UDC 658.26:621.311.1:004.85
DOI: 10.15587/1729-4061.2026.352199

DESIGN OF A SYSTEM FOR LOAD FORECASTING AND OPTIMAL CONTROL OVER ENTERPRISE ENERGY SOURCES UNDER UNSTABLE POWER SUPPLY

Pavlo Shevchyk*
ORCID: https://orcid.org/0009-0006-8359-962X
Vitalii Polovyi
Corresponding author
Doctor of Philosophy (PhD)*
E-mail: vitalii.y.polovyi@lpnu.ua
ORCID: https://orcid.org/0000-0003-3425-8774
Yana Ni
PhD
Department of Electronic Computers
Kharkiv National University of Radio Electronics
Nauky ave., 14, Kharkiv, Ukraine, 61166
ORCID: https://orcid.org/0000-0002-1352-700X
*Department of Applied Mathematics
Lviv Polytechnic National University
S. Bandery str., 12, Lviv, Ukraine, 79013

This study investigates the process underlying the autonomy of industrial enterprises that face critical instability of centralized power supply and dynamic market tariffs.

Given the significant losses of generating capacities and frequent interruptions in the supply of electricity, enterprises often use autonomous power sources. Optimal utilization of such sources requires economic substantiation and construction of appropriate models. The task of these models is to improve the economic efficiency and energy independence of an enterprise by automating optimal control over energy sources.

This work proposes a two-level system. The first level is responsible for forecasting energy consumption. The second one is a deterministic optimization algorithm for automatic selection of a power source and economically justified control over the operational schedule of autonomous power sources.

When forecasting energy consumption, two models built on the basis of Random Forest and LightGBM were compared. The models yielded average absolute errors, as a percentage of the average, of 5–7% and 8–10%, which indicates their applicability for further decision-making.

Analysis of the optimization algorithm on real data revealed overall energy cost savings of 9–12% compared to unoptimized use of electricity from the grid. These results were achieved through timely switching between central and alternative power sources. Switching occurs when the use of the source becomes more economically advantageous, subject to technical constraints.

The resulting system could be used by enterprises that require uninterrupted power supply and exploit generators as alternative power sources

Keywords: enterprise energy consumption optimization, generator control, machine learning, LightGBM regression

Received 05.12.2025
Received in revised form 23.01.2026
Accepted 10.02.2026
Published 27.02.2026

***How to Cite:** Shevchyk, P., Polovyi, V., Ni, Y. (2026). Design of a system for load forecasting and optimal control over enterprise energy sources under unstable power supply. Eastern-European Journal of Enterprise Technologies, 1 (8 (139)), 6–14. https://doi.org/10.15587/1729-4061.2026.352199*

1. Introduction

Since the beginning of the full-scale invasion of Ukraine, the country's energy infrastructure has suffered serious losses. For example, between September 2022 and September 2024, more than 25 large-scale missile and drone attacks on energy facilities were carried out. As a result, almost 30 GW of generating capacity was lost – 21 GW in 2022–2023 and over 9 GW in 2024 [1]. Approximately 70% of all generating capacity was destroyed or damaged: the capacities of Centrenergo suffered significant damage, two hydroelectric power plants were destroyed, and DTEK lost 80% of its assets [2–4].

As of 2025, the war was in its active phase. In this regard, critical infrastructure remained under constant threat of new attacks. At the same time, the need for electricity, especially in industry, remained consistently high as enterprises play an important role in ensuring the country's defense capability. Before the war, Ukraine generated about 158.4 billion kWh of electricity annually, but now the vast majority of generating capacities have been destroyed or are temporarily unusable [1, 4].

In this context, it is relevant to devise solutions that could allow enterprises not only to effectively plan electricity consumption but also quickly switch to backup or alternative power sources, such as gas generation. Automated intelligent control systems based on machine learning methods could play an important role. Such systems solve the tasks of forecasting future loads, optimizing switching between energy sources, reducing peak loads and total consumption, as well as predicting future loads [5, 6].

2. Literature review and problem statement

The tasks of forecasting and optimizing energy consumption could be solved by using various machine learning models. For example, in study [7], a long short-term memory (LSTM) model was applied to forecast electricity consumption in the residential sector. The best results were achieved for aggregated data, for which the mean absolute percentage error (MAPE) metric values ranged from 8.18% to 8.64%. Data aggregation helps reproduce long-term dependencies in electricity consumption. However, this approach eliminates less significant local dependencies in the data because, for example, it ignores local peaks, which can lead to errors in planning autonomous power sources.

Work [8] demonstrated that LSTM outperforms conventional methods such as autoregressive integrated moving average (ARIMA), support vector regression (SVR) in medium-term forecasts during peak load times. The forecasting error for non-aggregated data according to the MAPE metric is approximately 20%, although this is expected due to the stochastic nature of non-aggregated data but is not suitable for the economic optimization problem. The paper describes possible steps to improve the accuracy, for example, increasing the context length and introducing dummy variables. However, the main goal of these changes is to more accurately predict peaks and long-term dependencies, rather than local variations in the data over a short period. A possible way to overcome this problem in the broader context of the presence of instability in the data may be the use of ensemble models, which usually better model conditional nonlinearities and behavioral dependencies for variable data.

An example of such a solution is study [9]. It shows that tree-based models such as eXtreme Gradient Boosting (XGBoost) provide highly accurate voltage stability forecasts in decentralized power systems (maximum accuracy of 96.8%) with a significant share of integrated renewable energy sources. Compared with classical analytical and statistical models, the model considered in the paper more effectively identifies unstable operating modes even under complex and changing system conditions. The resulting improved predictive ability allows for early intervention and supports improved voltage stability control, thereby reducing the risk of instability and system collapse. Although the model focuses on network stability and quality rather than load forecasting, it demonstrates that tree-based models with a wide set of features perform well under dynamic conditions.

It is important to note that in addition to technical load optimization, the price of different electricity sources should also be taken into account to verify the economic feasibility of the solution. Another way to solve the problem is to perform clustering on electricity consumption profiles before forecasting. The main assumption is that the variability of the data in the clusters will generally be less than the variability of the original data. Clustering was applied in [10]. It uses a cluster structure to forecast the load of electric vehicles to improve the efficiency of the network and the flexibility of charging systems. The Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) algorithm is used to build clusters. The clusters are used to classify charging cycles into residential, work, and instantaneous. Qualitative results indicate that the removal of emissions reduced the mean squared error (MSE) to 5.021 for the residential sector and 0.711 for the commercial sector. In addition, the coefficient of determination R^2 reached 96.87% for the residential sector and 94.89% for the commercial sector. This indicates a better fit of the model and an overall high forecasting accuracy. It is important to note that historical clusters lose their relevance under conditions of network instability or unstable consumption. Such conditions require the model to quickly adapt to changes depending on data, which is not addressed in the work.

Another option may be to use reinforcement learning, which can help obtain a more stable solution by choosing optimal strategies under conditions of uncertainty. This approach is reported in [11]. In the paper, algorithms were developed to optimize electricity usage for residential demand management, taking into account load and pricing uncertainties, user privacy, and distribution network constraints. One approach integrates deep reinforcement learning with federated learning, using convex relaxation to handle non-convex power flow constraints and transforming neural network updates into semi-definite programming problems. Simulation results on Institute of Electrical and Electronics Engineers (IEEE) test systems demonstrated a 33% reduction in peak load, a 13% reduction in expected user costs, and scalability for large networks. The results highlight the potential of learning-based strategies in demand management.

Similarly, paper [12] describes a system implemented using deep reinforcement learning algorithms that effectively controlled various system components, from battery-based energy storage systems to solar panels and generators. The proposed microgrid model achieved optimal performance with near-optimal policies. As a result, a 13% cost reduction was achieved compared to manual power system control. However, it should be noted that the formulation of such tasks is usually more complex, and the approach requires a simulator of the environment in which the agent is trained. In addition, if consumption profiles change over time, it is necessary to retrain the agent, which is resource-intensive and may not be feasible when daily forecasts are needed.

Another type of adaptive systems based on fuzzy rules can also solve the problem of multiple power sources. For example, work [13] describes a decision-making system that uses fuzzy logic to switch between power sources, but the main disadvantage of such systems is the static nature of the rules and the inability to adapt quickly. In addition, the work uses a multilayer perceptron to predict energy generation with an error of 8%. It is worth noting that the model is not always suitable for forecasting, especially for non-stationary time series, which is not discussed in the scope of that work since it focuses only on simple approaches to forecasting. In such cases, it may be more appropriate to use recurrent neural networks or regression models such as ARIMA, SVR, or Random Forest.

Overall, our review of the literature [7–13] demonstrates that modern approaches to artificial intelligence and machine learning can significantly improve forecasting accuracy, operational efficiency, and decision-making in power systems. These approaches are used in various areas, from demand forecasting and voltage stability assessment to demand management and microgrid control. However, regions facing structural constraints, infrastructure damage, and increased reliability requirements require adaptive, resilient, and autonomous solutions that go beyond theoretical productivity gains. This gap between technological potential and actual implementation is noticeable in countries experiencing systemic failures, such as Ukraine, where energy autonomy and the ability to respond quickly are becoming increasingly critical.

The current state of the Ukrainian power system is analyzed in [14]. It is worth noting that the work does not report mathematical models or simulations. It focuses on formulating the steps necessary to improve the situation. Another analytical paper [15] investigates the load schedules in the Ukrainian electricity system. It calculates that the system has a deficit of approximately 1 GW. As a result, recommendations are formulated to overcome this deficit; however, without specifying formal models or simulations that would confirm the feasibility of the proposed measures.

A more practical work [16] considers a hybrid model for forecasting electricity demand in Ukraine, which combines ARIMA and LSTM models. For the interval 2013–2020, the forecasting error was 2.78% for the MAPE metric. And for separate periods 2019+2020, 2019, 2020, the error was 3.17%, 2.79%, and 3.56%, respectively, for the same metric. This indicates its effectiveness for strategic planning of the enterprise. It should also be noted that such models are focused on medium-term forecasting of electricity demand, which can be considered a certain limitation for operational daily planning.

Studies [14–16] indicate that the Ukrainian power system faces serious operational challenges, including peak load imbalance, limited maneuvering and storage capacities. In addition, forced consumption restrictions and insufficient flexibility to ensure stable electricity supply necessitate the transition to decentralized and autonomous control mechanisms. In addition, lightweight and adaptive models capable of working with incomplete data and in offline mode are critically important under conditions of instability [13, 17].

In this context, hybrid local systems that integrate edge artificial intelligence (Edge AI), local data storage, and minimal dependence on centralized control are particularly relevant [14, 17]. As follows from the review of papers [7–13], an unresolved problem is that a significant part of modern research conventionally operates under the assumption of the stability of the main power source and is focused on optimizing consumption and stabilizing the load. The obvious reason for this is the low priority of such problems in countries with minimal risks of sudden loss of power supply.

The above gives grounds to argue that it is advisable to conduct research aimed at devising adaptive models of enterprise power system management that take into account the instability of power supply sources and technical limitations of internal generation.

3. The aim and objectives of the study

The aim of the study is to design a two-level system for forecasting electricity consumption of and optimal control over the enterprise's electricity supply sources. The system consists of a model for forecasting electricity consumption for a day in advance and an algorithm for generating an optimal schedule for managing electricity sources. Such a system should solve the task of increasing the economic efficiency and energy autonomy of the enterprise by optimizing the use of its own electricity supply sources, taking into account the potential instability of the main energy supply source.

To achieve this aim, the following objectives were set:

– to design a logical-information structure of the load forecasting system and optimal management of electricity supply sources;

– to construct a mathematical model for forecasting energy consumption 24 hours in advance based on historical data and to assess the accuracy of the forecast;

– to develop an algorithm for optimal management of the enterprise's electricity supply sources and assess its economic efficiency.

4. Materials and methods

4. 1. The object and hypothesis of the study

The object of the study is the process of ensuring the autonomy of industrial enterprises under conditions of critical instability of centralized power supply and dynamic market tariffs.

The principal hypothesis assumes that the use of short-term energy consumption forecasts, formed by regression ensemble models, within the framework of a deterministic model of mixed-integer linear programming (MILP), could make it possible to minimize the total costs of the enterprise for energy supply. It is assumed that the economic effect would be achieved by dynamic switching between the centralized power grid and local generating capacities, subject to strict adherence to technological constraints.

It is assumed that the obtained forecast will be sufficiently accurate and stable for use and will not cause false optimization in the algorithm described using the MILP apparatus. It is also assumed that switching between power sources is carried out almost instantly.

Simplifications are the use of binary indicators to control generators and ignoring the costs of depreciation and maintenance of equipment.

4. 2. Methods for the energy consumption forecasting model

To forecast electricity consumption 24 hours in advance, the ensemble machine learning methods Random Forest and LightGBM were selected. The choice of these methods is due to their ability to effectively handle nonlinear dependencies in time series and their resistance to outliers in the data.

At the stage of initial prototyping, Random Forest was used for the forecasting task, which forms the forecast as the average value of the results of a large number of decision trees [18, 19]

$$\hat{y}=\frac{1}{N}\sum_{i=1}^{N} y_i\left(x\right), \tag{1}$$

where N is the number of trees, $y_i(x)$ is the prediction of the i-th tree.

Random Forest is algorithmically simple and suitable for parallel computation of trees but may suffer from reduced accuracy in irregular patterns [20, 21].

LightGBM is based on the gradient boosting method and decision tree construction using coarse bins, which significantly reduces training time while maintaining high accuracy [22]. LightGBM consistently minimizes the loss function

$$\hat{y}_t\left(x\right)=\hat{y}_{t-1}+\eta f_t\left(x\right), \tag{2}$$

where η is the learning rate, $f_t(x)$ is the regressor tree at iteration t.

Among the advantages of LightGBM are its resistance to overfitting, interpretability due to the importance of features, and the ability to model complex dependencies without scaling the data. However, LightGBM can become less stable and less productive with an excessively large number of features due to increased noise and a decrease in the importance of useful features [21, 22].

4. 3. Metrics for assessing the accuracy of the electricity consumption forecasting model

Two main metrics were used to assess the accuracy of the model's forecasting: the mean absolute error (MAE)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right|, \tag{3}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, n is the number of observations, and the root mean square error (*RMSE*)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}. \tag{4}$$

RMSE allows us to estimate how large the deviations of forecasts from actual values are, taking into account the quadratic error scale, which is more sensitive to large deviations.

4. 4. Mathematical apparatus for the optimal control algorithm

To build a model for optimal control over energy supply sources, the MILP methods [23] were chosen. This approach allows us to formalize the problem of minimizing electricity costs as an objective function in the presence of a system of linear constraints.

The use of MILP is appropriate for modeling the operating modes of the system since it allows us to operate with both continuous variables (such as load) and binary ones (such as logical states of equipment). This makes it possible to find the global optimum within a given planning horizon.

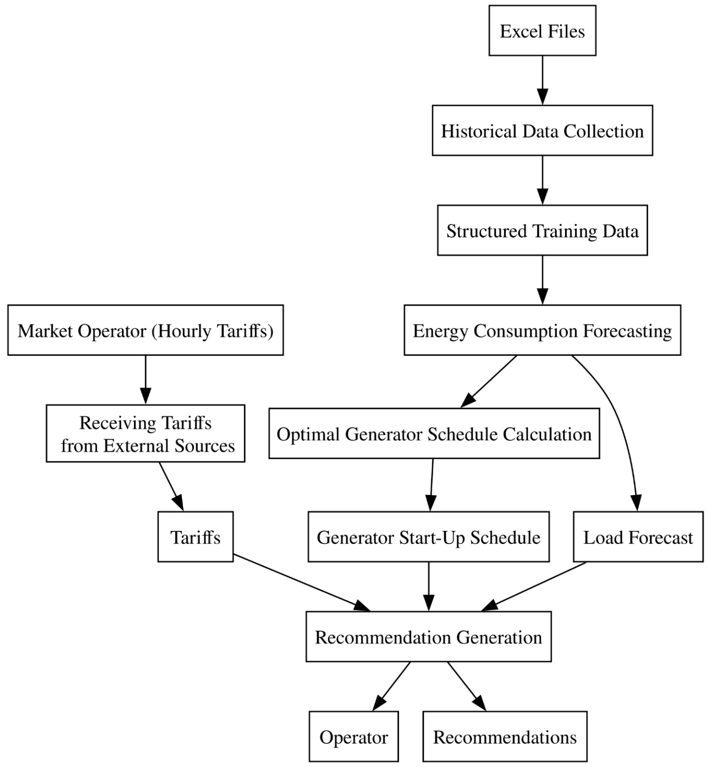

Fig. 2. Functional diagram IDEF0

5. Results of designing the system of forecasting energy consumption and management of energy supply of an enterprise

5. 1. Results of construction of the logical-informational structure of the system

Designing the information system of forecasting and optimal management of sources of power supply of an enterprise requires a clear formalization of its internal structure. For this purpose, a hierarchy of processes was built in the form of a function tree, which is based on the classical methodology of structural analysis, in particular, the data flow diagram (DFD – Fig. 1) and the functional diagram (IDEF0 – Fig. 2) [24, 25].

The basis of the system is a generalized function that reflects the main task of the system, namely, optimizing the use of power sources taking into account the forecasted electricity consumption and changes in energy tariffs. The upper part of the process tree is responsible for automated decision-making regarding the time to turn on generators based on the analysis of historical data, current tariffs, and load forecasts. This part is divided into several parts below, each of which corresponds to a separate subsystem.

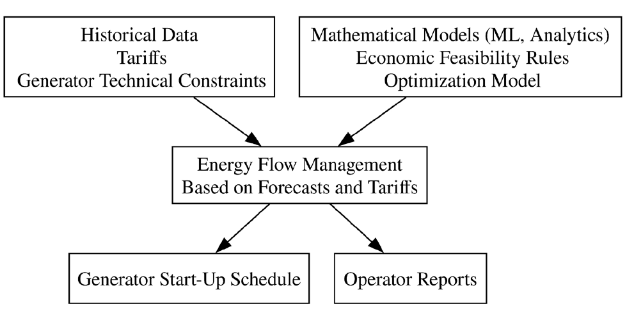

Fig. 1. Data flow diagram

The first part is responsible for collecting and preparing data. The system is integrated with a historical data source (Excel file), from which the values of previous energy consumption are read. Next, the data is checked for completeness and integrity, possible anomalies are eliminated, and normalization is performed, necessary for further operation of machine learning algorithms.

The second module of the system forecasts electricity consumption. A forecast is built for a given period using data from the previous step.

The next stage is to obtain information about tariffs. The system is connected to an external price source. The work uses a web resource with current electricity prices, which automatically collects and processes tariff data. The data make it possible to take into account the dynamics of energy costs when making decisions on the economic feasibility of using backup power sources [16].

The fourth part is responsible for optimizing the generator switching schedule. To do this, based on the forecast and tariffs, several alternative scenarios for switching on generators and the time of power supply from the external network are formed. Each scenario is evaluated for economic feasibility and compliance with technical constraints. Finally, the most advantageous option is selected.

The last part forms the results for the user. The work schedule is represented in the form of visualization, and the result itself is stored in the database. The schedule can also be provided in the form of messages for the enterprise's power system dispatcher.

The described sequence can be depicted in the form of a function tree, where the top is the global objective, and each lower level details individual parts of the implementation. This approach allows one to logically structure the system for the further implementation of individual logical modules. This is useful for simultaneous forecasting at several enterprises or adding alternative sources of generation.

5. 2. Results of constructing a mathematical model for forecasting electricity consumption of an enterprise

The mathematical model of forecasting is a functional mapping $F:\boldsymbol{X}\to\mathbb{R}$. This mapping correlates the vector of input features $X\in\boldsymbol{X}$ with the predicted power value $\widehat{P}_{t+1}$ for the next forecasting step

$$\widehat{P}_{t+1}=F\left(X,\Theta\right), \tag{5}$$

where Θ is the vector of model parameters.

At the design stage, two model configurations were formed.

For the model using the Random Forest method, a feature vector X is defined

$$X=\left\{x_{hour},x_{dw},x_{lag1}\right\}, \tag{6}$$

where x_{hour} is the hour of the day, x_{dw} is the day of the week, x_{lag1} is the first-order lag feature corresponding to the load of the previous hour.

Taking into account the properties of LightGBM, an extended feature vector X was developed for the model based on this method, which makes it possible to take into account seasonal and calendar dependencies

$$X=\left\{x_{hour},x_{dw},x_{month},x_{lag1},x_{lag24},x_{last\_y},x_{is\_w}\right\}, \tag{7}$$

where x_{hour} is the hour of the day, x_{dw} is the day of the week, x_{month} is the month, x_{lag1} is the lag feature corresponding to the load of the previous hour, x_{lag24} is the lag feature corresponding to the load for the previous 24 hours, x_{last_y} is the value for the same day of the previous year, x_{is_w} is the weekend indicator.

The results of the comparison in percentages of the average of the two models for load forecasting based on the *MAE* (3) and *RMSE* (4) metrics are given in Table 1.

Table 1

Comparison of model prediction accuracy

Model	*MAE*	*RMSE*
Random Forest regression	8–10%	12–15%
LightGBM regression	5–7%	10–12%

For both models, based on numerous experiments conducted for different days and times of day using bootstrapping, the efficiency was estimated as a percentage of the average load value.

5. 3. Results of development of an algorithm for optimal control over energy sources of an enterprise

To make economically sound decisions that take into account technical requirements for equipment, the proposed model uses the MILP apparatus. This approach makes it possible to simultaneously minimize daily operating costs while adhering to critical constraints, in particular, the minimum time of continuous operation of generator $T_{\min}$.

The optimization task is to determine the operating mode of autonomous power sources and the purchase of energy from the day-ahead market (DAM) for each hour $t\in\{1, 2, ..., 24\}$.

The model uses the following variables that determine the optimal schedule: $E_{gen}^{act}\left(t\right)$ (kWh) – actual generated energy from generators at hour t; $E_{mo}^{act}\left(t\right)$ (kWh) – actual purchased energy from DAM at hour t, $N(t)$ (pcs.) – number of generators started at hour t; $Z(t)\in\{0,1\}$ – binary indicator of generator operation (1 – operating, 0 – off); $Y(t)\in\{0,1\}$ is a binary indicator of generator start-up at hour t, $I(t)\in\{0,1\}$ is a binary indicator of network availability, E_{excess} (kWh) is the excess energy generated by generators at hour t.

The input parameters are the predicted load $P(t)$ for hour v obtained using LightGBM, the hourly DAM tariff $p_{mo}(t)$, the generator power P, the maximum number of generators $N_{\max}$, the minimum generator operating time $T_{\min}$, and the specific generation cost C_{gen} (UAH/kWh). The specific generation cost is defined as

$$C_{gen}=K_{gas}\cdot P_{gas}, \tag{8}$$

where K_{gas} is the specific gas consumption of the generator (for example, 0.25 m^3/kWh), and P_{gas} is the gas price per unit volume (for example, 15 UAH/m^3).

The objective function $\mathbb{C}\left(t\right)$ minimizes the total costs during the day, taking into account operating costs and fixed start-up costs C_{start}

$$\min\mathbb{C}\left(t\right)=\sum_{i=1}^{24}\left[\begin{array}{l}E_{gen}^{act}\left(t\right)\cdot C_{gen}+\\+E_{mo}^{act}\left(t\right)p_{mo}\left(t\right)+C_{start}Y\left(t\right)\end{array}\right]. \tag{9}$$

To ensure the physical and technical correctness of the solution, the MILP model must satisfy a number of key constraints.

The first constraint is due to the fact that under real conditions generators usually operate only at full capacity. This means that it is possible to obtain a generation surplus, which can be either a real surplus or a formal surplus associated with forecast inaccuracy. In each case, this must be taken into account in the model. For this purpose, the surplus variable $E_{excess}\geq 0$ (kWh) is used, which we take into account in the additional constraint on the total energy

$$E_{gen}^{act}\left(t\right)+E_{mo}^{act}\left(t\right)=P\left(t\right)+E_{excess}\left(t\right),\forall t. \tag{10}$$

The second constraint is a generation limit, which limits the actual generation, which cannot exceed the nominal power of the running generators, and their number is limited by the available number $N_{\max}$

$$E_{gen}^{act}\left(t\right)\leq N\left(t\right)\cdot P,\ N\left(t\right)\leq N_{\max},\forall t. \tag{11}$$

The third constraint is the minimum generation level, to ensure that the equipment operates at an effective load and avoid obtaining low efficiency, if the generators operate at $Z(t)=1$, they must generate the minimum required amount of energy $E_{\min}^{start}$

$$E_{gen}^{act}\left(t\right)\geq Z\left(t\right)\cdot E_{\min}^{start},\forall t. \tag{12}$$

The fourth constraint concerns network availability, modeled by a binary network availability indicator $I(t)$, which limits energy purchase $E_{mo}^{act}\left(t\right)$ in the event of a power outage

$$E_{mo}^{act}\left(t\right)\leq P\left(t\right)\cdot I\left(t\right),\forall t. \tag{13}$$

The latter sets the technical operation limits, namely the minimum operating time $T_{\min}$ and startup costs using binary indicators $Z(t)$ and $Y(t)$.

The relationship between the quantity and the operation indicator is that if the generators are operating, then $Z(t)=1$, and if not – $Z(t)=0$, then $N(t)=0$

$$N\left(t\right)\leq N_{\max}\cdot Z\left(t\right),\forall t. \tag{14}$$

The start onset $Y(t)$ is set to 1 only at the hour when the transition from the off state $Z(t) = 0$ to the on state $Z(t) = 1$ occurs

$$Y(t) \geq Z(t) - Z(t-1),\ \forall t \in \{2,\ldots,24\},\ Y(1) = Z(1). \quad (15)$$

The $T_{\min}$ constraint is a key constraint that ensures that if the generators started at hour t ($Y(t) = 0$), they will remain on for the next $T_{\min}$ hours

$$\sum_{k=t}^{t+T_{\min}-1} Z(k) \geq T_{\min} \cdot Y(t),\ \forall t \in \{1,\ldots,24 - T_{min} + 1\}. \quad (16)$$

Note that when $C_{start} = 0$ the optimization function takes the following form

$$\min \mathbb{C} = \sum_{t=1}^{24} \left[E_{gen}^{act}(t) \cdot C_{gen} + E_{mo}^{act}(t) \cdot p_{mo}(t) \right]. \quad (17)$$

However, all technical constraints, including $T_{\min}$ and the associated $Z(t)$ and $Y(t)$, must be maintained.

Using this MILP allows the system to globally plan the operation schedule for 24 hours, deciding to start generators during periods of high DAM tariffs. However, the technical requirements of continuous operation of $T_{\min}$ may force generators to operate even during hours with low tariffs.

In the next step, model (8)–(17) is represented as two interconnected algorithms.

Algorithm 1 (Fig. 3) describes this high-level workflow in detail, outlining how the system integrates a machine learning-based forecasting module with a deterministic optimization module to make daily strategic decisions.

The process is performed hourly or daily, using historical consumption data and current market prices as inputs to further form a globally optimal schedule for the next 24 hours. The LightGBM regression (2) is chosen as the forecasting model here, given its higher accuracy compared to Random Forest.

The central element of the control system is the optimization kernel, which uses MILP to determine the most economically advantageous schedule while strictly adhering to technical constraints. Algorithm 2 (Fig. 4) formalizes this mathematical problem. It defines a set of solution variables, the objective function of which is to minimize the total operating costs and a system of linear constraints that satisfies equations (6)–(17). This formulation preserves the binary logic for generator operation $Z(t)$ and start-up $Y(t)$, clearly ensuring compliance with the minimum operating time $T_{\min}$ requirement, thereby guaranteeing compliance with physical requirements along with economic feasibility.

The results of computer simulation of algorithm (6)–(17) using the forecast obtained using the LightGBM model at $C_{start} = 0$ for 15 hours are given in Table 2.

Algorithm 1 Generate Optimal Daily Schedule

Require: Historical_Data, Tariffs=$\{p_{mo}(t)\}$,
Tech_Params = $\{C_{gen}$, $\mathbf{P}$, N_{max}, T_{min}, E_{min}^{start}, $I(t)\}$
Ensure: Optimal Schedule $N(t)$, Cost $\mathbb{C}(t)$
$T \leftarrow 24$ ▷ Define time horizon
Features ← GENERATE_FEATURES(Current_Date, Historical_Data)
$P(t) \leftarrow$ ARRAY[T]
▷ Stage 2: Load Forecasting
for $t \leftarrow 1$ to T **do**
$P(t) \leftarrow$ LightGBM_Model.PREDICT(Features[t])
end for
▷ Stage 3: Optimization: MILP Solver
$\{N(t), E_{gen}(t), E_{mo}(t), \mathbb{C}(t)\} \leftarrow$ SOLVE_MILP_PROBLEM($P(t)$, Tariffs, Tech_Params)
▷ Stage 4: Output
return $\{N(t), E_{gen}(t), E_{mo}(t), \mathbb{C}(t)\}$

Fig. 3. Pseudocode for generating an optimal schedule

Algorithm 2 MILP Problem Formulation: Optimization Core

Require: $P(t)$, Tariffs=$\{p_{mo}(t)\}$, Tech_Params=$\{C_{gen}$, $\mathbf{P}$, N_{max}, T_{min}, $E_{min}^{start}\}$
▷ 1. DECLARE VARIABLES
$E_{gen}^{act}(t), E_{mo}^{act}(t), E_{excess}(t), C_{start} \in \mathbb{R} \geq 0$
$N(t) \in \mathbb{Z} : 0 \leq N(t) \leq N_{max}$
$I(t), Z(t), Y(t) \in \{0, 1\}$
▷ 2. OBJECTIVE FUNCTION
MINIMIZE $\mathbb{C}(t) = \sum_{t=1}^{T} \left[E_{gen}^{act}(t) \cdot C_{gen} + E_{mo}^{act}(t) \cdot p_{mo}(t) + C_{start} \cdot Y(t)\right]$
▷ 3. DEFINE CONSTRAINTS $\forall t \in [1, T]$
SUBJECT TO CONSTRAINTS:
C1: Energy Balance $E_{gen}^{act}(t) + E_{mo}^{act}(t) = P(t) + E_{excess}(t)$
C2: Max Generation Limit $E_{gen}^{act}(t) \leq N(t) \cdot \mathbf{P}$
C3: Min Generation Level $E_{gen}^{act}(t) \geq Z(t) \cdot E_{min}^{start}$
C4: Grid Availability $E_{mo}^{act}(t) \leq P(t) \cdot I(t)$
C5: Link N(t) and Z(t) $N(t) \leq N_{max} \cdot Z(t)$
▷ Startup and Minimum Run Time Logic
C6a: Startup Y(1) $Y(1) = Z(1)$
for $t \leftarrow 2$ to T **do**
C6b: Startup Y(t) $Y(t) \geq Z(t) - Z(t-1)$
end for
for $t \leftarrow 1$ to $T - T_{\min} + 1$ **do**
C7: Min Run Time $\sum_{k=t}^{t+T_{\min}-1} Z(k) \geq T_{\min} \cdot Y(t)$
end for
▷ 4. SOLUTION
RETURN SOLVER.FIND_OPTIMAL_SOLUTION(OBJECTIVE, CONSTRAINTS)

Fig. 4. Formulation of mixed-integer linear programming in pseudocode

Table 2

Results of software simulation of the optimization algorithm

t	$P(t)$	$N(t)$	$E_{gen}^{act}(t)$	$C_{gen}^{act}(t)$	$E_{mo}^{act}(t)$	$p_{mo}(t)$	$C_{tot}(t)$	$C_{ref}(t)$	$\Delta C(t)$
0	5605.38	5	6768.00	28200.00	0.00	4.70	28200.00	26345.95	–1854.05
1	7115.46	6	8121.60	33840.00	0.00	4.30	33840.00	30596.48	–3243.52
2	7403.22	0	0.00	0.00	7403.22	3.80	28132.24	28132.24	0.00
3	7403.22	0	0.00	0.00	7403.22	3.20	24391.49	24391.49	0.00
4	7590.00	0	0.00	0.00	7590.00	1.00	7690.05	7690.05	0.00
5	7588.02	0	0.00	0.00	7588.02	0.10	758.80	758.80	0.00
6	7597.92	6	8121.60	33840.00	0.00	5.00	33840.00	37989.60	4149.60
7	7302.24	6	8121.60	33840.00	0.00	6.80	33840.00	49655.23	15815.23
8	7227.00	6	8121.60	33840.00	0.00	6.80	33840.00	49143.60	15303.60
9	7393.98	6	8121.60	33840.00	0.00	4.49	33840.00	33184.18	–655.82
10	7464.60	6	0.00	0.00	7464.60	2.09	15586.08	15586.08	0.00
11	7535.88	0	0.00	0.00	7535.88	0.78	5855.38	5855.38	0.00
12	7481.10	0	0.00	0.00	7481.10	0.10	748.11	748.11	0.00
13	7420.38	0	0.00	0.00	7420.38	0.25	1855.10	1855.10	0.00
14	7503.54	0	0.00	0.00	7503.54	0.10	750.35	750.35	0.00

Here, $C_{gen}^{act}(t) = E_{gen}^{act}(t) \cdot C_{gen}$ (UAH) is the generation cost, $C_{ref} = P(t) \cdot p_{mo}(t)$ (UAH) is the theoretical cost with full procurement from DAM, $C_{tot}(t) = C_{gen}^{act}(t) + C_{mo}^{act}(t)$ (UAH) is the total cost, $C_{mo}^{act} = E_{mo}^{act}(t) \cdot p_{mo}(t)$ (UAH) is the real cost with full procurement from DAM, $\Delta C(t) = C_{ref}(t) - C_{tot}(t)$ (UAH) is the difference between the generation price and the price with full procurement from DAM.

6. Discussion of results related to the system for generating the optimal electricity supply schedule

The designed logical-information structure of the system provides a direct connection between the forecasting module and the optimization unit based on MILP. This allows for prompt correction of energy consumption plans. The use of DFD notation at the design stage allowed for the formalization of data flows between different sources, which minimizes delays in the formation of the vector of input parameters for machine learning models.

The results of the comparison of forecasting models are given in Table 1, showing that LightGBM produced better results in both the MAE (5–7%) and RMSE (10–12%) metrics, with smaller deviations compared to Random Forest. This indicates that the use of gradient boosting in LightGBM together with an expanded set of constructed features allows for good generalization of the data used for forecasting. The assessment of the forecasting results is comparable to the MAPE metric reported in [7] in the range of 8.18–8.64%. The error is somewhat higher than in [9], which, however, does not provide the value of the MAPE metric, but the achieved maximum accuracy of 96.8% indicates quantitatively better results. The obtained accuracy is also higher than in [8], in which the MAPE is approximately 20%. However, it should be noted that the data used in this work and the data from [7–9] are different in terms of the volume and characteristics of the processes to which they correspond.

The MILP apparatus used for the mathematical description of the process of controlling energy sources, in contrast to conventional control strategies based on threshold values, makes it possible to operate on variables to control the switching on and off of generators. This ensures a better correspondence of the obtained modes to real processes.

The logic of the algorithm allows the system not only to react to changes in the network but also proactively calculate the optimal use of resources in advance. Thanks to this, the system can make economically sound decisions about launching local generating capacities even before market price peaks.

The developed algorithm takes into account operational constraints. The introduction of parameters for the minimum duration of equipment operation makes it possible to avoid premature wear of generating capacities. Thanks to this, the algorithm balances between saving money and preserving the technical resource of the equipment. This makes the proposed solution suitable for implementation in the production sector where equipment reliability and direct economic benefit are important.

The algorithm, based on the predicted hourly load, strategically makes decisions on the choice of energy source in order to minimize the total costs of energy supply. The results of our work (Table 2) show that the use of local generators is optimal during periods of peak tariffs. In particular, when the purchase price on the market is the highest (for example, 6.80 UAH/kWh in the hours between 07:00 and 08:00), the generation strategy significantly reduces the total costs of energy supply for these hours. Conversely, during nighttime off-peak hours, the DAM tariffs are reduced to 0.10 UAH/kWh, making grid energy consumption a more economical choice.

The optimization of generator scheduling is performed at the global level by the MILP apparatus, which must comply with critical technical constraints, including the minimum continuous operation time $T_{\min}$ (11). This requirement sometimes leads to the fact that the generation schedule shows a negative value (in particular at 00:00, 01:00 and 09:00), which means a temporary excess of generation costs compared to the market tariff. This phenomenon is evidence of strategic, non-local decision-making by the model, which prioritizes compliance with the $T_{\min}$ constraint (11). This mechanism allows the system to achieve significant savings during peak consumption hours. This means that the total savings from such strategic scheduling significantly exceed periodic local losses.

The efficiency of the algorithm was assessed by comparing the costs of electricity when purchasing only from the market with the costs when operating a limited number of generators. Over the considered period of time, the balanced use of generating capacities and purchased energy from the market led to a reduction in costs by UAH 29515.04, which is 9.44% relative to the base costs. The result confirms the importance of optimization in managing electricity sources when tariffs fluctuate and problems with the availability of energy sources.

Despite the obvious optimization of costs under the established technical operating conditions, the proposed hybrid model has limitations and disadvantages. The main limitation is the dependence on the accuracy of the load forecast $P(t)$ for the day in advance. Although the use of LightGBM showed high forecast accuracy, forecast errors and instability can cause suboptimal solutions of the optimization system. Such forecast behavior can cause unjustified generator starts or excessive energy purchase from the market. Another drawback of the model is that although the RDM tariffs $p_{mo}(t)$ are determined hourly for 24 hours, they are considered as deterministic inputs at the time of optimization. This, in turn, limits the possibility of dynamic adjustment.

Further research could focus on extending the proposed model by integrating energy storage systems. This approach would allow for the storage of potentially excess electricity generated during periods subject to minimum operating time restrictions. Theoretically, this approach could increase the overall economic efficiency of the system beyond what is already achieved. In addition, it is possible to consider the possibility of implementing power modulation algorithms for generator sets instead of binary indicators to obtain a more accurate balance of energy demand and consumption. However, further complicating the model with additional variables would significantly increase the computational complexity required to find the optimal solution.

7. Conclusions

1. The information-logical structure and architecture of the enterprise energy consumption management information system have been developed. The proposed structure describes an automated cycle from data collection to load forecasting and decision-making on optimal management of electricity sources. Formalization of the architecture in the form of structural diagrams has made it possible to clearly define the boundaries of interaction between the forecasting model and the optimization algorithm.

2. The structure of the forecasting model has been designed, which uses a wide set of constructed features. The model built makes it possible to model complex nonlinear dependencies in energy consumption data. Experiments have shown that the LightGBM model demonstrated accuracy, as a percentage of the average, of 5–7% by the *MAE* metric and 10–12% by the *RMSE*. Such accuracy is sufficient for further use in the optimization algorithm.

3. An optimal control algorithm has been developed, which, based on the predicted load, makes it possible to automate the choice between energy consumption sources and determine the optimal number of running generators if necessary. The algorithm takes into account not only the technical limitations of the generators but also the dynamic change in the cost of electricity in real time. Based on the experiments, it is shown that the control algorithm achieves global cost optimization while adhering to the technical constraints. The algorithm uses market tariffs during hours when they are low and switches to local energy generation during hours of peak tariffs. The evaluation of the algorithm's effectiveness demonstrated a significant economic benefit over the considered time period, which led to a cost reduction of 9.44% compared to the theoretical base costs. The decision-making system includes, in particular, local cost overruns caused by the need to strictly adhere to operational constraints. Our result confirms the financial effectiveness of the proposed optimization system in managing electricity costs under variable tariffs and unstable electricity supply.

Conflicts of interest

The authors declare that they have no conflicts of interest in relation to the current study, including financial, personal, authorship, or any other, that could affect the study, as well as the results reported in this paper.

Funding

The study was conducted without financial support.

Data availability

The data will be provided upon reasonable request.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technologies when creating the current work.

Authors' contributions

Pavlo Shevchyk: conceptualization, data curation, formal analysis, investigation, methodology, software, resources; **Polovyi Vitalii**: conceptualization, methodology, format analysis, supervision, validation, writing – original draft, writing – reviewing & editing; **Yana Ni**: visualization, supervision, validation, writing – reviewing & editing.

References

1. DiXi Group calculated mass attacks on the energy system of Ukraine (2024). Ukrainian Energy. Available at: https://ua-energy.org/en/posts/03-10-2024-0c9b86dc-4300-49c0-afb3-307d43ca6d1d
2. Bytva za svitlo. Zvit pro diyalnist (2024). DTEK. Available at: https://dtek.com/content/upload/report/DTEK%202024%20Action%20Report%20Ukrainian%20250221.pdf

3. Otsinka priamykh zbytkiv ta nepriamykh vtrat enerhetychnoho sektoru Ukrainy vnaslidok povnomasshtabnoho vtorhnennia rosii (2024). Kyivska shkola ekonomiky. Available at: https://kse.ua/wp-content/uploads/2024/06/KSE_Vpliv-vii--ni-na-energetiku_UA-1.pdf
4. A Rebirth in Flame: Ukraine's Beleaguered Energy System (2025). CEPA. Available at: https://cepa.org/article/a-rebirth-in-flame-ukraines-beleaguered-energy-system
5. Kumar, M., Pal, N. (2023). Machine Learning-based Electric Load Forecasting for Peak Demand Control in Smart Grid. Computers, Materials & Continua, 74 (3), 4785–4799. https://doi.org/10.32604/cmc.2022.032971
6. Halimuzzaman, M. (2025). AI-driven optimization of hybrid renewable energy systems: A review of techniques, challenges, and future direction. Pacific Journal of Advanced Engineering Innovations, 2 (1), 22–32. https://doi.org/10.70818/pjaei.v02i01.093
7. Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., Zhang, Y. (2019). Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network. IEEE Transactions on Smart Grid, 10 (1), 841–851. https://doi.org/10.1109/tsg.2017.2753802
8. Fayyazbakhsh, A., Kienberger, T., Vopava-Wrienz, J. (2025). Comparative Analysis of Load Profile Forecasting: LSTM, SVR, and Ensemble Approaches for Singular and Cumulative Load Categories. Smart Cities, 8 (2), 65. https://doi.org/10.3390/smartcities8020065
9. Ibrar, M., Hassan, M. A., Shaukat, K., Alam, T. M., Khurshid, K. S., Hameed, I. A. et al. (2022). A Machine Learning-Based Model for Stability Prediction of Decentralized Power Grid Linked with Renewable Energy Resources. Wireless Communications and Mobile Computing, 2022 (1). https://doi.org/10.1155/2022/2697303
10. Mudgal, Y., Tiwari, R., Krishnan, N., Téllez, A. A. (2025). Advanced Cluster-Based Load Forecasting and Peak Demand Management for Electric Vehicle Charging Networks. IEEE Access, 13, 105664–105677. https://doi.org/10.1109/access.2025.3578875
11. Bahrami, S., Chen, Y. C., Wong, V. W. S. (2021). Deep Reinforcement Learning for Demand Response in Distribution Networks. IEEE Transactions on Smart Grid, 12 (2), 1496–1506. https://doi.org/10.1109/tsg.2020.3037066
12. Suanpang, P., Jamjuntr, P., Jermsittiparsert, K., Kaewyong, P. (2022). Autonomous Energy Management by Applying Deep Q-Learning to Enhance Sustainability in Smart Tourism Cities. Energies, 15 (5), 1906. https://doi.org/10.3390/en15051906
13. Danylchenko, D. O., Fedorchuk, S. O., Maslak, M. V. et al. (2024). Intelektualni systemy elektropostachannia na osnovi zaluchennia aktyvnykh spozhyvachiv. Kharkiv: NTU "KhPI", 310. Available at: https://repository.kpi.kharkov.ua/entities/publication/1511a424-3c71-467d-a8f6-9ad66d58ae33
14. Sirenko, Yu. V., Volvach, T. S., Savoiskyi, O. Yu., Kozin, V. M. (2025). Analysis of the state of ukraine's energy system and measures to improve the situation. Visnyk of Kherson National Technical University, 1 (2 (93)), 229–237. https://doi.org/10.35546/kntu2078-4481.2025.2.1.30
15. Bobrov, O., Yaroshenko, Y. (2024). Analysis of electrical load schedules of UPS of Ukraine. Electrical and Information Systems, 106, 27–34. https://doi.org/10.32782/eis/2024-106-5
16. González Grandón, T., Schwenzer, J., Steens, T., Breuing, J. (2024). Electricity demand forecasting with hybrid classical statistical and machine learning algorithms: Case study of Ukraine. Applied Energy, 355, 122249. https://doi.org/10.1016/j.apenergy.2023.122249
17. Petko, S. (2024). Application of AI in the Energy Sector of Ukraine in the Conditions of War. Challenges to national defence in contemporary geopolitical situation (CNDCGS'2024). Brno, 22–26. Available at: https://ir.kneu.edu.ua/items/9520f314-cf25-4a12-9369-29155c54d6d4
18. Breiman, L. (2001). Random Forests. Machine Learning, 45 (1), 5–32. https://doi.org/10.1023/a:1010933404324
19. Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., Lawler, J. J. (2007). Random forests for classification in ecology. Ecology, 88 (11), 2783–2792. https://doi.org/10.1890/07-0539.1
20. Liaw, A., Wiener, M. (2002). Classification and regression by randomForest. R News, 2, 18–22. Available at: https://journal.r-project.org/articles/RN-2002-022/RN-2002-022.pdf
21. Hastie, T., Tibshirani, R., Friedman, J. (2009). The Elements of Statistical Learning. In Springer Series in Statistics. Springer New York. https://doi.org/10.1007/978-0-387-84858-7
22. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W. et al. (2017). LightGBM: A highly efficient gradient boosting decision tree. NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. Available at: http://dl.acm.org/doi/10.5555/3294996.3295074
23. Bertsimas, D., Tsitsiklis, J. N. (1997). Introduction to linear optimization. Athena Scientific, 587.
24. 1320.2-1998 – IEEE Standard for Conceptual Modeling Language – Syntax and Semantics for IDEF1X97 (IDEFobject). Available at: https://ieeexplore.ieee.org/document/8883275
25. Yourdon, E. (1989). Modern structured analysis. Englewood Cliffs: Prentice Hall, 688.