Forecasting-Driven Demand Response Management in Smart Grids Using Neural Networks

Rushikesh V. Pandya[1], Manan M. Desai[2], Urvisha P. Vala[3]

[1]Assistant Professor, Electrical Engineering Department, School of Engineering and Technology, Dr. Subhash University, Junagadh, Gujarat, India.

[2]Head of Department, Electrical Engineering Department, School of Engineering and Technology, Dr. Subhash University, Junagadh, Gujarat, India.

[3]Student, Electrical Engineering Department, School of Engineering and Technology, Dr. Subhash University, Junagadh, Gujarat, India

DOI: https://doi.org/10.51584/IJRIAS.2025.1005000129

Received: 09 June 2025; Accepted: 13 June 2025; Published: 25 June 2025

ABSTRACT

Demand Responses Management plays a pivotal role on the advancement of smart grid by promoting efficient energy utilization and optimizing overall system performance. A critical aspect of DRM is the accurate forecasting of household electricity demand, which is essential for effective power system planning, operational efficiency, and informed decision-making. Precise demand predictions contribute to cost reduction, risk mitigation, and strategic energy management. This study proposes an neural network-based forecasting Focus specifically designed for load side management in smart grid. It explores application of advanced Artificial Neural Network (ANN) architectures to model and predict electricity consumption patterns. The performance of the proposed ANN models is rigorously evaluated and benchmarked against three conventional statistical methods.

Keywords: Bode; neural mesh; Sharp grid

INTROMISSION

Sharp grids play a pivotal role in safe the safe, efficient, and true operation of power system, significantly donate to the reduction of energy losses across the energy network. However, neo smart grids face numerous goods and technical challenges in their mission to get electricity securely and cost-benefit to end users. Between the primary challenges are load-flow data, energy plan, and system control. [7]To achieve best planning within operation, accurate forecasting models are essential.

Over the past tenner, load forecasting has rapidly emerged as a critical area of research in the smart grid domain. A wide array of forecasting techniques has been proposed, with mathematical models showing promising accuracy. These models are designed to minimize assess errors between predicted and actual energy demand rate.

Demand forecasting in smart grids is crucial for several reasons:

- Cut unit production costs and maintaining the efficiency of energy infrastructure.
- Supporting the scheduling of high-risk care operations and energy reserve management.
- Providing essential data for design and enhancing power supply effectiveness.

Prose Review on Forecasting Models for Load Forecasting

Time series forecasting involves analyzing historical data to predict future values based on time-dependent patterns. These models can be classified by their forecasting horizon [2]:

The primary difference among these forecasting horizons lies in the magnitude of the variables considered, rather than in the[3] specific forecasting techniques used (Hernandez et al. [3]). However, time series models also have limitations. Not all models are universally applicable, and it is the responsibility of analysts to understand the boundaries of their tools.

Dynamic models, which incorporate time-dependent variables and random inputs, are used to model dynamic system behavior. They include:

- **Auto-Regressive and Step Average Models:** These combine past values and error terms to predict future outcomes. Auto-Regressive Integrated Moving Average (ARIMA) models extend ARMA by including differencing to account for non-stationary data.
- **State-Space Models:** Suitable for complex dynamic systems, these models use a mathematical framework involving input, output, and state variables—typically described by first-order differential equations. They are widely used in ecological, biological, and engineering systems.

A comprehensive review by Czapaj et al. [4] analyzed 48 reprint spanning 265 Foretell models from 1997 to 2018. The study assessed models based on Mean Absolute Percentage Error (MAPE) and introduced a novel approach for selecting suitable forecasting models. The top 11 technique identified include. Note that the repeated instance of ANN refers to distinct architectures. While FGRM and GRM utilize explanatory variables, the remaining eight are categorized as auto-regressive models.

Prose Review on Auto-Regressive Approaches for Load Forecasting

The results of previous reviews strongly support the[7] considerable potential of auto-regressive (AR) approaches in forecasting power demand. Employing such models enables transmission system operators to achieve higher forecasting accuracy and operational efficiency. Among the AR models, the **Auto-Regressive Integrated Moving Average (ARIMA)** model is particularly favored for its simplicity and effectiveness in short-term forecasting.

Despite its advantages, ARIMA has two significant limitations (Khashei et al. [8]):

- **Linear Assumption:** ARIMA assumes that a vector future value is a linear function of its past values and associated random errors. This assumption fails when the underlying system exhibits non-linear behavior—a characteristic common in real-world energy systems (Zhang et al. [9]). In such cases, ARIMA's predictive performance deteriorates.
- **Data Requirement:** ARIMA models demand extensive historical data for optimal performance. A minimum of 50–100 data points is generally necessary to produce reliable forecasts.

To overcome these constraints, **hybrid models** have been proposed. By combining multiple forecasting techniques, it is possible to harness their respective strengths and mitigate individual weaknesses. One such hybrid framework integrates the core concepts of **ARIMA, artificial neural networks (ANNs), and fuzzy logic models**. In this architecture:

- ANNs address the **non-linearity** of the data.
- Fuzzy logic techniques improve **adaptability** and help in managing cases with limited data. This integration results in a more robust and accurate forecasting system (Khashei et al. [8]).

Prose Review on ANN Approaches for Load Forecasting

Recent developments have seen a surge in the use of **artificial neural networks (ANNs)** for load Foretell, offering significant advancements over traditional statistical methods.

Need and Novelty of Our Approach

The continued evolution of [2]machine learning (ML)-based forecasting techniques addresses diverse challenges in energy demand prediction, particularly with advances in neural network design and application.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue V May 2025

The primary objective of this study is Several challenges persist in this domain:

- Ensuring prediction accuracy.
- Achieving optimal model training performance.
- Reducing forecast error margins.

To address these[3] challenges, we propose and analyze distinct models:

- **Auto-Atavistic Model:** Relies on previous energy consumption values within a defined time window.
- **Hybrid Auto-Regressive Model:** Integrates both historical consumption and input feature data for enhanced prediction accuracy.

Each model is assessed across various **multi-layer ANN architectures**, and evaluated using two **real-world smart grid datasets**. Our comparative analysis provides practical recommendations for selecting optimal forecasting strategies in smart grid environments.

The selected forecasting frameworks aim to:

- Promote energy conservation.
- Support load balancing and demand-response mechanisms.
- Enable effective financial planning for distributed energy producers and utility companies.

ANN Bode Approach for Demand Response Management

Artificial Neural Network (ANN) techniques represent a specialized of machine learning methods widely used in predictive analytics across multiple domains. Their strength lies in their ability to model **non-linear relationships** and approximate **complex functions** with high precision, making them particularly suitable for **energy load forecasting** in smart grids.

Among the various ANN architectures, **Recurrent Neural Networks (RNNs)** stand out due to their inherent capability to **capture temporal dependencies** within sequential data. Unlike traditional feedforward networks, RNNs incorporate memory by retaining information from previous time steps and using it to influence current outputs. This characteristic enables RNNs to learn complex time-dependent patterns, which is essential in modeling fluctuating energy demands (Figure 2) [2].

In this study, RNNs are applied to estimate the **unsure smart grid energy load** at a given time step yty_tyt. The model utilizes **historical input variables** ptp_tpt, observed over previous time steps t=1,2,…,Tt = 1, 2, \dots, Tt=1,2,…,T, for r=1,2,…,Rr = 1, 2, \dots, Rr=1,2,…,R features (e.g., hour of day, day type, temperature, number of appliances). These inputs are processed to generate an estimate y^t\hat{y}_ty^t, representing the **average energy demand** at time ttt.

Each layer lll of the neural network, where l=1,2,…,Ll = 1, 2, \dots, Ll=1,2,…,L, may employ a **distinct activation function** fl(·)f_l(\cdot)fl(·), designed to capture non-linear transformations of the data. The number of neurons in each layer is denoted by SlS_lSl, and the activation functions serve to model the output of each neuron within their respective layers.

The overall network can be represented as a **composite function** F(pt)F(p_t)F(pt), which transforms the input vector into a final output prediction y^t\hat{y}_ty^t. This transformation can be mathematically expressed as:

The network architecture effectively maps the **multi-dimensional input space** into a **single-valued energy demand forecast**, leveraging the sequential learning capabilities of RNNs. This methodology allows for a **data-driven prediction framework** capable of handling real-time smart grid inputs and facilitating accurate demand response management.

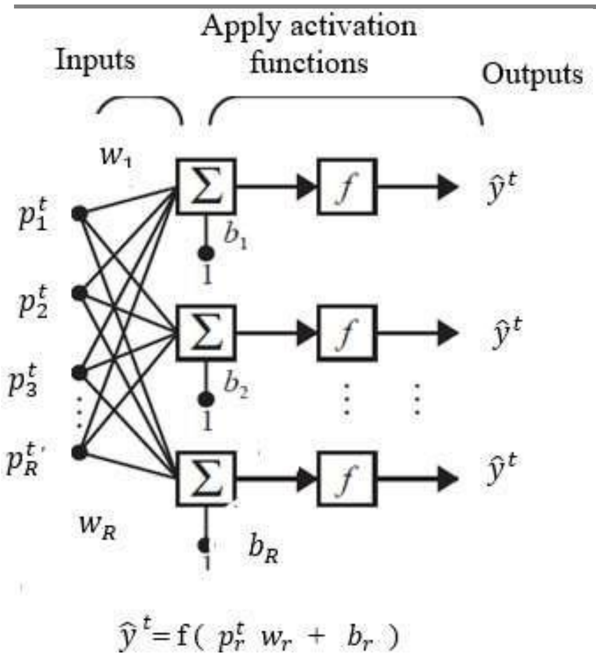

Figure 1. Neural Network Sample Architecture[3].

Steepest Descent Methods

The **steepest descent method** is a classical optimization technique used for training artificial neural networks. In this scheme, **mass updates** are applied after each fore pass through the network, based on the gradient of the error function. The weight update equations are as follows:

Data Put, Data Formatting, and Component Analysis

This study utilizes two real-world smart grid energy consumption data sets collected from residential households in Toronto, Canada:

Tower 1

- Contains **daily energy consumption data** from **1082 households**.
- Covers a **36-month period**: 2019, 2020, and 2021.
- Each household contributes **one observation**.
- Features include:
  - Household location (coordinates),
  - Number of electric appliances,
  - Day order (ordinal number of the day in a year),
  - Season,
  - Day of the week,
  - Time period of measurement.

Tower 2

- Spans **60 months**: 2017 to 2021.
- Features are similar to Data Set 1, with additional granularity:
  - Hourly measurements within defined time periods,
  - More detailed time segmentation (period × hour combinations).

Feature Encoding

- **Seasons** are defined as:
  - Season 1: November–February

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue V May 2025

  - Season 2: March–June
  - Season 3: July–October
- **Days of the week**: encoded as integers (Monday = 1, ..., Sunday = 7).
- **Time periods**:
  - Period 1: 00:00–05:59
  - Period 2: 06:00–11:59
  - Period 3: 12:00–17:59
  - Period 4: 18:00–23:59
- In **Data Set 2**, six measurement hours (1–6) are considered within each period.

Data Audio and Cleaning

To prepare the datasets for ANN training, raw data were transformed[6] into **input-output time series samples**. This preprocessing step included addressing missing or erroneous data. In **Data Set 1**, 14 entries out of 1082 recorded **zero kilowatt-hour (KWH)** consumption values, likely due to **sensor or recording errors**.

These anomalous entries were corrected by replacing them with **context-aware average values**, derived from:

- Comparable weekdays,
- Corresponding seasons,
- Similar time periods.

This cleaning step ensured consistent and accurate inputs for model training.

Relation Analysis

A **Static its matrix** was computed for **Residential Area 1** to examine the linear relationships between energy consumption and other features. The results [8]revealed that:

- **Number of appliances** exhibits the **strongest positive correlation** with energy consumption (KWH).
- Other features (e.g., day order, season, day of the week, time period) also show **positive correlations**, but with comparatively **lower impact**.

These insights validate the relevance of appliance count as a **key predictor** in the forecasting model and justify its inclusion in all ANN-based models evaluated in this study.

Trend and Seasonality

Depending on the behavior of seasonal variations in relation to[8] the series level, **additive** or **multiplicative** decomposition approaches may be applied:

- An **additive decomposition** is suitable when seasonal fluctuations and residual variance remain **constant over time**, regardless of the level of the series.
- A **multiplicative decomposition** is appropriate when these fluctuations are **proportional** to the series level, indicating increasing variance as the trend increases.

The general formulations for time series decomposition are as follows:

- **Additive model**:

  yt=Tt+St+Rty_t = T_t + S_t + R_tyt=Tt+St+Rt

- **Multiplicative model**:

  yt=Tt×St×Rty_t = T_t \times S_t \times R_tyt=Tt×St×Rt

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue V May 2025

where:

- yty_tyt is the observed time series at time ttt,
- TtT_tTt is the **trend component**, representing long-term progression,
- StS_tSt is the **seasonal component**, capturing periodic fluctuations,
- RtR_tRt is the **residual (or irregular)** component, containing noise and unexplained variance.

In this study, an **additive seasonal decomposition** was applied to **Toronto Data Set 1**, and the results (Figure 3) revealed a **significant seasonal effect** on electricity consumption. Notably, there is a **sharp increase in energy demand** during **Season 1 (November–February)**, aligning with the **cold winter months in Canada**, which typically drive higher heating-related electricity usage.

To statistically validate the presence of seasonality, a **Dickey-Fuller test** was conducted on the electricity consumption series. The resulting **p-value of 0.07** exceeds the standard 0.05 threshold, indicating **non-stationarity** and **confirming the presence of a seasonal component** in the data.

These findings substantiate the inclusion of **seasonal indicators** in the neural network input variables and highlight the importance of incorporating **seasonal dynamics** into forecasting models for smart grid energy demand.

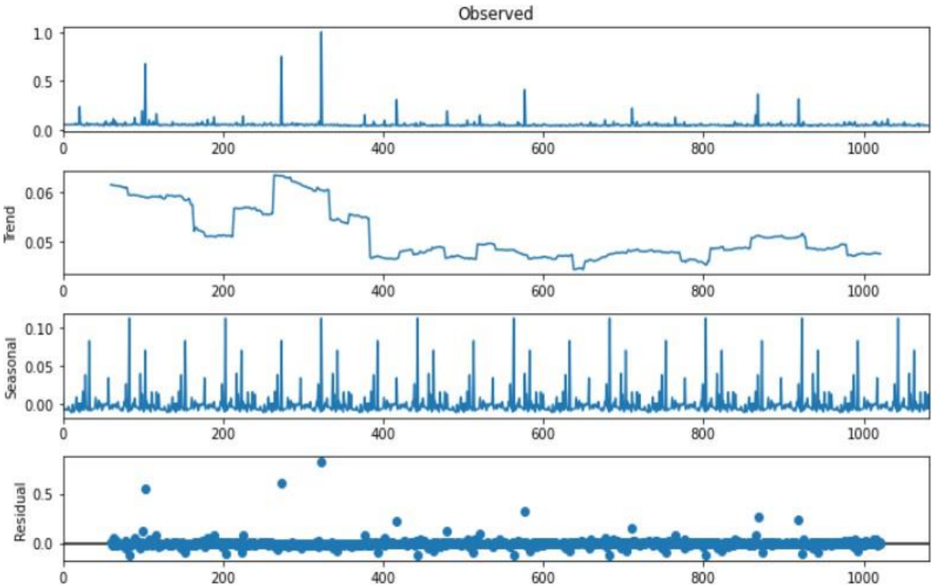

Figure 3. Additive Trend-Seasonality Decomposition Results[4].

ANN Forecasting Experimental Results

This section presents a comparative evaluation of the artificial neural network (ANN) forecasting model against three benchmark and optimization algorithms, including **Adam**, **Adagrad**, and **Adamax**, to assess the forecasting performance.

Auto-Regressive Model

predict future[5] values based on past observations, assuming that current values are linearly dependent on previous time steps and an error term. The AR model of order rrr uses the past rrr values to forecast the current observation. Its formulation is as follows:

Where yt−ry_{t-r}yt−r represents lagged values (up to 6 hours) of the dependent variable.

Due to the data structure of **Data Set 1** (with only one observation per household), the AR model is not applicable. Therefore, AR modeling is **only implemented on Data Set 2**, with delays ranging from 1 to 6 hours to capture within-period consumption dynamics.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue V May 2025

Hybrid Model

Hybrid models combine multiple model [5]types to leverage their respective strengths. Our proposed hybrid model integrates features from both the time series and auto-regressive approaches, aiming to capture both external variables and temporal dependencies.

The hybrid model is expressed as:

yt=F(pt,yt−r)=fL(…fl(…f1(pt,yt−r)))y_t = F(p_t, y_{t-r}) = f_L(\ldots f_l(\ldots f_1(p_t, y_{t-r})))yt=F(pt ,yt−r)=fL(…fl(…f1(pt,yt−r)))

As with the AR model, the hybrid model is **applied exclusively to Data Set 2**, given the availability of multi-hour measurements.

The experimental results are based on a **three-layer ANN**, where we systematically varied:

- The number of neurons per layer: S1S_1S1, S2S_2S2, and S3S_3S3,
- Activation functions: f1f_1f1, f2f_2f2,
- Optimizers: Adam, Adagrad, and Adamax.

Performance was measured using **Root Mean Squared Error (RMSE)**, a widely accepted metric to assess predictive accuracy. The RMSE is computed as:

RMSE=1N∑t=1T(yt−y^t)2\text{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^{T} (y_t - \hat{y}_t)^2}RMSE=N1 t=1∑T(yt−y^t)2

Where:

- yty_tyt is the observed value,
- y^t\hat{y}_ty^t is the predicted value,
- NNN is the total number of observations.

The training set was fixed at **75% of the total data**, with the remaining 25% used for validation. Although the optimal training-validation ratio remains a subject of ongoing debate, this choice aligns with standard practice and the primary focus of this study.

Summary of Findings (see Table 3):

- The **best-performing architecture** used **ReLU** as both f1f_1f1 and f2f_2f2, combined with the **Adam optimizer**, yielding a **lowest RMSE of 1.201** with the following configuration:

  - S1=64S_1 = 64S1=64,
  - S2=32S_2 = 32S2=32,
  - S3=16S_3 = 16S3=16.

- **Adam** consistently outperformed **Adagrad** and **Adamax** in terms of stability and RMSE across configurations.

- Architectures using **tanh** and **log-sigmoid** activation functions exhibited **higher average RMSEs**, with the best tanh-log-sigmoid configuration yielding an RMSE of **1.793**.

Visualization (Figure 4) displays the comparison between actual consumption (blue curve) and predicted values (green curve) for **Data Set 1** using the optimal ANN configuration described above.

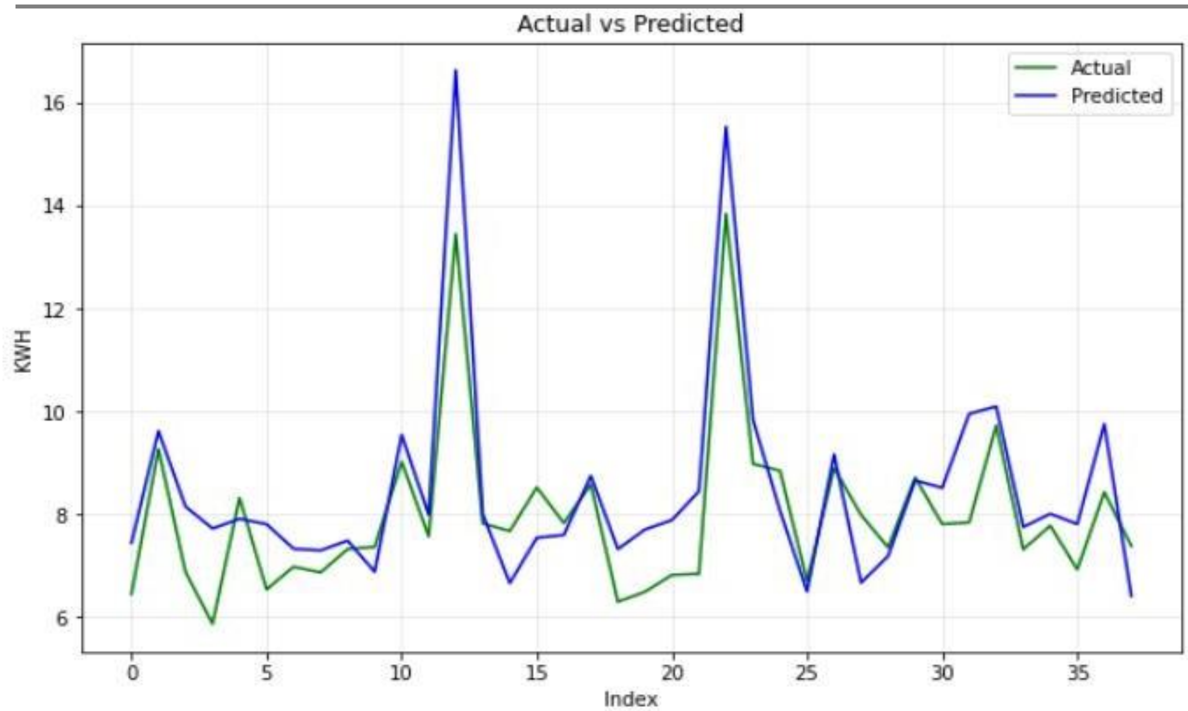

Figure 4. Actual and Predicted Measurement and Prediction on Data Set 1[7].

For the **time series model**, the most effective configuration—using **ReLU** as both activation functions f1f_1f1 and f2f_2f2—achieves the **lowest RMSE of 0.977**. This result is obtained using a deep architecture with:

- S1=256S_1 = 256S1=286 cell in the first layer,
- S2=128S_2 = 128S2=127 cell in the second layer,

the optimal configuration also leverages **ReLU** for f1f_1f1 and f2f_2f2. It yields the **minimum RMSE of 0.807**, with a more compact architecture:

- S1=32S_1 = 32S1=32,
- S2=16S_2 = 16S2=16,

The **hybrid model**, which integrates both time-series and autoregressive features, again employs **ReLU** activations and achieves the **lowest overall RMSE of 0.697**. This performance is recorded using the same deep architecture as the best-performing time series model, with:

- S1=256S_1 = 256S1=256,
- S2=128S_2 = 128S2=128,
- S3=64S_3 = 64S3=64,
- and the **Adamax parser**.

Across all models and architectural variations, **Adam** and **Adamax** optimizers demonstrate superior stability and performance compared to **Adagrad**, particularly when the number of neurons varies in the following ranges:

- S1∈[32,256]S_1 \in [32, 256]S1∈[32,256],
- S2∈[16,128]S_2 \in [16, 128]S2∈[16,128],
- S3∈[8,64]S_3 \in [8, 64]S3∈[8,64].

Overall, the hybrid model logical delivers the lowest RMSE, indicating[7] its robustness and effectiveness in forecasting energy consumption by capturing both temporal and feature-based dynamics.summarizes the forecasting results for **Residential Area 4** using the three model types: time series, auto-regressive, and hybrid. The findings reveal that the final architecture—configured with **ReLU** as both activation functions (f1f_1f1 and f2f_2f2)—consistently delivers the most accurate predictions. In the **time series model**, the optimal configuration achieves an RMSE of **1.640**, utilizing:

- S1=32S_1 = 32S1=32 cell in the first layer,
- S2=16S_2 = 16S2=16 cell in the second layer,
- S3=8S_3 = 8S3=8 cell in the third layer, and the **Adam parser**.

the best performance is attained with an RMSE of **0.158**, using:

- S1=128S_1 = 128S1=128,
- S2=64S_2 = 64S2=64,
- S3=32S_3 = 32S3=32,S1=256S_1 = 256S1=256,
- S2=128S_2 = 128S2=128,
- S3=64S_3 = 64S3=64,
- and again, the **Adamax parser**.

Across all four neural network architectures and all three model types, **Adam** and **Adamax** optimizers demonstrate **greater stability and reliability** than **Adagrad**. This consistency is observed even as the architecture scales:

- S1∈[32,256]S_1 \in [32, 256]S1∈[32,256],
- S2∈[16,128]S_2 \in [16, 128]S2∈[16,128],
- S3∈[8,64]S_3 \in [8, 64]S3∈[8,64].

In summary, the auto-regressive model yields the lowest RMSE (0.158), while[7] the **hybrid model remains the most consistently robust** across varied configurations, affirming its capability to usefull capture both historical patterns and dynamic dependencies in the residential load data.

SUMMARY OF RESULTS

For both data sets 1 and 2, the final accuracy and precision in our three-layer Artificial Neural Network models is consistently achieved using the final architecture, which employs as the activation function for both hidden layers (**f1:ReLU,f2:ReLUf_1: \text{ReLU}, f_2: \text{ReLU}f1:ReLU,f2:ReLU**). For **data set 1**, the **Adam optimizer** generally provides the most accurate predictions, whereas for **data set 2**, the **Adamax optimizer** is predominantly more effective.

Specifically, in **data point set 2**:

- **Time Series Model**: The **Adamax** optimizer delivers the best performance for **house areas 1, 2, 3, 5, and 6**, while **residential area 4** achieves its best accuracy using the **Adam** optimizer.

Among all model types, the **hybrid model exhibits the lowest RMSE values**, making it the most accurate. This model effectively integrates both past observations and time-dependent patterns, providing a more robust forecast of residential electricity demand.

Additionally, **Stochastic Gradient Descent** proved unable for both data sets. Due to the sparsity of gradients and the presence of many zero elements in the subsets, the randomness inherent in SGD undermines its convergence and accuracy in this context.

CONCLUSIONS

This study proposed an artificial neural network (ANN)–based forecasting near for **demand-side management** in smart grid systems. The experimental model was evaluated on two real-world electricity consumption data sets using various three-layer neural network architectures integrated into **time series**, **auto-regressive**, and **hybrid** statistical models.

While the **time series model** offers the fastest computation time, it does so at the cost of reduced prediction accuracy. In contrast, the **hybrid model**—though computationally more intensive—delivers **significantly higher accuracy**, making it ideal for practical implementations in energy forecasting systems.

By selecting the optimal neural architecture and statistical model, energy providers can better manage electricity supply and demand. This approach also offers potential benefits in energy conservation, strategic financial planning, and supporting consumers initiating small-scale energy production.

REFERENCES

1. Raza , M.Q.; Khosravi, A. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. Renew. Sustain. Energy Rev. **2015**, 50, 1352–1372.
2. Hippert, H.S.; Pedreira, C.E.; Souza, C.R. Neural Networks for Short-Term Load Forecasting: A review and Evaluation. IEEE Trans. Power Syst. **2001**, 16, 44–51.
3. Hernandez, L.; Baladron, C.; Aguiar, J.; Belen, C.; Sanchez-Esguevillas, A.J.; Lloret, J.; Massana, J. A Survey on Electric Power Demand Forecasting: Future Trends in Smart Grids, Microgrids and Smart Buildings, A Survey on Electric Power Demand Forecasting: Future Trends in Smart Grids, Microgrids and Smart Buildings. IEEE Commun. Surv. Tutorials **2014**, 16, 1460–1495.
4. Czapaj, J.S.M.; Kaminski, R. A review of auto-regressive methods applications to short-term demand forecasting in power systems. Energies **2022**, 15, 6729.
5. Kheirkhah, A.; Azadeh, A.; Saberi, M.; Azaron, A.; Shakouri, H. Improved estimation of electricity demand function by using of artificial neural network, principal component analysis and data envelopment analysis. Comput. Ind. Eng. **2013**, 64, 425–441.
6. Wang, Y.; Bielicki, J.M. Acclimation and the response of hourly electricity loads to meteorological variables. Energy **2018**, 142, 473–485.
7. Rana, M.; Koprinska, I. Forecasting electricity load with advanced wavelet neural networks. Neurocomputing **2016**, 182, 118–132.
8. Khashei, M.; Bijari, M.; Ardali, G.A.R. Improvement of auto- regressive integrated moving average models using fuzzy logic and artificial neural networks (anns). Neurocomputing **2009**, 72, 956–967.
9. Zhang, G.; Patuwo, B.E.; Hu, M.Y. Forecasting with artificial neural networks: The state of the art. Int. J. Forecast. **1998**, 14, 35–62.
10. Dewangan, F.; Abdelaziz, A.Y.; Biswal, M. Load Forecasting Models in Smart Grid Using Smart Meter Information: A Review. Energies **2023**, 16, 1404.