Contents list available at CBIORE journal website

International Journal of Renewable Energy Development

Journal homepage: https://ijred.cbiore.id

Research Article

Hybrid WRF–SARIMA model to improve day-ahead wind speed forecast accuracy

Maritza Bernabe[a], Erasmo Cadenas[a*], Erika López-Espinoza[b], Rafael Campos-Amezcua[c]

[a]*Faculty of Mechanical Engineering, Universidad Michoacana de San Nicolás de Hidalgo, Michoacán, México*

[b]*Institute of Atmospheric Sciences and Climate Change, National Autonomous University of Mexico, México*

[c]*National Center for Technological Research and Development, National Technological Institute of Mexico, Morelos, México*

Abstract. Accurate wind speed forecasts are critical for integrating wind energy into power grids, reducing imbalance costs in electricity markets, and optimizing wind farm operations. Day-ahead forecasts are typically generated using numerical weather prediction (NWP) models. This work proposes a hybrid model for 24-hour wind speed forecasting, which combines the Weather Research and Forecasting (WRF) model with the Seasonal Autoregressive Integrated Moving Average (SARIMA) model. The proposed model improves the accuracy of the WRF wind speed forecast through the SARIMA technique by identifying significant autocorrelations in the forecast errors. The study was conducted in La Ventosa, Mexico, a region with significant development in the wind power sector. Wind speed data measured at heights of 17.5 m and 40 m were used during periods of low and high wind speeds. The model's performance was evaluated using the metrics mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE). The results showed that the hybrid WRF-SARIMA model outperformed the WRF model. Forecast errors for MAE were reduced between 29% and 45%, for MSE between 40% and 67%, and for RSME between 22% and 43%. The WRF-SARIMA leverages the benefits of physical NWP models while incorporating the interpretability and reduced computational cost of traditional statistical models. In this way, the proposed model improves wind speed forecast accuracy, especially in the operational contexts of wind energy management.

Keywords: Wind energy, WRF, SARIMA, Wind speed forecast, Day-ahead.

Received: 4th Sept 2025; Revised: 17th Oct 2025; Accepted: 28th Nov 2025; Available online: 10th Dec 2025

1. Introduction

Throughout history, the availability of energy has constrained technological development and human activities. Economic and technological growth lead to increased energy consumption (Hagens, 2020). The current energy demand, along with the adverse effects of fossil fuels, have contributed to the growth of wind energy. In 2024, 117 GW of new wind power capacity was connected to power grids globally, bringing the total installed wind capacity to 1136 GW (Zhao, 2025). In Mexico, wind power generation in 2023 reached 20,700 GWh, accounting for 5.97% of total electricity generation, nearly double the 2017 figure of 10,456 GWh (Secretaria de Energia [SENER], 2024).

Power system operators must maintain a balance between energy demand and generation. The variable nature of wind poses a challenge for operators, particularly as the share of wind power in the grid increases. Wind speed forecast has proven to be a valuable and reliable tool for addressing this issue.

Wind speed forecasts cover various time horizons: 1) forecast of seconds or minutes for unit maintenance and control, 2) 1 to 6 hours for the operation of small power systems, 3) 1 to 72 hours for the operation of interconnected power systems, and 4) 1 to 7 days for maintenance planning (El-Fouly *et al.,* 2008). The 24-hour forecasts are fundamental due to their connection to electricity markets and their support for decision-making in power systems.

Wind speed forecasting models can be divided into two broad categories: statistical methods, which analyze historical data to generate forecasts, and physical methods, which derive forecasts from numerical weather prediction models. In recent years, the literature has shown a predominance of hybrid models over single models, indicating that combining multiple forecasting techniques can significantly improve accuracy (Makridakis, 1989). Hybrid models encompass various combinations of different statistical methods, as well as combinations of a physical model with statistical techniques. Statistical forecasting models include autoregressive moving average models (Aasim *et al.,* 2019; Tian *et al.,* 2020; Zhang *et al.,* 2020; Hu *et al.,* 2021), neural networks (Sun *et al.,* 2022; Zhang *et al.,* 2020), long short-term memory (Lawal *et al.,* 2021; Tyass *et al., 2023*), and deep learning (Zhang *et al.,* 2024). In addition, techniques such as signal decomposition (Liu *et al.,* 2021; Moreno *et al.,* 2021; Wang *et al.,* 2021; Li *et al.,* 2022), error correction (Ding *et al.,* 2022; Duan *et al.,* 2021; Wang *et al.,* 2018), and optimization (Liu *et al.,* 2020; Liu *et al.,* 2024) are considered. Statistical models typically focus on the short and very short-term forecasts. However, they can also release the 24-hour wind speed forecast, for example, using models based on deep learning (Moreno *et al.,* 2024; Hong *et al.,* 2024; Thu *et al.,* 2023) or traditional statistical methods (Liu *et al.,* 2021; Costa *et al.,* 2021). Regarding the SARIMA model, Liu *et al.* (2021)

* Corresponding author
Email: ecadenas@umich.mx (E. Cadenas)

https://doi.org/10.61435/ijred.2026.61746
ISSN: 2252-4940/© 2026.The Author(s). Published by CBIORE

demonstrated that a 24-hour SARIMA model could outperform the long short-term memory and gated recurrent unit models. These studies indicate that statistical models can be effective for 24-hour forecasts.

Day-ahead forecasts, important for electricity markets, are primarily produced by physical models, also known as numerical weather prediction (NWP) models, which numerically solve the equations for the conservation of mass, energy, and momentum that describe atmospheric processes. Different physical models have been used for wind speed forecasting, including the European Centre for Medium-Range Weather Forecasts (ECMWF) (Zhang *et al.,* 2023; Yang *et al.,* 2023), High-Resolution Rapid Refresh (HRRR) (Myers *et al.,* 2024), China Meteorological Administration mesoscale model (CMA-MESO) (Liu *et al.,* 2024), and the most widely used, Weather Research and Forecasting (WRF) model (Zhao *et al.,* 2021).

The WRF model is a limited-area model that utilizes initial and boundary conditions from global models. The model's physical parameterizations can be adjusted to accurately represent the regional characteristics of the simulation domain. For wind speed simulation at low altitudes, the Planetary Boundary Layer (PBL) parameter has proven to be the most sensitive (Jacondino *et al.,* 2021; Mi *et al.,* 2023). Various PBL schemes have been tested to enhance wind speed forecast (Yang *et al.,* 2023; Dzebre *et al.,* 2020). NWP models have limitations; for instance, wind speeds derived from numerical simulations are particularly biased close to the surface. Additionally, truncations in the integration processes, simplifications of atmospheric processes, and inaccuracies in initial and boundary conditions, contribute to greater forecast errors (Al-Yahyai *et al.,* 2010). To overcome this problem, hybrid models that use statistical methods to correct forecast errors have been proposed. Chen *et al.* (2019) propose a linear regression method to correct the forecast bias of a WRF ensemble. Tsai *et al.* (2021) use the decaying average algorithm to adjust a wind speed forecast made with a WRF ensemble. Xu *et al.* (2021) employ a long short-term memory model to correct wind speed forecast errors obtained with a WRF model. Liu *et al.* (2023) and Zhao *et al.* (2024) use temporal convolutional networks and long short-term memory models to reduce forecast errors from a WRF model. Wang *et al.* (2019) apply five models: linear regression, support vector machine, back-propagation neural network, random forest, and radial basis function neural network as a strategy to correct the wind speed forecast errors of an NWP model. These investigations demonstrate that using statistical models as an error correction strategy effectively improves the wind speed forecasts of physical models. However, many of these approaches lack interpretability and require extensive training data, and do not prioritize forecast error analysis as a fundamental component of the error correction process. It is not sufficient to provide the statistical model with enough data to obtain accurate predictions; it is necessary to evaluate whether the data can be interpreted and predicted by the model.

In contrast to machine learning-based models, SARIMA models do not require large amounts of data to train the models, achieving comparable improvements with reduced computational complexity. A SARIMA-based approach provides an explicit, robust methodology for analyzing trends and seasonal components of data, an important aspect in the energy sector where more interpretability is required. They possess a solid mathematical foundation and, in the case of univariate small data sets, they prevail over the machine learning ones (Kontopoulou *et al.,* 2023), even for 24-hour forecasts (Liu *et al.,* 2021)

Every approach has different strengths and weaknesses depending on the specific application and the characteristics of the dataset (Iaousse *et al.,* 2023). The advantages of SARIMA models are: 1) smaller data sets, 2) interpretability, 3) solid mathematical foundation, and 4) lower computational cost. The disadvantage is the model's linearity, but, as has been found, this does not limit its ability to predict complex and oscillatory patterns.

This study proposes a WRF-SARIMA model for 24-hour-ahead wind speed forecasts. The main contribution is the implementation of an error-correction strategy using an interpretable, automatable approach, validated under real-world conditions. Initially, the focus is on identifying sequential dependencies or remaining patterns in the forecast errors generated from the on-site measurements and the WRF model. Due to the characteristics of the datasets available for this study, it is concluded that the SARIMA technique is the best option for modeling the identified patterns in the forecast errors. Time series with seasonal and trend components can be effectively analyzed using SARIMA models. These models are generally faster and easier to implement than more complex alternatives, such as machine learning-based approaches. The SARIMA models are used to predict forecast errors and incorporate them into the final forecast.

To assess the accuracy of the proposed WRF-SARIMA model, its performance was compared with that of the standalone WRF model using mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE) as evaluation metrics.

The remainder of this paper is organized as follows: Section 2 presents the available data and details of the WRF simulations. Section 3 describes the methodology of the proposed model. Section 4 covers the results and discussions, while Section 5 presents the conclusions.

2. Available data

2.1 Study area and in-situ measurement data

This study was conducted in La Ventosa, Oaxaca (16°32'40.8"N, 94°57'09.0"W), an open terrain with a high level of turbulence (Lopez, 2018), which is located in southeastern Mexico on the Isthmus of Tehuantepec (Figure 1). This region has shown significant development in the wind power sector due to strong winds. These winds, known as *nortes* or *tehuanos*, result from atmospheric conditions and local topography. The pressure difference between the Gulf of Mexico and the Gulf of Tehuantepec generates strong northerly winds that flow through a mountain gap (Romero, 2003).

Wind speed data (Figure 2) were collected at heights of 17.5 m and 40 m above the ground using a Gill WindMaster 3D ultrasonic anemometer mounted on a lattice mast, with measurements taken at a frequency of 1 Hz. The original

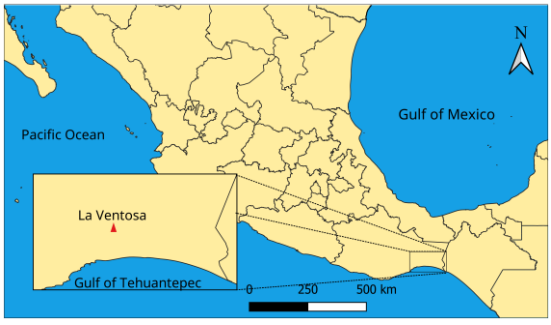

Fig 1. Location of La Ventosa, Oaxaca, Mexico. A region with a high level of turbulence and strong winds.

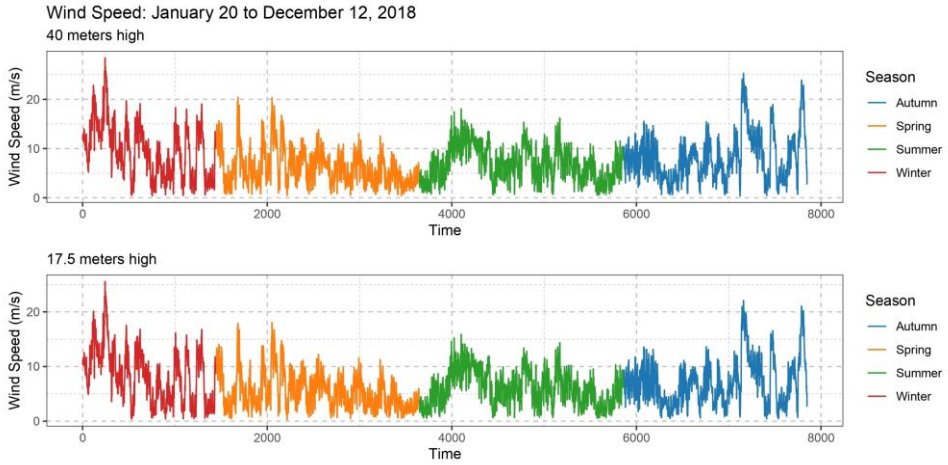

Fig 2. Wind speed time series measured at heights of 17.5 m and 40 m, in La Ventosa, from January 20 to December 12, 2018.

Table 1
Average wind speed in La Ventosa

Season	Wind speed at 40 m (m/s)	Wind speed at 17.5 m (m/s)
Spring	6.34	5.48
Summer	7.10	6.16
Autumm	8.29	7.11
Winter	9.67	8.46
Mean	7.66	6.64
Maximum	28.41	25.54
Standard deviation	4.43	3.98

measurement period spanned from January 20 to December 12, 2018, with a temporal resolution of 1 second. For this work, the data were averaged hourly, resulting in 7848 values.

Maximum wind speed was observed during the autumn and winter seasons. Table 1 shows the seasonal and annual average, the maximum value, and the standard deviation of wind speed, estimated from the measured data. The estimated average wind speed in winter was 9.67 m/s at the height of 40 m and 8.46 m/s at the height of 17.5 m. In autumn, the average was 8.29 m/s at the height of 40 m and 7.11 m/s at the height of 17.5 m. The annual average wind speed was 7.66 m/s at the height of 40 m and 6.64 m/s at the height of 17.5 m. Figure 3 shows that a Weibull distribution model fits the measured wind speed data at heights of 17.5 m and 40 m.

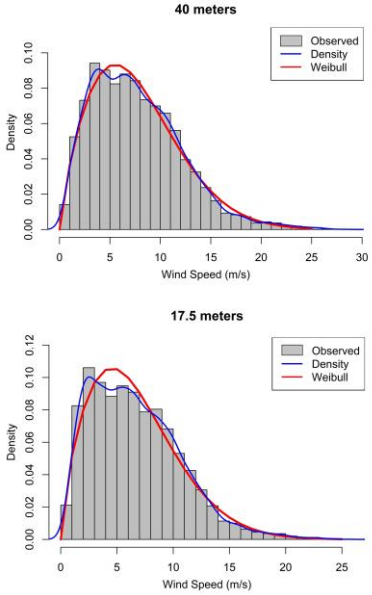

Fig 3. Probability distribution of measured wind speed in La Ventosa, showing the fit to a Weibull distribution for both 17.5 m and 40 m heights.

2.2 WRF simulations

The WRF model is a system developed for research and numerical weather predictions. It has two modules: the WRF Pre-processing System (WPS) and the Advanced Research WRF (ARW). The ARW core uses a non-hydrostatic approximation with an Arakawa staggered C-grid and vertical coordinates that follow the terrain topography near the surface. Numerical simulations were conducted to forecast wind speed data for June 2018 and for November 13 to December 12, 2018. The outputs utilized in this study were provided by the Ocean-Atmosphere Interaction Group (Grupo Interacción Océano-Atmósfera [IOA], 2025) of the Institute of Atmospheric Sciences and Climate Change (Instituto de Ciencias de la Atmósfera y Cambio Climático [ICAyCC], 2025) at the National Autonomous University of Mexico. The IOA group has investigated the WRF model's sensitivity to various initial and boundary conditions (Jurado, 2017) and has incorporated accurate land-use and vegetation data to enhance weather forecast accuracy (Rivera-Martínez, 2018; López-Espinoza *et al.*, 2020).

The WRF simulations were conducted using version 3.9, with a domain covering the entire Mexican territory, ranging from 4.12° N to 38.42° N and from 74.87° W to 123.36° W, with a horizontal resolution of 15 km, 338 x 262 grid points, 50 vertical levels, and an hourly temporal resolution. The initial and boundary conditions come from the Global Forecast System Model (GFS) database at a horizontal resolution of 0.25°, beginning at 00:00 Universal Time Coordinated (UTC). Table 2 shows the physical parameterization schemes used by the WRF model. The forecast is updated daily.

The first six hours of simulation with the WRF model were discarded due to spin-up, while retaining the subsequent 24 hours to capture a diurnal cycle (local time) of the wind speed. Figure 4 shows the bilinear interpolation performed to obtain the wind speed values at the site of interest, using the four nearest WRF grid points ws_{ij} to the measurement site ws.

Table 2
WRF physics parameterization schemes

	Physics options
Microphysics Scheme	WRF Single–moment 3–class Scheme
Planetary Boundary Layer	Yonsei University Scheme (YSU)
Cumulus Parameterization	Kain–Fritsch Scheme
Shortwave radiation	Dudhia scheme
Longwave radiation	RRTM scheme
Land Surface	Unified Noah Land Surface Model

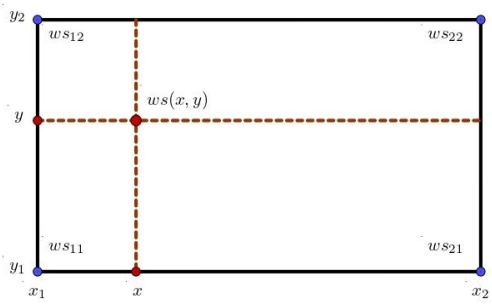

Fig 4. Schematic of the bilinear interpolation used to derive the WRF wind speed at the exact mast location from the four nearest grid points.

$$ws(x,y) = \frac{1}{(x_2-x_1)(y_2-y_1)}[x_2-x, x-x_1]\begin{bmatrix} ws_{11} & ws_{12} \\ ws_{21} & ws_{22} \end{bmatrix}\begin{bmatrix} y_2-y \\ y-y_1 \end{bmatrix} \quad (1)$$

where $ws_{ij} = ws(x_i, y_j)$.

3. Methodology

3.1 SARIMA model

The ARMA models combine autoregressive (AR) and moving average (MA) schemes. They are widely used for wind speed forecasting due to their robustness and easy application. A time series Y_t is modeled as a combination of past values Y_{t-k} and past errors e_{t-k}

$$Y_t = c + \phi_1 Y_{t-1} + \cdots + \phi_k Y_{t-k} + e_t - \theta_1 e_{t-1} - \cdots - \theta_k e_{t-k} \quad (2)$$

The SARIMA models are employed when the time series exhibits seasonal patterns. Non-seasonal variables are represented by p, d, q, and seasonal variables by P, D, Q. The general expression for the $SARIMA(p,d,q)(P,D,Q)_s$ model is given as:

$$(1-\phi_1 B - \& \ldots - \phi_p B^p)(1-\Phi_1 B^s - \cdots - \Phi_P B^{Ps})(1-B)^d(1-B^s)^D Y_t = (1-\theta_1 B - \cdots - \theta_q B^q)(1-\Theta_1 B^s - \cdots - \Theta_Q B^{Qs})e_t \quad (3)$$

where ϕ_i = ith non-seasonal autoregressive parameter,
θ_j = jth non-seasonal moving average parameter,
Φ_m = mth seasonal autoregressive parameter,
Θ_n = nth seasonal moving average parameter,
d = number of non-seasonal differences,
D = number of seasonal differences,
e_t = the error term at time t,
s = number of periods per season.

such that $BY_t = Y_{t-1}$ and $Be_t = e_{t-1}$.

The simple exponential smoothing (SES) model prediction equation is mathematically equivalent to an *ARIMA(0,1,1)* model without the constant term. Similarly, Holt's forecast method is comparable to an *ARIMA(0,2,2)* model. George Box and Gwilym Jenkins developed the most popular modeling strategy, the Box-Jenkins methodology (Makridakis, 2008).

3.2 Hybrid WRF-SARIMA model

This study develops a multi-step wind speed forecasting model that combines the WRF wind speed forecast with SARIMA models. The proposed methodology, shown in Figure 5, consists of five stages: collecting wind speed time series, analyzing forecasting errors, modeling and forecasting using SARIMA, combining the wind speed forecasts obtained with the WRF and SARIMA models, and evaluating the final proposed model. Each stage is elaborated in detail in the following subsections. Figure 6 illustrates the flowchart for the proposed

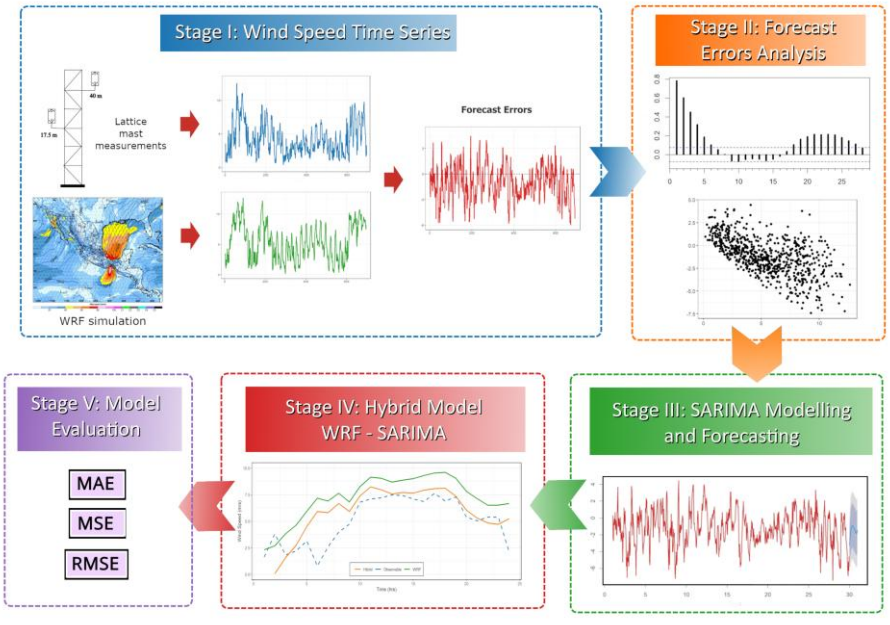

Fig 5. Diagram of the proposed hybrid WRF-SARIMA methodology, showing the five main stages: data acquisition, error analysis, modeling, forecast combination, and model evaluation.

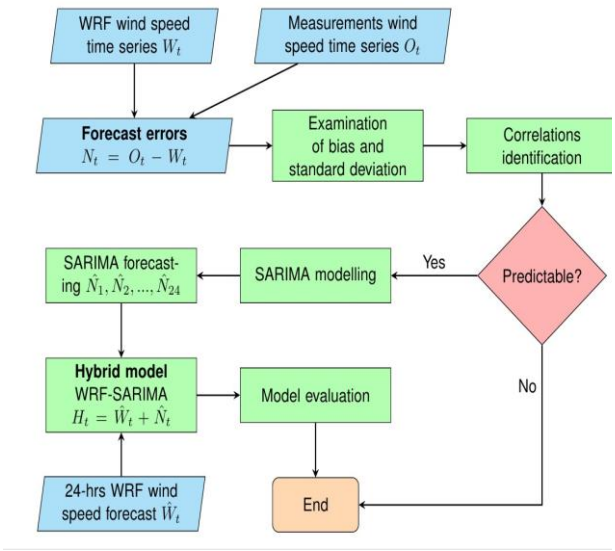

Fig 6. Detailed flowchart of the WRF-SARIMA model, illustrating the sequence from data input to the final forecast.

WRF-SARIMA model. A consideration was selecting steps that could be automated to enable effective real-world deployment.

3.2.1 Stage I. Wind speed time series

This stage includes the pre-processing steps that ensure that the mast-measured wind speed time series, denoted by O_t, and the WRF wind speed time series, denoted by W_t, have the same temporal resolution covering the same period of time and location. The forecast error time series, denoted by N_t, is defined as the difference between time series O_t and W_t:

$$N_t = O_t - W_t \tag{4}$$

Consequently, the time series N_t has the same temporal resolution covering an identical period and location as the time series O_t and W_t.

3.2.2 Stage II. Forecast errors analysis

It includes the statistical tests performed on the forecast error time series N_t to determine whether it exhibits serial correlation and, therefore, whether the time series is predictable. Analyzing the correlation properties of the time series N_t is crucial for determining the feasibility of predicting forecast errors. The *autocorrelation function* (ACF) quantifies the relationship between values in a time series and helps to identify cyclical and seasonal patterns. The ACF is the set of autocorrelation coefficients r_k at lag k.

$$\boldsymbol{r_k} = \frac{\sum_{i=k+1}^{n}(N_i - \bar{N})(N_{i-k} - \bar{N})}{\sum_{i=1}^{n}(N_i - \bar{N})^2} \tag{5}$$

If autocorrelation values are found outside the range of critical values, it indicates that the time series N_t exhibits linear correlations.

Additional tests, such as the von Neumann ratio of the mean-square-successive difference test, the Ljung-Box test, and heteroskedasticity graphs, are used to support or reject the hypothesis of serial correlation in the time series.

3.2.3 Stage III. SARIMA modeling and forecasting

Once the series N_t has been found to be predictable based on previous tests, the time series is modeled and forecasted using SARIMA models, a robust technique with a clear methodology and an explicit explanation of how the models operate.

The algorithmic process proposed in this study is a variation of that proposed by Box-Jenkins (Makridakis *et al.*, 2008), incorporating additional conditions. The methodology is divided into four steps: Stationarity, Examination of ACF and PACF, Model selection, and Forecasting. The values predicted by the $SARIMA(p,d,q)(P,D,Q)_{24}$ model are saved for use in the next stage.

3.2.4 Stage IV. Combination of forecasts

$\widehat{W}_t$ denotes the 24-hour WRF wind speed forecast. Then, $\widehat{N}_t$ and $\widehat{W}_t$ are both time series corresponding to hourly 24-hour forecasts, and each consists of 24 elements. These values serve as inputs to the new hybrid model H_t, defined as:

$$H_t = \widehat{W}_t + \widehat{N}_t \tag{6}$$

The time series H_t represents the 24-step ahead wind speed forecast generated by the hybrid WRF-SARIMA model. This series contains 24 elements and combines the WRF model's wind speed forecast with the SARIMA model's predicted forecast error. It is important to note that this methodology is fully automatable, making its operational implementation much easier.

3.2.5 Stage IV. Model evaluation

The performance of the proposed model is evaluated using the metrics of bias, mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE).
The time series $\hat{O}_t$ represents in-situ measurements corresponding to the same time period as $\widehat{W}_t$ and $\widehat{N}_t$.

Let $e_t = \hat{O}_t - \widehat{W}_t$ and $\varepsilon_t = \hat{O}_t - H_t$ then:

$$e_t = \varepsilon_t + \widehat{N}_t \tag{7}$$

If ε_t and $\widehat{N}_t$ have the same sign ($\varepsilon_t \widehat{N}_t \geq 0$), then:

$$|\varepsilon_t| \leq |e_t| \tag{8}$$

That is, the residuals ε_t of the hybrid model are smaller in magnitude than those e_t of the WRF model under certain conditions. The performance of the hybrid WRF-SARIMA model is compared to that of the WRF model and two other hybrid models: WRF-Simple Exponential Smoothing (WRF-SES) and WRF-Holt's Method (WRF-Holt).

3.3 Evaluation criteria

The accuracy measures used to evaluate the model's performance include bias, mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE).

$$BIAS = \frac{1}{n}\sum_{t=1}^{n} e_t \tag{9}$$

Mean absolute error (MAE)

$$MAE = \frac{1}{n}\sum_{t=1}^{n} |e_t| \tag{10}$$

Mean squared error (MSE)

$$MSE = \frac{1}{n}\sum_{t=1}^{n} e_t^2 \tag{11}$$

Root mean squared error (RMSE)

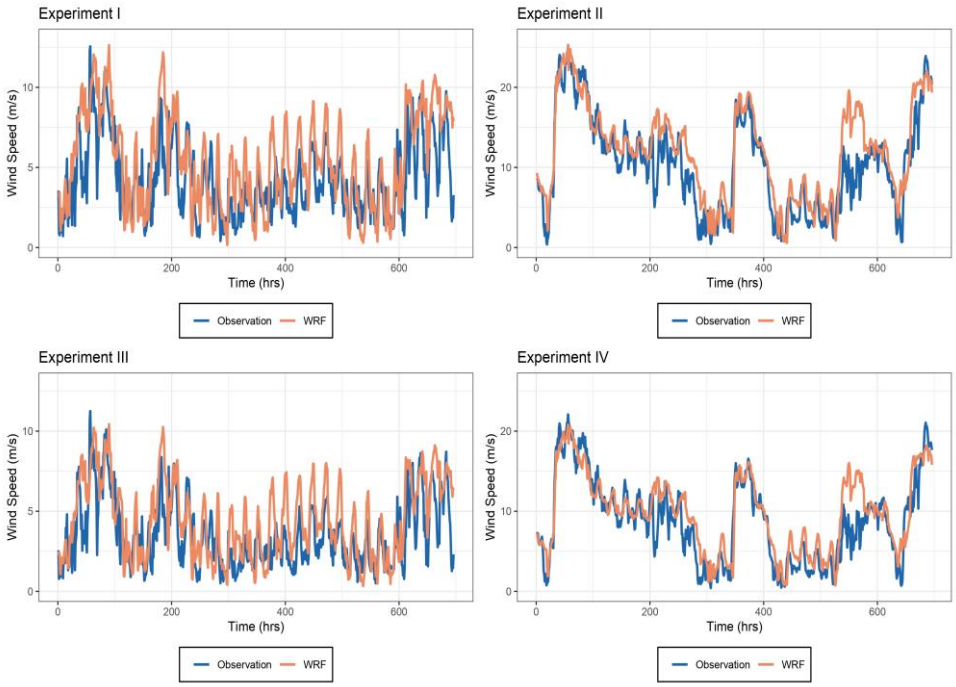

Fig 7. Comparison of wind speed time series: mast measurements vs. WRF model simulations.

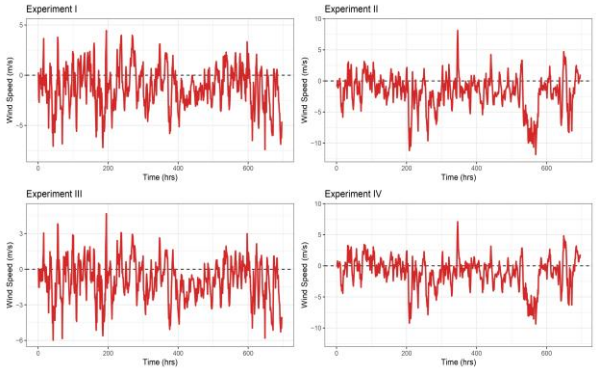

Fig 8. Time series of forecast errors for all four experiments, showing the discrepancy between the WRF simulations and the mast measurements.

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2} \quad (12)$$

The residual e_t is defined as:

$$e_t = O_t - F_t \quad (13)$$

where O_t denotes the in-situ wind speed measurements for a period t, F_t the forecasted wind speed from various models for the same period t, and n the total number of elements.
Bias refers to whether a model's forecast tends to underestimate or overestimate. An overestimated forecast predicts a nonexistent wind resource, while an underestimation overlooks valuable wind resources. MAE and MSE assess the magnitude of forecast residuals, while RMSE measures their spread, placing more weight on larger residuals than on smaller ones.

4. Results and discussions

Four experiments were conducted to apply the proposed model. Experiment I took place from June 1 to 29, 2018, at a height of 40 meters. Experiment II spanned from November 12 to December 10, 2018, at 40 meters. Experiment III covered June 1 to 29, 2018, but at a height of 17.5 meters. Experiment IV was performed from November 12 to December 10, 2018, at 17.5 meters. Each experiment lasted 30 days at hourly temporal resolution, resulting in 720 values, excluding a 24-hour period for the forecast test.

Figure 7 illustrates the wind speed time series obtained from the mast measurements and the WRF simulations for the four experiments. Lower wind speeds characterized Experiments I and III, while higher wind speeds characterized Experiments II and IV.

4.1 Forecast errors analysis

Figure 8 shows the time series N_t of the wind speed forecast errors for the four experiments. Following the methodology, the time series N_t is calculated as the difference between the mast measurements O_t and the WRF simulations W_t.

$$N_t = O_t - W_t \quad (14)$$

In Experiment I, the time series N_t exhibited a minimum value of -7.40 m/s and a maximum value of 4.44 m/s. In Experiment II, the minimum value was -11.83 m/s, while the maximum reached 8.12 m/s. For Experiment III, the minimum value was -5.98 m/s, and the maximum was 4.67 m/s. Lastly, in Experiment IV, the minimum value was -9.35 m/s, whereas the maximum was 7.09 m/s.

Initially, it was determined whether the forecast was overestimated or underestimated by calculating the mean of the time series $\bar{N}$ or bias, and the standard deviation σ. The results in Table 3 indicated negative bias values, meaning the WRF model overestimated wind speed forecasts. The standard

Table 3
Mean and standard deviation of forecast errors

Time series	Bias	Std. Dev. σ	z-Statistic
Experiment I	-1.19	2.06	-15.17
Experiment II	-1.80	2.78	-17.09
Experiment III	-0.91	1.70	-14.10
Experiment IV	-1.09	2.44	-11.78

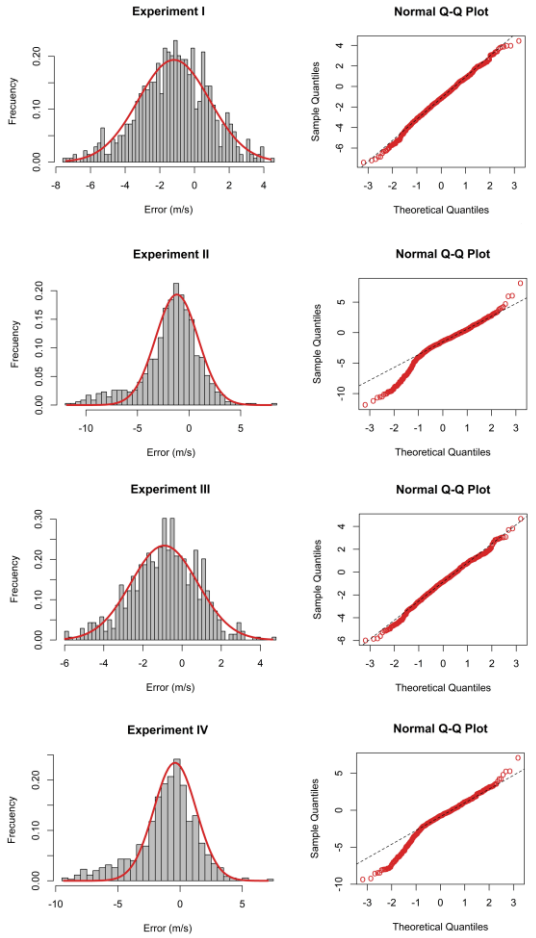

Fig 9. Histogram and Q-Q plot to verify the normality of the forecast errors.

deviation for the experiments was: ±2.06 m/s for Experiment I, ±2.78 m/s for Experiment II, ±1.70 m/s for Experiment III, and ±2.44 m/s for Experiment IV.

A *z-test* was conducted to determine whether the mean $\overline{N}$ is statistically different from zero. Before performing the z-test, normality tests were conducted on the time series. The results indicated that at a significance level of $\alpha = 0.05$, the data were

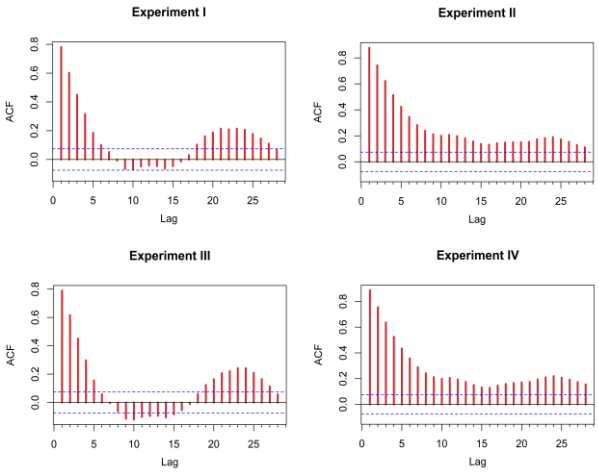

Fig 10. Autocorrelation function (ACF) of the time series of forecast errors, revealing significant correlations at multiple lags, including a seasonal pattern at lag 24 (diurnal cycle).

Table 4
Values of autocorrelation of the forecast errors for Experiment I

Experiment I			
$r_1 = 0.78$	$r_8 = -0.01$	$r_{15} = -0.05$	$r_{22} = 0.21$
$r_2 = 0.60$	$r_9 = -0.06$	$r_{16} = -0.01$	$r_{23} = 0.21$
$r_3 = 0.45$	$r_{10} = -0.07$	$r_{17} = 0.03$	$r_{24} = 0.21$
$r_4 = 0.32$	$r_{11} = -0.05$	$r_{18} = 0.10$	$r_{25} = 0.18$
$r_5 = 0.19$	$r_{12} = -0.04$	$r_{19} = 0.16$	$r_{26} = 0.15$
$r_6 = 0.10$	$r_{13} = -0.05$	$r_{20} = 0.19$	$r_{27} = 0.11$
$r_7 = 0.05$	$r_{14} = -0.06$	$r_{21} = 0.21$	$r_{28} = 0.07$

normally distributed in the four experiments. Figure 9 shows the Q-Q plots. Also, in Figure 9, a normal density function was fitted to the histograms of the four experiments.

The *z-score* for a significance level of $\alpha = 0.05$ is z=1.96. Since the calculated z-scores for the experiments are much lower than z=1.96, the null hypothesis is rejected, indicating that the mean forecast error for the four experiments is statistically different from zero. Next, an analysis of sequential dependencies was performed to verify whether the time series of forecast errors met the condition of predictability. The *autocorrelations* r_k, were determined for the first 28 lags. The critical values interval was: $\pm 1.96/\sqrt{696} = \pm 0.0743$.

The ACF plots, shown in Figure 10, indicate the following results for each experiment: Experiment I, the autocorrelations $r_1, \ldots, r_6, r_{18}, \ldots, r_{27}$ fall outside the interval of the critical values; in Experiment II, the autocorrelations $r_1, \ldots, r_{28}$ fall outside the interval of the critical values; in Experiment III the autocorrelations $r_1, \ldots, r_5, r_{19}, \ldots, r_{27}$ also fall outside the interval of the critical values; in Experiment IV, the autocorrelations $r_1, \ldots, r_{28}$ also fall outside the interval of the critical values. From the ACF plots, the value r_{24} is significant in each experiment, exhibiting seasonal patterns over a 24-hour period. Table 4 shows the autocorrelation values for experiment I.

A *von Neumann ratio of the mean-square-successive difference test* was computed to verify the presence of serial correlation in consecutive values. The z-score for a significance level of $\alpha = 0.05$ is z=1.96. In Experiment I, a value z=20.70 and a coefficient VN= 0.43 were obtained; in Experiment II, a value z=23.22 and a coefficient VN= 0.24 were obtained; in Experiment III, a value z=20.91 and a coefficient VN= 0.42 were obtained; in Experiment IV, a value z=23.47 and a coefficient VN= 0.22 were obtained. In all four experiments, the null hypothesis was rejected in favor of the alternative, indicating the presence of serial correlation at lag 1 in the time series. The p-values from the Ljung-Box test for a lag of h=24 allowed rejection of the null hypothesis in all four experiments. This supports the alternative hypothesis, which indicates that the time series of forecast errors is not independently distributed and exhibits correlations. Additionally, the plots in Figure 12 of the standardized errors against the WRF predicted values clearly illustrate a pattern in the dispersion of the errors.

These results conclude that the time series N_t for each of the four experiments satisfies the condition of being predictable. Therefore, it is plausible to model and forecast the time series.

4.2 SARIMA modeling and forecasting

Figure 11 illustrates the flowchart for modeling and 24-hour forecasting for the time series using the $SARIMA(p,d,q)(P,D,Q)_{24}$ models. The methodology is divided into four steps:

1. Stationarity: Use the Augmented Dickey-Fuller test to verify whether the series N_t is stationary. Calculate the difference $N_t - N_{t-1}$ for a no-stationarity time series; from here, the

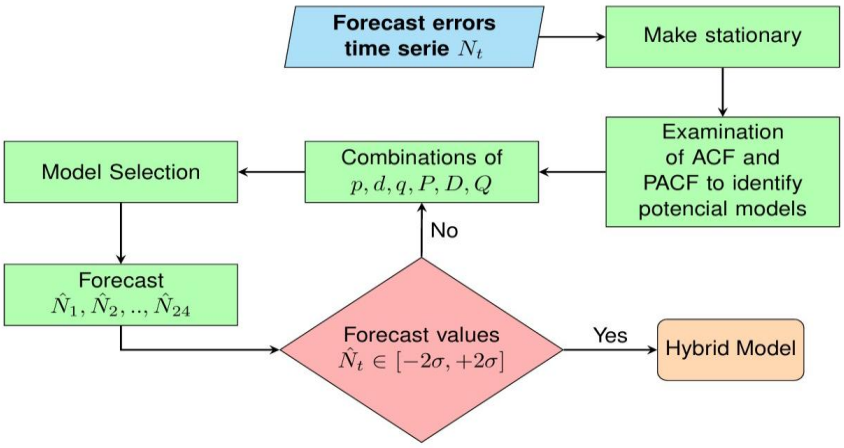

Fig 11. SARIMA modeling and forecasting flowchart. The steps to select the best-fitting model for predicting errors.

value for d is obtained. A seasonal difference $N_t - N_{t-24}$ is calculated for a seasonal non-stationary time series, obtaining the value of D. A maximum of two differences is recommended.

2. Examination of ACF and PACF: Autocorrelation coefficients (ACF) and partial autocorrelation coefficients (PACF) are calculated to identify potential $SARIMA(p,d,q)(P,D,Q)_{24}$ models. The r_k values that fall outside of the interval of critical values $\pm 1.96/\sqrt{n}$ are a guideline for determining the appropriate ranges for the non-seasonal variables p and q, as well as the seasonal variables P and Q.
3. Model selection: Given the appropriate ranges, all possible combinations of the variables *p, d, q, P, D, Q* are determined. Each combination is a potential model. The arima function from the forecast package in R (Hyndman et al., 2008) is used to estimate the parameters $\boldsymbol{\phi}$, $\boldsymbol{\theta}$, $\boldsymbol{\Phi}$, and $\boldsymbol{\Theta}$ for the potential models. Using the Corrected Akaike's Information Criterion (AICc), the 20 best-performing models are selected. The final SARIMA model is selected according to the Parsimony principle, favoring the model with fewer parameters.
4. Forecasting: A 24-step ahead forecast of the time series N_t is performed using the SARIMA model selected in the previous step. The predicted values are denoted by $\widehat{N_1}, \widehat{N_2}, \ldots, \widehat{N_{23}}, \widehat{N_{24}}$. To ensure that the predicted values $\widehat{N_t}$ are consistent with the time series N_t, the condition that the predicted values $\widehat{N_t}$ are within the interval $[-2\sigma, +2\sigma]$ must be satisfied, where σ denotes the standard deviation of the series N_t. Given the case where $\widehat{N_t} \notin [-2\sigma, +2\sigma]$, another potential model is selected, omitting the model that did not meet the condition.

The dataset N_t, consisting of 29 days, was used to train the ARIMA model for each experiment, followed by an out-of-sample forecast for the 30th day. The modeling and forecasting were conducted according to the methodology described previously.

An Augmented Dickey-Fuller test was performed to determine the stationarity of the time series for each experiment. The values of d=0 and D=0 were obtained for Experiment I, d=1 and D=0 for Experiment II, d=0 and D=0 for Experiment III, d=1 and D=0 for Experiment IV. Since the time series exhibits a 24-hour seasonal pattern in the four experiments, a seasonal period of s=24 was determined. To identify potential SARIMA models, the ACF and partial ACF were examined. The autocorrelations r_k that fall outside the critical value interval helped determine the ranges for the variables p, q, P, and Q. In Experiment I, a potential model search was conducted with the variables p and P ranging from 0 to 6, and q and Q also ranging from 0 to 6. In Experiment II, the ranges for p and P were from 0 to 9 and from 0 to 9 for q and Q. In Experiment III, the variables p and P were tested over the range 0 to 5, and the variables q and Q were tested over the range 0 to 5. In Experiment IV, the ranges for p and P were 0 to 9, and similarly for q and Q. The AICc was used to determine the 20 best-fitting SARIMA models for the time series (Table 5);

Table 5
The 20 best-fitting SARIMA models for Experiment I

Model	AICc	Sum	Model	AICc	Sum
SARIMA(1,0,5)(1,0,2)$_{24}$	2110	9	SARIMA(2,0,4)(0,0,2)$_{24}$	2118	8
SARIMA(1,0,5)(2,0,2)$_{24}$	2112	10	SARIMA(0,0,4)(1,0,2)$_{24}$	2121	7
SARIMA(0,0,5)(1,0,2)$_{24}$	2112	8	SARIMA(1,0,3)(1,0,2)$_{24}$	2121	7
SARIMA(1,0,4)(1,0,2)$_{24}$	2113	8	SARIMA(2,0,3)(1,0,2)$_{24}$	2122	8
SARIMA(1,0,5)(0,0,2)$_{24}$	2113	8	SARIMA(1,0,2)(2,0,2)$_{24}$	2122	7
SARIMA(0,0,5)(2,0,2)$_{24}$	2114	9	SARIMA(0,0,4)(2,0,2)$_{24}$	2123	8
SARIMA(1,0,4)(2,0,2)$_{24}$	2116	9	SARIMA(2,0,1)(2,0,0)$_{24}$	2123	5
SARIMA(1,0,4)(0,0,2)$_{24}$	2116	7	SARIMA(1,0,2)(1,0,2)$_{24}$	2124	6
SARIMA(2,0,4)(1,0,2)$_{24}$	2116	9	SARIMA(1,0,3)(2,0,2)$_{24}$	2124	8
SARIMA(0,0,5)(0,0,2)$_{24}$	2117	7	SARIMA(0,0,4)(0,0,2)$_{24}$	2125	6

Sum: Sum of variables *p, d, q, P, D, Q*.

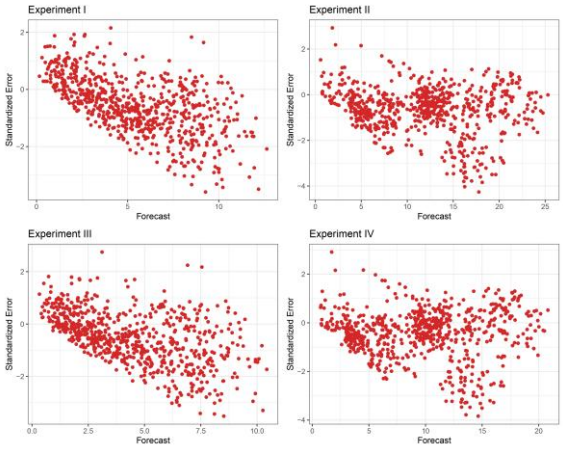

Fig 12. Heteroscedasticity plots: standardized errors against forecasted values, illustrating non-random dispersion patterns that confirm the predictability of the forecast errors.

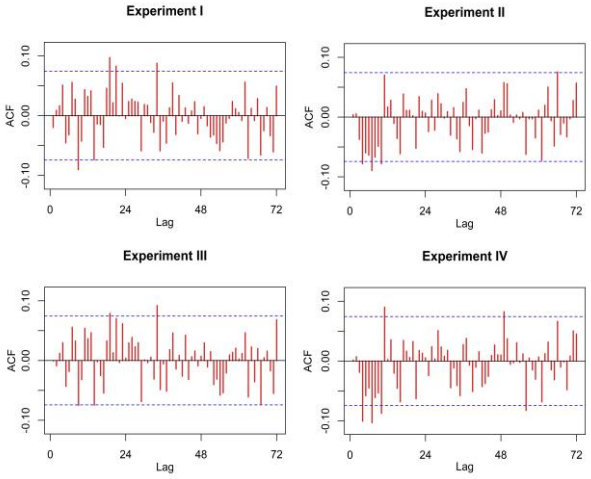

Fig 13. Residual autocorrelation plot of the best-fitting SARIMA model, exhibiting white noise behavior.

among these, the final model was the one with the fewest parameters, according to the Parsimony principle.

To explore a cost-effective alternative, the auto.arima function from the forecast package in R (Hyndman et al., 2008) was tested, yielding satisfactory results. In the function, the variables *p, q, P*, and *Q* are selected by minimizing AICc; variables d and D are selected using the KPSS unit test. The function returns the best-fitting model based on the AICc.

The model selected in Experiment I was SARIMA(2,0,1)(2,0,0)$_{24}$, which has the following equation:

$$N_t = 1.58N_{t-1} - 0.64N_{t-2} + 0.06N_{t-24} - 0.10N_{t-25} + 0.04N_{t-26} + 0.06N_{t-48} - 0.10N_{t-49} + 0.04N_{t-50} - 0.77\epsilon_{t-1} + \epsilon_t \quad (15)$$

In Experiment II, the selected model was SARIMA(0,1,2)(0,0,2)$_{24}$, which has the following equation:

$$N_t = N_{t-1} + 0.05\,\epsilon_{t-1} - 0.10\,\epsilon_{t-2} + 0.09\,\epsilon_{t-24} + 0.01\,\epsilon_{t-25} - 0.01\,\epsilon_{t-26} + 0.10\,\epsilon_{t-48} + 0.01\,\epsilon_{t-49} - 0.01\,\epsilon_{t-50} + \epsilon_t \quad (16)$$

In Experiment III, the selected model was SARIMA(3,0,1)(0,0,2)$_{24}$, which has the following equation:

$$N_t = 1.50N_{t-1} - 0.50N_{t-2} - 0.08N_{t-3} - 0.71\epsilon_{t-1} + 0.07\epsilon_{t-24} - 0.05\epsilon_{t-25} + 0.06\epsilon_{t-48} - 0.05\epsilon_{t-49} + \epsilon_t \quad (17)$$

And, in Experiment IV, the selected model was SARIMA(0,1,2)(0,0,2)$_{24}$, which has the following equation:

$$N_t = N_{t-1} + 0.09\,\epsilon_{t-1} - 0.09\,\epsilon_{t-2} + 0.10\,\epsilon_{t-24} + 0.01\,\epsilon_{t-25} - 0.01\,\epsilon_{t-26} + 0.09\,\epsilon_{t-48} + 0.01\,\epsilon_{t-49} - 0.01\,\epsilon_{t-50} + \epsilon_t \quad (18)$$

Figure 13 displays the autocorrelation plots for the best-fitting SARIMA models. For white noise, about 95% of the autocorrelation values are expected to fall within the interval (Makridakis *et al.*, 2008). This indicates that the residuals of the chosen SARIMA models demonstrate random behavior. Additionally, Figure S1 presents the residuals of the best-fitting SARIMA model, along with the autocorrelation plot and histogram for each experiment. This further confirms that the residuals exhibit random, or white noise, behavior, thereby validating the selection of the model.

The chosen model was used to perform a 24-hour forecast for each experiment. Then, the time series $\widehat{N_t}$ contains the 24 forecast values representing the forecast errors predicted by the SARIMA model. It was also determined whether the forecasted values $\widehat{N_t}$ satisfy the condition $\widehat{N_t} \in [-2\sigma, +2\sigma]$; that is, whether the predicted values fall within the interval of two standard deviations of N_t, in order to ensure that the forecasted

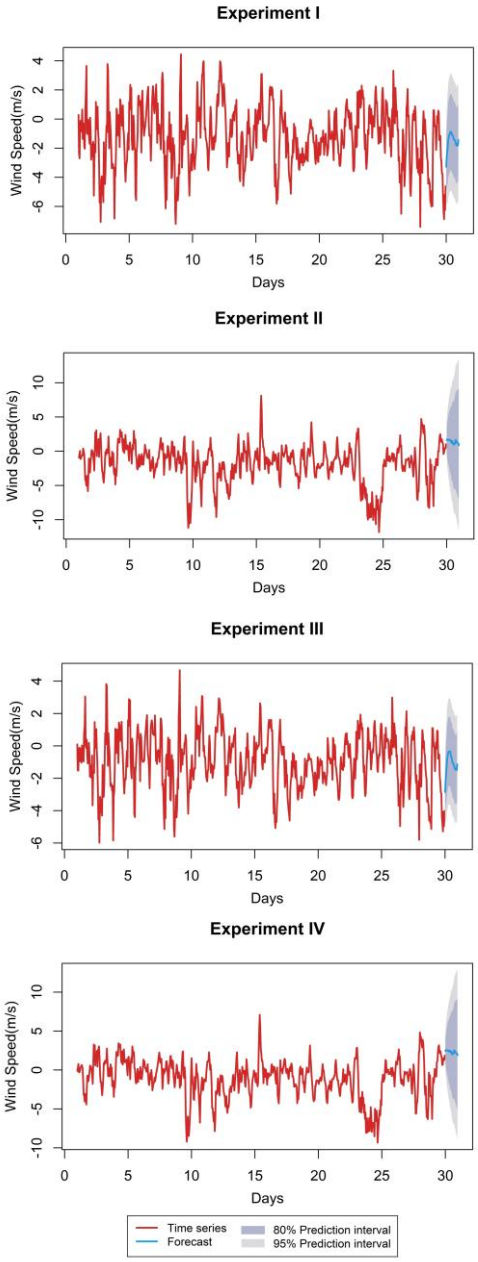

Fig 14. 24-hour-ahead forecast for the error series using the selected SARIMA model, including 80% and 95% prediction intervals.

values $\widehat{N_t}$ are consistent with N_t. Figure 14 shows the time series N_t and the forecasted values $\widehat{N_t}$ with interval predictions at 80% and 95%, for all four experiments.

4.3 Combination of forecasts

The building of the hybrid model H_t was performed using the equation:

$$H_t = \widehat{W_t} + \widehat{N_t} \qquad t \in \{1,..,24\} \quad (19)$$

where $\widehat{W_t}$ denotes the WRF wind speed forecast for the 30th day, and $\widehat{N_t}$ indicates the forecast errors predicted by the SARIMA model.

Both time series contain 24 values; consequently, H_t also contains 24 values. Therefore, the time series $H_t$represents the 24-hour wind speed forecast from the hybrid WRF-SARIMA model.

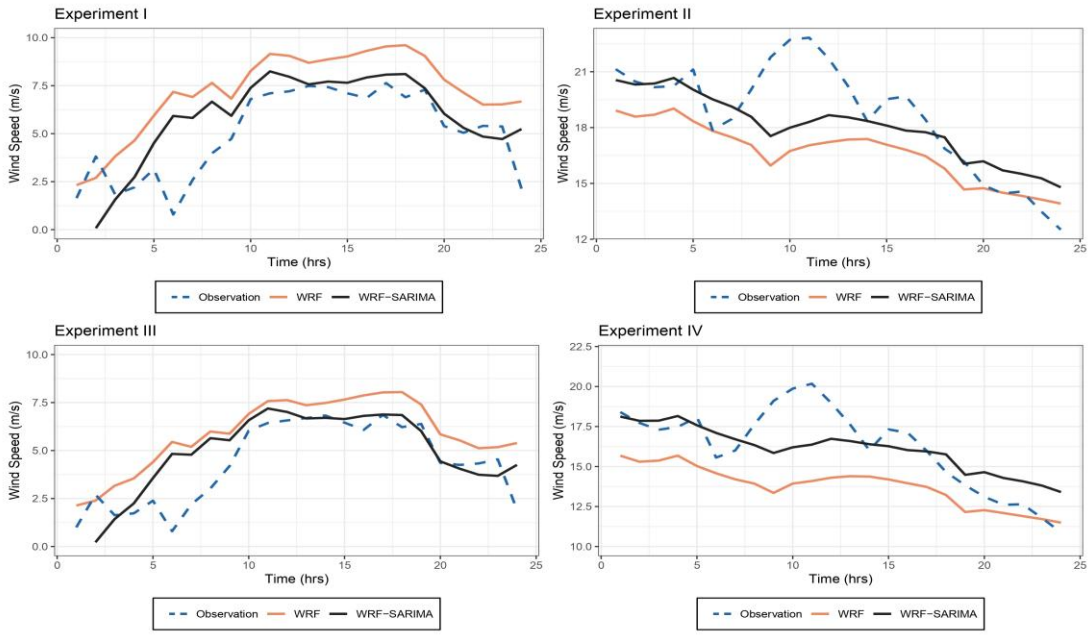

Fig 15. Comparison of 24-hour-ahead wind speed forecasts: mast measurements vs. WRF model vs. hybrid WRF-SARIMA model.

Figure 15 illustrates the wind speed predicted by the hybrid WRF-SARIMA model, the wind speed predicted by the WRF model, and the wind speed from the mast measurements. It is important to note that the proposed methodology can correct both overestimated WRF forecasts, as observed in Experiments I and III, and underestimated WRF forecasts, as seen in Experiments II and IV.

4.4 Model evaluation

The performance of the hybrid WRF-SARIMA model was compared with two other hybrid models, WRF-SES and WRF-Holt. The WRF-SES model combines forecasts from the WRF model with forecast errors predicted by the ARIMA(0,1,1)/SES model. Similarly, the WRF-Holt model integrates WRF forecasts with those generated by the ARIMA(0,2,2)/Holt method. The error metrics for each model are presented in Table 6.

Table 6
Error comparisons among forecasting models

Experiment	Model	Bias	MAE	MSE	RMSE
Experiment I	WRF	-2.22	2.32	6.97	2.64
	WRF-SARIMA	**-0.69**	**1.34**	**3.53**	**1.88**
	WRF-SES	2.56	2.69	8.59	2.93
	WRF-Holt	2.65	2.78	9.02	3.00
	% Improvement	70%	**42%**	**49%**	**28%**
Experiment II	WRF	1.99	2.16	7.75	2.78
	WRF-SARIMA	**0.60**	**1.53**	**4.11**	**2.03**
	WRF-SES	0.91	1.60	4.65	2.16
	WRF-Holt	0.88	1.60	4.61	2.15
	% Improvement	70%	**29%**	**47%**	**27%**
Experiment III	WRF	-1.57	1.59	3.54	1.88
	WRF-SARIMA	**-0.48**	**1.02**	**2.13**	**1.46**
	WRF-SES	2.54	2.59	7.55	2.75
	WRF-Holt	2.61	2.66	2.93	2.82
	% Improvement	69%	**36%**	**40%**	**22%**
Experiment IV	WRF	2.46	2.51	9.14	3.02
	WRF-SARIMA	**0.14**	**1.38**	**3.00**	**1.73**
	WRF-SES	0.55	1.39	3.42	1.85
	WRF-Holt	0.51	1.40	3.41	1.85
	% Improvement	94%	**45%**	**67%**	**43%**

% Improvement: WRF-SARIMA model over the WRF baseline.

The error metrics indicated that the hybrid WRF-SARIMA model outperformed the WRF, WRF-SES, and WRF-Holt models across the four experiments. In Experiment I, the hybrid WRF-SARIMA model demonstrated 42% improvement over the WRF model for the MAE metric, 49% for the MSE, and 28% for the RMSE. In Experiment II, the WRF-SARIMA model showed 29% improvement in MAE, 47% in MSE, and 27% in RMSE compared to the WRF model. In Experiment III, the WRF-SARIMA model demonstrated 36% improvement in MAE, 40% in MSE, and 22% in RMSE compared with the WRF model. Finally, in Experiment IV, the WRF-SARIMA model achieved 45% improvement in MAE, 67% in MSE, and 43% in RMSE over the WRF model. The improvement percentages of the WRF-SARIMA model over the WRF baseline are also shown in Table 6.

Let e_t be the residuals of the WRF model and ε_t be the residuals of the hybrid WRF-SARIMA model. The comparison of both residuals, illustrated in Figure 16, indicates that the

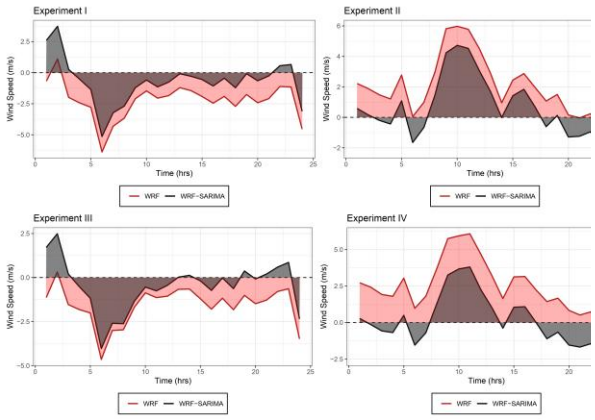

Fig 16. Forecast residual comparisons between the WRF and WRF-SARIMA models, demonstrating the improvement in accuracy.

residuals ε_t have a lower absolute magnitude than the residuals e_t:

$$|\varepsilon_t| \leq |e_t| \quad (22)$$

The results show that the hybrid model reduces forecast error magnitude, improving wind speed forecast accuracy.

4.5 Discussions

The data presented significant challenges in developing this study. Due to high investment costs and the private nature of most wind power developments, obtaining accurate wind speed measurements above 10 meters is difficult. The data utilized in this study are derived from various government-funded projects. The wind speed measurements from La Ventosa, recorded at heights of 17.5 meters and 40 meters, are both reliable and of high quality. Unfortunately, we could not access data from another location with the same quality. Consequently, the proposed hybrid model was tested exclusively at this one site.

Regarding the WRF data, we are thankful for the opportunity to use the outputs of the forecast model developed by the IOA research group at ICAyCC-UNAM, which is continuously tested and improved. However, we could test the model for only two distinct one-month periods due to limitations in our storage capacity. We selected one month characterized by high wind speeds and another month with low wind speeds for the experiments. Given the relatively short dataset, there is a risk of overfitting. To avoid it, we used the principle of Parsimony, which favors simple models over complex ones; paid special attention to the correct identification of seasonal periods; included the condition that the predicted values lie within two standard deviations, and checked the residuals to verify that they were indeed white noise. This dataset was the minimal tested set for which the SARIMA model was able to identify seasonal patterns and trends to generate reliable forecasts. An extended time period would allow the model to identify patterns in the time series more clearly, leading to a better model selection. We acknowledge these limitations; however, we are confident that the robust methodology employed, along with the positive outcomes demonstrated during testing at La Ventosa, a site characterized by complex conditions, strongly indicates that the model can be effectively replicated in diverse regions and conditions.

The generalization of the proposed model to other sites relies on the generalization of both the WRF model and the SARIMA model. The WRF model has been applied and tested in different locations (Chen et al., 2019; Dzebre *et al*., 2020; Jacondino *et al.*, 2021; Liu *et al.*, 2023; Tsai *et al*., 2021; Xu *et al*., 2021; Yang *et al.*, 2023; Zhao *et al*., 2024), whether for wind speed forecasting or for wind power forecasting, proving to be a reliable tool for 24-hour forecasts. On the other hand, the replicability of SARIMA models, according to the proposed methodology, depends on the availability of wind speed measurements at the site of interest. We compared the proposed model's performance with that of other 24-hour wind speed forecasting models reported in the literature. The proposed WRF-SARIMA model showed improvements of 29% to 45% in MAE, and 22% to 43% in RMSE. Xu *et al.* (2021) reported improvements of 24% and 31% in MAE, and 27% and 28% in RMSE, using a long short-term memory model to correct wind speed forecast errors generated with a WRF model. Zhao *et al.* (2024) reported improvements of 15% in MSE and 14% in RMSE by employing temporal convolutional networks and long short-term memory models to reduce forecast errors from a WRF model. Wang *et al.* (2019) reported improvements of 1.7% to 3.9% in RMSE by combining machine-learning models to correct wind speed forecast errors from an NWP model. The above shows that the improvement percentages of our model are within the same range as those reported using machine learning-based models.

The power supplied by the wind turbine is calculated from wind speed data and the turbine's power curve; thus, accurate wind speed forecasts lead to accurate wind power generation forecasts. Improvements in wind speed accuracy are operationally very valuable to power system operators because they enable better estimates of available power, allow dispatch adjustments, and improve grid frequency stability. Also, reduce imbalance penalties in electricity markets when producers do not meet their scheduled commitments. It is also important to recognize that SARIMA models are linear and are, therefore, effective at identifying linear patterns in time series data. However, their performance may be limited when faced with more complex patterns. To address this limitation, future research could explore combining nonlinear and linear statistical models as an error-correction strategy rather than relying solely on SARIMA. Additionally, it is crucial to consider the computational costs associated with this approach. We have noticed that the immediately preceding values significantly affect the forecasts generated by the ARIMA models. As a result, the transition point from the training phase to the forecasting phase is crucial in shaping the forecast trend. For future work, it would be helpful to identify the optimal starting point for the forecast to improve the SARIMA model's performance. One potential approach is to identify change points in the mean, variance, and trend. Our findings align with existing research on the benefits of statistical techniques for improving NWP wind-speed forecasts. The combination of the WRF and SARIMA models proves to be a viable and effective approach for improving the accuracy of the WRF wind speed forecast. Furthermore, we aim to ensure that the proposed methodology can be automated, facilitating its application in real-time scenarios. This leads to better integration of wind energy into power grids, reduced costs for wind farm operators, and minimized penalties for imbalances.

5. Conclusions

The study presents a hybrid 24-hour wind speed forecasting model that combines the Weather Research and Forecasting (WRF) model with the Seasonal Autoregressive Integrated Moving Average (SARIMA) model. The hybrid approach demonstrated consistent performance for both low and high wind speed periods at different heights above the terrain. The findings indicate that the WRF-SARIMA hybrid model outperforms the standalone WRF model. Error metrics showed notable reductions, with a 29% to 44% decrease in MAE, a 39% to 67% decline in MSE, and 22% to 42% improvement in RMSE. By analyzing and modeling the structure of forecast errors through the SARIMA technique, this methodology offers tools for identifying predictable patterns within these errors, thereby facilitating forecast corrections. Overall, this approach enhances the understanding of forecast errors and suggests more effective adjustments to improve forecast accuracy. The model represents a promising alternative for wind speed forecasting due to its automation potential and ease of implementation.

Acknowledgments

The authors would like to express their sincere gratitude to SECIHTI for the financial support provided to Maritza Bernabé in pursuing graduate studies, specifically through the research scholar grant No. 1009386. Additionally, the authors would like to extend their heartfelt appreciation to the IOA group at ICAyCC-UNAM for their support and accessibility in the

development of this study. They also extend their appreciation to the CONACYT project INFR-2014-01-224823 and the PAPIIT project IA107416. And to Coordinación de la Investigación Científica UMSNH. M. Bernabé wishes to express special gratitude to Alissa for her time and kindness, highly appreciated.

Author Contributions: M.B.: Formal analysis, Investigation, methodology, software, visualization, writing—original draft, E.C.: conceptualization, methodology, supervision, writing—review and editing, E.L.: conceptualization, resources, supervision, writing—review and editing, R.C.: resources, supervision, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: The authors received no financial support for the research, authorship, and/or publication of this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

Aasim, Singh, S. A., Mohapatra, A. (2019). Repeated wavelet transform based arima model for very short-term wind speed forecasting, *Renewable Energy,* 136, 758–768. https://doi.org/10.1016/j.renene.2019.01.031

Al-Yahyai, S., Charabi, Y., Gastli, A. (2010). Review of the use of numerical weather prediction (NWP) models for wind energy assessment. *Renewable and Sustainable Energy Reviews*, 14(9), 3192–3198. https://doi.org/10.1016/j.rser.2010.07.001

Chen, S.-H., Yang, S.-C., Chen, C.-Y., Van Dam, C., Cooperman, A., Shiu, H., MacDonald, C., Zack, J. (2019). Application of bias corrections to improve hub height ensemble wind forecasts over the Tehachapi wind resource area. *Renewable Energy*, 140, 281–291. https://doi.org/10.1016/j.renene.2019.03.043

Costa, M. A., Ruiz-Cárdenas, R., Mineti, L. B., Prates, M. O. (2021). Dynamic time scan forecasting for multi-step wind speed prediction. *Renewable Energy*, 177, 584–595. https://doi.org/10.1016/j.renene.2021.05 .160

Ding, L., Bai, Y., Liu, M.-D., Fan, M.-H., Yang, J. (2022). Predicting short wind speed with a hybrid model based on a piecewise error correction method and Elman neural network. *Energy*, 244, 122630. https://doi.org/10.1016/j.energy.2021.122630

Duan, J., Zuo, H., Bai, Y., Duan, J., Chang, M., Chen, B. (2021). Short-term wind speed forecasting using recurrent neural networks with error correction. *Energy*, 217, 119397. https://doi.org/10.1016/j.energy.2020.119397

Dzebre, D. E., Adaramola, M. S. (2020). A preliminary sensitivity study of planetary boundary layer parameterisation schemes in the weather research and forecasting model to surface winds in coastal Ghana. *Renewable Energy*, 146, 66–86. https://doi.org/10.1016/j.renene.2019.06.133

El-Fouly, T. H. M., El-Saadany, E. F., & Salama, M. M. A. (2008). One day ahead prediction of wind speed and direction. IEEE Transactions on Energy Conversion, 23 (1) 191–201.

Grupo Interacción Océano-Atmósfera (IOA) del Instituto de Ciencias de la Atmósfera y Cambio Climático. Universidad Nacional Autónoma de México. http://grupo-ioa.atmosfera.unam.mx/, accessed: 2025-03-11.

Hagens, N. (2020). Economics for the future–beyond the superorganism. *Ecological Economics*, 169, 106520. https://doi.org/10.1016/j.ecolecon.2019.106520

Hong, Y.-Y., Rioflorido, C. L. P. P., Zhang, W. (2024). Hybrid deep learning and quantum-inspired neural network for day-ahead spatiotemporal wind speed forecasting. *Expert Systems with Applications*, 241, 122645. https://doi.org/10.1016/j.eswa.2023.122645

Hu, W., Yang, Q., Chen, H.-P., Yuan, Z., Li, C., Shao, S., Zhang, J. (2021). New hybrid approach for short-term wind speed predictions based on preprocessing algorithm and optimization theory. *Renewable Energy,* 179, 2174–2186. https://doi.org/10.1016/j.renene.2021.08.044

Hyndman, R. J., Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, 27(3), 1–22. https://doi.org/10.18637/jss.v027.i03

Iaousse, M., Jouilil, Y., Bouincha, M., & Mentagui, D. (2023). A comparative simulation study of classical and machine learning techniques for forecasting time series data. *iJOE*, 19(08), 57. https://doi.org/10.3991/ijoe.v19i08.39853

Instituto de Ciencias de la Atmósfera y Cambio Climático (ICAyCC). Universidad Nacional Autónoma de México, https://www.atmosfera.unam.mx/, accessed: 2025-03-11.

Jacondino, W. D., DaSilva Nascimento, A. L., Calvetti, L., Fisch, G., Au gustus Assis Beneti, C., Da Paz, S. R. (2021). Hourly day-ahead wind power forecasting at two wind farms in northeast Brazil using wrf model. Energy, 230, 120841. https://doi.org/10.1016/j.energy.2021.120841

Jurado de Larios, O. (2017). Sensibilidad del modelo wrf ante condiciones iniciales y de frontera: Un estudio de caso en el Valle de México [Ph.D. thesis]. Universidad Nacional Autónoma de México, Ciudad de México, México.

Kontopoulou, V. I., Panagopoulos, A. D., Kakkos, I., & Matsopoulos, G. K. (2023). A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks. *Future Internet*, 15(8), 255. https://doi.org/10.3390/fi15080255

Lawal, A., Rehman, S., Alhems, L. M., Alam, M.M. (2021). Wind speed prediction using hybrid 1D CNN and BLSTM network, *IEEE Access*, 9 156672–156679. https://doi.org/10.1109/ACCESS.2021.3129883

Li, D., Jiang, F., Chen, M., Qian, T. (2022). Multi-step-ahead wind speed forecasting based on a hybrid decomposition method and temporal convolutional networks. *Energy*, 238, 121981. https://doi.org/10.1016/j. energy.2021.121981

Liu, M., Ding, L., Bai, Y. (2021). Application of a hybrid model based on empirical mode decomposition, novel recurrent neural networks, and the ARIMA to wind speed prediction. *Energy Conversion and Management*, 233, 113917. https://doi.org/10.1016/j.enconman.2021.113917

Liu, W., Bai, Y., Yue, X., Wang, R., Song, Q. (2024). A wind speed forcasting model based on rime optimization based vmd and multi-headed self-attention-lstm. *Energy*, 294, 130726. https://doi.org/10.1016/j.energy .2024.130726

Liu, X., Li, Z., Shen, Y. (2024). Study on downscaling correction of near-surface wind speed grid forecasts in complex terrain. *Atmosphere*, 15(9), 1090. https://doi.org/10.3390/atmos15091090

Liu, X., Lin, Z., Feng. Z. (2021). Short-term offshore wind speed forecast by seasonal ARIMA: a comparison against GRU and LSTM. *Energy*, 227, 120492. https://doi.org/10.1016/j.energy.2021.120492

Liu, X., Zhang, L., Wang, J., Zhou, Y., Gan, W. (2023). A unified multi-step wind speed forecasting framework based on numerical weather prediction grids and wind farm monitoring data. *Renewable Energy*, 211, 948–963. https://doi.org/10.1016/j.renene.2023.05.006

Liu, Z., Jiang, P., Zhang, L., Niu, X. (2020). A combined forecasting model for time series: Application to short-term wind speed forecasting. *Applied Energy*, 259, 114137. https://doi.org/10.1016/j.apenergy.2019.11 4137

López-Espinoza, E. D., Zavala-Hidalgo, J., Mahmood, R., Gómez-Ramos, O. (2020). Assessing the impact of land use and land cover data representation on weather forecast quality: A case study in central México. *Atmosphere*, 11(11), 1242. https://doi.org/10.3390/atmos11111242

Lopez-Villalobos, C. A., Rodriguez-Hernandez, O., Campos-Amezcua, R., Hernandez-Cruz, G., Jaramillo, O. A., Mendoza, J. L. (2018). Wind turbulence intensity at La Ventosa, Mexico: a comparative study with the IEC61400 standards. *Energies*, 11(11), 3007. https://doi.org/10.3390/en11113007

Makridakis, S. (1989). Why combining works? *International Journal of Forecasting*, 5(4), 601–603.

Makridakis, S., Wheelwright, S. C., Hyndman, R. J. (2008). *Forecasting Methods and Applications*, John Wiley & Sons.

Mi, L., Shen, L., Han, Y., Cai, C., Zhou, P., Li, K. (2023). Wind field simulation using wrf model in complex terrain: A sensitivity study with orthogonal design. *Energy*, 285, 129411. https://doi.org/10.1016/j.energy .2023.129411

Moreno, S. R., Mariani, V. C., Dos Santos Coelho, L. (2021). Hybrid multi-stage decomposition with parametric model applied to wind

speed forecasting in Brazilian northeast. *Renewable Energy*, 164, 1508–1526. https://doi.org/10.1016/j.renene.2020.10.126

Moreno, S. R., Seman, L. O., Stefenon, S. F., Dos Santos Coelho, L., Mariani, V. C. (2024). Enhancing wind speed forecasting through synergy of machine learning, singular spectral analysis, and variational mode decomposition. *Energy*, 292, 130493. https://doi.org/10.1016/j.energy.2024.130493

Myers, T. A., Van Ormer, A., Turner, D. D., Wilczak, J. M., Bianco, L., Adler, B. (2024). Evaluation of hub-height wind forecasts over the New York Bight. *Wind Energy*, 27(10), 1063–1073. https://doi.org/10.1002/we.2936

Rivera-Martínez, S. (2018). *Análisis del uso de suelo y vegetación en México entre 1968 y 2011 para su uso en un modelo de pronóstico meteorológico* [Ph.D. Thesis], Universidad Nacional Autónoma de México, Ciudad de México, México.

Romero-Centeno, R., Zavala-Hidalgo, J., Gallegos, A., O'Brien, J. J. (2003). Isthmus 40 of Tehuantepec wind climatology and ENSO signal. *Journal of Climate*, 16(15), 2628–2639. https://doi.org/10.1175/1520-0442(2003)016<2628:IOTWCA>2.0.CO;2

Team, R. C. (2025). R: A language and environment for statistical computing. https://www.r-project.org/

Thu, N. T. H., Bao, P. Q., Van, P. N. (2023). A hybrid model of decomposition, extended Kalman filter and autoregressive-long short-term memory network for hourly day ahead wind speed forecasting. *J. Appl. Sci. Eng*, 27, 3063 3071. https://doi.org/10.6180/jase.202409_27(9).0004

Tian, Z., Wang, G., Ren, Y. (2020). Short-term wind speed forecasting based on autoregressive moving average with echo state network compensation. Wind Engineering, 44(2), 152–167. https://doi.org/10.1177/0309524X19849867

Tsai, C.-C., Hong, J.-S., Chang, P.-L., Chen, Y.-R., Su, Y.-J., Li, C.-H. (2021). Application of bias correction to improve wrf ensemble wind speed forecast. Atmosphere, 12(12), 1688. https://doi.org/10.3390/atmos12121688

Tyass, I., Khalili, T., Mohamed, R., Abdelouahed, B., Raihani, A., Mansouri, K. (2023). Wind speed prediction based on statistical and deep learning models. *International Journal of Renewable Energy Development*, 12(2), 288. https://doi.org/10.14710/ijred.2023.48672

Secretaría de Energía (SENER) (2024). *Informe pormenorizado sobre el desempeño y las tendencias de la industria eléctrica nacional 2023*. Gobierno de México.

Wang, H., Han, S., Liu, Y., Yan, J., Li, L. (2019). Sequence transfer correction algorithm for numerical weather prediction wind speed and its application in a wind power forecasting system. *Applied Energy*, 237, 1–10. https://doi.org/10.1016/j.apenergy.2018.12.076

Wang, L., Li, X., Bai, Y. (2018). Short-term wind speed prediction using an extreme learning machine model with error correction. *Energy Conversion and Management*, 162, 239–250. https://doi.org/10.1016/j.enconman.2018.02.015

Wang, S., Wang, J., Lu, H., Zhao, W. (2021). A novel combined model for wind speed prediction–combination of linear model, shallow neural networks, and deep learning approaches. *Energy*, 234, 121275. https://doi.org/10.1016/j.energy.2021.121275

Xu, W., Liu, P., Cheng, L., Zhou, Y., Xia, Q., Gong, Y., Liu, Y. (2021). Multi-step wind speed prediction by combining a wrf simulation and an error correction strategy. *Renewable Energy*, 163, 772–782. https://doi.org/10.1016/j.renene.2020.09.032

Yang, J., Sengupta, M., Xie, Y., Shin, H. H. (2023). Developing a 20-year high resolution wind data set for Puerto Rico. *Energy*, 285, 129439. https://doi.org/10.1016/j.energy.2023.129439

Yang, X., Dai, K., Zhu, Y. (2023). Calibration of gridded wind speed forecasts based on deep learning. *Journal of Meteorological Research*, 37(6), 757–774. https://doi.org/10.1007/s13351-023-3001-1

Zhang, D., Hu, G., Song, J., Gao, H., Ren, H., Chen, W. (2024). A novel spatiotemporal wind speed forecasting method based on the microscale meteorological model and a hybrid deep learning model. *Energy*, 288, 129823. https://doi.org/10.1016/j.energy.2023.129823

Zhang, Y., Pan, G., Chen, B., Han, J., Zhao, Y., Zhang, C. (2020). Short-term wind speed prediction model based on ga-ann improved by vmd. *Renewable energy*, 156, 1373–1388. https://doi.org/10.1016/j.renene.2019.12.047

Zhang, Y.-M., Wang, H. (2023). Multi-head attention-based probabilistic CNN-BiLSTM for day-ahead wind speed forecasting. Energy, 278, 127865. https://doi.org/10.1016/j.energy.2023.127865

Zhao, F. (2025). *Global Wind Report* | GWEC. Global Wind Energy Council.

Zhao, J., Guo, Y., Lin, Y., Zhao, Z., Guo, Z. (2024). A novel dynamic ensemble of numerical weather prediction for multi-step wind speed forecasting with deep reinforcement learning and error sequence modelling. *Energy*, 131787. https://doi.org/10.1016/j.energy.2024.131787

Zhao, J., Guo, Z., Guo, Y., Lin, W., Zhu, W. (2021). A self-organizing forecast of day ahead wind speed: Selective ensemble strategy based on numerical weather predictions. *Energy*, 218, 119509. https://doi.org/10.101 6/j.energy.2020.119509

Supplementary files

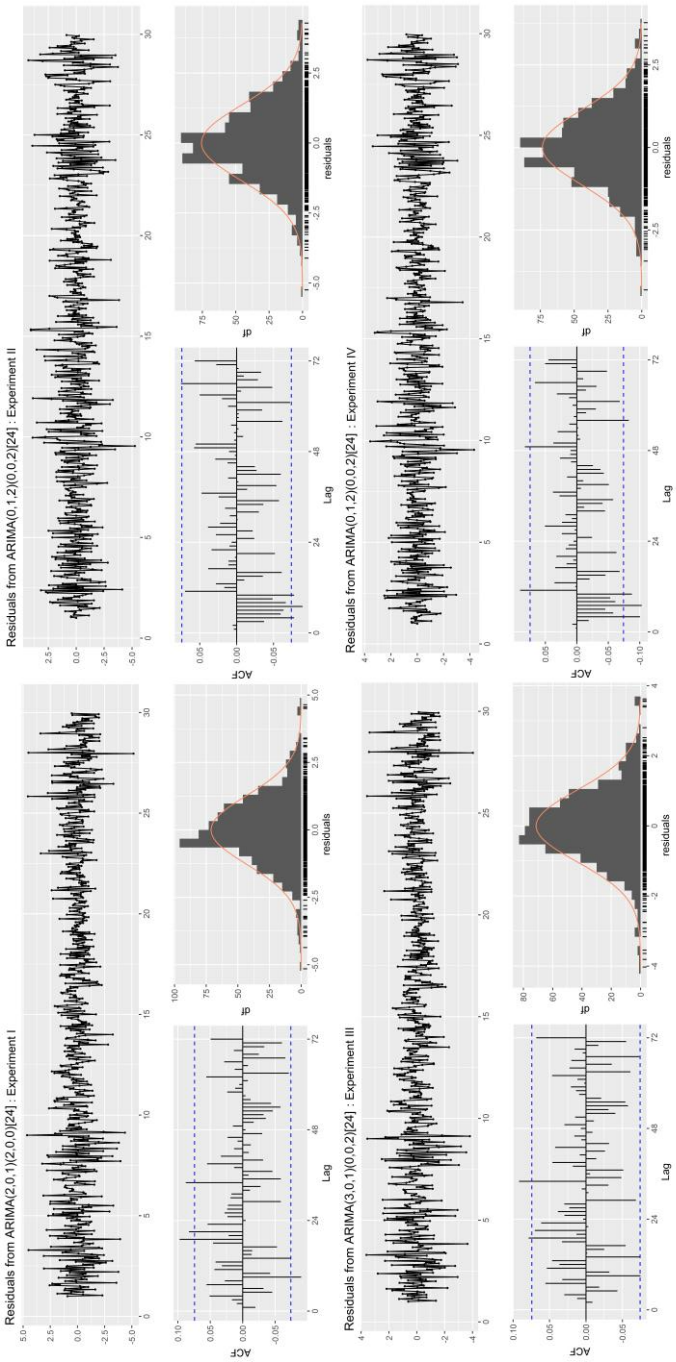

Fig S1. Residual check of the best-fitting SARIMA model, exhibiting white noise behavior