KMAN PUBLISHING INSTITUTE

Article history:
Received 23 June 2026
Revised 16 September 2026
Accepted 24 September 2026
Initial Publish 27 October 2026
Published online 01 May 2027

Journal of Resource Management and Decision Engineering

Volume 6, Issue 5, pp 1-21

A Novel Hybrid Model Based on an Optimized Multilayer Perceptron Neural Network and the Cuckoo Search Algorithm for Forecasting Monthly Household Electricity Consumption: A Case Study of Ilam County

Sadegh. Feizollahi[1*], Roham. Mokhtari Dezaki[2], Ahmadreza. Fatahi Nezhad[2], Mohammad. Heidari Goujani[2]

[1] Department of management, Il.C., Islamic Azad university, Ilam, Iran
[2] Department of management, ShK.C., Islamic Azad university, Shahr-e Kord, Iran

*** Corresponding author email address**: sadegh.feizollahi@iau.ac.ir

A r t i c l e   I n f o

Article type:
Original Research

How to cite this article:
Feizollahi, S., Mokhtari Dezaki, R., Fatahi Nezhad, A., & Heidari Goujani, M. (2027). A Novel Hybrid Model Based on an Optimized Multilayer Perceptron Neural Network and the Cuckoo Search Algorithm for Forecasting Monthly Household Electricity Consumption: A Case Study of Ilam County. *Journal of Resource Management and Decision Engineering*, *6*(5), 1-21.
https://doi.org/10.61838/kman.jrmde.439

A B S T R A C T

The objective of this study is to propose a novel method for forecasting the monthly electricity consumption of subscribers in Ilam County. The results of this study can assist energy-supply companies in obtaining an overall understanding of household electricity consumption patterns. For forecasting purposes, a combination of a multilayer perceptron (MLP) neural network and the Cuckoo Search Optimization Algorithm is employed. In the proposed method, the estimation process is performed in three stages: “preprocessing,” “feature selection,” and “estimation.” The minimum redundancy–maximum relevance (MRMR) algorithm is used for feature selection. This algorithm identifies the optimal features for forecasting electricity consumption such that they exhibit minimum redundancy and maximum relevance to the target variable. In the estimation stage, a combination of three multilayer perceptron neural networks is employed, and their configurations are adjusted using the Cuckoo Search Optimization Algorithm so as to minimize the estimation error. The results of the performance evaluation of the proposed method indicated that the accuracy of forecasting subscribers’ monthly electricity consumption can be maximized using six features. Furthermore, comparison of the proposed method with other existing techniques showed that the proposed hybrid system can perform the forecasting process with a mean absolute error of 5.66 kWh, representing an error reduction of at least 29.4% compared with previous methods.

Keywords: *multilayer perceptron network, electricity consumption forecasting, machine learning, electricity consumption in Ilam.*

1. Introduction

The rapid growth of electricity demand, increasing urbanization, expansion of residential infrastructure, proliferation of electrical appliances, and growing dependence of households on energy-intensive technologies have transformed electricity consumption forecasting into a central issue in contemporary energy management. Accurate estimation of future electricity demand is essential for generation planning, load balancing, distribution management, infrastructure investment, demand-side management, and the economic operation of power systems. At the household level, forecasting is particularly challenging because residential electricity use is influenced by heterogeneous and dynamic factors, including household behavior, seasonal patterns, climatic conditions, appliance ownership, occupancy schedules, income levels, and broader socioeconomic trends. These characteristics make residential electricity consumption highly variable and often nonlinear, thereby limiting the effectiveness of conventional statistical forecasting approaches. In this context, data-driven and machine-learning methods have increasingly attracted attention because of their ability to identify complex relationships within historical consumption data and produce more flexible predictive models (Ardabili et al., 2019; Mosavi & Bahmani, 2019).

The importance of electricity forecasting is also closely associated with sustainability objectives and the transition toward smarter energy systems. Reliable electricity-demand prediction can help utilities reduce unnecessary reserve margins, improve the allocation of generation resources, optimize grid operation, and decrease both economic and environmental costs. In smart-city environments, energy information generated from buildings and households can be transformed into predictive knowledge that supports intelligent control and more efficient use of resources. Machine-learning models have therefore become increasingly important for demand and consumption forecasting in the context of sustainable and smart cities because they can process historical data, discover nonlinear patterns, and adapt to complex variations in consumption behavior (Ardabili et al., 2019). The broader energy literature likewise indicates that the interaction among electricity demand, economic growth, fossil-fuel use, and environmental pressures has made accurate energy forecasting increasingly relevant for long-term energy policy and environmental planning, particularly in countries with substantial dependence on fossil-based energy systems (Lotfalipour et al., 2010).

Iran provides an especially relevant context for research on electricity-demand forecasting. The national energy system has historically been characterized by a strong dependence on fossil fuels, while the need to diversify the energy mix and expand renewable-energy deployment has received increasing policy attention. At the same time, electricity consumption continues to rise, creating pressure on generation, transmission, and distribution infrastructure. Research on Iran's energy sector has highlighted both the significant potential of renewable resources, particularly solar photovoltaic power, and the institutional, technological, economic, and policy barriers that continue to affect their wider deployment (Gorjian et al., 2019). Under these conditions, improving the accuracy of electricity-demand forecasting becomes important not only for operational planning but also for facilitating the integration of renewable resources, reducing mismatches between supply and demand, and supporting long-term energy-transition strategies. More accurate forecasts can assist electricity providers in anticipating short- and medium-term demand, improving capacity allocation, and reducing the need for costly corrective interventions.

Household electricity consumption constitutes a particularly important component of electricity demand because residential users display substantial temporal variability. Consumption may change considerably across months due to temperature fluctuations, heating and cooling requirements, changes in occupancy, and seasonal household activities. This temporal variability means that monthly electricity forecasting requires models capable of extracting meaningful patterns from sequential consumption records while remaining robust to noise and irregular observations. Recent studies have consequently examined a variety of data-driven approaches to residential electricity forecasting. For example, household monthly electricity consumption has been forecast using pattern-based approaches that identify similarities between current and historical consumption structures, demonstrating the potential value of historical-pattern information in improving monthly demand estimation (Hosseini, 2025). Similarly, research applying long short-term memory architectures to residential electricity forecasting has emphasized the effectiveness of sequence-based deep-learning models for capturing temporal dependencies in household consumption data (Chen, 2025).

The forecasting of electricity demand has evolved considerably from traditional regression and time-series models toward artificial intelligence, machine learning, and deep-learning techniques. Conventional statistical models are often based on assumptions regarding linearity, stationarity, or predefined functional relationships among variables. However, electricity-consumption data frequently exhibit nonlinear behavior, abrupt variations, seasonal effects, and interactions among multiple explanatory factors. Machine-learning approaches are advantageous in such contexts because they can approximate complex functions without requiring a complete prior specification of the relationships embedded in the data. Reviews of machine-learning applications in energy-consumption prediction have identified artificial neural networks, support vector machines, ensemble methods, decision trees, and hybrid optimization-based models among the most frequently employed techniques for improving predictive performance (Mosavi & Bahmani, 2019). The growing availability of digital metering data has further strengthened the applicability of these methods by providing more detailed and reliable consumption records.

Artificial neural networks are among the most widely used machine-learning techniques for electricity-demand forecasting. Their principal advantage lies in their capacity to approximate nonlinear mappings between input variables and target outputs. In electricity-consumption forecasting, a neural network can learn the relationship between historical load patterns and future consumption without requiring an explicit mathematical representation of the underlying process. Multilayer perceptron networks, in particular, have been extensively used for prediction problems because of their flexible structure and universal approximation capability. Nevertheless, the effectiveness of an MLP depends strongly on its architecture, number of hidden neurons, initialization of connection weights, bias parameters, and training procedure. An inadequately configured network may exhibit poor generalization, slow convergence, overfitting, or convergence toward local optima. Therefore, determining an appropriate architecture and parameter configuration remains one of the major challenges in designing neural-network forecasting systems.

Deep recurrent neural networks have also been employed for building-level electricity forecasting because of their ability to model sequential dependencies. Research on commercial and residential buildings has shown that recurrent neural-network architectures can exploit temporal information in electricity-consumption data and generate effective forecasts across different building types (Rahman et al., 2018). Such findings illustrate the broader potential of neural approaches for energy-demand modeling. However, deep architectures generally require substantial amounts of training data and may involve greater computational complexity than simpler feedforward models. In applications involving relatively limited household datasets or monthly observations, an optimized multilayer perceptron may therefore provide an attractive balance between predictive flexibility and computational efficiency. This consideration becomes particularly important when the forecasting framework is intended for practical implementation by local utilities with limited computational resources.

Another important challenge in electricity-demand forecasting concerns the selection of appropriate input variables. Historical consumption records may contain numerous lagged observations, but not all of them contribute equally to the prediction of future electricity use. Including irrelevant or highly redundant features can increase computational complexity, reduce model interpretability, and even degrade predictive accuracy. Feature selection is therefore a critical preprocessing step in machine-learning-based forecasting. Among the available feature-selection approaches, the maximum relevance–minimum redundancy method seeks to select variables that have strong relationships with the target variable while exhibiting minimal redundancy with one another. This strategy has been applied in machine-learning environments to improve predictive efficiency by balancing relevance and redundancy during feature selection (Zhao et al., 2019). For electricity-demand forecasting, such an approach is especially useful because lagged consumption values may be strongly correlated with one another, making it necessary to identify a compact subset that preserves the most informative temporal characteristics.

The integration of feature-selection algorithms with predictive models can improve both efficiency and accuracy. By removing unnecessary input dimensions, the learning algorithm can focus on the most informative patterns in the data, potentially reducing training time and improving generalization. In household electricity forecasting, this issue is particularly important because monthly consumption histories naturally generate multiple candidate lag variables. A forecasting framework that treats all historical observations as equally informative may therefore include redundant information. Applying MRMR before model construction offers a systematic mechanism for identifying the most relevant historical consumption features while minimizing duplication of information. This feature-selection stage can consequently simplify the subsequent neural-network model and reduce the dimensionality of the optimization problem.

In addition to feature selection, optimization of neural-network parameters represents another major avenue for improving electricity-consumption forecasting. Metaheuristic optimization algorithms have been increasingly applied to model training because they can search complex, nonlinear, and nonconvex solution spaces without relying solely on gradient information. Long-term electricity-consumption forecasting research has shown that evolutionary and population-based optimization techniques can be effectively integrated with forecasting systems to enhance predictive accuracy (Kaboli et al., 2016). Such approaches are particularly useful when the problem includes simultaneous optimization of structural and numerical parameters, such as the number of hidden neurons and the connection weights of an artificial neural network.

Among population-based optimization techniques, the Cuckoo Search Algorithm has received considerable attention because of its relatively simple structure, limited number of control parameters, and capacity for global search. The algorithm is inspired by the brood-parasitic behavior of certain cuckoo species and typically combines random nest selection with Lévy-flight-based exploration. Its optimization mechanism allows candidate solutions to explore different regions of the search space while retaining high-quality solutions between iterations. An adaptive form of the Cuckoo Search Algorithm has demonstrated the ability to improve optimization performance by modifying search behavior and balancing exploration and exploitation (Mareli & Twala, 2018). These characteristics make Cuckoo Search potentially suitable for optimizing neural-network parameters, where the search space may be highly nonlinear and contain multiple local minima.

The combination of neural networks with metaheuristic optimization may overcome several limitations of conventional neural-network training. Standard gradient-based algorithms, although computationally efficient, may be sensitive to initialization and can become trapped in local optima. Metaheuristic methods provide an alternative mechanism for exploring the parameter space more globally. By optimizing both the neural-network topology and its weight vector, a Cuckoo Search-based framework can potentially identify configurations that yield lower forecasting errors than conventionally trained neural networks. This is particularly relevant for multilayer perceptrons because predictive performance can vary substantially depending on the hidden-layer structure. Instead of manually selecting the number of neurons or relying exclusively on trial-and-error procedures, an optimization algorithm can search for an appropriate network configuration using an explicit error criterion.

Another promising strategy for improving forecasting accuracy is the use of ensemble learning. A single predictive model may produce errors because of model-specific bias, training variability, or sensitivity to particular observations. Combining multiple learning models can reduce the influence of these individual errors and improve robustness. In neural-network forecasting, averaging the outputs of several independently optimized networks provides a simple ensemble mechanism through which extreme or inaccurate predictions from one network can be moderated by the outputs of the others. Such an approach can be especially valuable when the individual networks differ in topology and weight configuration, because diversity among the constituent models is generally an important source of ensemble effectiveness. Accordingly, a forecasting framework that combines feature selection, metaheuristic optimization, and neural-network aggregation has the potential to address multiple sources of prediction error simultaneously.

Climatic variability provides another important motivation for improving electricity-demand forecasting. Household electricity consumption is often sensitive to temperature and weather patterns because cooling and heating loads constitute significant components of residential demand. Recent research examining electricity consumption and its relationship with climate change has shown that forecasting electricity demand requires attention to changing environmental conditions and their potential influence on consumption behavior (Zhafran & Miefthawati, 2024). Although historical electricity-consumption values can implicitly reflect seasonal and climatic patterns, the variability associated with changing weather reinforces the need for predictive models capable of learning nonlinear and temporally dynamic relationships. Flexible machine-learning frameworks are particularly appropriate for this purpose because they can update their learned relationships when new observations become available.

Recent developments in household electricity forecasting also demonstrate a continued transition toward models that exploit sequential structure more effectively. LSTM-based studies have emphasized that the temporal dependence of household demand can be captured through models specifically designed for sequence learning (Chen, 2025), whereas pattern-based monthly forecasting has shown that recurring consumption profiles themselves contain substantial predictive information (Hosseini, 2025). These approaches support the general proposition that historical household electricity consumption contains meaningful temporal regularities. However, they also highlight an unresolved methodological question: how can a forecasting system identify the most informative historical observations while simultaneously optimizing the predictive model used to transform those observations into future estimates? Addressing both dimensions jointly may provide a more efficient alternative to frameworks that concentrate exclusively on either temporal modeling or parameter optimization.

The literature therefore suggests three complementary methodological requirements for improving household electricity forecasting. First, the model should select informative historical features and eliminate redundant input information. Second, the nonlinear forecasting model should be configured using an optimization mechanism capable of exploring a complex parameter space. Third, the final forecasting system should be sufficiently robust to reduce model-specific errors. The MRMR algorithm provides an appropriate mechanism for satisfying the first requirement, while the Cuckoo Search Algorithm provides a population-based strategy for optimizing neural-network topology and weights. An ensemble of independently optimized multilayer perceptron networks can address the third requirement by integrating multiple forecasts into a single output. The combination of these components offers a coherent hybrid architecture in which each stage addresses a distinct source of forecasting inefficiency.

The need for such an integrated framework is particularly evident in local-scale electricity systems, where utilities may require reliable forecasting but have access to comparatively limited datasets. In these settings, extremely data-intensive deep-learning models may not always be necessary or practical, whereas conventional forecasting approaches may fail to capture nonlinear temporal patterns sufficiently well. An optimized MLP-based ensemble provides a potentially effective intermediate solution because it preserves the nonlinear learning capacity of neural networks while using feature selection to limit input dimensionality and metaheuristic optimization to enhance network configuration. The use of monthly data further makes this approach relevant for medium-term utility planning, where forecasting accuracy can support budgeting, procurement, maintenance scheduling, infrastructure planning, and demand-management decisions.

In Ilam, the availability of historical household electricity-consumption records provides an opportunity to develop and evaluate such a framework under real-world conditions. Local electricity-demand patterns may reflect climatic seasonality, household behavior, and region-specific socioeconomic characteristics. A forecasting model developed using actual household records can therefore provide more operationally meaningful evidence than models tested exclusively on synthetic or benchmark datasets. Moreover, evaluating the proposed model against conventional neural networks, Gaussian process regression, and classification and regression trees can clarify whether the integration of feature selection, Cuckoo Search optimization, and ensemble averaging provides measurable improvements over alternative machine-learning methods.

Accordingly, the aim of this study is to develop and evaluate a novel hybrid forecasting model that integrates MRMR-based feature selection, Cuckoo Search-based optimization, and an ensemble of multilayer perceptron neural networks to improve the accuracy of monthly household electricity-consumption forecasting in Ilam County.

2. Methods and Materials

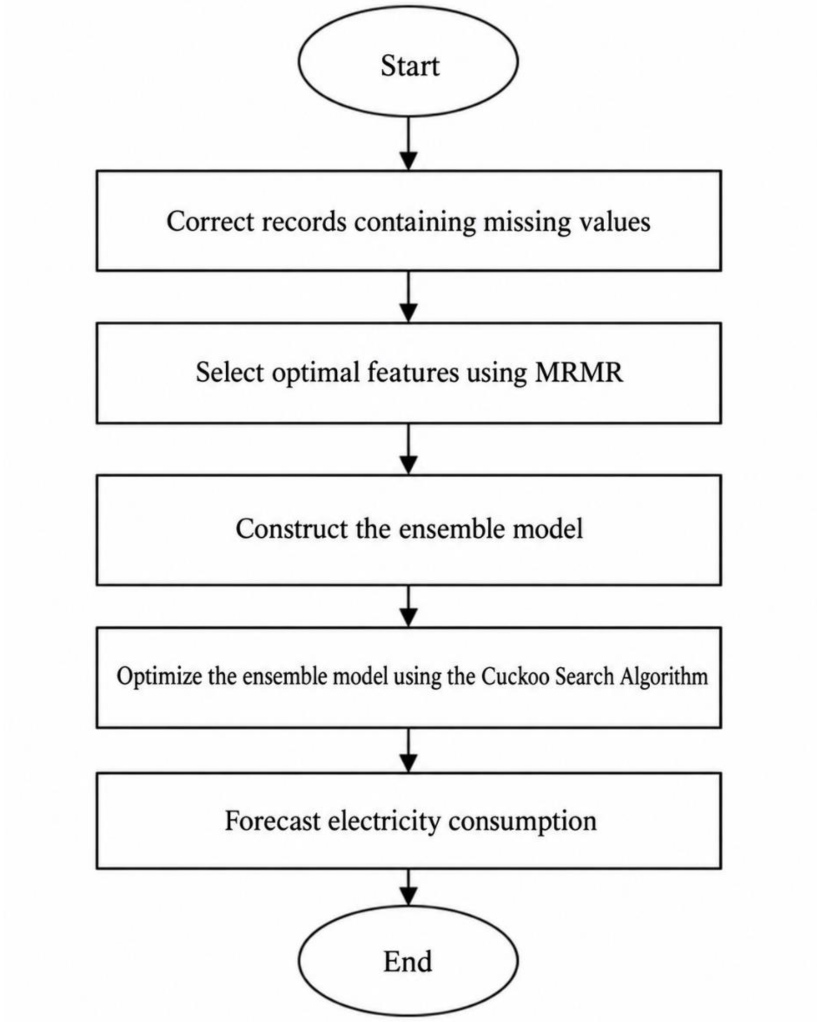
In this study, a hybrid framework based on machine learning and optimization algorithms was employed to forecast monthly household electricity consumption in Ilam. The data comprised actual electricity consumption records of households in Ilam over a 10-year period (2015–2025), obtained from the Ilam Regional Electricity Company. During the preprocessing stage, missing values were replaced using temporal averaging. Subsequently, the MRMR algorithm was employed to select the optimal subset of features associated with electricity consumption, thereby reducing redundancy while preserving maximum relevance to the target variable. In the next stage, an ensemble model based on three multilayer perceptron (MLP) neural networks was designed. The architecture and weight vectors of each network were optimized simultaneously and independently using the Cuckoo Search metaheuristic algorithm. The fitness function used in this algorithm was the mean absolute error (MAE) between the predicted and actual values. Finally, the model's final output was obtained by averaging the outputs of the three neural networks. The performance of the proposed model was compared with baseline methods, including a conventional artificial neural network (ANN), Gaussian process regression (GPR), and a classification and regression tree (CART).

The model proposed in this study is an intelligent hybrid framework for electricity consumption forecasting and was designed in three main stages. First, the raw data are preprocessed and the optimal features are selected using the MRMR algorithm so that redundancy is reduced while maximum relevance to electricity consumption is retained. In the second stage, an ensemble model based on three multilayer perceptron (MLP) neural networks is developed, in which the parameters and topology of each network are independently and simultaneously optimized by the Cuckoo Search metaheuristic algorithm to minimize forecasting error. Finally, the model's final output is calculated by averaging the results of the three networks in order to improve forecasting accuracy and stability.

In the first stage of the proposed method, the raw electricity consumption records of subscribers are preprocessed. For this purpose, records containing missing values are corrected, after which the MRMR algorithm is employed to select the optimal features for describing subscribers' electricity consumption records. Following feature selection, an MLP-based ensemble system is used to organize the estimation model. In this ensemble system, the Cuckoo Search Algorithm is used to determine the optimal configurations of the MLP models. Finally, the final output of the proposed method is determined based on the mean of the outputs of the three aforementioned multilayer perceptron neural networks. These stages are illustrated in Figure 1.

Figure 1

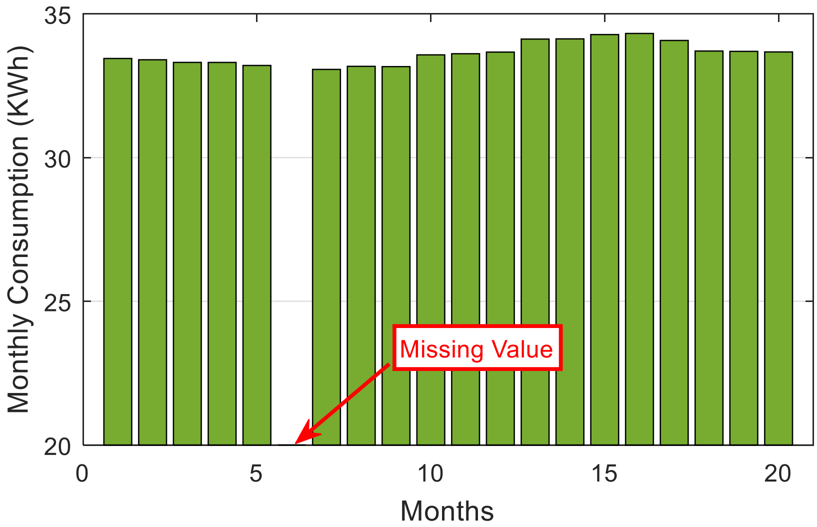
Flowchart of the proposed method



The purpose of preprocessing electricity consumption data is to prepare them for subsequent processing stages. However, the structure of the samples contained in the database must first be described. In the database used in this study, each data record contains the electricity consumption history of one subscriber at regular one-month intervals. Thus, each record indicates how many watt-hours of electricity a subscriber consumed in different months. An example of a database record is presented in Figure 2.

Figure 2

Example of a database record



As illustrated in the figure above, some samples in the database may contain missing values. In the example shown, information regarding the subscriber's electricity consumption in the sixth month is unavailable. In such cases, the proposed method corrects missing values according to the following rules:

If consumption records are available for both the preceding and following months of the current missing value, the mean electricity consumption of the preceding and following months is substituted for the missing value. For example, in Figure 2, the mean consumption in the fifth and seventh months is considered the electricity consumption for the sixth month.

If consumption records are unavailable for both the preceding and following months of the current missing value, the entire data record is disregarded because it contains multiple instances of incomplete information.

Using the simple rules described above, the proposed method corrects records containing incomplete information. After this procedure, a set of time series containing subscribers' electricity consumption histories becomes available. In the proposed method, these time series are represented as follows:

$$S_t = \{E_{t-10}, E_{t-9}, \dots, E_{t-1}, E_t\} \qquad (1)$$

In the above equation, $t$denotes the time index, and $E_t$represents the amount of electricity consumed at that time point. Based on time series S, electricity consumption in period $t+1$can be forecast. Accordingly, if the relationship between time series $S$and electricity consumption in period $t+1$is represented by a function such as Z, forecasting can be performed using subscribers' historical electricity consumption records according to the following equation:

$$E_{t+1} = Z(S_t) \qquad (2)$$

In the proposed method, an ensemble system based on neural networks is used to represent function Z. It should be noted that determining the length of sequence $S$is of considerable importance. In other words, not all members of the sequence $\{E_{t-10}, E_{t-9}, \dots, E_{t-1}, E_t\}$may be necessary for forecasting E_{t+1}. The use of unnecessary input data can not only increase processing time but also reduce estimation accuracy. Therefore, the MRMR algorithm is employed in the proposed method to identify the optimal features for electricity consumption estimation. In other words, at this stage, the MRMR algorithm selects an optimal subset of the members of sequence S, thereby maximizing forecasting accuracy.

MRMR is a feature-selection approach that seeks to select features exhibiting high correlation with the target variable (output) and low correlation with one another. For continuous features, the F-statistic can be used to calculate correlation with the target variable (relevance), whereas Pearson’s correlation coefficient can be used to calculate correlations among features (redundancy). Features are then selected sequentially by applying a greedy search to maximize an objective function based on relevance and redundancy. Two common objective functions are used in the MRMR algorithm:

- Mutual Information Difference (MID) criterion
- Mutual Information Quotient (MIQ) criterion

These functions represent, respectively, the difference between and the quotient of relevance and redundancy. In the proposed method, the Mutual Information Quotient criterion is used to define the objective function of the MRMR algorithm. The MRMR algorithm identifies an optimal set of features that simultaneously possesses the following two characteristics:

The selected features exhibit maximum pairwise differences and the highest correlation with the target variable; therefore, they can effectively describe the target variable. The objective of the MRMR algorithm is to identify an optimal feature set Sthat simultaneously maximizes V_S, representing the relevance of $S$to the target variable y, and minimizes W_S, representing the redundancy of S. Considering mutual information I, $V_S$and $W_S$can be defined as follows:

$$V_S = \frac{1}{|S|}\sum_{x\in S} I(x, y) \qquad (3)$$

$$W_S = \frac{1}{|S|^2}\sum_{x,z\in S} I(x, z) \qquad (4)$$

In the above equations, $|S|$denotes the number of features in S. Identifying an optimal set Srequires considering all $2^{|D|}$possible combinations, where Drepresents the complete set of features in the database. Instead, the MRMR algorithm ranks features using the Mutual Information Quotient (MIQ) through a forward-addition scheme. This procedure has a computational complexity of $O(|D|\cdot|S|)$and is calculated as follows:

$$MIQ_x = \frac{V_x}{W_x} \qquad (5)$$

In the above equation, $V_x$and $W_x$represent the relevance and redundancy of a feature, respectively:

$$V_x = I(x, y) \qquad (6)$$

$$W_x = \frac{1}{|S|}\sum_{z\in S} I(x, z) \qquad (7)$$

Accordingly, the MRMR algorithm ranks the data features as follows:

1. Select the feature with the highest relevance, $\max_{x\in D} V_x$, and add it to the initially empty set S.
2. Store features with nonzero relevance and zero redundancy in S'.
3. If S'contains no feature with nonzero relevance and zero redundancy, proceed to Step 4.
4. Otherwise, select the feature with the highest relevance,

$$\max_{x\in S',\, W_x=0} V_x,$$

and add the selected feature to set S.
5. Repeat Step 2 until redundancy is no longer zero for all features in S'.
6. Select the feature in $S'$with the highest MIQ value, nonzero relevance, and nonzero redundancy, and add it to set S:

$$\max_{x\in S'} MIQ_x = \max_{x\in S'} \frac{I(x, y)}{\frac{1}{|S|}\sum_{z\in S} I(x, z)} \qquad (8)$$

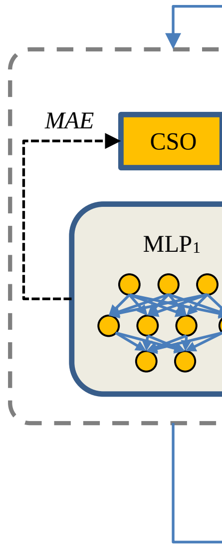
7. Repeat Step 4 until relevance becomes zero for all features in the remaining candidate set.
8. Add the features with zero relevance to Sin random order.

By implementing the above steps, the Nhighest-ranked features are selected and passed to the next stage of the proposed method, namely, construction of the ensemble model.

As noted previously, in the second stage of the proposed method, a combination of multilayer perceptron networks and the Cuckoo Search Optimization Algorithm is employed to construct the ensemble model. The ensemble model used in the proposed method consists of three multilayer perceptron networks, for each of which the optimal topology and weight vector are determined using the Cuckoo Search Algorithm. It should be noted, however, that more than three multilayer perceptron networks can be employed in this ensemble model. The operation of the proposed ensemble model is illustrated in Figure 3.

Figure 3

Proposed ensemble model based on neural networks and the Cuckoo Search Algorithm



According to the procedure illustrated in the figure above, each multilayer perceptron network is independently optimized using the Cuckoo Search Algorithm, and during the optimization process, the optimal neural-network topology and weight vector are determined. Because the optimization process for each neural network is independent of the other networks, this stage of the proposed method can be implemented using parallel-processing techniques. Accordingly, the processing time required to construct the ensemble model can be substantially reduced. Thus, at this stage of the proposed method, one Cuckoo Search optimizer is executed for each multilayer perceptron network. Using training error as its fitness function, each optimizer attempts to identify the most appropriate topology and weight vector for its corresponding neural network. The remainder of this section describes the process of optimizing multilayer perceptron neural networks using the Cuckoo Search Algorithm. For this purpose, the structure of the solution vector and population in the optimization algorithm is first described, followed by an explanation of the stages through which the algorithm determines the optimal solution.

This section describes the structure of the solution vector employed in the Cuckoo Search Algorithm. The solution vector in the optimization algorithm used in the proposed method determines the topology of the multilayer perceptron network as well as the weights of the connections among neurons and the network biases. Therefore, each solution vector in the optimization algorithm consists of two interrelated components. The first component specifies the network topology, whereas the second component specifies the neuron weights and biases for the topology defined in the first component. Consequently, the solution vectors in the Cuckoo Search Algorithm have variable lengths determined by the topology specified for the multilayer perceptron network. Because the number of possible multilayer perceptron network topologies may be unlimited, constraints must be imposed on the network-topology component of the solution vector. Accordingly, the following constraints are applied to the first component of the solution vector to restrict the search space:

Each multilayer perceptron network contains one hidden layer. Accordingly, the first component of the solution vector contains one element, and the numerical value of this element represents the number of neurons assigned to the hidden layer of the multilayer perceptron network.

The hidden layer of the multilayer perceptron network contains a minimum of 4 and a maximum of 15 neurons.

It should be noted that the first component of the solution vector, which determines the topology, describes only the number of neurons in the hidden layer because the dimensions of the input and output layers of the neural network are determined by the number of input features and output variables, respectively.

Because the number of neurons in the neural network is determined by the first component of the solution vector, the length of the second component is also determined according to the topology specified in the first component. For a neural network with I input neurons, H hidden neurons, and P output neurons, the length of the second component of the solution vector in the Cuckoo Search Algorithm is $H \times (I + 1) + P \times (H + 1)$.

The optimality of each solution in the Cuckoo Search Algorithm is evaluated using a fitness function. For this purpose, after each solution vector specifies the topology and weights of the neural network, the network outputs are generated for the training samples and compared with the actual target values. The mean absolute error criterion is then used to evaluate the performance of the neural network and the optimality of the generated solution. Accordingly, the fitness function of the Cuckoo Search Algorithm is defined as follows:

$$MAE = \sum_{i=1}^{N} |T_i - Z_i| \qquad (9)$$

In the above equation, N represents the number of training samples, and T_i denotes the target value for the ith training sample. Moreover, Z_i represents the output generated by the neural network for the ith training sample. In the proposed method, the Cuckoo Search Optimization Algorithm is employed to determine a neural-network structure capable of minimizing Equation (9).

The initial population in the Cuckoo Search Algorithm is generated randomly, and the search bounds for the second component of the solution vector are set to $[-, 1 + 1]$. Thus, each weight associated with the connections among neurons and each neural-network bias assumes a value within this interval.

In the proposed method, the Cuckoo Search Algorithm is responsible for determining the most appropriate neural-network topologies and their optimal weight values. This process is described below. Cuckoo Search is an optimization algorithm that, in addition to providing faster convergence than many other optimization methods, does not require the specification of numerous parameters. The general framework of this algorithm involves initiating the search from a set of random solutions and subsequently improving them through different operators. The algorithm operates according to the following three rules:

Each cuckoo lays one egg at a time and deposits its egg in a randomly selected nest.

The best nests with the highest-quality eggs are retained for the next generation. In other words, the best solutions remain present in the next iteration of the search algorithm.

The number of available host nests is fixed, and a host bird discovers a cuckoo egg with a probability of $P_a \in (0, 1)$.

The pseudocode of the Cuckoo Search Algorithm is as follows:

Algorithm 1. Cuckoo Search

1. Define the algorithm fitness function as $f(x) = (x_1, \ldots, x_d)$.
2. Generate an initial population comprising n nests, each containing K eggs.
3. While $t < T$, or until one of the stopping criteria is satisfied, repeat the following operations:
    1. Select a random cuckoo i using a Lévy-flight procedure.
    2. Evaluate the fitness of the selected cuckoo using the fitness function.
    3. Select a random nest j from the n nests.
    4. If the solutions contained in j dominate those contained in i, replace i with the solutions contained in j.
    5. Abandon a proportion P_a of the worst nests and replace them with new nests.
    6. Retain the best solutions/nests.
    7. Increment t by one.
4. End.

According to the above pseudocode, the search algorithm begins by generating a random set of solutions for the nests. During each iteration, a cuckoo such as i is first selected randomly, and the fitness of its solution is calculated. In the next step, a second random nest, such as j, is selected, and the fitness of its solution is compared with that of nest i. If nest j dominates nest i, the set of solutions in nest j is considered superior to that in nest i, and the solutions of i are replaced accordingly. In the subsequent step, a proportion P_a of nests with the poorest fitness values is abandoned, and new sets of solutions are generated randomly. At the end of each iteration, the best solution sets are retained. These steps continue until the number of iterations reaches the predetermined value T. Accordingly, each neural network in the proposed ensemble model uses the above procedure to optimize its structure and weight vector. Ultimately, the structure used to forecast electricity consumption in the test samples is the one capable of minimizing the fitness function presented in Equation (9).

The final stage of the proposed method involves forecasting electricity consumption for the next time period using the neural-network ensemble model. In the proposed method, an averaging technique is employed to determine the final output of the ensemble system. The purpose of this technique is to reduce the error of individual learning models by combining their outputs so that the error associated with each neural network can be compensated for by the outputs of the other neural networks. Accordingly, in the final stage of the proposed method, each optimized neural network in the ensemble system forecasts the outputs of the test samples, after which the mean of the outputs generated by these models is considered the final estimated value.

3. Findings and Results

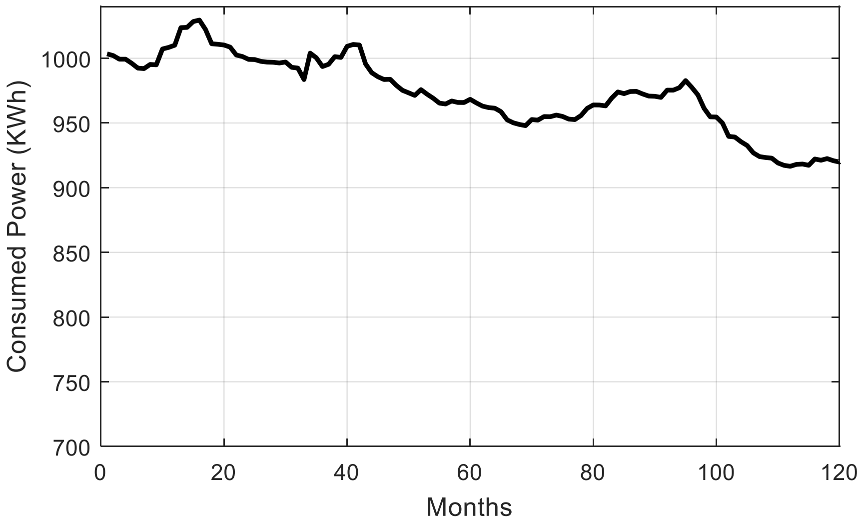
At this stage, the proposed method was evaluated, and its results were compared with those of other forecasting models. The proposed method was implemented using MATLAB software tools. The experiments were conducted using a real-world database containing the electricity consumption records of 10 households.

The database used in the proposed method was collected from the Ilam Electricity Department. This database contains the electricity consumption data of 10 households in Ilam. The electricity consumption histories of these households cover the period from approximately April–May 2015 to April–May 2025.

The database contains 120 data records for each household. These records describe the electricity consumption of each household during different months. Of the data available for each household, 96 samples were used for the learning phase and construction of the forecasting model, whereas the remaining samples were used for testing. Figure 4 shows variations in the monthly electricity consumption of one household in the database.

Figure 4

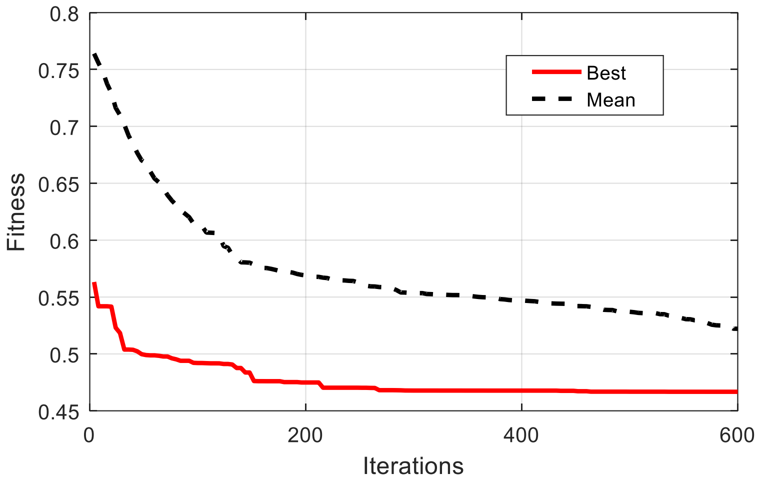
Variations in the monthly electricity consumption of one household in the database



The proposed method performs forecasting by combining the results generated by three multilayer perceptron neural networks. In this method, the Cuckoo Search Optimization Algorithm is employed to determine the optimal weight vector and structure of each neural network. The parameters used in the Cuckoo Search Algorithm to optimize the structure of the neural networks in the proposed method were as follows: number of iterations = 200, population size = 40, and nest-abandonment probability = 0.25. Figure 5 illustrates the variation in the fitness of the Cuckoo Search Algorithm during optimization of the neural-network structure.

Figure 5

Variation in the fitness of the Cuckoo Search Algorithm during neural-network structure optimization

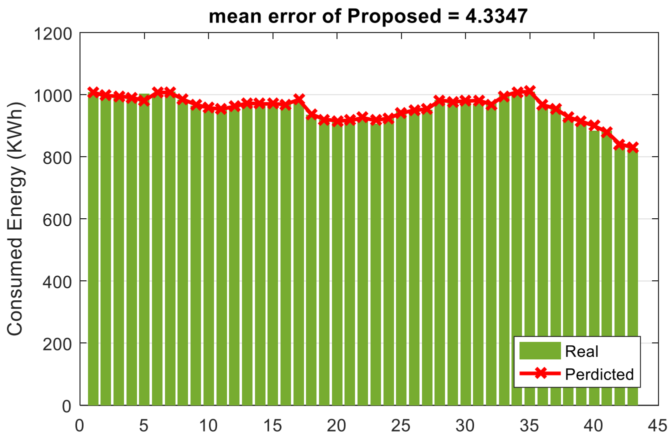


According to the figure above, the Cuckoo Search Algorithm can minimize forecasting error (fitness) in the learning model by efficiently searching the problem space and determining the optimal combination of neural-network weight values. The downward trend in the fitness curves indicates the continuous improvement of the discovered solutions and the movement of the population toward the global optimum.

Evaluation of the proposed method showed that the three optimal multilayer perceptron neural-network models for forecasting subscribers' electricity consumption contained 7, 8, and 10 neurons in their hidden layers, respectively. The training errors of these learning models were 0.3945, 0.4747, and 0.3912, respectively. No significant relationship was observed between the number of neurons in the neural networks and training error. Therefore, it cannot be concluded that a larger or smaller number of neurons in a multilayer perceptron neural network has a significant effect on the training error of the model. After determining the optimal structure and weight vector for these learning models, electricity consumption was forecast by each model, and the final output was subsequently calculated by averaging the outputs of all three models. Figure 6 presents the actual values and the values forecast by the proposed method, based on the combined results of three multilayer perceptron neural networks, for electricity consumption in one of the test samples.

Figure 6

Performance of the proposed method in forecasting household electricity consumption

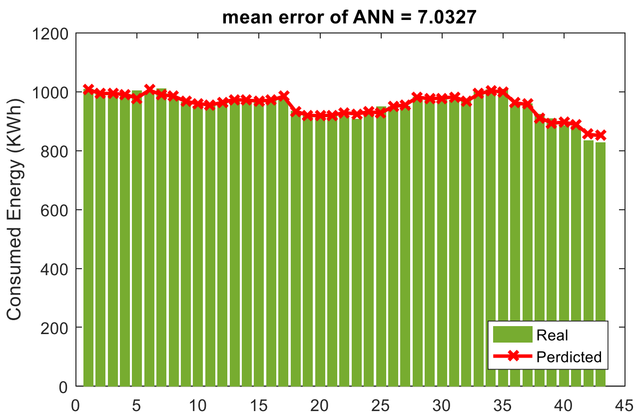


Based on the results obtained, the mean absolute error of the proposed method for the sample shown in Figure 6 was 4.33. To compare the performance of the proposed method, its error in forecasting household electricity consumption was compared with those of other forecasting algorithms, including an artificial neural network trained using the Levenberg–Marquardt algorithm, Gaussian process regression, and a classification and regression tree.

Figure 7 presents the actual values and the values forecast by the artificial neural network for monthly electricity consumption in one of the test samples. This neural network contains one hidden layer with 15 neurons.

Figure 7

Performance of the artificial neural network in forecasting household electricity consumption



Based on these results, the mean absolute error of this neural network for the test sample shown in Figure 7 was 7.03. Accordingly, the proposed method can reduce forecasting error by 38% compared with a conventional neural network. Thus, the proposed method, through the use of the Cuckoo Search Algorithm, can improve the performance of neural networks compared with conventional training functions.

Figures 8 and 9 present, respectively, the actual values and the values forecast by Gaussian process regression and the classification and regression tree for the same test sample.

Figure 8

Performance of the Gaussian process regression model in forecasting household electricity consumption

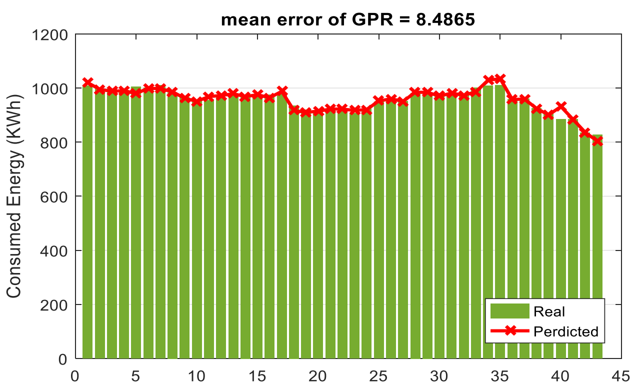
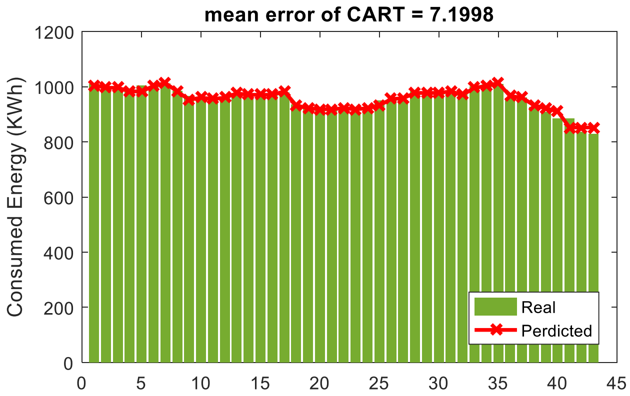


Figure 9

Performance of the classification and regression tree in forecasting household electricity consumption



These results indicate that the mean absolute error of the GPR model for the test sample presented in the figures above was 8.48. In addition, the classification and regression tree forecast electricity consumption for this sample with a mean error of 7.199. Based on these findings, the proposed method can substantially reduce forecasting error. This improvement can be attributed to two factors:

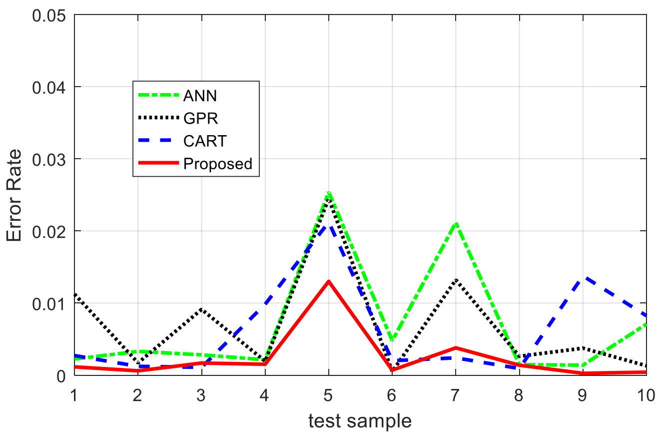
In the proposed method, a combination of multiple models is used to minimize forecasting error. In this setting, the learning models reduce one another's errors through the averaging technique.

Moreover, the use of the Cuckoo Search Optimization Algorithm makes it possible to obtain neural-network models with a high level of performance.

Figure 10 illustrates variations in forecasting error across different households (samples) in the database. In this figure, the horizontal axis represents the households contained in the database, whereas the vertical axis represents the forecasting error rate of the different algorithms in estimating each household's electricity consumption.

Figure 10

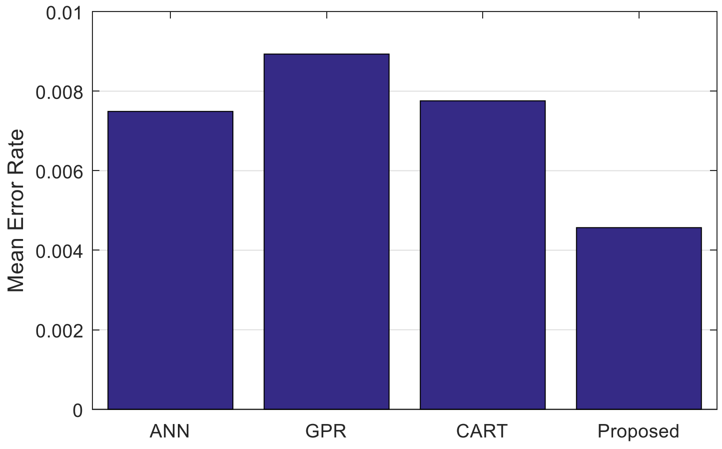
Variations in forecasting error across different households in the database



According to Figure 10, the proposed method exhibits a lower error rate for most samples. The minimum and maximum error rates of the proposed method across all database samples were 0.00021 and 0.01456, respectively. To examine this criterion more precisely, the mean error rates of the different algorithms can be evaluated. Figure 11 presents the error rates of the proposed method and the other algorithms used to forecast the electricity consumption of all households.

Figure 11

Mean error rate of the proposed method compared with the other methods



The error-rate criterion was calculated by dividing the difference between the learning model output (forecast electricity consumption) and the actual consumption value by the actual consumption value:

$$ErrorRate = \frac{1}{N}\sum_{i=1}^{N}\frac{|y_i - \hat{y}_i|}{\hat{y}_i} \qquad (10)$$

In the above equation, N denotes the number of test samples, whereas y_i and $\hat{y}_i$ represent the forecast and actual electricity consumption values for the ith test sample, respectively.

Based on the results obtained, the proposed method can effectively reduce the error rate in forecasting household electricity consumption. This outcome results from the use of the Cuckoo Search Optimization Algorithm to optimize the learning components of the proposed ensemble system. Table 1 presents the mean absolute error values of the proposed algorithm and the other learning models in forecasting household electricity consumption.

Table 1

Mean Absolute Error Obtained from the Performance Evaluation of the Proposed Method

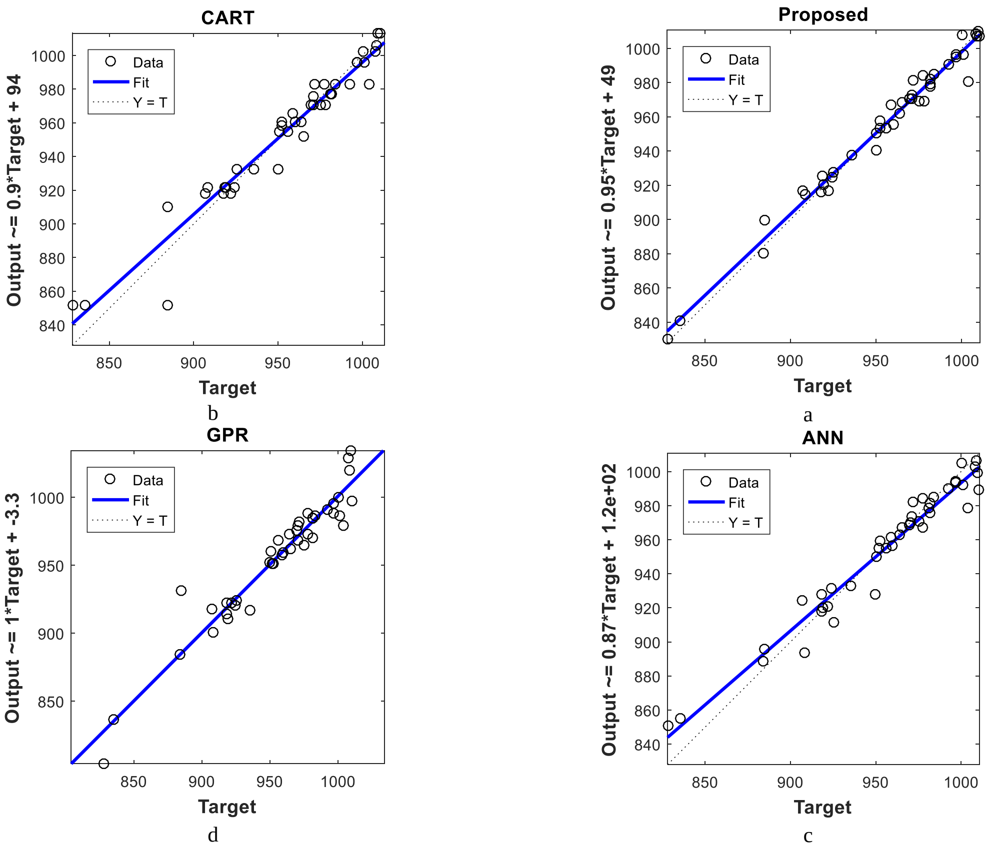
Household ID	CART	ANN	GPR	Proposed Method
1	7.8176	8.2415	13.6571	5.4260
2	7.1998	7.0327	8.4865	4.3347
3	4.4696	9.0560	12.3304	3.3795
4	11.1135	8.9369	25.8376	7.9625
5	1.6389	3.6016	6.0095	1.2004
6	17.5555	12.0860	49.1733	11.8833
7	14.8493	10.2491	16.7727	8.8393
8	1.9676	2.1811	2.2893	1.0898
9	11.3315	10.8027	19.7347	6.8769
10	13.8658	7.9653	17.0244	5.5715
Mean	9.1809	8.0152	17.1315	5.6563

The results obtained from the database evaluation presented in Table 1 indicate that the proposed method can reduce forecasting error in all cases. Therefore, the use of an optimization strategy can serve as an effective technique for improving the accuracy of multilayer perceptron neural networks. Based on the obtained results, the mean absolute error of the proposed method in forecasting electricity consumption across all database samples was 5.66, indicating a reduction of at least 29.4% in mean absolute error compared with the methods used for comparison. The regression plots of the proposed method and the other methods are presented in Figure 12.

(a) Proposed method
(b) CART
(c) ANN
(d) GPR

Figure 12

Regression plots of (a) the proposed method, (b) CART, (c) ANN, and (d) GPR

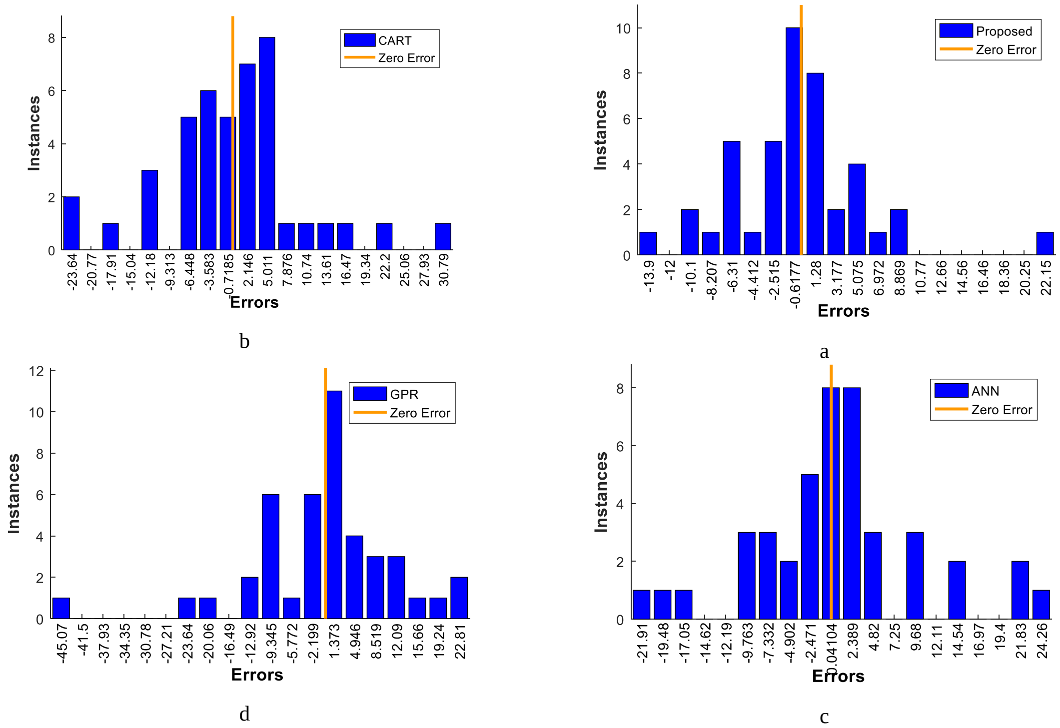


Examination of the regression plots presented in the figure above shows that the outputs of the proposed method are more consistent with the expected outputs; consequently, the forecast values exhibit less deviation from the fitted line. This figure demonstrates that the proposed method is capable of modeling subscribers' consumption patterns more accurately and that the consumption values forecast by the proposed method are closer to the actual consumption pattern of each household. The error histograms of the proposed method and the other methods used for comparison are presented in Figure 13.

(a) Proposed method
(b) CART
(c) ANN
(d) GPR

Figure 13

Error histograms of (a) the proposed method, (b) CART, (c) ANN, and (d) GPR



Comparison of the error histograms of the different methods in Figure 13 indicates that the error range of the proposed method is narrower than that of the other methods. This characteristic confirms that the use of an ensemble technique and optimization of the structure of multilayer perceptron neural networks in the proposed method can enhance the reliability of the forecast outputs. According to these plots, for more than two-thirds of the test samples, the error of the proposed method lies within the range of −0.6 to +1.2, indicating that, in most cases, the proposed method can forecast household electricity consumption with an error of less than 1.5 kWh. Achieving this level of error demonstrates the effectiveness of the proposed method in modeling household electricity consumption patterns and suggests that the proposed method can also be employed in real-world applications.

4. Discussion and Conclusion

The findings of this study demonstrated that the proposed hybrid forecasting framework, which integrates MRMR-based feature selection, Cuckoo Search optimization, and an ensemble of multilayer perceptron neural networks, can substantially improve the accuracy of monthly household electricity-consumption prediction. The most important result was that the proposed model achieved a mean absolute error of 5.6563 across the 10 households included in the database, outperforming the comparison models, including ANN, GPR, and CART. The corresponding mean absolute errors for ANN, GPR, and CART were 8.0152, 17.1315, and 9.1809, respectively. Therefore, the proposed method reduced the mean absolute prediction error by at least 29.4% relative to the competing approaches. In addition, the proposed method yielded the lowest prediction error for every household examined, indicating that the superiority of the model was not confined to a small subset of cases. These findings suggest that integrating feature-selection and optimization procedures with ensemble neural learning can improve both the accuracy and robustness of household electricity-consumption forecasting.

The superior performance of the proposed model is consistent with the broader literature emphasizing the effectiveness of machine-learning methods for energy-consumption prediction. Machine-learning approaches are particularly appropriate for electricity-demand forecasting because electricity-use patterns are typically nonlinear, temporally dependent, and influenced by interacting behavioral and environmental factors. Reviews of energy-consumption prediction have shown that neural networks and other machine-learning approaches can outperform conventional models when the underlying relationships cannot be adequately represented through simple linear structures (Mosavi & Bahmani, 2019). Similarly, the application of machine-learning models to building energy-demand and consumption prediction has demonstrated their usefulness for sustainable energy management and smart-city applications (Ardabili et al., 2019). The present findings reinforce these conclusions by showing that a machine-learning system specifically optimized for household monthly consumption can achieve lower forecasting errors than conventional ANN, GPR, and CART models.

An important finding of this study was that the highest forecasting performance was achieved using six selected features rather than the complete set of historical consumption variables. This result highlights the importance of feature selection in electricity-consumption forecasting. The MRMR algorithm is designed to retain variables that are strongly associated with the target while reducing redundancy among the selected predictors. In this study, applying this principle allowed the forecasting model to avoid the unnecessary inclusion of highly correlated or weakly informative historical consumption values. This finding is consistent with the logic underlying maximum relevance–minimum redundancy feature-selection approaches, which seek to improve predictive efficiency by balancing the relevance of individual variables with the redundancy among them (Zhao et al., 2019). The result therefore suggests that increasing the number of historical inputs does not necessarily improve forecasting accuracy and may instead introduce redundant information that complicates learning.

The relevance of this finding is particularly evident in monthly household electricity forecasting, where adjacent lagged consumption values may contain overlapping information. Consumption levels in consecutive months can be highly correlated because household routines, appliance ownership, and seasonal conditions often change gradually. Consequently, the simultaneous use of all lagged values may unnecessarily increase model dimensionality without contributing independent predictive information. By selecting six informative features, the proposed model appears to have preserved the most useful temporal information while reducing unnecessary complexity. This feature-selection strategy may also have contributed to the lower computational burden and improved generalization observed in the proposed system. The importance of efficiently representing temporal information is also reflected in recent studies applying sequence-oriented models to household energy forecasting, where historical patterns are treated as the primary source of predictive information (Chen, 2025; Hosseini, 2025).

The results also showed that the three optimized MLP networks contained 7, 8, and 10 hidden neurons, respectively, and achieved training errors of 0.3945, 0.4747, and 0.3912. No meaningful relationship was observed between the number of hidden neurons and training error. This finding indicates that increasing network size does not automatically lead to greater predictive accuracy. Neural-network performance depends not only on the number of neurons but also on the interaction between topology, connection weights, biases, input representation, and training strategy. A relatively compact architecture may therefore perform as effectively as or better than a larger architecture if its parameters are appropriately optimized. This observation supports the use of optimization-based model configuration rather than the manual assumption that more complex networks necessarily produce more accurate forecasts.

The Cuckoo Search Algorithm appears to have played an important role in the performance improvement achieved by the proposed model. The fitness curves showed a gradual decline during the optimization process, indicating continuous improvement in candidate solutions and convergence toward lower forecasting error. Cuckoo Search is particularly suitable for such problems because it combines global exploration with local refinement and requires relatively few control parameters. Previous research has shown that adaptive Cuckoo Search can provide effective optimization performance by improving the balance between exploration and exploitation (Mareli & Twala, 2018). The current findings are consistent with this evidence, suggesting that Cuckoo Search can be successfully employed to optimize both the topology and weight vectors of multilayer perceptron networks used for household electricity forecasting.

The performance of the proposed model further supports earlier research demonstrating the usefulness of population-based optimization in electricity-demand prediction. Long-term electricity-consumption forecasting using artificial cooperative search has shown that metaheuristic optimization can improve predictive performance by searching for model configurations that are difficult to obtain through conventional procedures (Kaboli et al., 2016). The present study extends this principle to monthly household electricity forecasting by simultaneously optimizing neural-network architecture and connection weights. The substantial reduction in error relative to the conventional ANN is particularly informative. For the representative test sample, the proposed model produced an MAE of 4.33 compared with 7.03 for the conventional ANN, corresponding to an error reduction of approximately 38%. This result indicates that optimization of the neural-network configuration can materially improve predictive performance beyond the use of a neural network alone.

Another important result concerns the advantage of ensemble averaging. The final prediction of the proposed system was obtained by averaging the outputs of three independently optimized MLP models. This procedure was intended to reduce model-specific errors and improve stability. The results suggest that this ensemble strategy was effective because the proposed model consistently produced lower errors across households than the individual comparison models. Even if one constituent neural network generates a relatively inaccurate prediction for a particular observation, averaging across multiple independently optimized networks can reduce the influence of that error. This mechanism may explain why the proposed method not only achieved lower average error but also exhibited a narrower error distribution in the histogram analysis. The observed stability is particularly important for practical forecasting applications because utilities require models that perform consistently across consumers rather than models that achieve high accuracy only for selected households.

The regression plots further confirmed the superiority of the proposed model. Predicted values generated by the hybrid system were more closely aligned with expected values and exhibited less deviation from the regression line than the values produced by CART, ANN, and GPR. This indicates that the proposed model captured the actual structure of household electricity-use patterns more effectively. The finding is consistent with research showing that residential electricity demand contains temporal dependencies that can be modeled using data-driven neural approaches (Rahman et al., 2018). Although recurrent neural networks are specifically designed for sequential information, the current results demonstrate that a properly optimized MLP ensemble can also capture useful temporal relationships when relevant lagged features are carefully selected.

The error histogram provides additional evidence regarding the reliability of the proposed method. For more than two-thirds of the test observations, prediction errors were located approximately within the range of −0.6 to +1.2, and in most cases the model was able to estimate household electricity use with an error below 1.5 kWh. This relatively concentrated error distribution indicates that the model was not only accurate on average but also stable across test observations. Such stability is important in residential demand forecasting because household electricity use may vary substantially over time. Forecasting models that produce occasional large deviations can create operational difficulties even when their average performance is acceptable. The proposed model's narrower error range therefore represents a practically relevant advantage.

The results can also be interpreted in light of the growing emphasis on temporal pattern recognition in residential electricity forecasting. Recent research on household monthly electricity consumption has shown that identifying similarities between historical consumption patterns can support accurate prediction (Hosseini, 2025). Likewise, LSTM-based forecasting has demonstrated the importance of learning temporal dependencies from historical residential electricity data (Chen, 2025). The proposed model follows a different computational strategy but is conceptually aligned with these studies because it uses historical electricity-consumption values as the principal predictive information. The contribution of the current approach lies in combining temporal feature selection with neural-network optimization and ensemble averaging, rather than relying exclusively on a single sequential-learning model.

The findings should also be considered within the broader context of environmental and energy-management challenges. Accurate household electricity forecasting can support more efficient generation scheduling, grid management, and demand-side planning. These functions are especially relevant in energy systems where electricity production remains closely associated with fossil-fuel consumption and environmental impacts. Previous research on Iran has demonstrated important relationships among economic growth, fossil-fuel consumption, and carbon dioxide emissions, emphasizing the environmental implications of energy use (Lotfalipour et al., 2010). Improved electricity-demand forecasting can indirectly contribute to more efficient resource utilization by reducing mismatches between generation and consumption and supporting better planning of energy resources.

The relevance of improved forecasting is further strengthened by Iran's energy-transition challenges. Although Iran possesses considerable renewable-energy potential, particularly for solar photovoltaic generation, the development of renewable energy has faced economic, policy, infrastructural, and institutional barriers (Gorjian et al., 2019). Greater penetration of variable renewable generation increases the importance of accurate demand forecasting because supply from renewable sources can fluctuate considerably. Improved knowledge of expected household demand can therefore facilitate better coordination between consumption and generation. Although the present study did not directly model renewable-energy integration, its findings suggest that accurate local demand forecasting may provide a useful analytical component for future smart-grid and renewable-energy management systems.

Climate variability is another factor that reinforces the importance of robust electricity forecasting. Residential electricity use is strongly influenced by heating and cooling requirements, and these may change as climatic conditions vary. Research examining electricity consumption in relation to climate change has emphasized that changes in environmental conditions can substantially affect demand patterns (Zhafran & Miefthawati, 2024). The proposed model used historical consumption rather than explicit climate variables; nevertheless, seasonal climatic effects may be partially embedded in historical monthly consumption patterns. The successful forecasting performance therefore suggests that the model was able to recover some of these recurring temporal structures indirectly. Incorporating explicit meteorological information in future versions of the model may further improve accuracy under changing climatic conditions.

The comparison with GPR and CART also provides useful methodological evidence. GPR produced the largest average MAE, whereas CART also exhibited substantially higher errors than the proposed model. These results suggest that the household consumption patterns present in the dataset may be too nonlinear or heterogeneous to be adequately represented by these individual learning approaches. The proposed model benefits from multiple layers of adaptation: MRMR reduces irrelevant information, Cuckoo Search optimizes network structure and weights, and averaging combines multiple forecasts. The observed improvement therefore likely reflects the cumulative effect of these components rather than the contribution of a single technique. This finding is consistent with the broader development of energy forecasting toward hybrid models in which complementary computational procedures are integrated to improve predictive performance (Ardabili et al., 2019; Mosavi & Bahmani, 2019).

Overall, the results indicate that an optimized ensemble of multilayer perceptron neural networks can provide an effective solution for monthly household electricity-consumption forecasting when combined with appropriate feature selection. The proposed system achieved lower forecasting errors, greater consistency, and closer agreement with actual consumption values than the comparison methods. Its performance supports the use of metaheuristic neural optimization and compact feature subsets in household energy modeling and demonstrates that relatively small datasets can still support accurate forecasting when the model architecture is carefully designed and optimized.

This study has several limitations that should be considered when interpreting the findings. First, the dataset consisted of electricity-consumption records from only 10 households in Ilam, which limits the extent to which the results can be generalized to larger and more heterogeneous populations. Second, the model primarily relied on historical electricity-consumption values and did not explicitly incorporate potentially influential variables such as outdoor temperature, humidity, household size, income, dwelling characteristics, appliance ownership, occupancy schedules, electricity tariffs, or behavioral factors. Third, the database covered monthly observations, and therefore the model was not evaluated for shorter forecasting intervals such as hourly or daily consumption. Fourth, the proposed model was evaluated using a specific set of comparison algorithms, and alternative deep-learning, boosting, or probabilistic forecasting methods were not examined. Finally, although the proposed optimization procedure improved prediction performance, the computational cost associated with independently optimizing multiple neural networks may become substantial when the number of households, input variables, or candidate network structures increases.

Future studies should evaluate the proposed framework using substantially larger datasets collected from different cities, climatic regions, and household groups in order to determine its generalizability. Researchers could also incorporate meteorological, demographic, behavioral, economic, and building-related variables into the feature-selection stage and assess whether these factors improve prediction beyond historical consumption alone. The proposed model could be extended to hourly, daily, and seasonal forecasting tasks and compared with LSTM, GRU, transformer-based architectures, random forests, gradient boosting, and other advanced ensemble models. Future research could also investigate alternative feature-selection procedures and compare different metaheuristic optimization algorithms for configuring neural-network topology and weights. Another useful direction would be to develop adaptive or online versions of the proposed system capable of updating model parameters when new consumption data become available. Probabilistic forecasting and prediction intervals should also be considered to quantify uncertainty and provide electricity utilities with information about the range of likely future demand rather than only point estimates.

Electricity distribution companies can use optimized household forecasting models to improve monthly demand planning, allocate generation and distribution resources more efficiently, and identify periods in which unusually high consumption is likely to occur. Utilities may implement the proposed approach as part of smart-metering or energy-management platforms in which recent consumption records are automatically processed and used to update household-level forecasts. Feature-selection procedures should be incorporated into operational forecasting systems to reduce unnecessary data processing and focus computational resources on the most informative consumption indicators. Utilities should also consider using ensembles of independently optimized forecasting models rather than relying on a single neural network, particularly when prediction reliability is important for planning decisions. At the managerial level, household-level forecasts can support targeted demand-management programs, infrastructure maintenance planning, load-shifting initiatives, and more efficient procurement of electricity. The model may also be integrated into decision-support systems for local grid management, allowing planners to identify consumption trends early and implement preventive measures before demand pressures affect network performance.

Authors' Contributions

Authors contributed equally to this article.

Declaration

In order to correct and improve the academic writing of our paper, we have used the language model ChatGPT.

Transparency Statement

Data are available for research purposes upon reasonable request to the corresponding author.

Acknowledgments

We would like to express our gratitude to all individuals helped us to do the project.

Declaration of Interest

The authors report no conflict of interest.

Funding

According to the authors, this article has no financial support.

Ethics Considerations

In this research, ethical standards including obtaining informed consent, ensuring privacy and confidentiality were considered.

References

Ardabili, S., Mosavi, A., & Várkonyi-Kóczy, A. R. (2019). Building Energy Information: Demand and Consumption Prediction with Machine Learning Models for Sustainable and Smart Cities. In (pp. 191-201). Springer.

Chen, Y. (2025). Research on Home Electricity Prediction Based on LSTM. *Applied and Computational Engineering*. https://doi.org/10.54254/2755-2721/2025.tj23583

Gorjian, S., Zadeh, B. N., Eltrop, L., Shamshiri, R. R., & Amanlou, Y. (2019). Solar Photovoltaic Power Generation in Iran: Development, Policies, and Barriers. *Renewable and Sustainable Energy Reviews*, *106*, 110-123.

Hosseini, B. (2025). Forecasting Household Monthly Electricity Consumption Using the Similar Pattern Algorithm. *Academia Green Energy*. https://doi.org/10.20935/acadenergy7500

Kaboli, S. H. A., Selvaraj, J., & Rahim, N. A. (2016). Long-Term Electric Energy Consumption Forecasting via Artificial Cooperative Search Algorithm. *Energy*, *115*, 857-871.

Lotfalipour, M. R., Falahi, M. A., & Ashena, M. (2010). Economic Growth, CO2 Emissions, and Fossil Fuels Consumption in Iran. *Energy*, *35*(12), 5115-5120.

Mareli, M., & Twala, B. (2018). An Adaptive Cuckoo Search Algorithm for Optimization. *Applied Computing and Informatics*, *14*(2), 107-115.

Mosavi, A., & Bahmani, A. (2019). Energy Consumption Prediction Using Machine Learning: A Review. *Preprints*.

Rahman, A., Srikumar, V., & Smith, A. D. (2018). Predicting Electricity Consumption for Commercial and Residential Buildings Using Deep Recurrent Neural Networks. *Applied Energy*, *212*, 372-385.

Zhafran, F., & Miefthawati, N. P. (2024). Forecasting Electricity Consumption and Its Relationship with Climate Change in Pekanbaru City. *Protek: Jurnal Ilmiah Teknik Elektro*. https://doi.org/10.33387/protk.v11i3.7309

Zhao, Z., Anand, R., & Wang, M. (2019). Maximum Relevance and Minimum Redundancy Feature Selection Methods for a Marketing Machine Learning Platform.