Wiley
International Journal of Intelligent Systems
Volume 2025, Article ID 8890906, 18 pages
https://doi.org/10.1155/int/8890906

WILEY

Research Article

Long Short-Term Memory Wavelet Neural Network for Renewable Energy Generation Forecasting

Eliana Vivas,[1] **Héctor Allende-Cid**,[1,2,3] **Lelys Bravo de Guenni**,[4] **Aurelio F. Bariviera**,[5] **and Rodrigo Salas**[6,7,8]

[1]*School of Computer Engineering, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile*
[2]*Knowledge Discovery Department, Fraunhofer-Institute of Intelligent Analysis and Information Systems (IAIS), Sankt Augustin, Germany*
[3]*Lamarr Institute for Machine Learning and Artificial Intelligence, Dortmund, Germany*
[4]*Department of Statistic, University of Illinois at Urbana, Champaign, Illinois, USA*
[5]*Department of Business, ECO-SOS, Universitat Rovira i Virgili, Reus, Spain*
[6]*School of Biomedical Engineering, Universidad de Valparaíso, Valparaíso, Chile*
[7]*Center of Interdisciplinary Biomedical and Engineering Research for Health (MEDING), Universidad de Valparaíso, Valparaíso, Chile*
[8]*Millennium Institute for Intelligent Healthcare Engineering (iHealth), Santiago, Chile*

Correspondence should be addressed to Héctor Allende-Cid; hector.allende@pucv.cl

Received 4 November 2023; Revised 13 December 2024; Accepted 20 December 2024

Academic Editor: Eugenio Vocaturo

Renewable energy forecasting is crucial for pollution prevention, management, and long-term sustainability. In response to the challenges associated with energy forecasting, the simultaneous deployment of several data-processing approaches has been used in a variety of studies in order to improve the energy–time-series analysis, finding that, when combined with the wavelet analysis, deep learning techniques can achieve high accuracy in energy forecasting applications. Consequently, we investigate the implementation of various wavelets within the structure of a long short-term memory neural network (LSTM), resulting in the new LSTM wavelet (LSTMW) neural network. In addition, and as an improvement phase, we modeled the uncertainty and incorporated it into the forecast so that systemic biases and deviations could be accounted for (LSTMW with luster: LSTMWL). The models were evaluated using data from six renewable power generation plants in Chile. When compared to other approaches, experimental results show that our method provides a prediction error within an acceptable range, achieving a coefficient of determination (R^2) between 0.73 and 0.98 across different test scenarios, and a consistent alignment between forecasted and observed values, particularly during the first 3 prediction steps.

Keywords: deep learning; energy generation forecasting; long short-term memory neural network; renewable energy; time-series forecasting; wavelet analysis

1. Introduction

Forecasting electric power is important in the management and balancing of existing power systems [1]. The fact that electric power systems are part of the growing trend of improving the environment around the world makes it harder to generate power and get it to where it needs to go because of the size and complexity of these changes. The load

projections aid in identifying ways to optimize the operational mechanisms over a certain time period, ensuring demand even under unfavorable system conditions [2].

Along with significant improvements in forecasting theory [3, 4], technology in the energy forecasting research domain [5] has also advanced swiftly. In the context of energy, the current research has focused on the development of models with the highest forecasting precision. It is important to keep in mind that the right modeling approach can help find the best ways to improve the design and implementation of power forecasting systems. Notably, a robust and reliable modeling system can be an effective and economical solution to reduce the number of experiments required and, ideally, limit the experimental runs only to the boundary cases [6].

In this way, machine learning (ML) techniques have demonstrated remarkable success in forecasting processes. Compared to traditional approaches for modeling very complex systems, ML techniques based on data have showed improved predictive power [6, 7]. This powerful tool has been widely used in similar processes, such as weather index forecasting, diverse type of renewable power forecasting, economic indicators forecasting, and others.

Over the years, notable data-driven techniques such ARMA models [8], support vector regression (SVR) [9], graph convolutional networks (GCNs) [10], fuzzy models [11, 12], neurofuzzy models [13, 14], and graph-based artificial neural networks (graph-ANN) [15] have been utilized to capture spatial and temporal dependencies within environmental data and energy networks obtaining remarkable results in short- and long-term prediction [16–19].

Also, techniques based on deep neural networks [20, 21] have been widely adopted, being wavelet-based models popular in signal decomposition, due to their resilience and accuracy [22]. The outstanding multiresolution properties of the wavelet transform (WT) enable us to manage variability with remarkable accuracy [23]. Among the context of neural networks, the wavelet neural network (WNN) and the long short-term memory (LSTM) neural network stand out. The WNN combines the WT with ANNs' capabilities, becoming a standard tool in math and engineering [24]. WNNs compress better than other ANNs, helping to learn nonlinear data features [25]. The LSTM, on the other hand, is a form of the recurrent artificial neural network proposed by Hochreiter and Schmidhuber [26]. This network performs effectively for time series learning training tasks, particularly when nonlinear signal modeling situations are included, according to Di-Giorgi et al. [27] and López-Gonzales et al. [28]. Several authors have utilized this approach for power forecasting [1, 29, 30].

Finally, because of the high performance of the implementation of the WT in the context of time-series forecasting, especially when combined with LSTM neural networks [1, 31], in this research, we explore the implementation of diverse wavelets inside the structure of an LSTM neural network creating a novel model-denominated LSTM wavelet (LSTMW) neural network. In this new model, we modified the update gate inside the structure of an LSTM implementing high-compression activation functions.

In addition, an improvement phase was also implemented for the predictions in this model (LSTMW with luster: LSTMWL), in this phase, residuals from previously fitted power values are added to the power prediction using lagged regression models, so the systemic biases and deviations could be accounted for in the forecast values. All the proposed techniques were implemented at six power stations in Chile and were compared with powerful forecasting methods. The novelty and major contributions of this study are as follows:

- The introduction of two novel LSTM neural network models that integrate wavelet functions into the update gates.
- The uncertainty was incorporated into the forecast to account for the systemic biases and deviations in the forecast values.
- The conclusive arguments employed underscore the strength of the proposed approach, achieved through the integration of various metrics to evaluate forecast precision. These metrics were thoughtfully designed to accommodate the complex temporal and spatial variations inherent in energy data time series. The application of cross-validation across six distinct power generation plants further enhances the validity of these metrics.
- In the context of forecasting, the performance of the proposed models is compared to that of various potential approaches, such as hybrid and conventional models equipped with high-quality performance associated with a grid search hyperparameter implementation.

The paper is organized as follows: In Section 2, we present the state of the art and relevant literature review. In Section 3, we review the materials and procedures used, as well as the proposed approach and other models involved. The results and discussion follow in Section 4. Section 5 presents the conclusions and future directions for this investigation.

2. State of the Art and Relevant Literature Reviews

In recent years, the field of energy generation data forecasting, particularly in the context of renewable energy, has been the subject of numerous studies and innovative approaches. Due to the growing body of the literature in this area, several literature reviews have been published that provide a comprehensive view of the current methods, models, and techniques. These reviews are essential to identify common patterns, research gaps, and areas for future development.

The following (Table 1) is a summary of some relevant reviews in the field, focusing on studies that address the combination of artificial intelligence, statistics, and hybrid methods for prediction in energy generation systems. This collection provides an overview of the most discussed approaches in the literature and serves as the foundation for

Table 1: Reviews of the literature on methods related to energy generation forecasting.

Reference	Year	Notable methods	Description
[32]	2020	Combining forecasting and multiple regression model	This article presents a succinct review of influential studies in energy forecasting, outlines current research trends, stresses the importance of reproducibility in research, and highlights key open data sources. It also suggests ways to improve the quality of published research and provides insights into the future developments in energy forecasting
[33]	2021	Hybrid forecasting models	This paper reviews recent deep learning–based solar and wind energy forecasting, focusing on data, preprocessing methods, and evaluation techniques. It summarizes key studies in tables, highlights current challenges, and suggests future research directions
[34]	2024	Tree-based models and ensemble methods	This review covers advanced forecasting techniques for the energy industry, focusing on statistical, machine learning (ML), and deep learning (DL) methods and their ensembles. It explores various time-series forecasting techniques and evaluation criteria, including MAE, RMSE, MAPE, R2, CVRMSE, and execution time (ET). The paper also presents a standard methodology commonly used in energy forecasting
[35]	2023	Data-driven models based on AI	This review discusses machine learning and deep learning algorithms for wave energy forecasting and optimizing wave energy converters (WECs), highlighting how combining these methods improves forecasting and WEC configuration. It also emphasizes the strong capability of these algorithms in forecasting wave resources and energy characteristics
[36]	2024	AI application, data cleaning, and reprocessing	This study surveys AI methods, including traditional AI and metaheuristics, for forecasting the availability of solar and wind energy systems. It presents a taxonomy of AI in renewable energy, explores metaheuristic algorithms optimizing AI methods, and formulates a conceptual framework for hybrid AI applications. The study also discusses challenges, insights, and future directions
[1]	2020	Long short-term memory and wavelet	This study reviews models for electric power forecasting, comparing classical, statistical, and machine learning models. It highlights the effectiveness of hybrid models and finds that forecasting errors decrease with shorter time horizons, ML models with multiple exogenous variables improve accuracy, and forecasting performance has significantly improved in the past 5 years

the innovative developments and proposals presented in this work.

The reviewed literature emphasizes notable advancements in energy forecasting, particularly with the incorporation of advanced ML and deep learning methods. The studies presented in this chapter offer a strong basis for understanding both the advantages and challenges of different forecasting techniques.

3. Materials and Methods

3.1. Study Area and Power Energy Generation Data. The hourly power (MW) generated at six Chilean stations was selected for the analysis. Given the pivotal role of data quality in the success of projects pertaining to energy resource management and planning, our station selection process was meticulous. We were careful to select stations whose records had the least amount of missing data for the time periods under consideration and were obtained from reliable sources. Each location shows differences in both seasonal patterns and power generation amounts. The common period of record used for the six stations was between April 2020 and April 2021 (8761 records). A summary of the power generation data from each station is shown in Figure 1. The differences in average generation between the hydroelectric plants are inherent to their distinct design and generation capacities.

Figures 2 and 3 show both the hourly power generation line charts and boxplots for the six locations. Each box height represents the interquartile range. The whiskers extend to the most extreme data point, which is no more than 1.5 times the interquartile range from the box. The hourly variations in generation levels from the six examined stations are visually compared. The data series exhibit seasonal patterns with fluctuations, and their intraday levels are not constant throughout time. There are variations in power generation levels between the highest and lowest points at each station.

We conducted a detailed analysis of the autocorrelation function (ACF) and partial ACF (PACF) to the hourly power generation time series to identify the most relevant lags to include as inputs for our algorithm. This step ensures that the model effectively captures the temporal dependencies and recurring patterns present in the data. We remove the diurnal cycle from the time series to determine the predictor's lag. The ACF and PACF plots are shown in Figure 4.

3.2. LSTMW Neural Network. Our initial proposed model is called the LSTMW neural network. It represents an extension of the conventional LSTM neural network. In the LSTMW, the cell state update incorporates a wavelet function for activation instead of the commonly used hyperbolic tangent (Tanh) family. The node responsible for the cell state update employs wavelet coefficients derived from an expansion function, with specific coefficients exhibiting significant values.

The fundamental concept behind wavelets revolves around interpreting data based on scale. Smaller scales represent higher frequencies, while larger scales correspond to lower frequencies [37]. Likewise, the WT achieves excellent frequency resolution for low-frequency events and reliable time resolution for high-frequency events [37]. Moreover, the WT adapts to different frequency characteristics present in a wide range, allowing the examination of time series with nonstationary dynamics at various frequencies [38, 39].

This approach is novel, as the existing literature on energy demand prediction models typically involves techniques that do not incorporate the WT within the network algorithm. Instead, they perform the decomposition of time series into high- and low-frequency components (wavelet) as a preprocessing step external to the neural network algorithm architecture [40–48].

3.2.1. LSTMW Neural Network Scheme. The proposed LSTMW cell architecture (depicted in Figure 5 and defined by Equations (1)–(8)) comprises interconnected memory blocks, forming a recurrent network structure. Each block consists of one or more self-connected memory cells and three multiplicative units: the input (i_t), output (o_t), and forget (f_t) gates in time t. These gates operate as nonlinear summation components, collecting activations from within and outside the block and controlling cell activation via multiplicative operations (represented by small circles with "x" in Figure 5).

The input (i_t) and output (o_t) gates are computed based on the previous output (h_{t-1}), the input vector (x_t), and the weight matrices from the input (w_i) or output (w_o) layers of the cell. In addition, the corresponding bias vector is added. On the other hand, the forget gate (f_t) multiplies the cell's previous state (h_{t-1}). The cell state (c_t) is determined by combining the forget gate (f_t) and the previous cell state (c_{t-1}). This result is then summed with the input gate (i_t) and the cell update state ($\tilde{c}t$), which represents a wavelet (Φ) layer. The wavelet layer is calculated using the previous output (h_{t-1}), the input vector (x_t) that is translated and dilated (a_w), and the weight matrix specific to the cell (w_w). In addition, the corresponding bias term (b_w) is included in the calculation.

The LSTMW output (h_t) is calculated by the following equations:

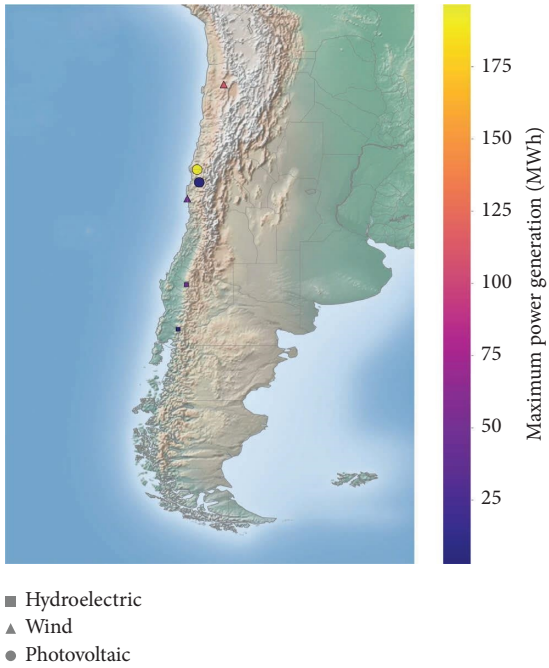
$$u_t = \left[h_{t-1}, x_t\right], \tag{1}$$

$$f_t = \sigma w_f \cdot u_t + b_f, \tag{2}$$

$$i_t = \sigma\left(w_i \cdot u_t + b_i\right), \tag{3}$$

$$o_t = \sigma\left(w_o \cdot u_t + b_o\right), \tag{4}$$

$$a_w = \frac{u_t - \omega_{(\xi)}}{\omega_{(\xi)}}, \tag{5}$$



Station name	Generation	Cases	Mean	Std
HP MAMPIL	Hydroelectric	8 761	12.04	14.07
HP PALMAR	Hydroelectric	8 761	3.12	1.27
PE PUNTA PALMERAS	Wind	8 761	12.23	13.25
PE SIERRA GORDA ESTE	Wind	8 761	39.15	32.93
PFV EL ROMERO	Photovoltaic	8 761	56.84	73.58
PFV TAMBO REAL	Photovoltaic	8 761	0.70	0.93

Figure 1: Location map of the renewable energy stations, including generation type and hourly power statistics. The abbreviations HP, PE, and PFV refer to the types of energy generation: hydroelectric, wind power, and photovoltaic, respectively.

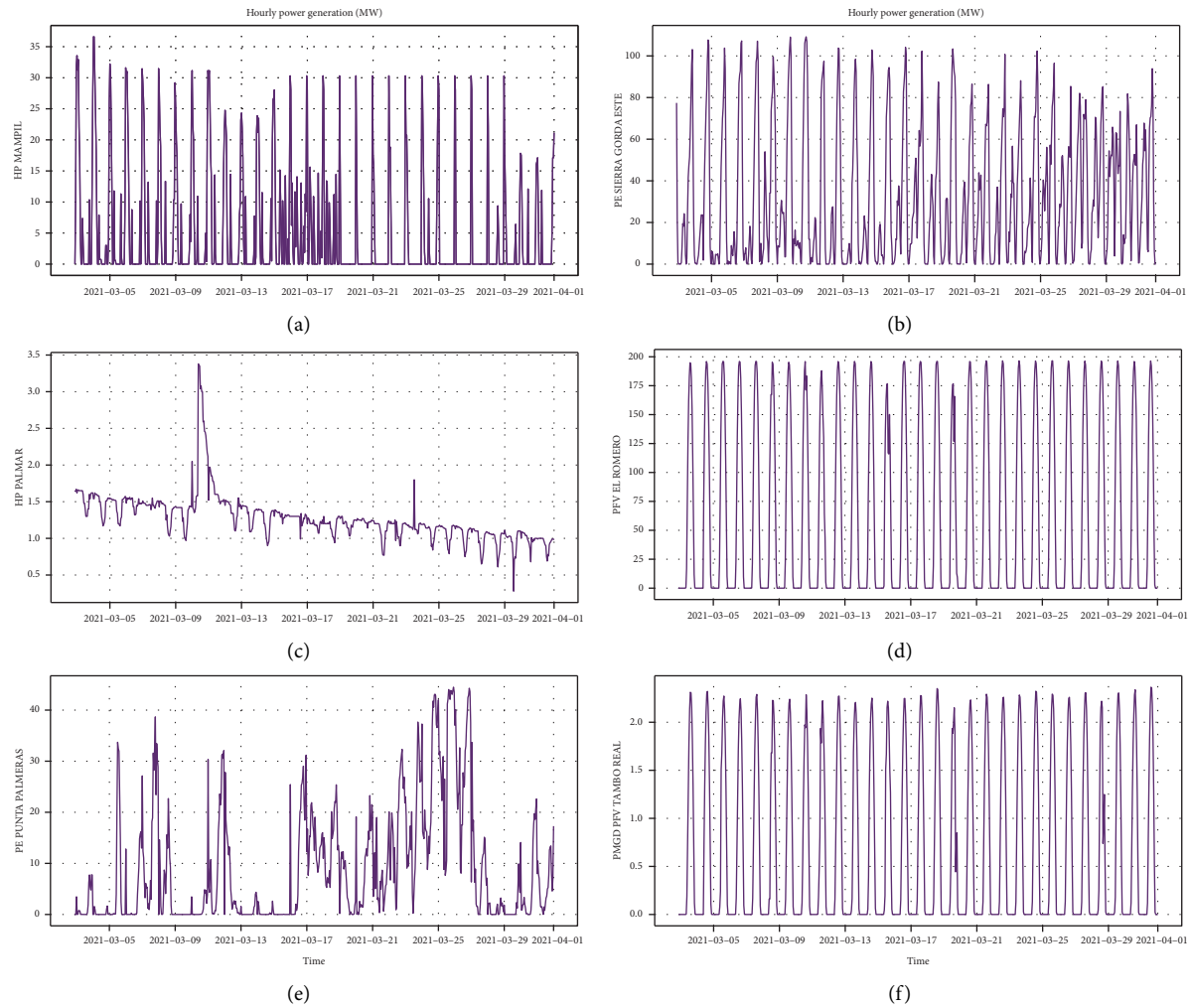


Figure 2: Hourly time series of renewable power generation data samples for each of the six stations: Mampil (a), Sierra Gorda Este (b), Palmar (c), El Romero (d), Punta Palmeras (e), and Tambo Real (f).

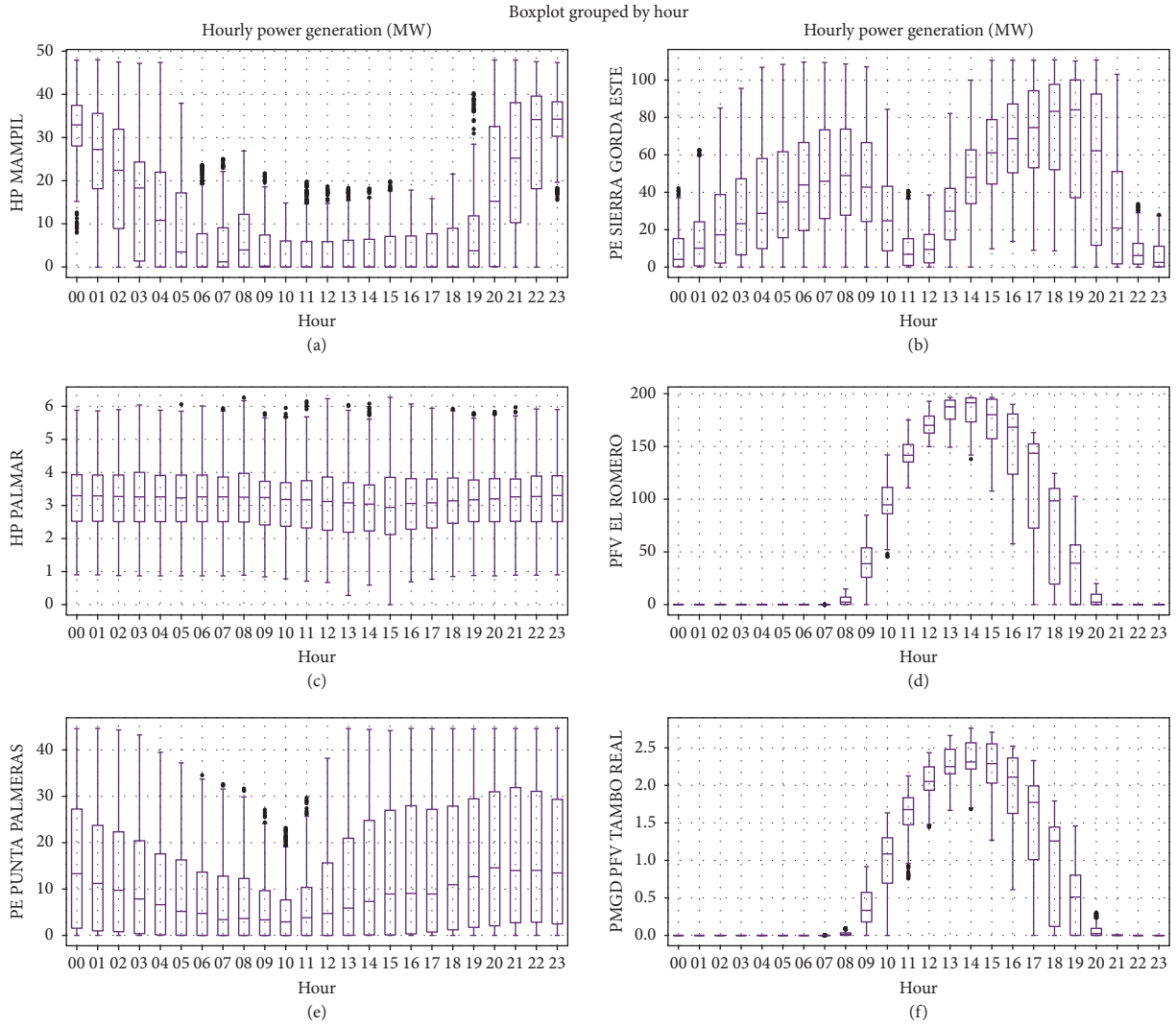


FIGURE 3: Box-plots of the hourly power generation of the six stations: Mampil (a), Sierra Gorda Este (b), Palmar (c), El Romero (d), Punta Palmeras (e), and Tambo Real (f).

$$\tilde{c}_t = \psi(a_w) \cdot w_w + b_w, \tag{6}$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t, \tag{7}$$

$$h_t = o_t \circ c_t, \tag{8}$$

where w_f, w_i, and w_o are weight matrices; b_f, b_i, b_o, and b_w are bias vectors; h_t is the output of LSTMW at the time t; σ is the sigmoid activation function; ψ is the wavelet activation function; and a_w is the current input translated by the parameter ω_ξ and dilated by the parameter ω_ζ. These parameters are adjusted during the training phase. Finally, $\circ$ denotes the Hadamard product.

3.2.2. Translation and Dilation Parameters. Various methods have been proposed for optimized initialization of the wavelet parameters [39]. We implemented the heuristic initialization method; this initialization for the translation and dilation parameters was introduced by Zhang and Benveniste [49] and summarized in Alexandridis and Zapranis's study [37] as follows:

$$\omega_{(\xi)} = 0.5(N_t + M_t), \tag{9}$$

$$\omega_{(\zeta)} = 0.2(M_t - N_t), \tag{10}$$

where M_t and N_t are defined as the maximum and minimum of input u_t:

$$M_t = \max_{t=1,\ldots,T}(u_t), \tag{11}$$

$$N_t = \min_{t=1,\ldots,T}(u_t). \tag{12}$$

3.2.3. Wavelet Activation Function. In the literature, some mother wavelets are suggested [37, 50]:

- The Gaussian derivative is given by

$$\psi(x) = x^2 e^{-(1/2)x^2}. \tag{13}$$

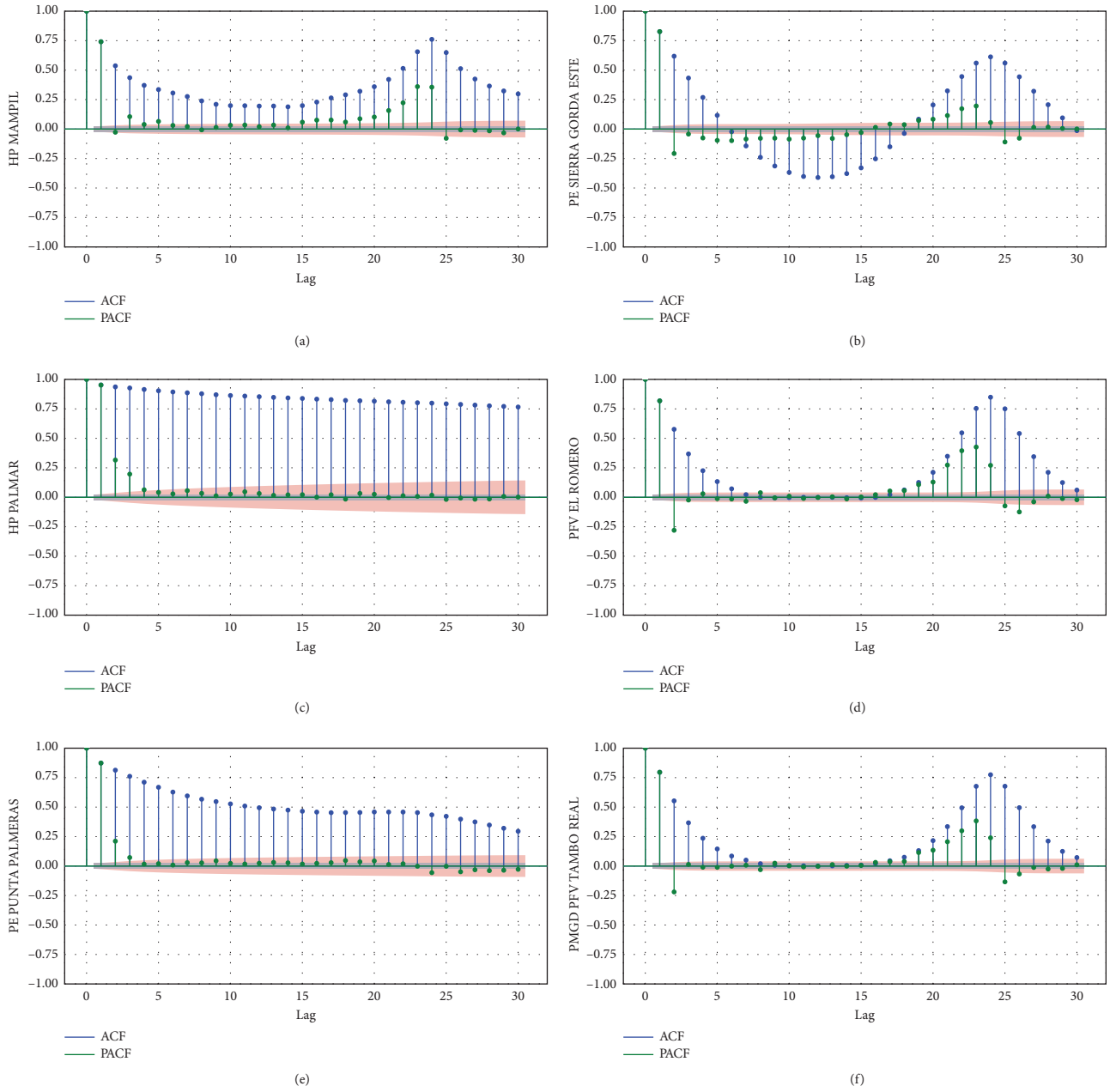


Figure 4: Autocorrelation function (ACF), and the partial ACF (PACF) plot of the hourly power generation for each station: Mampil (a), Sierra Gorda Este (b), Palmar (c), El Romero (d), Punta Palmeras (e), and Tambo Real (f). The blue line represents the ACF, and the green line represents the PACF.

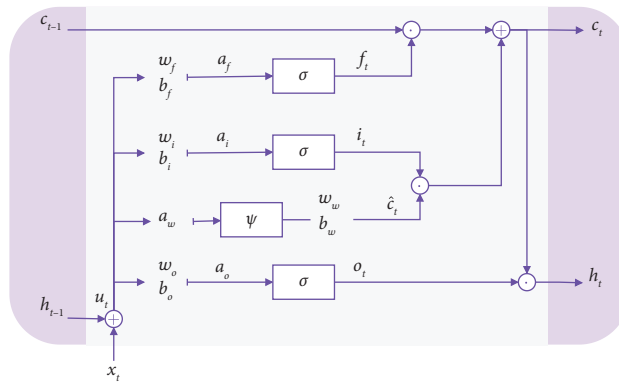


Figure 5: Proposed long short-term memory wavelet neural network scheme (LSTMW).

- The second derivative of the Gaussian, known as the Mexican hat,

$$\psi(x) = 1 - x^2 e^{-(1/2)x^2}. \tag{14}$$

- The Morlet wavelet is obtained from a function that exhibits proportionality to both the cosine function and the Gaussian probability density function. This wavelet is nonorthogonal, possesses an infinite extent of support, and exhibits its peak energy concentration in proximity to the origin within a narrow frequency band [50].

$$\psi(x) = e^{-(1/2)x^2} \cos(5x). \tag{15}$$

- A complex version of the modified Morlet wavelet [51] is

$$\psi(x) = e^{-(1/2)x^2} \cos(\omega_0 x). \tag{16}$$

- The Sinc (Shannon) wavelet originates from a function that shares proportionality with the cosine function. Similar to the previously mentioned wavelets, the Sinc wavelet is nonorthogonal and exhibits infinite support. However, it differs in that its maximum energy concentration occupies a broader frequency spectrum around the origin when compared to the two aforementioned wavelets [50].

$$\psi(x) = \frac{\sin(\pi x^2)}{\pi x^2}. \tag{17}$$

Selection of the mother wavelet depends on the application and is not limited to the foregoing choices [37].

We use the Mexican hat, Shannon, Morlet, and modified Morlet functions as mother wavelets, which have proven to be useful and perform satisfactorily in various applications.

3.2.4. Prediction Enhancement. In addition, as part of an enhancement process, we introduced a methodology (referred to as LSTMWL: LSTMW with luster) to capture and integrate uncertainty into the forecast, aiming to address systemic biases and deviations. This approach aligns with the statistical postprocessing stage employed in improving the prediction accuracy of LSTM models combined with wavelet decomposition, as described in Vivas et al.'s study [52]. In this method, the residuals obtained from the LSTMW adjustment, representing the discrepancy between observed and fitted values, were subjected to modeling through lagged regression models. These models were used to forecast the residuals, which were then incorporated into the energy forecast. By incorporating the uncertainty through this procedure, we aimed to account for systemic biases and deviations that may affect the forecasted values. The previous research has demonstrated favorable outcomes when incorporating separate noise prediction into the forecasted values [52–55].

3.3. Evaluation Criteria. Numerous papers dealing with energy forecasting have assessed their results in terms of the traditional indicators for statistical accuracy: coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). These measures are described as follows:

Performance criteria	equation
Root mean square error (RMSE)	$\mathrm{RMSE} = \sqrt{\frac{1}{m_k}\sum_{k=1}^{m}(t_k - y_k)^2}$,
Mean absolute error (MAE)	$\mathrm{MAE} = \frac{1}{m_k}\sum_{k=1}^{m}\lvert t_k - y_k\rvert$,
Adjusted coefficient of determination $\left(R^2_{\mathrm{adj}}\right)$	$R^2_{\mathrm{adj}} = 1 - \left(1 - R^2\right)\frac{m-1}{m-p-1}$,

(18)

where t_k is the actual value of energy, y_k is the model forecast value, m is the total number of observations, R^2 is the coefficient of determination, and p the number of inputs in the model.

3.4. Comparative Analysis. This methodology was compared to other well-known time-series models. In Table 2, a summary of each of the comparative models implemented is presented.

3.5. Grid Search Cross-Validation. We generated grid-search cross-validation based on 10 iterations; within each iteration, we implemented a parameter grid to search for the optimal model performance, taking three evaluation metrics into consideration. Our tuning involved repeatedly fitting and evaluating the same forecaster with various hyperparameters.

The repeatability of accuracies was evaluated with respect to 10 data groups that were selected using sequences of a fixed size of N-720 by sequence, where N represents the

Table 2: Overview of benchmark forecasting algorithms implemented.

Name	Summary
Multilayer	Perceptron (MLP): A multilayer perceptron or feedforward deep network is made up of layers with connections that convey data from one layer to the next. MLPs are trained using gradient descent to minimize loss functions by adjusting network weights based on gradients [56]
Long short-term memory (LSTM)	The recurrent artificial neural network known as LSTM, introduced by Hochreiter and Schmidhuber [26], is an effective tool for training tasks using time-series data. It excels, especially when nonlinear signal modeling is necessary. Many researchers have successfully applied LSTM for power forecasting, as demonstrated by Sharma and Kakkar [4] and Liu et al. [30]. In our study, we implemented an LSTM architecture based on Lipton et al.'s study [57], as well as a hybrid LSTM scheme similar to the approach described by Li et al. [47]
Wavelet neural network (WNN)	During the training phase of a WNN, the error function is reduced by changing parameters iteratively [39]. Here, we have implemented an efficient heuristic approach for initializing parameters that ensures that the inputs fall inside wavelet function active boundaries. The topology and details of the WNN can be consulted in Alexandridis and Zapranis [39]
Recurrent wavelet neural network (RWNN)	In recurrent WNNs, the output is determined not only by the network's current inputs but also by the network's past outputs or conditions. Recurrent networks, often known as feedback networks, have feedback. There are several kinds of recurrent networks based on the feedback link. The wavelet network input in recurrent wavelet network structures consists of delayed samples of the system output. The topology and details of the RWNN can be consulted in Saleh et al.'s study [58]
Deep wavelet (DWave)	We implemented a long short-term memory (LSTM) network combined with wavelet decomposition. Wavelet decomposition was used for data preprocessing, as suggested by Li et al. [47], and implemented in Vivas, Allende-Cid, and Salas's study [1]. We use the acronym DWave for this hybrid model

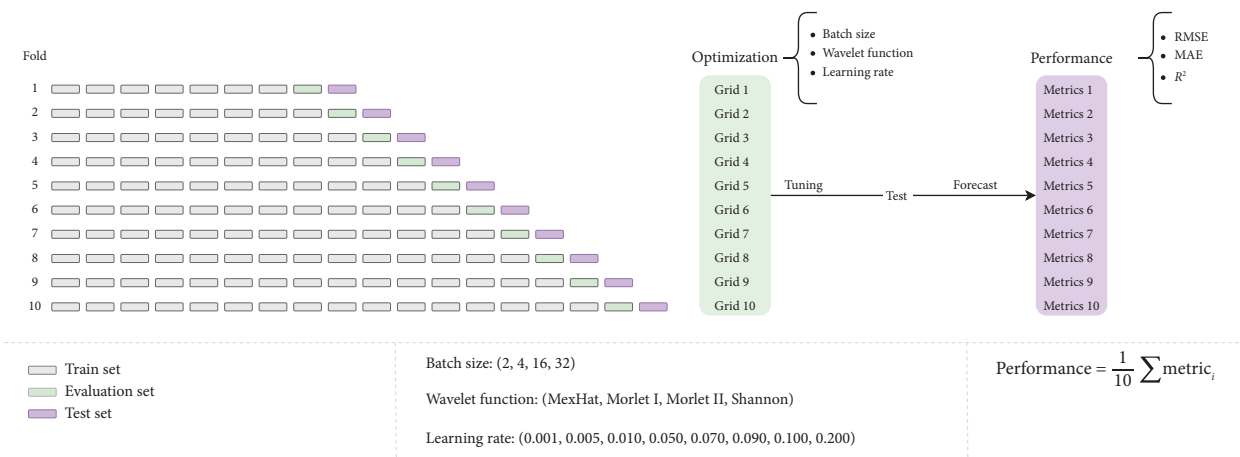


Figure 6: Grid search cross-validation scheme.

number of hours of the previous sequence. In this regard, the network underwent testing using various data subsets to ascertain its capability to consistently estimate power values across different sample groups within a provided database (Figure 6).

4. Results

Figure 7 illustrates that by implementing the grid, it is possible to determine which combination of batch size (Batch), learning rate (Learning), and wavelet activation function produces the best predictions for the validation set. The criterion for selecting the best model was the model with the best R^2 fit; however, coherence is observed in the behavior of the residual metrics for the corresponding configuration of hyperparameters in the grid.

In Figure 8, it is observed that all models tend to exhibit a good level of forecast accuracy for photovoltaic power stations, with the MLP network reaching the lowest level of precision in the majority of cases, and the proposed model LSTMWL being the best predictor in the majority of instances.

In general terms, it can be seen in Figure 8 that the proposed models (highlighted in purple and blue) achieved improvements in forecast accuracy (R^2), and the expected

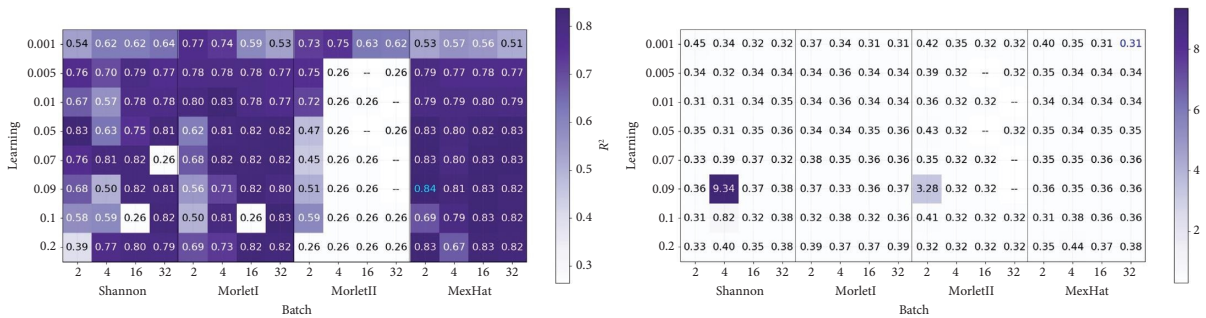
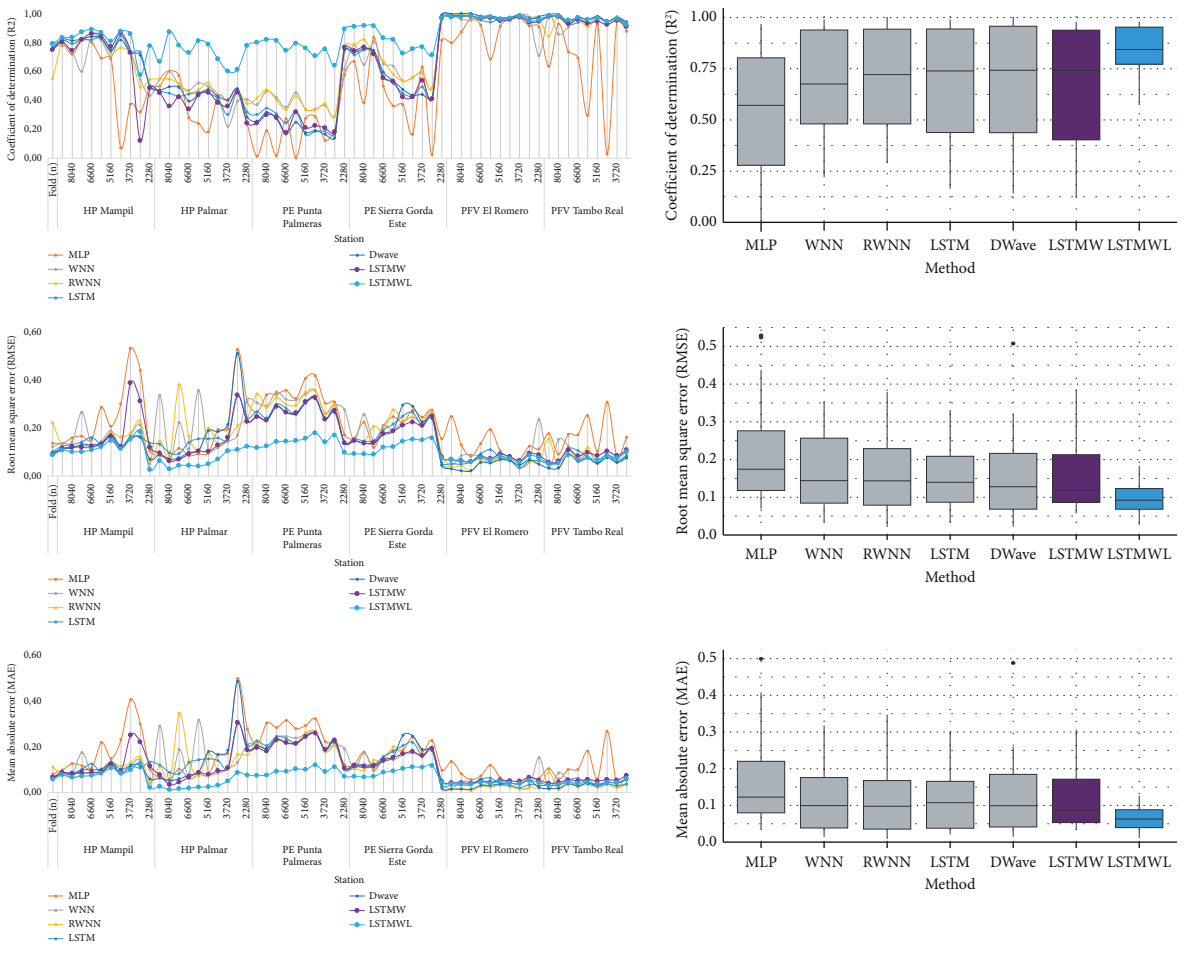


FIGURE 7: Performance of the proposed model (LSTMW) on the validation dataset after applying grid search cross-validation, evaluated using the R^2 metric.



(a) (b)

FIGURE 8: Performance achieved in the forecast (test set): (a) by renewable energy station and validation batch, according to the implemented neural network; (b) based on R^2, MAE, and RMSE according to the implemented model grouping all the outputs of the different stations and cross-validation folds.

dispersion in accuracy is also more controlled. In terms of the behavior of the residues, improvements are also detected in relation to the rest of the models, mostly in the proposal in which a postprocessing phase for the residual values was incorporated (LSTMWL).

In agreement with Figure 8, the results observed in Tables 3, 4, and 5 indicate that the proposed model (LSTMWL) provide better forecasts (higher R^2 and smaller residuals) in the majority of stations; however, when the series is highly seasonal, as in the case of solar

Table 3: Average performance achieved in the forecast (R^2 for the test set) per generating station, according to the implemented neural network.

Station	MLP	WNN	LSTM	Neural networks RWNN	Dwave	LSTMW	LSTMWL
HP Mampil	0.60	0.74	0.79	0.73	0.79	0.73	**0.82**
HP Palmar	0.42	0.46	0.43	0.48	0.46	0.41	**0.73**
PE Punta Palmeras	0.16	0.38	0.25	0.38	0.22	0.24	**0.76**
PE Sierra Gorda Este	0.45	0.63	0.58	0.65	0.59	0.59	**0.82**
PFV El Romero	0.88	0.97	0.97	0.98	**0.98**	0.96	0.97
PFV Tambo Real	0.70	0.92	0.96	0.94	**0.96**	0.94	0.96

Note: The values indicating the best performance are marked in bold.

Table 4: Average performance achieved in the forecast (RMSE for the test set) per generating station, according to the implemented neural network.

Station	MLP	WNN	LSTM	Neural networks RWNN	Dwave	LSTMW	LSTMWL
HP Mampil	0.25	0.16	0.14	0.16	0.13	0.17	**0.12**
HP Palmar	0.16	0.17	0.15	0.16	0.16	0.13	**0.06**
PE Punta Palmeras	0.34	0.31	0.27	0.30	0.26	0.27	**0.14**
PE Sierra Gorda Este	0.21	0.22	0.19	0.21	0.20	0.18	**0.12**
PFV El Romero	0.13	0.06	0.07	0.05	**0.05**	0.08	0.07
PFV Tambo Real	0.16	0.10	0.08	0.08	**0.06**	0.09	0.07

Note: The values indicating the best performance are marked in bold.

Table 5: Average performance achieved in the forecast (MAE for the test set) per generating station, according to the implemented neural network.

Station	MLP	WNN	LSTM	Neural networks RWNN	Dwave	LSTMW	LSTMWL
HP Mampil	0.18	0.11	0.10	0.10	0.10	0.12	**0.08**
HP Palmar	0.14	0.13	0.14	0.13	0.14	0.10	**0.03**
PE Punta Palmeras	0.27	0.22	0.23	0.22	0.22	0.21	**0.09**
PE Sierra Gorda Este	0.17	0.16	0.16	0.15	0.16	0.15	**0.09**
PFV El Romero	0.08	0.03	0.04	**0.02**	0.03	0.05	0.04
PFV Tambo Real	0.11	0.05	0.04	0.04	**0.03**	0.05	0.04

Note: The values indicating the best performance are marked in bold.

energy series, the RWNN model is a more accurate predictor. The LSTMWL model demonstrated superior performance for less seasonal and dispersed series (PE Punta Palmeras).

4.1. Statistical Significance Tests. To have greater precision regarding the differences, various hypothesis tests were implemented. The Kruskal–Wallis test was used to examine the statistical significance of the differences in the performance of the implemented models; this test has been used in a variety of ways in other forecasting studies to examine differences between series [59, 60].

- The Kruskal–Wallis test, created by Kruskal and Wallis [61], is a nonparametric method for determining whether a set of data is derived from the same population. It is an intuitively nonparametric version of ANOVA. In contrast to the parametric ANOVA, the Kruskal–Wallis test does not need the assumptions of the traditional ANOVA. Consequently, it can be used for tasks involving samples with complex distribution characteristics [60]. In addition to the Kruskal–Wallis test, the below-described second nonparametric test was implemented.
- The Friedman test is a nonparametric alternative to the repeated measures ANOVA. It determines if there is a statistically significant difference between the means of three or more groups containing the same subjects. This test has also been used to compare forecasts in other works [62, 63].

Both tests are conducted under the following hypotheses:

- Null hypothesis (H_0): The average accuracy metric of each model is the same
- Alternative hypothesis (H_a): At least one of the average accuracy metric of the models differs from the rest

The p values reported in Tables 6, 7, 8, 9, 10, and 11 are used to assess the statistical significance of differences in performance metrics (RMSE, MAE, and R^2) between the

Table 6: Hypothesis test to compare average R^2 by power plant.

Station	Friedman	p value	Kruskal	p value
HP Mampil	43.987	$1.42E-03$	36.017	$3.94E-02$
HP Palmar	29.151	0.0006112	17.553	0.04073
PE Punta Palmeras	33.016	0.0001327	66.428	0.6743
PE Sierra Gorda Este	53.836	$2.03E-05$	35.258	$5.37E-02$
PFV El Romero	28.434	0.000807	20.736	0.01388
PFV Tambo Real	31.488	0.000244	22.462	0.007523

Table 7: Hypothesis test to compare average *RMSE* by power plant.

Station	Friedman	p value	Kruskal	p value
HP Mampil	42.210	$3.01E-03$	35.041	$5.86E-02$
HP Palmar	34.449	$7.45E-02$	27.037	0.001379
PE Punta Palmeras	49.348	$1.43E-04$	18.615	0.02867
PE Sierra Gorda Este	42.927	$2.22E-03$	33.593	0.0001052
PFV El Romero	27.405	0.001198	16.802	0.05192
PFV Tambo Real	33.171	0.0001247	22.34	0.007861

Table 8: Hypothesis test to compare average MAE by power plant.

Station	Friedman	p value	Kruskal	p value
HP Mampil	43.769	$1.56E-03$	38.662	$1.33E-02$
HP Palmar	30.803	0.00032	20.993	0.01268
PE Punta Palmeras	47.665	$2.95e-07$	23.324	0.005508
PE Sierra Gorda Este	43.332	$1.87E-03$	31.988	0.0002001
PFV El Romero	21.39	0.01103	63.158	0.7079
PFV Tambo Real	26.158	0.001925	10.527	0.3095

Table 9: p values from the hypothesis test of the mean difference in R^2 using the Wilcoxon signed-rank test.

	DWave	LSTM	LSTMW	LSTMWL	MLP	RWNN
LSTM	0.25538	—	—	—	—	—
LSTMW	0.00290	0.05704	—	—	—	—
LSTMWL	**$7.4e-07$**	**$1.9e-08$**	**$1.7e-11$**	—	—	—
MLP	$1.6e-06$	$5.1e-06$	0.00010	**$1.7e-11$**	—	—
RWNN	0.02547	0.00022	$3.8e-06$	**$1.1e-08$**	$1.3e-08$	—
WNN	0.14394	0.04034	0.00487	**$1.8e-09$**	$1.4e-06$	0.54853

Note: The smallest p values are marked in bold.

Table 10: p values from the hypothesis test of the mean difference in RMSE using the Wilcoxon signed-rank test.

	DWave	LSTM	LSTMW	LSTMWL	MLP	RWNN
LSTM	0.00074	—	—	—	—	—
LSTMW	0.05422	0.16977	—	—	—	—
LSTMWL	**$2.7e-05$**	**$7.0e-08$**	**$1.7e-11$**	—	—	—
MLP	$7.0e-09$	$1.4e-07$	$2.1e-10$	**$1.7e-11$**	—	—
RWNN	$6.7e-05$	0.27431	0.13994	**$7.5e-08$**	$3.1e-05$	—
WNN	$8.9e-05$	0.02314	0.01981	**$8.4e-09$**	0.00084	0.73766

Note: The smallest p values are marked in bold.

power generation plants and algorithms tested. A p value less than 0.05 indicates that the difference between two algorithms is statistically significant, while a p value greater than 0.05 suggests no significant difference. The Friedman and Wilcoxon signed-rank tests were applied to determine whether the differences in model performance were

Table 11: p values from the hypothesis test of the mean difference in MAE using the Wilcoxon signed-rank test.

	DWave	LSTM	LSTMW	LSTMWL	MLP	RWNN
LSTM	0.108	—	—	—	—	—
LSTMW	0.189	0.287	—	—	—	—
LSTMWL	**2.1*e* − 06**	**5.3*e* − 07**	**1.7*e* − 11**	—	—	—
MLP	1.6*e* − 07	8.6*e* − 06	8.8*e* − 09	**2.2*e* − 11**	—	—
RWNN	0.229	0.014	0.063	**2.2*e* − 06**	8.3*e* − 07	—
WNN	0.868	0.391	0.429	**9.3*e* − 07**	1.8*e* − 05	0.874

Note: The smallest p values are marked in bold.

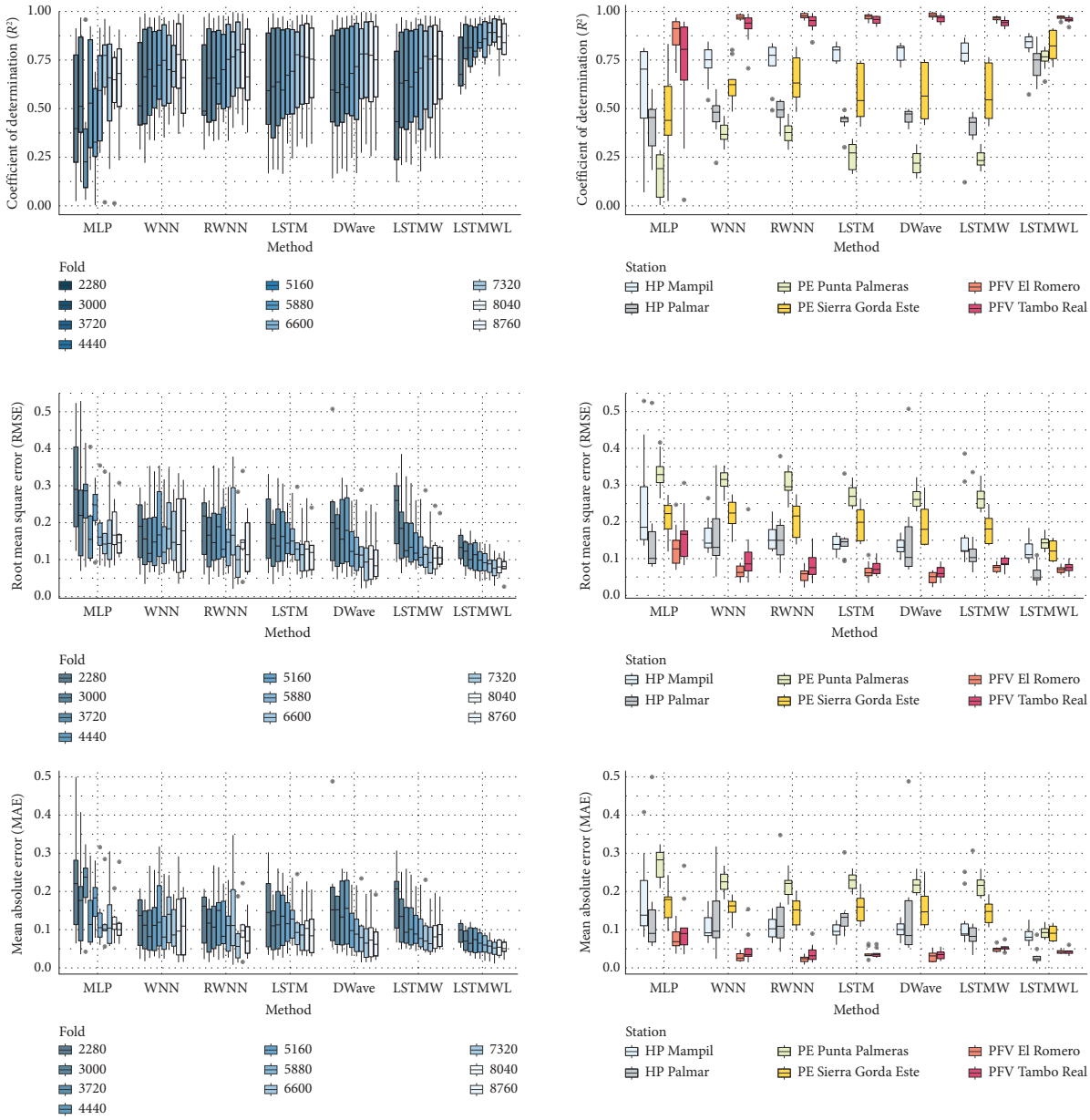


Figure 9: Forecast performance of the implemented models according to the cross-validation fold. The performances of all the stations were consolidated (test set). Forecast performance of the implemented models by renewable energy station. The performances of all the cross-validation folds were consolidated (test set).

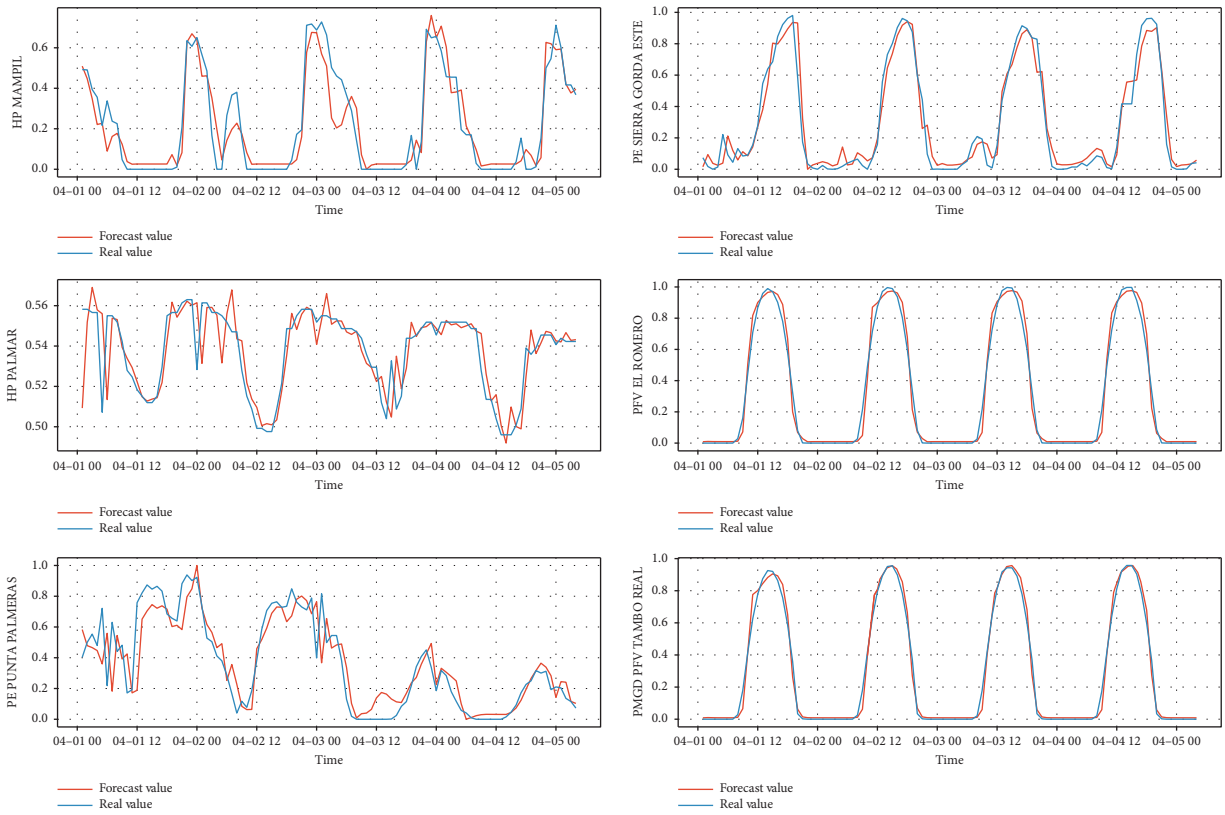


FIGURE 10: Actual and predicted output of the LSTMWL neural network (test set). Corresponding to the cross-validation fold made up of 7320 records.

meaningful, with pairwise comparisons revealing where significant performance differences exist, especially highlighting the superior performance of the proposed LSTMW and LSTMWL models.

In the tests, significant differences are found between at least two groups (Tables 6, 7, and 8). To find out what they are, two-by-two comparisons are made using the Wilcoxon signed-rank test (Tables 9, 10, and 11).

Although both the Friedman and Kruskal–Wallis tests are significant in contrasting the difference in the average performance of the models, pairwise comparisons using the Wilcoxon test (Tables 9, 10, and 11) do not reveal significant differences in all cases.

Remarkably, there are differences between the performance of the forecasts obtained through the proposed models (LSTMW and LSTMWL) and most other models, highlighting the particularity of the model when the post-processing phase is implemented (LSTMWL).

Likewise, various boxplots are shown to illustrate the distributions of the performance metrics. The results of all the stations were consolidated in a single graph according to the fold of the cross-validation by the model. In Figure 9, it can be seen that the neural networks that implement the wavelet analysis (WNN, RWNN, LSTMW, DWave, and LSTMWL) are typically more dispersion-stable than models that only employ traditional activation functions (MLP and LSTM).

It can also be seen in Figure 9 that the precision of the models varies according to the size of the fold; the smallest batches tend to show a greater number of atypical performances. However, it is not a result that applies to all the models. The series is seasonal, so the presence or absence of a fold in the historical record can alter the metrics results, depending on the model's capability to predict that segment.

Figure 10 illustrates the behavior of the series of real values and the values that were predicted by the LSTMWL model for the cross-validation fold made up of 7320 records. The forecast was made for 720 records in each of the stations. In general terms, the model manages to accurately forecast the electrical energy in all stations.

As shown in Figure 11, most models perform well up to 3 steps ahead, but the quality of forecasts deteriorates beyond that, with an increasing gap between predicted and actual values. This quality decline is primarily due to the propagation of errors from earlier prediction stages. The incorporation of exogenous sources of variability could potentially mitigate this issue by improving error management and enhancing forecast accuracy over longer horizons.

Likewise, as shown in Figure 11, the real photovoltaic generation values are zero during periods without sunlight due to the complete dependence of photovoltaic generation on solar availability. This behavior is expected, as there is no energy output during nighttime.

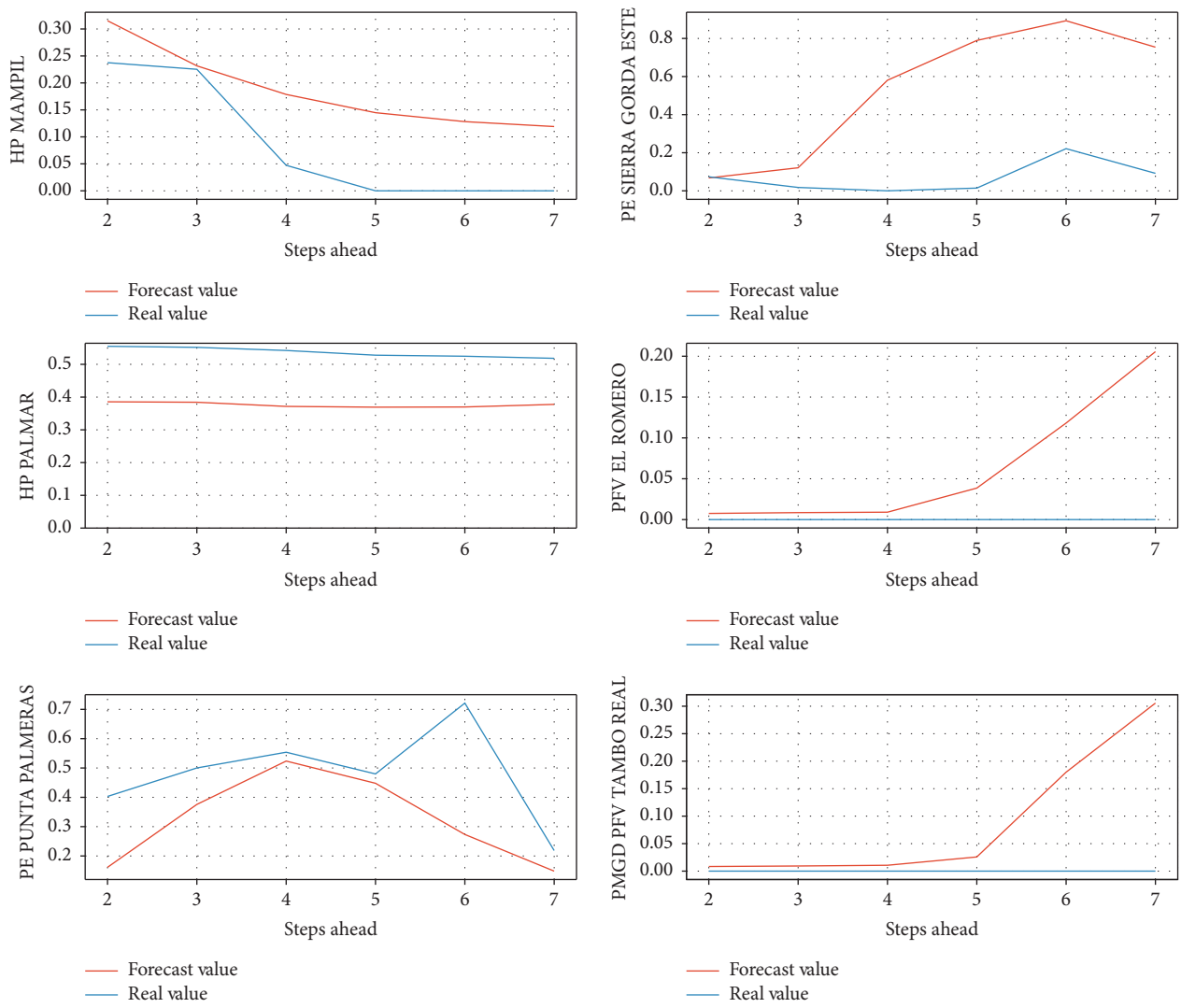


Figure 11: Forecast performance across multiple prediction steps.

The proposed algorithm accurately captures this pattern up to the third forecast step. However, as the forecast horizon extends, predictions may diverge due to the inherent uncertainty associated with long-term forecasting—a common challenge in energy prediction.

5. Conclusions

New models based on wavelet analysis and deep learning were proposed to build predictive energy models in Chile. In one of our proposed models (LSTMWL), uncertainty was incorporated into the forecast to account for systemic biases and deviations in the forecast values. These models have been tested in different case studies, and their performances were compared with those of other powerful models. We found that in most cases, our proposed models perform better.

It was observed that the proposed models perform better in predicting small amplitude, short-duration signal variability, which can often be masked by larger fluctuations. This behavior was particularly evident in the case of the PE Punta Palmeras Station, where the models excelled in capturing subtle variations. However, for predicting large amplitude signals, such as those observed at the PFV El Romero Station, the proposed models showed limitations. In this case, the recurrent WNN (RWNN) outperformed the others, demonstrating that the proposed models may not be as effective for handling larger, more pronounced signal variations. This highlights the need for further refinement of the models to handle a wider range of signal types and amplitudes effectively.

Although the proposed methodology achieved high accuracy levels, future work could be directed in two directions. On the one hand, the methodology can be extended to forecast with more time steps ahead. On the other hand, the influence of external predictors can be incorporated as external sources of local climate variability. The improvement phase can be optimized by implementing other algorithms to model the residuals, such as prediction enhancement with LSTM networks. Likewise, it would be

interesting to evaluate the possibility of including the construction of confidence and prediction intervals to estimate the possible values in which future energy generation will occur with a certain probability.

Data Availability Statement

The data utilized in this study are not publicly accessible. However, summary statistics or additional information can be shared by the authors upon reasonable request.

Disclosure

The funders played no part in shaping the study's design, data collection, analysis, or interpretation, nor did they contribute to the manuscript's writing or the decision to publish the findings.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

We express our appreciation to ANID (National Agency for Research and Development) through the Doctoral Scholarship Program, project number 2021-21211763, which provides funding for the first author's doctoral studies. We also acknowledge the partial financial support for Rodrigo Salas' work through FONDECYT grant number 1221938, ANID-Millennium Science Initiative Program ICN2021-004, ANID ANILLO ATE 220020, and FONDEF IDEA I+D 2021 ID21I10093.

Acknowledgments

We express our appreciation to ANID (National Agency for Research and Development) through the Doctoral Scholarship Program, project number 2021-21211763, which provides funding for the first author's doctoral studies. We also acknowledge the partial financial support for Rodrigo Salas' work through FONDECYT grant number 1221938, ANID-Millennium Science Initiative Program ICN2021-004, ANID ANILLO ATE 220020, and FONDEF IDEA I+D 2021 ID21I10093.

General Statement

Correction added on 22.03.2025, after first publication: Dr Héctor Allende-Cid was designated as corresponding author.

References

[1] E. Vivas, H. Allende-Cid, and R. Salas, "A Systematic Review of Statistical and Machine Learning Methods for Electrical Power Forecasting With Reported Mape Score," *Entropy* 22, no. 12 (2020): 1412.

[2] J. W. Taylor, L. M. De Menezes, and P. E. McSharry, "A Comparison of Univariate Methods for Forecasting Electricity Demand up to a Day Ahead," *International Journal of Forecasting* 22, no. 1 (2006): 1–16.

[3] R. Banos, F. Manzano-Agugliaro, F. Montoya, C. Gil, A. Alcayde, and J. Gómez, "Optimization Methods Applied to Renewable and Sustainable Energy: A Review," *Renewable and Sustainable Energy Reviews* 15, no. 4 (2011): 1753–1766.

[4] A. Sharma and A. Kakkar, "Forecasting Daily Global Solar Irradiance Generation Using Machine Learning," *Renewable and Sustainable Energy Reviews* 82 (2018): 2254–2269.

[5] S. G. Kim, J. Y. Jung, and M. K. Sim, "A Two-Step Approach to Solar Power Generation Prediction Based on Weather Data Using Machine Learning," *Sustainability* 11, no. 5 (2019): 1501.

[6] M. Aghbashlo, W. Peng, M. Tabatabaei, et al., "Machine Learning Technology in Biodiesel Research: A Review," *Progress in Energy and Combustion Science* 85 (2021): 100904.

[7] F. Leite Coelho da Silva, K. da Costa, P. Canas Rodrigues, R. Salas, and J. L. López-Gonzales, "Statistical and Artificial Neural Networks Models for Electricity Consumption Forecasting in the Brazilian Industrial Sector," *Energies* 15, no. 2 (2022): 588.

[8] B. Singh and D. Pozo, "A Guide to Solar Power Forecasting Using Arma Models," in *2019 IEEE PES Innovative Smart Grid Technologies Europe* (ISGT-Europe, 2019), 1–4.

[9] A. Zendehboudi, M. Baseer, and R. Saidur, "Application of Support Vector Machine Models for Forecasting Solar and Wind Energy Resources: A Review," *Journal of Cleaner Production* 199 (2018): 272–285, https://www.sciencedirect.com/science/article/pii/S095965261832153X, https://doi.org/10.1016/j.jclepro.2018.07.164.

[10] W. Liao, B. Bak-Jensen, J. R. Pillai, Z. Yang, and K. Liu, "Short-Term Power Prediction for Renewable Energy Using Hybrid Graph Convolutional Network and Long Short-Term Memory Approach," *Electric Power Systems Research* 211 (2022): 108614, https://www.sciencedirect.com/science/article/pii/S0378779622004795, https://doi.org/10.1016/j.epsr.2022.108614.

[11] H. Yang, P. Jiang, Y. Wang, and H. Li, "A Fuzzy Intelligent Forecasting System Based on Combined Fuzzification Strategy and Improved Optimization Algorithm for Renewable Energy Power Generation," *Applied Energy* 325 (2022): 119849, https://www.sciencedirect.com/science/article/pii/S0306261922011175, https://doi.org/10.1016/j.apenergy.2022.119849.

[12] A. Veloz, R. Salas, H. Allende-Cid, H. Allende, and C. Moraga, "Identification of Lags in Nonlinear Autoregressive Time Series Using a Flexible Fuzzy Model," *Neural Processing Letters* 43 (2016): 641–666.

[13] Y. Morales, M. Querales, H. Rosas, H. Allende-Cid, and R. Salas, "A Self-Identification Neuro-Fuzzy Inference Framework for Modeling Rainfall-Runoff in a Chilean Watershed," *Journal of Hydrology* 594 (2021): 125910.

[14] M. Querales, R. Salas, Y. Morales, H. Allende-Cid, and H. Rosas, "A Stacking Neuro-Fuzzy Framework to Forecast Runoff from Distributed Meteorological Stations," *Applied Soft Computing* 118 (2022): 108535.

[15] W. Liao, B. Bak-Jensen, J. R. Pillai, Y. Wang, and Y. Wang, "A Review of Graph Neural Networks and Their Applications in Power Systems," (2021), https://api.semanticscholar.org/CorpusID:231698364.

[16] E. Arslan Tuncar and B. S¸afak Sağlam & Oral, "A Review of Short-Term Wind Power Generation Forecasting Methods in Recent Technological Trends Energy Reports," (2024), 197–209, https://www.sciencedirect.com/science/article/pii/S2352484724003603, https://doi.org/10.1016/j.egyr.2024.06.006.

[17] T. González Grandón, J. Schwenzer, T. Steens, and J. Breuing, "Electricity Demand Forecasting With Hybrid Classical

Statistical and Machine Learning Algorithms: Case Study of ukraine," *Applied Energy* 355 (2024): 122249, https://www.sciencedirect.com/science/article/pii/S0306261923016136, https://doi.org/10.1016/j.apenergy.2023.122249.

[18] L. Baur, K. Ditschuneit, M. Schambach, C. Kaymakci, T. Wollmann, and A. Sauer, "Explainability and Interpretability in Electric Load Forecasting Using Machine Learning Techniques—A Review," *Energy and AI* 16 (2024): 100358, https://www.sciencedirect.com/science/article/pii/S2666546824000247, https://doi.org/10.1016/j.egyai.2024.100358.

[19] C. Ubal, G. Di-Giorgi, J. E. Contreras-Reyes, and R. Salas, "Predicting the Long-Term Dependencies in Time Series Using Recurrent Artificial Neural Networks," *Machine Learning and Knowledge Extraction* 5, no. 4 (2023): 1340–1358.

[20] F. Wang, Y. Yu, Z. Zhang, J. Li, Z. Zhen, and K. Li, "Wavelet Decomposition and Convolutional Lstm Networks Based Improved Deep Learning Model for Solar Irradiance Forecasting," *Applied Sciences* 8, no. 8 (2018): 1286.

[21] Y. Liu, L. Guan, C. Hou, et al., "Wind Power Short-Term Prediction Based on Lstm and Discrete Wavelet Transform," *Applied Sciences* 9, no. 6 (2019): 1108.

[22] V. Nourani, A. H. Baghanam, J. Adamowski, and O. Kisi, "Applications of Hybrid Wavelet–Artificial Intelligence Models in Hydrology: A Review," *Journal of Hydrology* 514 (2014): 358–377, https://doi.org/10.1016/j.jhydrol.2014.03.057.

[23] E. Vivas, H. Allende-Cid, R. Salas, and L. Bravo, "Polynomial and Wavelet-Type Transfer Function Models to Improve Fisheries' Landing Forecasting With Exogenous Variables," *Entropy* 21, no. 11 (2019): 1082, https://doi.org/10.3390/e21111082.

[24] Q. J. Guo, H. B. Yu, and A. D. Xu, "Modified Morlet Wavelet Neural Networks for Fault Detection," *2005 International Conference on Control and Automation* 2 (2005): 1209–1214.

[25] S. Sivakumar, A. A. Gopalai, K. H. Lim, D. Gouwanda, and S. Chauhan, "Joint Angle Estimation With Wavelet Neural Networks," *Scientific Reports* 11, no. 1 (2021): 1–15.

[26] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation* 9, no. 8 (1997): 1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735.

[27] G. Di-Giorgi, R. Salas, R. Avaria, C. Ubal, H. Rosas, and R. Torres, "Volatility Forecasting Using Deep Recurrent Neural Networks as Garch Models," *Computational Statistics* (2023): 1–27.

[28] J. L. López-Gonzales, R. Salas, D. Velandia, and P. Canas Rodrigues, "Air Quality Prediction Based on Singular Spectrum Analysis and Artificial Neural Networks," *Entropy* 26, no. 12 (2024): 1062.

[29] Y. Jung, J. Jung, B. Kim, and S. Han, "Long Short-Term Memory Recurrent Neural Network for Modeling Temporal Patterns in Long-Term Power Forecasting for Solar Pv Facilities: Case Study of South Korea," *Journal of Cleaner Production* 250 (2020): 119476.

[30] J. Liu, X. Wang, Y. Zhao, B. Dong, K. Lu, and R. Wang, "Heating Load Forecasting for Combined Heat and Power Plants Via Strand-Based Lstm," *IEEE Access* 8 (2020): 33360–33369.

[31] Y. Zeng, J. Chen, N. Jin, X. Jin, and Y. Du, "Air Quality Forecasting With Hybrid Lstm and Extended Stationary Wavelet Transform," *Building and Environment* 213 (2022): 108822.

[32] T. Hong, P. Pinson, Y. Wang, R. Weron, D. Yang, and H. Zareipour, "Energy Forecasting: A Review and Outlook," *IEEE Open Access Journal of Power and Energy* 7 (2020): 376–388.

[33] G. Alkhayat and R. Mehmood, "A Review and Taxonomy of Wind and Solar Energy Forecasting Methods Based on Deep Learning," *Energy and AI* 4 (2021): 100060.

[34] A. Mystakidis, P. Koukaras, N. Tsalikidis, D. Ioannidis, and C. Tjortjis, "Energy Forecasting: A Comprehensive Review of Techniques and Technologies," *Energies* 17, no. 7 (2024): 1662.

[35] A. Shadmani, M. R. Nikoo, A. H. Gandomi, R. Q. Wang, and B. Golparvar, "A Review of Machine Learning and Deep Learning Applications in Wave Energy Forecasting and Wec Optimization," *Energy Strategy Reviews* 49 (2023): 101180.

[36] B. O. Abisoye, Y. Sun, and W. Zenghui, "A Survey of Artificial Intelligence Methods for Renewable Energy Forecasting: Methodologies and Insights," *Renewable Energy Focus* (2023): 100529.

[37] A. K. Alexandridis and A. D. Zapranis, *Wavelet Neural Networks: With Applications in Financial Engineering, Chaos, and Classification* (John Wiley and Sons, 2014).

[38] I. Daubechies, *Ten Lectures on Wavelets* (SIAM, 1992).

[39] A. K. Alexandridis and A. D. Zapranis, "Wavelet Neural Networks: A Practical Guide," *Neural Networks* 42 (2013): 1–27, https://doi.org/10.1016/j.neunet.2013.01.008.

[40] D. Benaouda, F. Murtagh, J. L. Starck, and O. Renaud, "Wavelet-Based Nonlinear Multiscale Decomposition Model for Electricity Load Forecasting," *Neurocomputing* 70, no. 1-3 (2006): 139–154, https://doi.org/10.1016/j.neucom.2006.04.005.

[41] A. Laouafi, M. Mordjaoui, F. Laouafi, and T. E. Boukelia, "Daily Peak Electricity Demand Forecasting Based on an Adaptive Hybrid Two-Stage Methodology," *International Journal of Electrical Power and Energy Systems* 77 (2016): 136–144, https://doi.org/10.1016/j.ijepes.2015.11.046.

[42] D. Benaouda and F. Murtagh, "Electricity Load Forecast Using Neural Network Trained From Wavelet-Transformed Data," in *2006 IEEE International Conference on Engineering of Intelligent Systems* (IEEE, 2006), 1–6.

[43] S. Mujeeb, T. A. Alghamdi, S. Ullah, A. Fatima, N. Javaid, and T. Saba, "Exploiting Deep Learning for Wind Power Forecasting Based on Big Data Analytics," *Applied Sciences* 9, no. 20 (2019): 4417, https://doi.org/10.3390/app9204417.

[44] S. Saroha and S. Aggarwal, "Wind Power Forecasting Using Wavelet Transforms and Neural Networks With Tapped Delay," *CSEE Journal of Power and Energy Systems* 4, no. 2 (2018): 197–209, https://doi.org/10.17775/CSEEJPES.2016.00970.

[45] J. F. Chen, Q. H. Do, T. V. A. Nguyen, and T. T. H. Doan, "Forecasting Monthly Electricity Demands by Wavelet Neuro-Fuzzy System Optimized by Heuristic Algorithms," *Information* 9, no. 3 (2018): 51, https://doi.org/10.3390/info9030051.

[46] I. P. Panapakidis, G. C. Christoforidis, N. Asimopoulos, and A. S. Dagoumas, "Combining Wavelet Transform and Support Vector Regression Model for Day-Ahead Peak Load Forecasting in the Greek Power System," in *2017 IEEE International Conference on Environment and Electrical Engineering and 2017 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)* (IEEE, 2017), 1–6.

[47] P. Li, K. Zhou, X. Lu, and S. Yang, "A Hybrid Deep Learning Model for Short-Term Pv Power Forecasting," *Applied Energy* 259 (2020): 114216, https://doi.org/10.1016/j.apenergy.2019.114216.

[48] A. Laouafi, M. Mordjaoui, A. Medoued, T. E. Boukelia, and A. Ganouche, "Wind Power Forecasting Approach Using Neuro-Fuzzy System Combined With Wavelet Packet Decomposition, Data Preprocessing, and Forecast Combination

Framework," *Wind Engineering* 41, no. 4 (2017): 235–244, https://doi.org/10.1177/0309524X17709726.
[49] Q. Zhang and A. Benveniste, "Wavelet Networks," *IEEE Transactions on Neural Networks* 3, no. 6 (1992): 889–898.
[50] A. Banakar and M. F. Azeem, "Artificial Wavelet Neuro-Fuzzy Model Based on Parallel Wavelet Network and Neural Network," *Soft Computing* 12, no. 8 (2008): 789–808.
[51] P. Addison, J. Watson, and T. Feng, "Low-Oscillation Complex Wavelets," *Journal of Sound and Vibration* 254, no. 4 (2002): 733–762.
[52] E. Vivas, L. B. de Guenni, H. Allende-Cid, and R. Salas, "Deep Lagged-Wavelet for Monthly Rainfall Forecasting in a Tropical Region," *Stochastic Environmental Research and Risk Assessment* 37, no. 3 (2023): 831–848, https://doi.org/10.1007/s00477-022-02323-x.
[53] P. Unnikrishnan and V. Jothiprakash, "Daily Rainfall Forecasting for One Year in a Single Run Using Singular Spectrum Analysis," *Journal of Hydrology* 561 (2018): 609–621, https://doi.org/10.1016/j.jhydrol.2018.04.032.
[54] L. Bravo de Guenni, M. García, A. G. Munoz, et al., "Predicting Monthly Precipitation Along Coastal Ecuador: ENSO and Transfer Function Models," *Theoretical and Applied Climatology* 129, no. 3-4 (2017): 1059–1073, https://doi.org/10.1007/s00704-016-1828-4.
[55] D. McInerney, M. Thyer, D. Kavetski, J. Lerat, and G. Kuczera, "Improving Probabilistic Prediction of Daily Streamflow by Identifying Pareto Optimal Approaches for Modeling Heteroscedastic Residual Errors," *Water Resources Research* 53, no. 3 (2017): 2199–2239, https://doi.org/10.1002/2016WR019168.
[56] A. Graves, "Supervised Sequence Labelling," in *Supervised Sequence Labelling With Recurrent Neural Networks* (Springer, 2012), 5–13.
[57] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A Critical Review of Recurrent Neural Networks for Sequence Learning," *arXiv preprint arXiv:1506.00019* (2015): https://doi.org/10.48550/arXiv.1506.00019.
[58] A. L. Saleh, A. A. Obed, H. H. Qasim, et al., "Wavelet Neural Networks for Speed Control of Bldc Motor," in *Automation and Control IntechOpen* (2020).
[59] E. E. Bas¸akın, Ö. Ekmekcioğlu, H. Çıtakoğlu, and M. Özger, "A New Insight to the Wind Speed Forecasting: Robust Multi-Stage Ensemble Soft Computing Approach Based on Pre-Processing Uncertainty Assessment," *Neural Computing and Applications* 34, no. 1 (2022): 783–812.
[60] S. X. Lv and L. Wang, "Deep Learning Combined Wind Speed Forecasting With Hybrid Time Series Decomposition and Multi-Objective Parameter Optimization," *Applied Energy* 311 (2022): 118674.
[61] W. H. Kruskal and W. A. Wallis, "Use of Ranks in One-Criterion Variance Analysis," *Journal of the American Statistical Association* 47, no. 260 (1952): 583–621, https://www.jstor.org/stable/2280779.
[62] N. Ayoobi, D. Sharifrazi, R. Alizadehsani, et al., "Time Series Forecasting of New Cases and New Deaths Rate for Covid-19 Using Deep Learning Methods," *Results in Physics* 27 (2021): 104495.
[63] S. Das, S. Mishra, and M. Senapati, "Improving Time Series Forecasting Using Elephant Herd Optimization With Feature Selection Methods," *Journal of Management Analytics* 8, no. 1 (2021): 113–133.