ISSN 2007-9737

Importance of the Parameterization Schemes in the WRF Model for Wind Speed Forecasting: Case Study of Tepuxtepec, Michoacán

Itzagueri García-Rodríguez[1], Alma Rosa Mendez-Gordillo[2,*], Rafael Campos-Amezcua[3], Sixtos A. Arreola-Villa[2], Erasmo Cadenas-Calderón[1]

[1] Universidad Michoacana de San Nicolás de Hidalgo,
Facultad de Ingeniería Mecánica,
Mexico

[2] Universidad Autónoma de Coahuila,
Facultad de Ingeniería Mecánica y Eléctrica Unidad Norte,
Mexico

[2] TecNM/Centro Nacional de Investigación y Desarrollo Tecnológico,
Mexico

{itzagueri.garcia, ecadena}@umich.mx, {alma_mendez, svilla}@uadec.edu.mx,
rafael.ca@cenidet.tecnm.mx

Abstract. The planning of wind energy dispatch constantly faces the challenges of wind speed intermittency and variability. Therefore, it is crucial to have models that generate reliable forecasts to support the development of this renewable energy source. This study evaluates the performance of three configurations of the Weather Research and Forecasting (WRF) model for hourly wind speed forecasting in Tepuxtepec, Michoacán, Mexico, with a 24-hour forecast horizon. Three parameterization schemes were compared: WRF_WMK, WRF_TBG, and WRF_MYB. These schemes were selected based on the geographic, climatic, and meteorological characteristics of the region, as well as the need to assess the WRF model's performance under different physical configurations. Simulations were conducted for four representative dates—one per season—considering the annual temperature cycle that influences wind behavior.

The simulations used MERRA-2 reanalysis data as input and were evaluated against measurements from NASA's POWER project. The comparison between simulated and observed wind speeds was performed using four error metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Bias, and the Correlation Coefficient (r). Additionally, Prediction Intervals (PIs) at 80% and 95% confidence levels were calculated to assess the reliability of the forecasts. The results showed that the WRF_TBG configuration outperformed the others, reducing RMSE by up to 60% compared to WRF_WMK. The forecasted values were within the 80% PI for up to 80% of the total values, and within the 95% PI for up to 100%. Seasonal evaluation revealed that the model performed best in winter and worst in summer, likely due to the influence of intense convective processes during the latter season.

Keywords. WRF model, parameterization schemes, wind speed forecasting.

1 Introduction

Currently, the primary application of wind energy is electricity generation through wind turbines, which efficiently and sustainably convert the kinetic energy of the wind into electrical energy. According to the International Energy Agency's (IEA) World Energy Outlook 2024, global renewable energy generation capacity is projected to increase from the current 4,250 GW to nearly 10,000 GW by 2030 [26]. Wind and solar energy are expected to lead this expansion and represent a substantial portion of global electricity generation in the coming

decade. This projection underscores the strategic role of wind energy in the global energy transition.

In terms of installed wind capacity, China leads the world with 441,895 MW, followed by the United States and Germany. Significant progress has also been made in Mexico, with a total installed capacity of 7,318 MW, positioning the country 17th globally [56, 45]. This growth reflects Mexico's commitment to diversifying its energy matrix and incorporating renewable sources to meet electricity demand while promoting emission reductions and greater energy independence [27, 3].

The first step in increasing installed capacity is to identify sites with wind potential through a comprehensive assessment that considers factors such as wind speed and consistency, among other important characteristics that enable the effective harnessing of this resource. In Mexico, the Isthmus of Tehuantepec is one of the regions with the most significant wind resources worldwide. This region has driven the development of numerous electricity generation projects, solidifying the country's role as a key player in the international wind energy market. In the search for other sites with promising wind resources, this study analyzes the locality of Tepuxtepec in the state of Michoacán. Due to its topographic characteristics and wind speed, Tepuxtepec represents a potential site for investment in wind energy development.

In recent years, the need for accurate wind speed forecasts has become more critical due to growing energy demand and the increasing contribution of renewable sources to electricity generation. One of the main challenges in wind energy generation lies in its strong dependence on wind speed [28]. Multiple factors influence this meteorological variable, making it one of the most difficult to integrate into electrical power generation systems. Accurate wind speed forecasts are essential for energy dispatch planning, as they allow for the generation of electricity production estimates when combined with other techniques [10].

It is important to note that different types of generation forecasts are defined by their time horizon and the needs of the end user. Medium-term forecasts (MTF), with horizons ranging from 2 to 20 days, are used by energy generators for decisions related to supply security and grid maintenance. System operators use short-term forecasts (STF), which range from 6 to 48 hours, to prevent congestion, schedule generation, and manage loads. Very short-term forecasts (VSTF), with horizons from minutes to six hours, are essential for real-time operation and dispatch of generation units, and for managing transmission constraints [36, 4]. This study generated hourly wind speed forecasts with a 24-hour forecast horizon, in alignment with the STF category, making them particularly useful for short-term planning and operational decision-making in the energy sector.

Forecasting models are generally categorized as global or regional. Global models [49, 20] simulate atmospheric conditions on a worldwide scale using low-resolution grids. Regional or mesoscale models, on the other hand, focus on smaller areas (on the order of hundreds of kilometers) with significantly higher spatial resolution. To enhance local applicability, a technique known as downscaling is applied to improve the resolution of global forecasts. There are two main downscaling approaches: statistical and dynamic [19, 14]. For atmospheric features smaller than 10 km, models must include non-hydrostatic processes to accurately simulate mesoscale meteorology. Such processes include solar radiation, turbulence, evaporation, condensation, convection, and surface heat and moisture fluxes.

The Weather Research and Forecasting (WRF) model is among the most widely used regional models for dynamic downscaling. It is extensively used in operational meteorology and research applications [43, 52]. Its main advantage lies in its adaptability to a wide range of spatial and temporal scales, from operational short-term forecasting to long-term climate studies. This flexibility is enabled by a wide variety of physical parameterization schemes [48]. These schemes allow users to tailor the model configuration to the specific characteristics of the study area and the atmospheric processes under investigation.

Therefore, the appropriate selection of physical parameterizations is a key step in achieving reliable forecasts in diverse geographic contexts [46, 17, 2].

Parameterization schemes in WRF represent physical processes that cannot be explicitly resolved at the chosen resolution [21]. An improper selection of these schemes can lead to significant biases, particularly in variables such as temperature, precipitation, and wind speed [33, 18]. Thus, the configuration must be tailored to the regional characteristics and the meteorological objective of the study [51]. The use of appropriate schemes has been shown to reduce uncertainty in simulation results, producing more realistic atmospheric behavior and improving forecast skill [16, 5]. This is the main contribution of the present study, which evaluates multiple WRF parameterizations for forecasting wind speed at an unstudied site using modest computational resources.

Recent research emphasizes the growing use of advanced modeling techniques, including machine learning (ML), to improve wind speed forecasting in areas with high wind potential. ML models can estimate spatio-temporal wind fields with high accuracy, especially in regions with sparse observational data. These approaches have been employed to identify optimal zones for wind energy development by forecasting wind speed across diverse geographical areas [1, 7, 44]. Within this context, the WRF model remains a powerful tool due to its ability to simulate atmospheric dynamics in any location of interest. Its flexibility is based on proper domain configuration and the selection of appropriate parameterizations [42].

To evaluate the performance of the model, time series of wind speed data were generated for multiple representative dates throughout the year. These simulations were compared to reference data from the National Aeronautics and Space Administration (NASA) POWER project [40]. NASA's database was chosen for its open access, georeferenced data retrieval capabilities, and its availability of complete wind speed records for the entire year. The selection of the three parameterization schemes was driven by the need to assess WRF's performance in a region with complex topography and variable weather conditions. Each configuration allows for analyzing wind behavior under different atmospheric scenarios, based on the specific features of the site. Forcing data were obtained from the NCEP GFS Global Forecast Historical Archive [41], and5 simulations were carried out for each selected date. Model performance was assessed using standard statistical error metrics by comparing forecasted values against the NASA reference data.

The remainder of this article is structured as follows: Section 2 describes the methodology, including the study area, data sources, date selection process, and WRF configurations. Section 3 presents and analyzes the results, comparing the different model configurations and evaluating the prediction accuracy using error metrics and prediction intervals. Section 4 summarizes the main conclusions and discusses their implications for wind energy project planning in Tepuxtepec and similar regions. Finally, future research directions and practical applications are proposed.

2 Methodology and Data

2.1 Study Site

The municipality of Contepec, located in the northeast of the state of Michoacán, Mexico, is geographically situated at coordinates 20°01′30″ N and -100°14′15″ W, at an altitude of 2,490 meters above sea level.

From a climatic perspective, the region has a temperate climate with rainfall concentrated in the summer. Average temperatures in the area range from 8.6°C to 22.4°C [11].

The Tepuxtepec Dam is also found in this area. The potential land uses are agricultural and livestock, and the physiography of the region consists of mountains, hills, plains, and valleys [25]. These topographic and climatic characteristics were taken into account when selecting configurations for each parameterization scheme. Figure 1 shows a map of the central region of the Mexican Republic, highlighting the study site for this research. The inset photo illustrates the flat terrain, where part of the Tepuxtepec Dam can be seen.

Fig. 1. Site of study: Tepuxtepec, Michoacán, Mexico.

2.2 NASA POWER Reference Databases

Various sources of information were examined to select the database for wind speed records. Two meteorological stations near the study site were identified: the Tepuxtepec station, located at 19°59′48″ N and -100°13′49″ W, and the Temascales station, situated at 20°03′03″ N and -100°08′58″ W [25]. However, the records from these stations were incomplete, leading to the exploration of the Prediction Of Worldwide Energy Resources (POWER) database, developed by the Langley Research Center (LaRC) at the National Aeronautics and Space Administration (NASA). This database provided recent measurements and complete, easily accessible records. The NASA POWER (MERRA-2) data represent an excellent choice for wind resource analysis, particularly in areas with limited measurement infrastructure. The NASA database was chosen because it offers recent measurements, enables precise location selection for the study site, and provides complete and easily accessible records [40]. NASA-POWER wind data have been validated in various wind energy studies for reliability in areas lacking ground stations [31, 57]. The data are derived from MERRA-2 and undergo quality control processes ensuring consistency.

The data obtained spans from January 1, 2024, to December 31, 2024, with an hourly temporal resolution, and at 10 meters above the surface. The WRF model also provides wind speed at this height, allowing for direct validation without extrapolating wind speed to different heights. In addition to wind data, accurate surface information was required for the model setup. Land cover data were obtained from INEGI's Serie VI database, which includes updated satellite-based vegetation classification validated through field campaigns [13]. These data were used in the WRF Preprocessing System.

2.3 Selection of Dates for Wind Speed Forecast Generation

In order to ensure appropriate forecast date selection, the wind was considered a meteorological variable that fluctuates over time due to various factors, including atmospheric pressure, temperature, and changes in topography. These factors lead to wind direction and speed shifts throughout the day and across seasons [58]. At the study site, differences in heating between the land and the water body create local winds, including daytime and nighttime breezes. The sample was segmented using the K-means technique, part of machine learning, specifically unsupervised learning [15]. This approach aided in identifying patterns and clusters in the wind data, enabling a more informed selection of representative dates for forecasting. The number of clusters was generated using measurements from 2024 extracted from the NASA POWER project. These clusters were selected based on the results of the elbow method and the silhouette index, which graphically identifies the optimal number of clusters. These techniques facilitated the determination of the most appropriate grouping of data points, ensuring robust and meaningful segmentation for analyzing wind patterns. To reduce computational cost while ensuring data representativeness, the K-means algorithm was applied to the 2024 hourly wind speed time series. The number of clusters was evaluated using the elbow method, which shows a consistent reduction in the sum of squared errors (SSE), stabilizing from four clusters onward (Figure 2). Although the decrease in SSE becomes progressively smaller after K = 3, the difference between K = 4 and higher values is minimal, indicating that additional clusters do not contribute significantly to improved representation.

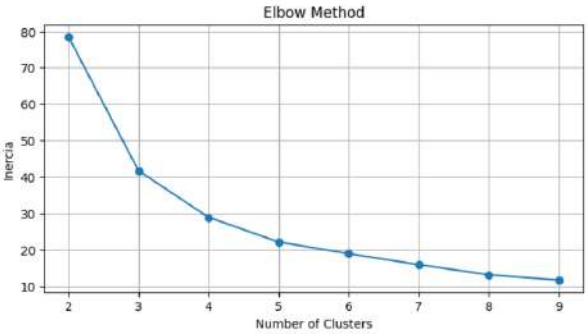

Fig. 2. K-means: Elbow Method used to determine the optimal number of clusters.

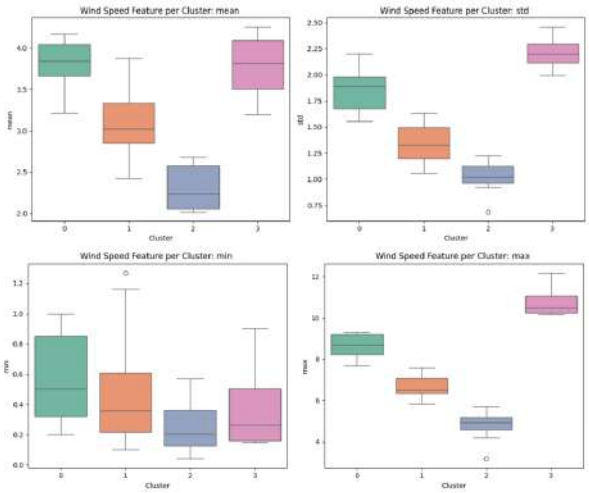

Fig. 3. Box-plots of descriptive statistics show differences in mean, standard deviation, minimum, and maximum wind speeds across the clusters.

Therefore, K = 4 was selected to balance statistical performance and climatological reasoning. This choice also aligns with the natural division of the data into four seasons, allowing for the selection of one representative day per season. The boxplots of descriptive statistics (Figure 3) confirm that each cluster exhibits distinct wind patterns in terms of mean, variability, and range. Thus, the reduced set of four representative days ensures seasonal coverage and preserves the essential variability required for evaluating wind forecast performance across diverse atmospheric conditions. Although the initial interpretation suggested three clusters, four were chosen, guided by both numerical criteria and climatological knowledge. This choice allowed for capturing seasonal variability in the data by selecting one representative day for each season.

Table 1. Representative simulation dates selected for each season based on K-means clustering.

Season	Simulation Dates
Winter	January 21, 2024
Spring	April 11, 2024
Summer	July 16, 2024
Autumn	October 6, 2024

The decision to use four clusters was further validated by the box-plots of descriptive statistics in Figure 3, which clearly show differences in mean, standard deviation, minimum, and maximum wind speeds across the clusters. These patterns highlight distinct wind behaviors in each cluster, confirming that the reduced dataset is both representative and appropriate for the objectives of this study. Consequently, a proposal of days grouped by week was created, representing each cluster type with days in homogeneous groups. From these, only one day per cluster was selected to represent the others. Table 1 presents the chosen dates for the four seasons.

2.4 Parameterizations Schemes and WRF Modeling

This study focuses on atmospheric modeling using the open-source WRF model, emphasizing the role of physical parameterization schemes in improving the quality of wind forecasts. By selecting configurations tailored to the site's orographic and land-use characteristics, we seek to increase the physical realism of the model outputs. The configuration used in this study was designed for a 24-step-ahead forecast horizon, balancing detail and computational feasibility. This approach facilitates high-quality, site-specific forecasts without relying on other techniques that base their construction on anemometric measurements and improves reproducibility in

Fig. 4. Delimitation of the control domain for simulating Tepuxtepec, Michoacán, Mexico, at a resolution of 10 km.

Table 2. Parameterization scheme combinations used in the WRF model, specifying the microphysics, planetary boundary layer (PBL) and cumulus convection options.

Schemes	Microphysics	PBL	Convection
WRF_WMK	WMS6	MYNN2.5	Kain-Fritsch
WRF_TBG	Thompson	Boulac	Grell-Freitas
WRF_MYB	Morrison	YSU	BMJ

similar regions. While this study focused on physical model optimization, future research will integrate hybrid and machine learning models with WRF outputs to improve short-range forecasts, as suggested by recent work [9, 12]. The WRF model, version 4.0 [53], was utilized with a single-domain configuration and a horizontal resolution of 10 km. Figure 4 illustrates the spatial configuration of the nested simulation domains used in the WRF model. The outer domain provides synoptic-scale conditions, while the inner domains progressively refine the resolution to better capture local atmospheric processes. The innermost domain focuses on the Tepuxtepec area, allowing a detailed representation of wind dynamics influenced by local topography and land use. This nested approach is essential for achieving accurate mesoscale forecasting in complex terrain. Initial and boundary conditions were obtained from the NCEP GFS Historical Archive (NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive) [41], extracting data solely for the designated simulation dates.

In the WRF model, microphysics describes the representation of particle-level processes occurring in clouds, such as condensation, freezing, melting, deposition, and droplet coalescence, which determine the formation and evolution of precipitation [47]. The Planetary Boundary Layer (PBL) is the part of the atmosphere in direct contact with the Earth's surface, where intense turbulent processes occur. In this layer, heat, moisture, and momentum exchange between the surface and the air, influencing cloud formation, pollutant dispersion, and the evolution of wind near the ground. In the WRF model, PBL parameterization is crucial, especially for forecasts at heights such as 10 meters above ground level, as it directly impacts the accuracy of meteorological predictions [47].

Convection parameterization in meteorological models is crucial for representing small-scale processes that cannot be explicitly resolved due to limitations in spatial resolution. This parameterization is particularly relevant in regions where vertical air movements significantly influence local weather development. Convection is related to the dynamic processes of vertical air transport (updrafts and downdrafts), which generate convective clouds (e.g., cumulus) and are linked to energy and moisture transfer, as well as storm development [47].

The three selected configurations `WRF_WMK`, `WRF_TBG`, and `WRF_MYB` were chosen based on their physical diversity and proven applicability to complex terrain.`WRF_TBG` combines Thompson microphysics, BouLac PBL, and Grell-Freitas cumulus schemes, offering accurate wind representation as reported in similar studies [21].

Three schemes with different parameterization configurations were evaluated for each proposed date (Table 2). These schemes were designed based on the site's topographic and climatic characteristics, and abbreviations were created using the first letter of each parameterization name for easier identification. This method was effective because the parameterizations for microphysics, planetary boundary layer (PBL), and convection

did not share overlapping initials, allowing for clear distinction.

- **WRF_WMK configuration.** It utilizes the WSM6 scheme for microphysics, which simulates the formation and evolution of hydrometeors, including water vapor, rain, and more. This microphysics scheme is recognized for balancing complexity and computational efficiency [23]. MYNN 2.5 for the planetary boundary layer (PBL) is based on a hierarchy of second-order closures and is designed to represent turbulence in the boundary layer accurately. This approach proves beneficial in high-resolution simulations where a detailed representation of turbulence is essential [39]. The Kain-Fritsch scheme for deep convection addresses the effects of convective storms not explicitly resolved by the model. It employs a one-dimensional column model that simulates the development and dissipation of convective clouds, adjusting the temperature and humidity profiles to reflect convective transport [32].
- **WRF_TBG configuration.** It employs the Thompson scheme for microphysics, which simulates the formation and evolution of hydrometeors, including water vapor and rain. This microphysics scheme is a single-moment parameterization [50]; BouLac for the planetary boundary layer (PBL) uses a first-order closure parameterization that calculates the boundary layer height and turbulent diffusion coefficients based on turbulent kinetic energy (TKE) [8]; and Grell-Freitas for convection. This deep convection parameterization scheme is an advanced version of the Grell scheme that incorporates a stochastic representation of convective clouds and their interactions with the environment. It is suitable for high-resolution models and enables a smooth transition between resolved and unresolved convection scales [22].
- **WRF_MYB configuration.** It employs the Morrison double-moment scheme for microphysics, which enables a more accurate depiction of microphysical processes in clouds. This scheme is particularly beneficial in research examining the impact of aerosols on cloud formation and precipitation [37]. The YSU (Yonsei University) scheme is used for the planetary boundary layer (PBL); it is a first-order parameterization that extends turbulent mixing throughout the entire boundary layer, including an explicit representation of the upper inversion layer. This scheme is recognized for its effectiveness in simulating atmospheric phenomena across various geographic and climatic conditions [24]. Additionally, the Betts-Miller-Janjic (BMJ) scheme for convection adjusts temperature and humidity profiles toward a targeted state based on observational data, which helps facilitate a realistic portrayal of convective processes. This scheme is particularly suited for operational applications due to its computational efficiency and capability to simulate convection in diverse climatic environments [29].

Simulations were executed on a PowerEdge T320 server with an Intel Xeon 1.8 GHz dual-core processor and 50 GB RAM. Although this configuration offers moderate processing power, it still constrained the number and resolution of simulations. As a result, a representative-date strategy based on K-means clustering was implemented to balance computational feasibility with atmospheric variability coverage.

2.5 Error Metrics

In order to validate the models, the following error metrics were used: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Bias, and the Correlation Coefficient (r). These metrics are widely utilized because each offers complementary insights into the model's error [55]:

- **Root Mean Square Error (RMSE).** Significantly penalizes large errors, making it

effective in highlighting major discrepancies between forecasts and actual data:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[y_{i_{(Pred)}} - y_{i_{(Obs)}}\right]^2}. \quad (1)$$

- **Mean Absolute Error (MAE).** Measures the average error without excessively weighting extreme errors, providing a clear view of the average error:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_{i_{(Pred)}} - y_{i_{(Obs)}}\right|. \quad (2)$$

- **Bias.** Indicates whether the model tends to overestimate or underestimate the variable systematically:

$$Bias = \frac{1}{N}\sum_{i=1}^{N}\left[y_{i_{(Pred)}} - y_{i_{(Obs)}}\right]. \quad (3)$$

- **Correlation Coefficient (r):** Assesses how effectively the model represents the temporal variability of the actual variable. A high value signifies that, despite any magnitude errors, the model aligns with the trend of the observed data:

$$r = \sum_{i=1}^{N}\frac{\left(y_{i_{(\text{Pred})}} - \bar{y}_{(\text{Pred})}\right)\left(y_{i_{(\text{Obs})}} - \bar{y}_{(\text{Obs})}\right)}{\sqrt{\left(y_{i_{(\text{Pred})}} - \bar{y}_{(\text{Pred})}\right)^2}\sqrt{\left(y_{i_{(\text{Obs})}} - \bar{y}_{(\text{Obs})}\right)^2}}. \quad (4)$$

These metrics are appropriate for validating the performance of the WRF model because, together, they enable the assessment of both the magnitude of errors and the model's capability to replicate the variability and trend of wind speed. This step is crucial in numerical forecasting, as demonstrated in this study.

3 Results and Discussion

3.1 Comparison of Configurations

The results obtained from the simulations with the three proposed configurations (WRF_WMK, WRF_TBG, and WRF_MYB) demonstrate significant differences in performance, assessed through RMSE, MAE, bias, and correlation with NASA data (Figure 6). The WRF_TBG configuration (Thompson, BouLac, Grell-Freitas) distinguishes itself by providing accurate estimates for most analyzed dates. In Figure 5, the graphs of the wind speed time series for the proposed forecast dates can be observed; each time series corresponds to records obtained from the NASA database and those derived from the three parameterization schemes of the WRF Model.

Since all forecasts have deviations, it is appropriate to calculate probability bands of occurrence [54] in addition to the predicted values. This is useful because it gives the user an estimate of the best and worst-case scenarios that may arise. Therefore, prediction intervals were calculated for the WRF_TBG model, which exhibited the best performance, to assess the accuracy percentage of the top model within these intervals.

These results demonstrate that WRF_TBG reduces absolute errors and offers more excellent temporal stability, especially in regions with complex topography, like Tepuxtepec. The difference in model performance underscores the importance of carefully selecting parameterizations, as specific combinations of microphysics, boundary layer [5, 35], and convection schemes directly affect the quality of the forecast.

Table 3 compares error metrics for the three WRF model configurations (WRF_WMK, WRF_TBG, and WRF_MYB). WRF_TBG consistently stood out by showing the lowest error values and the highest correlation with NASA data. For example, on January 21, 2024, WRF_TBG achieved an RMSE of 2.12 and a correlation of 0.47, significantly outperforming WRF_WMK, which had an RMSE of 5.15 and a negative correlation (-0.21). This behavior indicates that WRF_TBG provides estimates closer to the actual values and a greater ability to capture wind variability.

On April 11, 2024, WRF_TBG once again emerged as the best option, exhibiting an RMSE of 2.81 and an MAE of 2.33, compared to the higher values of WRF_WMK (RMSE of 3.18 and

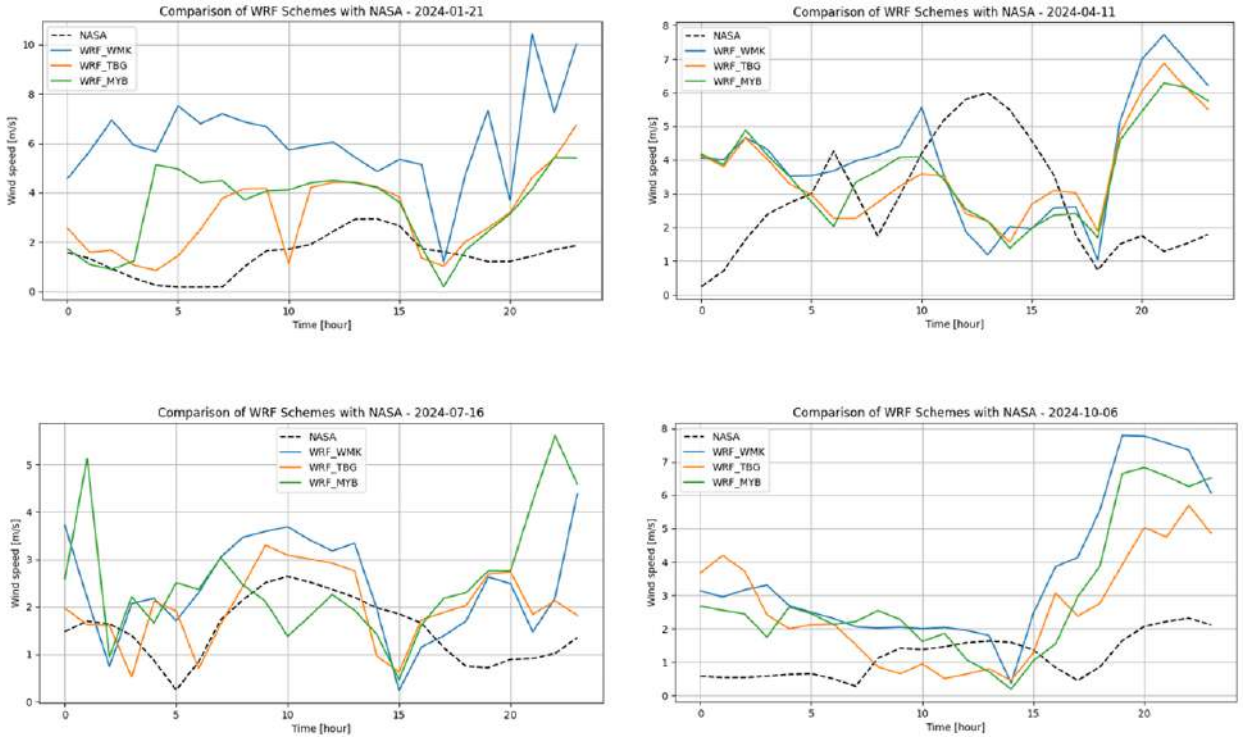

Fig. 5. Hourly wind speed time series comparison between the NASA-POWER database and WRF model outputs using three different parameterization schemes (WRF_WMK, WRF_TBG, and WRF_MYB) for four representative dates. Each subplot corresponds to a specific date selected through K-means clustering, allowing the evaluation of model performance under different seasonal conditions.

MAE of 2.66). This trend persisted through the summer (July 16, 2024), when WRF_TBG recorded an RMSE of 0.95 and an MAE of 0.77, the only configuration with errors below 1. This indicates a superior ability to capture atmospheric conditions during intense convection. Finally, in October, WRF_TBG demonstrated its superiority with an RMSE of 2.05 compared to WRF_WMK's 3.10.

The superiority of WRF_TBG can be explained by the specific characteristics of its parameterizations. Thompson (Microphysics): it enables a more precise representation of the formation and evolution of hydrometeors, enhancing the simulation of clouds and precipitation, which are crucial elements in surface wind prediction. BouLac (Planetary Boundary Layer): it i based on turbulent kinetic energy (TKE), and provides a thorough treatment of turbulence, which is essential in complex terrains like Tepuxtepec. Grell-Freitas (Convection): it effectively represents the transition between resolved and unresolved convection scales, enhancing the fit under variable convection conditions.

The results of WRF_TBG are particularly noteworthy regarding RMSE and correlation. For example, on January 21, 2024, WRF_TBG achieved an RMSE of 2.12 and a correlation of 0.47, compared to 5.15 and -0.21, achieved by WRF_WMK. This trend remained consistent across most of the analyzed seasons, indicating that the proper selection of parameterizations directly affects forecast quality.

3.2 Prediction Intervals (PI)

Prediction intervals are typically based on the Mean Squared Error (MSE) because it estimates the variance of the forecast error [54]. The square

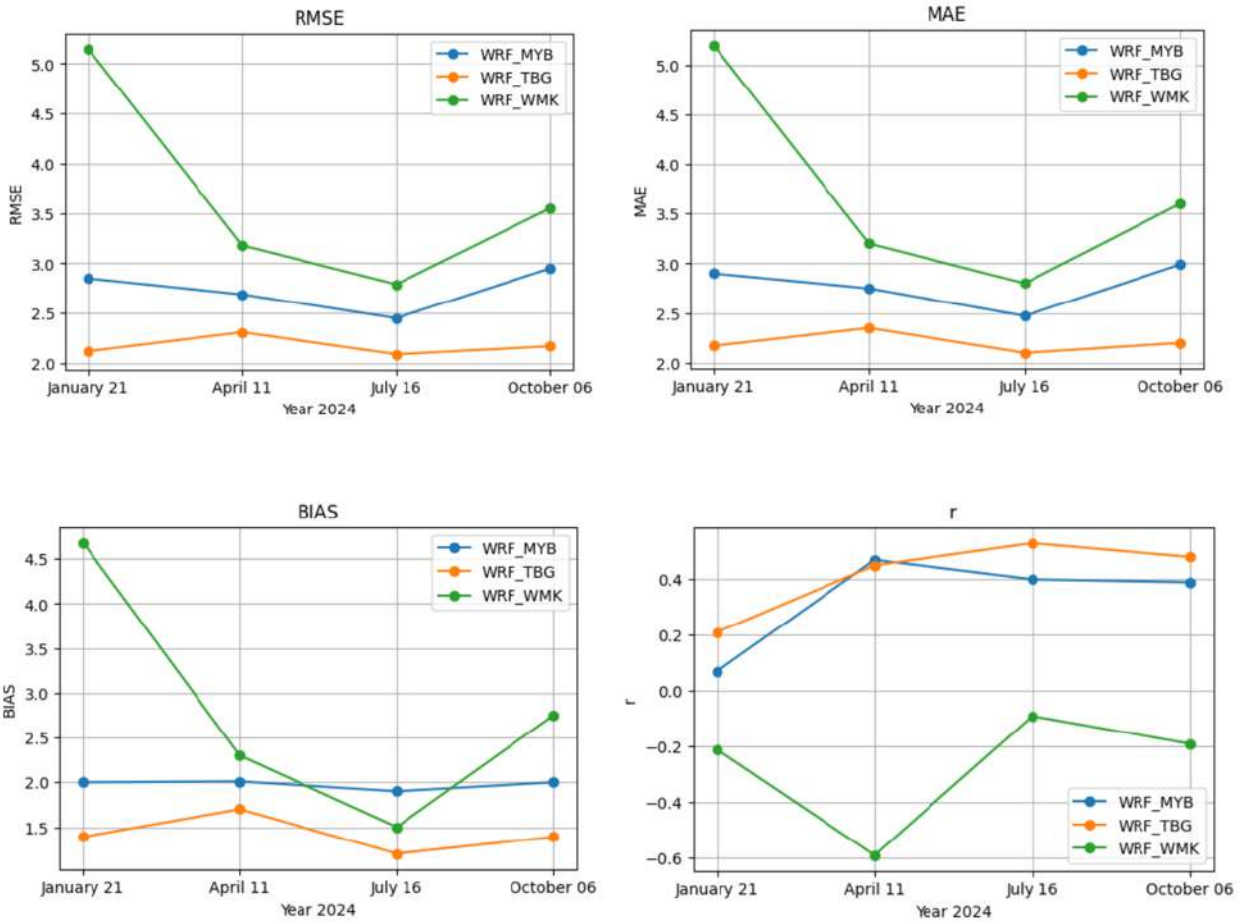

Fig. 6. Evaluation of forecast accuracy using four statistical metrics: RMSE, MAE, BIAS, and Pearson correlation coefficient (r), for each parameterization scheme (WRF_MYB, WRF_TBG, and WRF_WMK) across four representative simulation dates. The graphs illustrate the temporal variation in model performance, highlighting the consistency and relative advantages of each configuration.

root of the MSE provides an estimate of the standard deviation of the forecast error.

The standard assumption for constructing prediction intervals is that the forecast errors are normally distributed with a mean of zero.

Given this assumption, an approximate prediction interval for the subsequent observation is calculated using the following expression:

$$F_{n+h} \pm z \cdot \sqrt{MSE_h}. \tag{5}$$

where n is the index of the last observed point in the time series, h represents the prediction horizon, that is, the number of steps into the future for which the estimate is being made, z is the width and probability of the prediction interval, and MSE_h for multi-step forecasts.

The 80% and 95% prediction intervals provide valuable information about the reliability of the forecast (Figure 7). It was noted that WRF_TBG displayed the least dispersion in its prediction intervals, further affirming its stability and robustness for the study site. The graphs

ISSN 2007-9737

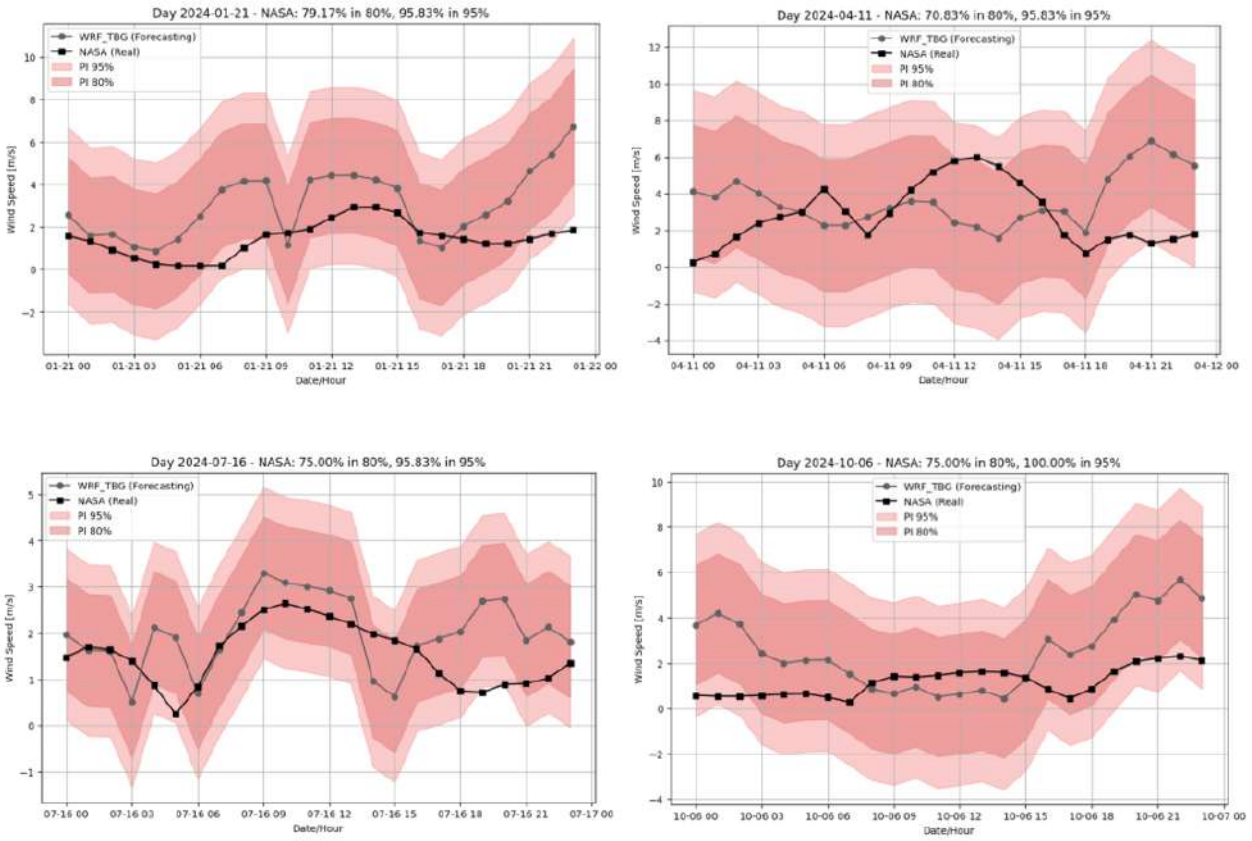

Fig. 7. Hourly wind speed prediction intervals (80% and 95%) generated by the WRF_TBG parameterization for four representative dates. The shaded areas represent the uncertainty bands, and the black lines indicate the NASA-POWER reference data. Coverage percentages show the proportion of NASA values falling within each interval, illustrating the reliability of the forecasts under different seasonal conditions.

created for each season illustrate how the series predicted by WRF_TBG stays within the prediction intervals, particularly during winter and autumn, when uncertainty was significantly reduced.

Overall, it is observed that for the 80% interval, the percentage of observations falling within it ranges approximately from 70% to 80%, while for the 95% interval, the range extends from 95% to 100%. This indicates that assuming the normality of errors and calculating the standard deviation from the MSE, the model effectively captures most of the actual measurements within the proposed prediction intervals, especially in the 95% case.

Although the WRF model can be configured for different forecast horizons, this study focused on 24-hour forecasts to align with the representative-event-based design. Future work will explore shorter and extended forecast windows. This characteristic suggests that the WRF_TBG configuration is accurate in point values and consistent over time.

Beyond technical accuracy, accurate wind speed forecasting improves the economic performance of power systems by reducing imbalance costs, minimizing the need for spinning reserves, and avoiding penalties from deviation between forecasted and generated energy [34, 6]. It also increases revenue opportunities for wind producers through improved market participation and risk management [30], while supporting the broader

Table 3. Forecast performance metrics for each parameterization scheme and simulation date. Values are based on comparisons with NASA-POWER wind speed data.

SCHEMES	RMSE	MAE	BIAS	r
January 21, 2024				
WRF_WMK	5.15	4.72	4.68	-0.21
WRF_TBG	2.12	1.73	1.30	0.47
WRF_MYB	2.57	2.08	1.94	0.07
April 11, 2024				
WRF_WMK	3.18	2.66	1.16	-0.49
WRF_TBG	2.81	2.33	0.78	-0.55
WRF_MYB	2.81	2.40	0.79	-0.56
July 16, 2024				
WRF_WMK	1.32	1.16	0.90	0.41
WRF_TBG	0.95	0.77	0.48	0.34
WRF_MYB	1.85	1.44	0.99	-0.33
October 16, 2024				
WRF_WMK	3.10	2.54	2.44	0.50
WRF_TBG	2.05	1.79	1.33	0.26
WRF_MYB	2.51	2.06	1.79	0.52

integration of renewables and reducing emissions [38].

4 Conclusion and Future Work

The comparative analysis of WRF model configurations shows that choosing parameterization schemes significantly affects forecast accuracy. Among the options evaluated, the WRF_TBG configuration (Thompson, BouLac, Grell-Freitas) was the most suitable for the Tepuxtepec region. It exhibited the lowest RMSE and MAE values and the highest correlation with NASA's observed measurements (MERRA-2, POWER Project). The effectiveness of this configuration stems from the Thompson scheme's capability to model relevant microphysical processes in the region, the BouLac scheme's ability to accurately capture turbulence in the boundary layer, and the Grell-Freitas scheme's adaptability in representing convective processes.

Additionally, using 80% and 95% prediction intervals proved to be a key tool for assessing forecast uncertainty and validating the stability of WRF_TBG under variable atmospheric conditions.

These results are relevant to developing more accurate prediction models and offer essential information for assessing wind potential in the Tepuxtepec region, aiding in the planning of renewable energy generation projects.

The results demonstrated that the WRF_TBG configuration outperformed the other configurations, reducing the RMSE by an average of up to 60% compared to WRF_WMK.

The forecasted values were accurate within the 80% prediction interval for up to 80% of the total values and within the 95% prediction interval for a maximum of 100%. Seasonal evaluation indicated improved model performance in winter and diminished performance in summer, likely due to the impact of intense convection during the latter season.

As potential future research directions, it is suggested that hybrid models that integrate WRF with artificial intelligence techniques be investigated to enhance wind forecast accuracy in areas with complex topography, such as Tepuxtepec site.

Additionally, it is suggested that the model's performance be evaluated over longer forecast horizons and under extreme atmospheric conditions. Regarding practical applications, the results obtained can be crucial in planning and optimizing wind energy projects. They contribute to identifying sites with high energy potential and enhancing Mexico's renewable energy matrix.

Acknowledgments

The authors express their gratitude to CONAHCYT for the financial support provided to Itzagueri García-Rodríguez for his graduate studies and the research scholar grants 324425.

References

1. **Amato, F., Guignard, F., Walch, A., Mohajeri, N., Scartezzini, J.-L., Kanevski, M. (2022).** Spatio-temporal estimation of wind speed and wind power using extreme learning machines: predictions, uncertainty and technical potential. Stochastic Environmental Research and Risk Assessment, Vol. 36, No. 8, pp. 2049–2069. DOI: 10.1007/s00477-022-02219-w.
2. **Argüeso, D., Businger, S. (2018).** Wind power characteristics of oahu, hawaii. Renewable Energy, Vol. 128, pp. 324–336. DOI: 10.1016/j.renene.2018.05.080.
3. **Asociación Empresarial Eólica (AEE) (2024).** La eólica en el mundo. https://aeeolica.org/sobre-la-eolica/la-eolica-en-el-mundo/.
4. **Avella-Rodríguez, F. A. (2017).** Tendencias recientes en el pronóstico de velocidad de viento para generación eólica. Master's thesis, Universidad Nacional de Colombia.
5. **Banks, R. F., Tiana-Alsina, J., Baldasano, J. M., Rocadenbosch, F., Papayannis, A., Solomos, S., Tzanis, C. G. (2016).** Sensitivity of boundary-layer variables to pbl schemes in the wrf model based on surface meteorological observations, lidar, and radiosondes during the hygra-cd campaign. Atmospheric Research, Vol. 176–177, pp. 185–201. DOI: 10.1016/j.atmosres.2016.02.024.
6. **Barthelmie, R. J., Pryor, S. C., Takle, E. S., Kjellström, J. O. T., Vanderwende, J. P., Bach, K. E. (2008).** Wind energy development and its impact on atmospheric boundary layer processes. Bulletin of the American Meteorological Society, Vol. 89, No. 9, pp. 1303–1314. DOI: 10.1175/2008BAMS2397.1.
7. **Bouche, D., Flamary, R., d'Alché Buc, F., Plougonven, R., Clausel, M., Badosa, J., Drobinski, P. (2022).** Wind power predictions from nowcasts to 4-hour forecasts: a learning approach with variable selection. DOI: 10.48550/ARXIV.2204.09362.
8. **Bougeault, P., Lacarrere, P. (1989).** Parameterization of orography-induced turbulence in a mesobeta–scale model. Monthly Weather Review, Vol. 117, No. 8, pp. 1872–1890. DOI: 10.1175/1520-0493(1989)117¡1872:pooiti¿2.0.co;2.
9. **Chen, X., Liu, Y. (2021).** A hybrid model based on wrf and machine learning for short-term wind speed forecasting. Renewable Energy, Vol. 178, pp. 319–335. DOI: 10.1016/j.renene.2021.06.023.
10. **Clean Energy Ministerial (2020).** Grid integration series: variable renewable energy forecasting - scaling up renewable energy project. www.cleanenergyministerial.org/resources-cesc/grid-integration/grid-integration-studies/.
11. **Colegio Nacional de Abogados Municipalistas (2024).** Enciclopedia de los Municipios y Delegaciones de México. www.derechomunicipal.org.mx/single-post/2018/07/24/enciclopedia-de-los-municipios-y-delegaciones-de-méxico.
12. **Cornejo-Bueno, L., López, G., Prieto, L., Salcedo-Sanz, S. (2018).** Hybrid machine learning forecasting of solar radiation values using ensemble methods and wavelet transforms. Energy Conversion and Management, Vol. 162, pp. 1–9. DOI: 10.1016/j.enconman.2018.01.068.
13. **de Estadística y Geografía (INEGI), I. N. (2017).** Conjunto de datos vectoriales de uso del suelo y vegetación, serie vi (escala 1:250 000). https://www.inegi.org.mx/app/biblioteca/ficha.html?upc=889463791711. Escala 1:250 000. México.
14. **Desamsetti, S., Dasari, H. P., Langodan, S., Titi, E. S., Knio, O., Hoteit, I. (2019).** Efficient dynamical downscaling of general circulation models using continuous data assimilation. Quarterly Journal of the Royal Meteorological Society, Vol. 145, No. 724, pp. 3175–3194. DOI: 10.1002/qj.3612.

15. **Doan, Q.-V., Amagasa, T., Pham, T.-H., Sato, T., Chen, F., Kusaka, H. (2023).** Structural k-means (s k-means) and clustering uncertainty evaluation framework (cuef) for mining climate data. Geoscientific Model Development, Vol. 16, No. 8, pp. 2215–2233. DOI: 10.5194/gmd-16-2215-2023.

16. **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, Vol. 146, pp. 66–86. DOI: 10.1016/j.renene.2019.06.133.

17. **Feng, C., Sun, M., Cui, M., Chartan, E. K., Hodge, B.-M., Zhang, J. (2019).** Characterizing forecastability of wind sites in the united states. Renewable Energy, Vol. 133, pp. 1352–1365. DOI: 10.1016/j.renene.2018.08.085.

18. **Fernández-González, S., Martín, M. L., García-Ortega, E., Merino, A., Lorenzana, J., Sánchez, J. L., Valero, F., Rodrigo, J. S. (2018).** Sensitivity analysis of the wrf model: Wind-resource assessment for complex terrain. Journal of Applied Meteorology and Climatology, Vol. 57, No. 3, pp. 733–753. DOI: 10.1175/jamc-d-17-0121.1.

19. **Ferrelli, F., Bustos, M. L., Piccolo, M. C., Huamantinco Cisneros, M. A., Perillo, G. M. E. (2016).** Downscaling de variables climáticas a partir del reanálisis ncep/ncar en el sudoeste de la provincia de buenos aires (argentina). Papeles de Geografía, , No. 62, pp. 21. DOI: 10.6018/geografia/2016/239051.

20. **García-Moya, J. A. (2003).** Los modelos numéricos de predicción del tiempo. **de Meteorología, I. N.**, editor, Libro de ponencias del 1er Encuentro sobre Meteorología y Atmósfera de Canarias, pp. p. 104–106.

21. **Gonçalves, L. d. J. M., Kaiser, J., Palmeira, R. M. d. J., Gallo, M. N., Parente, C. E. (2024).** Evaluation of a high resolution wrf model for southeast brazilian coast: The importance of physical parameterization to wind representation. Atmosphere, Vol. 15, No. 5, pp. 533. DOI: 10.3390/atmos15050533.

22. **Grell, G. A., Freitas, S. R. (2014).** A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. Atmospheric Chemistry and Physics, Vol. 14, No. 10, pp. 5233–5250. DOI: 10.5194/acp-14-5233-2014.

23. **Hong, S., Lim, J.-O. J. (2006).** The wrf single-moment 6-class microphysics scheme (wsm6). Asia-pacific Journal of Atmospheric Sciences, Vol. 42, pp. 129–151.

24. **Hong, S.-Y., Noh, Y., Dudhia, J. (2006).** A new vertical diffusion package with an explicit treatment of entrainment processes. Monthly Weather Review, Vol. 134, No. 9, pp. 2318–2341. DOI: 10.1175/mwr3199.1.

25. **Instituto Nacional de Estadística y Geografía (INEGI) (2024).** Contepec, Estado de Michoacán de Ocampo. https://www.inegi.org.mx/app/biblioteca/ficha. Online; accessed: 10-Apr-2025.

26. **International Energy Agency (IEA) (2024).** World Energy Outlook. www.iea.org/reports/world-energy-outlook-2024/executive-summary.

27. **International Renewable Energy Agency (IRENA) (2024).** Record Growth in Renewables but Progress Needs to be Equitable. www.irena.org/News/pressreleases/2024/Mar/Record-Growth-in-Renewables-but-Progress-Needs-to-be-Equitable-ES.

28. **James F. Manwell, Jon G. McGowan, Anthony L. Rogers (2009).** Wind Energy Explained: Theory, Design and Application. Wiley.

29. **Janjić, Z. I. (1994).** The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. Monthly Weather Review, Vol. 122, No. 5, pp. 927–945. DOI: 10.1175/1520-0493(1994)122¡0927:tsmecm¿2.0.co;2.

30. **Jeon, J.-H., Lee, S.-Y., Lee, H.-J., Roh, J.-H. (2022).** A study on the application of wind

power forecasting for the operation of power systems with high renewable penetration. Energies, Vol. 15, No. 4, pp. 1281. DOI: 10.3390/en15041281. Accessed: 2025-04-10.

31. **Jr., P. W. S., Westberg, D., White, M. A., Zhang, T., Cox, S. J., Lee, S.-H., Wei, C., Fan, Y. (2018).** Nasa power project: Improving meteorological data for renewable energy applications. Energy Procedia, Vol. 149, pp. 423–430. DOI: 10.1016/j.egypro.2018.08.195. Presented at the 7th International Conference on Renewable Energy Research and Applications (ICRERA).
32. **Kain, J. S. (2004).** The kain–fritsch convective parameterization: An update. Journal of Applied Meteorology (1988-2005), Vol. 43, No. 1, pp. 170–181.
33. **Lee, K., Park, B., Kim, J., Hong, J. (2024).** Day-ahead wind power forecasting based on feature extraction integrating vertical layer wind characteristics in complex terrain. Energy, Vol. 288, pp. 129713. DOI: 10.1016/j.energy.2023.129713.
34. **Lew, D., Brinkman, G., Draxl, E., Freedman, J., Zhang, Y., Bloom, A. (2011).** The value of wind power forecasting. Technical Report TP-5500-55530, National Renewable Energy Laboratory (NREL). Accessed: 2025-04-10.
35. **Mantovani Júnior, J. A., Aravéquia, J. A., Carneiro, R. G., Fisch, G. (2023).** Evaluation of pbl parameterization schemes in wrf model predictions during the dry season of the central amazon basin. Atmosphere, Vol. 14, No. 5, pp. 850. DOI: 10.3390/atmos14050850.
36. **Montecinos-Geisse, S. (2021).** Pronóstico de generación del recurso eólico y solar para el despacho de carga. Technical report, Deutsche Gesellschaft für.
37. **Morrison, H., Thompson, G., Tatarskii, V. (2009).** Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. Monthly Weather Review, Vol. 137, No. 3, pp. 991–1007. DOI: 10.1175/2008mwr2556.1.
38. **Muftahov, R., Marchenko, A., Sidorov, D. (2016).** Economic efficiency of wind power forecasting in the context of energy market integration. Energy Procedia, Vol. 100, pp. 117–122. DOI: 10.1016/j.egypro.2016.10.153. Accessed: 2025-04-10.
39. **Nakanishi, M., Niino, H. (2004).** An improved mellor-yamada level-3 model with condensation physics: Its design and verification. Boundary-Layer Meteorology, Vol. 112, No. 1, pp. 1–31. DOI: 10.1023/B:BOUN.0000020164.04146.98.
40. **National Aeronautics and Space Administration (NASA) (2024).** Langley Research Center (LaRC) Prediction Of Worldwide Energy Resources (POWER). https://power.larc.nasa.gov/data-access-viewer/.
41. **National Centers For Environmental Prediction/National Weather Service/NOAA/U.S. Department Of Commerce (2015).** Ncep gfs 0.25 degree global forecast grids historical archive. DOI: 10.5065/D65D8PWK.
42. **NCAR (2024).** Weather Research & Forecasting Model (WRF). https://www.mmm.ucar.edu/models/wrf.
43. **NOAA (2021).** Numerical weather prediction. http://www.ncdc.noaa.gov/data-access/model-data/model-datasets/numerical-weather-prediction. Accessed: 2025-04-10.
44. **Perumpalot, V., Drisya, G. V., Kumar, K. S. (2018).** Cross-location wind speed forecasting for wind energy applications using machine learning based models. DOI: 10.48550/ARXIV.1808.03480.
45. **Remedia.bio (2024).** Top 10 de países con más energía eólica instalada. https://remedia.bio/index.php/2024/09/23/conoce-el-top-10-de-paises-con-mas-energia-eolica-instalada/.
46. **Roque Rodríguez, A. E., Ferrer Hernández, A., Borrajero Montejo, I., Sierra Lorenzo, M. (2018).** Elaboración de pronóstico energético a corto plazo para parques eólicos. Ingeniería Energética, Vol. 39, pp. 115 – 122.

47. **Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Wang, W., Huang, X.-Y., Duda, M. (2008).** A description of the advanced research wrf version 3. Technical report. DOI: 10.5065/D68S4MVH.

48. **Solano-Farias, F., García-Valdecasas Ojeda, M., Donaire-Montaño, D., Rosa-Cánovas, J. J., Castro-Díez, Y., Esteban-Parra, M. J., Gámiz-Fortis, S. R. (2024).** Assessment of physical schemes for wrf model in convection-permitting mode over southern iberian peninsula. Atmospheric Research, Vol. 299, pp. 107175. DOI: 10.1016/j.atmosres.2023.107175.

49. **Stolz, W. (2008).** Predicción del clima y modelos numéricos. Revista de Ciencias Ambientales, Vol. 35, No. 1, pp. 34. DOI: 10.15359/rca.35-1.7.

50. **Thompson, G., Field, P. R., Rasmussen, R. M., Hall, W. D. (2008).** Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. part ii: Implementation of a new snow parameterization. Monthly Weather Review, Vol. 136, No. 12, pp. 5095–5115. DOI: 10.1175/2008mwr2387.1.

51. **Vincent, C. L., Hahmann, A. N. (2015).** The impact of grid and spectral nudging on the variance of the near-surface wind speed. Journal of Applied Meteorology and Climatology, Vol. 54, No. 5, pp. 1021–1038. DOI: 10.1175/jamc-d-14-0047.1.

52. **Weather Research and Forecasting (WRF) (2024).** Source codes and graphics software downloads. www.mmm.ucar.edu/wrf/users/download/.

53. **Weather Research and Forecasting (WRF) (2024).** Wrf model version 4.0. updates. https://www2.mmm.ucar.edu/wrf/users/wrf_files/wrfv4.0/updates-4.0.

54. **Wheelwright, S., Makridakis, S., Hyndman, R. J. (1998).** Forecasting: Methods and applications, 3rd edition.

55. **Willmott, C. J., Matsuura, K. (2005).** Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. Climate Research, Vol. 30, No. 1, pp. 79–82.

56. **Xataka (2024).** Gráfico de los países con más energía eólica instalada en el mundo. www.xataka.com/energia/grafico-paises-energia-eolica-instalada-mundo-espana-sale-bastante-bien-parada.

57. **Xie, B., Li, Y., Zervos, A., Wang, S., Liu, W. (2020).** Evaluation of nasa power reanalysis meteorological data for wind resource assessment in china. Energies, Vol. 13, No. 3, pp. 620. DOI: 10.3390/en13030620.

58. **Zhou, S., Liu, D., Zhu, M., Tang, W., Chi, Q., Ye, S., Xu, S., Cui, Y. (2022).** Temporal and spatial variation of land surface temperature and its driving factors in zhengzhou city in china from 2005 to 2020. Remote Sensing, Vol. 14, No. 17, pp. 4281. DOI: 10.3390/rs14174281.

Article received on 05/03/2025; accepted on 16/06/2025.
**Corresponding author is Alma Rosa Mendez-Gordillo.*