26th International Congress on Modelling and Simulation, Adelaide, SA, Australia, 30 Nov–4 Dec 2025
mssanz.org.au/modsim2025

The role of artificial intelligence techniques in forecasting renewable energy sources

SW Johnson[a], Asef Nazari[a], Mohammad Ghasemi[a], Dhananjay Thiruvay[a], Rosemary Van Der Meer[a], Fatemeh Shiri[b]

[a] *School of Information Technology, Deakin University, Geelong, 3216, VIC, Australia,* [b]*School of Computing Technologies, RMIT*

Email: asef.nazari@deakin.edu.au

Abstract: Short-term forecasts of renewable generation are essential for the operation of microgrids, particularly in coastal environments characterised by rapid weather changes. This study assesses a set of recent machine learning (ML) tools, including Support Vector Regression (SVR), Long Short-Term Memory (LSTM), a Transformer model, and a stacking ensemble for forecasting solar irradiation and wind speed in the Anglesea area in Victoria, Australia. A standard persistence model is used for comparison. The forecasting process began with thorough preprocessing, including imputation, Min–Max scaling and cyclical time encodings, alongside hyperparameter tuning. The mentioned tools were deployed to forecast the solar irradiation and wind speed. From the obtained results, LSTM demonstrates superior performance in volatile wind segments, whereas the Transformer effectively identifies structured solar irradiation patterns but encounters difficulties with unpredictable wind behaviour. SVR maintains competitiveness in stable conditions and provides consistent accuracy. The findings indicate that the integration of diverse learners enhances robustness in the context of the accuracy enhancement of weather-dependent resources' parameters forecasting.

Keywords: *Artificial Intelligence, Forecasting, Renewable energy, Microgrid, Long Short-Term Memory*

1. INTRODUCTION

When in relation to microgrid operation, short-term forecasting of renewable production is essential because it requires balancing supply, storage and demand in almost real time. The Anglesea coastal microgrid's solar irradiance and wind speed differ a lot due to the rapidly changing cloud cover, sea breezes and storm fronts (Voyant & Notton, 2018). Therefore, persistence or linear statistical models are inaccurate or cost-ineffective.

More commonly used methods, such as ARIMA, are often utilised for their ability to be easily understood and for their general reliability, but the assumptions of linearity and stationarity often fail when applied to weather-related time series (Ugbehe et al., 2025). Throughout this study, the selected AI and ML methods effectively follow the nonlinear patterns and time dependencies (Yang et al., 2024), but each model still comes with certain drawbacks despite their abilities. Support Vector Regression (SVR) offers a more direct and simple approach but is held back by its memory constraints (Santamaría-Bonfil et al., 2016). Long Short-Term Memory (LSTM) networks effectively learn sequence structures but require meticulous tuning (Li et al., 2024). Transformers are capable of extracting long-range patterns but can encounter difficulties when faced with noisy and highly irregular data, in addition to requiring greater computational resources (Hanif & Mi, 2024).

Utilising one year of hourly solar-irradiance and wind-speed observations, we benchmarked multiple models, including an SVR, a bidirectional LSTM, a two-layer Transformer and an ensemble that combines their outputs (workflow shown in Figure 1). All models share the same preprocessing consisting of median imputation for gaps, outlier filtering, Min-Max normalisation and cyclical time features. This balanced treatment enables a controlled comparison with a persistence baseline and reveals whether the ensemble indeed provides the most reliable forecasts, which would be similar to findings by other researchers (Shojaei and Mokhtar, 2024).

2. LITERATURE REVIEW

Accurate renewable-energy forecasting has evolved throughout the years in repeating jumps of new developments, each solving some downfalls of the last. Transparent linear models were succeeded by kernel-based machine-learning techniques, which then encouraged the evolution of deep-learning architectures and, most recently, hybrid and ensemble frameworks that can utilise each model's strengths. Nonetheless, microgrids in rapidly changing environments, especially coastal, still demand solutions that are accurate, timely and operationally practical. Even with all these advancements, only a handful of studies have thoroughly compared these models within a coastal-microgrid setting, leaving their relative strengths largely untested (Sina et al., 2023).

Microgrid operation relies heavily on accurately forecasting demand and weather-related renewable resources, a task that becomes especially challenging in certain coastal settings like Anglesea in Victoria, Australia, where the weather is highly unpredictable. Traditional statistical models, such as ARIMA and linear regression, are fast and efficient but rely on data with high levels of predictability, that leads to a reduction in their accuracy with more variable patterns common in renewable energy data (Ahmad et al., 2020). Another method that is less popular in a practical setting, but is more commonly used as a baseline, is the persistence model. However, due to its simplicity, it is also expected to perform poorly when tasked with the erratic behaviour of coastal environments.

2.1. Short-term Forecasting Models

Support Vector Regression (SVR) is a more common machine learning approach for forecasting nonlinear relationships with minimal feature engineering due to its versatility and ease of configuration. SVR has outperformed classical models using kernel methods across numerous energy applications (Sha'aban, 2024). Its understandability and small data requirements make it ideal for use in more complicated models. For instance, Ko and Lee (2013) used SVR to configure the hidden nodes and initial parameters of a radial-basis-function neural network (RBFNN), then refined that network with a dual extended Kalman filter (DEKF). This led to the creation of an SVR-DEKF-RBFNN ensemble that combines SVR's global generalisation, the RBFNN's local function-approximation power and DEKF's real-time adaptive learning. Hybrids like the previously mentioned show that, when SVR in incorporated correctly into a larger system, it can perform outstandingly well. However, the method alone still lacks an internal memory mechanism for sequential patterns, limiting its effectiveness when used on its own.

To address time patterns in data, researchers have begun to utilise neural networks. They found Long Short-Term Memory (LSTM) neural network performs especially well in time series data. LSTM architectures retain long-range dependencies using gated memory. They have also shown strong results in location-specific demand forecasting (Yang et al., 2024). A related architecture, the Gated Recurrent Unit (GRU), simplifies the LSTM design by combining certain gates and states, allowing for a reduction in computational cost while retaining accuracy (Li et al., 2024). This study does not evaluate GRUs, focusing instead on LSTMs due to their more popular role in renewable energy forecasting. Data quality is often an overlooked factor and can easily lead to lower model performance. This is generally affected by variables, including faulty sensors or telemetry errors, that can cause empty data entries or noisy readings. The imputation technique selected is also an essential part of its accuracy, especially in the case of battery systems where errors tend to compound over time (Pazhoohesh et al., 2021).

Transfer learning has become more popular due to its ability to reduce large, task-specific datasets and the common problem of overfitting. By fine-tuning pretrained models, researchers were able to increase accurate energy forecasts with limited local data, which could prove useful in scenarios that lack an abundance of data (Sharma et al., 2024). However, it also comes with its downsides, including transferring possible faulty or low-quality learning to other models (Sharma et al., 2024). Transformers, originally for large language models, have been repurposed for time series thanks to self-attention, which can focus on long-range dependencies without relying on recurrent connections. They have surpassed other models in several solar and wind studies (Piantadosi et al., 2024; Huang et al., 2023), but their computational demands are not ideal for typical edge-deployed microgrid controllers unless optimised. Some deep learning models also raise concerns due to their low explainability. LSTMs offer limited insight, and Transformers are inherently opaque (Machlev et al., 2022). Add-on interpretability techniques, tools like SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations), and attention visualisation can reveal some of the models reasoning behind certain decisions (Machlev et al., 2022). An example of an architecture that reduces this issue is the Temporal Fusion Transformer, which can incorporate explainability directly (Lim et al., 2021). Ensemble learning, in particular stacking, refers to when a meta-learner is made up of several base models, which can significantly boost forecast accuracy. Shojaei and Mokhtar (2024) showed that a two-level stacking architecture using deep-learning and k-nearest-neighbour predictors remained reliable even during highly diverse conditions. Decomposition ensemble hybrids that model trend, seasonal trends and noise separately have also proved to adapt well to highly variable wind-speed data. A recent survey confirms ensembles superior performance above single-model architectures (Sina et al. 2023).

2.2. A Hybrid Approach to Forecasting

Although deployment has improved, maintaining multiple models still raises training and operational costs. This is especially true in low resource environments (Li et al., 2023). Techniques like weight pruning, precision reduction and smarter model switching can help balance accuracy and scalability. Operators also need forecasts that are both fast and easy to understand. Machlev et al. (2022) recommend utilising simple if-then rule layers or automatic alert systems so human operators can quickly understand and respond to the output of a model and act on it with a higher level of confidence.

Accuracy of models has been the priority of the past, but now there are also other factors we must consider. Deep models are being altered with operational goals in mind, including reducing emissions, storage efficiency and reliability through multi-objective optimisation (Aguilar et al., 2024). Federated learning is another example of this, it enables collaborative model training across microgrids without sharing raw data (Li et al., 2023), though it introduces its own challenges, such as synchronisation and communication issues.

From reviewing the previous works in the literature, the following gaps are identified.

1- Linear models fail to capture the nonlinear, rapidly changing weather patterns of renewable-focused microgrids, especially those in coastal environments.
2- Many AI models struggle with sudden wind-ramp events and changing cloud cover, which is almost always a factor of coastal systems.
3- Previous work rarely combines multiple complementary AI techniques into an architecture that is designed for highly variable conditions.
4- Complicated and more in-depth models often exceed the computational and time limitations of what is practical in microgrid environments.

In this investigation, we address these gaps by providing the following contributions:

1- We benchmark persistence, SVR, LSTM and Transformer models on the same coastal microgrid data with identical preprocessing and feature design.
2- A ridge-regression meta-learner integrating SVR, LSTM and a Transformer achieves the lowest errors and delivers consistent performance during erratic weather conditions.
3- We show that LSTMs are especially efficient at forecasting erratic wind patterns, Transformers at structured solar patterns, while the ensemble offers the most reliable all-around performance and can be directly applied to real-time microgrid operations.

SVR, LSTM, Transformers and ensemble methods have all advanced renewable energy forecasting. In the case of a microgrid like Anglesea, the desired model is the one that can strike the right balance between accuracy, interpretability and runtime efficiency while also adapting to the quality of the data, the system's hardware limits and the needs of its operators.

3. METHODOLOGY

Long Short-Term Memory (LSTM) is a specific type of recurrent neural network (RNN) where ordinary hidden units are replaced by memory cells that have input, forget, and output gates. These gates allow the network to choose whether to keep or discard information throughout the timesteps. This helps to solve the vanishing-gradient problem that usually limits more common RNNs. Renewable outputs often reflect daily weather patterns, because of this, long-range memory allows LSTMs to perform well in short-term energy forecasting. In our framework, the LSTM is fed a 24-hour sliding window of past solar irradiance and wind speed data which has eight cyclical time features tethered to each value. A bidirectional system analyses each window of the past and future data that allows events later in the windows, such as rapidly changing wind gusts, to change how past patterns are processed. Once the LSTM has extracted the key time related patterns, its output is passed to a fully connected layer that turns those patterns into hourly predictions for both solar irradiance and wind speed. The LSTM is trained with a gradient optimiser which includes a learning-rate decay and early stopping to avoid over-fitting. The resulting model captures daily irradiance ramps, evening wind shifts and other coastal weather patterns far more effectively than classic feed-forward alternatives. Thus, the LSTM provides a strong example of comparison for the performance of all other models.

Figure 1 provides an overview of the entire workflow, from raw data through feature engineering, the three base learners and the ridge-stacked ensemble.

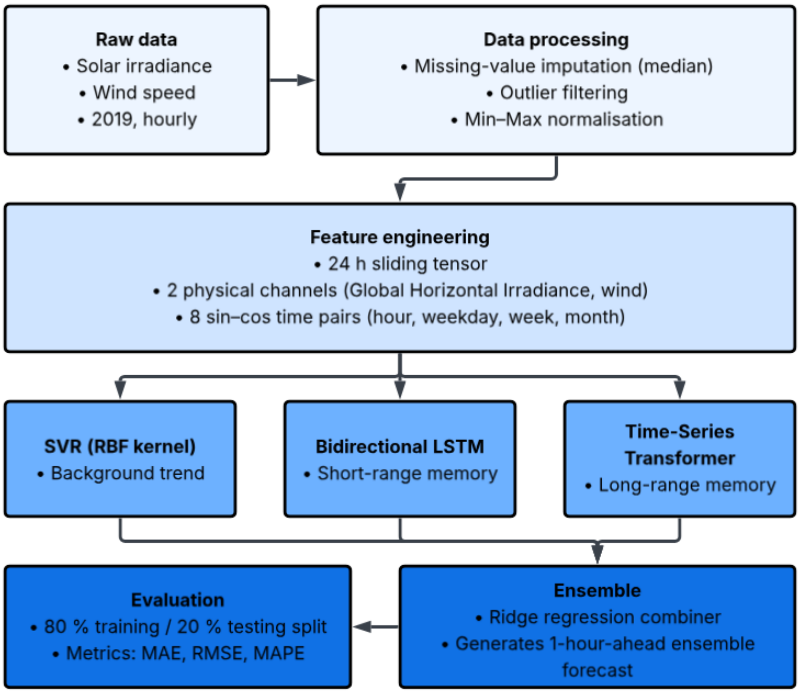

Figure 1 Workflow for renewable-energy forecasting ensemble

Hourly solar-irradiance and wind-speed series for Anglesea covering 1 January 2019 – 31 December 2019 were downloaded from the Renewables.ninja repository (Pfenninger and Staffell 2016) on 1 March 2025. Missing entries were filled with a median value, and values that are considered as physically impossible, such as negative irradiance, were all deleted. All values are then rescaled to [0,1] using Min–Max normalisation to keep gradients consistent and uniform across models. All models were then trained to perform one-step-ahead forecasting at an hourly resolution, where each forecast corresponds to the next hour's solar irradiance or wind speed.

Each forecast is a result of a tensor that combines values with a calendar-like context. A 24-hour sliding window encapsulates the recent trends of irradiance and wind, which isolates daily patterns and ramp events. The tables rhythms repeat in cycles consisting of hours, weekdays, weeks, and months, but ordinary numbers break that smooth pattern (the jump from 23 to 0 looks huge). To resolve this issue, each calendar field is converted into a pair of sine-and-cosine values. This technique turns the clock into a circle, so 23 sits right next to 0. Once we added the smooth calendar pairs to the irradiance and wind readings, every timestamp supplies 14 input columns in total, giving the models enough detail to track daily patterns much more efficiently.

Three base learners were trained on identical feature tensors. A support-vector regressor with a radial-basis kernel provides a strong estimate of underlying trends, while a bidirectional LSTM captures the non-linear, time-related patterns that can be observed during morning irradiance jumps and evening wind ramps. Alongside them is a Transformer with two self-attention layers and four heads each, that can inspect the full 24-hour window in one sweep, identifying broader moments in the solar and wind data. Each model produced a one-hour-ahead forecast, aligned to the hourly resolution of the dataset, and the three numbers were submitted into a ridge-regression combiner with a mild ℓ_2 penalty, which prevents any singular learner from overpowering others and keeps all learners relevant. A positive coefficient rewards more accurate predictions from models, and a negative one counters that effect. All models use the first 80% of the year for training and the final 20% for testing. This way, the evaluation reflects a more realistic deployment with no visibility of future actual values.

The resultant values were then evaluated with the three metrics defined in Equations (1) to (3). Equation (1), the mean absolute error (MAE), measures how far, on average, the set of predictions are from the true values. Equation (2), the root mean square error (RMSE), applies extra emphasis on occasional large errors that by squaring the errors before averaging. Equation (3), the mean absolute percentage error (MAPE), shows forecast accuracy as an average percentage that makes it easy to understand and compare to other data sets of different sizes. By using all three evaluation methods we can ensure a fair comparison of every base learner and the final ensemble product.

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|yt - \hat{y}t| \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(yt - \hat{y}t)^2} \quad (2)$$

$$MAPE = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{yt - \hat{y}t}{yt}\right| \quad (3)$$

Where, yt is the observed (true) value at time step t, $\hat{y}t$ is the predicted value at the same step, and n counts the number of forecasts compared with the actual value.

4. RESULTS & DISCUSSION

4.1. Error Summary on the 2019 Block

Table 1 displays the MAE, RMSE and MAPE related to the final 20% of the year. Reported MAE and RMSE values are in kilowatts (kW) for both solar PV output and wind power output, while MAPE is expressed as a percentage. The persistence, which is a standard baseline in solar-and-wind work (Voyant & Notton 2018), yields 12.68% MAPE for wind and a 33.35% for solar, emphasising the volatile nature of the coastal environment. SVR, an approach popularised from its effectiveness in short-term tasks (Ko & Lee 2013), cuts wind MAPE to 8.64%. The bidirectional LSTM lowers every solar metric, reaching 5.68% MAPE, which is expected due to its ability of gated-memory networks that can capture non-linear seasonal patterns (Li et al. 2024). The two-layer Transformer follows the irradiance curve almost as closely, similar to the success of attention mechanisms reported by Hanif & Mi (2024) but shows larger errors when there are ramps in wind speed. When the three learners are merged into an ensemble, the attributes of each are inherited. Solar MAPE drops to 4.56% which is a vast improvement over the persistence baseline, while wind RMSE also outperforms all other models.

Table 1 comparing the methods based on assessment criteria.

Model	Wind MAE	Wind RMSE	Wind MAPE	Solar MAE	Solar RMSE	Solar MAPE
Naive	0.0378	0.0603	12.68%	0.0694	0.1055	33.35%
SVR	0.0266	0.0489	**8.64%**	0.0174	0.0376	9.83%
LSTM	**0.0253**	0.0460	9.17%	0.0162	0.0349	5.68%
Transformer	0.0379	0.0577	13.79%	0.0193	0.0394	7.34%
Stacking	0.0261	**0.0449**	11.66%	**0.0104**	**0.0191**	**4.56%**

4.2. Day-level behaviour of the models

Figure 2 examines a day with clear sky conditions. All machine-learning curves align with the observed ramp, yet the ensemble remains closest to the midday peak. Interestingly the SVR performs well early, but underperforms later in the day. Figure 3 captures a wind-shift event. SVR responds too sharply to the 03:00 gust, the Transformer under-reacts and the LSTM times the drop but exaggerates the midday low. By averaging these errors, the ensemble stays on the correct trajectory while controlling the overshoot, reducing the daily RMSE by more than any other single learner.

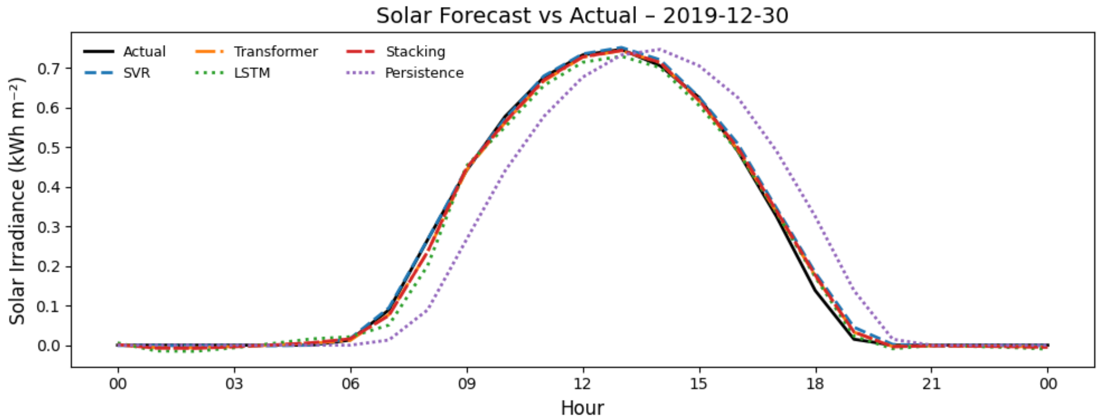

Figure 2 Hour-ahead solar-irradiance forecasts for 30 December 2019

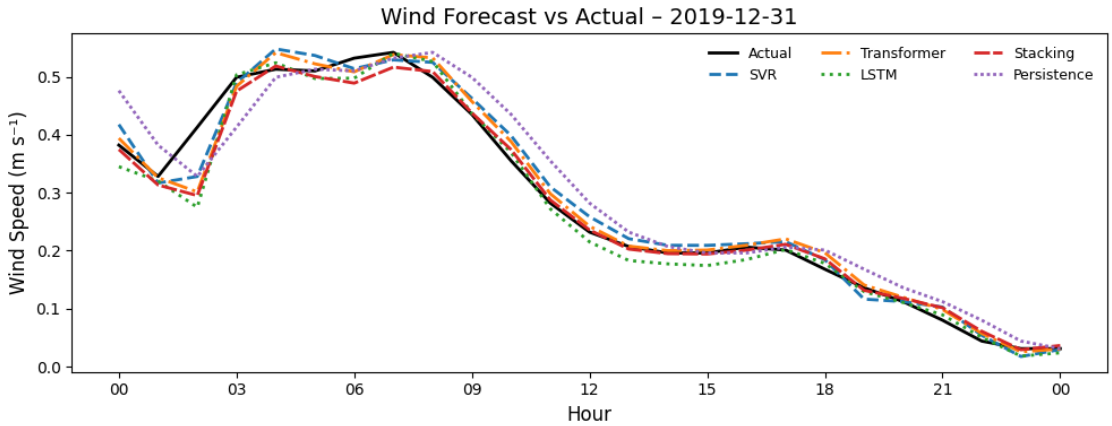

Figure 3 Hour-ahead wind-speed forecasts for 31 December 2019

4.3. Context within recent literature and operational relevance

The ensemble keeps solar-irradiance error relatively under control and inside the single-digit range which is generally considered adequate for routine micro-grid scheduling (Ahmad et al. 2020). This performance is due to the ensemble allowing each learner to address a different aspect of the coastal data and then using that with a simple ridge layer whose weights remain easy to inspect, which is an attribute highlighted as essential in recent discussions of explainable AI for power systems (Machlev et al. 2022). Because the combiner adds only a few coefficients, the full stack remains light enough to run on a possible low-power controller at Anglesea while delivering forecasts accurate enough for real-time storage and effective fuel consumption.

5. CONCLUSION & FUTURE WORK

This study's main intention was to improve short-term renewables forecasting for the Anglesea, Victoria, Australia microgrid by combining three base-level learners. An SVR for analysing smooth background trends, a bidirectional LSTM for local sequential structure and a light two-layer Transformer for long-range multi-variable patterns, all combined into a ridge-stacking ensemble. Trained on a year of hourly solar irradiance and wind-speed data (Pfenninger and Staffell 2016), the ensemble achieved a high level of performance on all the selected evaluation metrics. The result confirms what the literature review revealed initially. No single architecture captures every feature of weather-driven variability, and a transparent ensemble architecture can combine base learners' forecasts without sacrificing understandability. The ensemble's positive weights were able to amplify the LSTM and Transformer, where each excelled, while small negative weights damped the SVR's tendency to lag when faced with rapid wind ramps. The entire system was trained chronologically and tested on 20% of the data that had not been used for training, which allowed the reported gains to translate directly into a practical scenario. Multiple areas of possible future work have also become clear. A more diverse and extensive assortment of data, including sky cloud cover imagery or even previous power output levels, could aid the model's response to fast-moving fronts and ramp events. Another area of possible improvement could be to incorporate ranges into forecasts instead of single numbers. Grid operators could benefit from becoming aware of a likely spread of outcomes, not just the singular most possible value. Swapping the current ridge combiner for a method, such as quantile regression or Bayesian averaging, would let the model show an output with levels of confidence for a range of possible forecasts. Models would also benefit greatly from the incorporation of explainability tools. A strong example of this is attention heat-maps or some sort of gate activation monitoring ability, which could better explain what features the ensemble is using. This could give engineers more confidence that forecasts made by the model rest on real physical patterns rather than weak similarities. Exploring these possibilities could lead to a well-rounded and highly adaptable forecaster. One that can process highly detailed data, report confidence levels and explain how it reached each prediction. Beyond improving hourly accuracy in coastal environments, these future works will push the ensemble's ability to work in diverse conditions, which could bring the field closer to a reliable and transparent renewable-energy forecasting for micro-grids.

REFERENCES

Aguilar, D., Quinones, J. J., Pineda, L. R., Ostanek, J. & Castillo, L. (2024) Optimal scheduling of renewable energy microgrids: A robust multi-objective approach with machine learning-based probabilistic forecasting. *Applied Energy*, 369, 123548.

Ahmad, T., Zhang, H. & Yan, B. (2020) A review on renewable energy and electricity requirement forecasting models for smart grid and buildings. *Sustainable Cities and Society*, 55, 102052.

Hanif, M. F. & Mi, J. (2024) Harnessing AI for solar energy: Emergence of transformer models. *Applied Energy*, 369, 123541.

Huang, S., Yan, C. & Qu, Y. (2023) Deep learning Transformer based wind power forecasting approach. *Frontiers in Energy Research*, 10, 1055683.

Ko, C. & Lee, C. (2013) Short-term load forecasting using SVR (support vector regression)-based radial basis function neural network with dual extended Kalman filter. *Energy*, 49, 413–422.

Li, H., Li, S., Wu, Y., Xiao, Y., Pan, Z. & Liu, M. (2024) Short-term power load forecasting for integrated energy system based on a residual and attentive LSTM-TCN hybrid network. *Frontiers in Energy Research*, 12, 1384142.

Li, Y., He, S., Li, Y., Shi, Z. & Zeng, Z. (2023) Federated multiagent deep reinforcement learning approach via physics-informed reward for multimicrogrid energy management. *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–11.

Lim, B., Arik, S. O., Loeff, N. & Pfister, T. (2021) Temporal Fusion Transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4), 1748–1764.

Machlev, R., Heistrene, L., Perl, M., Levy, K. Y., Belikov, J., Mannor, S. & Levron, Y. (2022) Explainable Artificial Intelligence (XAI) techniques for energy and power systems: Review, challenges and opportunities. *Energy and AI*, 9, 100169.

Pazhoohesh, M., Allahham, A., Das, R. & Walker, S. (2021) Investigating the impact of missing data imputation techniques on battery energy management system. *IET Smart Grid*, 4(2), 162–175.

Pfenninger S and Staffell I (2016) Renewables.ninja: open database of hourly global wind and solar generation profiles. Available at https://www.renewables.ninja (accessed 1 March 2025)

Piantadosi, G., Dutto, S., Galli, A., De Vito, S., Sansone, C. & Di Francia, G. (2024) Photovoltaic power forecasting: A Transformer based framework. Energy and AI, 18, 100444.

Santamaría-Bonfil, G., Reyes-Ballesteros, A. & Gershenson, C. (2016) Wind speed forecasting for wind farms: A method based on support. *Renewable Energy*, 85(1), 790–809.

Sha'aban, Y. A. (2024) Predictive models for short term load forecasting in the UK's electrical grid. *PLOS ONE*, 19, 1–23.

Sharma, S. D., Coursey, A., Quinones-Grueiro, M. & Biswas, G. (2024) Comparison of Transfer Learning Techniques for Building Energy Forecasting. *IFAC-PapersOnLine*, 58(4), 180–185.

Shojaei, T. & Mokhtar, A. (2024) Forecasting energy consumption with a novel ensemble deep learning framework. *Journal of Building Engineering*, 96, 110452. Sina, L. B., Secco, C. A., Blazevic, M. & Nazemi, K. (2023) Hybrid forecasting methods—a systematic review. *Electronics*, 12, 2019.

Sorkun, M. C., Paoli, C. & Incel, Ö. D. (2017) Time Series Forecasting on Solar Irradiation using Deep Learning. *10th International Conference on Electrical and Electronics Engineering (ELECO)*, Istanbul.

Ugbehe, P. O., Diemuodeke, O. E. & Aikhuele, D. O. (2025) Electricity demand forecasting methodologies and applications: a review. *Sustainable Energy Research*, 12(19), 1–32.

Voyant, C. & Notton, G. (2018) Solar irradiation nowcasting by stochastic persistence: A new parsimonious, simple and efficient forecasting tool. *Renewable and Sustainable Energy Reviews*, 92, 343–352.

Yang, Y., Lou, H., Wu, J., Zhang, S. & Gao, S. (2024) A survey on wind power forecasting with machine learning approaches. *Neural Computing and Applications*, 36, 12753–12773.

