Human Technology
https://ht.csr-pub.eu
ISSN: 1795-6889
Volume 21(2), September 2025, 431–447

OPERATIONAL HVAC ENERGY LOAD PREDICTION: EDGE-ORIENTED FORECASTING MODELS

Arkadiusz Tomczyk
Lodz University of Technology
Poland
ORCID 0000-0001-9840-6209

Wiktor Wojciechowski
Warsaw School of Economics
Poland
ORCID 0000-0002-4235-2739

Jakub Walczak
Lodz University of Technology
Poland
ORCID 0000-0002-5632-9484

Piotr Lipiński
Lodz University of Technology
Poland
ORCID 0000-0002-8269-3802

Agnieszka Wosiak
Lodz University of Technology
Poland
ORCID 0000-0001-6124-1236

Michał Morawski
Lodz University of Technology
Poland
ORCID 0000-0002-8902-1259

Piotr Napieralski
Lodz University of Technology
Poland
ORCID 0000-0003-1427-7791

Abstract: *Contemporary office building standards require the installation of advanced HVAC systems, which, alongside industrial processes, are significantly responsible for energy consumption. While the literature describes numerous multi-parameter, deep artificial neural networks that are highly complex but poorly explainable (black box), commercial control systems must be based on simple-to-implement, easily adaptable intelligent predictive models whose reliability stems from their monitorability and explainability. The study verified several edge-oriented approaches using machine learning techniques for short- and medium-term predictions (1 day) with relatively high granularity (15 minutes). High granularity of time intervals, combined with high prediction reliability, not only provide a practical opportunity to optimize energy consumption, but also to increase the efficiency of renewable energy sources exploitation. Operational, high granularity predictions can also be used to detect anomalies and attacks, as required by modern cybersecurity rules.*

Keywords: *operational prediction, HVAC control, energy load forecast, edge computing, anomaly detection, cybersecurity.*

DOI: https://doi.org/10.14254/1795-6889.2025.21-2.10

INTRODUCTION

Problem Statement

In the Polish electricity sector, households account for only around 22% of total electricity consumption in 2024 (Polish Energy Regulatory Office, 2024). Since enterprises are responsible for the majority of demand, addressing their energy use is of critical importance. Recent macro-level studies confirm that understanding the socio-economic drivers of electricity demand is fundamental for accurate load-forecasting frameworks (Hlongwane, 2025; Krutsyak, 2019). A particular challenge and, at the parallel, the most anticipated scenario, lies in synchronizing energy demand with renewable energy supply. Nevertheless, the fluctuation of electricity demand throughout the day, combined with the unpredictable, weather-dependent (sun, wind, etc.) nature of renewable energy sources, make it difficult to manage and build effective interdependence without precise monitoring and planning of demand and supply processes. This issue considerably affects GHG emission and overall costs due to two main reasons. First, given the dynamic structure of Poland's energy mix, shifting demand or surpassing demand can allow for better correlation with green energy supply and contribute to a meaningful reduction in the carbon footprint. Second, aligning and balancing electrical loads can substantially lower electricity costs for enterprises. It is not surprising that energy producers and suppliers set tariffs in such a way as to counteract unexpected changes in supply and demand, which cause power grid overload and can lead to a complete power system failure (power outages). These aspects merit closer examination.

Poland remains among the European countries with the highest reliance on coal in its energy mix (International Energy Agency, 2024). However, the energy mix is not constant throughout the year, or even within a single day. Variations are primarily the result of changing contribution of renewable energy sources (RES), whose output fluctuates with daily and weather-related conditions. Consequently, conscious and well-planned balancing of energy demand throughout the day could lead to a substantial decrease in greenhouse gas emissions. By planning periods of increased RES supply and simultaneously planning load shifts to periods of increased RES availability, it is possible to achieve ecological synergies between green energy supply and demand. While certain operational regimes of enterprises are difficult or even infeasible to modify, significant opportunities remain. In particular, systems with high thermal inertia, such as heating, ventilation, and air conditioning (HVAC), offer considerable flexibility for demand-side management (Aslam, 2024). The global warming observed around the world means that air conditioning systems are becoming particularly common and increasing the load on energy networks. The development of alternative heat sources (heat pumps, etc.) is also shifting the load from coal and gas to electricity, which, with proper management, can be produced using renewable sources. In addition, the protection of company data must be taken into account, especially if it concerns energy consumption in various industrial processes. Data is a key company resource that must be strictly protected. Therefore, processing data on low-level edge devices rather than sending it to cloud servers seems to be a safer and more acceptable solution for businesses. However, edge devices have limited resources and low computing power, which eliminates the use of large neural network models, even if they exhibit good predictive properties. The question remains, however, to what extent uncomplicated machine learning techniques allow for the prediction of high-granularity data.

Therefore, in this article, we argue that a reliable multi-step short-term (operational) electricity consumption forecast can equip small and medium-sized enterprises (SMEs) with simplified predictive models that do not require large resources and offer high operational and optimization efficiency. These, in turn, can not only reduce costs but also reduce the carbon footprint thanks to rationalization of energy consumption by exploiting the thermal dynamics of a building and shifting load of HVAC system. We validate this proposition through a simulation study of sample HVAC electricity consumption in an office building located in central Poland.

Related Works

HVAC systems have been extensively studied by researchers in the context of both cost reduction and environmental impact mitigation. In modern buildings, they exhibit relatively high thermal inertia, which makes them well-suited for optimization and load shifting within building operation regimes. When thermal building dynamics are taken into account, optimizing power consumption can yield significant economic and environmental benefits.

The thermal dynamics of buildings can be modeled using three principal approaches (Balali, 2023):

1. **White-box approaches** — physics-based models such as detailed thermal dynamics equations (Xu, 2008) or simplified resistance–capacitance (RC) networks (Fraisse, 2002). These methods require extensive knowledge of both the building and the HVAC system.
2. **Black-box approaches** — data-driven models that discover relationships useful for forecasting, typically relying on large datasets. These methods disregard underlying physics but depend heavily on the availability of sufficient data.
3. **Grey-box approaches** — hybrid methods that combine physical knowledge with data-driven techniques, offering a compromise between model transparency and data requirements.

Since obtaining comprehensive and detailed information about a building is rarely feasible, thermal dynamics are often modeled using black-box or grey-box approaches. This is mainly due to the fact that the physical parameters of a building are used during its design, whereas after construction, the physical conditions may change and, what is more, the characteristics of a building change over time and depending on how it is used, which does not always correspond to the assumptions. Provided that the thermal inertia of a building (including HVAC response delays) can be modeled with sufficient accuracy, the HVAC central unit can anticipate future demand driven by occupants or outdoor conditions. This enables proactive control strategies that significantly reduce heating and cooling costs (Verbeke, 2018). When coupled with an environmentally friendly energy mix, such approaches can also lower the environmental footprint of HVAC operations. This principle lies at the core of Smart Grids and demand-side management in electricity networks.

Beyond the direct benefits to HVAC system owners, large-scale implementation helps smooth the electricity demand–supply curve, thereby stabilizing the grid. A stable grid is more predictable, reducing the need for costly on-demand generation and preventing avoidable GHG emissions (Koebrich, 2025). Hence, precise modeling of thermal dynamics is critical. Using grey- and black-box approaches, the dynamics can be reasonably captured by combining outdoor conditions (such as air temperature, solar irradiation, and wind) with the power consumption of the HVAC system. Olu-Ajayi et al. (2025) point out that, in addition to weather conditions, the authors attempt to take into account the occupancy of the building, which intuitively seems to be a significant modelling factor. Since weather variables can typically be forecasted with acceptable accuracy over short- to medium-range horizons (Lam, 2023), the main challenge lies in forecasting the electricity consumption of the HVAC central unit multiple steps ahead. Characteristics of the HVAC control systems show that forecasts should have adequate granularity, both in the short and mid-term horizons. However, the definition of adequate granularity is imprecise, even in the domain of energy forecasting. Although formally adequate granularity predictions are located between seconds and days (Chen, 2022), the vast majority of authors (Olu, 2023) understand it as hourly or even longer time windows what diminish practical application of research. The need to shorten prediction granularity was raised, among others, by (Amasyali, 2018; Wang, 2017). It is worth noting, however, that demand response or operational system optimization requires rapid computational iterations of control algorithms. In addition, energy suppliers in Europe use a quarter of an hour as the billing unit, integrating instantaneous energy consumption characteristics from a fifteen-minute interval, following European legislation in the minimum operational granularity of energy market. These arguments clearly indicate that a thorough analysis of the ability of operational, 15 minutes granularity, short and mid-term forecasting models to predict energy (power) consumption should be carried out, which is the contribution of this study. It is worth verifying the forecasting capabilities not only for the nearest forecasting horizon but also examining the accuracy of medium range forecasts — several hours up to one day. It should be noted that among the many available forecasting models and techniques, we will pay particular attention to solutions that are simple and resource-efficient, as implementing solutions industrially, while maintaining appropriate security standards, requires their installation in a computing fog that does not have the same resources as typical servers. Addressing this challenge represents a pivotal milestone toward practical reducing GHG emissions. That also provides SMEs with sufficient time to plan and respond to changing circumstances efficiently.

One of the most elementary forecasting methods is the persistence model, which is widely used as a benchmark for comparison with more sophisticated approaches. It is defined as in (1), meaning that the forecasted values are identical to the most recently observed ones (Jung, 2014).

$$\overline{P(t+h)} = P(t), \tag{1}$$

where $\overline{P(t+h)}$ is a forecast at time *t* made *h* steps ahead (i.e., forecast of a value observed at time t+h made at time *t*) and $P(t)$ is an observed value at the time t.

Although extremely simple, in some cases it is sufficient. Its strength lies in computational simplicity, but its accuracy depends strongly on the characteristics of the time series being modeled. For short-term forecasts in stable environments (e.g., monotonic, smooth, or slowly varying series), the persistence model is often adequate (Das, 2018). For forecasting steady-system active energy consumption, short-term persistence forecasts can be valid, whereas for highly oscillatory power signals in real and dynamic systems, the method rarely performs well and predictions it does for multi-step forecasting usually does not reflect the reality. Though its performance is highly constrained, we selected persistence model as one of the reference forecasting methods, following the common practice in the domain.

To overcome the limitations of the elementary persistence model, many statistical approaches have been developed. Among the most widely applied are the autoregressive integrated moving average model (ARIMA) and its seasonally extended variant, SARIMA (Szostek, 2024; Kumari, 2024). Both models are linear and autoregressive, relying on lagged values and lagged forecast residuals to improve forecasts. However, they also have important limitations. In addition, both ARIMA and SARIMA demand careful parameter selection, which can be challenging to automate effectively. ARIMA has been one of the traditional models used for electricity consumption forecasting; nevertheless, both the basic and seasonal variants can sometimes produce unsatisfactory results when the data exhibit strong trends or irregular seasonal patterns (Klyuev, 2022).

Among other popular data-driven artificial intelligence models, two widely applied and successful approaches are based on random forests (Ho, 1995) and, even more successfully, on gradient boosting regressors (Friedman, 2001). Both methods follow a supervised, ensemble-based machine learning paradigm (Afzaal, 2023). They have been extensively applied in forecasting tasks, particularly in climate modeling (Khan, 2024) and electricity consumption forecasting (Di Persio, 2023; Magalhães, 2024). A major advantage of these techniques is the transparency of parameters and reasonable size, allowing models to be installed on edge devices.

Even more complex and sophisticated methods are based on artificial neural networks (ANNs), which have been widely adopted across industries and research domains. Even in their simplest form, the two-layer nonlinear perceptron, they offer remarkable flexibility and expressive power proved by means of universal approximation theorem (Selmic, 2002) for smooth functions. However, this potential comes at the cost of complexity, since ANNs involve numerous hyperparameters that must be carefully optimized, even though that optimization (in many cases) can be automatized.

The popular approach to predicting energy consumption time series is based on LSTM networks. Sendra et al. (Sendra, 2020) address the issue of next-day energy demand for HVAC systems, which significantly supports Demand Side Management in Smart Grids. The most relevant example of forecasts, reaching 24 hours with 30-minute granularity, is the paper by Fan (2019). The authors verify RNN, LSTM, and GRU recurrent models for predictive control solutions for intelligent HVAC system operations. The main contribution was the use of the better properties of modelling dependencies in time series by LSTM and GRU models and reducing the impact of the last prediction error, which was propagated recursively to subsequent prediction horizons in reference solutions. As a result, more distant forecasts were increasingly further from the truth. Although the authors contest the usefulness of feed forward networks (FFN) or traditional machine learning methods in similar tasks, they focus exclusively on recurrent models and do not compare them with FFN. That requires skepticism about the superiority of developed RNNs methods with respect to FNN, especially taking into consideration higher sensitivity of RNN to hyper-parameters selection than FNNs. The paper by Song (2024) provides a comprehensive overview of deep learning methods for time series prediction, using energy consumption data as an example. Several dozen AI models, including statistical, recurrent, convolutional, transformer-type, and fully connected networks, were compared for predictions reaching several days ahead, but with one-hour granularity. In addition, attention was drawn to computational and operational complexity, which is an obstacle to learning and adapting predictive models in a commercial, specifically in the context of edge-oriented deployments. To diminish the importances of both obstacles, we selected a multi-layered perceptron as ANN representative for multi-step time series forecasting. It provides less parameters resulting in shorter training and retraining (for online procedures) phases. Furthermore, it comes with easier and more robust tuning procedures and, what is essential for commercial automatic and semi-automatic systems, it provides more interpretability than recurrent ANN.

METHODS

To ensure credibility of experimental studies and simulation, we partially adapted the three-pillar methodological framework proposed by Walczak (2025) to focus on core elements: open dataset, complementary quality measures, and adequate experimentation protocol.

Used Datasets

In our studies, we used the custom dataset acquired from main HVAC device operating in an office building located in Lodz (central Poland). The dataset contains four electricity characteristics all of which are described in detail in Tab. 1. However, in further studies, we only use two basic characteristics of active power, namely that which reflects the actual operational load on the network and active energy as the total energy consumption.

Table 1. Characteristics of the Dataset Used in Simulation Studies

Name	Unit
Active energy	kWh
Reactive energy	kVarh
Active power	kW
Apparent power	VA

The entire dataset covers the period between 1st July 2023 and 30th June 2025 with the temporal resolution of one minute. Measured values have been interpolated by monitoring system to align with full minute time index (e.g. 22:10 instead of 22:09:47). The dataset comes with the predefined train-test split following blocking strategy (Roberts et al., 2017) suitable for data with temporal dependencies: training split encompass data between July 2023 until the end of December 2024, the remaining contiguous part constitutes the test split.

Particular attention was paid to forecasting the active power characteristic, as it represents the instantaneous electricity demand, which makes it extremely useful for operational demand response and facilitating grid stability. Furthermore, unlike energy, active power is not monotonic, it exhibits peaks and ramps, which makes forecasting more challenging.

Adopted Quality Measures

In our studies, we carried out an extensive comparison of reference forecasting methods, using complementary quality measures elaborated in this section. The first measure applied was the Mean Absolute Error (MAE), defined as

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|P(t) - \overline{P(t)}\right|, \qquad (2)$$

where $P(t)$ denotes the observed value of the time series at time t, and $\overline{P(t)}$ is the forecasted value at the same time, and n is the length of the time series.

MAE is more robust to outliers than the Mean Squared Error (MSE), and therefore results tend to be more stable. However, MAE (like MSE) is unbounded, and its magnitude alone does not directly indicate whether a forecast is accurate or note. To address this limitation, we followed Chicco (2021) and additionally used the coefficient of determination R^2 (3) and the symmetric Mean Absolute Percentage Error (4), sMAPE, which are more informative and bounded. Specifically, R^2 is bounded within $(-\infty, 1]$, whereas sMAPE — $[0, 2]$. It is important to note that MAE and sMAPE represent error measures, thus, lower values indicate better forecasts. In contrast, R^2 reflects the proportion of explained variance, so higher values indicate better forecasts.

$$R^2 = 1 - \frac{\sum_{t=1}^{n}\left(\overline{P(t)} - P(t)\right)^2}{\sum_{t=1}^{n}(P(t) - \dot{P})}, \qquad (3)$$

where $\dot{P}$ denotes the arithmetic average of the observed values.

The symmetric Mean Absolute Percentage Error is given by (4):

$$sMAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{\left|P(t)-\overline{P(t)}\right|}{0.5\cdot\left|P(t)+\overline{P(t)}\right|}. \tag{4}$$

Data Clearing

After carefully examining the dataset, we identified suspicious values in the power readings. Considering the nominal power of the HVAC central unit from which the data was collected, the time series displays two peaks that exceed the valid operational range. The values from these peaks surpass the maximum power declared by the manufacturer by an order of magnitude. Therefore, as these values are more likely result of monitoring system malfunction rather than genuine anomalous observations, we trimmed the portion of data between 5th and 16th May, treating this segment of the characteristic as missing values. Missing data can be supplemented with interpolation or extrapolation techniques, but for the purposes of the experiments described, only recorded data was used.

Forecasting Method

In this section, we present the approaches employed for predicting the values of active power and active energy h steps ahead, where one step corresponds to a 15-minute interval (resolution of time series, aligned with European policy for electricity billing). The experiments considered forecasting horizons of h = 1, 4, 8, 48, and 96 (i.e., 15 minutes, 1 hour, 2 hours, 12 hours, and 24 hours ahead, respectively). Two backward window lengths were examined: L = 96 (one day) and L = 480 (five days), both with a 15-minute sampling granularity. The latter case was included to allow the model to capture weekly patterns. Five days are sufficient because the network can infer the day of the week from the observed sequence of values (e.g., after five working days comes Saturday, after three working days followed by two non-working days comes Monday, etc.). A separate model was fitted for each (L, h) pair. To select the optimal model (i.e., the optimal hyperparameters), the training dataset was partitioned into a training subset and a validation subset in an 80% to 20% ratio, and the generalization performance was assessed on the validation set.

As a reference, a naive persistence model (eq. 1) was used for both active power and energy. It forecasts the value observed at time t as the prediction for t+h. Such a baseline is often competitive at very short horizons and whenever clear cyclic patterns align with the horizon h (e.g., quarter-hourly or daily repetitions). It does not account for trends, regime changes, or exogenous shocks, so its accuracy typically deteriorates as h grows or conditions shift.

The first trained model was a two-layer perceptron (MLP), as it constitutes a universal approximator when equipped with a sufficiently large number of neurons in the hidden layer. In the experiments, different numbers of hidden neurons (the K hyperparameter) were considered: 1024 and 2048, to ensure adequate flexibility and to avoid potential underfitting. The number of network inputs was equal to the number of historical measurements used for prediction (the L hyperparameter). Two training strategies were adopted. In the first, termed *single*, the network predicted only a single value corresponding to the specified forward step and therefore had a single output. In the second, termed *multi*, the network predicted all intermediate samples up to the specified forward horizon, with the final value taken as the prediction. This approach was intended to investigate whether the additional intermediate prediction tasks could improve performance (serving as a form of regularization). In this case, the network had h outputs. It should be noted that for h = 1, both strategies are equivalent.

To eliminate the influence of potential fluctuations in the time series and to normalize the network inputs, thereby facilitating more effective training, the input data was standardized. Standardization involved subtracting the mean and then dividing by the standard deviation, with these statistics computed separately for each processed input sample. Furthermore, to mitigate overfitting (which is possible due to the large value of the K hyperparameter), a strategy was adopted to select the model with the best generalization ability. Generalization performance was evaluated on the validation set by monitoring the minimum mean squared error (MSE) achieved during training. Since the outcome of artificial neural network training depends on the initialization of weights, each experiment (with fixed hyperparameters) was repeated five times to enable a more reliable assessment of model performance. For all training runs, the maximum number of epochs was set at 5,000. However, to accelerate computation, early stopping was employed with a patience parameter of 100 epochs, meaning training would terminate if no improvement was observed within this window. Optimization was carried out using the Adam algorithm with a learning rate of 0.001. Three loss functions were considered for training: the standard MSE loss, the MAE loss, and the Fair loss with scaling parameter equal to 1.0.

To complement the perceptron approach, forecasts of active power and active energy were also generated using linear regressions whose coefficients were estimated using the ordinary least squares (OLS) and two ensemble methods: random forest (RF) and gradient boosting regressors (XGB). The primary advantages of OLS is straightforward interpretability and low computational complexity. However, OLS is limited in that it cannot capture the nonlinear relationships that are likely present in the analyzed time series. Therefore, we also included ensemble methods. For both active power and energy, we specified the models with lagged first differences of the series as regressors, consistent with the perceptron configuration.

For active energy, an alternative target was adopted for the classical models. Specifically, for OLS, RF and XGB, the target was reformulated as the cumulative increase from t to $t + h$, i.e., $\hat{P}(t,h) = P(t+h) - P(t)$. The energy forecasts were then reconstructed by adding the predicted increase to the last available meter reading. In contrast, the MLP variant directly predicted the future meter level.

Hyperparameters for the ensemble models were selected on the validation subsample using RMSE as the primary criterion. The final configuration was kept constant across all horizons and both targets: RF used 100 trees, and XGB used 100 boosting rounds with maximal depth set to six and learning rate equals 0.1. OLS included an explicit intercept and did not require hyperparameter tuning.

RESULTS AND DISCUSSION

Tab. 2 and 3 present the prediction accuracy metrics of active power and energy, respectively.

Table 2. Active power – prediction accuracy metrics

Method	Horizon	MAE	MAPE	SMAPE	R^2
Naive persistence model (k-1)	1	0.01034	0.13908	0.10654	0.86692
	4	0.03023	0.49737	0.32733	0.49484
	8	0.04425	0.77747	0.49707	0.23183
	48	0.10322	1.70806	0.92971	-1.58971
	96	0.05081	0.86738	0.54141	-0.02348
Linear regression (OLS)	1	0.01384	0.21397	0.19148	0.90544
	4	0.03012	0.50079	0.37292	0.71834
	8	0.03700	0.63892	0.45294	0.61772
	48	0.04749	0.78777	0.52507	0.42785
	96	0.04514	0.76717	0.51685	0.47959
Random forest (RF)	1	0.01225	0.18640	0.15368	0.92992
	4	0.02124	0.35811	0.26593	0.83047
	8	0.02795	0.49600	0.34683	0.74579
	48	0.04006	0.75012	0.47411	0.57044
	96	0.04059	0.76784	0.48381	0.56427
Gradient boosting (XGB)	1	0.00857	0.12402	0.10781	0.95241
	4	0.01892	0.31529	0.23543	0.84181
	8	0.02547	0.44521	0.31507	0.75427
	48	0.03739	0.68144	0.44207	0.57140
	96	0.03725	0.68431	0.44657	0.57482
MLP	1	0.00980	0.14428	0.12601	0.93539
	4	0.02063	0.32686	0.27564	0.82232
	8	0.02674	0.47454	0.35102	0.75881
	48	0.03570	0.64144	0.42152	0.56776
	96	0.03625	0.60030	0.42868	0.54276

Note. Own calculations

Table 3. Active energy – prediction accuracy metrics

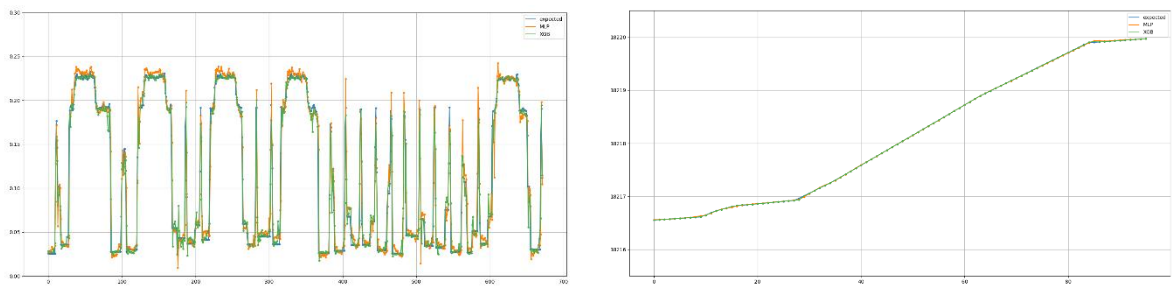
Method	Horizon	MAE	MAPE	SMAPE	R^2
Naive persistence model (k-1)	1	0.02683	1.5e-06	1.5e-06	1.00000
	4	0.10734	5.9e-06	5.9e-06	1.00000
	8	0.21469	0.00001	0.00001	1.00000
	48	1.29025	0.00007	0.00007	0.99990
	96	2.58569	0.00014	0.00014	0.99963
Linear regression (OLS)	1	0.00346	1.9e-07	1.9e-07	1.00000
	4	0.02083	1.1e-06	1.1e-06	1.00000
	8	0.04532	2.5e-06	2.5e-06	1.00000
	48	0.30777	0.00002	0.00002	0.99999
	96	0.51115	0.00003	0.00003	0.99998
Random forest (RF)	1	0.00222	1.2e-07	1.2e-07	1.00000
	4	0.01171	6.4e-07	6.4e-07	1.00000
	8	0.02735	1.5e-06	1.5e-06	1.00000
	48	0.20484	0.00001	0.00001	0.99999
	96	0.33686	0.00002	0.00002	0.99999
Gradient boosting (XGB)	1	0.00225	1.2e-07	1.2e-07	1.00000
	4	0.01174	6.4e-07	6.4e-07	1.00000
	8	0.02700	1.5e-06	1.5e-06	1.00000
	48	0.20218	0.00001	0.00001	1.00000
	96	0.32721	0.00002	0.00002	0.99999
MLP	1	0.00503	2.7e-07	2.7e-07	1.00000
	4	0.01738	9.5e-07	9.5e-07	1.00000
	8	0.03399	1.9e-06	1.9e-06	1.00000
	48	0.19941	0.00001	0.00001	1.00000
	96	0.34444	0.00002	0.00002	0.99999

Note. Own calculations

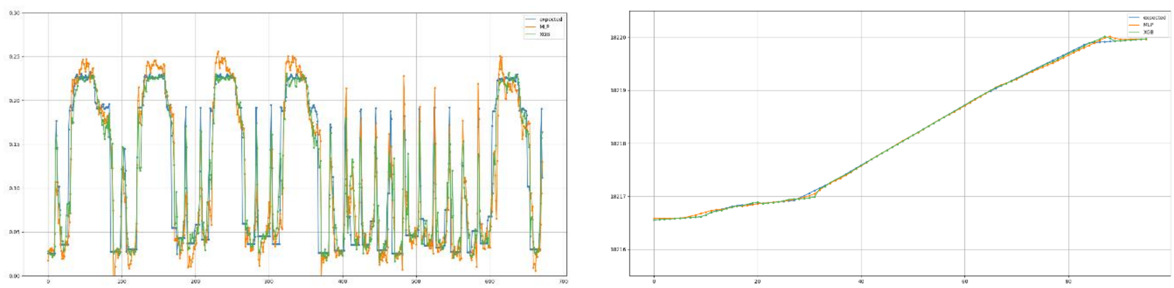
On the active power task, XGB delivered the most accurate predictions across most horizons, particularly for short and medium ranges (15–120 minutes), where iterative residual fitting captures nonlinear interactions among recent differences. Random Forest consistently trailed XGB by a small margin and outperformed OLS at all horizons, benefiting from the longer backward window at larger h, which exposes weekly structure in the differenced domain. OLS, while linear and highly interpretable, nonetheless provided a solid improvement over the naive persistence benchmark, especially as the horizon lengthened and persistence deteriorated (including negative R^2 at 12–24 hours).

On the active energy forecast task, all methods achieved near-perfect fit ($R^2 \approx 1$) across horizons, including the naive persistence model. This outcome follows from the reconstruction $\overline{P(t+h)} = P(t) + \hat{P}(t,h)$, the last reading P(t) carries most of the level, and the model estimates only the residual increment. Consequently, differences among XGB, Random Forest, OLS, and MLP were numerically small and operationally negligible. In this setting, OLS is generally sufficient for deployment unless marginal gains from tree ensembles or the perceptron are required by specific operational constraints.

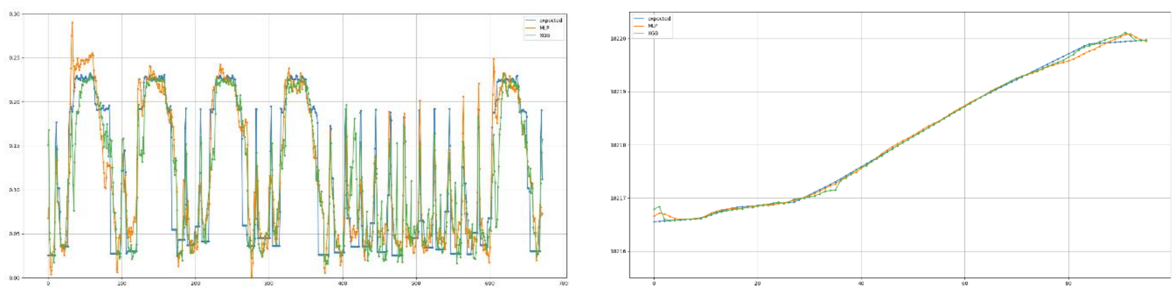
The MLP matched or slightly outperformed Random Forest performance at several horizons and surpassed OLS. For MLP, the experiments demonstrated that $K = 2048$ neurons constituted an excessive number for the data under consideration, with the best results obtained for $K = 1024$ hidden neurons. It is possible that, depending on the characteristics of the data, this number could be further reduced (resulting in computationally less complex models) without a loss in prediction quality. In all experiments, the *multi* strategy performed at least as well as the *single* strategy. The few cases where the best result was obtained with the *single* strategy occurred when $h = 1$, in which both strategies are equivalent. The observed differences arise from the randomness of neural network weight initialization. More accurate prediction of distant samples (larger h values) requires a greater number of historical samples (larger L values). With respect to the choice of loss function, it was observed that in the tasks under consideration, MSE loss generally performed slightly worse than the other two functions. It should be noted that Fair loss, like MAE loss, is based on the absolute differences between the expected and predicted values.



$h = 1$ $h = 1$



$h = 4$ $h = 4$



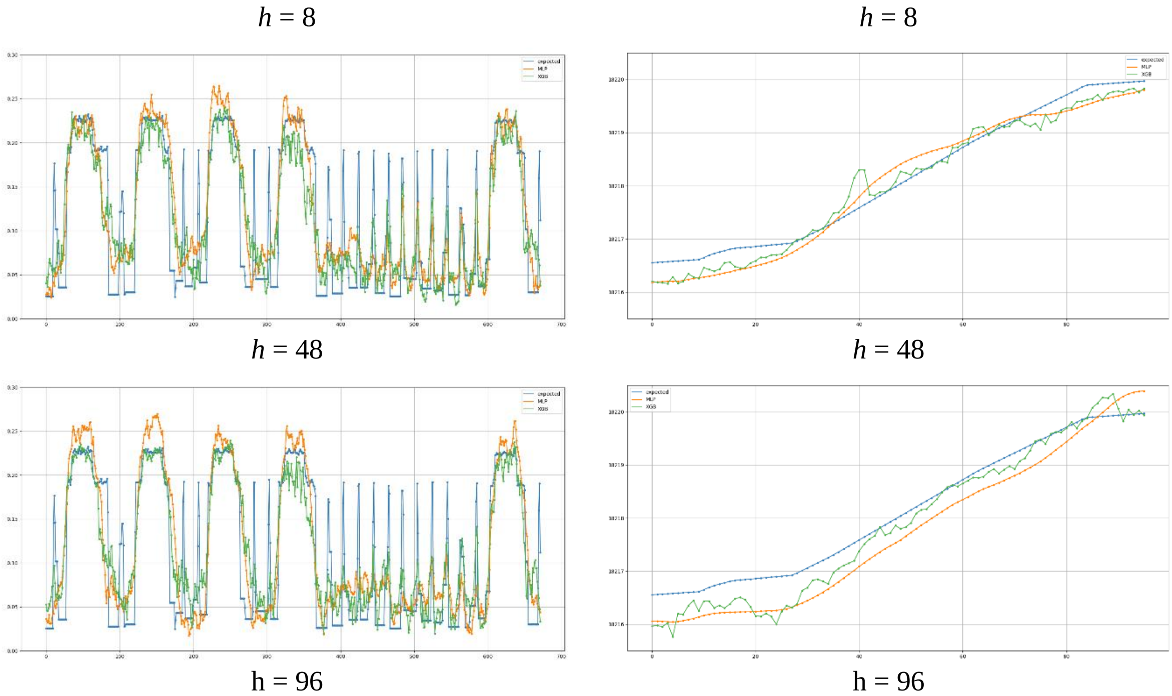


Figure 1. Sample predictions (XGBoost and MLP) of active power (left, 7 days) and active energy (right, 1 day) for different horizons *h* (XGBoost – green, MLP – orange, expected-blue).

The graphs presented in Figure 1 clearly demonstrate the relatively high accuracy of predictions, even for distant samples – a quarter of an hour distant by one day (96 samples). When forecasting active energy, no significant differences should be expected, as the function is clearly monotonic. This trend is best captured by a multilayer perceptron. On the other hand, the variability of active power is much better captured by XGBoost. Local peaks are visible even in distant forecasts, which is crucial when forecasting instantaneous loads on the electrical network. The reliability of forecasts is reinforced by the values of the error measures. For active energy, they are at a level that raises no doubts, while for active power, they show that the forecasts exhibit a trend of change, and incremental and successive updating of the forecasts allows for very high quality to be achieved.

In summary, for active power forecasting, XGB is recommended as the top performer, with RF a robust and stable alternative and MLP competitive when configured with a moderate hidden size. OLS remains a strong, transparent baseline that clearly surpasses persistence at longer horizons.

CONCLUSIONS

Based on reported results, several conclusions can be made. Specifically, forecasting active energy is solved by all considered models good for all forecast horizons. The worse quality measures reach less than 1% of sMAPE and more than 0.99 for coefficient of determination. That aligns with almost perfect forecast. It is not surprising as energy is monotonic function (it has no seasonality and trend is roughly constant). That confirms the high application potential of considered methods, in particular taking into account edge-oriented deployments of highly constrained memory and computational resources.

Active power, though in practice it is the quantity of energy use, is much harder to predict. It is yet not surprising either, as high oscillation and seasonality of data make finding underlying relationships challenging. Certainly, the difficulty increases while growing forecast horizon. In terms of sMAPE for the longest forecast horizon (i.e., 96 quarters ahead), the best result was achieved by MLP with metric value attained at the level of 42.9% whereas the worst for the naive persistence model (51.1%).

In overall, the presented forecast methods aligning with limited computational and memory resources available for edge deployments, provide predictions of sufficient accuracy to facilitate balancing the operational regime of HVAC systems.

IMPLICATIONS FOR RESEARCH, APPLICATION, OR POLICY

High-granularity energy-consumption forecasting, including mid-range projections, is essential for understanding demand drivers and managing modern power systems. Popular “building intelligence” solutions, like e.g., smart homes, renewable-ready grids, and effective demand–supply control, depend on operational forecasts at 15-minute resolution (or finer); without such capabilities, reliable automation and optimization are not feasible.

Building-efficiency standards (e.g., net-zero/zero-emission requirements) demand precise monitoring and control of equipment to meet targets and pass routine audits. Without the forecasting approach discussed in this article, it is difficult to conduct the energy reviews and audits specified by ISO 50001 (energy management systems).

An interesting implication of high-granularity prediction, in the context of further research, is the possibility of comparing predicted energy values with current time series. This makes it possible to detect anomalies and attacks on energy systems. Aggregated energy data does not allow for the monitoring of anomalies and attacks on energy networks, which can be effectively detected and diagnosed by comparing finely granulated forecasts with actual electricity consumption characteristics. At the same time, the compliance of energy systems with European NIS and CRA directives imposes new requirements on energy network monitoring and management systems that go beyond the highly aggregated approach often found in the literature.

In summary, high-granularity monitoring and planning is an important research challenge with significant practical and strategic implications, both in the context of safety certification, energy audits, and the widespread shift towards ecology and increasing the share of renewable sources in the energy mix.

ACKNOWLEDGEMENTS

This research was carried out as a part of project FENG.01.01-IP.02-0063/23 (ADSR), entitled: "*Intelligent algorithms and interoperable applications for the integration of BEMS/EnMS energy management systems with BACS/BMS/HMS, ESS, and ACP systems, to optimize energy use as part of Demand Side Response (DSR) energy flexibility programs*" as well as a part of project FENG.01.01-IP.01-A07N/24 (FIF), entitled: "Environment and the process of designing and manufacturing smart, autonomous, miniaturized FIF edge devices to enable zero-emission and energy efficiency improvements", co-financed by the National Centre for Research and Development in Poland.

REFERENCES

Afzaal, H., Farooque, A. A., Esau, T. J., Schumann, A. W., Zaman, Q. U., Abbas, F., & Bos, M. (2023). Artificial neural modeling for precision agricultural water management practices. In *Precision Agriculture* (pp. 169-186). Academic Press.

Amasyali, K., & El-Gohary, N. M. (2018). A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews*, 81, 1192-1205.

ASHRAE (2023). ANSI/ASHRAE Standard 55: Thermal Environmental Conditions for Human Occupancy (2023). American Society of Heating, Refrigerating and Air-Conditioning Engineers

Aslam, W., Andrianesis, P., & Caramanis, M. (2023). Optimal hvac energy and regulation reserve scheduling in power markets. *IEEE Transactions on Sustainable Energy*, 15(1), 201-214.

Balali, Y., Chong, A., Busch, A., & O'Keefe, S. (2023). Energy modelling and control of building heating and cooling systems with data-driven and hybrid models—A review. *Renewable and Sustainable Energy Reviews*, *183*, 113496.

Bertolini, M., Duttilo, P., & Lisi, F. (2025). Accounting carbon emissions from electricity generation: a review and comparison of emission factor-based methods. *Applied Energy*, *392*, 125992.

Chen, Y., Guo, M., Chen, Z., Chen, Z., & Ji, Y. (2022). Physical energy and data-driven models in building energy prediction: A review. *Energy Reports*, *8*, 2656-2671

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *Peerj computer science*, *7*, e623.

Das, U. K., Tey, K. S., Seyedmahmoudian, M., Mekhilef, S., Idris, M. Y. I., Van Deventer, W., ... & Stojcevski, A. (2018). Forecasting of photovoltaic power generation and model optimization: A review. *Renewable and Sustainable Energy Reviews*, *81*, 912-928.

Di Persio, L., & Fraccarolo, N. (2023). Energy consumption forecasts by gradient boosting regression trees. *Mathematics*, *11*(5), 1068.

European Commission. (2017). Commission Regulation (EU) 2017/2195 of 23 November 2017 establishing a guideline on electricity balancing. Official Journal of the European Union, L 312/6.
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R2195

Fan, C., Wang, J., Gang, W., & Li, S. (2019). Assessment of deep recurrent neural network-based strategies for short-term building energy predictions. *Applied energy*, *236*, 700-710.

Fraisse, G., Viardot, C., Lafabrie, O., & Achard, G. (2002). Development of a simplified and accurate building model based on electrical analogy. *Energy and buildings*, *34*(10), 1017-1031.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.

Hlongwane, N. W. (2025). Determinants of electricity consumption in South Africa: Insights from linear and nonlinear modeling approaches. *Economics, Management and Sustainability*, *10*(1), 6–26. https://doi.org/10.14254/jems.2025.10-1.1

Ho, T. K. (1995, August). Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition* (Vol. 1, pp. 278-282). IEEE.

International Energy Agency. (2024, March 21). *Poland energy mix*. https://www.iea.org/countries/poland/energy-mix

Jung, J., & Broadwater, R. P. (2014). Current status and future advances for wind speed and power forecasting. *Renewable and Sustainable Energy Reviews*, *31*, 762-777.

Khan, A. A., Chaudhari, O., & Chandra, R. (2024). A review of ensemble learning and data augmentation models for class imbalanced problems: Combination, implementation and evaluation. *Expert Systems with Applications*, *244*, 122778.

Kim, H., Kim, M., Kim, H., & Park, S. (2020). Decomposition analysis of CO2 emission from electricity generation: Comparison of OECD countries before and after the financial crisis. *Energies*, *13*(14), 3522.

Klyuev, R. V., Morgoev, I. D., Morgoeva, A. D., Gavrina, O. A., Martyushev, N. V., Efremenkov, E. A., & Mengxu, Q. (2022). Methods of forecasting electric energy consumption: A literature review. *Energies*, *15*(23), 8919.

Koebrich, S., Cofield, J., McCormick, G., Saraswat, I., Steinsultz, N., & Christian, P. (2025). Towards objective evaluation of the accuracy of marginal emissions factors. *Renewable and Sustainable Energy Reviews*, *215*, 115508.

Krutsyak, M. O. (2019). Forecasting demand on the Ukrainian electricity market using socio-economic variables. *Economics, Management and Sustainability, 4*(1), 46–57. https://doi.org/10.14254/jems.2019.4-1.5

Kumari, S., & Muthulakshmi, P. (2024). SARIMA model: an efficient machine learning technique for weather forecasting. *Procedia Computer Science*, *235*, 656-670.

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., ... & Battaglia, P. (2023). Learning skillful medium-range global weather forecasting. *Science*, *382*(6677), 1416-1421.

Liu, T., Xu, C., Guo, Y., & Chen, H. (2019). A novel deep reinforcement learning based methodology for short-term HVAC system energy consumption prediction. *International Journal of Refrigeration*, *107*, 39-51.

Ministerstwo Klimatu i Środowiska. (2024a). *Rozporządzenie Ministra Klimatu i Środowiska z dnia 24 kwietnia 2024 w sprawie ogłoszenia jednolitego tekstu rozporządzenia Ministra Klimatu i Środowiska w sprawie sposobu kształtowania i kalkulacji taryf oraz sposobu rozliczeń w obrocie energią elektryczną* [Regulation of the Minister of Climate and Environment on the announcement of the consolidated text of the regulation on the method of shaping and calculating tariffs and settlement methods in electricity trading]. Retrieved from https://www.ure.gov.pl/download/9/13359/Rozporzadzenietaryfoweelektroenergetyczne.pdf

Ministerstwo Klimatu i Środowiska. (2024b). *Rozporządzenie Ministra Klimatu i Środowiska z dnia 24 września 2024 w sprawie w sprawie wskaźnika emisji gazów cieplarnianych dla energii elektrycznej w 2025. Retrieved from* https://isap.sejm.gov.pl/isap.nsf/download.xsp/WDU20240001439/O/D20241439.pdf

Magalhães, B., Bento, P., Pombo, J., Calado, M. D. R., & Mariano, S. (2024). Short-term load forecasting based on optimized random forest and optimal feature selection. *Energies*, *17*(8), 1926.

Nicholson, S., & Heath, G. (2021). *Life cycle emissions factors for electricity generation technologies* (No. FY17 AOP 4.6. 8.3). National Renewable Energy Laboratory-Data (NREL-DATA), Golden, CO (United States); National Renewable Energy Lab.(NREL), Golden, CO (United States).

Olu-Ajayi, R., Alaka, H., Owolabi, H., Akanbi, L., & Ganiyu, S. (2023). Data-Driven Tools for Building Energy Consumption Prediction: A Review. *Energies*, *16*(6), 2574.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., ... & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography, 40*(8), 913-929.

Selmic, R. R., & Lewis, F. L. (2002). Neural-network approximation of piecewise continuous functions: application to friction compensation. *IEEE transactions on neural networks, 13*(3), 745-751.

Sendra-Arranz, R., & Gutiérrez, A. (2020). A long short-term memory artificial neural network to predict daily HVAC consumption in buildings. *Energy and Buildings*, *216*, 109952.

Song, X., Deng, L., Wang, H., Zhang, Y., He, Y., & Cao, W. (2024). Deep learning-based time series forecasting. *Artificial Intelligence Review*, *58*(1), 23.

Szostek, K., Mazur, D., Drałus, G., & Kusznier, J. (2024). Analysis of the Effectiveness of ARIMA, SARIMA, and SVR Models in Time Series Forecasting: A Case Study of Wind Farm Energy Production. *Energies (19961073)*, *17*(19).

Urząd Regulacji Energetyki. (2024, March 4). *2024*. https://www.ure.gov.pl/pl/energia-elektryczna/charakterystyka-rynku/12744,2024.html

Urząd Regulacji Energetyki. (2022, February 21). *Jak optymalizować koszty energii elektrycznej w sektorze publicznym?* https://www.ure.gov.pl/pl/urzad/informacje-ogolne/aktualnosci/9851,Jak-optymalizowac-koszty-energii-elektrycznej-w-sektorze-publicznym.html

Verbeke, S., & Audenaert, A. (2018). Thermal inertia in buildings: A review of impacts across climate and building use. *Renewable and sustainable energy reviews, 82,* 2300-2318.

Walczak, J. (2025). Evidence-based Methodological Framework for Machine Learning Studies. *Authorea Preprints*.

Wang, Z., & Srinivasan, R. S. (2017). A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models. *Renewable and sustainable energy reviews*, 75, 796-808.

Xu, X., & Wang, S. (2008). A simplified dynamic model for existing buildings using CTF and thermal network models. *International Journal of Thermal Sciences*, *47*(9), 1249-1262.

Zhang, X., Pipattanasomporn, M., Chen, T., & Rahman, S. (2019). An IoT-based thermal model learning framework for smart buildings. *IEEE Internet of Things Journal*, *7*(1), 518-527.

Authors' Note

All correspondence should be addressed to
Arkadiusz Tomczyk
Lodz University of Technology
Poland
https://orcid.org/0000-0001-9840-6209
arkadiusz.tomczyk@p.lodz.pl

Human Technology
ISSN 1795-6889
https://ht.csr-pub.eu