

SHORT TERM LOAD FORECASTING USING GROUP METHOD OF DATA HANDLING OPTIMIZED BY ARTIFICIAL BEE COLONY TECHNIQUE

Musa, M.^{1*}, Hassan, A.¹ and Hassan, S.M.²

¹Department of Electrical and Electronics Engineering, Abubakar Tafawa Balewa University, P.M.B. 0248, Bauchi Nigeria

²Department of Mechatronics and Systems Engineering, Abubakar Tafawa Balewa University, P.M.B. 0248, Bauchi Nigeria

Corresponding email: mmsardauna@gmail.com

ABSTRACT

The need for stable, secure and cost-effective power requires a balance between supply and demand of electricity. Hence, the need for an effective and accurate forecasting technique to maintain continuous supply of stable power to consumers. The limitations of some common forecasting techniques such as Artificial Neural Network (ANN) and Support Vector Machine (SVM) are falling into local minimum, overfitting and computational complexity. This paper presents Short Term Load Forecasting (STLF) technique using Group Method of Data Handling (GMDH) optimized by Artificial Bee Colony (ABC) algorithm. Historical load data and weather data were used as input variables to the model using MATLAB software. Performance of the proposed technique was evaluated and compared with GMDH load forecasting technique. The Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) were employed to investigate the performance of the presented models. Four years load data and weather data collected from Transmission Company of Nigeria (TCN) and Nigerian Metrological Agency (NIMET) were used for training, testing and validation of the models. Results showed that GMDH and GMDH optimized by ABC recorded MAPE of 4.6633 and 4.2083 respectively while the RMSE of 1.6861 and 1.217 respectively. The results obtained showed that the GMDH optimized by ABC model has the lowest MAPE and RMSE when compared with GMDH. The evaluation results indicated that the GMDH-ABC model provides more accurate load forecasts compared to former model. Hence, this research showed that GMDH-ABC approach is more accurate and effective than GMDH STLF technique. It is recommended that further studies on this work should focus on employing other optimization algorithms to optimize GMDH model. Also, it is suggested that long term load forecasting model using GMDH-ABC be developed by incorporating additional parameters such as season of the year, GDP, population growth, festive period and holidays.

Keywords: Electricity, Artificial Neural Network, Support Vector Machine, Short Term Load Forecasting

1 INTRODUCTION

Continuous uninterrupted supply of electricity to the load side is required by modern power systems. This calls for a sound understanding of how to forecast load demand in the present and the future with the least degree of error. To do this, researchers and experts have been working to create the most effective and cutting-edge method currently available for forecasting future electricity demand. For the planning and operation of the power system network, load forecasting is essential [1]. Dispatching, unit commitment, fuel allocation, and off-line network analysis are just a few of the processes and decisions that are controlled by load forecasting. This provides adequate time for the power utility provider to reduce the discrepancy between generation capacity and load demand as well as an understanding of the consumers' future demand. Given the high costs associated with power generating, demand prediction reduces the cost of energy production and aids in the creation of an

organized power system utility. Thus, load forecasting is very vital to generation, transmission and distribution utilities of the power sector [2]. A utility company's operations and planning depend on precise models for projecting electric power load. In forecasting load, accuracy is essential. Yet, load forecasting is a difficult endeavor that necessitates the examination of numerous direct and indirect influences on the procedure. Although there are many advantages to employing load forecasting systems, there are some difficulties that limit their accuracy. The forecasting process is complicated and occasionally chaotic in nature. The statistics are further influenced by weather-related factors, which makes forecasting more difficult. As a result, the load at any given hour depends on more than just the preceding hour; it also depends on the weather, demographic information, the number of appliances in the forecasting region, type of customers, customer number, and econometric data, etc. It is vital to keep the load forecasting error as low as feasible even in these diverse situations. Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE) are the common measures of forecast accuracy that are typically expressed in percentages. For an algorithm to be viable for load forecasting, its MAPE and RSME values must remain within a range of a few percentage points [1]. The remainder of this paper is organized as follows: Section II presents the review of related works, while Section III contains the methodology for the development of the proposed forecasting model. In section IV, experimental results are presented and discussed. Finally, Section V carries the conclusion.

2 LITERATURE REVIEW

Considering the importance of load forecasting in managing power system, many researchers developed load forecasting techniques using various algorithms ranging from statistical models, Artificial Intelligences (AI) based models as well as hybrid models. Auto Regressive Integrated Moving Average (ARIMA) based STLF technique was developed by [2]. In their work, historical hourly load and peak load of West Bengal were utilized to train and test the model. In the same vein, [3] presented a comparative study of time series analysis for predicting Nigeria's load demand. They utilized Harvey, Auto Regressive Moving Average (ARMA), and Exponential Smoothing time series models to analyze real-world data from the National Control Center Osogbo, and R-Statistics was used to simulate the models. They revealed that Harvey model was determined to have the lowest RSME and the best prediction accuracy. [4], worked on STLF using Kalman Filter technique. They used real life load data obtained from state load dispatch Centre India to develop the prediction model and provide one day ahead forecast. Likewise, STLF using Multiple Linear Regression (MLR) technique was proposed by [5] in which they utilized historical load data of Donwod region of New York city to develop the model using MLR and least square estimation. Also, [6] proposed regression analysis based MTLF for Nigeria's electricity distribution system. To anticipate the energy demand based on historical load consumption in the Abeokuta area, the work used three different regression techniques: linear, compound growth, and quadratic. In comparison to compound growth and quadratic, the linear regression model performed fairly better. [7] proposed STLF utilizing statistical approaches using Andra-Pradash State Electricity load data to predict load using Auto Moving Average (AMA), ARMA, and Kalman Filter techniques. Her finding revealed the supremacy of Kalman Filter over MA and ARMA. Though conventional statistical models such as MA, ARIMA, MLR and Kalman Filter have been found to work well for linear prediction problems, however, they perform poorly for complex non-linear load time series [8]. In another development, [9] proposed SVM-based electrical load forecasting approach using actual load data from a sub-station of Punjab State Electricity. For the simulation study with the approach, they separated the data into validation, training, and testing. They used the actual load data from the chosen area to evaluate the simulation results and discovered that they were within acceptable bounds. In a bid to employ ANN in load forecasting, [10] proposed STLF using ANN Approach with periodic and non-periodic factors. They developed the model and used previous loads data collected from Tai'an City Substation in Shandong Province of China to predict the power demand. Their findings revealed that extraneous variables affect load pattern as it was confirmed that the load decreased dramatically for a considerable amount of time during the Chinese lunar spring festival. Similarly, [11] proposed STLF using ANN. Beyond the simple minimum-maximum scaling procedures, the paper introduced two novel data pre-processing strategies applied to STLF utilizing ANN. In comparison to more conventional approaches, the two main pre-processing strategies focus on the importance of specific neural network input variables in connection to output variables. The technique yields better results in terms of Mean Squared Error, Mean Absolute Error,

and MAPE compared to previous research using data from the Greek interconnected system. Additionally, STLF using GMDH, Long Short-Term Memory (LSTM) and Adaptive Neuro-Fuzzy Inference (ANFIS) were developed by [12]. They employed LSTM, GMDH and ANFIS to predict load on smart grid consisting of wind turbine, photovoltaic and battery storage energy systems. They found LSTM to be a bit more accurate than the other two techniques. However, LSTM requires much better hardware requirement than the GMDH and ANFIS. Furthermore, SVM optimized by Particle Swarm Optimization (PSO) based STLF method was developed by [13]. In the work, PSO was applied to optimize the SVM model's parameters and thereafter, a load forecasting model was developed. It was determined that the SVM-PSO-based short-term load prediction model met the forecasting and accuracy requirements. However, this method is liable to early convergence. [14] presented STLF utilizing ANN and PSO algorithms. In order to predict load using the PSO method, some parameters of neural network such as learning rate and number of hidden layers, were determined in this work. The neural network was then utilized to anticipate the load using these adjusted parameters. MATLAB software was used to test the technique utilizing data from an Iranian grid. The outcomes of the simulation show how accurate the proposed model can forecast the load. The drawback of this method is the ANN's performance degradation brought on by over-fitting and memorization. Also, ANN optimized by ABC based STLF approach was developed by [15] with a view to enhancing prediction performance. In this work, ABC algorithm was employed to optimize the learning process of the ANN. The proposed hybrid prediction approach was validated by predicting the power demand in the Bushehr Province of Iran. They came to the conclusion that the outcome demonstrated how the technique enhanced forecasting accuracy. The drawback of ANN-ABC is that, it can become stuck in local optimum and does not ensure that the global optimal solution will be found. [16] proposed power demand forecasting based on weighted gray relational degree and enhanced ABC-SVM. The weight of the load-related physical data was obtained using the entropy weight approach, and previous as well as forecast load data chosen based on the weighted relational degree entered to the SVM to create a forecasting model. The enhanced ABC was then used to optimize the SVM parameters. Their finding revealed that the technique might significantly increase the model's accuracy. Another important work on STLF was proposed by [17] using hybrid GMDH-LSTM model optimized by Improved Carnivorous Plant Algorithm (ICPA). In this study, they decomposed the real load data into multiple Intrinsic Mode Functions (IMF) components with different characteristics using Improved Complete Ensemble Empirical Mode Decomposition Adaptive Noise (ICEEMDAN), while Information Entropy (IE) was added to prevent increased forecasting error and prolonged forecasting time due to too many IMF components. Finally, they utilized GMDH and LSTM to predict the restructured components separately. To get more accurate forecasts, the predictions' outcomes were then weighted and added. However, this technique is complex considering many processes involved before obtaining the final results.

3 METHODOLOGY

This section presents the methodology for the development of the proposed model. This includes data collection, data cleaning/processing, data division, model training, testing and validation as well as model performance evaluation. Figure 1 shows the GMDH-ABC Implementation Flowchart.

3.1 Data collection

Four years data were collected for each of the input parameters. Electric Load Data for the Bauchi Region of Jos Electricity Distribution Company (JEDC) were TCN and the weather were obtained from NIMET.

3.2 Data processing

Data processing entails identifying/detecting irrelevant, inaccurate and incomplete parts of the data with a view to replacing, modifying and/or deleting them where necessary. In this stage, irrelevant, Incomplete or inaccurate data parts were detected and necessary replacements, modifications/deletions were made.

3.3 Data division

The data collected were separated into three components. The components are: train, test and validate data sets. The model training was conducted using train data set while test data set and validate

data set were utilized for testing and evaluating the validity of the estimated model respectively. The train data set account for 60% of the data, 20% for testing and 20% for validation.

3.4 Loading the initial GMDH Structure and training the Model

Initial GMDH model structure was generated and optimized by ABC algorithm. The model was trained using the load and weather data.

3.5 Validating the trained Model

At this stage, the trained GMDH-ABC model was validated to see how accurately it can forecast the output values.

3.6 Performance Evaluation

Finally, MAPE and RMSE were used to assess the performance of the proposed approach and compare it with the performance of GMDH model.

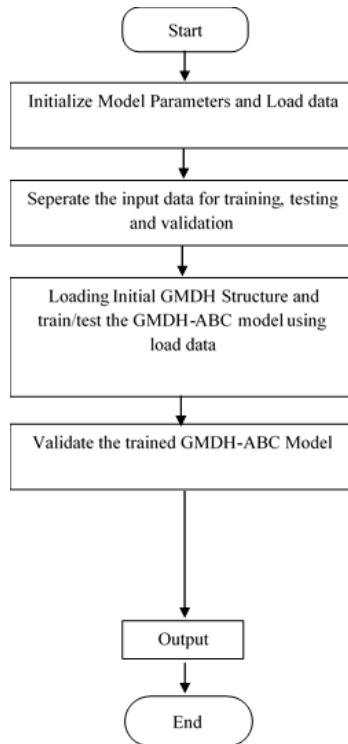


Figure 1. GMDH-ABC Flowchart

4 RESULTS AND DISCUSSION

This section presents the results obtained during the implementation of the proposed model as well as discussions of the results. The STLF is modelled using GMDH-ABC and then compared with results obtained using GMDH model. The GMDH-ABC model was employed to predict day ahead load demand. Results obtained are shown in tables and figures.

4.1 Performance GMDH STLF model

Figure 2 shows the plot of actual load versus the model's predicted load for a day. The plots indicate less significant variations except at some few (5) points. The performance of the developed GMDH model in terms of MAPE and RMSE is shown in Table 1 below. The MAPE obtained for training, testing and validation data is 4.6656, 4.6777 and 4.6814 respectively whereas the RMSE obtained for training, testing and validation is 1.6861, 1.6885 and 1.7097 respectively. From the performance results, it can be seen that there are no significant differences of RMSE and MAPE for training, testing and validation respectively. This indicates that there was no over fitting. Also, the model was found to be more accurate than Compound Growth (CG) and KF models with RMSE 1.6984, and 11.3400 respectively as found in the work of [6, 7].

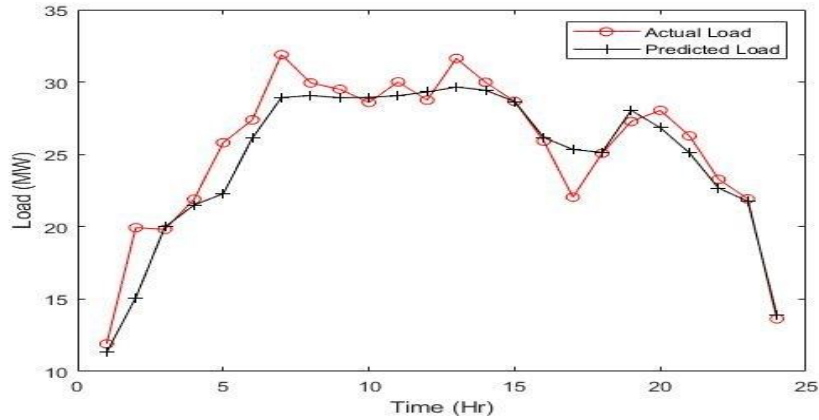


Figure 2. Plot of Actual Load versus GMDH Model Predicted Load

Table 1. Performance of Developed GMDH Model

Data	RMSE	MAPE
Training	1.6860	4.6656
Testing	1.6885	4.6777
Validation	1.7097	4.6814

4.2 Performance GMDH-ABC STLF model

Figure 3 shows the plot of actual load versus the predicted load for a day. The plots indicate less significant variations except at some few (3) points which is better than GMDH. The performance of the developed GMDH-ABC model in terms of MAPE and RMSE is shown in Table 2. The MAPE obtained for training, testing and validation data is 4.2042, 4.2534 and 4.2616 respectively whereas the RMSE obtained for training, testing and checking data is 1.2170, 1.2220 and 1.2229 respectively. Like in the case of GMDH, there are no significant differences of RMSE and MAPE for training, testing and checking data respectively thereby indicating absence of over fitting. Similarly, the GMDH-ABC model was found to be more accurate than CG, KF and SVM models with RMSE 1.6984, 11.3400 and 2.100 respectively as found in the work of [6, 7, 9].

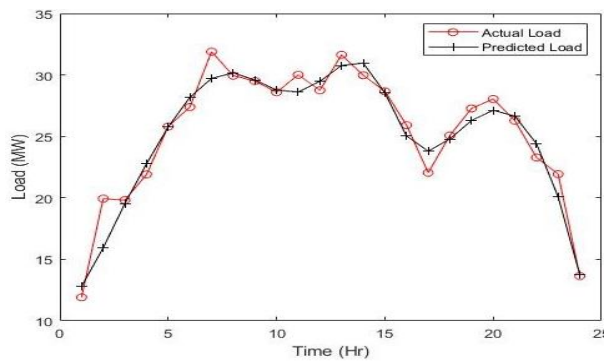


Figure 3. Plot of Actual Load versus GMDH-ABC Model Predicted Load

Table 2. Performance of Developed GMDH-ABC Model

Data	RMSE	MAPE
Training	1.2170	4.2042
Testing	1.2220	4.2534
Validation	1.2229	4.2616

4.3 Comparison of Models Performance

Figure 4 shows the plots of actual versus predicted load for the two individual models. From the plots, it can be seen that GMDH-ABC's predicted load variation from the actual load is less compared to GMDH. The performance result for the two developed models is presented in table 3. The

RMSE obtained for GMDH and GMDH-ABC is 1.6861 and 1.217 respectively while the MAPE obtained is 4.6633 and 4.2083 respectively. It can be seen that the difference between the MAPE of the proposed model with GMDH revealed about 10% decrease in forecasting error. Also, the model performed better than SVM with 2.100 RMSE as presented in the work of [9]. Thus, the GMDH-IABC model outperformed both GMDH and SVM model in terms of forecasting accuracy. Therefore, GMDH-ABC Model was found to be simple, accurate and effective load forecasting technique.

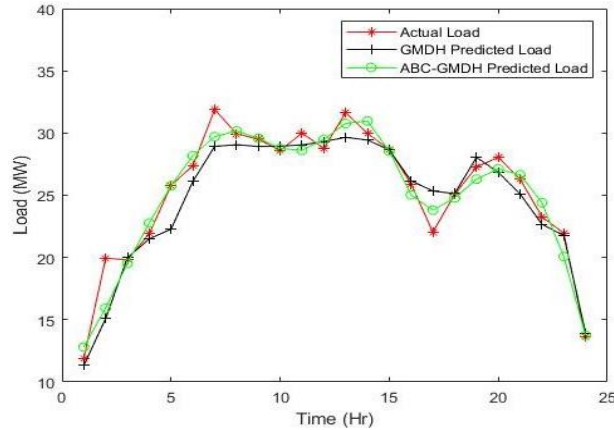


Figure 4. Plots of Actual versus Predicted Load for GMDH and GMDH-ABC Models

Table 3. Comparison of Models' Performance

MODEL	RMSE
GMDH	1.6861
GMDH-ABC	1.217
CG	1.6984
KF	11.34
SVM	2.1

4.4 GMDH-ABC Model Day Ahead Forecast

The GMDH-ABC STLF Model was successfully developed and the technique was employed to predict day ahead forecast. Figure 5 shows the plot of the day-ahead forecast. From the figure, it can be seen that peak load demands are recorded around 0600-0800hours, 1200-1400hours and 1800-2100hours which correspond to the obvious morning, afternoon and evening peak respectively. Moreover, the model's forecasting accuracy was found to be superior than that of ANN as presented in the work of [10] with 4.2600 MAPE as against GMDH-ABC's 4.2083. This confirmed the effectiveness of the developed forecasting technique.

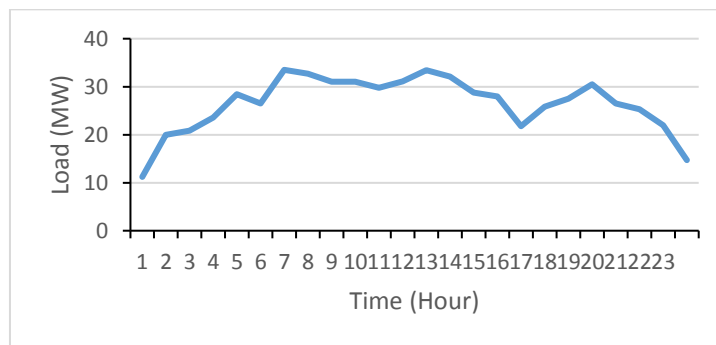


Figure 5. Plots of Day Ahead Forecast by GMDH-ABC

5 CONCLUSIONS

This paper presented the development of GMDH-ABC based STLF technique. The model was developed in MATLAB environment using load data and weather (temperature and humidity) data of Bauchi Region of JEDC. The model was successfully developed to predict short term load demand. Its

forecasting performance were evaluated by computing the MAPE and RMSE between the actual and predicted values. The results indicated that the GMDH-ABC model is an effective tool for load forecast used by electric power utility companies. The developed model was employed to predict day ahead load demand for Bauchi Region of JEDC. The forecast recorded good performance with very good prediction accuracy. The simulation results also show that a simple GMDH-ABC based prediction model outperformed GMDH as it had a MAPE of 4.2083 as compared with GMDH's 4.6633 indicating about 10% decrease in forecasting error. It is therefore concluded that this research which attempt to provide effective and accurate short-term load forecasting in the Nigerian power system using GMDH-ABC is highly successful.

REFERENCES

1. Abdullah, M. and Naeem, M. (2020). A Comprehensive Review of the Load Forecasting Techniques using Single and Hybrid Predictive Models. *IEEE Journals 10.1109. IEEE Access*, vol. 8, pp. 134911-134939.
2. Goswami, K., Ayandeep, G. and Arindam, K.S. (2018). Day-ahead Forecasting and Peak Load Management using Multivariate Auto Regression Technique. *2018 IEEE Applied Signal Processing Conference*, pp. 279-282.
3. Okakwu, I.K., Oluwasogo, E.S., Ibhaze, A.E. and Imoize, A.I. (2019). Comparative Study of Time Series Analysis for Forecasting Energy Demand in Nigeria. *Nigerian Journal of Technology*, vol. 38, no. 2, pp. 465-469.
4. Rajasekhar, S. and Chaitanya, M.N. (2019). Load forecasting using Kalman filtering technique, *International Journal for Modern Trends in Science and Technology*, vol. 5, issue 10, pp. 7-10.
5. Dhaval, B. and Deshpande, A. (2020). Short-term load forecasting using multiple linear regression. *International Journal of Electrical and Computer Engineering*, vol. 10, no. 4, pp. 3911-3917.
6. Olabode, O., Olatide, A., Titus, A. and Ignatius O. (2020). Medium Term Load Forecasting in a Nigerian Electricity Distribution Region using Regression Analysis Technique. *International Conference in Mathematics, Computer Engineering and Computer Science*, pp.1-5.
7. Haleema, S.D. (2010). Short Term Load Forecasting Using Statistical Methods: A Case Study of Load Data. *International Journal of Engineering and Technology*, pp. 516-520.
8. Akhtar, S., Shahzad, S., Zaheer, A., Ullah, H.S., Kilic, H., Gono, R., Jasiński, M. and Leonowicz, Z. (2023). Short-Term Load Forecasting Models: A Review of Challenges, Progress, and the Road Ahead. *Energies*, vol. 16, no. 10, pp. 1-29
9. Parag, N., Vinod, K., Manish, K. and Jyoti, G. (2020). Electrical Load Forecasting using Support Vector Machine Algorithm. *International Journal of Recent Technology and Engineering*, vol. 8, issue 6, pp. 4811-4816.
10. Jiuyun, S., Huanhe, D., Ya, G., Yong, F. and Yuan, K. (2021). Short-Term Load Forecasting Using an Artificial Neural Network Approach with Periodic and Non-periodic Factors: A Case Study of Tai'an, Shandong Province, China. *Journal of Computational Intelligence and Neuroscience*, pp. 1-8.
11. Athanasios, I.A., Dimitrios, B., Aspasia, D., Vasileios, M.L. and Lefteri, H. (2021). Enhanced Short Term Load Forecasting Using Artificial Neural Networks. *Energies*, vol. 14, no. 7788, pp. 1-14.
12. Salehimehr, S., Taheri, B. and Sedighzadeh, M. (2022). Short-term load forecasting in smart grids using artificial intelligence methods: A survey. *The Journal of Engineering, J. Eng.*, pp. 1133-1142.
13. Xiaoqun, L., Kang, X.M. and Cao, N. (2019). Short-term Load Forecasting Model and a Load Early Warning Model for Charging Station Based on Support Vector Machine Optimized by

- Particle Swarm Optimization. *2019 International Conference on Intelligent Transportation, Big Data & Smart City*, pp. 305-308.
14. Zahra, S.C. and Hossein, A. (2021). Short Term Load Forecasting Using Artificial Neural Network and Particle Swarm Optimization Algorithm. *Journal of Mathematical Problems in Engineering*, pp. 1-10.
 15. Baesmat, K.H., Iman, M. and Haidar, S. (2021). Improving the Performance of Short-Term Load Forecast Using a Hybrid Artificial Neural Network and Artificial Bee Colony Algorithm. *IEEE Canadian Journal of Electrical and Computer Engineering*, vol. 44, issue 3, pp. 275-282.
 16. Ruxe, L., Shumin, L. and Miaona, Y. (2021). Load Forecasting Based on Weighted Gray Relational Degree and Improved Support Vector Machine-Artificial Bee Colony. *Journal of Electrical Engineering and Technology*, vol. 16, no. 4, pp. 2191-2200.
 17. Gonggui, C., Jie, B., Tewe, C., Wei, W., Zongfu, W., Hongyu, L. and Mi, Z. (2022). Short Term Long Forecasting Using Hybrid GMDH-LSTM Model optimized by ICPA. *Engineering Letters*, vol. 30, no. 4, pp. 1-18.