Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
Vol. 28 (2025), No. 1, pp. 95–105
DOI : 10.47974/JSMS-1316

Interdisciplinary role of statistics in energy consumption prediction within smart grid big data analytics

Kamal Sharma *
Muma College of Business
University of South Florida
8350 N. Tamiami Trail Sarasota
Florida
FL 34243
U.S.A.

and

Department of Mechanical Engineering
GLA University
Mathura 281406
Uttar Pradesh
India

Bhuvan Unhelkar †
Muma College of Business
University of South Florida
8350 N. Tamiami Trail Sarasota
Florida
FL 34243
U.S.A.

S. Siva Shankar §
Department of Computer Science and Engineering
KG Reddy College of Engineering and Technology
Moinabad 500075
Telangana
India

* *E-mail:* **kamal.sharma@gla.ac.in (Corresponding Author)**
† *E-mail:* **bunhelkar@usf.edu**
§ *E-mail:* **drsivashankars@gmail.com**

Tulika Chakrabarti [‡]
Department of Chemistry
Sir Padampat Singhania University
Udaipur 313601
Rajasthan
India

Prasun Chakrabarti [@]
Department of Computer Science and Engineering
Sir Padampat Singhania University
Udaipur 313601
Rajasthan
India

B. Sivaneasan [#]
Specialist Adult Educator Engineering
Electrical Power Engineering Programme
Singapore Institute of Technology
1 Punggol Coast Road
Singapore 828608

Martin Margala [$]
School of Computing and Informatics
University of Louisiana
Lafayette
LA 70503
U.S.A.

Abstract

The study of energy consumption in buildings is crucial due to environmental and financial impacts. Smart metering is suggested to improve energy efficiency, but predicting electrical consumption accurately remains challenging, considering various factors. To tackle this, we propose a Multi-objective Golden Eagle Optimization (MGE) based Hybrid Deep Dynamic Convolutional Fuzzy Network (HDDCFN) method. This approach optimizes

[‡] *E-mail:* **tulika.chakrabarti@spsu.ac.in**
[@] *E-mail:* **drprasun.cse@gmail.com**
[#] *E-mail:* **sivaneasan@singaporetech.edu.sg**
[$] *E-mail:* **Martin.margala@louisiana.edu**

algorithm parameters and prediction models using multi-objective optimization, considering historical usage and weather conditions. The research is conducted using Matlab.

Subject Classification: *68T07.*

Keywords: *Energy time series, Energy prediction, Interdisciplinary, Smart meters, Smart grid (SG), Energy time series, Optimization and artificial intelligence.*

I. Introduction

Over the last decade, urbanization has surged, drawing people to Smart Cities for their advanced amenities. Residential energy consumption has risen due to urbanization, technology, and increased comfort needs. SG play a vital part in ensuring reliable and economical power for Smart Cities, transforming physical infrastructure into digital systems [1]. These grids improve efficiency and reliability by minimizing outages and introducing self-healing capabilities. Smart Grids utilize smart meters and advanced metering infrastructure (AMI) to measure energy consumption, generating vast amounts of high-velocity data. Big data technologies have the potential to revolutionize utility providers' capabilities, enabling a better understanding of customer energy consumption patterns, demand forecasting, outage prevention, and energy optimization [2][3]. Researchers are developing energy management systems using big data approaches to predict energy demand and supply consistently. Creating models to analyze smart meter data and predict energy consumption is crucial. This study proposes energy-efficient big data analytics for residential smart grid metering systems. It develops a time-series model for historical energy consumption and meteorological conditions, storing results in a database. It introduces the Multi-objective Golden Eagle Optimization (MGE)-based Hybrid Deep Dynamic Convolutional Fuzzy Network (HDDCFN) for energy consumption prediction, optimizing algorithm parameters and prediction models. The study validates the system's performance measures and includes sections on existing research, proposed techniques, results, and conclusions with suggestions for future work.

II. Related work

Ensuring responsible and efficient energy consumption practices is increasingly vital in today's power system. The author of [4] proposed an intelligent data mining methodology for analyzing energy time-series data, achieving short-term hourly prediction accuracy of 81.89% and long-term prediction accuracy ranging from 72.81% to 81.89%. In [5] the author introduced EvoHyS, integrating statistical and machine learning approaches, enhancing accuracy with Mean Squared Error (MSE) evaluation. Also [6] developed a smart energy consumption analysis system, focusing on managing energy usage without extra costs and utilizing IoT and Big Data in various sectors. The [7-9] presented a temporal and spatial-based ensemble forecasting model using clustering algorithms and LSTM and GRU units for high forecasting accuracy. In [10] the author proposed a spatio-temporal visual analytic technique for urban energy consumption patterns, while Fatema, Israt, and colleagues suggested a single complete LSTM for energy forecasting. Also [11] developed a hybrid model combining CNN and BLSTM for precise energy consumption forecasting. Further [12-16] conducted a comparative examination, finding that ConvLSTM combined with BiLSTM performed the best [17][18]. Traditional ML [19] algorithms are prone to local minima traps, and neural networks often face complexity and overfitting issues.

III. Proposed methodology

The proposed model, depicted in Figure 1, comprises data pre-processing, training/testing, prediction, and optimization phases. Figure 2 shows the flowchart of the proposed methodology. It starts by

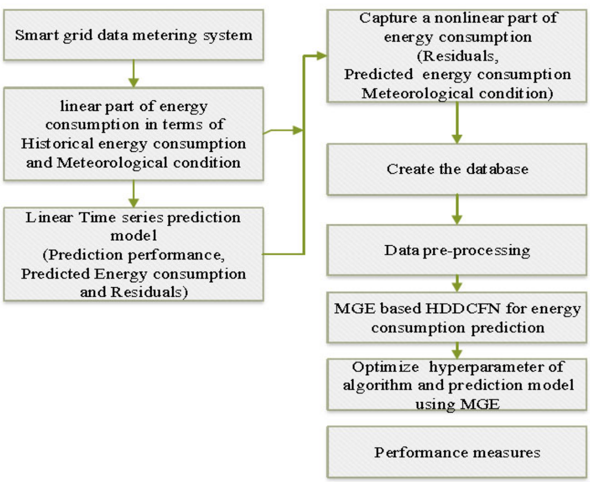

Figure 1

Proposed methodology

constructing a linear component for energy consumption prediction based on historical and meteorological factors, followed by integrating nonlinear aspects for database creation to ensure accuracy. A novel prediction model, MGE-based HDDCFN, is developed and evaluated using test and training sets. Training employs a hybrid deep learning algorithm, with hyperparameters optimized using the MGE method to enhance performance. Performance measures validate the effectiveness of the research.

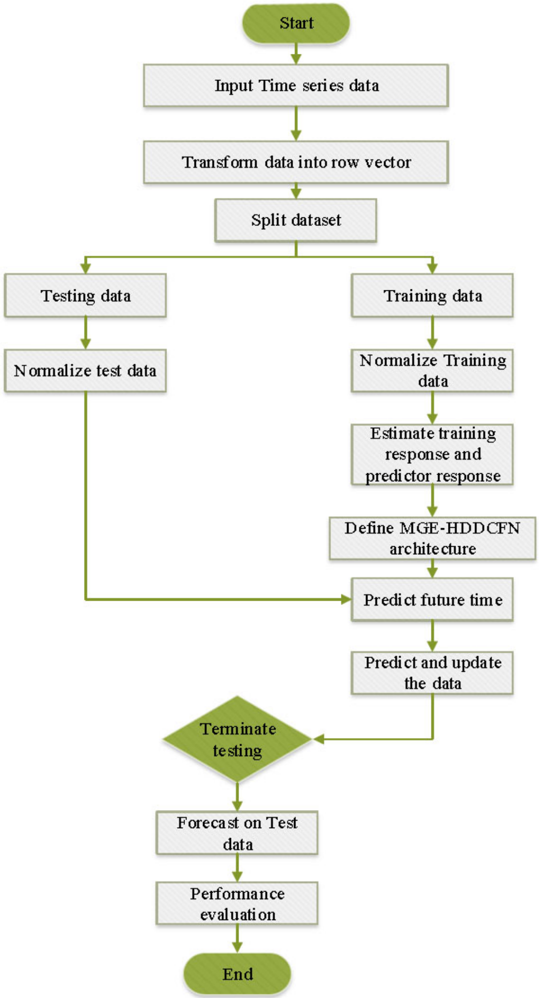

Figure 2

Flowchart of the proposed methodology

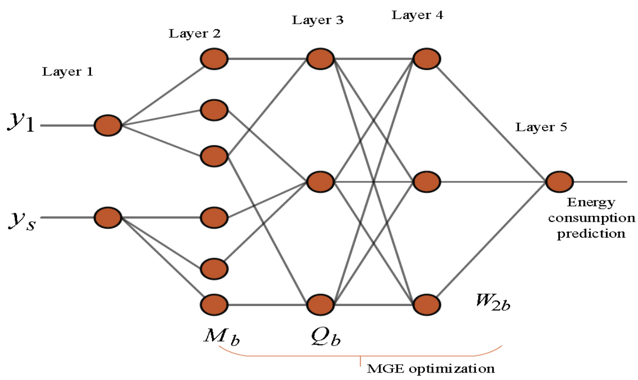

Figure 3

Structure of the MGE-HDDCNN method

- Proposed algorithm

  Figure 3 illustrates the new MGE-HDDCNN structure, utilizing Deep CNN to run a fuzzy system based on the TSK model.

- GE optimization

  The GE algorithm combines hunting and encircling behaviors to select parameters. Initialize the HDDCFN algorithm's settings and prime number parameters randomly. The HDDCFN algorithm prioritizes selecting the best fitness.

IV. Result and discussion

The experimental design of the suggested research project is described in this section. MATLAB is employed as a general-purpose language in this research study to carry out a number of tests.

- Dataset Description

  Data from ENTSOE and Red Electrica Espana were used for consumption and generation stats and settlement pricing. Open Weather API provided weather data for major Spanish cities. Training, validation, and testing sets were separated from the dataset, with performance analysis conducted on weekdays and weekends.

- Performance analysis

  The total hourly load curve is illustrated in Figure 4. The observation shows the overall load curve of the energy efficiency SG system.

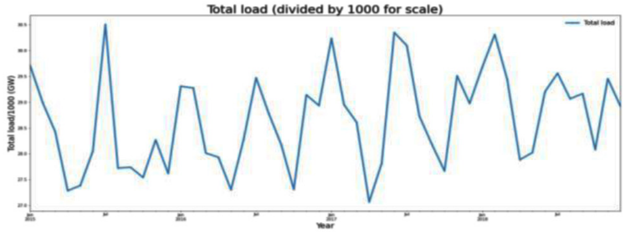

Figure 4

Total hourly load curve

Figure 5 displays Hourly load and apparent temperature, illustrating consumption patterns driven mainly by households and services.

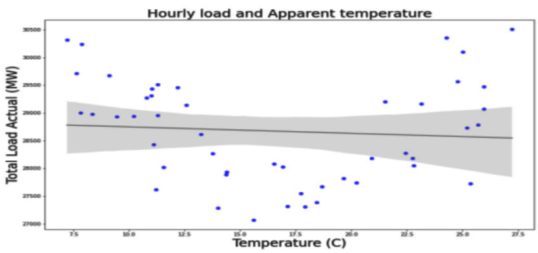

Figure 5

Hourly load and apparent temperature

The use of electricity for heating is restricted to a few service segments and high-income homes.

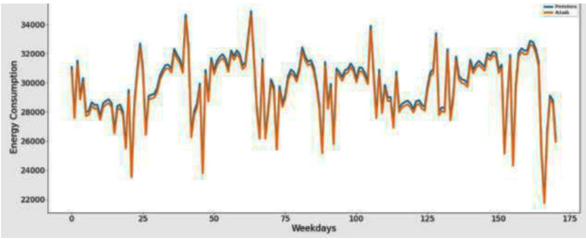

Figure 6

Load forecast for week day

Figures 6 and 7 depict hourly power demand trends for weekdays and weekends from 2015 to 2019.

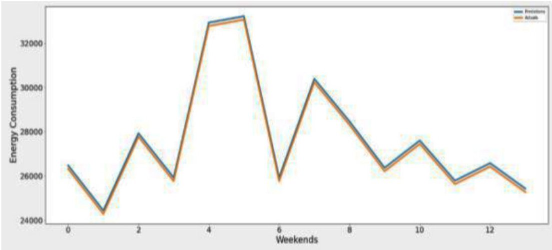

Figure 7

Load forecast for weekend day

The observation underscores the substantial disparity in power usage between weekdays and weekends, with significantly lower demand during weekends due to reduced activities. This forecast promises notable utility cost savings and enhanced facility utilization. Simulation results, along with metrics in Table 1, elucidate the suggested model's performance. It was trained and evaluated on unbiased data, using MAE and RMSE values for assessment.

Table 1

Elucidate the suggested model's performance Top of Form

Metrics	Weekday	Weekend
MAE	0.15	0.24
MSE	0.2	0.27
RMSE	0.43	0.76
Correlation (R)	0.842	0.861
Coefficient of Determination (R^2)	0.932	0.926

R^2 is a correlation coefficient that measures how well the prediction models performed. The correlation between the actual and expected values is greater the higher the R_2 value. In both situations, weekdays and weekends, the R^2 value for power demand is over 0.9, showing a significant connection between actual and expected values.

- Comparative analysis

  In Table 2, the effectiveness of the suggested approach and the traditional approach are compared.

Table 2

Comparison of performance measures using various approaches

Parameter	Algorithm	MAE	MSE	RMSE	R	R^2
Weekday	EvoHyS[17]	4.02	1.20	12.05	0.76	0.9
	LSTM-GRU [18]	5.44	0.310	2.669	-	-
	LSTM [19]	50.96	-	67.35	-	0.62
	Proposed	0.15	0.2	0.43	0.84	0.93
Weekend	EvoHyS[17]	3.05	1.05	23.05	0.682	0.90
	LSTM-GRU [18]	3.769	0.340	2.695	-	-
	LSTM [19]	31.5	-	31.5	-	0.59
	Proposed	0.24	0.27	0.76	0.861	0.92

The suggested model achieves the highest R2 values of 0.932 for weekdays and 0.926 for weekends.

V. Conclusion

This research presents a new deep learning framework for energy-saving decisions, vital due to increasing electricity use. Using advanced forecasting techniques, it optimizes energy consumption with high precision. Results, analyzed using MATLAB, demonstrate superior performance over existing methods. Future work includes refining the model for different climates and expanding to city-level energy estimates for improved accuracy.

References

[1] V. Vanitha and E. Vallimurugan, "A hybrid approach for optimal energy management system of Internet of Things enabled residential buildings in smart grid," *International Journal of Energy Research*, vol. 46, no. 9, pp. 12530–12548 (2022).

[2] S. Venkatachary, J. Prasad, R. Samikannu, A. Alagappan, and L. J. B. Andrews, "Cybersecurity infrastructure challenges in IoT-based virtual power plants," *Journal of Statistics and Management Systems*, vol. 23, no. 2, pp. 263–276 (2020).

[3] T. M. Skjølsvold, I. M. Henriksen, and M. Ryghaug, "Beyond the car: How electric vehicles may enable new forms of material politics at

the intersection of the smart grid and smart city," *Urban Geography,* pp. 1–18 (2022).

[4] H. E. Khodke, M. Bhalerao, S. N. Gunjal, G. S. Nirmal, S. Gore, and B. J. Dange, "An intelligent approach to empowering the research of biomedical machine learning in medical data analysis using PALM," *International Journal of Intelligent Systems and Applications in Engineering,* vol. 11, pp. 429–436 (2023).

[5] D. K. Behera, M. Das, S. Swetanisha, and J. Nayak, "XGBoost regression model-based electricity tariff plan recommendation in smart grid environment," *International Journal of Innovative Computing and Applications,* vol. 13, no. 2, pp. 79–87 (2022).

[6] V. F. Pires, I. Colak, and F. Kurokawa, "Smart grid as a key tool for the future of electrical distribution networks," *Energies*, vol. 15, no. 9, p. 3281 (2022).

[7] R. Josphineleela, S. Periasamy, N. Krishnaveni, D. S. Prasad, B. V. Rao, M. J. Garde, and S. Gore, "Exploration beyond boundaries: AI-based advancements in rover robotics for lunar missions like Chandrayaan," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 11, pp. 640–648 (2023).

[8] S. Gore, "Brain tumour segmentation and analysis using BraTS dataset with the help of improvised 2D and 3D UNet model," *Research-Square* (2023).

[9] K. Tharani, N. Kumar, V. Srivastava, S. Mishra, and M. P. Jayachandran, "Machine learning models for renewable energy forecasting," *Journal of Statistics and Management Systems,* vol. 23, no. 1, pp. 171–180 (2020).

[10] N. K. Singh and V. Mahajan, "End-user privacy protection scheme from cyber intrusion in smart grid advanced metering infrastructure," *International Journal of Critical Infrastructure Protection,* vol. 34, p. 100410 (2021).

[11] J. Chan, R. Ip, K. W. Cheng, and K. S. Chan, "Advanced metering infrastructure deployment and challenges," in 2019 IEEE PES GTD Grand International Conference and Exposition Asia (GTD Asia), pp. 435–439 (2019).

[12] J. S. Guerrero-Prado, W. Alfonso-Morales, and E. F. Caicedo-Bravo, "A data analytics/big data framework for advanced metering infrastructure data," *Sensors,* vol. 21, no. 16, p. 5650 (2021).

[13] P. Dadheech, D. Goyal, S. Srivastava, and C. M. Choudhary, "An efficient approach for big data processing using spatial Boolean queries," *Journal of Statistics and Management Systems*, vol. 21, no. 4, pp. 583–591 (2018).

[14] M. M. Rahman, M. B. Pasupuleti, and H. P. Adusumalli, "Advanced metering infrastructure data: Overviews for the big data framework," *ABC Research Alert*, vol. 7, no. 3, pp. 159–168 (2019).

[15] T. Wilcox, N. Jin, P. Flach, and J. Thumim, "A big data platform for smart meter data analytics," *Computers in Industry*, vol. 105, pp. 250–259 (2019).

[16] G. F. Huseien and K. W. Shah, "A review on 5G technology for smart energy management and smart buildings in Singapore," *Energy and AI* (2022).

[17] P. D. Diamantoulakis, V. M. Kapinas, and G. K. Karagiannidis, "Big data analytics for dynamic energy management in smart grids," *Big Data Research*, vol. 2, no. 3, pp. 94–101 (2015).

[18] M. Ghorbanian, S. H. Dolatabadi, and P. Siano, "Big data issues in smart grids: A survey," *IEEE Systems Journal*, vol. 13, no. 4, pp. 4158–4168 (2019).

[19] T. Zufferey, A. Ulbig, S. Koch, and G. Hug, "Forecasting of smart meter time series based on neural networks," in Data Analytics for Renewable Energy Integration: 4th ECML PKDD Workshop, DARE 2016, Riva del Garda, Italy, September 23, 2016, Revised Selected Papers, vol. 4, Springer International Publishing, pp. 10–21 (2017).

Received February, 2024