BEMS Reports, 2026; 12(1):9-14.
https://www.bemsreports.org

Brief Report

Energy Forecasting and Time Series Analysis Using Machine Learning

Utpal Mishra*

Visiting Scientist, Veterinary Science Center, School of Veterinary Medicine, University College Dublin, Belfield, Dublin 4, IRELAND.

ABSTRACT

We have moved from lacking a sufficient supply of electricity/power to producing it in abundance, so it is paramount to decipher how to bring it to optimal usage. This research lays a hand on forecasting energy, bringing in the consumption of electricity and city across the households, enabling stakeholders to accurately predict future energy consumption and generation and meet the demand to enhance sustainable practices. This research examines various Machine Learning algorithms and the very essence of Time Series Forecasting. Forecasting can be done in different span/time intervals as required but eventually depends on factors such as managing the load, trading electricity, and optimizing energy storage, which is crucial for strategic planning and helps to identify trends influenced by economic and social factors. Considering how we are moving forward, having Power System Forecasting is essential to make the optimal use of our resources, and with the generated data, using the approach of Machine Learning and Forecasting to understand the pattern can make a difference.

Keywords: Energy, Forecasting, Optimization, Regression, Strategic Planning, Sustainability, Time Series.

Correspondence:
Mr. Utpal Mishra
Visiting Scientist, Veterinary Science Center, School of Veterinary Medicine, University College Dublin, Belfield, Dublin 4, IRELAND.
Email: utpalmishra21official@gmail.com

Received: 10-05-2025 ;
Revised: 06-09-2025;
Accepted: 08-10-2025.

INTRODUCTION

Electricity is the paramount form of energy, essential for driving progress and innovation. Traditionally, fossil fuels have been the sole contributors to energy production. However, as we have witnessed their consumption and depletion rates, the need for alternative energy sources has become increasingly critical. Advancements in technology have illuminated various renewable energy sources, including solar, wind, hydro, thermal, geothermal, and more. Despite this progress, the focus has often been on developing new alternatives rather than optimizing existing energy resources. As a result, we need to catch up in effectively utilizing what is already available.[1,2]

Applying Machine Learning (ML) and Neural Networks (NN) forecasting techniques can enhance our understanding of energy production and consumption dynamics. Machine learning models are employed for:

i. Understanding weather conditions,

ii. Energy forecasting,

iii. Statistical analysis based on historical data,

iv. Creating suitable classification and regression models.[3,4]

Given the pressing need for efficient energy consumption amid a growing global population, analyzing historical data to train predictive models is imperative. These models can help determine when, where, and with what intensity energy production is impacted. Such analysis is essential for comprehending changing climatic conditions and maximizing energy capture.[5]

The interpretation of statistical data using machine learning can be simplified as enabling machines to learn from data patterns. Neural networks, which consist of interconnected nodes (including constant, variable, placeholder, and operating nodes), mimic the biological neural network's structure and function, allowing machines to perceive information similarly to the human brain.[6] In this context, machine learning can be utilized to address critical challenges in energy forecasting.

It can analyse historical data for energy forecasting and optimize energy flow by modelling a simulated brain that regulates the distribution of necessary energy across residential, commercial, and industrial sectors while also capturing excess power for future use. This approach aims to extend electrical supply to every corner of the globe, especially areas lacking sufficient energy access.

To comprehend the mechanisms required for energy production while forecasting and evaluating total energy consumption at various energy-generating locations, we need machine learning

OPEN ACCESS

Manuscript TECHNOMEDIA

DOI: 10.5530/bems.12.1.3

Publishing Partner : Manuscript Technomedia. [www.mstechnomedia.com]

regression models to fit the data and predict continuous values. Several regression models that have gained prominence in this context include:

i. Linear Regression,

ii. Support Vector Regression (SVR),

iii. Decision Tree Regression,

iv. Random Forest Regression,

v. K-Nearest Neighbours (KNN),

vi. Artificial Neural Networks (ANN).[7-9]

These models enable us to analyze historical energy data, identify patterns, and make accurate predictions regarding[44] future energy consumption and production. By employing these techniques, we can enhance the reliability of energy forecasting, ultimately supporting more effective energy management strategies.

LINEAR REGRESSION

The foremost algorithm in regression analysis is linear regression, which is known for its effectiveness in understanding profitability through insights and behaviour analysis of data. As the name implies, "regression" refers to fitting a line to data. At the same time, "linear" denotes the linear relationship between an independent variable (which describes characteristics or features of the data) and a dependent variable, also known as the target variable. Due to this univariate relationship, it is often referred to as simple linear regression.[10]

Extending this concept to encompass multiple independent variables leads to the term **Multilinear Regression**, which maps various feature variables to a single target variable.

Equation of Hypothesis

Linear Form

$$Y=Q_0+Q_1x_1+Q_2x_2+Q_nx_n$$

Matrix Form

$$Y= Q^Tx$$

Where:

a. Y: Actual output,

b. Q_0: Bias term,

c.: Q_i Model parameters,

d.: x_i Feature values ($x_0 = 1$).

The model fits the best line, known as the regression line, which encapsulates the influence of all features on the output, achieving an average loss or cost that is minimized. This is quantified by the threshold error between the predicted values Y_{pred} and the observed target values Y_{actual}.[11]

SUPPORT VECTOR REGRESSION

Support Vector Regression (SVR) is a robust algorithm that focuses on minimizing violations of the margin distance, fitting as many instances as possible within this margin. It can accommodate both linear and non-linear models by tuning three primary parameters:

i. **Tuning the C Parameter:** Also known as the regularization parameter, it is essential for managing noisy data.

ii. **Kernel Tuning:** Linear algebra is used to learn the hyperplane by adjusting the kernel.

iii. **Gamma Tuning:** This parameter determines the closeness of the training data points in the feature space.

Several types of kernels can be employed in SVR, including linear, polynomial, exponential, and Gaussian kernels:

i. Linear Kernel:

$$K(x) = B_0 + \sum_{i=1}^{n}(x, x_i)\alpha_i$$

ii. Polynomial Kernel:

$$K(x,x_i)=1+(x.x_i)^d$$

(Both linear and polynomial kernels are special cases of the Gaussian kernel.)

iii. Exponential Kernel:

$$K(x,x_i)=e^{((-\tau+(x.x\_i)2)}$$

iv. Gaussian Kernel:

$$G_{1D}(x;\sigma) = \frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-x^2}{2\sigma^2}}$$

$$G_{2D}(x,y;\sigma) = \frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(x^2+y^2)}{2\sigma^2}}$$

$$G_{ND}(x;\sigma) = \frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-|x|^2}{2\sigma^2}}$$

Where:

a. σ: Width of the Gaussian kernel

b. x: Input data point

Polynomial and exponential kernels employ the Kernel Trick to compute the separating hyperplane in higher-dimensional space. Training the model involves approximating the underlying function to achieve the desired outcome. The contraction coefficient (a vector of unknown variables) is calculated using the correlation matrix formula:

$$K_{i,j} = e^{-} \sum_{k=1}^{n} \theta_k |x_{ik} - x_{jk}|^2 + \epsilon\delta_{i,j}$$

The contraction coefficient (α) is determined using the dot product of the correlation matrix and the training dataset:

$$\alpha = K^{-1} y$$

Once the model is trained, testing is performed on a new data point:

$$y^* = \alpha . K_{i,j} = e^{-} \sum_{k=1}^{n} \theta_k |x_{ik} - x_{jk}|^2$$

A threshold value is set to minimize the difference between the actual and predicted outcomes, corresponding to the error.[12-14]

DECISION TREE

A Decision Tree can be interpreted as a[15] decision-making model structured like a tree. This model organizes the decision-making process by prioritizing attributes based on the information required to predict a continuous value at the output leaf of regression trees. The branches or splits in the tree are determined by choosing features and establishing conditions, which are assessed using measures such as Information Gain, Entropy, and the Gini Index. A common technique employed for splitting is known as Recursive Binary Splitting.

Recursive Binary Splitting: This is a greedy approach that aims to minimize the cost function of the attribute while selecting it as a binary node (splitting into two) at each branch based on information derived from the following tools:

i. **Entropy:** Entropy quantifies the degree of randomness or uncertainty in information processing, controlling how attributes are split based on the features they contain. Higher entropy indicates greater randomness and makes interpretation more difficult. The formula for entropy is:

$$\text{Entropy} = -\sum_{i=1}^{n} p_i \log(p_i)$$

ii. **Information Gain:** Information Gain measures the effectiveness of an attribute in reducing uncertainty, forming the core of Decision Tree creation:

$$\text{IG} = E(\text{parent}) - [\text{Average weight}] \times E(\text{child})$$

iii. **Gini Index:** Gini Impurity represents the probability of misclassification of a randomly selected data sample at each node based on the data sample distribution. A significant reduction in the Gini Index upon node splitting indicates the effectiveness of that split. The formula for Gini Index is:

$$Gini = 1 - \sum_{i=1}^{c} (p_i)^2$$

However, excessive splits can lead to increased complexity and overfitting when dealing with larger datasets. To mitigate this, methods such as setting a minimum training dataset size and a maximum depth for the decision tree can be implemented. Another effective technique is Pruning.

i. **Pruning Technique:** This technique optimizes the performance of Decision Trees by eliminating attributes with less significance or higher entropy. Two main approaches to execute this technique include:

ii.**Reduce Error Pruning:** This bottom-up method removes less significant child attributes without affecting accuracy.

Weakest Link/Cost Complexity Pruning: In this more complex method, a subtree is removed based on a learning parameter and the weight of its parent node.[16,17]

RANDOM FOREST

Random Forest builds on the Decision Tree concept, where the "forest" represents multiple Decision Trees considered simultaneously. The term "random" reflects that a random sample of the training dataset is used to construct the trees, and the selection of attributes for branching nodes is also random. This randomness is known as Bootstrapping.

Bootstrapping integrates randomness with replacing training samples for tree building and attribute selection for splitting under higher variance or lower entropy conditions. Due to this inherent randomness, the one with higher variance but lower cumulative variance is selected among the multiple Decision Trees formed. This predictive approach of choosing the optimal Decision Tree by averaging the variance is referred to as Bootstrap Aggregation or Bagging.[18-20]

K-NEAREST NEIGHBORS

K-Nearest Neighbors (KNN) is an easy-to-implement supervised machine learning algorithm used for both classification and regression tasks. As the name suggests, this algorithm relies on the proximity of data points, meaning that similar data items are located close to each other. To determine the closeness between data observations, the distance is calculated using the Minkowski Distance.

The **Minkowski Distance** is a generalized distance metric applied in a normed vector space, where the distance is measured between vectors. The term "normed" indicates that each vector has a non-negative length. The normed vector space is characterized by three properties: the zero vector, scalar factor, and triangular inequality, which governs the calculation of distance between points.

The formula for the Minkowski Distance is given by:

$$D(x,y) = (\sum_{i=1}^{n} |x_i - y_i|^p)^{1/}$$

This distance metric can be tailored for different applications by varying the parameter p:

i. **Manhattan Distance (p=1):** This distance is defined as the absolute sum of differences between Cartesian coordinates and is commonly used to evaluate grid paths. The formula is:

$$D(x,y) = \sum_{i=1}^{n} |x_i - y_i|$$

The L1 Norm, or Manhattan Distance, quantifies the absolute sum of differences in the vector space.

i. **Euclidean Distance (p=2):** This distance represents the straight-line distance between points. The formula is:

$$D(x,y) = (\sum_{i=1}^{n} |x_i - y_i|^p)^{1/2}$$

The L2 Norm, or Euclidean Distance, employs the Pythagorean Theorem to calculate distance. However, since the vectors are squared, this approach can skew results, as outliers disproportionately influence the distance.

i. Chebyshev Distance (p=∞): This distance is defined as the maximum absolute difference along any coordinate dimension:

$$D(x,y) = \max_i(|x_i - y_i|)$$

KNN's simplicity and effectiveness make it a popular choice for various machine learning applications, mainly when dealing with classification problems where the relationship between observations is crucial.[21-23]

EVALUATING METRICS

Loss Function

In regression tasks, loss functions measure the accuracy of predictions by quantifying the difference between the predicted and actual values. Below are five key loss functions for regression models,[24,25] each with a formula and explanation:

Mean Square Error (MSE)

Mean Square Error is the most commonly used loss function in regression. It calculates the average squared difference between the actual and predicted values, making larger errors more significant due to squaring.

$$J(Q) = \frac{1}{n}\sum_{i=1}^{n}(h(x_i) - y_i)^2$$

Where:

a. n: data size,

b. $h(x_i)$: hypothesis function,

$$h(Q) = Q_0 + Q_1x_1 + Q_2x_2 + \ldots + Q_nx_n$$

c. y: Actual output.

MSE is sensitive to outliers since errors are squared before they are averaged, giving more weight to larger discrepancies.[26]

Root Mean Square Error (RMSE)

Root Mean Square Error is the square root of the MSE. It offers a measure of how far, on average, the residuals are from zero, in the same units as the output variable.

$$J(Q) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(h(x_i) - y_i)^2}$$

RMSE is often used for its interpretability and is effective when large errors are particularly undesirable.[27]

Mean Absolute Error (MAE)

Mean Absolute Error measures the average of the absolute differences between the predicted and actual values. Unlike MSE, it does not penalize larger errors as heavily since no squaring is involved.

$$J(Q) = \frac{1}{n}\sum_{i=1}^{n}|h(x_i) - y_i|$$

MAE is robust to outliers and offers a more balanced metric for understanding prediction errors.[28]

CROSS VALIDATION

To ensure that the model has analysed and understood the data pattern without noise or without being overfitted/underfitting or with low bias, cross validation is required to statistically behold the stability of the model.

Cross Validation is a validation technique for the model to statistically examine the generalization pattern of the results on the independent dataset.

This model validation method provides a bit of flexibility over the splitting or groups or k-folds, which are as follows:

k-Fold (k=2)

It means the data is grouped into two i.e. the training and the test data. This type of grouping is opted if, we have enough data to make the model learn the pattern on a randomly trained training data. Any duplicacy and overlapping of grouped data should be avoid and final model - after testing - should be retrained on the complete dataset without any tuning in the hyperparameters.

k-Fold (k=3)

This is comparatively a better approach then binary grouping the dataset as the dataset is bifurcated into three, the training data, the validation data and the test data. To evaluate the quality of model fitted on trained data, model is validated (prior to testing) on a new sample (validation dataset). This pattern is chosen if the data size is sufficient enough to be grouped as such.

k-Fold (k)

For splitting the dataset, this is a prominent approach as the data available for to model the decipher the pattern is not never enough and model has to the problems of underfitting and increased loss.

In this method, the data is grouped into k folds and model is trained into k-1 times. Each time k-1 portion is trained and is validated over the remaining portion. Each time the model is trained is validated on a new piece of data which significantly reduces the underfitting and the overfitting problem. This method is chosen for a small sized data as the model is free from a high bias or a high variance.

FORECASTING

Forecasting can be defined as the process of analyzing and examining collected historical and current data to make predictions about future scenarios.29 In the context of optimal energy usage and supply, forecasting plays a crucial role in enhancing efficiency, reducing costs, and improving resource allocation.[30]

Forecasting techniques are applied across various domains, particularly in fields characterized by fluctuating parameters, such as energy consumption, temperature, and weather.[31] These techniques utilize statistical analysis to derive meaningful insights from observed data, which is subsequently fed into data processing phases to prepare datasets and apply regression algorithms.[32]

TIME SERIES FORECASTING

Time Series Forecasting is a technique used to predict[33] events based on sequences of observations over time.[34] It assumes that historical data patterns will continue. In energy forecasting, time series analysis helps identify trends, seasonality, and other cyclical patterns in energy usage.[35]

Critical aspects of time series forecasting include

i. Identification of Patterns: By analyzing historical data, various patterns such as trends (increasing or decreasing), seasonality (cyclical fluctuations), and noise (random variations) can be identified.[36]

ii. Stationarity: A stationary time series has constant mean and variance, making it easier to model.37 Statistical techniques are often employed to achieve stationarity by removing trends and seasonal effects.

The Augmented Dickey-Fuller Test (ADF Test) is commonly used to test for stationarity within a time series.[38] The null hypothesis of this test indicates the[39] presence of a unit root (non-stationarity), while the alternative hypothesis suggests the absence of unit roots (stationarity). A stationary time series will have consistent mean and variance, enabling more reliable predictions.[40]

A transformation method known as differencing is applied to remove dependencies in the series. This technique stabilizes the mean by eliminating trends and seasonal components, facilitating a more accurate analysis.[41]

Time series forecasting techniques have applications in various fields, including:

i. Weather Prediction: Forecasting energy demand based on temperature and weather patterns.[42]

ii. Pattern Recognition: Identifying usage trends and anomalies in energy consumption.[43]

iii. Economics: Analyzing and predicting market dynamics and energy pricing.[44]

iv. Earthquake Prediction: Utilizing time series methods to anticipate seismic events and their impact on energy infrastructure.[45]

ENERGY FORECASTING

Energy forecasting specifically refers to predicting future energy consumption, resource availability, and electricity prices.[46] This type of forecasting informs policymakers, utility companies, and consumers about expected energy trends, facilitating better planning and management of energy resources.[39]

CONCLUSION

In conclusion, Power System Forecasting is a critical process that enables utilities and energy managers to make informed decisions regarding energy generation, distribution,[47] and consumption. By utilizing various forecasting techniques tailored to different time spans-short-term, medium-term, and long-term-stakeholders can enhance the reliability and efficiency of power systems.

Short-term forecasting facilitates real-time management of electricity loads and trading, ensuring a balance between supply and demand. Medium-term forecasting aids in the optimization of energy storage solutions, especially for renewable sources, thus contributing to the sustainability of the energy sector. Long-term forecasting plays a vital role in strategic planning, enabling energy providers to anticipate future consumption trends influenced by socio-economic factors.

The application of statistical methods, such as regression models and time series analysis, allows for the effective prediction of energy needs, paving the way for a more resilient and responsive power system.[47] As the energy landscape continues to evolve, the integration of advanced forecasting models will be paramount in addressing the challenges posed by fluctuating energy demands and the increasing reliance on renewable resources.

CONFLICT OF INTEREST

The author declares that there is no conflict of interest.

REFERENCES

1. Coyle EJ, Moulton RJ. "Machine learning for energy forecasting: A review", *Renewable and Sustainable Energy Reviews*. 2020;119:109553. doi: 10.1016/j.rser.2019.109553.
2. Zhang H, Jiang Y. "Artificial neural networks for energy forecasting: A review", *Energy Reports*. 2021;7:1853-62. doi: 10.1016/j.egyr.2021.04.036.
3. Ahmed M, Khosravi A. "A review of machine learning approaches for energy consumption prediction", *Energy Reports*. 2019;5:2050-65. doi: 10.1016/j.egyr.2019.11.014.
4. Fathi A, Abedini A. 'Deep learning for energy management: A survey', *Renewable and Sustainable Energy Reviews*. 2020;131:110040. doi: 10.1016/j.rser.2020.110040.
5. Zelle D, O'Neil L. "Understanding linear regression: A comprehensive guide", *. J Stat Softw. 2021;98(2):1-28. doi: 10.18637/jss.v098.i02.
6. Marqués M, Pérez J. Neural networks: A comprehensive foundation, *Springer*. 2nd ed; 2019. doi: 10.1007/978-3-030-34879-3.
7. Kelleher JD, Tierney B. "Data Science: An Introduction to Statistics and Machine Learning", *. MIT Press; 2018. doi: 10.7551/mitpress/11105.001.0001.
8. Bishop CM. "Pattern Recognition and Machine Learning", *springer*; 2006. doi: 10.1007/978-0-387-31073-2.
9. Shalev-Shwartz S, Ben-David S. "Understanding Machine Learning: From Theory to Algorithms", *. Cambridge University Press; 2014. doi: 10.1017/CBO9781107298019.
10. The future of work in America: people and places, today and tomorrow. McKinsey Global Institute; 2020. doi: 10.1037/e307642020-001.
11. Hastie T, Tibshirani R, Friedman JH. "The Elements of Statistical Learning: Data Mining, Inference, and Prediction", *springer*. New York: Springer New York; 2009. doi: 10.1007/978-0-387-84858-7.
12. Smola AJ, Schölkopf B. "A tutorial on support vector regression", *Statistics and Computing*. Statistics and Computing. 2004;14(3):199-222. doi: 10.1023/B:STCO.0000035301.49549.88.
13. Chen S, Wang L. "Support vector regression for wind power forecasting: A review", *Renewable and Sustainable Energy Reviews*. 2019;113:109250. doi: 10.1016/j.rser.2019.109250.
14. Alpaydin E. "Introduction to Machine Learning", *. MIT Press; 2020. doi: 10.7551/mitpress/11436.001.0001.
15. Handling missing values in decision forests in the encrypted network traffic; 2018. Available from: https://core.ac.uk/download/186449328.pdf.
16. Breiman L. "Bagging predictors", *Machine Learning*. Mach Learn. 1996;24(2):123-40. doi: 10.1007/BF00058655.
17. Quinlan JR. "Induction of Decision Trees", *machine learning*. Mach Learn. 1986;1(1):81-106. doi: 10.1007/BF00116251.
18. Liu H, Motoda H. Springer; 1998. *Feature Selection for Knowledge Discovery and Data Mining*. doi: 10.1007/978-1-4615-5515-8.
19. Zhang H, Singer B. "Recursive partitioning and applications", *. Springer Science+Business Media; 2010. doi: 10.1007/978-1-4419-6913-6.
20. Zhang J, Yang Y. "A survey of Random Forest", *. Int J Comput Appl. 2015;113(14):1-8. doi: 10.5120/19857-8489.
21. Cover TM, Hart PE. Nearest neighbor pattern classification. IEEE Trans Inform Theory. 1967;13(1):21-7. doi: 10.1109/TIT.1967.1053964.
22. Altman NS. 'An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression', *The American Statistician*. 1992;46(3):175-85. doi: 10.2307/2684359.
23. Zhang Y, Wang J. "K-Nearest Neighbors Algorithm Based on Data Fusion", *. J Appl Math. 2016;2016. doi: 10.1155/2016/3976548.
24. Zhang Z. "A comprehensive review of loss functions in machine learning", *Artificial Intelligence Review*. 2020;53(3):2255-74. DOI: [10.1007/s10462-019-09780-6]. doi: 10.1007/s10462-019-09780-6.
25. Ahmed M, Khawaja A. "Performance evaluation of machine learning regression models", *. J Comp Sci. 2021;47:101-20. DOI: [10.1016/j.jocs.2020.101120]. doi: 10.1016/j.jocs.2020.101120.
26. Seber GA, Lee AJ. John Wiley & Sons. DOI: [10.1002/9781118279127]; 2012. *Linear Regression Analysis*. doi: 10.1002/9781118279127.
27. Willmott CJ, Matsuura K. "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance", *Climate Research*. Clim Res. 2005;30(1):79-82. DOI: [10.3354/cr030079]. doi: 10.3354/cr030079.
28. Huber PJ. "Robust Estimation of a Location Parameter", *. Ann Math Statist. 1964;35(1):73-101. DOI: [10.1214/aoms/1177703732]. doi: 10.1214/aoms/1177703732.
29. **Forecasting**; 2021. _Energy Forecasting Techniques_. DOI. doi: 10.1016/j.ijepes.2021.105123.
30. **Power System Forecasting**; 2020. _Short-term vs Long-term Energy Forecasting_. DOI. doi: 10.1109/TPWRS.2020.2986345.
31. **Energy Forecasting**; 2019. _Energy Resource Predictions_. DOI. doi: 10.1109/ACCESS.2019.2931145.
32. *Shiihara M, Higuchi R, Morita S, Furukawa T, Yazawa T, Uemura S, et al.*Time Series Forecasting**. Diagnosis by 64-row multidetector computed tomography for longitudinal superficial extension of distal cholangiocarcinoma. Journal of Surgical Research. 2019. _Techniques and Methods for Time Series Analysis_;235:487-93. doi: 10.1016/j.jss.2018.10.045.
33. Palvel S. Time series forecasting with prophet, *Medium*; n.d. [cited Oct 7 2024] Available from: https://subashpalvel.medium.com/time-series-forecasting-with-prophet-f6c0ccb80a7d.
34. **Statistical Analysis**; 2021. _Understanding Statistical Analysis for Time Series Forecasting_. DOI. doi: 10.1016/j.patcog.2021.107204.
35. **Pattern Recognition**; 2022. _Trends in Pattern Recognition for Energy Management_. DOI. doi: 10.1016/j.jenvman.2022.114136.
36. *Xia A, Sun C, Fu Q, Liao Q, Huang Y, Zhu X, et al.*Forecasting Techniques**. Biofuel production from wet microalgae biomass: comparison of physicochemical properties and extraction performance. Energy. 2020. _Review of Forecasting Techniques for Energy Demand_;212. doi: 10.1016/j.energy.2020.118581.
37. *Zhou Y, Zhang S, Wu L, Tian Y*Forecasting Models**. Predicting sectoral electricity consumption based on complex network analysis. Applied Energy. 2019. _Modeling and Forecasting Energy Consumption_;255. doi: 10.1016/j.apenergy.2019.113790.
38. **Statistical Tests**; 2020. _Statistical Tests for Time Series Analysis_. DOI. doi: 10.1016/j.swe.2020.100105.
39. Beneito P. The innovative performance of in-house and contracted Randamp;D in terms of patents and utility models. Res Policy. 2006;35(4):502-17. doi: 10.1016/j.respol.2006.01.007.
40. **Energy Analysis**; 2021. _Energy Consumption Analysis Using Time Series_. DOI. doi: 10.1016/j.apenergy.2021.117456.
41. **Transformations**; 2022. _Data Transformations for Forecasting_. DOI. doi: 10.1016/j.jss.2022.110123.
42. *Ivorra L, Cardoso PG, Chan SK, Cruzeiro C, Tagulao KA*Weather Forecasting**. Can mangroves work as an effective phytoremediation tool for pesticide contamination? An interlinked analysis between surface water, sediments and biota. Journal of Cleaner Production. 2021. _Impact of Weather on Energy Demand Forecasting_;295. doi: 10.1016/j.jclepro.2021.126334.
43. *Eleftheriadis S, Duffour P, Mumovic D*Usage Trends**. BIM-embedded life cycle carbon assessment of RC buildings using optimised structural design alternatives. Energy and Buildings. 2018. _Identifying Trends in Energy Consumption_;173:587-600. doi: 10.1016/j.enbuild.2018.05.042.
44. **Market Dynamics**; 2019. _Analyzing Market Dynamics in Energy Pricing_. DOI. doi: 10.1016/j.eneco.2019.104357.
45. *Riede F, Barnes GL, Elson MD, Oetelaar GA, Holmberg KG, Sheets P*Seismic Events**. Prospects and pitfalls in integrating volcanology and archaeology: a review. Journal of Volcanology and Geothermal Research. 2020. _Predicting Seismic Events Using Time Series Analysis_;401. doi: 10.1016/j.jvolgeores.2020.106977.
46. **Applications**; 2021. _Applications of Time Series Forecasting in Various Fields_. DOI. doi: 10.1016/j.swe.2021.100120.
47. Liu X, Liu X, Zhang X, Baziar A. "Hybrid Machine Learning and Modified Teaching Learning-Based English Optimization Algorithm for Smart City Communication", *sustainability*. 2023;15(15):11535. doi: 10.3390/su151511535.

Cite this article: Mishra U. Energy Forecasting and Time Series Analysis Using Machine Learning. BEMS Reports. 2026;12(1):9-14.