluded because of a practically singular correlation matrix (KMO increased from 0.561 to 0.653). The sequence of stages is given in Table 2, and the preprocessing pipeline with the quality control results is shown in Figure 2.

Table 2: Data Preprocessing Stages in the ASCSAF System

	Content	Result
Reconciliation of aggregates	Daily totals and peaks vs monthly reporting	24 of 24 months, discrepancy 0
Logical control	Check of E ≤ Pmax ×24, ranges, duplicates	12 violations in the E series

Calendar correction	Bridge days, unmarked working Saturdays	7 labels
Imputation	Regression of E on Pmax by day class	12 values (1.6 % of the series)
Feature engineering	Lags, moving averages, harmonics, degree-days, lag of monthly data	31 features, 671 observations
Preparation for factorisation	Standardisation, exclusion of collinear temperatures	KMO 0.561 → 0.653

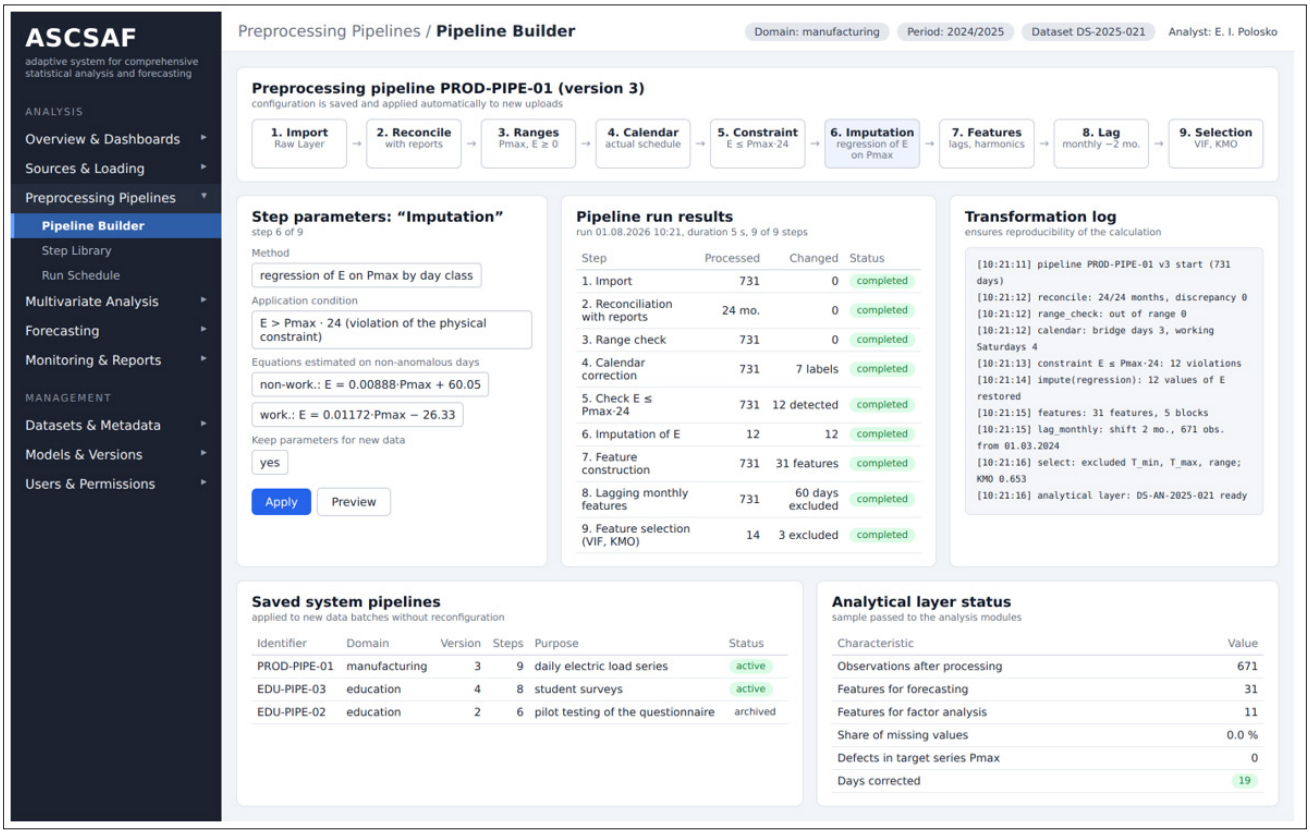

Figure 2: ASCSAP Screen: «Preprocessing» Pipelines» Section

The correctness of the decisions was verified by control calculations (Table 3). A two-month lag of the monthly features gave the best accuracy and guarantees their availability on the forecast date. Erroneous values of E degrade the forecast more than excluding the feature altogether; therefore, the check E ≤ Pmax×24 should be built into the metering system. Weather data, by contrast, are not critical: when the weather features were excluded or weather forecast errors were simulated (±1.5...2.5 °C), the RMSE changed by less than 1 %, which removes the system's dependence on external weather services. The factor structure is robust to the anomalies (KMO 0.653 vs 0.650).

Table 3: Effect of Preprocessing Decisions on Forecast Accuracy (Working Days)

Data preparation variant	RMSE, kW	MAPE, %
Baseline: E restored, 2-month lag	3,496	4.75
1-month lag of monthly features	3,631	4.95
3-month lag of monthly features	3,534	4.85
Original anomalous values of E	4,437	–
Previous-day E feature excluded	3,686	–

Correlation and Factor Analysis

The distribution of the peak is bimodal: on non-working days the average peak is about 21 % of the working-day level; the working-day flag explains 91.4 % of the variance ($\eta^2 = 0.914$), so model accuracy is evaluated below on working days. The intra-annual dynamics are significant (Kruskal–Wallis H = 184.5): the highest average level is in October (6.5 % above the annual average), the lowest in June (6.4 % below). The working-day level in 2025 decreased by 2.1 % relative to 2024 ($p = 0.007$). A load factor of 0.467 indicates a reserve for load levelling. On working days, the peak is most closely related to the level of the previous working day ($r = 0.703$); the autocorrelation decreases from 0.704 at a lag of one working day to 0.248 at a lag of seven, i.e. the load level is stable within the working week. The relationship with temperature is non-linear ($\rho = -0.525$ with $r = -0.356$): above +20 °C the

average peak is about 9 % lower than at 0...+10 °C, and below − 5 °C the growth stops. After removing seasonality, the relationship with heating degree-days is significant (partial r = 0.226, p < 0.001), whereas cooling is insignificant (r = −0.018): weather affects the peak mainly in winter through heating and ventilation. This conclusion is limited to the observed conditions: there were only 42 working days with temperatures above +20 °C.

Factor analysis of 11 indicators over 485 working days (KMO = 0.653, Bartlett's χ^2 = 2416.2, p < 0.001) with varimax rotation identified three factors explaining 62.6 % of the variance (Table 4).

Table 4: Factor Loadings After Varimax Rotation

Indicator	F1	F2	F3
Casting tonnage	**0.949**	0.079	0.004
Truck output per working day	**0.917**	0.071	0.032
Output per working day	**0.828**	0.107	0.095
Mean peak of recent working days	0.280	**0.833**	0.061
Peak of previous working day	0.241	**0.794**	0.006
Mean daily temperature	0.244	**−0.744**	0.057
Precipitation	−0.018	−0.036	**0.746**
Atmospheric pressure	−0.049	0.102	**−0.739**
Wind speed	0.137	0.204	**0.706**
Share of explained variance, %	**26.95**	**20.36**	**15.28**

F1 – production program, F2 – seasonal-inertial regime (achieved load level and temperature), F3 – weather instability. F2 is most closely related to the peak (r = 0.633 vs 0.258 for F1); F3 is practically unrelated (r = −0.076). The results of the module are shown in Figure 3. Previous-day electricity consumption has low communality (0.197), and its exclusion changes forecast accuracy by no more than 0.3 %. An auxiliary analysis of 24 monthly observations (KMO = 0.624) identified temperature and production factors (84.1 % of the variance): the monthly peak is related to the monthly average temperature (r = −0.781) and to the peak of the foundry (r = 0.634), whereas the relationship with output is insignificant, which with 24 observations does not prove its absence.

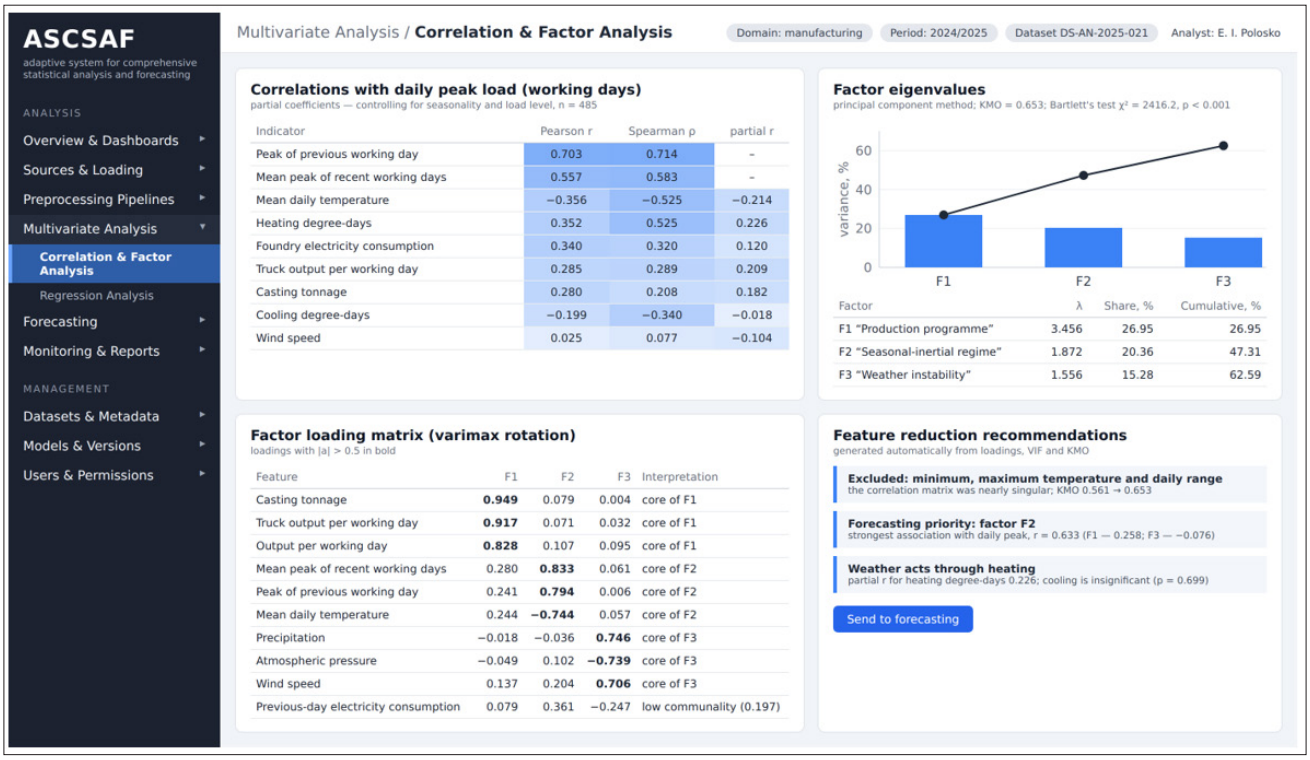

Figure 3: ASCSAF Screen: «Multivariate Analysis» Section

Forecasting Models

The ASCSAF model library covers the main classes of methods, which makes the system applicable to series of different length and structure: baseline models, linear and regularised regressions [5], decision-tree ensembles [6], and fully connected and recurrent neural networks (LSTM, GRU). The evaluation used rolling retraining without future information: at the beginning of each month the model was trained on all available data and forecast the following month (February – December 2025, 334 days, of which 232

were working days). Hyperparameters were tuned by grid search on data for May – July 2025; for neural network models the results were averaged over three runs. MAPE was not calculated over the entire sample, since the small load values of non-working days distort it. The results with adaptive correction are given in Table 5 and Figure 4.

Table 5: Comparison of Forecasting Models (February – December 2025)

Model	R^2 all	RMSE work, kW	MAPE work, %
Naïve (7 days earlier)	0.597	11,260	12.16
Calendar-group	0.957	3,875	5.53
Ridge regression (α = 0.3)	0.968	3,496	4.75
Random forest	0.967	3,752	4.94
Gradient boosting	0.963	3,950	5.43
MLP neural network	0.926	5,457	7.36
LSTM recurrent network	0.944	4,431	5.92
GRU recurrent network	0.954	4,170	5.64
Ensemble: regression + ridge + forest	0.970	3,520	4.71
Combined: ridge (work.) + forest (non-work.)	**0.970**	**3,496**	**4.75**

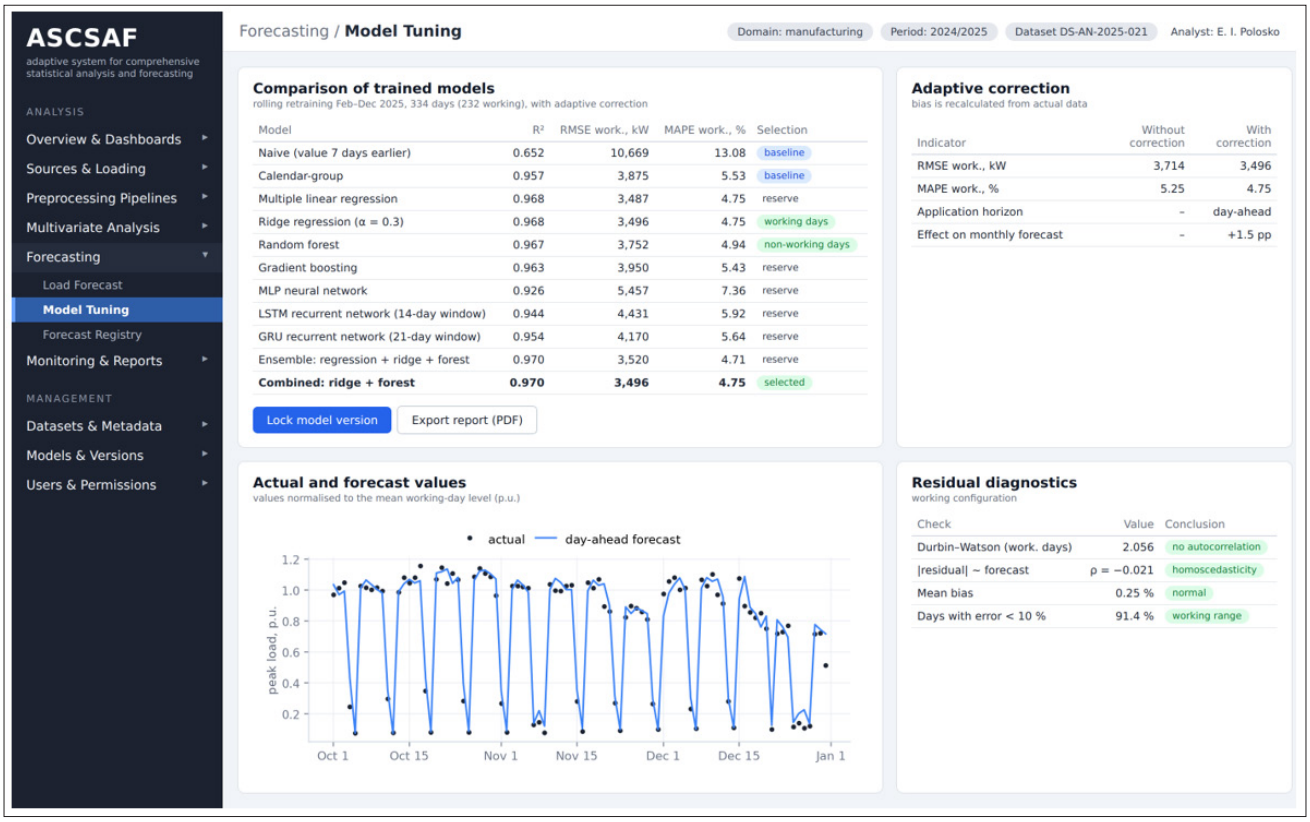

Figure 4: ASCSAF Screen: «Forecasting» Section

According to the Diebold–Mariano test [7], ridge regression is statistically indistinguishable from linear regression and the ensembles ($p = 0.484$); the forest, boosting and the calendar-group model are significantly worse. Neural network models are inferior to regression models: with 400–600 observations and 31 features, some of which are constant within a month, the training sample is insufficient for them. The recurrent LSTM and GRU networks (windows of 14 and 21 days) outperform the MLP but are inferior to the regression models: without correction their MAPE is 9.43 and 8.14 %, with correction 5.92 and 5.64 %. Ridge regression was chosen because of multicollinearity (VIF > 10 for nine features): with the same accuracy, its coefficients are more stable. For non-working days, the random forest is used, since the linear model produced negative values in long holiday series. The ensemble is formally more accurate by 0.04, but the difference is insignificant and its interpretability is worse, so ridge regression was preferred. Adaptive bias correction based on actual data reduces RMSE from 3,714 to 3,496 kW and MAPE from 5.25 to 4.75 %, but it is applicable only on the daily horizon: for a month it increases the error by about 1.5. The residuals are not autocorrelated (Durbin–Watson 2.056), the bias is 0.25 %, and the error magnitude does not depend on the forecast level ($\rho = -0.021$, $p = 0.749$).

Forecasting Results

On the daily horizon, the mean absolute percentage error was 4.75 %; an error below 10 % was obtained on 91.4 % of working days. On the monthly horizon (recursive day-by-day forecast with climatic-normal weather), MAPE is 7.14 %: from 1.69 % in September to 16.60 % in December 2025, when the load was reduced by a management decision: from 20 November the peak decreased by about 17 %, and from 16 December by almost 30 % of the previous level. Such a decision cannot be predicted from past data; the task of the system is to quickly take the new level into account, which is ensured by the day-ahead forecast with adaptation. The monthly forecast systematically underestimates the peak (median − 3.99 %), so the declaration should be adjusted. For January 2026, the forecast working-day level is 3–5 % lower than in January 2024 and 2025; taking the underestimation into account, the recommended declaration exceeds the point forecast of the peak by about 4 %, and the upper bound of the 90 % interval by almost 10 %. For September 2026 (a nine-month horizon, retrospective MAPE 7.86 %), in the baseline scenario the monthly peak is about 2 % lower than in September 2025, and the scenario corridor ranges from −17 % to the September 2025 level. The scenarios differ in the assumed load level: the inertial scenario keeps the September 2025 level, the baseline scenario follows the annual trend of −2.12 %, and the reduced scenario assumes the December 2025 level. The reduced scenario is a risk estimate in case of a sustained reduction of the production programme; the forecast should be updated using the data for Q1–Q2 2026, when the horizon shortens to three to four months.

Recommendations for Peak Smoothing

Measures are generated by the ASCSAF recommendation module using formalised rules: each rule compares the calculated indicators (the share of days with the peak in one time window, the number of days that set the monthly peak, the forecast bias, etc.) with domain-specific thresholds and outputs a measure with a justification. On 94.5 % of days the peak falls within the morning peak of 08:30–10:30, which exceeds the daytime level by only 2.88 %, and the monthly peak is set by about three days out of 21. A capacity reserve relative to the morning peak is observed both at night and in the evening, and it is higher at night; therefore, the constraint on load shifting is organisational rather than grid-related. The following measures are proposed: staged start-up of energy-intensive equipment at the beginning of a shift (up to 2.9 % of the peak); shifting some melts to the night – the peak of the foundry (22.6 % of the plant peak [8]) is determined by the melting schedule rather than by tonnage ($r = -0.301$); limiting secondary loads on days when the day-ahead forecast exceeds the current monthly peak (about 2 %); night-time preheating of premises in winter, taking into account the thermal inertia of buildings and sanitary requirements; adjustment of the declaration and data quality control. Under a two-part tariff [9], the capacity charge decreases in proportion to the reduction of the declared peak, i.e. potentially by 1.4–5.0 % (Figure 5). The effects of individual measures partially overlap and are not automatically additive; shifting melts and changing the ventilation mode require a preliminary technological check, whereas staged start-up and load limitation on risk days do not require major investment.

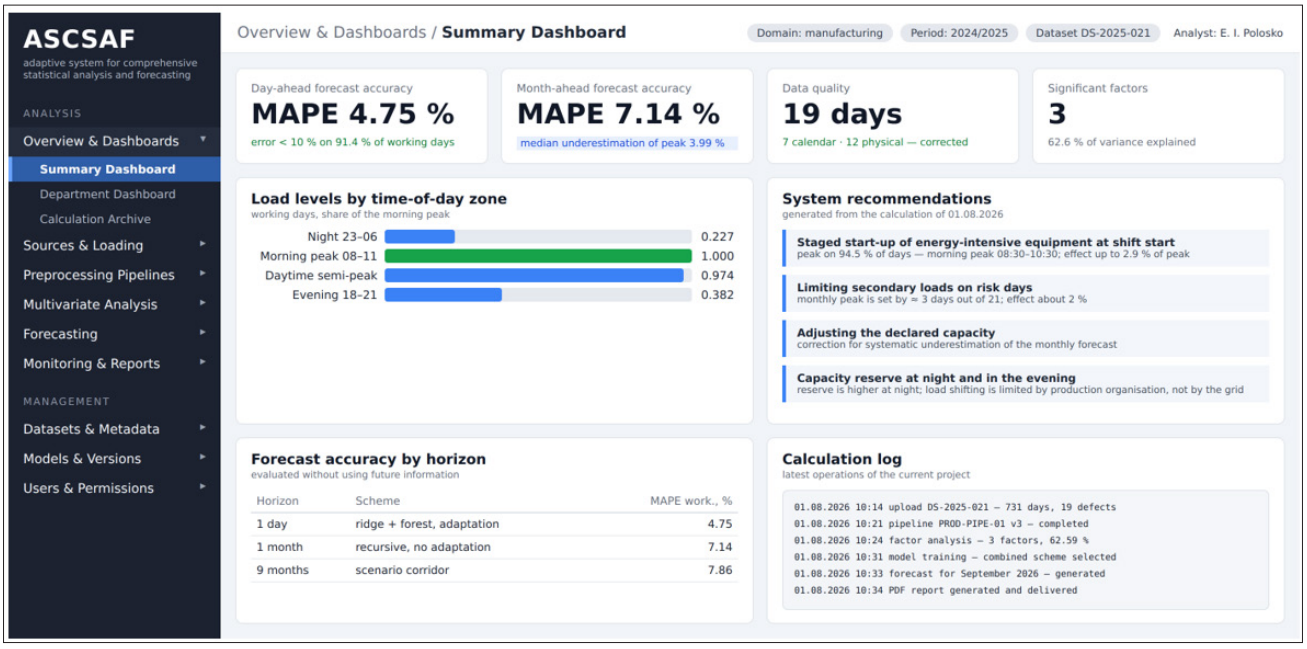

Figure 5: ASCSAF screen: «Overview & Dashboards» Section

Conclusion

1. Data preprocessing revealed 7 calendar mismatches and 12 violations of the condition $E \le P_{max} \cdot 24$; without their elimination, the forecast error increases by about 25 %. The peak on working days is determined primarily by load inertia and the heating component; three factors explain 62.6 % of the variance.
2. The combined scheme of ridge regression and random forest with adaptive correction provides MAPE of 4.75 % one day ahead and 7.14 % one month ahead.
3. Monthly production features should be used with a two-month lag; model accuracy is practically independent of the quality of

the weather forecast. Organisational measures can reduce the declared capacity by 1.4–5.0 % with a proportional reduction of the capacity charge.

References

1. Gursky S K. Adaptive forecasting of time series in the electric power industry / S K. Gursky. – Minsk : Nauka i tekhnika. 1983.
2. Hong T. Probabilistic electric load forecasting: A tutorial review / T. Hong, S. Fan // International Journal of Forecasting. 2016. 32: 914-938.
3. Vishniakou U A, Paloska K I. Software tools for automating calculations for factor, regression and correlation analyses. System Analysis and Applied Information Science. 2026. 1: 82-87.
4. Vishniakou UA, Khatskevich GA, Paloska KI. Comprehensive analysis for assessing the use of Internet of Things networks in the educational process of a university of informatics and radioelectronics. Statistics and Economics. 2025. 22: 52-60.
5. Hoerl AE, Kennard RW. Ridge regression: biased estimation for nonorthogonal problems / A. E. Hoerl, // Technometrics. 1970. 12: 55-67.
6. Breiman L. Random forests. Machine Learning. 2001. 45: 5-32.
7. Diebold FX, Mariano RS. Comparing predictive accuracy. Journal of Business & Economic Statistics. 1995. 13: 253-263.
8. Paloska KI, Golda Optimisation of energy consumption at OJSC “MAZ” using IoT sensors and neural networks for predictive analysis. Digital Transformation. 2026. 32: 45-50.
9. Electricity Supply Rules : approved by Resolution of the Council of Ministers of the Republic of Belarus of . 2011.

Copyright: © 2026 Vishniakou UA, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.