



Universidad Michoacana de San Nicolás de Hidalgo  
**Facultad de Ingeniería Mecánica**  
**Doctorado en Ciencias en Ingeniería Mecánica**

Tesis:

**Modelo híbrido WRF-SARIMA para mejorar  
los pronósticos de la velocidad del viento**

Para obtener el grado de:  
**Doctora en Ciencias en Ingeniería Mecánica**

Presenta:

**Maritza Talhia Bernabé Morales**

Asesor:

**Dr. Erasmo Cadenas Calderón**

Co-asesor:

**Dr. Víctor López Garza**

Morelia, Michoacán, mayo de 2026



**Ciencia y  
Tecnología**  
Secretaría de Ciencia, Humanidades,  
Tecnología e Innovación



# Agradecimientos

*Al Dr. Erasmo, por haber creído en mi, por su guía, paciencia y por la confianza que depositó en mi persona durante estos seis años.*

*A la Dra. Erika Danaé, por su amabilidad, tiempo y paciencia en el tiempo que estuve bajo su tutela.*

*A mi mamá y a mi hermana por su cariño y apoyo incondicional, por acompañarme en momentos difíciles y celebrar conmigo cada logro.*

*A las y los profesores de la Facultad de Ingeniería Mecánica, por su dedicación al brindarme las herramientas académicas para mi desarrollo profesional.*

*Al Instituto de Ciencias de la Atmósfera y Cambio Climático UNAM, por las facilidades brindadas para la realización de esta investigación.*

*A mis compañeras y compañeros del posgrado.*

*Al SECIHTI por la beca No.1009386.*

## AGRADECIMIENTOS

---

*Dedicado con especial cariño a mi abuelita Hermelinda,  
quien no tuvo la oportunidad de asistir a la escuela  
y a quien le hubiera enorgullecido este logro.*

*Finalmente, no dejemos de hablar de Gaza.  
Free Palestine*

# Índice general

|  |          |
|--|----------|
| <b>Resumen</b>   | VII      |
| <b>Abstrac</b>   | IX       |
| <b>Introduccion</b>  | XII      |
| <b>Objetivo General</b>  | XV       |
| Objetivos Específicos . . . . .  | XV       |
| <b>Hipótesis</b>   | XVI      |
| <b>Justificación</b>   | XVII     |
| <b>1. Fundamentos de la energía eólica y el pronóstico de la velocidad del viento</b>            | <b>1</b> |
| 1.1. La energía eólica como fuente renovable . . . . .   | 1        |
| 1.2. Fundamentos del viento . . . . .  | 1        |
| 1.3. Aerogeneradores . . . . .   | 3        |
| 1.4. Potencial eólico . . . . .  | 3        |
| 1.5. Horizontes de pronóstico de la velocidad del viento . . . . .                               | 6        |
| 1.6. Legislación en México . . . . .   | 7        |
| 1.7. ¿Por qué pronosticar la velocidad del viento? . . . . .                                     | 7        |
| 1.8. Modelos estadísticos: Modelos Box-Jenkins para series de tiempo . . .                       | 10       |
| 1.8.1. Autoregressive Moving Average (ARMA) . . . . .  | 10       |
| 1.8.2. Autoregressive Integrated Moving Average (ARIMA) . . . . .                                | 11       |
| 1.8.3. Seasonal Autoregressive Integrated Moving Average (SARIMA)                                | 12       |
| 1.8.4. Función de autocorrelación (ACF) y función de autocorrelación<br>parcial (PACF) . . . . . | 13       |
| 1.8.5. Fases de la metodología Box-Jenkins . . . . .   | 14       |
| 1.9. Modelos físicos: Modelo Weather Research and Forecasting (WRF) . .                          | 16       |
| 1.10. Métricas de desempeño del modelo . . . . .   | 18       |

|  |           |
|--|-----------|
| <b>2. Análisis de la literatura sobre modelos de pronósticos de la velocidad del viento</b>          | <b>21</b> |
| 2.1. Clasificación de los modelos de pronóstico de la velocidad del viento . . . . .                 | 21        |
| 2.2. Numeralia sobre publicaciones . . . . .   | 24        |
| 2.3. Análisis crítico de enfoques basados en SARIMA y WRF . . . . .                                  | 27        |
| 2.3.1. Modelos ARIMA y SARIMA . . . . .  | 28        |
| 2.3.2. Modelos físicos reportados . . . . .  | 32        |
| 2.3.3. Modelos híbridos para la corrección de errores . . . . .                                      | 33        |
| 2.4. Identificación de la brecha de conocimiento . . . . .   | 34        |
| <b>3. Metodología del modelo híbrido WRF-SARIMA para la corrección de errores de pronóstico</b>      | <b>37</b> |
| 3.1. Área de estudio y datos de mediciones en el sitio . . . . .                                     | 37        |
| 3.2. Pronósticos del modelo WRF . . . . .  | 40        |
| 3.3. Modelo Híbrido WRF-SARIMA . . . . .   | 44        |
| 3.3.1. Fase I. Obtención de la serie de tiempo de los errores de pronóstico del modelo WRF . . . . . | 44        |
| 3.3.2. Fase II. Análisis de los errores de pronóstico . . . . .                                      | 47        |
| 3.3.3. Fase III. Modelado y pronóstico con SARIMA . . . . .  | 49        |
| 3.3.4. Fase IV. Combinación de pronósticos . . . . .   | 51        |
| 3.3.5. Fase V. Evaluación del modelo . . . . .   | 52        |
| <b>4. Resultados y discusiones</b>   | <b>53</b> |
| 4.1. Descripción de los experimentos . . . . .   | 53        |
| 4.2. Análisis de errores de pronóstico . . . . .   | 55        |
| 4.3. Modelado y pronóstico SARIMA . . . . .  | 60        |
| 4.4. Combinación de pronósticos . . . . .  | 62        |
| 4.5. Evaluación del modelo . . . . .   | 64        |
| 4.6. Discusiones . . . . .   | 67        |
| 4.6.1. Limitaciones asociadas a los datos . . . . .  | 67        |
| 4.6.2. Generalización del modelo híbrido . . . . .   | 68        |
| 4.6.3. Comparación con otros modelos . . . . .   | 68        |
| 4.6.4. Implicaciones operativas para la generación eólica . . . . .                                  | 69        |
| 4.6.5. Limitaciones del enfoque SARIMA . . . . .   | 69        |
| <b>Conclusiones y trabajo futuro</b>   | <b>71</b> |
| 4.7. Trabajo futuro . . . . .  | 72        |
| <b>Artículo publicado</b>  | <b>74</b> |

# Índice de figuras

|   |    |
|---|----|
| 1.1. Elementos de un aerogenerador. . . . .   | 4  |
| 1.2. Elementos internos de un aerogenerador. Tomado de <i>Energía eólica</i> por Talayero, A., et.al. 2011. Prensas de la Universidad de Zaragoza. España. . . . .                          | 5  |
| 1.3. Curva de potencia de un aerogenerador de 5 MW de NREL. . . . .   | 6  |
| 1.4. Sección 9.10 de las Bases del Mercado Eléctrico que regula los pronósticos de generación intermitente. . . . .   | 8  |
| 1.5. Sección 3.1 del Manual de Pronósticos que regula los pronósticos de generación intermitente. . . . .   | 9  |
| 1.6. Diagrama de flujo de la metodología Box-Jenkins para el modelado de series de tiempo. Adaptado de <i>Forecasting Methods and Applications</i> (p.314), by S. Makridakis, 1998. . . . . | 15 |
| 1.7. Diagrama del funcionamiento del modelo WRF. . . . .  | 18 |
| 2.1. Búsqueda de las palabras <i>wind speed forecast</i> en la base de datos Scopus sin utilizar filtros. . . . .   | 24 |
| 2.2. Artículos científicos relacionados con los pronósticos de la velocidad del viento publicados entre 1990-2024 . . . . .   | 25 |
| 2.3. Artículos publicados entre 1990-2024 específicamente sobre modelos de pronóstico de la velocidad del viento. . . . .   | 26 |
| 2.4. Artículos publicados entre 2009-2024 sobre modelos de pronóstico de la velocidad del viento que incluyen un horizonte de pronóstico de 24 horas. . . . .                               | 27 |
| 3.1. Localización de La Ventosa, Oaxaca. . . . .  | 38 |
| 3.2. Velocidad del viento en La Ventosa. . . . .  | 38 |
| 3.3. Captura de pantalla de la pagina web del Atlas Eólico Mexicano del INEEL. . . . .  | 39 |
| 3.4. Distribuciones de probabilidad del la rapidez del viento en La Ventosa. . . . .  | 40 |
| 3.5. Fotografía de la visualización del modelo WRF al interior del Instituto de Ciencias de la Atmósfera y Cambio Climático UNAM. . . . .   | 41 |
| 3.6. Interpolación bilineal de la velocidad del viento. . . . .   | 43 |

## ÍNDICE DE FIGURAS

---

|   |    |
|---|----|
| 3.7. Diagrama de la metodología aplicada. . . . .   | 45 |
| 3.8. Diagrama de flujo del modelo WRF-SARIMA. . . . .   | 46 |
| 3.9. Diagrama de flujo del proceso de modelado y pronósticos con el modelo SARIMA. . . . .                              | 49 |
| 4.1. Series de tiempo de la rapidez del viento de las mediciones y de las simulaciones WRF. . . . .                     | 54 |
| 4.2. Serie de tiempo $N_t$ de los errores de pronóstico. . . . .  | 56 |
| 4.3. Histograma y gráfica Q-Q plot de la serie de tiempo $N_t$ . . . . .  | 57 |
| 4.4. Función de autocorrelación (ACF) de la serie de tiempo $N_t$ . . . . .   | 58 |
| 4.5. Gráfica de los errores estandarizados contra los valores pronosticados. . . . .                                    | 59 |
| 4.6. Verificación de residuos del modelo SARIMA con mejor ajuste, que exhibe un comportamiento de ruido blanco. . . . . | 62 |
| 4.7. Pronóstico e intervalos de predicción de la serie de tiempo $N_t$ . . . . .  | 63 |
| 4.8. Pronóstico de 24 horas de la velocidad del viento. . . . .   | 64 |
| 4.9. Comparación de los residuales de pronóstico del modelo WRF y del modelo WRF-SARIMA. . . . .                        | 66 |
| 4.10. Portada del artículo <i>Hybrid WRF-SARIMA model to improve day-ahead wind speed forecast accuracy</i> . . . . .   | 75 |

# Resumen

Los pronósticos de la velocidad del viento son esenciales para el crecimiento y la consolidación de la energía eólica. Debido a su conexión con los mercados eléctricos, los pronósticos con 24 horas de antelación adquieren una relevancia particular. En específico, el modelo Weather Research and Forecasting (WRF) ha ido ganando popularidad como el modelo seleccionado para realizar pronósticos de la velocidad del viento con horizontes de al menos un día, razón por la cual existe un gran interés en mejorar sus pronósticos. Este trabajo propone un modelo híbrido de pronóstico de la velocidad del viento para un horizonte de 24 horas, que combina el modelo WRF con el modelo Autorregresivo Integrado de Media Móvil Estacional (SARIMA). Usando la metodología SARIMA, se logra mejorar los pronósticos del modelo WRF al identificar autocorrelaciones significativas en los errores de pronóstico. El estudio se realizó en La Ventosa, Oaxaca, una región donde se han desarrollado importantes proyectos eólicos. La evaluación del desempeño del modelo se realizó usando la métricas del error absoluto medio (MAE), el error cuadrático medio (MSE) y la raíz del error cuadrático medio (RMSE). Los resultados mostraron que el modelo híbrido WRF-SARIMA presentó un desempeño superior al del modelo WRF. Los errores de pronóstico se redujeron entre un 29 % y un 45 % en la métrica MAE, entre un 40 % y un 67 % en el MSE y entre un 22 % y un 43 % en el RMSE. El modelo híbrido propuesto mejora los pronósticos de la velocidad del viento para un horizonte de 24 horas, aprovechando las ventajas del modelo WRF para pronósticos de mediano plazo e incorporando la interpretabilidad y la robustez del modelo SARIMA, teniendo en consideración aspectos como la automatización y el costo computacional, importantes en el contexto operativo de la gestión de la energía eólica.

*Palabras clave:* Energía, Eólica, WRF, SARIMA, Pronóstico.

# Abstrac

Wind speed forecasts are essential for the growth and consolidation of wind energy. Their connection with electricity markets makes 24-hour-ahead forecasts particularly relevant. The Weather Research and Forecasting (WRF) model has become increasingly popular for wind speed forecasts at horizons of one day or more, therefore, there is significant interest in improving its forecasts. This work proposes a hybrid wind speed forecasting model for a 24-hour horizon. It combines the WRF model with the Seasonal Autoregressive Integrated Moving Average (SARIMA) model. Using the SARIMA methodology, significant autocorrelations can be identified in forecast errors, allowing for improved WRF model forecasts. The study was conducted in La Ventosa, Oaxaca, a region with significant wind power projects. Model performance was evaluated using mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE) metrics. The results showed that the hybrid WRF-SARIMA model outperformed the WRF model alone. Forecast errors were reduced by between 29 % and 45 % for the MAE metric, between 40 % and 67 % for the MSE, and between 22 % and 43 % for the RMSE. The proposed hybrid model improves 24-hour wind speed forecasts by leveraging the WRF model's strengths for medium-term forecasts and incorporating the interpretability and robustness of the SARIMA model, while accounting for automation and computational cost, which are important in the operational context of wind energy management.

*Keywords:* Wind, Energy, WRF, SARIMA, Forecasting.

# Nomenclatura

Tabla 1: Nomenclatura de siglas y modelos

| Sigla    | Significado  |
|----------|--|
| ACF      | Autocorrelation Function   |
| ARIMA    | Autoregressive Integrated Moving Average                                       |
| ARMA     | Autoregressive Moving Average  |
| BPNN     | Backpropagation Neural Network   |
| CENACE   | Centro Nacional de Control de Energía  |
| CFE      | Comisión Federal de Electricidad   |
| DBN      | Deep Belief Network  |
| DL       | Deep Learning  |
| ECMWF    | European Centre for Medium-Range Weather Forecasts                             |
| ELM      | Extreme Learning Machine   |
| EMD      | Empirical Mode Decomposition   |
| FNN      | Feedforward Neural Network   |
| FOA      | Fruit Fly Optimization Algorithm   |
| GFS      | Global Forecast System   |
| GW       | Gigawatt   |
| ICEEMDAM | Improved Complete Ensemble Empirical Mode<br>Decomposition with Adaptive Noise |

Tabla 2: Nomenclatura de siglas y modelos

| Sigla  | Significado                                       |
|--------|---|
| LGB    | Light Gradient Boosting                           |
| LR     | Linear Regression                                 |
| LSSVM  | Least Squares Support Vector Machine              |
| LSTM   | Long Short-Term Memory                            |
| MAE    | Mean Absolute Error                               |
| ML     | Machine Learning                                  |
| MSE    | Mean Squared Error                                |
| NWP    | Numerical Weather Prediction                      |
| PACF   | Partial Autocorrelation Function                  |
| RF     | Random Forest                                     |
| RBFNN  | Radial Basis Function Neural Network              |
| RMSE   | Root Mean Squared Error                           |
| SARIMA | Seasonal Autoregressive Integrated Moving Average |
| SSA    | Singular Spectrum Analysis                        |
| SVM    | Support Vector Machine                            |
| SVR    | Support Vector Regression                         |
| TCN    | Temporal Convolutional Network                    |
| TFT    | Temporal Fusion Transformer                       |
| UTC    | Coordinated Universal Time                        |
| VMD    | Variational Mode Decomposition                    |
| WRF    | Weather Research and Forecasting Model            |
| WT     | Wavelet Transform                                 |

# Introducción

Existe una estrecha relación entre economía, tecnología y energía que no es tan visible para el ojo público. Cada actividad en la naturaleza requiere de energía, y esta es tan fundamental que su disponibilidad determina los límites del desarrollo tecnológico y de las actividades humanas. Un crecimiento económico y tecnológico también implica un aumento de la demanda de energía. Sin embargo, las fuentes de energía tradicionales, como los combustibles fósiles, son costosas tanto económicamente como ecológicamente, además de ser rápidamente agotables, con consecuencias negativas graves tales como el cambio climático (Hagens, 2020). Todo lo anterior obliga a adoptar una posición crítica frente a las fuentes de energía altamente contaminantes y, al mismo tiempo, a una transición más rápida hacia las energías limpias. Entre las energías renovables, las energías eólica y solar son de las más atractivas. La energía eólica, en particular, ha mostrado un rápido crecimiento a nivel global en los últimos años. En 2024, 117 GW de nueva capacidad instalada de energía eólica fueron conectados a las redes eléctricas del mundo, elevando la capacidad eólica total instalada a 1,136 GW (Zhao, 2025). Un contraste importante con la capacidad instalada de 60.4 GW en el año 2019, que mostró una capacidad acumulada global de 651 GW ese mismo año (GWEC, 2020).

En México, la energía eólica y la energía solar son las energías que más han crecido, pasando de una generación de energía eólica de 10,456 GW en 2017 a 20,700 GW en 2023, representando 5.97% de la generación total de electricidad en México en 2023 (SENER, 2024). Desde 2017, la Secretaría de Energía de México establece las regulaciones de pronósticos para los mercados de electricidad, donde todos los días se establecen los pronósticos de generación de energía eléctrica para las próximas 24 horas (GOBMEX, 2017).

Particularmente, la energía eólica difiere de las fuentes de energía tradicionales en el sentido de que esta es determinada por la velocidad del viento, el cual varía tanto en el tiempo como en el espacio. Esta naturaleza variable e intermitente del viento es uno de los principales problemas que enfrentan los operadores de las redes eléctricas.

cas. Estas irregularidades pueden poner en riesgo el correcto funcionamiento de la red, especialmente con un incremento en el porcentaje de energía eólica dentro de la misma. Es en este contexto donde los pronósticos de la velocidad del viento desempeñan un papel importante en el proceso de integración de la energía eólica dentro de las redes eléctricas convencionales. Un pronóstico confiable ayuda a reducir la incertidumbre en la producción de energía eólica y permite a los operadores planear con anticipación las variaciones de la misma.

Los pronósticos de velocidad del viento abarcan varios horizontes temporales: 1) pronóstico de segundos o minutos para el mantenimiento y control de unidades, 2) de 1 a 6 horas para la operación de pequeños sistemas eléctricos, 3) de 1 a 72 horas para la operación de sistemas eléctricos interconectados, y 4) de 1 a 7 días para la planificación del mantenimiento (El-Fouly et al., 2008). Los pronósticos de 24 horas son fundamentales por su conexión con los mercados eléctricos y por su apoyo a la toma de decisiones en los sistemas eléctricos.

Los modelos de pronóstico de la velocidad del viento se pueden dividir en dos grandes categorías: modelos estadísticos, que analizan datos históricos para generar pronósticos, y modelos físicos, que derivan pronósticos a partir de modelos numéricos de predicción del tiempo. En los últimos años, la literatura ha mostrado un predominio de los modelos híbridos sobre los modelos individuales, lo que indica que la combinación de múltiples técnicas de pronóstico puede mejorar significativamente la precisión (Makridakis, 1989). Los modelos híbridos abarcan diversas combinaciones de modelos, como la combinación entre modelos estadísticos, así como combinaciones de un modelo físico con técnicas estadísticas.

Los pronósticos de 24 horas, importantes para los mercados eléctricos, se generan principalmente mediante modelos físicos, también conocidos como modelos de predicción numérica del tiempo (NWP), que resuelven numéricamente las ecuaciones de conservación de masa, energía y momento que describen los procesos atmosféricos. El modelo más ampliamente utilizado actualmente es el modelo Weather Research and Forecasting (WRF).

Este trabajo propone un modelo híbrido WRF-SARIMA para pronósticos de velocidad del viento con 24 horas de anticipación. La principal contribución es la implementación de una estrategia de corrección de errores mediante con enfoque interpretable y automatizable, validado en condiciones reales. En un principio, la metodología se enfoca en identificar dependencias secuenciales o patrones existentes en los errores de pronóstico del modelo WRF. Las series temporales con componentes esta-

cionales y de tendencia pueden analizarse eficazmente mediante modelos SARIMA. Estos modelos suelen ser más rápidos y fáciles de implementar que alternativas más complejas, como los enfoques basados en el aprendizaje automático. En la presente metodología, los modelos SARIMA se utilizan para predecir los errores de pronóstico e incorporarlos al pronóstico final.

Para evaluar el desempeño del modelo híbrido se utilizaron diferentes métricas de error: el error absoluto medio (MAE), el error cuadrático medio (MSE) y la raíz del error cuadrático medio (RMSE).

# Objetivo General

**Mejorar los pronósticos de la velocidad del viento del modelo Weather Research and Forecasting mediante el desarrollo de un modelo híbrido WRF-SARIMA.**

## Objetivos Específicos

Para la realización del presente trabajo se establecieron los siguientes objetivos específicos:

- Regionalizar el modelo Weather Research and Forecasting (WRF) para adaptarlo al sitio seleccionado.
- Analizar los errores de pronóstico de la velocidad del viento generados por el modelo WRF con la finalidad de extraer patrones remanentes.
- Generar un modelo híbrido mediante el uso del modelo SARIMA como una técnica de corrección de error de pronóstico del modelo WRF.
- Validar el modelo híbrido WRF-SARIMA para evaluar su desempeño.
- Comparar el desempeño del modelo híbrido contra el modelo WRF.

# Hipótesis

**El modelo híbrido WRF – SARIMA, que combina el modelo físico Weather Research and Forecasting con el modelo estadístico Seasonal Autoregressive Integrated Moving Average, puede mostrar un mejor desempeño para el pronóstico de la velocidad del viento en comparación con los pronósticos obtenidos únicamente del modelo WRF.**

# Justificación

La energía es una pieza fundamental en el desarrollo de nuestra civilización. Cada año aumenta la demanda y, por lo tanto, la generación de energía a nivel mundial. Pero este aumento no es inocuo, sino que conlleva un costo, con consecuencias serias como lo es el cambio climático y los altos niveles de contaminación derivados del uso de combustibles fósiles. Este panorama ha obligado a replantear nuestra forma de obtención de energía, encontrando en las energías renovables una opción más favorable, tanto ecológicamente como económicamente.

Actualmente vivimos una transición energética que involucra un cambios en todos los sectores relacionados con la industria energética. De acuerdo con la Agencia Internacional de Energía (IEA, por sus siglas en inglés), las energías renovables es el sector, dentro del rubro de la energía, en el que más dinero se ha invertido en los últimos años, seguido solamente por el sector de distribución y almacenamiento. Muy atrás queda el sector fósil y el sector nuclear. Todo lo anterior en concordancia con las redes eléctricas del futuro, o las 5 Ds de la red eléctrica, las cuales hacen referencia a las tendencias de descarbonización, descentralización, diversificación, democratización y digitalización.

Dentro de las energías renovables, la energía solar y eólica son las que resultan más atractivas. A nivel mundial, la energía eólica ha superado el terawatt de capacidad instalada total, con números que siguen en aumento cada año. En el caso de México, actualmente se contabilizan 70 parques eólicos con alrededor de 3,247 aerogeneradores distribuidos a lo largo del territorio nacional. En este contexto, ¿por qué entonces se debería prestar atención a los pronósticos de la velocidad del viento?

La energía eólica depende de la velocidad del viento, la cual varía tanto en tiempo como en espacio. Esto nos genera dos problemáticas, la primera de ellas es la intermitencia, es decir, que se cuentan con periodos de generación de energía eólica seguidos por periodos de inactividad y por lo tanto de nula generación. El segundo es la variabilidad, lo cual se explica debido a las variaciones en las velocidades del viento a lo largo de un periodo de tiempo, lo que deriva en una variación de

generación de energía eólica. Estos dos problemas pueden poner en riesgo todo el sistema de distribución de energía, especialmente con un aumento en el porcentaje de energía eólica dentro de la red.

Es en este punto donde los pronósticos de la velocidad del viento cobran especial relevancia, ya que un pronóstico confiable de la velocidad del viento ayuda a reducir la incertidumbre en la producción de energía eólica y ayuda a planear con anticipación las variaciones de la misma.

Una vez establecida la importancia de los pronósticos para la generación de energía, es necesario señalar la relevancia de ciertos horizontes de pronóstico sobre otros. México, al igual que otras naciones, cuenta con mercados de compra-venta de energía. Nos centraremos en específico en el Mercado del Día en Adelanto, donde se realizan ofertas de energía para cada hora del día siguiente. Esta es la razón principal de que se haya seleccionado un horizonte de pronóstico de 24 horas adelante para los pronósticos de la velocidad del viento.

Aunque existen un abanico de modelos de pronóstico, para este tipo de horizontes de 24 horas se suelen utilizar modelos físicos, los cuales basan sus pronósticos en un análisis de las condiciones atmosféricas. Para esta investigación se decidió utilizar el modelo físico Weather Research and Forecasting (WRF). Adicionalmente, se decidió utilizar modelos Estacionales Autorregresivos de Medias Mviles (SARIMA) como una estrategia para mejorar el desempeño del modelo WRF, generándose como producto final un modelo híbrido WRF-SARIMA, no reportando con anterioridad en la literatura.

# Capítulo 1

## Fundamentos de la energía eólica y el pronóstico de la velocidad del viento

### 1.1. La energía eólica como fuente renovable

La energía eólica forma parte de las energías limpias o renovables, en algunas ocasiones estas energías son también llamadas energías intermitentes, por sus características de generación de electricidad, aunque este es un concepto aún en discusión. La energía eólica ha mostrado un crecimiento importante en los últimos años, siendo uno de los sectores que mayor incremento de inversión ha recibido. La energía eólica ha demostrado ser una solución viable para la obtención de energía, buscando mitigar los efectos del cambio climático.

El convertir la energía cinética del viento en electricidad es posible mediante el uso de aerogeneradores. El viento mueve las aspas del aerogenerador. En el rotor, mediante varios procesos, ocurre la generación de electricidad, la cual es enviada en primera instancia a una subestación, y finalmente a las redes de transmisión eléctrica. Los aerogeneradores pueden ser colocados tanto en tierra como en mar. En tierra se privilegia zonas con altas velocidades del viento y poca turbulencia. Mientras que en el mar se busca vientos fuertes y constantes.

### 1.2. Fundamentos del viento

Para lograr un buen entendimiento de la energía eólica es necesario conocer las características del viento. Los vientos son esencialmente movimientos de masa de aire, los cuales pueden ser un fenómeno a escala global o a escala regional. Los fenómenos

## CAPÍTULO 1. FUNDAMENTOS DE LA ENERGÍA EÓLICA Y EL PRONÓSTICO DE LA VELOCIDAD DEL VIENTO

---

a escala global son resultado de la irradiación irregular del sol sobre la superficie del planeta. En las regiones cercanas al ecuador, una mayor temperatura provoca que el aire sea menos denso, por lo tanto este aire caliente asciende y viaja hacia los polos, donde el aire se encuentra a menor temperatura (Talayero & Telmo, 2011).

Los fenómenos a escala regional o local son provocados por las condiciones orográficas, como pueden ser la existencia de montañas o valles, las diferencias de temperatura entre la tierra y el mar, e incluso la existencia de estructuras como edificios o árboles.

Los lugares donde el aire desciende son consideradas zonas de alta presión, y los lugares donde el aire asciende son zonas de baja presión. Este gradiente de presiones provoca el movimiento del aire de zonas de alta presión hacia zonas de baja presión, lo que al final determina la velocidad y dirección del viento.

El movimiento del viento es principalmente horizontal, y su velocidad y fuerza es determinado por las siguientes fuerzas (Villarubia, 2007):

- Gradiente horizontal de presión: Fuerza con dirección perpendicular a las isobaras de presión, que va desde las zonas de alta presión hacia las zonas de baja presión.
- Fuerza de Coriolis: Fuerza provocada por la rotación de la Tierra, que provoca una desviación de la masa de aire hacia la derecha en el hemisferio norte, y hacia la izquierda en el hemisferio sur.
- Fuerza centrífuga: Fuerza debido a la curvatura de las isobaras, actúa de acuerdo al radio de curvatura de la trayectoria.
- Fuerzas de rozamiento: Fuerza presente en las capas bajas de la atmósfera, influenciada por factores por el tipo de suelo y la orografía.

Además de los viento globales existen vientos locales provocados por condiciones regionales, como son la brisa marina y la brisa terrestre, debido a que la tierra se calienta más rápido que el mar se produce una corriente de aire ascendente que provoca una circulación de mar a tierra, en la noche el fenómeno se invierte y el aire viaja de la tierra hacia el mar; las brisas de valle y de montaña son originadas debido a que el aire montaña arriba se enfría más rápido por lo que tiende a descender hacia el valle por la ladera del la montaña (viento catabático), por el día el proceso se invierte y el aire tiende a subir por la ladera (viento anabático) (Villarubia, 2007).

### 1.3. Aerogeneradores

Los aerogeneradores son los dispositivos que convierten la energía cinética del viento en electricidad. Los aerogeneradores modernos son máquinas orientables hacia el viento. Poseen un rotor, el cual está compuesto por las palas unidas mediante el buje (Figura 1.1). Incluye una barquilla que contiene varios elementos como el eje, el multiplicador, el freno, el generador eléctrico y transformador, y en su parte externa también incluye instrumentos de medición como la veleta y el anemómetro (Figura 1.2). También incluye la cimentación o torre, cuyo objetivo es aprovechar las altas velocidades del viento presentes a mayores alturas (Talayero & Telmo, 2011).

El funcionamiento de un aerogenerador de forma sintetizada es el siguiente: El viento hace girar las palas del aerogenerador, éstas están conectadas a la turbina mediante el buje, que a su vez está conectado a un eje lento; una multiplicadora aumenta el número de revoluciones y trasfiere el movimiento a un eje rápido. En el rotor se transforma la energía cinética en energía mecánica y en el generador, la energía mecánica en energía eléctrica. El convertidor lo convierte a corriente alterna de baja tensión. Se eleva la tensión y pasa a la subestación donde se transforma en corriente de alto voltaje. Finalmente, se envía a las redes de transmisión.

Los aerogeneradores se pueden clasificar de acuerdo a diferentes criterios (Villarubia, 2013):

- De acuerdo a la posición de su eje: Eje horizontal y eje vertical.
- Según el número de palas: Monopala, Tripala y Multipala.
- De acuerdo al lugar de emplazamiento: Onshore para aerogeneradores ubicados en tierra y offshore para aquellos ubicados en mar.
- Según el tipo de viento: Aerogeneradores para viento clase I, clase II o clase III.
- De acuerdo a su potencia nominal: Microturbinas eólicas (< 5 kW), turbinas de baja potencia (5-100 kW), turbinas de media potencia (100-150 kW) y alta potencia (5-7 MW onshore y 8-15 MW offshore).
- Por la intensidad de la turbulencia: Categorías A, B y C.

### 1.4. Potencial eólico

La potencia de un aerogenerador es la cantidad de energía que produce, la cual se mide en kilowatts o megawatts. La capacidad de generación de energía de un aere-

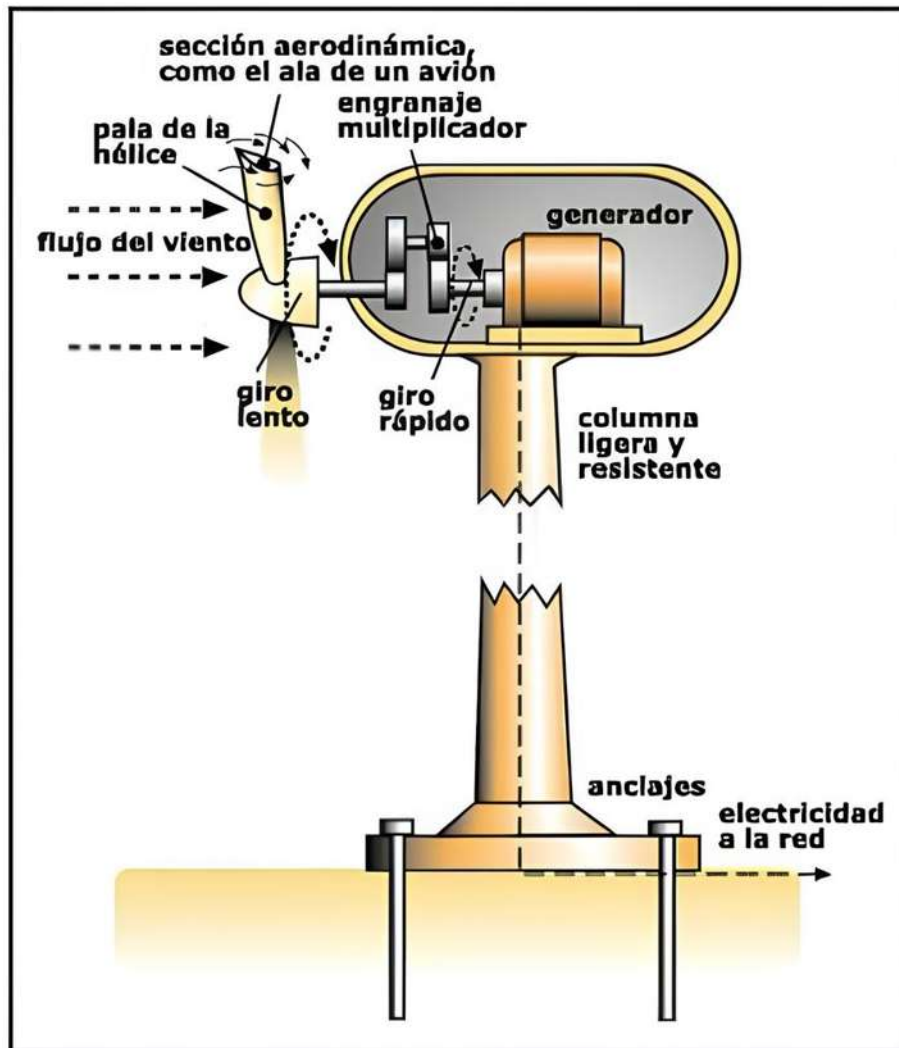


Figura 1.1: Elementos de un aerogenerador.

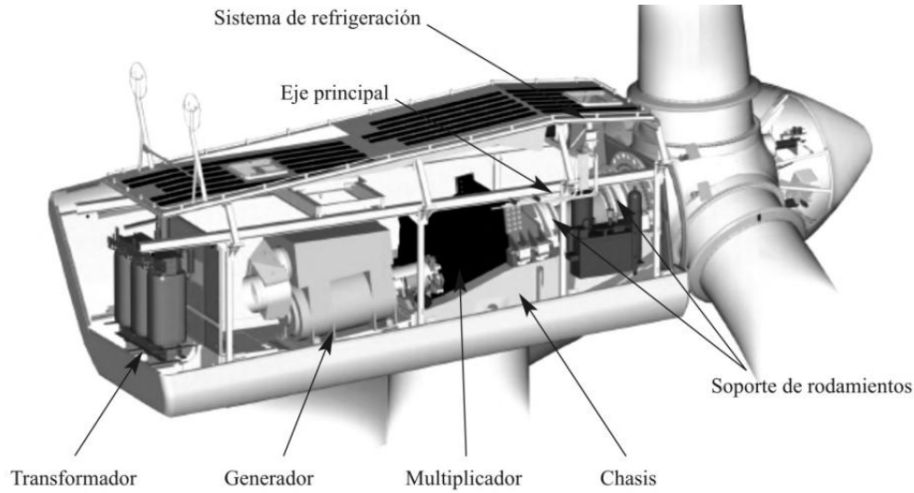


Figura 1.2: Elementos internos de un aerogenerador. Tomado de *Energía eólica* por Talayero, A., et.al. 2011. Prensas de la Universidad de Zaragoza. España.

generador depende de la disponibilidad del recurso eólico, por lo tanto, estos no operan a una potencia constante. Los aerogeneradores modernos comienzan a generar energía cuando se alcanza velocidades del viento de alrededor de 2 m/s. Conforme esta velocidad aumenta así lo hace su potencia de generación hasta llegar a su valor nominal, que suele ser de los 12 m/s a los 25 m/s. Al sobrepasar estas velocidades el aerogenerador se detiene para evitar daños mecánicos y eléctricos (San Miguel et al., 2023). La Figura 1.3 muestra la curva de potencia de un aerogenerador de 5 MW perteneciente al National Renewable Energy Laboratory (NREL). La potencia de arranque va de 0-3 m/s, el aumento de potencia tiene lugar de los 3 m/s a los 11.4 m/s, de los 11.4 m/s a los 25 m/s se alcanza su potencia nominal, y a velocidades mayores que 25 m/s el aerogenerador se detiene, es decir, alcanza su velocidad de corte.

La evaluación del recurso eólico se realiza usando la siguiente ecuación:

$$E_w = \sum_{t=1}^N P_w(U_t)(\Delta t) \quad (1.1)$$

donde  $E_w$  = la energía generada,

$P_w$  = la potencia del aerogenerador,

$U_t$  = el valor de la velocidad del viento en el tiempo  $t$ ,

$\Delta t$  = el intervalo de tiempo de la serie temporal  $U_t$ .

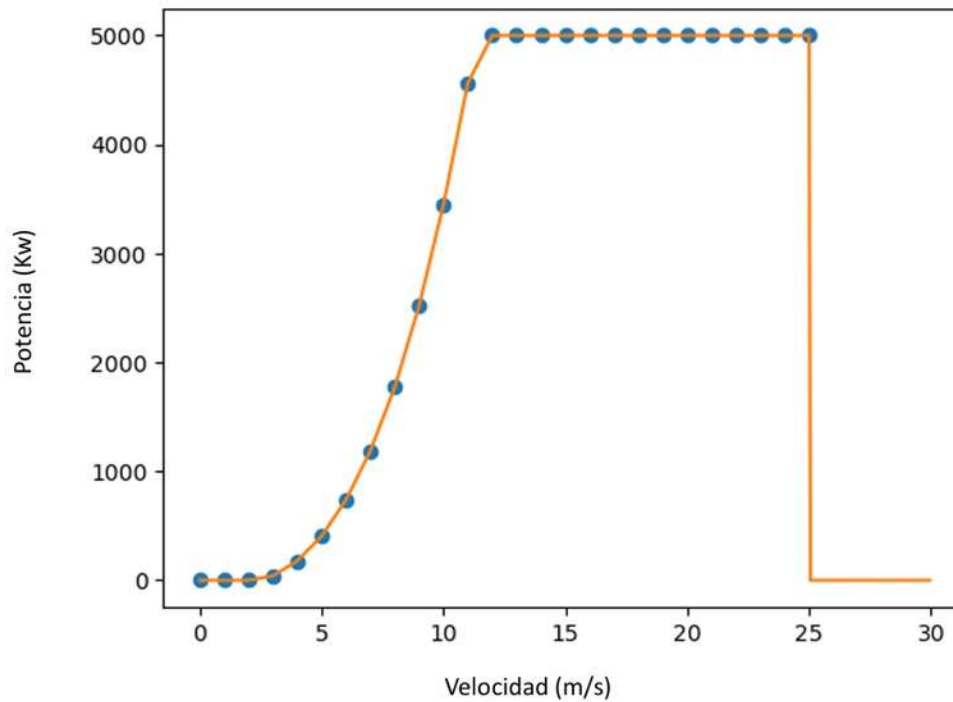


Figura 1.3: Curva de potencia de un aerogenerador de 5 MW de NREL.

## 1.5. Horizontes de pronóstico de la velocidad del viento

Los horizontes de pronóstico de la velocidad del viento cumplen diferentes objetivos dependiendo de para qué se requieran. Según [El-Fouly et al. \(2008\)](#) se pueden distinguir los siguientes horizontes de pronóstico:

- **Nowcasting (segundos o minutos):** Control y mantenimiento de las unidades de los parques eólicos.
- **1–6 horas:** Operación de Pequeños Sistemas Eléctricos.
- **1-72 horas:** Operación de Sistemas Eléctricos Interconectados. Despacho y selección de unidades (Unit Commitment). Comercio de electricidad en donde existen mercados de electricidad.
- **3–7 días:** Planificar el mantenimiento de los parques eólicos, programar el mantenimiento de la red y la operaciones de almacenamiento de energía.

Para pronósticos a corto plazo menores de 24 horas se suelen utilizar modelos estadísticos. Para pronósticos de múltiples pasos mayores a 24 horas se suele optar

## CAPÍTULO 1. FUNDAMENTOS DE LA ENERGÍA EÓLICA Y EL PRONÓSTICO DE LA VELOCIDAD DEL VIENTO

---

por modelos físicos. Los pronósticos de 24 horas poseen una especial relevancia por su estrecha relación con los mercados del día en adelante.

### 1.6. Legislación en México

En el caso de México, el Centro Nacional de Control de Energía (CENACE) es el encargado del control operativo del Sistema Eléctrico Nacional y es el encargado de operar el Mercado Eléctrico Mayorista. Sus funciones incluyen también garantizar el acceso a la Red Nacional de Trasmisión y proponer la ampliación y modernización de la misma. El pronóstico de la demanda a corto plazo del CENACE incluye los siguientes procesos:

- Pronóstico de la demanda a nivel gerencial.
- Pronóstico de ofertas de suministro básico a nivel de zonas de carga para el mercado del día en adelante.
- Pronóstico de ofertas de suministro básico a nivel de zonas de carga para la asignación de unidades de generación por confiabilidad.
- Integración de pronósticos de generación intermitente de centrales.
- Pronóstico de generación intermitente de CENACE.
- Balance de energía.

En el caso específico de la generación intermitente, que incluye a la energía eólica, sus pronósticos están normados por la Base 9.10.2 de las Bases del Mercado Eléctrico, publicadas el 8 de septiembre del 2015 en el Diario Oficial de la Federación ([Figura 1.4](#)), y por el Capítulo 3.1 del Manual de Pronósticos, publicado el 23 de noviembre de 2017 en el Diario Oficial de la Federación ([Figura 1.5](#)).

Es importante señalar que los pronósticos de generación intermitente, que incluye a la generación eólica, se actualizan diariamente con una resolución temporal horaria ( $\Delta t = 1$  hora), y entre sus usos se incluye el mercado del día en adelante. He aquí la importancia de los pronósticos de 24 horas de la velocidad del viento en la operación de la generación de energía eólica.

### 1.7. ¿Por qué pronosticar la velocidad del viento?

La energía eólica es diferente de las fuentes de energía de origen fósil en el sentido de que su producción depende de la velocidad del viento, el cual es variable tanto

# CAPÍTULO 1. FUNDAMENTOS DE LA ENERGÍA EÓLICA Y EL PRONÓSTICO DE LA VELOCIDAD DEL VIENTO

---

## **9.10 Pronósticos**

### **9.10.1 Pronósticos de demanda de energía eléctrica**

- (a) En el mercado de PRIMERA ETAPA, el CENACE preparará todos los pronósticos de carga.
- (b) En el mercado de SEGUNDA ETAPA, las Entidades Responsables de Carga deberán presentar al CENACE pronósticos de carga de cada hora:
  - (i) antes de las 9:00 horas de cada día para al menos los siguientes 7 días o un periodo mayor a solicitud del CENACE; y,
  - (ii) en cualquier momento si ocurre un cambio significativo en los pronósticos descritos en el punto anterior, para al menos los siguientes 7 días o un periodo mayor a solicitud del CENACE.
- (c) El CENACE desarrollará pronósticos de demanda de energía eléctrica independientes para su uso en:
  - (i) la planificación de mediano plazo, deberán ser actualizados según sea necesario; y,
  - (ii) la Asignación de Unidades de Central Eléctrica para Confiabilidad en el Día en Adelanto, la Asignación Suplementaria de Unidades de Central Eléctrica para Confiabilidad y la Asignación de Unidades de Central Eléctrica en Tiempo Real; estos pronósticos deberán ser permanentemente actualizados en tiempo real.

### **9.10.2 Pronósticos de generación intermitente**

- (a) Los representantes de las Centrales Eléctricas intermitentes están obligados a presentar pronósticos de generación de energía eléctrica por hora al CENACE, en los términos que definan los Manuales de Prácticas de Mercado.
- (b) Los representantes de las Centrales Eléctricas intermitentes están obligados a presentar pronósticos de generación de energía eléctrica en tiempo real al CENACE, en los términos que definan los Manuales de Prácticas de Mercado.
- (c) El CENACE desarrollará pronósticos independientes de generación para su uso en:
  - (i) la planificación de mediano plazo; y,
  - (ii) la Asignación de Unidades de Central Eléctrica para Confiabilidad en el Día en Adelanto, la Asignación Suplementaria de Unidades de Central Eléctrica para Confiabilidad y la Asignación de Unidades de Central Eléctrica en Tiempo Real.

Figura 1.4: Sección 9.10 de las Bases del Mercado Eléctrico que regula los pronósticos de generación intermitente.

# CAPÍTULO 1. FUNDAMENTOS DE LA ENERGÍA EÓLICA Y EL PRONÓSTICO DE LA VELOCIDAD DEL VIENTO

## Capítulo 3 Pronósticos de generación

### 3.1 Disposiciones generales

- 3.1.1** Los Participantes del Mercado que representen Unidades de Central Eléctrica firmes no-despachables o intermitentes serán responsables de estimar los pronósticos de generación correspondientes a cada una de sus unidades y presentarlos como Ofertas de Venta o programas de generación, según corresponda, en el Mercado de Energía de Corto Plazo.
- 3.1.2** Los Participantes del Mercado representantes de Unidades de Central Eléctrica serán responsables de proveer los insumos necesarios que permitan al CENACE emitir sus propios pronósticos de generación, mediante los formatos que el CENACE determine para ello.
- 3.1.3** El CENACE realizará pronósticos horarios de generación firme no-despachable o intermitente para asegurar una eficiente planificación de la operación del Sistema Interconectado Nacional, así como para la asignación y despacho de Unidades de Central Eléctrica en el Mercado de Energía de Corto Plazo, conforme al esquema de la Figura 1.
- 3.1.4** Los pronósticos de generación elaborados por el CENACE, así como su uso, horizonte temporal, periodicidad y frecuencia de actualización se resumen en la Tabla 2.

**Tabla 2. Resumen de la cronología de los pronósticos de generación**

| Tipo       | Pronóstico                   | Uso   | Límite para entrega de insumos | Horizonte de tiempo | Granularidad | Frecuencia de actualización |
|------------|------------------------------|---|--------------------------------|---------------------|--------------|-----------------------------|
| Generación | Intermitente-horario         | <ul style="list-style-type: none"> <li>• Planificación de mediano plazo</li> <li>• AU-MDA</li> <li>• AU-GC</li> </ul> | 10:00 hrs de cada día          | 7 días              | 1 hora       | Diaria                      |
|            | Intermitente-quinceminutal   | <ul style="list-style-type: none"> <li>• AU-GC</li> <li>• AU-TR</li> </ul>  | continuo                       | 2.5 horas           | 15 minutos   | 15 minutos                  |
|            | Firme no despachable-horario | <ul style="list-style-type: none"> <li>• Planificación de mediano plazo</li> <li>• AU-MDA</li> <li>• AU-GC</li> </ul> | 10:00 hrs de cada día          | 7                   | 1 hora       | Diaria                      |

Figura 1.5: Sección 3.1 del Manual de Pronósticos que regula los pronósticos de generación intermitente.

## CAPÍTULO 1. FUNDAMENTOS DE LA ENERGÍA EÓLICA Y EL PRONÓSTICO DE LA VELOCIDAD DEL VIENTO

---

en el tiempo como en el espacio. Esta característica es la que provoca que sea llamada también una fuente de energía intermitente, el cual es uno de los principales problemas que enfrentan los operadores eléctricos. Estas variaciones pueden poner en riesgo el correcto funcionamiento de los sistemas, especialmente con una mayor cantidad de energía eólica dentro de la red.

El aumento cada vez mayor del porcentaje de energía eólica dentro de los sistemas eléctricos está obligando a que se busque que la operación de los parques eólicos sea pensada más como plantas de generación controladas y menos como fuentes de energías intermitentes. Para lograr este objetivo es indispensable anticipar la producción de los parques eólicos, para lo cual el pronóstico de la velocidad del viento es fundamental, ya que de ello depende la estimación de la potencia eólica que entregarán los aerogeneradores.

Los pronósticos son esenciales para el mantenimiento de las unidades de los parques eólicos, optimizar el flujo de potencia entre las unidades convencionales y las eólicas, participar activamente en los mercados eléctricos, planificar la generación de los sistemas eléctricos, y organizar el almacenamiento eficiente de la energía (El-Fouly et al., 2008).

Además, una adecuada predicción permite reducir la incertidumbre operativa, mejorar la integración de la energía renovable en la red y minimizar los costos asociados a desbalances o a la necesidad de recurrir a fuentes de respaldo convencionales. Por lo tanto, el desarrollo y la implementación de métodos de pronóstico avanzados no solo impactan en la operación técnica, sino también en la rentabilidad y la viabilidad a largo plazo de los proyectos eólicos.

### **1.8. Modelos estadísticos: Modelos Box-Jenkins para series de tiempo**

#### **1.8.1. Autoregressive Moving Average (ARMA)**

Para series de tiempo que muestran un comportamiento estacionario, un modelo ARMA puede ser suficiente para su modelado. Una serie de tiempo estacionaria significa que sus datos fluctúan alrededor de una media sin una tendencia particularmente significativa y su varianza permanece esencialmente constante con el paso del tiempo.

## CAPÍTULO 1. FUNDAMENTOS DE LA ENERGÍA EÓLICA Y EL PRONÓSTICO DE LA VELOCIDAD DEL VIENTO

---

Un modelo  $ARMA(p, q)$  es expresado como:

$$Y_t = c + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (1.2)$$

donde  $\phi_i$  =  $i$ -ésimo parámetro autorregresivo,  
 $\theta_j$  =  $j$ -ésimo parámetro de medias móviles,  
 $e_t$  = el termino error en el tiempo  $t$ ,  
 $e_{t-k}$  = el termino error en el tiempo  $t - k$ .

O su ecuación simplificada con notación backshift:

$$(1 - \phi_1 B - \dots - \phi_p B^p) Y_t = (1 - \theta_1 B - \dots - \theta_q B^q) e_t \quad (1.3)$$

donde  $BY_t = Y_{t-1}$ .

Las variables  $p$  y  $q$  indican que el modelo  $ARMA(p, q)$  posee un número  $p$  de parámetros autorregresivos y un número  $q$  de parámetros de medias móviles.

### 1.8.2. Autoregressive Integrated Moving Average (ARIMA)

En el caso de series de tiempo no estacionarias, es decir, que no poseen una media y una varianza constante, primero es necesario remover esa no estacionariedad lo cual se logra a través de la aplicación de una transformada matemática y el cálculo de la diferencia a sus datos, según sea el caso.

La operación diferencia es definida como  $Y'_t = Y_t - Y_{t-1}$ .

La ecuación para el modelo  $ARIMA(p, d, q)$ , en notación backshift es:

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d Y_t = (1 - \theta_1 B - \dots - \theta_q B^q) e_t \quad (1.4)$$

donde  $\phi_i$  =  $i$ -ésimo parámetro autorregresivo,  
 $\theta_j$  =  $j$ -ésimo parámetro de medias móviles,  
 $e_t$  = el termino error en el tiempo  $t$ ,  
 $e_{t-k}$  = el termino error en el tiempo  $t - k$ .  
 $d$  = el número de diferencias realizadas

Las variables  $p$ ,  $d$  y  $q$  indican que el modelo  $ARIMA(p, d, q)$  posee un número  $p$  de parámetros autorregresivos, un número  $q$  de parámetros de medias móviles y un número  $d$  de diferencias.

## CAPÍTULO 1. FUNDAMENTOS DE LA ENERGÍA EÓLICA Y EL PRONÓSTICO DE LA VELOCIDAD DEL VIENTO

Existen diferentes transformaciones matemáticas para estabilizar la varianza, entre las cuales se encuentra la raíz cuadrada ( $X_t = \sqrt{Y_t}$ ), la raíz cúbica ( $X_t = \sqrt[3]{Y_t}$ ), logaritmo ( $X_t = \log(Y_t)$ ) y la inversa negativa ( $X_t = -1/Y_t$ ). Todas estas funciones son miembros de la familia de las transformaciones de potencia:

$$X_t = \begin{cases} -Y_t^p, & p < 0, \\ \log(Y_t), & p = 0, \\ Y_t^p, & p > 0. \end{cases} \quad (1.5)$$

Escogiendo  $p = 1/2$  se obtiene la transformación de la raíz cuadrada,  $p = 1/3$  se obtiene la raíz cúbica, y  $p = -1$  se obtiene la inversa negativa.

### 1.8.3. Seasonal Autoregressive Integrated Moving Average (SARIMA)

Para series de tiempo influenciadas por factores estacionales, es decir, que presentan un patrón que se repite continuamente después de un intervalo fijo de tiempo, un modelo SARIMA es una mejor opción.

La ecuación para el modelo  $SARIMA(p, d, q)(P, D, Q)_s$ , en notación backshift, es expresada como:

$$\begin{aligned} (1 - \phi_1 B - \dots - \phi_p B^p)(1 - \Phi_1 B^s - \dots - \Phi_P B^{Ps})(1 - B)^d (1 - B^s)^D Y_t \\ = (1 - \theta_1 B - \dots - \theta_q B^q)(1 - \Theta_1 B^s - \dots - \Theta_Q B^{Qs}) e_t \end{aligned} \quad (1.6)$$

donde  $\phi_i$  =  $i$ -ésimo parámetro autoregresivo,

$\theta_j$  =  $j$ -ésimo parámetro de medias móviles,

$e_t$  = el termino error en el tiempo  $t$ ,

$e_{t-k}$  = el termino error en el tiempo  $t - k$ .

$\Phi_m$  =  $m$ -ésimo parámetro estacional autoregresivo,

$\Theta_n$  =  $n$ -ésimo parámetro estacional de medias móviles,

$s$  = número de periodos por estación,

$d$  = el número de diferencias no estacionales,

$D$  = el número de diferencias estacionales.

Las variable  $p, d, q, P, D, Q$  y  $s$  indican que el modelo  $ARIMA(p, d, q)(P, D, Q)_s$  posee un número  $p$  de parámetros autorregresivos no estacionales, un número  $q$  de

parámetros de medias móviles no estacionales, un número  $P$  de parámetros autorregresivos estacionales, un número  $Q$  de parámetros de medias móviles estacionales, un número  $d$  de diferencias no estacionales, un número  $D$  de diferencias estacionales y que el patrón estacional se repite con una periodicidad  $s$ .

#### 1.8.4. Función de autocorrelación (ACF) y función de autocorrelación parcial (PACF)

Sean  $Y_t, Y_{t-1}, \dots, Y_{t-k}, \dots$  puntos de una serie de tiempo. La función de autocorrelación (ACF) es la gráfica del coeficiente de autocorrelación  $r_k$  como una función del desfase  $k$ , el cual se calcula de acuerdo a la [Ecuación 1.7](#).

$$r_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2} \quad (1.7)$$

El ACF mide la relación lineal entre dos valores  $Y_t$  y  $Y_{t-k}$  a una distancia  $k$  de separación, es decir, mide cómo es la relación de un punto en los datos de la serie de tiempo con sus puntos anteriores. Es adimensional, es decir, independiente de la escala de medición de la serie temporal.

El ACF ayuda a identificar si en una serie de tiempo los valores anteriores contienen información sobre el valor siguiente, si existe una relación entre un valor y sus valores previos. También proporciona una verificación útil sobre la estacionalidad, los ciclos y otros patrones presentes en series temporales.

La función de autocorrelación parcial (PACF) es la gráfica del coeficiente de autocorrelación parcial, utilizada para medir el grado de asociación entre  $Y_t$  y  $Y_{t-k}$ , cuando los efectos de los desfases  $1, 2, 3, \dots, k-1$  han sido eliminados. Es decir, identificar cómo es la relación entre los valores actuales de una serie temporal con valores anteriores de la misma serie temporal, manteniendo constantes los efectos de los desfases intermedios entre ellos.

La función de autocorrelación (ACF) para los errores de pronóstico es muy útil para determinar si existe algún patrón en los errores (residuales) después de que se ha aplicado un modelo de pronóstico. Esta es una herramienta útil para indicar si un modelo de pronóstico es candidato a mejorarse ([Makridakis et al., 2008](#)).

Si los valores de autocorrelación caen fuera de los límites de los valores críticos se puede inferir que existe información extra en la serie de tiempo que no está siendo

capturada por el modelo.

$$\text{valores críticos} = 2/\sqrt{n}$$

donde  $n$  = número de elementos de la serie de tiempo.

La serie de tiempo muestra un comportamiento aleatorio, también llamado ruido blanco, cuando no se muestran autocorrelaciones significativas.

### 1.8.5. Fases de la metodología Box-Jenkins

La metodología Box-Jenkins es utilizada para la selección de los modelos Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA), y Seasonal Autoregressive Integrated Moving Average (SARIMA). Fue desarrollada por George Box y Gwilym Jenkins a principios de 1970s, y frecuentemente se hace referencia a ella simplemente como modelos los ARMA, ARIMA o SARIMA (Makridakis et al., 2008).

En la [Figura 1.6](#) se resume los pasos básicos de la metodología Box-Jenkins para modelar series de tiempo. Consiste de tres fases:

- *Fase 1 - Identificación:* Se puede dividir a su vez en dos fases, la primera consiste en todos los procesos de preparación de los datos, como la limpieza de los datos, la eliminación de los valores atípicos, y la verificación de la estacionariedad de la serie de tiempo. En los casos de series de tiempo no estacionarias se realiza una diferencia en sus datos para estabilizar su varianza y una transformación matemática para estabilizar la media. En la segunda parte, mediante el análisis de las gráficas de los coeficientes de autocorrelación (ACF) y autocorrelación parcial (PACF) se identifican los potenciales modelos ARMA, ARIMA o SARIMA, según sea el caso.
- *Fase 2 - Estimación y prueba:* Fase en la que se estiman los valores de los diferentes parámetros de los potenciales modelos. Usando criterios como el de máxima verosimilitud se ajusta un potencial modelo a los datos, se calcula el valor de los parámetros para dicho modelo y se realiza un análisis a los residuales del modelo. Actualmente el cálculo de los parámetros se realiza usando herramientas computacionales que agilizan los procesos de cómputo.
- *Fase 3 - Aplicación:* Fase donde se selecciona el modelo que se usará para el pronóstico de la serie de tiempo. Un criterio para la selección de dicho modelo

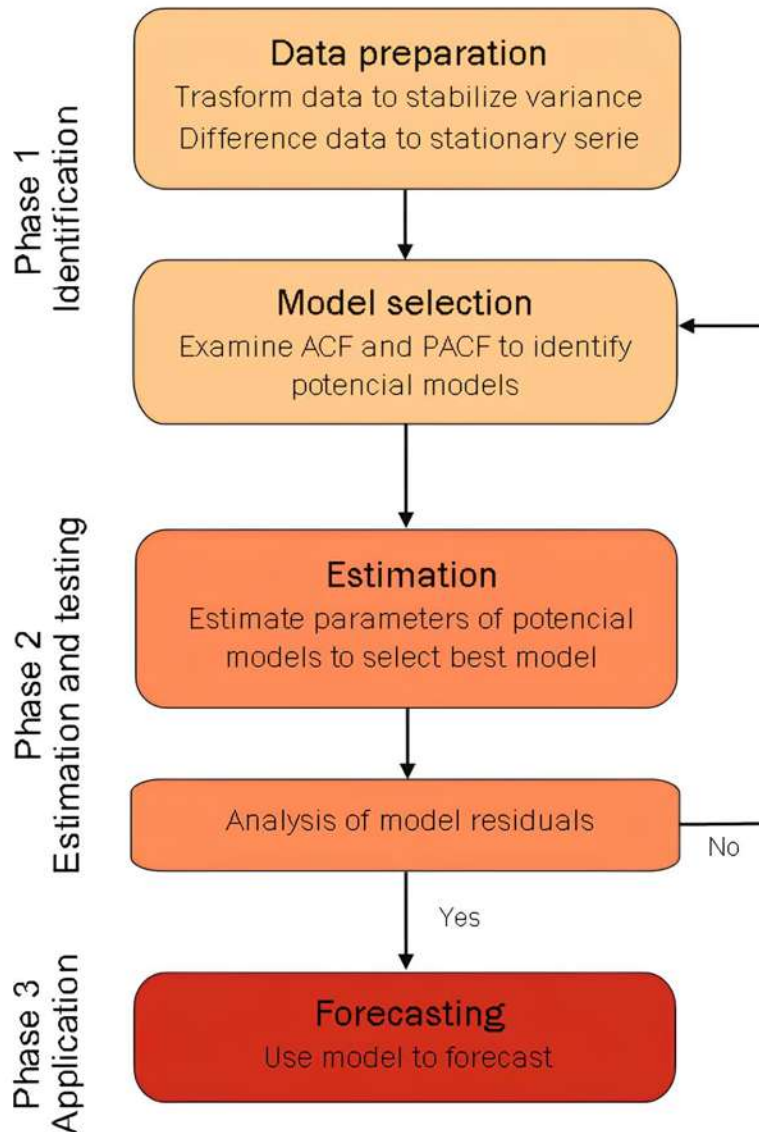


Figura 1.6: Diagrama de flujo de la metodología Box-Jenkins para el modelado de series de tiempo. Adaptado de *Forecasting Methods and Applications* (p.314), by S. Makridakis, 1998.

es el análisis de los residuales después del ajuste. Si los residuales muestran un comportamiento aleatorio o de ruido blanco el modelo es escogido, de lo contrario se regresa a la fase de la selección del modelo. Es importante mencionar que pueden existir diferentes modelos que cumplan con este criterio, por lo tanto, diferentes modelos pueden seleccionarse para el pronóstico.

Se puede decir que la metodología Box-Jenkins se define como un conjunto de pasos que incluyen la identificación, el ajuste, la verificación de los modelos y el pronóstico que se obtiene directamente de la ecuación de un modelo ya ajustado.

## 1.9. Modelos físicos: Modelo Weather Research and Forecasting (WRF)

El modelo Weather Research and Forecasting es un modelo numérico de predicción meteorológica. Este sistema fue desarrollado para cumplir funciones tanto operativas como de investigación, y es resultado de varias colaboraciones entre instituciones y universidades de los Estados Unidos y otros países (Skamarock et al., 2019).

El modelo WRF es un modelo regional, es decir, sus pronósticos se limitan a una determinada región o dominio del globo terrestre. En la [Figura 1.7](#) se muestran de forma sintetizada los tres módulos principales que conforman el proceso de compilación del modelo WRF:

- *El sistema de pre-procesamiento WPS (WRF Pre-processing System):* Incluye los procesos previos a la compilación del modelo WRF. Este conformado por tres programas con funciones específicas:
  - *Geogrid:* En este módulo se especifica el dominio del modelo, es decir, se define la región territorial en donde se llevarán a cabo las simulaciones. En el archivo `namelist.wps` se establecen las coordenadas geográficas de inicio y fin del dominio, el número de puntos de malla y la resolución espacial, así como la proyección cartográfica que se usará. También se define si se usará dominios anidados, los cuales permiten aumentar la resolución en áreas específicas. La ejecución de `geogrid.exe` genera los archivos `geom_em.d01.nc`, que contienen información sobre la topografía, el uso de suelo y características de la superficie.

## CAPÍTULO 1. FUNDAMENTOS DE LA ENERGÍA EÓLICA Y EL PRONÓSTICO DE LA VELOCIDAD DEL VIENTO

---

- *Ungrib*: El proceso siguiente es la preparación de las condiciones iniciales y de frontera del modelo, las cuales derivan de un modelo global. Los modelos globales más usados son el modelo GFS y el modelo ECMWF. Lo anterior asegura que la simulación regional sea congruente con la dinámica atmosférica global. Estos datos se suelen encontrar en formato `.grib`, por lo que el programa `ungrib.exe` extrae las variables meteorológicas de los archivos globales.
  - *Metgrid*: El programa `metgrid.exe` asigna a cada punto de la malla ya definida los valores de las variables meteorológicas extraídos, mediante el uso de interpolaciones espaciales y temporales, generando los archivos `met_em.d01.nc`. Estos archivos son la base para inicializar el modelo WRF, garantizando la coherencia entre el modelo global y la simulación regional.
- *El modulo ARW (Solver Advanced Research WRF):*
- *Configuración*: Antes de ejecutar el modelo se realiza su configuración mediante el archivo `namelist.input`. En este archivo se establecen las fechas de inicio y fin del pronóstico, la frecuencia de salida de los resultados, y muy importante, los esquemas físicos de parametrización de los procesos atmosféricos. La correcta selección de estos parámetros deriva en que el modelo realice una simulación realista de la atmósfera en el sitio de interés.
  - *Ejecución*: El paso siguiente es la ejecución de `real.exe` que transforma los archivos `met_em.d01.nc` en un archivo `wrfinput.d01` que contiene las condiciones iniciales, y un archivo `wrfbdy.d01` que incluye a las condiciones de frontera, un formato legible para el modelo. El programa `wrf.exe` lleva a cabo la integración numérica de las ecuación físicas que describen los procesos atmosféricos en el dominio definido, generando archivos `wrfout.d01.nc` que contienen el pronóstico final. Los tiempos de cálculo dependerán de los recursos disponibles que tenga a su alcance el modelo.
- *El post-procesamiento*: Es el compendio de herramientas que se utilizan para analizar los datos de pronóstico del modelo WRF, y varían de acuerdo a los objetivos de su uso. Incluyen a software como GrADS para la elaboración de mapas, programas como Integrated Data Viewer para simulaciones en 3D, programas como MatLab para la creación de gráficas y datos, y lenguajes como Python

## CAPÍTULO 1. FUNDAMENTOS DE LA ENERGÍA EÓLICA Y EL PRONÓSTICO DE LA VELOCIDAD DEL VIENTO

para la manipulación de la información en diversas formas. Se puede incluir también la validación de modelo, al hacer la comparación con observaciones locales y la interpretación de los pronósticos aplicada a diferentes problemáticas.

Una descripción más detallada del proceso de compilación se puede encontrar en [Skamarock et al. \(2019\)](#).

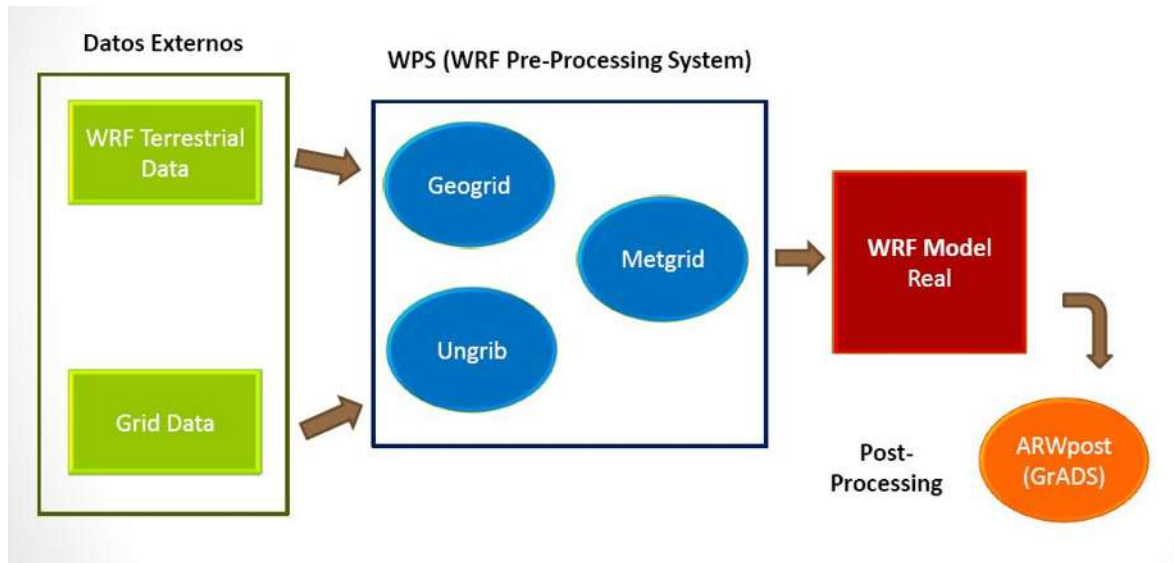


Figura 1.7: Diagrama del funcionamiento del modelo WRF.

Entre las ventajas que presenta el modelo WRF se encuentran que es de código abierto, lo cual lo hace fácilmente accesible para la comunidad científica, fomentando la creación de comunidades que brindan soporte técnico y desarrollan herramientas adicionales, principalmente de software libre. El modelo WRF puede personalizarse al definir los parámetros físicos y condiciones iniciales de acuerdo a la región de interés. Adicionalmente, el modelo tiene un potencial de alta resolución, llegando a escalas de hasta un kilómetro, aunque esto depende de los recursos computacionales disponibles.

### 1.10. Métricas de desempeño del modelo

Las métricas de desempeño del modelo son parte de las herramientas que permiten la validación de un modelo de pronóstico, por lo tanto, son una herramienta muy útil para determinar la idoneidad de un modelo. Existe una variedad de métricas para

## CAPÍTULO 1. FUNDAMENTOS DE LA ENERGÍA EÓLICA Y EL PRONÓSTICO DE LA VELOCIDAD DEL VIENTO

---

medir el desempeño de un modelo de pronóstico. En este trabajo se ha optado por las métricas más comunes utilizadas para los modelos de pronóstico de la velocidad del viento.

$$BIAS = \frac{1}{n} \sum_{t=1}^n e_t \quad (1.8)$$

El error absoluto medio (Mean absolute error)

$$MAE = \frac{1}{n} \sum_{t=1}^n |e_t| \quad (1.9)$$

El error cuadrático medio (Mean squared error)

$$MSE = \frac{1}{n} \sum_{t=1}^n e_t^2 \quad (1.10)$$

La raíz del error cuadrático medio (Root mean squared error)

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2} \quad (1.11)$$

Definiendo el error del modelo como:

$$e_t = R_t - F_t \quad (1.12)$$

$R_t$ ,  $F_t$  y  $e_t$  representan a series de tiempo.  $R_t$  corresponde a la serie de tiempo de las observaciones locales y  $F_t$  a la serie de tiempo del pronóstico. El valor de  $n$  dependerá del número de elementos de las series de tiempo, para un número  $n$  de observaciones y elementos en la serie pronosticada habrá  $n$  valores de error.

Un valor pequeño en la métrica BIAS significa que los errores positivos y negativos tienden a contrarrestarse entre sí. Esta métrica ayuda a determinar si el pronóstico está siendo subestimado o sobreestimado. Un pronóstico sobreestimado predice un recurso eólico que no es real y un pronóstico subestimado deja fuera valioso recurso eólico. Aunque ninguna de estas situaciones es deseada, se suele preferir un pronóstico subestimado sobre uno sobreestimado, pero ambas situaciones conllevan problemas al sistema de distribución de la red eléctrica al amenazar el equilibrio entre la oferta y la demanda de electricidad.

Las métricas MAE y MSE son útiles para conocer de forma general la magnitud del error del modelo al hacer a cada uno de los errores positivos. La métrica MAE logra

## CAPÍTULO 1. FUNDAMENTOS DE LA ENERGÍA EÓLICA Y EL PRONÓSTICO DE LA VELOCIDAD DEL VIENTO

---

este objetivo al tomar el valor absoluto de cada error y luego promediándolos, la métrica MSE lo logra al elevar cada error al cuadrado y calculando de igual manera su promedio. En la métrica MSE los errores más grandes poseen un mayor peso que los errores más pequeños, útil cuando la prioridad es identificar errores grandes. Se puede interpretar como un indicador que mide la propagación de los errores. La métrica RMSE es más interpretable que MSE al estar a la misma escala que los datos, pero es importante señalar que en esta métrica, al igual que en MSE, los errores más grandes tienen un mayor peso.

Las métricas de error porcentual como es el error porcentual medio (MPE) y el error porcentual absoluto medio (MAPE) se han dejado fuera al no poder ser calculadas cuando las series temporales de valores reales contienen ceros, una situación frecuente en las series de tiempo de velocidad del viento.

## Capítulo 2

# Análisis de la literatura sobre modelos de pronósticos de la velocidad del viento

### 2.1. Clasificación de los modelos de pronóstico de la velocidad del viento

Los modelos de pronóstico de la velocidad del viento se suelen dividir en dos grandes categorías, los modelos estadísticos y los modelos físicos:

- *Modelos estadísticos:* Los modelos estadísticos, como su nombre lo dice, hacen uso de técnicas estadísticas para realizar sus pronósticos. Su característica principal es su necesidad de tener disponibles datos históricos para entrenar los modelos, los cuales suelen ser representados mediante series de tiempo. Su funcionamiento se basa en encontrar patrones y tendencias en la información que posteriormente serán utilizados en la construcción de los pronósticos. Frecuentemente usados para horizontes de 1 o pocos pasos. Esta categoría puede dividirse a la vez en dos categorías, los modelos tradicionales y los modelos basados en aprendizaje automático:
  - *Modelos tradicionales:* Los modelos tradicionales se fundamentan en teoría matemática y/o inferencia estadística. Entre sus fortalezas están el ser explicativos, interpretables y útiles para conjuntos de datos de menor tamaño. Algunos ejemplos son los modelos de regresión lineal, los modelos de promedios móviles (MA), los modelos autorregresivos (AR), los modelos de suavizado exponencial simple (SES), los modelos de Holt-Winters,

## CAPÍTULO 2. ANÁLISIS DE LA LITERATURA SOBRE MODELOS DE PRONÓSTICOS DE LA VELOCIDAD DEL VIENTO

---

los modelos autorregresivos de medias móviles (ARMA), y los modelos estacionales autorregresivos integrados de medias móviles (SARIMA).

- *Modelos basados en aprendizaje automático (ML)*: Los modelos basados en aprendizaje automático aprenden los patrones directamente de los datos de entrenamiento, requieren de menos supuestos que los modelos estadísticos y son capaces de procesar grandes cantidades de información. Algunos ejemplos son los modelos basados en redes neuronales, bosques aleatorios, gradient boosting, entre otros.
- *Modelos físicos*: Los modelos físicos también son conocidos como modelos numéricos de predicción del tiempo, Numerical Weather Prediction (NWP) models en inglés. Estos modelos utilizan las ecuaciones físicas que describen los procesos atmosféricos para realizar sus pronósticos. Entre las ecuaciones fundamentales podemos encontrar la ecuación de conservación de energía, conservación de masa y conservación del momentum. Su enfoque se basa en entender a la predicción del tiempo como un problema de condiciones iniciales. Requieren de recursos computacionales considerables ya que realizan una discretización (total o regional) de la atmósfera e integran las ecuaciones en cada uno de los puntos. Este tipo de modelos son los que mejor han funcionado para horizontes de pronóstico de múltiples pasos. La velocidad del viento es sólo una de las múltiples variables que pueden pronosticar (temperatura, radiación, precipitaciones, etc). Estos modelos se pueden clasificar a la vez en modelos globales y modelos regionales:
  - *Modelos globales*: Los modelos globales describen los procesos atmosféricos que suceden en todo el globo terrestre, su objetivo es lograr una descripción de la atmósfera lo más cercana a la realidad. Debido a que son modelos que cubren todo el planeta la resolución de sus mallados suelen rondar los 50 Km. Ejemplo de estos modelos son el modelo Global Forecast System (GFS) y el modelo del European Centre for Medium-Range Weather Forecasts (ECMWF).
  - *Modelos regionales o de mesoescala*: Estos modelos describen la situación atmosférica en un área geográfica específica, esta es la razón por la cuál también son llamados modelos de área limitada (LAM). Se ayudan de los resultados de los modelos globales para obtener sus condiciones iniciales y de frontera, y lograr resolución mayores que pueden llegar hasta los

## CAPÍTULO 2. ANÁLISIS DE LA LITERATURA SOBRE MODELOS DE PRONÓSTICOS DE LA VELOCIDAD DEL VIENTO

---

2 km. Son capaces de modelar fenómenos locales. Ejemplo de estos modelos son el modelo High-Resolution Rapid Refresh (HRRR), el modelo HARMONIE-AROME, el modelo Meteorological Administration Mesoscale (CMA-MESO) y el más ampliamente usado, el modelo Weather Research and Forecasting (WRF).

En los últimos años se ha vivido un auge de los **modelos híbridos**. La revisión bibliográfica muestra un predominio de propuestas basadas en la combinación de dos o más modelos sobre el uso de un sólo modelo. [Makridakis \(1989\)](#) explicaba las razones de por qué la combinación de múltiples técnicas de pronóstico mejorar significativamente el desempeño en general de los modelos de pronóstico. Entre las razones se encuentra que, contrario a lo que dictaría la lógica, la combinación de modelos reduce errores sistemáticos y aleatorios, cada modelo posee sus propios errores y al combinar los modelos los errores se compensan entre si y disminuyen su tamaño. Otra razón es que la combinación ofrece una robustez metodológica, un solo modelo puede fallar bajo ciertas condiciones, así que tener más de un modelo reduce esta vulnerabilidad. La combinación de modelos aprovecha las fortalezas de cada modelo y reduce vulnerabilidades, además de que ofrecen mayor flexibilidad para adaptarse a diferentes contextos.

Los modelos híbridos también se pueden clasificar en dos categorías de acuerdo al tipo de modelos que combinan:

- *Modelo híbrido estadístico:*
  - *Modelo tradicional + machine learning (ML):* Modelos que combinan algunas de las técnicas estadísticas tradicionales con modelos basadas en aprendizaje automático.
  - *Modelo exclusivo ML:* Consisten en propuestas de modelos que se basan específicamente en diferentes técnicas de aprendizaje automático. Pueden combinar dos o mutiples modelos de ML.
- *Modelo híbrido físico + estadístico:* Los podemos encontrar más comunmente como una combinación de un modelo físico con una o más técnicas estadísticas, ya sean modelos estadísticos tradicionales o basados en aprendizaje automático. Por sus características, es el tipo de modelos más usado para pronósticos de múltiples pasos.

## CAPÍTULO 2. ANÁLISIS DE LA LITERATURA SOBRE MODELOS DE PRONÓSTICOS DE LA VELOCIDAD DEL VIENTO

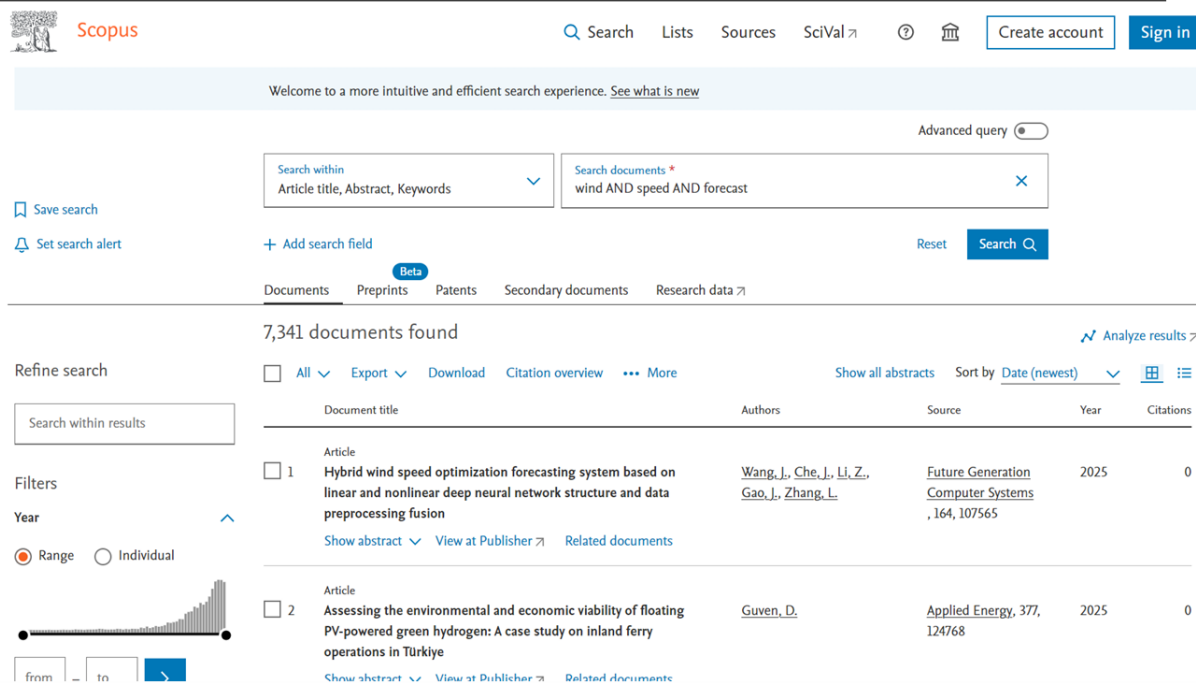


Figura 2.1: Búsqueda de las palabras *wind speed forecast* en la base de datos Scopus sin utilizar filtros.

De acuerdo a la clasificación mostrada podemos decir que el modelo propuesto en el presente trabajo se encuentra en la categoría de los modelos híbridos físicos + estadísticos, ya que combina al modelo físico WRF, con el modelo estadístico SARIMA.

### 2.2. Numeralia sobre publicaciones

La búsqueda realizada en la base de datos Scopus (Elsevier, 2024) usando las palabras clave WIND SPEED FORECAST (Figura 2.1) arrojó un total de 7,341 documentos publicados a la fecha del 1 de diciembre de 2024. Reduciendo la búsqueda solamente a artículos científicos y publicaciones en inglés, el número de documentos se redujo a 4,809 documentos.

En la Figura 2.2 se muestran los artículos científicos en inglés publicados entre 1990 a 2024, la gráfica muestra no sólo un aumento en el número de publicaciones sino que también este crecimiento ha seguido un patrón exponencial, mostrando que se trata de un campo que se ha ido consolidado a lo largo de los años. Lo anterior en concordancia al aumento en la generación y la capacidad instalada de energía eólica en el mundo, especialmente en la última década.

## CAPÍTULO 2. ANÁLISIS DE LA LITERATURA SOBRE MODELOS DE PRONÓSTICOS DE LA VELOCIDAD DEL VIENTO

---

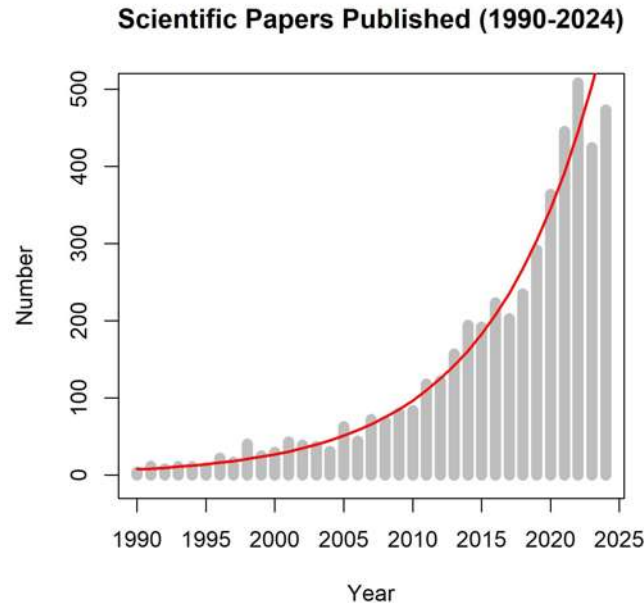


Figura 2.2: Artículos científicos relacionados con los pronósticos de la velocidad del viento publicados entre 1990-2024

Esta búsqueda incluye no sólo a los artículos que hablan sobre pronósticos sino también a los artículos relacionados en diferentes sentidos con la velocidad del viento. Al realizar un nuevo filtrado manual se redujeron a 2,082 artículos que tratan en específico sobre modelos de pronóstico de la velocidad del viento. La búsqueda muestra que existe un abanico de modelo de pronóstico, que van desde los modelos más tradicionales basados en técnicas estadísticas, pasando por los modelos numéricos, hasta los modelos más complejos basados en aprendizaje profundo.

En la [Figura 2.3](#) se muestra el crecimiento en la publicación de artículos sobre modelos de pronóstico de la velocidad del viento en los últimos tres décadas, en el cual se observa, al igual que la figura anterior, un crecimiento exponencial, siendo el 2022 el año con más publicaciones sobre el tema. Lo anterior permite deducir que los modelos de pronóstico de la velocidad del viento es un campo de la investigación que se encuentra en expansión, caracterizado por una diversidad de enfoques y propuestas multidisciplinarias.

Para los fines del presente trabajo, es de especial interés los pronósticos de la velocidad del viento con un horizonte de 24 horas debido a su estrecha relación con los mercados de la electricidad de un día en adelante. Por esta razón se realizó un se-

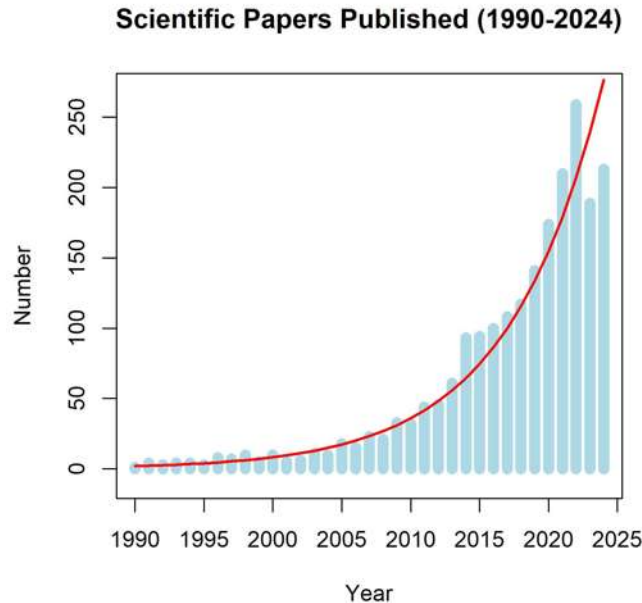


Figura 2.3: Artículos publicados entre 1990-2024 específicamente sobre modelos de pronóstico de la velocidad del viento.

gundo filtrado manual sólo incluyendo los modelos de pronóstico con horizonte de predicción igual o mayor a 24 horas. El nuevo filtrado mostró que sólo 229 artículos cumplieron con este criterio.

Tomando en consideración sólo los artículos publicados a partir del 2019, se obtuvieron 141 artículos. La [Figura 2.4](#) sigue mostrando un crecimiento en el número de publicaciones entre los años 2009 al 2024, este crecimiento es de manera exponencial aunque con una tendencia menos clara. Entre estos artículos, 97 artículos mencionan al modelo Weather Research and Forecasting (WRF) y 11 artículos mencionan al modelo Seasonal Autoregressive Integrated Moving Average (SARIMA) o alguna de sus variaciones (AR, MA, ARMA, ARIMA). Lo anterior muestra una fuerte predominancia de los modelos basados en WRF para horizontes de pronóstico de al menos 24 horas. Dada la confiabilidad que han mostrado los modelos numéricos para predicciones de múltiples pasos no es sorpresa esta cantidad de publicaciones. WRF en particular tiene la ventaja de ser un modelo de código abierto y accesible para la investigación.

La revisión de estas estadísticas permiten deducir que el campo de investigación de los pronósticos de la velocidad del tiempo a la fecha es una área se mantiene en

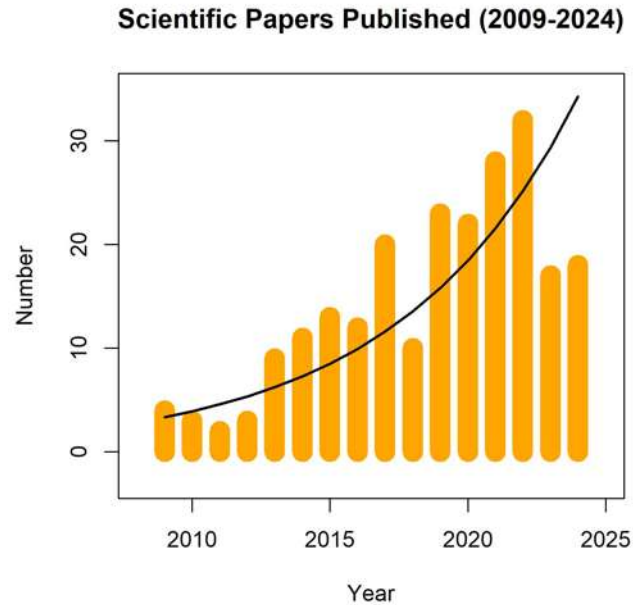


Figura 2.4: Artículos publicados entre 2009-2024 sobre modelos de pronóstico de la velocidad del viento que incluyen un horizonte de pronóstico de 24 horas.

continúa en expansión y contiene una producción científica considerable, con la finalidad de enfrentar los desafíos asociados con la variabilidad del viento para brindar seguridad en el proceso actual de transición energética. En las siguientes secciones se profundiza acerca de las áreas de mejora y la brecha de conocimiento identificada.

### 2.3. Análisis crítico de enfoques basados en SARIMA y WRF

A continuación se muestra un análisis de los diferentes modelos reportados en la literatura para el pronóstico de la velocidad del viento, organizados de acuerdo al tipo de modelos que los conforman, y cuya fecha de publicación es posterior al año 2018<sup>1</sup>.

---

<sup>1</sup>Con excepción del modelo de [Guo et al. \(2011\)](#)

### 2.3.1. Modelos ARIMA y SARIMA

Respecto a los modelos estadísticos, la escena se encuentra dominada por los modelos híbridos. Estos los podemos encontrar como la combinación de modelo tradicionales con modelos basados en aprendizaje automático (ML), o como la combinación de solamente modelos basados en ML. En la [Tabla 2.1](#) se muestra los modelos estadísticos que utilizan al modelo ARIMA (o alguna de sus variantes) como parte de sus procesos para el pronóstico de la velocidad del viento.

Tabla 2.1: Modelos estadísticos de pronóstico de la velocidad del viento.

| Técnicas                         | Referencias                          | Horizonte de pronóstico | Características                   | Año  |
|----------------------------------|--------------------------------------|-------------------------|-----------------------------------|------|
| ARIMA, DBN, ELM, LGB, LSTM y SVR | <a href="#">Du et al. (2024)</a>     | 1, 2 y 3 pasos          | Modelo híbrido (Tradicional + ML) | 2024 |
| ARIMA y VMD                      | <a href="#">Ding et al. (2022)</a>   | 7 pasos                 | Tradicional + ML                  | 2022 |
| ARIMA y FNN                      | <a href="#">Sun &amp; Jin (2022)</a> | 3, 8, 12 y 24 hrs.      | Tradicional + ML                  | 2022 |
| SARIMA                           | <a href="#">Liu et al. (2021b)</a>   | 24 hrs.                 | Tradicional                       | 2021 |
| ICEEMDAN, BPNN y ARIMA           | <a href="#">Duan et al. (2021)</a>   | 1 paso                  | Tradicional + ML                  | 2021 |
| VMD y ARIMA                      | <a href="#">Moreno et al. (2021)</a> | 1 y 12 pasos            | Tradicional                       | 2021 |
| ARIMA y BPNN                     | <a href="#">Hu et al. (2021)</a>     | 1, 2 y 3 pasos          | Tradicional + ML                  | 2021 |
| EMD, LSTM y ARIMA                | <a href="#">Liu et al. (2021a)</a>   | 1 día                   | Tradicional + ML                  | 2021 |
| VMD, FOA, ARIMA y DBN            | <a href="#">Zhang et al. (2020)</a>  | 1 día                   | Tradicional + ML                  | 2020 |
| WT y ARIMA,                      | <a href="#">Singh et al. (2019)</a>  | 1, 3, 5, 7 y 10 minutos | Tradicional                       | 2019 |
| ELM, ICEEMDAN y ARIMA            | <a href="#">Wang et al. (2018)</a>   | 1 paso                  | Tradicional + ML                  | 2018 |
| SARIMA y LSSVM                   | <a href="#">Guo et al. (2011)</a>    | 12 pasos                | Tradicional + ML                  | 2011 |

[Du et al. \(2024\)](#) propone la combinación de 6 modelos (ARIMA, DBN, ELM, LGB, LSTM, SVR) mediante el uso de Temporal Fusion Transformers (TFT) para realizar sus pronósticos, usa 1 año de datos como datos de entrada, divididos en 60 % para el entrenamiento, 20 % para la validación y 20 % para la prueba, con un horizon-

## CAPÍTULO 2. ANÁLISIS DE LA LITERATURA SOBRE MODELOS DE PRONÓSTICOS DE LA VELOCIDAD DEL VIENTO

---

te de 1,2 y 3 pasos. Si bien este enfoque logra altas precisiones, su principal limitación es la elevada complejidad computacional y los grandes volúmenes de datos. Además, la combinación de múltiples modelos reduce significativamente la interpretabilidad. Desde el punto de vista operativo, implementar y mantener seis modelos simultáneamente resulta inviable para muchos parques eólicos.

[Ding et al. \(2022\)](#) utiliza Variational Mode Decomposition(VMD) como preprocesamiento para descomponer la señal, eliminar el ruido y eliminar la no estacionariedad, la predicción de frecuencias bajas se realiza con ARIMA, más un postprocesamiento para corregir el error. Una limitación importante de este estudio es su horizonte de pronóstico de solo 7 pasos, insuficiente para aplicaciones en mercados eléctricos que requieren 24 horas de anticipación.

[Sun & Jin \(2022\)](#) combinan ARIMA con una red neuronal FNN (multi-layer perceptron neural network) donde ARIMA define la capa de entrada y la FNN genera pronósticos de 24 horas con 24 neuronas de salida, con horizontes de pronósticos de 3, 8, 12, y 24 horas. Si bien se utilizan 11 años de datos históricos (1.73 millones de mediciones horarias a 10 metros de altura) no se explica si volúmenes tan extensos son necesarios o si con menores datos se lograrían resultados comparables.

El trabajo de [Liu et al. \(2021b\)](#) resulta particularmente relevante porque demuestra que un modelo tradicional como SARIMA puede superar a modelos de deep learning (LSTM y GRU) en pronósticos de 24 horas. Esta es una contribución significativa, ya que cuestiona la noción generalizada de que los modelos más complejos son siempre superiores. Sin embargo, el estudio se limita a datos offshore y no explora la combinación de SARIMA con modelos físicos. Una fortaleza clave es que utilizó sólo 29 días de entrenamiento, demostrando la eficiencia de SARIMA con conjuntos de datos reducidos.

[Duan et al. \(2021\)](#) propone un modelo híbrido en el que se usa un métodos de descomposición de señales (ICEEMDAN) como preprocesamiento, la predicción se realiza con una red neuronal BPNN y se realiza una corrección de error como postprocesamiento donde se usa el modelo SARIMA para pronosticar el error. La combinación de BPNN (caja negra) con ARIMA (interpretable) resulta en un modelo híbrido que solo es parcialmente interpretable. Desde el punto de vista operativo, su complejidad lo hace inviable para muchos parques eólicos.

[Moreno et al. \(2021\)](#) combinan Variational Mode Decomposition (VMD) con Singular Spectrum Analysis (SSA) como preprocesamiento, utilizando ARIMA para pronosticar cada subserie resultante. El estudio se evalúa con horizontes de 1 y 12 pasos,

## CAPÍTULO 2. ANÁLISIS DE LA LITERATURA SOBRE MODELOS DE PRONÓSTICOS DE LA VELOCIDAD DEL VIENTO

---

acercándose a un horizonte de 24 horas pero insuficiente para los mercados eléctricos. Adicionalmente la descomposición usada genera interrogantes, la secuencia y el propósito de cada descomposición no están claramente justificados, dificultando la interpretabilidad del modelo.

[Hu et al. \(2021\)](#) usa los modelos de persistencia, AR, ARMA y Back Propagation Neural Network (BPNN) para el pronóstico, el modelo se entrena con 1 mes de datos y el horizonte de pronóstico es de 1, 2 y 3 pasos adelante. Si bien la combinación de modelos lineales (AR, ARMA) con una red neuronal es una propuesta sólida, el horizonte máximo de 3 pasos limita importantemente su aplicabilidad en el área de la energía eólica.

[Liu et al. \(2021a\)](#) propone un modelo híbrido que usa Empirical Mode Decomposition (EMD) para descomponer la serie en altas y bajas frecuencias, usa LSTM para pronosticar las frecuencias altas y ARIMA para pronosticar las frecuencias bajas y los residuales, el modelo se entrena con datos de 5, 7 y 10 días a intervalo 15 minutos con un 80 % de los datos destinados al entrenamiento. Los conjuntos de datos de entrenamiento parecer ser insuficientes, especialmente para el modelado de las frecuencias altas con LSTM.

[Zhang et al. \(2020\)](#) propone un modelo híbrido que usa la descomposición Variational Mode Decomposition (VMD), optimización mediante Fruit Fly Optimization (FOA), ARIMA para componentes regulares y Deep Belief Network (DBN) para componentes irregulares. Se usan una semana de datos de entrenamiento y una proporción del 80 %. Debido a la combinación de modelos, el costo computacional es elevado, y es difícilmente interpretable, por ejemplo, no es posible rastrear que parte del modelo es el encargado de corregir los errores de pronóstico.

[Singh et al. \(2019\)](#) propone el uso de Wavelet Transform (WT) para descomponer las series de viento en cuatro subseries de frecuencias altas y bajas, los pronósticos se realizan con ARIMA. Se logran capturar las ráfagas y turbulencia del viento. Debido a su enfoque en ráfagas de viento Su horizonte de pronóstico es a muy corto plazo, lo que limita su aplicabilidad a pronósticos más largos.

El modelo de [Wang et al. \(2018\)](#) utiliza Extreme Learning Machine (ELM) para el pronóstico principal y emplea ARIMA para determinar el tamaño del vector de entrada del ELM, implementa una corrección de error donde se descompone la serie temporal de errores mediante ICEEMDAN y se pronostica cada componente con ARIMA. El estudio se limita a pronósticos de un solo paso y el costo computacional es muy alto dado su horizonte de pronóstico.

## CAPÍTULO 2. ANÁLISIS DE LA LITERATURA SOBRE MODELOS DE PRONÓSTICOS DE LA VELOCIDAD DEL VIENTO

---

Aunque publicado en 2011, vale la pena mencionar el modelo propuesto por [Guo et al. \(2011\)](#), donde se utiliza el modelo SARIMA para realizar el pronóstico, y el modelo Least Squares Support Vector Machines (LSSVM) para predecir los errores de pronóstico. Este estudio constituye uno de los primeros intentos documentados de utilizar un modelo para predecir los errores de pronóstico de otro modelo. El uso de LSSVM y SARIMA hace al modelo parcialmente interpretable, con un horizontes limitado a 12 pasos.

Recapitulando, los modelos estadísticos incluyen a los modelos autorregresivos de medias móviles ([Singh et al., 2019](#); [Zhang et al., 2020](#); [Hu et al., 2021](#)), neural networks ([Sun & Jin, 2022](#)), long short-term memory ([Lawal et al., 2021](#)), y deep learning ([Zhang et al., 2024](#)). Adicionalmente, se pueden encontrar técnicas como descomposición de señales ([Liu et al., 2021a](#); [Moreno et al., 2021](#)), corrección de errores ([Ding et al., 2022](#); [Duan et al., 2021](#); [Wang et al., 2018](#)) y optimización ([Liu et al., 2020](#)).

Tradicionalmente los modelos estadísticos se centran en pronósticos a corto plazo o a pronósticos de muy corto plazo. Sin embargo, se pueden encontrar propuestas para horizontes de pronóstico de 24 horas, por ejemplo, usando modelos basados en deep learning ([Moreno et al., 2024](#); [Hong et al., 2024](#); [Thu et al., 2023](#)) o métodos estadísticos tradicionales ([Liu et al., 2021b](#); [Costa et al., 2021](#)). Con respecto al modelo SARIMA, [Liu et al. \(2021b\)](#) demostró que un modelo de pronóstico SARIMA de 24 horas puede tener un mejor desempeño que el modelo long short-term memory y gated recurrent unit models. Estos estudios indican que los modelos estadísticos pueden ser eficaces para los pronósticos de 24 horas.

Varios modelos muestran una excesiva complejidad sin justificar las razones de ello, combinan tres o más técnicas sin mostrar como cada componente añadida una mejora significativa al pronóstico final, esto impide cuantificar la contribución real de cada componente. Varios modelos usan una cantidad importante de datos para su entrenamiento, lo cual contrasta con la realidad de muchos parques eólicos que no cuentan con esta cantidad de información. Muchos de los modelos basados en ML operan como cajas negras. Esto es problemático para el sector energético, donde los operadores necesitan comprender por qué un modelo genera ciertas predicciones para tomar decisiones informadas. Los horizontes de este tipo de modelo se siguen limitando al corto plazo, de los 13 estudios analizados solo 5 abordan horizontes de 24 horas.

### 2.3.2. Modelos físicos reportados

Los pronósticos de velocidad del viento con horizonte de 24 horas, importantes para los mercados de la electricidad son producidos principalmente por modelos físicos (modelos de predicción numérica del tiempo). Diferentes modelos NWP han sido utilizados en la literatura para este propósito. Por ejemplo, [Yang et al. \(2023b\)](#) y [Zhang & Wang \(2023\)](#) usan el modelo European Centre for Medium-Range Weather Forecasts (ECMWF), [Myers et al. \(2024\)](#) utiliza el modelo High-Resolution Rapid Refresh (HRRR) para la predicción de la velocidad del viento, [Liu et al. \(2024\)](#) hace lo propio con el modelo China Meteorological Administration Mesoscale Model (CMA-MESO), y sobre el modelo Weather Research and Forecasting (WRF) podemos encontrar el trabajo de [Zhao et al. \(2021\)](#).

Una limitación común en estos estudios es que utilizan los modelos NWP como predictores únicos, sin implementar una estrategia de corrección de errores como post-procesamiento, una oportunidad de mejora que la presente investigación busca atender.

Entre los modelos NWP disponibles, el Weather Research and Forecasting (WRF) destaca por ser el más ampliamente utilizado en el ámbito académico y operativo, principalmente debido a cuatro ventajas: (1) es de código abierto y gratuito, (2) cuenta con una comunidad activa de desarrolladores y usuarios, (3) permite ajustar las parametrizaciones físicas para adaptarse a regiones específicas, y (4) puede ser ejecutado en una infraestructura computacional moderada. Estas características contrastan con los otros modelos, cuyo acceso es limitado y de alto costo, lo que limita su uso para contextos de investigación. Por estas razones, el modelo WRF ha sido seleccionado como el modelo físico base en la presente investigación.

Como ya se mencionó, una de las principales ventajas del WRF es su capacidad para ajustar parametrizaciones físicas que representen las características regionales del dominio de la simulación. En particular, para pronósticos de velocidad del viento a bajas alturas (típicamente la altura del buje de aerogeneradores, entre 40 y 120 metros), el esquema de la capa límite planetaria (Planetary Boundary Layer, PBL) ha mostrado ser el parámetro más sensible ([Duarte Jacondino et al., 2021](#); [Mi et al., 2023](#)). Varios esquemas PBL han sido probados para mejorar los pronósticos de la velocidad del viento ([Yang et al., 2023a](#); [Dzebre & Adaramola, 2020](#)).

### 2.3.3. Modelos híbridos para la corrección de errores

El modelo WRF, y los modelo NWP en general, presentan limitaciones que afectan la precisión de sus pronósticos, por ejemplo, las velocidades del viento derivadas de simulaciones numéricas están particularmente sesgadas cerca de la superficie. Adicionalmente, las truncaciones en los procesos de integración numérica, las simplificaciones en la representación de los procesos atmosféricos, y las imprecisiones en las condiciones iniciales y de frontera (provenientes de modelos globales como el GFS) contribuyen a la acumulación de errores en los pronósticos ([Al-Yahyai et al., 2010](#)).

Para solucionar este problema, se han propuesto modelos híbridos que usan métodos estadísticos para corregir los errores de pronóstico. La [Tabla 2.2](#) muestra los modelos híbridos encontrados en la literatura que se basan en este enfoque.

Tabla 2.2: Modelos que utilizan técnicas estadísticas para corregir errores de pronóstico provenientes de un modelo físico.

| Modelo físico | Referencias                        | Horizonte de pronóstico | Técnica usada para la corrección de error | Año  |
|---------------|------------------------------------|-------------------------|---|------|
| WRF           | <a href="#">Chen et al. (2019)</a> | 36 hrs.                 | Regresión lineal                          | 2019 |
| WRF           | <a href="#">Tsai et al. (2021)</a> | 72 hrs.                 | Decaying average                          | 2021 |
| WRF           | <a href="#">Xu et al. (2021)</a>   | 48 hrs.                 | LSTM                                      | 2021 |
| WRF           | <a href="#">Liu et al. (2023)</a>  | 96 pasos                | TCN y LSTM                                | 2023 |
| WRF           | <a href="#">Zhao et al. (2024)</a> | 24 hrs.                 | TCN y LSTM                                | 2024 |
| ECMWF         | <a href="#">Wang et al. (2019)</a> | 1, 4, 12 y 24 hrs.      | LR, SVM, BPNN, RF y RBFNN                 | 2019 |

[Chen et al. \(2019\)](#) propone un método de regresión lineal para corregir el sesgo de pronóstico de un ensemble WRF de 20 miembros, en el cuál también se probaron diferentes parametrizaciones, con un periodo de simulación de 3 meses. Aunque la regresión lineal es interpretable y computacionalmente eficiente, su capacidad para capturar los patrones en los errores de pronóstico tiene limitaciones. Además no explica sobre la generalización de los resultados a diferentes estaciones del año.

[Tsai et al. \(2021\)](#) usa un algoritmo Decaying Average para ajustar la velocidad del viento pronosticada mediante un ensemble WRF de 20 miembros para la región de Taiwan, las simulaciones corresponde a un periodo de 5 meses. Si bien este enfoque es simple y fácil de implementar, no modela explícitamente las dependencias

## CAPÍTULO 2. ANÁLISIS DE LA LITERATURA SOBRE MODELOS DE PRONÓSTICOS DE LA VELOCIDAD DEL VIENTO

---

secuenciales ni los patrones estacionales presentes en los errores de pronóstico, lo que limita su efectividad.

[Xu et al. \(2021\)](#) propone un modelo híbrido que combina WRF, Variational Mode Decomposition (VMD), Principal Component Analysis (PCA) para el pronóstico primario, seguido del modelo Long Short-Term Memory (LSTM) para la corrección del error de pronóstico. El modelo presenta limitaciones, por ejemplo, una alta complejidad computacional debido a la combinación de múltiples técnicas, grandes volúmenes de datos para entrenar el LSTM y falta de interpretabilidad, ya que el LSTM opera como una caja negra. Además la selección de parámetros para VMD y PCA requiere de un ajuste empírico, lo que dificulta su automatización.

[Liu et al. \(2023\)](#) propone un modelo que combina el modelo WRF, con temporal convolution network (TCN) y 2D convolution long short-term memory network (ConvLSTM), estos últimos son también usados para la corrección del error. Al estar basado en deep learning, el modelo requiere grandes volúmenes de datos históricos, en contextos reales, esto lo puede hacer inviable para varios parques eólicos.

[Zhao et al. \(2024\)](#) propone un modelo donde se construye un ensamble de pronósticos WRF y se elimina los miembros con el menor rendimiento usando un proceso de decisión Markov y deep reinforcement learning, posteriormente, se realiza una corrección del error usando Temporal Convolutional Network (TCN) y Bidireccional Long Short-term Memory (BiLSTM). Si bien reportan mejoras del 15 % en MSE y 14 % en RMSE, el costo computacional es alto, estos modelos requieren de recursos significativamente mayores que los enfoques basados en modelos estadísticos tradicionales, con riesgo de sobreajuste por la cantidad de hiperparámetros.

Finalmente, [Wang et al. \(2019\)](#) aplicó cinco modelos: Regresión lineal (LR), Support Vector Machine (SVM), Back-Propagation Neural Network (BPNN), Random Forest (RF) y Radial Basis Function Neural Network (RBFNN), como una estrategia para corregir el error de pronóstico de la velocidad del viento de un modelo de predicción numérica del tiempo. Las mejoras reportadas (1.7 %-3.9 % en RMSE) son modestas para la complejidad de los modelos, además no se profundiza porque ciertos modelos funcionan mejor para cierto tipo de horizontes de pronóstico.

### 2.4. Identificación de la brecha de conocimiento

Estas investigaciones demuestran que usar modelos estadísticos como estrategia de corrección de errores mejora de manera efectiva los pronósticos de velocidad del

## CAPÍTULO 2. ANÁLISIS DE LA LITERATURA SOBRE MODELOS DE PRONÓSTICOS DE LA VELOCIDAD DEL VIENTO

---

viento de los modelos físicos. Sin embargo, se identifican varios vacíos:

- La mayoría de los enfoques basados en ML (LSTM, TCN, redes neuronales) actúan como cajas negras, disminuyendo la interpretabilidad y dificultando entender que patrones de error están siendo corregidos.
- Estos modelos requieren grandes volúmenes de datos históricos, lo que contrasta con la realidad de varios proyectos eólicos, especialmente en México.
- La implementación operativa de modelos basados en deep learning demanda recursos computacionales importantes, lo que puede ser limitante para parques eólicos de menor escala.
- Varios estudios aplican directamente las técnicas de corrección los errores sin un análisis previo que determine si dichos errores son efectivamente predecibles, es decir, si presentan autocorrelaciones significativas. No basta con proporcionar al modelo estadístico suficientes datos para obtener predicciones precisas, es necesario evaluar si los datos pueden ser interpretados y pronosticados por el modelo.

En contraste con los modelos basados en aprendizaje automático, los modelos SARIMA no requieren de grandes cantidades de datos para su entrenamiento, logrando mejoras importantes con menores recursos computacionales. Un enfoque basado en SARIMA proporciona una metodología explícita y robusta para analizar las tendencias y los componentes estacionales presentes en los datos, un aspecto importante para el sector energético donde se requiere mayor interpretabilidad. Los modelos SARIMA poseen una sólida base matemática y, en el caso de conjuntos de datos pequeños de una sola variable, estos modelos prevalecen sobre los modelos de aprendizaje automático ([Kontopoulou et al., 2023](#)), incluso para pronósticos de 24 horas ([Liu et al., 2021b](#)).

Cada enfoque (tradicional o basado en aprendizaje automático) tiene diferentes fortalezas y debilidades dependiendo de las aplicaciones específicas y de las características del conjunto de datos ([Iaousse et al., 2023](#)). Las ventajas de los modelos SARIMA son: 1) funciona para conjuntos de datos de menor tamaño, 2) interpretabilidad, 3) una base matemática sólida y 4) un menor costo computacional. La principal desventaja es la linealidad del modelo SARIMA, pero, como se ha demostrado, esto no limita su capacidad para predecir patrones complejos y oscilatorios.

## CAPÍTULO 2. ANÁLISIS DE LA LITERATURA SOBRE MODELOS DE PRONÓSTICOS DE LA VELOCIDAD DEL VIENTO

---

A pesar de estas ventajas, en la literatura revisada no se ha reportado un modelo híbrido que combine específicamente al modelo físico Weather Research and Forecasting (WRF) con el modelo estadístico SARIMA como estrategia de corrección de errores para pronósticos de velocidad del viento con horizonte de 24 horas. Los estudios existentes utilizan SARIMA como modelo principal de pronóstico (Liu et al., 2021b) o en combinación con otros modelos estadísticos (Guo et al., 2011), pero no como corrector de errores de un modelo físico como el WRF. Tampoco se ha evaluado este enfoque en condiciones de alta turbulencia como las presentes en La Ventosa, Oaxaca, una región con fuerte potencial eólico en México.

Por lo tanto, la presente investigación busca llenar estos vacíos mediante un modelo híbrido WRF-SARIMA que prioriza el análisis de los errores de pronóstico, ofrece interpretabilidad y bajo costo computacional, y funciona con conjuntos de datos de tamaño reducido, características relevantes para su implementación operativa, especialmente en contextos de generación eólica como el mexicano.

El aporte de este trabajo de investigación es:

- Un modelo híbrido WRF-SARIMA novedoso, interpretable y automatizable.
- Validación en un sitio real con condiciones eólicas complejas.
- Una alternativa de bajo costo computacional frente a los modelos basados en ML.

## Capítulo 3

# Metodología del modelo híbrido WRF-SARIMA para la corrección de errores de pronóstico

### 3.1. Área de estudio y datos de mediciones en el sitio

Este estudio se realizó en La Ventosa, Oaxaca, Mexico; población ubicada en el sureste de México en el Istmo de Tehuantepec (ver [Figura 3.1](#)). El lugar se ubica cerca de la costa, a 30 metros sobre el nivel del mar. En esta región se ha llevado a cabo un número importante de proyectos eólicos debido a la presencia de fuertes vientos, con velocidades promedio mayores a la media nacional. Estos vientos, conocidos como *nortes* o *tehuanos* son consecuencia de las condiciones atmosféricas y de la topografía local. Los vientos intensos se generan debido a la diferencia de presión entre el Golfo de México y el Golfo de Tehuantepec que pasan a través de un paso montañoso ([Romero-Centeno et al., 2003](#)). El sitio de la mediciones se localiza en  $16^{\circ}32'40.8''$  latitud norte y  $94^{\circ}57'20.8''$  longitud oeste. El cual se trata de un terreno abierto con un alto nivel de turbulencia ([Lopez-Villalobos et al., 2018](#)).

En [Figura 3.2](#) se muestran las mediciones de la magnitud de la velocidad del viento registradas para el período que va del 20 de enero al 12 de diciembre de 2018, a alturas de 17.5 m y 40 m sobre el suelo. Para la toma de las mediciones se utilizó un anemómetro ultrasónico 3D Gill WindMaster montado sobre un mástil de celosía, a frecuencias de 1 Hz. Originalmente los datos fueron registrados a una resolución temporal de 1 segundo, pero para los fines de este trabajo, los datos se promediaron a hora, obteniendo como resultando 7,848 valores a 17.5 m de altura y 7,848 valores a 40 m de altura.

### CAPÍTULO 3. METODOLOGÍA DEL MODELO HÍBRIDO WRF-SARIMA PARA LA CORRECCIÓN DE ERRORES DE PRONÓSTICO

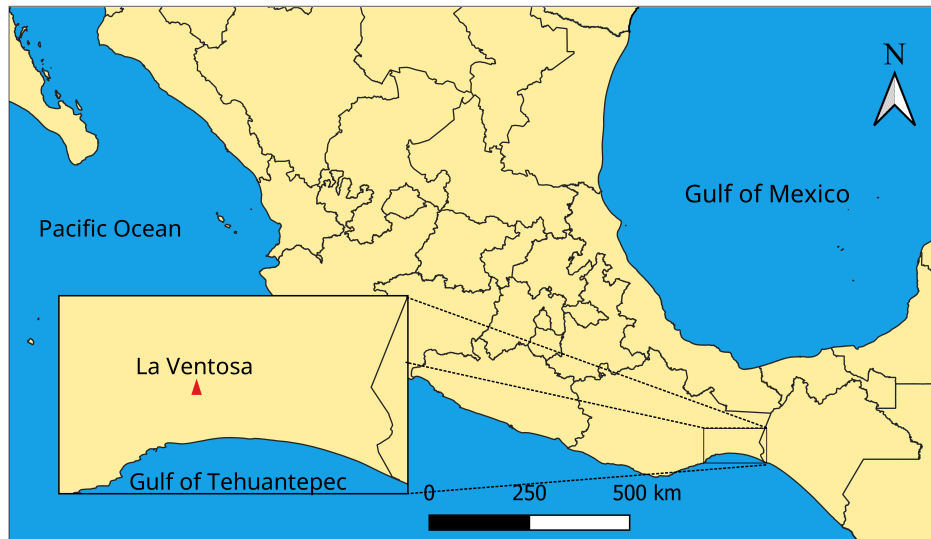


Figura 3.1: Localización de La Ventosa, Oaxaca.

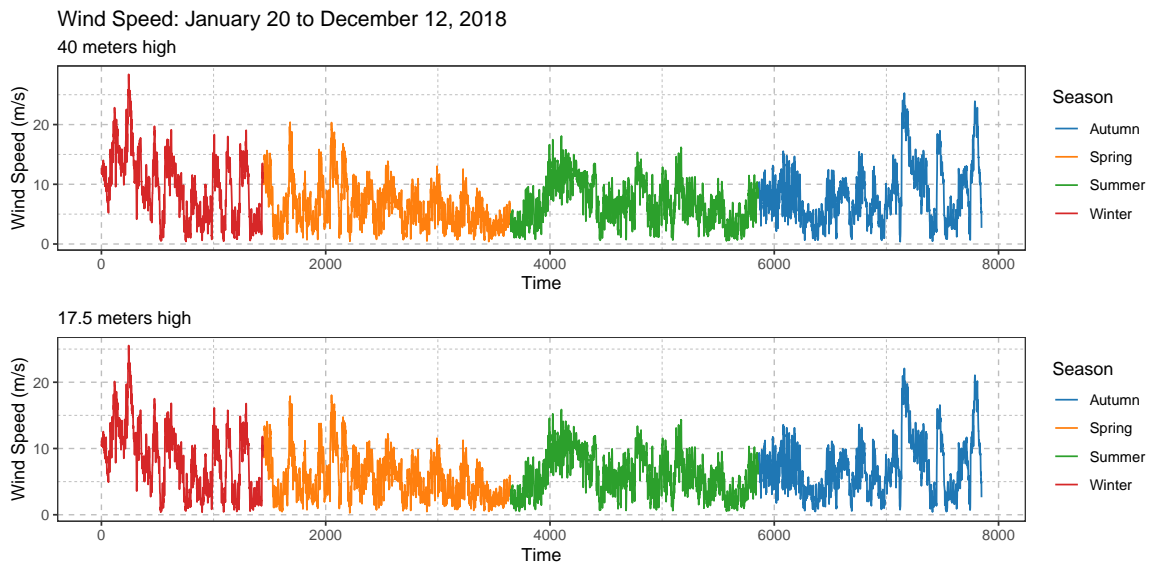


Figura 3.2: Velocidad del viento en La Ventosa.

## CAPÍTULO 3. METODOLOGÍA DEL MODELO HÍBRIDO WRF-SARIMA PARA LA CORRECCIÓN DE ERRORES DE PRONÓSTICO



The screenshot displays the 'M4 CERTE (operando) Descripción del proyecto' page. At the top, there are logos for SENER, INEEL, DTU, GEIC, IG, and AEM. The page header includes navigation links and a session status. The main content area is titled 'M4 CERTE (operando) Descripción del proyecto' and includes a table with project information and a photograph of a meteorological tower.

| Información del proyecto |   |
|--------------------------|---|
| Duración:                | 2017-11-23 19:30 adelante   |
| Datos históricos desde:  | 2017-11-23  |
| Posición:                | 16° 32' 49.27" N<br>94° 57' 20.83" O<br>31 m (sobre el nivel del mar)<br><a href="#">Mostrar en el mapa</a> |
| Contacto:                | Gerardo Pérez Martínez  |
| Página principal:        | <a href="http://www.ineel.mx">www.ineel.mx</a>  |
| Descripción:             | Torre meteorológica de 80m para la elaboración del AEM.   |

[Enlace a una imagen más grande](#)

Figura 3.3: Captura de pantalla de la página web del Atlas Eólico Mexicano del INEEL.

Las mediciones fueron realizadas como parte del proyecto Atlas Eólico Mexicano (AEM) teniendo como actores principales a la Universidad Tecnológica de Dinamarca (DTU) y al Instituto Nacional de Electricidad y Energías Limpias (INEEL), con financiamiento proveniente del Fondo para la Transición Energética y el Aprovechamiento Sustentable de la Energía (FOTEASE) de la Secretaría de Energía (SENER) en el año 2018 (ver [Figura 3.3](#)). Como parte del proyecto se colocaron 7 torres para medir la velocidad del viento a diferentes alturas, durante un año, en diferentes partes del país cuya zona estaba catalogada con de alto potencial eólico ([INEEL, 2025](#)).

De acuerdo a los datos, la velocidad máxima del viento se registró durante las estaciones del otoño y el invierno. En la [Tabla 3.1](#) se muestra el promedio estacional y anual, el valor máximo y la desviación estándar de la velocidad del viento, estimados a partir de los datos medidos. La velocidad promedio estimada en invierno fue de 9.67 m/s a la altura de 40 m y de 8.46 m/s a la altura de 17.5 m. En otoño, el promedio fue de 8.29 m/s a la altura de 40 m y de 7.11 m/s a la altura de 17.5 m. El promedio anual de velocidad del viento fue de 7.66 m/s a la altura de 40 m y de 6.64 m/s a la altura de 17.5 m.

En la [Figura 3.4](#) se muestra que un modelo de distribución de Weibull es el que mejor se ajusta a los datos de la velocidad del viento. Lo anterior se explica debido a que los valores extremos, como son las rachas fuertes, ocurren con poca frecuencia, mientras que las velocidades bajas o moderadas se presentan con mayor frecuencia. Esto contrasta con los resultados de [Jaramillo & Borja \(2004\)](#), quienes encontraron en

## CAPÍTULO 3. METODOLOGÍA DEL MODELO HÍBRIDO WRF-SARIMA PARA LA CORRECCIÓN DE ERRORES DE PRONÓSTICO

Tabla 3.1: Promedio de la velocidades del viento en La Ventosa

| Estación            | Vel. del viento a 40 m (m/s) | Vel. del viento a 17.5 m (m/s) |
|---------------------|------------------------------|--------------------------------|
| Primavera           | 6.34                         | 5.48                           |
| Verano              | 7.10                         | 6.16                           |
| Otoño               | 8.29                         | 7.11                           |
| Invierno            | 9.67                         | 8.46                           |
| Promedio            | <b>7.66</b>                  | <b>6.64</b>                    |
| Máximo              | 28.41                        | 25.54                          |
| Desviación estándar | $\sigma=4.43$                | $\sigma=3.98$                  |

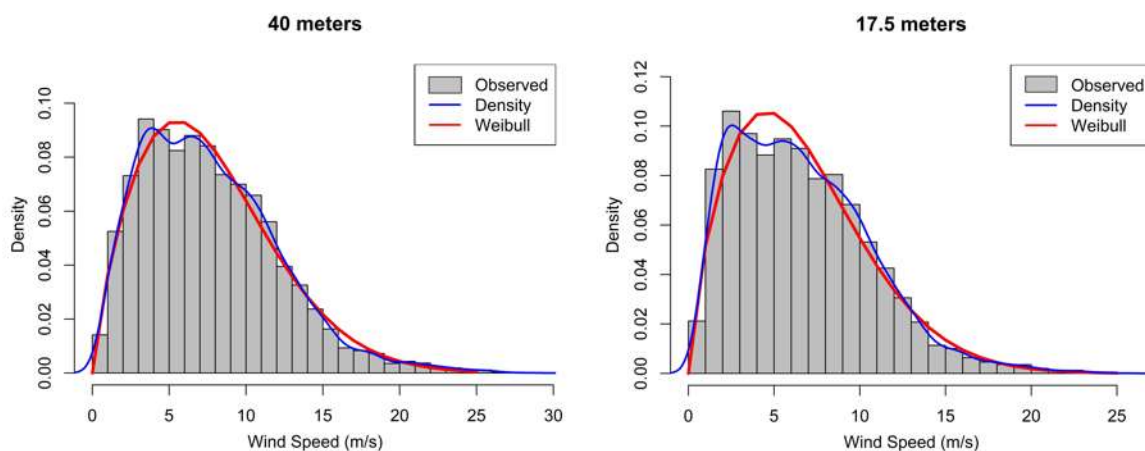


Figura 3.4: Distribuciones de probabilidad del la rapidez del viento en La Ventosa.

sus resultados que en La Ventosa la velocidad del viento poseen una distribución de probabilidad bimodal.

### 3.2. Pronósticos del modelo WRF

Como se mencionó en la [Sección 1.9](#), el modelo WRF es un sistema desarrollado para actividades de investigación y para la predicción numérica del tiempo. Los pronósticos WRF utilizados en este estudio fueron proporcionados por el Grupo de Interacción Océano-Atmósfera ([Grupo IOA, 2024](#)) del Instituto de Ciencias de la Atmósfera y Cambio Climático ([ICAyCC, 2024](#)) de la Universidad Nacional Autónoma de México (UNAM). El modelo se encuentra en funcionamiento desde el 2015 hasta la fecha (ver [Figura 3.5](#)).

El grupo IOA ha investigado la sensibilidad del modelo WRF a diversas condiciones

### CAPÍTULO 3. METODOLOGÍA DEL MODELO HÍBRIDO WRF-SARIMA PARA LA CORRECCIÓN DE ERRORES DE PRONÓSTICO

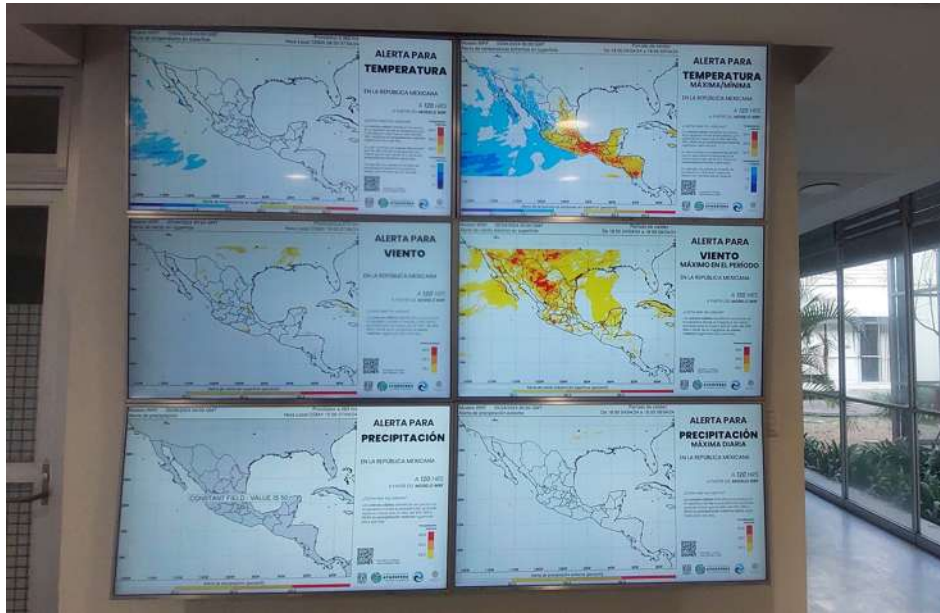


Figura 3.5: Fotografía de la visualización del modelo WRF al interior del Instituto de Ciencias de la Atmósfera y Cambio Climático UNAM.

iniciales y de frontera (Jurado de Larios, 2017) y ha incorporado datos precisos de uso de suelo y vegetación para mejorar la exactitud de los pronósticos meteorológicos (Rivera-Martínez, 2018; López-Espinoza et al., 2020).

Las simulaciones se realizaron utilizando la versión 3.9 de WRF, el cual usa una aproximación no hidrostática con una malla escalonada Arakawa tipo C y con coordenadas verticales que siguen la topografía del terreno. Las simulaciones comprende los periodos del 1 al 30 de junio de 2018 y del 13 de noviembre al 12 de diciembre de 2018.

El dominio usado cubre principalmente el territorio mexicano, abarcando desde  $4.12^\circ$  hasta  $38.42^\circ$  latitud norte y desde  $74.87^\circ$  hasta  $123.36^\circ$  longitud oeste, con una resolución horizontal y vertical de 15 km,  $338 \times 262$  puntos en la malla, 50 niveles verticales y una resolución temporal de 1 hora. Las condiciones iniciales y de frontera provienen de la base de datos del modelo de pronóstico global GFS con una resolución horizontal de  $0.25^\circ$ , inicializado el modelo a las 00:00 Tiempo Universal Coordinado (UTC) y actualizado cada 24 horas, para un pronóstico de 120 horas adelante.

En la Tabla 3.2 se muestra los esquemas de parametrización física utilizados para las simulaciones WRF.

En este trabajo las primeras seis horas de simulación fueron descartadas debido al

## CAPÍTULO 3. METODOLOGÍA DEL MODELO HÍBRIDO WRF-SARIMA PARA LA CORRECCIÓN DE ERRORES DE PRONÓSTICO

Tabla 3.2: Esquema de la parametrización física del modelo WRF.

|                          | 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  |

spin-up<sup>1</sup>, conservando las 24 horas posteriores para capturar un ciclo diurno completo de acuerdo a la hora local (GMT-6).

Es importante señalar que no existe un punto exacto en la malla de WRF que coincida exactamente con el sitio de las mediciones, por lo tanto, se realizó una interpolación bilineal utilizando los cuatro puntos en la malla más cercanos al sitio.

La [Ecuación 3.1](#) muestra la ecuación usada para la interpolación bilineal, donde  $ws(x, y)$  corresponde a la velocidad del viento resultante en el sitio de las mediciones, y  $ws_{11}$ ,  $ws_{12}$ ,  $ws_{21}$  y  $ws_{22}$  corresponden a las velocidades del viento en los cuatro puntos más cercanos en la malla.

$$ws(x, y) = \frac{1}{(x_2 - x_1)(y_2 - y_1)} [x_2 - x, x - x_1] \begin{bmatrix} ws_{11} & ws_{12} \\ ws_{21} & ws_{22} \end{bmatrix} \begin{bmatrix} y_2 - y \\ y - y_1 \end{bmatrix} \quad (3.1)$$

En la [Figura 3.6](#) se muestra una representación visual de la ecuación anterior.

Adicionalmente, a continuación se muestra el código<sup>2</sup> en MATLAB que se utilizó para realizar la interpolación bilineal:

```

archivo = 'E:\wrfout_d01_2018-11-01_00.nc';

TIMEdata = ncread(archivo, 'Times');
dim = size(TIMEdata);
XLATdata = ncread(archivo, 'XLAT');
lat = XLATdata(1, :, 1);
XLONGdata = ncread(archivo, 'XLONG');
long = XLONGdata(:, 1, 1);
u_variable = ncread(archivo_nc, 'U');
v_variable = ncread(archivo_nc, 'V');

latitud_limites = [16.5 16.6]; % N to S
longitud_limites = [-95.026 -94.882]; % W to E

```

<sup>1</sup>Dejar un tiempo de spin-up permite que las variables del modelo se ajusten y estabilicen antes de analizar los resultados.

<sup>2</sup>Elementos de este código son de la autoría de la Dra. Erika Danaé López Espinoza.

### CAPÍTULO 3. METODOLOGÍA DEL MODELO HÍBRIDO WRF-SARIMA PARA LA CORRECCIÓN DE ERRORES DE PRONÓSTICO

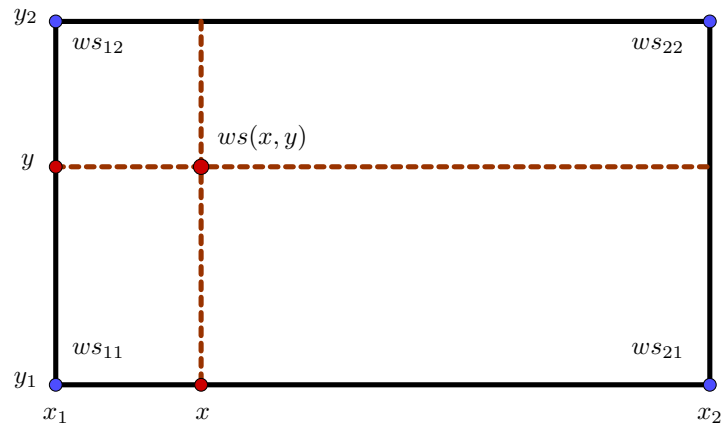


Figura 3.6: Interpolación bilineal de la velocidad del viento.

```

ubicacion =[16.546101 -94.953395]; %ubicacion del sitio de las mediciones

idxLat = find(latitud_lmites(1) <= lat & lat <= latitud_lmites(2));
idxLon = find(longitud_lmites(1) <= long & long <= longitud_lmites(2));

%Encontrar los 4 puntos mas cercanos al sitio de las mediciones
if (ubicacion(1) <= lat(idxLat))
y2=lat(idxLat)
y1=lat(idxLat-1)
idlat2=idxLat
idlat1=idxLat-1
else (ubicacion(1) > lat(idxLat))
y1=lat(idxLat)
y2=lat(idxLat+1)
idlat1=idxLat
idlat2=idxLat+1
end

if (punto(2) <= long(idxLon))
x2=long(idxLon)
x1=long(idxLon-1)
idlon2=idxLon
idlon1=idxLon-1
else (punto(2) > long(idxLon))
x1=long(idxLon)
x2=long(idxLon+1)
idlon1=idxLon
idlon2=idxLon+1
end

z=1; % interpolacion para el nivel z=1
u_11= squeeze(u_variable(idlon1,idlat1,z,1:pasos));

```

## CAPÍTULO 3. METODOLOGÍA DEL MODELO HÍBRIDO WRF-SARIMA PARA LA CORRECCIÓN DE ERRORES DE PRONÓSTICO

```
v_11= squeeze(v_variable(idlon1 ,idlat1 ,z ,1: paso));
punto_11=sqrt((u_11.*u_11) + (v_11.*v_11));
u_12= squeeze(u_variable(idlon1 ,idlat2 ,z ,1: paso));
v_12= squeeze(v_variable(idlon1 ,idlat2 ,z ,1: paso));
punto_12=sqrt((u_12.*u_12) + (v_12.*v_12));
u_21= squeeze(u_variable(idlon2 ,idlat1 ,z ,1: paso));
v_21= squeeze(v_variable(idlon2 ,idlat1 ,z ,1: paso));
punto_21=sqrt((u_21.*u_21) + (v_21.*v_21));
u_22= squeeze(u_variable(idlon2 ,idlat2 ,z ,1: paso));
v_22= squeeze(v_variable(idlon2 ,idlat2 ,z ,1: paso));
punto_22=sqrt((u_22.*u_22) + (v_22.*v_22));

velocidad_interpolada = (1/((x2-x1)*(y2-y1)))*((punto_11*(x2-xp)*
(y2-yp))+(punto_21*(xp-x1)*(y2-yp))+(punto_12*(x2-xp)*(yp-y1))
+(punto_22*(xp-x1)*(yp-y1)));

%Grafica de la magnitud del viento en el sitio de las mediciones
y en los 4 puntos mas cercanos
xpuntos = 1:1:24;
plot(xpuntos ,punto_11 ,xpuntos ,punto_12 ,xpuntos ,punto_21 ,xpuntos ,
punto_22 ,xpuntos ,velocidad_interpolada)
```

### 3.3. Modelo Híbrido WRF-SARIMA

Este trabajo desarrolla un modelo de pronóstico de velocidad del viento, con un horizonte de múltiples pasos, que combina el pronóstico del modelo base WRF con una corrección de error basada en el modelo SARIMA. La metodología propuesta, la cual se muestra en la [Figura 3.7](#), consta de cinco etapas: (1) Pronóstico base WRF y mediciones de la velocidad del viento, (2) análisis de los errores de pronóstico, (3) modelado y predicción de los errores mediante SARIMA, (4) combinación de los pronósticos, y (5) evaluación del modelo híbrido.

La [Figura 3.8](#) ilustra el diagrama de flujo del modelo híbrido WRF-SARIMA. Durante el desarrollo del modelo híbrido WRF-SARIMA, se consideró como criterio esencial que cada una de las cinco fases que lo integran sea susceptible de automatización, lo que permitiría una aplicación operativa más sencilla en el contexto de los parques eólicos.

#### 3.3.1. Fase I. Obtención de la serie de tiempo de los errores de pronóstico del modelo WRF

A continuación se describen los pasos que se realizaron para obtener la serie de tiempo de los errores de pronóstico:

CAPÍTULO 3. METODOLOGÍA DEL MODELO HÍBRIDO WRF-SARIMA PARA LA CORRECCIÓN DE ERRORES DE PRONÓSTICO

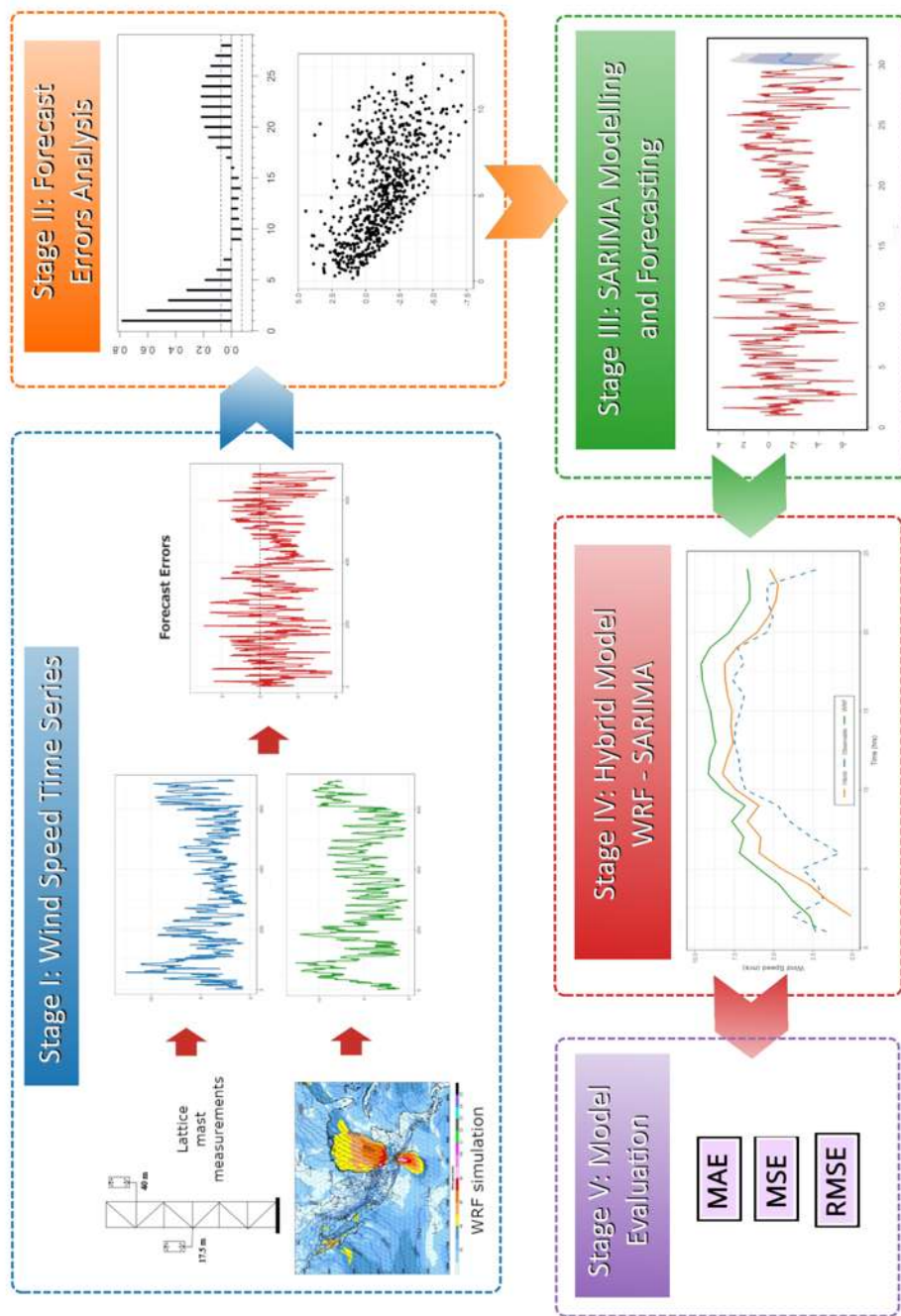


Figura 3.7: Diagrama de la metodología aplicada.

CAPÍTULO 3. METODOLOGÍA DEL MODELO HÍBRIDO WRF-SARIMA PARA LA CORRECCIÓN DE ERRORES DE PRONÓSTICO

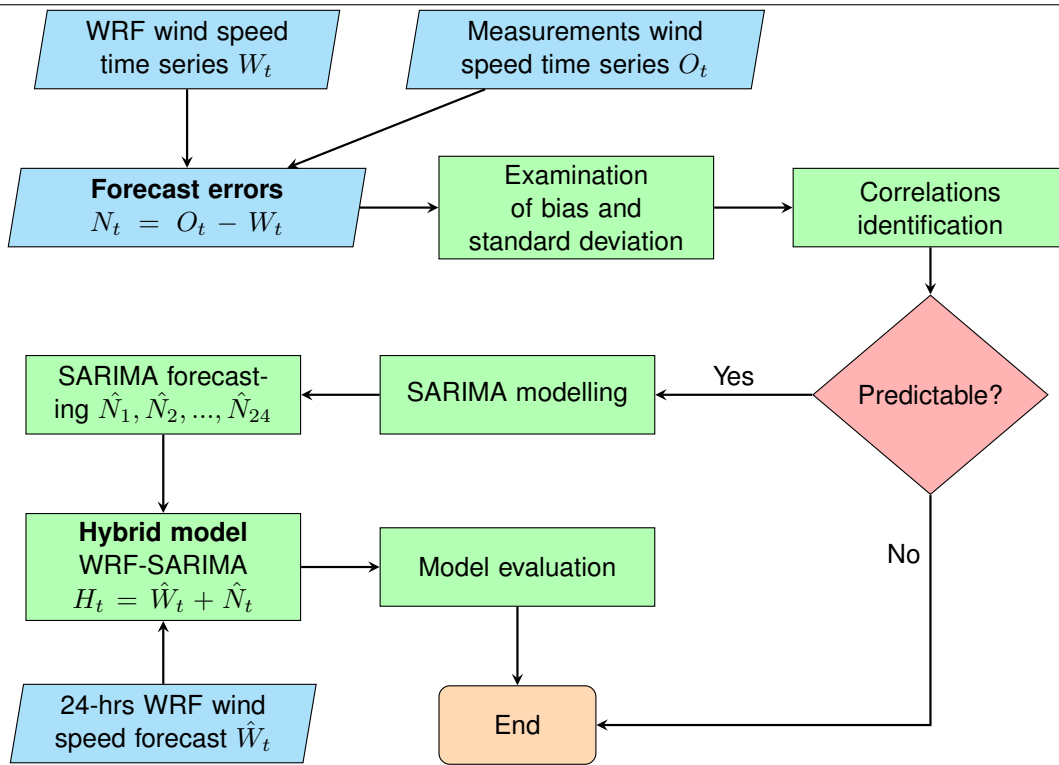


Figura 3.8: Diagrama de flujo del modelo WRF-SARIMA.

1. Se obtuvieron datos históricos de las mediciones de la velocidad del viento en el sitio seleccionado, identificando la ubicación exacta del sitio y la altura a la que fueron registradas.
2. Se realizó un pronóstico base con el modelo WRF.
3. Del pronóstico WRF, se extrajo la componente  $u$  y  $v$  de la velocidad del viento, para posteriormente calcular su magnitud mediante las formula  $\sqrt{u^2 + v^2}$ . Adicionalmente, se realizó una interpolación bilineal para hacer coincidir el sitio del pronóstico con el sitio de las mediciones, de acuerdo a la [Ecuación 3.1](#) anteriormente descrita.
4. La serie de tiempo de los errores de pronóstico  $N_t$  se obtuvo mediante la diferencia:

$$N_t = O_t - W_t \quad (3.2)$$

donde  $O_t$  representa a las mediciones (observaciones) de la velocidad del viento registradas en el sitio y  $W_t$  a la velocidad del viento pronosticada por el modelo WRF.

### CAPÍTULO 3. METODOLOGÍA DEL MODELO HÍBRIDO WRF-SARIMA PARA LA CORRECCIÓN DE ERRORES DE PRONÓSTICO

Tanto  $N_t$ ,  $O_t$  y  $W_t$  representan a series de tiempo. Las tres series tienen la misma resolución temporal de 1 hora, corresponden al mismo sitio, y tienen el mismo número de elementos, por lo tanto, cubren el mismo periodo de tiempo.

#### 3.3.2. Fase II. Análisis de los errores de pronóstico

En esta fase se incluyen las pruebas estadísticas realizadas a la serie de tiempo de los errores de pronóstico  $N_t$  para determinar la factibilidad de su predicción.

**Prueba 1:** Promedio de los errores  $\bar{N} = 0$ .

La primera prueba consiste en determinar si el promedio de los errores  $\bar{N}$  es igual a 0, como sería en el caso de una serie de sólo contiene errores aleatorios. Si la serie es aleatoria entonces no puede pronosticarse.

Para verificar este supuesto, se utilizó la *prueba-z*, cuyo valor  $z$  se calculó de acuerdo a la [Ecuación 3.3](#). La hipótesis nula  $H_0$  establece que  $\bar{N} = 0$ .

$$z = \frac{\bar{N}}{\sigma/\sqrt{n}} \quad \text{donde} \quad \bar{N} = \frac{\sum_{i=1}^n N_i}{n} \quad (3.3)$$

Es importante señalar que la prueba-z requiere de datos distribuidos normalmente. Por lo tanto, fue necesario realizar pruebas de normalidad a la serie temporal  $N_t$ , como son los gráficos Q-Q y la prueba de normalidad de Shapiro-Wilk, descrita por la [Ecuación 3.4](#).

$$\text{Shapiro} - \text{Wilk} = \frac{(\sum_{i=1}^n a_i N_{(i)})^2}{\sum_{i=1}^n (N_i - \bar{N})^2} \quad (3.4)$$

donde  $a_i$  son los pesos calculados de acuerdo con el tamaño de la muestra.

**Prueba 2:** Autocorrelación.

Las propiedades de autocorrelación de una serie de tiempo están estrechamente ligadas con su potencialidad para pronosticarse. Si en la serie de tiempo  $N_t$  se encuentran autocorrelaciones significativas, entonces la serie  $N_t$  puede predecirse.

En la [Subsección 1.8.4](#) se hace una descripción de cómo calcular estas autocorrelaciones.

**Prueba 3:** Prueba von Neumann ratio of the mean-square-successive difference.

La prueba *von Neumann ratio of the mean-square-successive difference*, calculada mediante la [Ecuación 3.5](#), se utiliza para determinar si los elementos de la serie son generados aleatoriamente, es decir, son valores independientes uno del otro.

La hipótesis nula  $H_0$  indica la ausencia de correlaciones seriales en valores consecutivos (para  $lag = 1$ ). En contraste, la hipótesis alternativa sugiere la presencia de

### CAPÍTULO 3. METODOLOGÍA DEL MODELO HÍBRIDO WRF-SARIMA PARA LA CORRECCIÓN DE ERRORES DE PRONÓSTICO

correlaciones seriales. Si no existen correlaciones seriales, entonces la serie fue generada aleatoriamente, y por lo tanto, no puede predecirse.

$$VN = \frac{\sum_{i=2}^n (N_i - N_{i-1})^2}{\sum_{i=1}^n (N_i - \bar{N})^2} \quad y \quad z = \frac{VN - E(VN)}{\sigma(VN)} \quad (3.5)$$

donde  $E(VN) = 2/n - 1$  y  $\sigma(VN) = \sqrt{4n^2(n-2)/(n-1)^3(n+1)}$ .

**Prueba 4:** Prueba Ljung-Box.

Las pruebas Portmanteau, tales como la prueba Ljung-Box, cuyos valores se calculan se acuerdo a la [Ecuación 3.6](#), son usado para determinar si la serie de tiempo es aleatoria o si por el contrario exhibe autocorrelaciones significativas para un  $lag = h$ . Si la serie posee autocorrelaciones significativas entonces es susceptible de ser pronosticada.

$$Q = n(n+2) \sum_{k=1}^h \frac{r_k^2}{n-k} \quad (3.6)$$

**Prueba 5:** Prueba run.

La prueba *run* verifica si la serie temporal  $N_t$  presenta un comportamiento aleatorio contando cuántas veces los elementos cambian de signo. El número de elementos positivos se denota por  $n_1$ , el número de elementos negativos se denota por  $n_2$ , y el número total de cambios de signo se denota por  $r$ .

$$z_r = \frac{r - E(r)}{\sigma(r)} \quad (3.7)$$

donde

$$E(r) = \frac{2n_1n_2}{n_1 + n_2} + 1, \quad \sigma(r) = \sqrt{\frac{(2n_1n_2)(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}$$

La hipótesis nula  $H_0$  establece que la serie fue generada aleatoriamente. La prueba run rechaza la hipótesis nula si

$$|z_r| > z_{1-\alpha/2} \quad (3.8)$$

donde  $\alpha$  es el nivel de significancia.

Si la serie fue generada aleatoriamente entonces no puede ser pronosticada.

**Prueba 6:** Gráficas de heterocedasticidad.

Las graficas de heterocedasticidad, es decir, de los errores estandarizados  $N_t$  frente a los valores pronosticados  $W_t$ , ilustran visualmente las tendencias y el grado de dispersión presentes en la serie  $N_t$ . Si las gráficas no presentan patrones, ni formas,

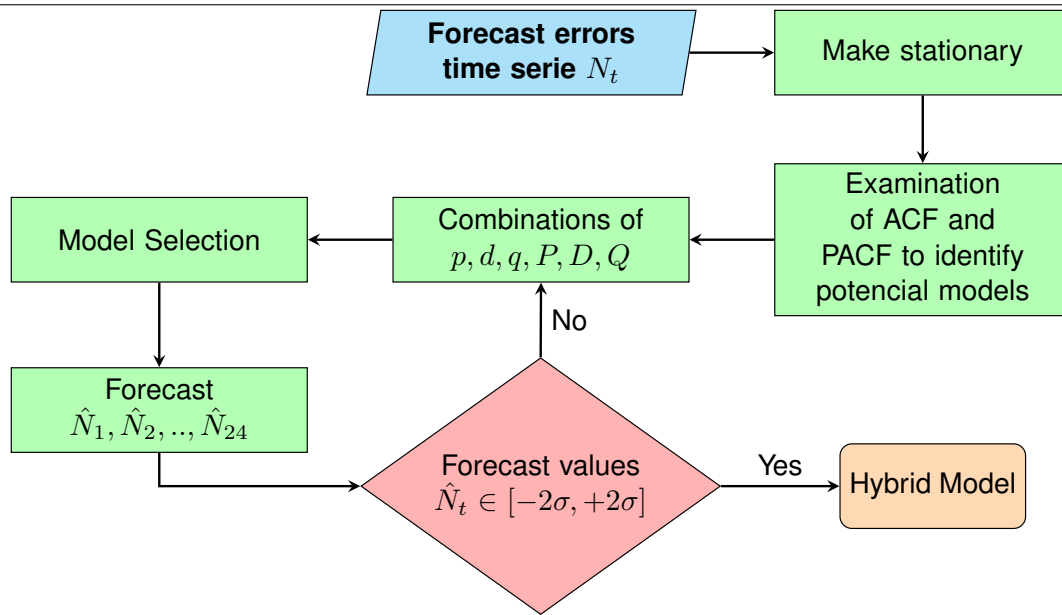


Figura 3.9: Diagrama de flujo del proceso de modelado y pronósticos con el modelo SARIMA.

ni tendencias, significa que los errores presentan con comportamiento azaroso, y por lo tanto no pueden ser pronosticados.

En conclusión, al demostrar estadísticamente que la serie  $N_t$  no es aleatoria y presenta correlaciones lineales significativas, se puede concluir entonces que la serie temporal  $N_t$  es predecible, lo que permite avanzar a la etapa del modelado y pronóstico.

### 3.3.3. Fase III. Modelado y pronóstico con SARIMA

En esta fase se llevó a cabo la predicción de los errores de pronóstico  $N_t$ , una vez que se determinó su factibilidad.

El proceso algorítmico propuesto en esta fase es una variación del propuesto por Box-Jenkins, descrito en la [Sección 3.3](#), incorporando la notación matemática de las fases anteriores y la condición de que los valores pronosticados estén dentro de un intervalo de dos desviaciones estándar.

La [Figura 3.9](#) ilustra el diagrama de flujo para el modelado y pronóstico de la serie temporal  $N_t$ , para un horizonte de pronóstico de 24 horas.

La metodología propuesta se divide en cuatro pasos:

- Estacionariedad: Se utilizó la prueba de Dickey-Fuller aumentada para verificar si la serie  $N_t$  es estacionaria. Para series temporales no estacionarias se

### CAPÍTULO 3. METODOLOGÍA DEL MODELO HÍBRIDO WRF-SARIMA PARA LA CORRECCIÓN DE ERRORES DE PRONÓSTICO

---

calculó la diferencia  $N_t - N_{t-1}$ , de aquí se obtuvo el valor de  $d$ . Una diferencia estacional  $N_t - N_{t-24}$  se calculó para series estacionales no estacionarias obteniendo el valor de  $D$ . Se recomienda un máximo de dos diferencias.

- Examen de ACF y PACF: Se calcularon los coeficientes de autocorrelación (ACF) y de autocorrelación parcial (PACF) para identificar potenciales modelos. Los valores  $r_k$  que caen fuera del intervalo de valores críticos sirvieron como una guía para determinar los rangos apropiados de las variables no estacionales  $p$  y  $q$ , y de las variables estacionales  $P$  y  $Q$ .
- Selección del modelo: Determinados los rangos apropiados de  $p, d, q, P, D$  y  $Q$ , se calcularon todas las posibles combinaciones. Cada combinación es un modelo potencial. Mediante asistencia computacional se calcularon los parámetros  $\phi, \theta, \Phi, \Theta$  de los modelos potenciales. Usando el Criterio de Información de Akaike corregido (AICc), se seleccionaron los 20 modelos con mejor desempeño. El modelo SARIMA final se eligió de acuerdo con el principio de Parsimonia, favoreciendo el modelo con menos parámetros.
- Pronóstico: Se realizó el pronóstico de la serie  $N_t$  (24 pasos adelante) utilizando el modelo seleccionado. Los valores pronosticados se denotan por  $\hat{N}_1, \dots, \hat{N}_{24}$ . Para asegurar que los valores pronosticados sean congruentes con la serie original  $N_t$ , se aplicó la condición de que los valores pronosticados  $\hat{N}_t$  estén dentro del intervalo  $[-2\sigma, +2\sigma]$ , donde  $\sigma$  denota la desviación estándar de la serie  $N_t$ . En el caso de que  $\hat{N}_t \notin [-2\sigma, +2\sigma]$ , se selecciona otro modelo potencial, omitiendo el modelo que no cumplió la condición.

Una opción para calcular los parámetros  $\phi, \theta, \Phi$  y  $\Theta$  fue utilizar la función `arima` del paquete `forecast` (Hyndman & Khandakar, 2008) en R (Team, 2025).

También se exploró la opción de usar el lenguaje Python para automatizar los pasos descritos anteriormente. A continuación se incluye una porción del código correspondiente al paso de la selección del modelo.

Listing 3.1: Función SARIMA

```
param_combinations = list(product(p_range, d_range, q_range, P_range,
D_range, Q_range))

def evaluate_sarima_model(train, order, seasonal_order):
    try:
        model = SARIMAX(train, order=order, seasonal_order=seasonal_order)
        model_fit = model.fit(dispatch=False)
```

## CAPÍTULO 3. METODOLOGÍA DEL MODELO HÍBRIDO WRF-SARIMA PARA LA CORRECCIÓN DE ERRORES DE PRONÓSTICO

---

```
        criterion = model_fit.aicc
        return criterion, order, seasonal_order
    except:
        return float('inf'), order, seasonal_order

max_combinations = 150 # Limitar el numero de combinaciones
for i, (p, d, q, P, D, Q) in enumerate(param_combinations
[:max_combinations]):
    order = (p, d, q)
    seasonal_order = (P, D, Q, s)
    score, order, seasonal_order = evaluate_sarima_model(train,
test, order, seasonal_order)
    results.append((score, order, seasonal_order))
    print(f'Combinacion{i+1}/{min(len(param_combinations),
max_combinations)}: ARIMA{order}x{seasonal_order}-AICc:{score:.4f}')

results_sorted = sorted(results, key=lambda x: x[0])
top = results_sorted[:20] # Los mejores 20 modelos
# Sumar los parametros
def calculate_param_sum(model):
    order = model[1]
    seasonal_order = model[2]
    return sum(order) + sum(seasonal_order[:3])

print('\nLos mejores modelos SARIMA:')
for i, (score, order, seasonal_order) in enumerate(top):
    print(f'{i+1}. ARIMA{order}x{seasonal_order}-AICc:{score:.4f}-
Suma parametros:{calculate_param_sum((score, order, seasonal_order))}')

# Seleccionar el modelo con menor suma de parametros entre los 20 mejores
best_model_info = min(top, key=lambda x: calculate_param_sum(x))
best_score, best_order, best_seasonal_order = best_model_info
```

### 3.3.4. Fase IV. Combinación de pronósticos

Es importante aclarar la diferencia entre dos series temporales provenientes del modelo WRF. Por un lado está la serie  $W_t$  que junto con las mediciones  $O_t$  se utilizó para calcular la serie de los errores de pronóstico  $N_t$ . Por otro lado está la serie  $\hat{W}_t$  que corresponde a un nuevo pronóstico de 24 horas adelante obtenido mediante el modelo WRF, esta nueva serie  $\hat{W}_t$  es la que se utiliza en esta fase.

Es importante también volver a señalar que la serie  $\hat{N}_t = \{\hat{N}_1, \dots, \hat{N}_{24}\}$  denota al pronóstico de los errores para un horizonte de 24 pasos adelante. Es decir,  $\hat{N}_1 = N_{t+1}, \dots, \hat{N}_{24} = N_{t+24}$ .

Ambas series  $\hat{W}_t$  y  $\hat{N}_t$  contiene 24 valores, y una resolución temporal de 1 hora, por lo tanto, ambas corresponden a pronósticos de 24 horas adelante.

### CAPÍTULO 3. METODOLOGÍA DEL MODELO HÍBRIDO WRF-SARIMA PARA LA CORRECCIÓN DE ERRORES DE PRONÓSTICO

---

El pronóstico final  $H_t$  del modelo híbrido WRF-SARIMA se define como:

$$H_t = \hat{W}_t + \hat{N}_t \quad (3.9)$$

Entonces, la serie de tiempo  $H_t$  representa al pronóstico de velocidad del viento generado por el modelo híbrido WRF-SARIMA para un horizonte de 24 horas.

Esta fase de la metodología también es automatizable.

#### 3.3.5. Fase V. Evaluación del modelo

Para la validación del modelo híbrido WRF-SARIMA se utilizaron nuevas observaciones de la velocidad del viento denotadas por la serie temporal  $\hat{O}_t$  que coinciden con el periodo de tiempo del pronóstico  $H_t$ .

El error se denota:

$$\varepsilon_t = \hat{O}_t - H_t \quad (3.10)$$

La evaluación del desempeño del modelo propuesto se realiza utilizando las métricas del error absoluto medio (MAE), el error cuadrático medio (MSE) y la raíz del error cuadrático medio (RMSE).

A continuación la demostración de porque es posible que los errores del modelo híbrido sean de menor magnitud que los errores del modelo WRF.

Sea  $e_t = \hat{O}_t - \hat{W}_t$  entonces:

$$e_t = (\varepsilon_t + H_t) - (H_t - \hat{N}_t) \quad (3.11)$$

$$e_t = \varepsilon_t + \hat{N}_t \quad (3.12)$$

Si  $\varepsilon_t$  y  $N_t$  tienen el mismo signo ( $\varepsilon_t N_t \geq 0$ ), entonces:

$$|\varepsilon_t| \leq |e_t| \quad (3.13)$$

Es decir, los residuales  $\varepsilon_t$  del modelo híbrido WRF-SARIMA son de menor magnitud que los  $e_t$  del modelo WRF bajo ciertas condiciones.

# Capítulo 4

## Resultados y discusiones

### 4.1. Descripción de los experimentos

Se realizaron cuatro experimentos para la aplicación el modelo propuesto:

- El Experimento I tuvo lugar del 1 al 29 de junio de 2018, a una altura de 40 metros.
- El Experimento II se desarrolló del 12 de noviembre al 10 de diciembre de 2018, también a una altura de 40 metros.
- El Experimento III cubrió el periodo del 1 al 29 de junio de 2018, pero a una altura de 17.5 metros.
- Y el Experimento IV se llevó a cabo del 12 de noviembre al 10 de diciembre de 2018, a 17.5 metros.

Cada experimento abarcó un periodo de 30 días a una resolución temporal horaria, resultando en 720 valores para cada experimento, excluyendo un período de 24 horas para el pronóstico.

La [Figura 4.1](#) ilustra una comparación entre las series temporales de velocidad del viento de las mediciones en el sitio y de las velocidades de las simulaciones WRF. En los Experimentos I y III se observan velocidades más bajas, mientras que los Experimentos II y IV velocidades más altas. A grandes rasgos se logra observar que las simulaciones WRF logran capturar las fluctuaciones y tendencias de la velocidad del viento. Esto nos podría llevar a suponer que el modelo WRF tiene un desempeño satisfactorio, pero hasta realizar un análisis a los errores se puede afirmar si el modelo WRF logro capturar todos los patrones existentes en las series de tiempo de la velocidad del viento.

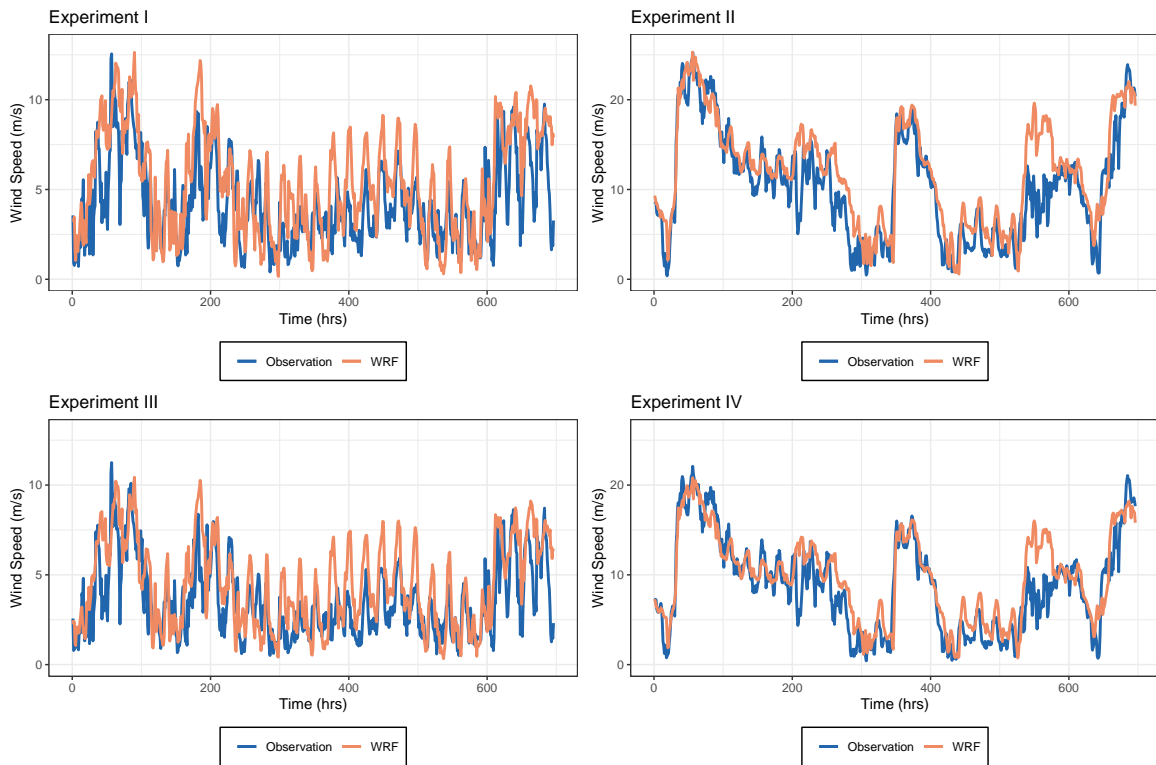


Figura 4.1: Series de tiempo de la rapidez del viento de las mediciones y de las simulaciones WRF.

Tabla 4.1: Promedio y desviación estándar de los errores de pronóstico.

| Experimento     | Promedio $\bar{N}$ / Sesgo | Desv. Est. $\sigma$ | Valor-z |
|-----------------|----------------------------|---------------------|---------|
| Experimento I   | -1.19                      | 2.06                | -15.17  |
| Experimento II  | -1.80                      | 2.78                | -17.09  |
| Experimento III | -0.91                      | 1.70                | -14.10  |
| Experimento IV  | -1.09                      | 2.44                | -11.78  |

## 4.2. Análisis de errores de pronóstico

De acuerdo a la metodología, la serie temporal  $N_t$  se calculó como la diferencia entre las mediciones del mástil  $O_t$  y las simulaciones  $W_t$  del modelo WRF.

En el Experimento I, la serie temporal  $N_t$  presentó un valor mínimo de -7.40 m/s y un valor máximo de 4.44 m/s. En el Experimento II, el valor mínimo fue -11.83 m/s, mientras que el máximo alcanzó 8.12 m/s. Para el Experimento III, el valor mínimo fue -5.98 m/s y el máximo fue 4.67 m/s. Finalmente, en el Experimento IV, el valor mínimo fue -9.35 m/s, mientras que el máximo fue 7.09 m/s. La desviación estándar para los experimentos fue:  $\pm 2.06$  m/s para el Experimento I,  $\pm 2.78$  m/s para el Experimento II,  $\pm 1.70$  m/s para el Experimento III, y  $\pm 2.44$  m/s para el Experimento IV.

La [Figura 4.2](#) muestra la gráfica de los errores de pronóstico para los cuatro experimentos. En la imagen se puede observar que los valores de los errores tienden a estar en el lado negativo, lo cual se confirmó al hacer el cálculo del sesgo (ver [Tabla 4.1](#)), mostrando que el modelo WRF sobrestimó los pronósticos de velocidad del viento.

**Prueba 1:** Promedio de los errores  $\bar{N} = 0$ .

De acuerdo con la metodología, se realizó una *prueba z* ([Ecuación 3.3](#)) para determinar si la media  $\bar{N}$  es estadísticamente diferente de cero. Antes de llevar a cabo la prueba *z*, se realizaron pruebas de normalidad sobre la serie temporal  $N_t$ , específicamente la prueba de Shapiro-Wilk ([Ecuación 3.4](#)) y los gráficos Q-Q (Ver [Figura 4.3](#)).

Los resultados indicaron que, con un nivel de significancia de  $\alpha = 0.05$ , los datos se distribuyeron normalmente en los cuatro experimentos. En la [Figura 4.3](#) se ajustó una función de densidad normal a los histogramas de los cuatro experimentos.

El valor-z para un nivel de significancia de  $\alpha = 0.05$  es  $z = 1.96$ . Dado que los valores *z* calculados para los experimentos son mucho menores que  $z = 1.96$  (ver [Tabla 4.1](#)), se rechaza la hipótesis nula, concluyendo que la media  $\bar{N}$  de los errores de pronóstico para los cuatro experimentos es estadísticamente diferente de cero.

**Prueba 2:** Autocorrelación.

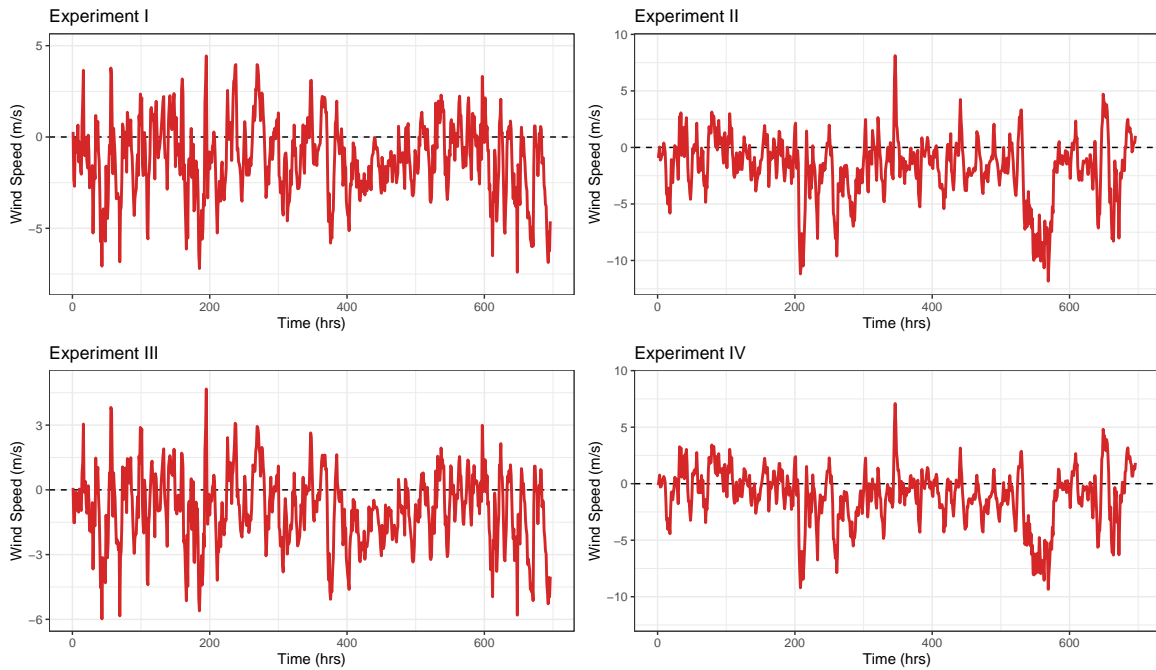


Figura 4.2: Serie de tiempo  $N_t$  de los errores de pronóstico.

A continuación, se realizó un análisis de dependencias secuenciales a la serie temporal  $N_t$  de errores de pronóstico. Las autocorrelaciones  $r_k$ , calculadas utilizando la [Ecuación 1.7](#), fueron determinadas para los primeros 28 rezagos. El intervalo de valores críticos fue:

$$\pm 1.96 / \sqrt{696} = \pm 0.0743 \quad (4.1)$$

Los gráficos de autocorrelación, mostrados en la [Figura 4.4](#), indican los siguientes resultados para cada experimento:

En el Experimento I, las autocorrelaciones  $r_1, \dots, r_6, r_{18}, \dots, r_{27}$  quedan fuera del intervalo de los valores críticos; en el Experimento II, las autocorrelaciones  $r_1, \dots, r_{28}$  están fuera del intervalo de los valores críticos; en el Experimento III, lo mismo pasa para las autocorrelaciones  $r_1, \dots, r_5, r_{19}, \dots, r_{27}$ ; y en el Experimento IV, las autocorrelaciones  $r_1, \dots, r_{28}$  también caen fuera del intervalo de los valores críticos.

A partir de los gráficos de la ACF, el valor  $r_{24}$  resulta significativo en cada experimento, mostrando patrones estacionales para un período de 24 horas.

En la [Tabla 4.2](#) se muestran los valores de autocorrelación  $r_k$  para el Experimento I.

La existencia de autocorrelaciones significativas en la serie  $N_t$  indica que es posible su modelado y pronóstico.

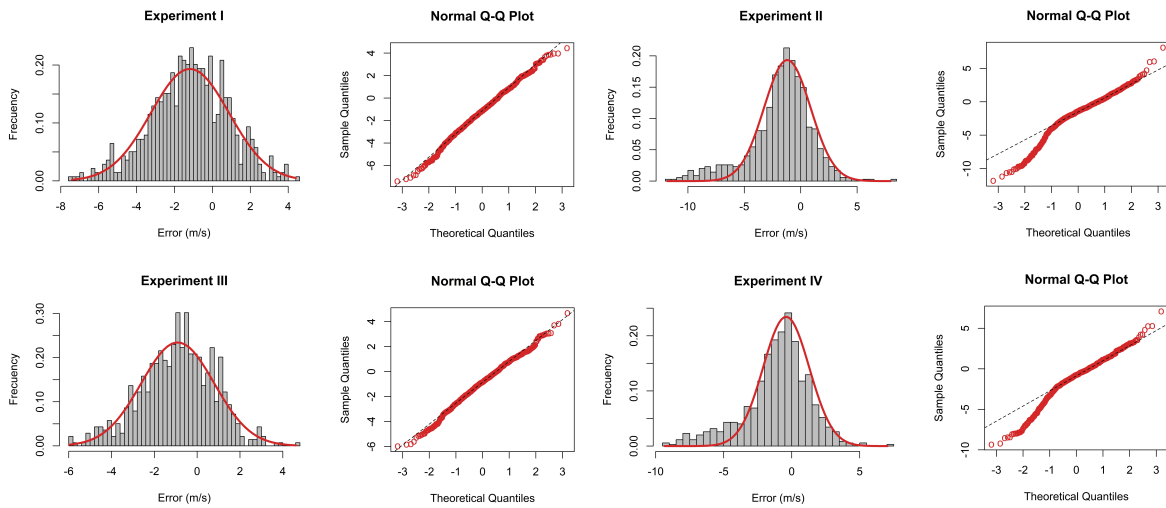


Figura 4.3: Histograma y gráfica Q-Q plot de la serie de tiempo  $N_t$ .

Tabla 4.2: Valores of autocorrelación  $r_k$

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

**Prueba 3:** Prueba von Neumann ratio of the mean-square-successive difference.

Se llevó a cabo una prueba *von Neumann ratio of the mean-square-successive difference* para verificar la existencia de correlaciones seriales en valores consecutivos, utilizando [Ecuación 3.5](#).

En el Experimento I se obtuvo un valor  $z = 20.70$  y un coeficiente  $VN = 0.43$ ; en el Experimento II se obtuvo un valor  $z = 23.22$  y un coeficiente  $VN = 0.24$ ; en el Experimento III se obtuvo un valor  $z = 20.91$  y un coeficiente  $VN = 0.42$ ; y en el Experimento IV se obtuvo un valor  $z = 23.47$  y un coeficiente  $VN = 0.22$ .

El puntaje  $z$  para un nivel de significancia de  $\alpha = 0.05$  es  $z = 1.96$ . Por lo tanto, en los cuatro experimentos se rechazó la hipótesis nula, favoreciendo la hipótesis alternativa, lo que indica la existencia de correlaciones seriales en el  $lag = 1$  de la serie temporal y aportando una prueba más que la serie  $N_t$  puede ser pronosticada.

**Prueba 4:** Prueba Ljung-Box.

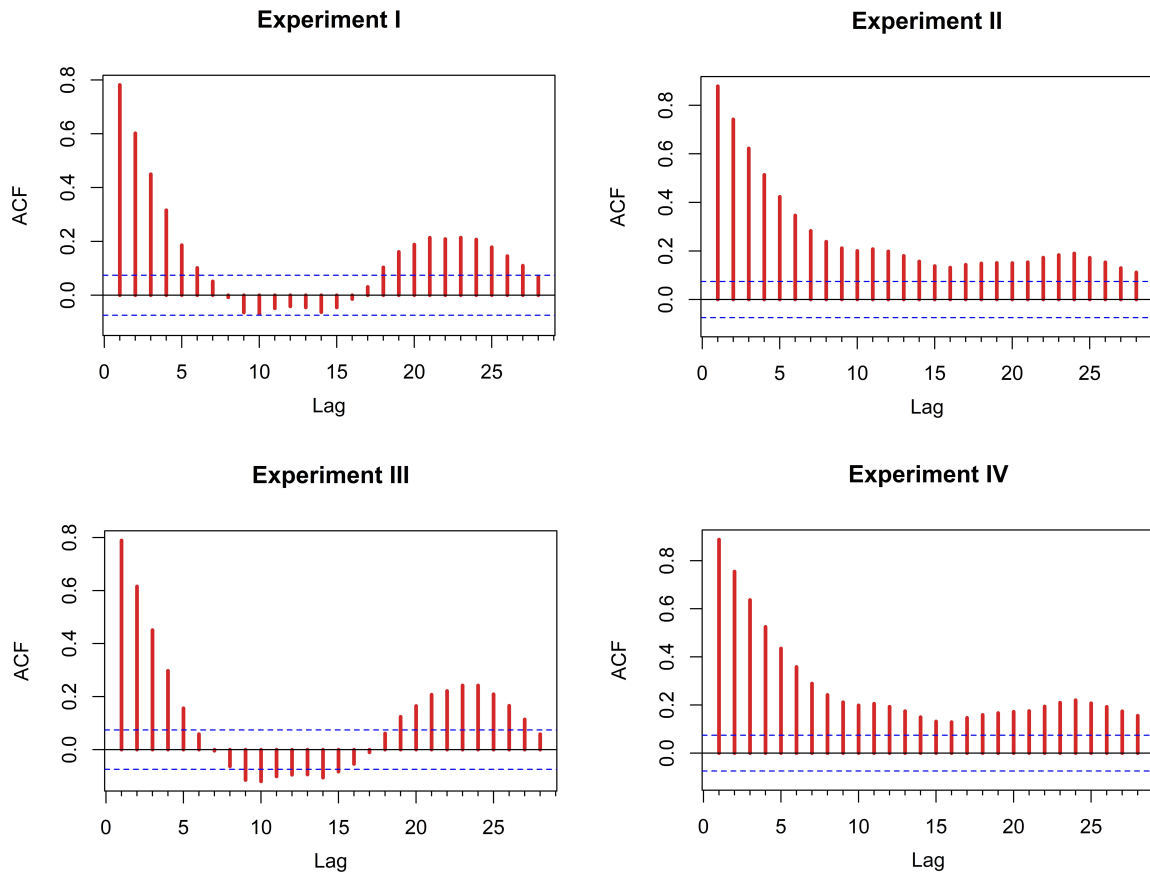


Figura 4.4: Función de autocorrelación (ACF) de la serie de tiempo  $N_t$ .

Los valores  $p$  calculados mediante la *prueba de Ljung-Box*, utilizando [Ecuación 3.6](#) para un rezago  $h = 24$ , permitieron rechazar la hipótesis nula en los cuatro experimentos. Esto respalda la hipótesis alternativa, que indica que los valores de la serie temporal  $N_t$  no están distribuidos independientemente y presentan correlaciones, y por lo tanto puede ser pronosticados.

**Prueba 5:** Prueba run.

La prueba *run* también se realizó para los cuatro experimentos, utilizando [Ecuación 3.7](#). En el Experimento I los resultados fueron  $r = 136$ ,  $n_1 = 192$ ,  $n_2 = 504$ , y un valor  $z = -16.12$ ; en el Experimento II los resultados fueron  $r = 123$ ,  $n_1 = 164$ ,  $n_2 = 532$ , y un valor  $z = -17.10$ ; en el Experimento III los resultados fueron  $r = 136$ ,  $n_1 = 209$ ,  $n_2 = 487$ , y un valor  $z = -16.12$ ; en el Experimento IV los resultados fueron  $r = 125$ ,  $n_1 = 226$ ,  $n_2 = 470$ , y un valor  $z = -16.96$ .

Los valores  $p$  obtenidos de los cuatro experimentos fueron lo suficientemente pequeños como para rechazar la hipótesis nula con un nivel de significancia de  $\alpha =$

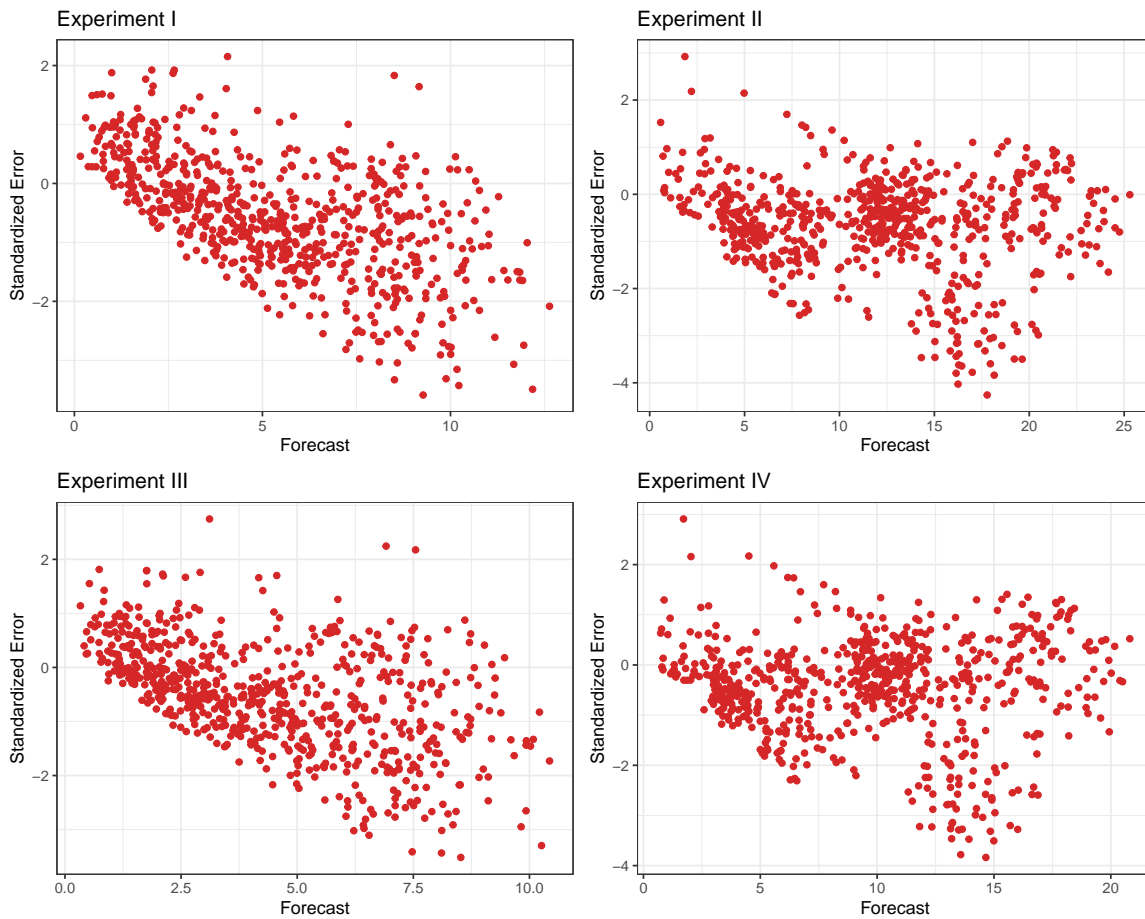


Figura 4.5: Gráfica de los errores estandarizados contra los valores pronosticados.

0.05, favoreciendo la hipótesis alternativa de que las series temporales  $N_t$  no son generadas aleatoriamente y tiene el potencial de ser pronosticados.

**Prueba 5:** Gráficas de heterocedasticidad.

Finalmente se realizaron gráficas de heterocedasticidad, las cuales se muestran en la [Figura 4.5](#), de los errores estandarizados  $N_t$  contra los valores pronosticados por WRF  $W_t$ .

Las gráficas muestran claramente que existe un patrón en la dispersión de los errores, lo que significa que no fueron generados aleatoriamente.

Las pruebas estadísticas aplicadas demostraron que la serie  $N_t$  no es ruido blanco, sino que presenta autocorrelaciones significativas. Por lo tanto, se concluye que dicha serie es predecible y que sus patrones remanentes pueden ser capturados mediante un modelo estadístico.

### 4.3. Modelado y pronóstico SARIMA

Se utilizó un conjunto de datos de 29 días de la serie temporal  $N_t$  para entrenar el modelo SARIMA en cada uno de los cuatro experimentos. Seguido de un pronóstico fuera de muestra para el día 30.

El modelado y pronóstico se realizaron siguiendo la metodología. Se llevó a cabo una prueba de Dickey-Fuller aumentada para determinar la estacionariedad de la serie temporal en cada experimento. De los resultados de la prueba Dickey-Fuller, los valores obtenidos fueron:  $d = 0$  y  $D = 0$  en el Experimento I,  $d = 1$  y  $D = 0$  en el Experimento II,  $d = 0$  y  $D = 0$  en el Experimento III, y  $d = 1$  y  $D = 0$  en el Experimento IV.

Dado que la serie temporal presenta un patrón estacional de 24 horas en los cuatro experimentos (ver [Figura 4.4](#)) se determinó un período estacional de  $s = 24$ .

Para identificar posibles modelos SARIMA, se examinaron las gráficas ACF y PACF. Las autocorrelaciones  $r_k$  que caen fuera del intervalo de valores críticos ayudaron a determinar el rango de las variables  $p, q, P$  y  $Q$ . En el Experimento I, la búsqueda de modelos potenciales se realizó con las variables  $p$  y  $P$  en el rango de 0 a 6, y  $q$  y  $Q$  también en el rango de 0 a 6. En el Experimento II, los rangos para  $p$  y  $P$  fueron de 0 a 9 y de 0 a 9 para  $q$  y  $Q$ . En el Experimento III, se trabajó en el rango de 0 a 5 para las variables  $p$  y  $P$ , y en el rango de 0 a 5 para las variables  $q$  y  $Q$ . En el Experimento IV, los rangos para  $p$  y  $P$  fueron de 0 a 9, y un rango similar para  $q$  y  $Q$ .

El *Criterio de Información de Akaike* (AIC) se utilizó para determinar los mejores 20 modelos SARIMA que mejor se ajustan a la serie temporal  $N_t$ . El modelo SARIMA final se selecciona de acuerdo con el principio de parsimonia, favoreciendo aquel con el menor número de parámetros (ver [Tabla 4.3](#)).

Para explorar una alternativa automatizable y de bajo costo computacional se probó la función `auto.arima` del paquete `forecast` ([Hyndman & Khandakar, 2008](#)) en el lenguaje R ([Team, 2025](#)), obteniendo resultados satisfactorios. En esta función las variables  $p, q, P$  y  $Q$  se eligen minimizando el AIC; las variables  $d$  y  $D$  se determinan mediante la prueba de unidad KPSS y la función devuelve el modelo con mejor ajuste basado en el AIC.

El modelo SARIMA seleccionado para el Experimento I fue  $SARIMA(2, 0, 1)(2, 0, 0)_{24}$ , el cual tiene la siguiente ecuación:

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

Tabla 4.3: Los 20 mejores modelos SARIMA para el Experimento I

| Modelo   | AICc | Suma | Modelo                                   | AICc        | Suma     |
|--|------|------|--|-------------|----------|
| SARIMA(1,0,5)(1,0,2) <sub>24</sub>                   | 2110 | 9    | SARIMA(2,0,4)(0,0,2) <sub>24</sub>       | 2118        | 8        |
| SARIMA(1,0,5)(2,0,2) <sub>24</sub>                   | 2112 | 10   | SARIMA(0,0,4)(1,0,2) <sub>24</sub>       | 2121        | 7        |
| SARIMA(0,0,5)(1,0,2) <sub>24</sub>                   | 2112 | 8    | SARIMA(1,0,3)(1,0,2) <sub>24</sub>       | 2121        | 7        |
| SARIMA(1,0,4)(1,0,2) <sub>24</sub>                   | 2113 | 8    | SARIMA(2,0,3)(1,0,2) <sub>24</sub>       | 2122        | 8        |
| SARIMA(1,0,5)(0,0,2) <sub>24</sub>                   | 2113 | 8    | SARIMA(1,0,2)(2,0,2) <sub>24</sub>       | 2122        | 7        |
| SARIMA(0,0,5)(2,0,2) <sub>24</sub>                   | 2114 | 9    | SARIMA(0,0,4)(2,0,2) <sub>24</sub>       | 2123        | 8        |
| SARIMA(1,0,4)(2,0,2) <sub>24</sub>                   | 2116 | 9    | <b>SARIMA(2,0,1)(2,0,0)<sub>24</sub></b> | <b>2123</b> | <b>5</b> |
| SARIMA(1,0,4)(0,0,2) <sub>24</sub>                   | 2116 | 7    | SARIMA(1,0,2)(1,0,2) <sub>24</sub>       | 2124        | 6        |
| SARIMA(2,0,4)(1,0,2) <sub>24</sub>                   | 2116 | 9    | SARIMA(1,0,3)(2,0,2) <sub>24</sub>       | 2124        | 8        |
| SARIMA(0,0,5)(0,0,2) <sub>24</sub>                   | 2117 | 7    | SARIMA(0,0,4)(0,0,2) <sub>24</sub>       | 2125        | 6        |
| Suma: Suma de las variables<br>$p, d, q, P, D$ y $Q$ |      |      |  |             |          |

En el Experimento II, el modelo seleccionado fue  $SARIMA(0, 1, 2)(0, 0, 2)_{24}$ , el cual tiene la siguiente ecuación:

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

En el Experimento III, el modelo seleccionado fue  $SARIMA(3, 0, 1)(0, 0, 2)_{24}$ , el cual tiene la siguiente ecuación:

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

Y en el Experimento IV, el modelo seleccionado fue  $SARIMA(0, 1, 2)(0, 0, 2)_{24}$ , el cual tiene la siguiente ecuación:

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

La [Figura 4.6](#) presenta los residuos del mejor modelo SARIMA para cada experimento, junto con la gráfica de autocorrelación y el histograma. Para un ruido blanco, se espera que alrededor del 95 % de los valores de autocorrelación caigan dentro del intervalo ([Makridakis et al., 2008](#)). Esto confirma que los residuos exhiben un comportamiento aleatorio, o de ruido blanco, validando de esta manera la selección del modelo.

Con el modelo SARIMA seleccionado se generó un pronóstico de 24 horas para cada experimento. La serie temporal resultante  $\hat{N}_t = \{\hat{N}_1, \dots, \hat{N}_{24}\}$  contiene los 24 valores

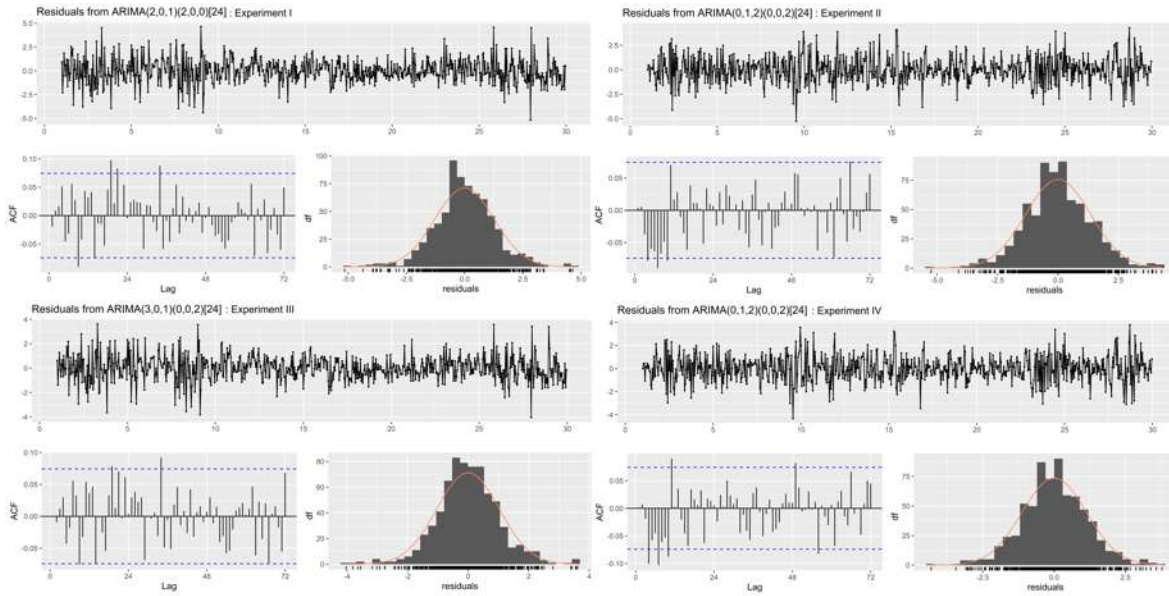


Figura 4.6: Verificación de residuos del modelo SARIMA con mejor ajuste, que exhibe un comportamiento de ruido blanco.

que representan los errores de pronóstico estimados por el modelo SARIMA. Estos valores serán combinados con el pronóstico del modelo WRF en la fase siguiente.

También se determinó si los valores pronosticados  $\hat{N}_t$  satisfacen la condición  $\hat{N}_t \in [-2\sigma, +2\sigma]$ ; es decir, si los valores predichos se encuentran dentro del intervalo de dos desviaciones estándar de  $N_t$ , con el fin de asegurar que los valores pronosticados  $\hat{N}_t$  sean congruentes con la serie original  $N_t$ .

En la [Figura 4.7](#) se muestra la serie temporal  $N_t$  y los valores pronosticados  $\hat{N}_t$ , así con sus intervalos de predicción al 80 % y 95 % para los cuatro experimentos.

#### 4.4. Combinación de pronósticos

El modelo híbrido  $H_t$  se construyó a partir de la suma del pronóstico del modelo WRF y los errores estimados por el modelo SARIMA, de acuerdo con la siguiente ecuación:

$$H_t = \hat{W}_t + \hat{N}_t, \quad t \in \{1, \dots, 24\} \quad (4.2)$$

donde  $\hat{W}_t$  representa el pronóstico de velocidad del viento generado por el modelo WRF para el día 30, y  $\hat{N}_t$  corresponde a los errores de pronóstico estimados por el

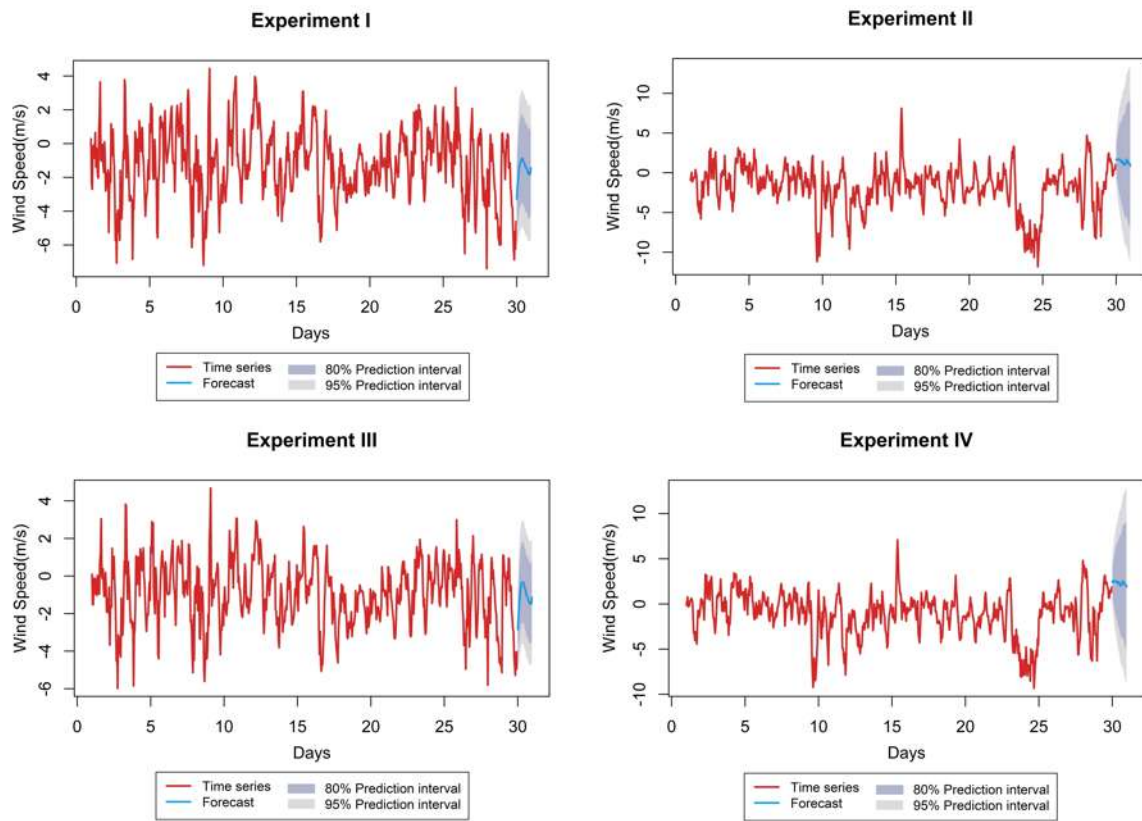


Figura 4.7: Pronóstico e intervalos de predicción de la serie de tiempo  $N_t$ .

modelo SARIMA también para el día 30.

Dado que ambas series temporales,  $\hat{N}_t$  y  $\hat{W}_t$ , constan de 24 valores horarios, la serie resultante  $H_t$  también contiene 24 elementos. En consecuencia,  $H_t$  constituye el pronóstico de velocidad del viento con horizonte de 24 horas generado por el modelo híbrido WRF-SARIMA.

La Figura 4.8 compara el pronóstico del modelo híbrido, el pronóstico del WRF y las mediciones reales para los cuatro experimentos.

Cabe señalar que la metodología propuesta es capaz de corregir tanto pronósticos sobreestimados del WRF (Experimentos I y III) como pronósticos subestimados (Experimentos II y IV), lo que demuestra su flexibilidad ante diferentes condiciones de la velocidad del viento.

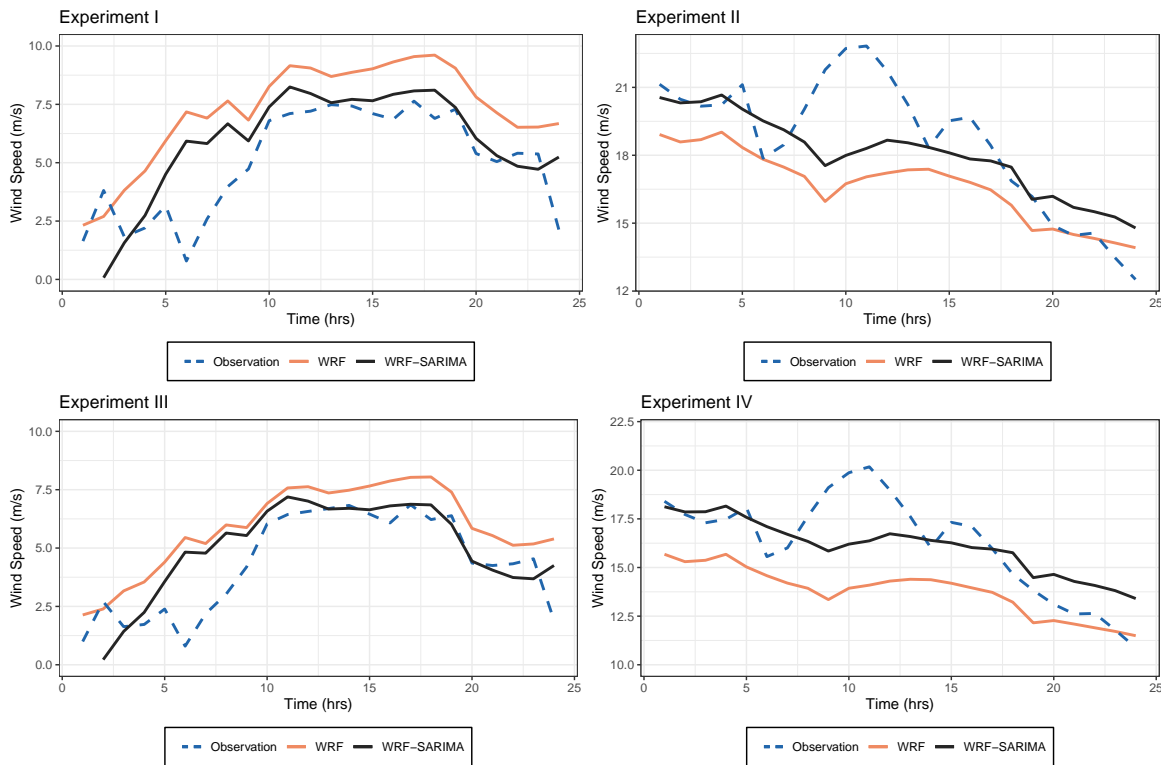


Figura 4.8: Pronóstico de 24 horas de la velocidad del viento.

## 4.5. Evaluación del modelo

El desempeño del modelo híbrido WRF-SARIMA se comparó con otros dos modelos híbridos: el modelo WRF-SES y el modelo WRF-Holt.

## CAPÍTULO 4. RESULTADOS Y DISCUSIONES

El modelo WRF-SES combina los pronósticos del modelo WRF con los errores de pronóstico estimados por el modelo ARIMA(0,1,1)/SES. De manera similar, el modelo WRF-Holt integra los pronósticos del WRF con aquellos generados por el método ARIMA(0,2,2)/Holt.

Los resultados de las métricas de error para cada modelo se presentan en la [Tabla 4.4](#). Así como % de mejora del modelo WRF-SARIMA en comparación con el modelo WRF.

Tabla 4.4: Comparación de los errores de pronóstico entre los diferentes modelos.

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

Los resultados de las métricas de error indicaron que el modelo híbrido WRF-SARIMA mostró mejor desempeño que el modelo WRF, el modelo WRF-SES y el modelo WRF-Holt en los cuatro experimentos.

En el Experimento I, el modelo híbrido WRF-SARIMA mostró una mejora del 42 % sobre el modelo WRF en la métrica MAE, una mejora del 49 % en la MSE y una mejora del 28 % en la RMSE. En el Experimento II, el modelo WRF-SARIMA presentó una mejora del 29 % en la métrica MAE, una mejora del 47 % en la MSE y una mejora del 27 % en la RMSE en comparación con el modelo WRF. En el Experimento III, el modelo WRF-SARIMA demostró una mejora del 36 % sobre el modelo WRF en la métrica MAE, una mejora del 40 % en MSE y una mejora del 22 % en RMSE. Final-

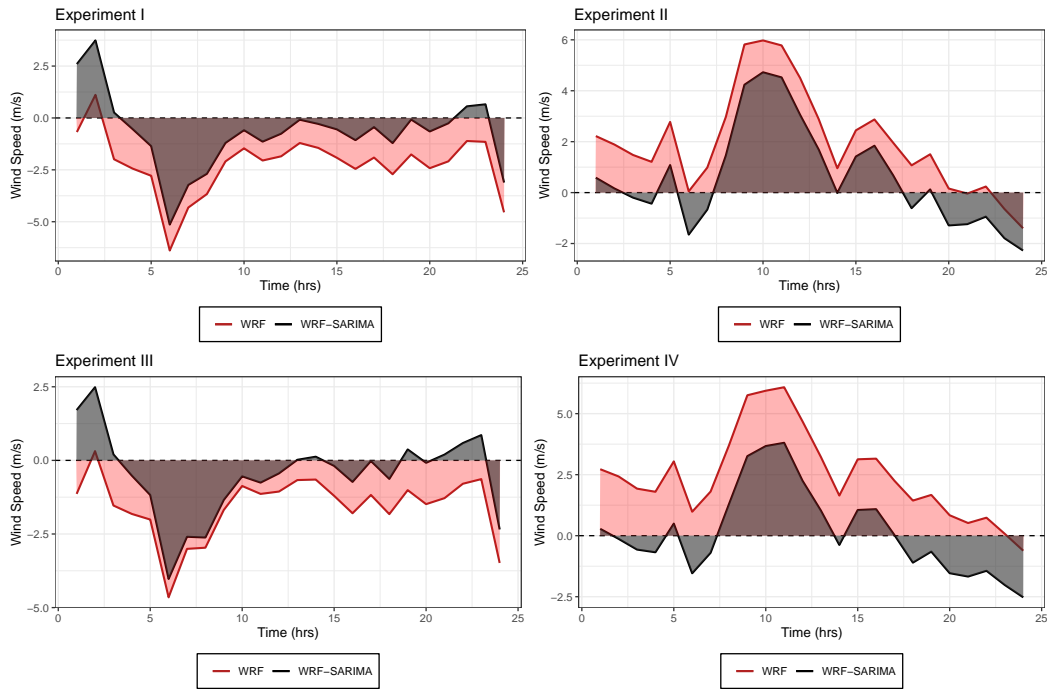


Figura 4.9: Comparación de los residuales de pronóstico del modelo WRF y del modelo WRF-SARIMA.

mente, en el Experimento IV, el modelo WRF-SARIMA alcanzó una mejora del 45 % sobre el modelo WRF en MAE, una mejora del 67 % en MSE y una mejora del 43 % en RMSE.

Sea  $e_t$  los residuos del modelo WRF y  $\varepsilon_t$  los residuos del modelo híbrido WRF-SARIMA. Como se observa en la [Figura 4.9](#), los residuos  $\varepsilon_t$  del modelo híbrido presentan una magnitud absoluta menor que los residuos  $e_t$  del modelo WRF, cumpliéndose que:

$$|\varepsilon_t| \leq |e_t| \quad (4.3)$$

Esta reducción, consistente en los cuatro experimentos, demuestra que el modelo híbrido disminuye la magnitud de los errores de pronóstico, lo que se traduce directamente en una mayor precisión de las predicciones de velocidad del viento.

Esta mejora es particularmente relevante en el contexto operativo de los parques eólicos, donde errores de menor magnitud implican una planificación más confiable de la generación y una reducción de las penalizaciones por desbalances en el mercado eléctrico.

## 4.6. Discusiones

### 4.6.1. Limitaciones asociadas a los datos

El desarrollo de este estudio implicó enfrentar desafíos significativos relacionados con la disponibilidad y calidad de los datos. Debido al carácter privado de la mayoría de los proyectos eólicos en el Istmo de Tehuantepec, la obtención de mediciones confiables de la velocidad del viento a alturas superiores a los 10 metros resultó particularmente difícil. Los datos empleados en esta investigación provinieron de diversos proyectos financiadas por instituciones públicas gubernamentales, lo que subraya la importancia del apoyo institucional para la generación de información de calidad en el sector energético.

Las mediciones de velocidad del viento en La Ventosa, registradas a 17.5 m y 40 m sobre el nivel del suelo, son confiables y de alta calidad. Sin embargo, no fue posible acceder a datos de igual calidad en otras ubicaciones, por lo que el modelo híbrido propuesto se validó exclusivamente en un solo sitio.

En cuanto a los datos de las simulaciones WRF, agradecemos la oportunidad de utilizar los resultados del modelo de pronóstico desarrollado por el grupo de investigación IOA en el ICAYCC-UNAM, cuyo modelo de pronóstico es continuamente actualizado y validado. No obstante, las limitaciones en nuestra capacidad de almacenamiento restringieron la evaluación del modelo a dos periodos de un mes cada uno: uno caracterizado por altas velocidades del viento y otro por bajas velocidades. Dado que se utilizaron conjuntos de datos relativamente cortos (30 días por experimento), existía el riesgo potencial de sobreajuste. Para mitigar este riesgo, se implementaron múltiples estrategias: Se aplicó el principio de Parsimonia, favoreciendo modelos SARIMA con el menor número posible de parámetros, se prestó especial atención a la correcta identificación de los periodos estacionales, garantizando que estos tuvieran sentido físico (ciclo diurno), se incorporó una condición de consistencia que exige que los valores pronosticados se encuentren dentro del intervalo de dos desviaciones estándar de la serie histórica, y se verificó que los residuos del modelo se comportaran como ruido blanco.

El conjunto de datos utilizado representa el mínimo para el cual el modelo SARI-MA logró identificar patrones estacionales y tendencias con suficiente claridad para generar pronósticos confiables. Es plausible que un periodo de entrenamiento más extenso (por ejemplo, varios meses o un año completo) permitiría al modelo identificar los patrones de la serie temporal con mayor precisión, conduciendo a una mejor

selección de modelo. Si bien se reconocen estas limitaciones, la robustez de la metodología empleada, junto con los resultados positivos obtenidos en La Ventosa, un sitio caracterizado por condiciones de alta turbulencia y por condiciones complejas, sugiere fuertemente que el modelo puede ser replicado efectivamente en diversas regiones y condiciones climáticas.

### 4.6.2. Generalización del modelo híbrido

La generalización del modelo propuesto a otros sitios depende de dos factores principales, por un lado de la posibilidad de generalizar el modelo WRF y por otro lado la generalización de la aplicación del modelo SARIMA.

El modelo WRF ha sido aplicado y probado en diferentes ubicaciones ([Chen et al., 2019](#); [Dzebre & Adaramola, 2020](#); [Duarte Jacondino et al., 2021](#); [Liu et al., 2021a, 2023](#); [Tsai et al., 2021](#); [Xu et al., 2021](#); [Yang et al., 2023b](#); [Zhang et al., 2024](#)), ya sea para pronósticos de velocidad del viento o para pronósticos de generación eólica, demostrando consistentemente ser una herramienta confiable para horizontes de 24 horas.

La replicabilidad de los modelos SARIMA, conforme a la metodología propuesta, depende fundamentalmente de la disponibilidad de mediciones de la velocidad del viento en el sitio de interés. Esta dependencia es común en cualquier enfoque de corrección de errores que requiera datos históricos para el entrenamiento.

### 4.6.3. Comparación con otros modelos

Se comparó el desempeño del modelo WRF-SARIMA con otros modelos híbridos reportados en la literatura para pronósticos de velocidad del viento a 24 horas. El modelo propuesto alcanzó reducciones del error absoluto medio (MAE) entre 29 % y 45 %, y de la raíz del error cuadrático medio (RMSE) entre 22 % y 43 %.

[Xu et al. \(2021\)](#) reportaron mejoras de 24 % y 31 % en MAE, y de 27 % y 28 % en RMSE, utilizando un modelo de memoria a largo plazo (LSTM) para corregir errores de pronóstico de velocidad del viento generados con un modelo WRF. Por su parte, [Zhao et al. \(2024\)](#) reportaron mejoras de 15 % en MSE y 14 % en RMSE empleando redes convolucionales temporales y modelos LSTM para reducir errores de pronóstico de un modelo WRF. Finalmente, [Wang et al. \(2019\)](#) reportaron mejoras de 1.7 % a 3.9 % en RMSE al combinar modelos de aprendizaje automático para corregir errores de pronóstico de velocidad del viento de un modelo NWP.

Estos resultados permiten concluir que los porcentajes de mejora alcanzados por el modelo WRF-SARIMA se sitúan dentro del mismo rango que los reportados por enfoques basados en aprendizaje automático, e incluso superan a algunos de ellos. Esta equivalencia en la mejora de los pronósticos es particularmente relevante si se considera que el modelo SARIMA ofrece ventajas adicionales: interpretabilidad, menor requerimiento de datos y menor costo computacional.

### **4.6.4. Implicaciones operativas para la generación eólica**

La potencia suministrada por un aerogenerador se calcula a partir de la velocidad del viento y de la curva de potencia de la turbina. En consecuencia, pronósticos confiables de la velocidad del viento se traducen directamente en estimaciones más confiables de la generación eólica esperada. Las mejoras en los pronósticos de velocidad del viento son operacionalmente muy valiosas para los operadores de sistemas eléctricos, ya que permiten:

- Estimar con mayor certeza la potencia disponible en el parque eólico.
- Facilitar ajustes en el despacho.
- Mejorar la estabilidad de la frecuencia en la red eléctrica.
- Reducir las penalizaciones económicas por desbalances en los mercados eléctricos cuando los productores no cumplen con sus compromisos programados.

### **4.6.5. Limitaciones del enfoque SARIMA**

Es importante reconocer que los modelos SARIMA son inherentemente lineales y, por lo tanto, efectivos para identificar patrones lineales y estacionales en datos de series temporales. Sin embargo, su desempeño puede verse limitado frente a patrones más complejos, como relaciones no lineales o cambios abruptos en la variación del viento. Para abordar esta limitación, investigaciones futuras podrían explorar la combinación de modelos estadísticos no lineales (por ejemplo, redes neuronales) con modelos lineales como SARIMA, integrando ambas aproximaciones en una estrategia de corrección de error más flexible.

Asimismo, se ha observado que los valores inmediatamente precedentes afectan significativamente los pronósticos generados por los modelos ARIMA. Como resultado, la transición entre la fase de entrenamiento y la fase de pronóstico, es decir, el punto

en el que se corta la serie histórica y comienza la predicción, es crítica en la determinación de la tendencia del pronóstico. Para trabajos futuros, sería útil desarrollar métodos que identifiquen el punto inicial óptimo para el pronóstico, posiblemente mediante la detección de puntos de cambio en la media, la varianza y la tendencia de la serie temporal.

Los hallazgos de este estudio se alinean con investigaciones previas sobre los beneficios de las técnicas estadísticas para mejorar los pronósticos de velocidad del viento derivados de modelos NWP. La combinación de los modelos WRF y SARIMA demuestra ser un enfoque viable y efectivo para mejorar los pronósticos de velocidad del viento, particularmente en horizontes de 24 horas relevantes para el mercado eléctrico.

Adicionalmente, la metodología propuesta ha sido diseñada con un énfasis explícito en su capacidad de automatización, lo que facilita su implementación en escenarios operativos en tiempo real. Esta característica es fundamental para su adopción práctica en la industria eólica. Esto contribuye a una mejor integración de la energía eólica en las redes eléctricas, a la reducción de costos operativos para los propietarios de parques eólicos y a la minimización de penalizaciones por desbalances energéticos.

# Conclusiones y trabajo futuro

El presente trabajo de investigación tuvo como objetivo desarrollar un modelo híbrido de pronóstico de velocidad del viento con horizonte de 24 horas, combinando el modelo físico Weather Research and Forecasting (WRF) con el modelo estadístico Seasonal Autoregressive Integrated Moving Average (SARIMA), con la finalidad de mejorar los pronósticos de la velocidad del viento. Para ello, se implementó una estrategia de corrección de errores fundamentada en la identificación de dependencias secuenciales y patrones remanentes en los errores de pronóstico del modelo WRF. El sitio de experimentación seleccionado fue La Ventosa, Oaxaca, una región caracterizada por su alto potencial eólico pero también por sus condiciones de alta turbulencia.

A partir de los resultados obtenidos, se derivan las siguientes conclusiones:

- El modelo WRF-SARIMA mostró un desempeño consistente tanto en periodos de bajas velocidades del viento como en periodos de altas velocidades, así como en las dos alturas de medición evaluadas (17.5 m y 40 m). Esta congruencia en diferentes condiciones operativas respalda su potencial de aplicabilidad en diferentes contextos.
- El proceso de análisis de los errores de pronóstico, junto con su posterior modelado mediante SARIMA, demostró ser una herramienta eficaz para identificar patrones remanentes y estructuras predecibles dentro de dichos errores. Este enfoque facilita correcciones más precisas y fundamentadas de los pronósticos originales.
- Las métricas de error indican que el modelo híbrido WRF-SARIMA presenta un mejor desempeño que el modelo WRF por sí solo, mostrando una reducción significativa en los errores. Las disminuciones del error medio absoluto (MAE) fueron en el rango de 29 % - 44 %, las reducciones del error cuadrático medio (MSE) estuvieron en el intervalo de 39 % - 67 %, y para la raíz del error

cuadrático medio (RMSE) las disminuciones porcentuales en los errores fueron entre 22 % - 42 %.

- Comparando el desempeño del modelo propuesto con el desempeño de otros modelos híbridos similares encontrados en la literatura, basados principalmente en aprendizaje automático, se observa que los porcentajes de mejora alcanzados por el modelo WRF-SARIMA son similares a los porcentajes de mejora reportados por dichos enfoques, indicando que el desempeño del presente modelo está a la par con los modelos publicados hasta ahora. Lo cual es significativo si se considera que el modelo SARIMA opera con menores requerimientos de datos y recursos computacionales.
- El modelo híbrido propuesto posee diferentes fortalezas: una base matemática sólida y bien establecida, capacidad de entrenamiento con conjuntos de datos reducidos, y una mayor interpretabilidad, aspecto importante en el sector energético. Su mayor debilidad es la linealidad del modelo SARIMA. Sin embargo, los resultados obtenidos demuestran que esta limitación no impide predecir patrones complejos y oscilatorios como los presentes en las series de velocidad del viento en La Ventosa.

En síntesis, el enfoque desarrollado mejora la comprensión de los errores de pronóstico y propone ajustes más efectivos para mejorar los pronóstico de la velocidad del viento. El modelo WRF-SARIMA representa una alternativa prometedora para el pronóstico de velocidad del viento, especialmente por su potencial de automatización y su facilidad de implementación operativa, características que lo distinguen de enfoques más complejos y menos interpretables.

### **4.7. Trabajo futuro**

A partir de las limitaciones identificadas y de las preguntas que surgieron durante esta investigación, se plantean las siguientes posibles líneas de trabajo futuro:

Como se observó durante el desarrollo del estudio, los valores inmediatamente precedentes al inicio del pronóstico tienen una influencia significativa en las predicciones generadas por los modelos SARIMA. Se propone investigar métodos para identificar el punto óptimo de transición entre la fase de entrenamiento y la fase de pronóstico, posiblemente mediante la detección automática de puntos de cambio en la media, la varianza y la tendencia de la serie temporal.

Dado que la turbulencia del viento representa un desafío para los modelos de pronóstico actuales, se propone explorar la incorporación de técnicas fractales que permitan caracterizar y predecir las fluctuaciones turbulentas que los modelos convencionales no logran capturar adecuadamente.

Si bien el presente estudio se centró en el horizonte de 24 horas por su relevancia para el mercado eléctrico, resulta de interés extender el análisis a otros horizontes temporales (por ejemplo, 6, 12, 48 y 72 horas) para evaluar el desempeño del modelo híbrido en diferentes ventanas de predicción y determinar el rango más óptimo para su aplicación.

En línea con los avances recientes en aprendizaje automático, es de interés explorar la aplicación de nuevos modelos de deep learning, como los transformers, y su posible combinación con el modelo WRF, ya sea como reemplazo del componente SARIMA o como complemento del mismo.

Se tiene previsto realizar una evaluación comparativa del desempeño de diferentes enfoques híbridos para la corrección de errores del WRF, incluyendo al menos: WRF-SARIMA (presentado en este trabajo), WRF-ANN (redes neuronales artificiales), y WRF-Cascada Multiplicativa. Esta comparación permitirá identificar las fortalezas relativas de cada enfoque en función de las características del sitio y los recursos disponibles.




En conjunto, estas líneas de investigación futura permitirán consolidar y extender el enfoque desarrollado, contribuyendo a la mejora continua de los pronósticos de velocidad del viento y, por consiguiente, a una integración más eficiente y confiable de la energía eólica en los sistemas eléctricos.

## Artículo publicado sobre la presente tesis

|                         |  |
|-------------------------|--|
| <b>Título</b>           | Hybrid WRF–SARIMA model to improve day-ahead wind speed forecast accuracy  |
| <b>Autores</b>          | Bernabe, M.; Cadenas, E.; Lopez-Espinoza, E.; Campos-Amezcu, R.  |
| <b>Año</b>              | 2026   |
| <b>Revista</b>          | International Journal of Renewable Energy Development  |
| <b>Volumen y Número</b> | Vol. 15, No. 1   |
| <b>Páginas</b>          | 76–88  |
| <b>DOI</b>              | <a href="https://doi.org/10.61435/ijred.2026.61746">https://doi.org/10.61435/ijred.2026.61746</a>  |
| <b>Resumen</b>          | El artículo propone un modelo híbrido que combina el sistema de predicción meteorológica WRF con el modelo estadístico SARIMA para mejorar los pronósticos de la velocidad del viento con un horizonte de 24 horas. Los resultados muestran que la integración de ambos enfoques reduce los errores de pronóstico y ofrece una herramienta más confiable para la planificación de recursos energéticos renovables. |



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

Maritza Bernabe<sup>a</sup> , Erasmo Cadenas<sup>a\*</sup> , Erika López-Espinoza<sup>b</sup> , Rafael Campos-Amezcu<sup>a</sup> 

<sup>a</sup>Faculty of Mechanical Engineering, Universidad Michoacana de San Nicolás de Hidalgo, Michoacán, México

<sup>b</sup>Institute of Atmospheric Sciences and Climate Change, National Autonomous University of Mexico, México

<sup>c</sup>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 RMSE 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.



@ The author(s). Published by CBIORÉ. This is an open access article under the CC BY-SA license (<http://creativecommons.org/licenses/by-sa/4.0/>).

Received: 4<sup>th</sup> Sept 2025; Revised: 17<sup>th</sup> Oct 2025; Accepted: 28<sup>th</sup> Nov 2025; Available online: 10<sup>th</sup> 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 (Secretaría de Energía [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)

Figura 4.10: Portada del artículo *Hybrid WRF–SARIMA model to improve day-ahead wind speed forecast accuracy*.

# Bibliografía

- 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>.
- Du, P., Yang, D., Li, Y., & Wang, J. (2024). An innovative interpretable combined learning model for wind speed forecasting. *Applied Energy*, 358, 122553. <https://doi.org/10.1016/j.apenergy.2023.122553>.
- 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>.
- Duarte Jacondino, W., da Silva Nascimento, A. L., Calvetti, L., Fisch, G., Augustus 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>.
- 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>.

## BIBLIOGRAFÍA

---

- 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.
- Elsevier (2024). Scopus database search: "wind speed forecast". <https://www.scopus.com>. Consulta realizada en noviembre de 2024. Resultados: 7341 documentos.
- GOBMEX (2017). *Manual de Pronosticos*. Diario Oficial de la Federación (Mexico). [https://www.dof.gob.mx/nota\\_detalle.php?codigo=5505475&fecha=23/11/2017#gsc.tab=0](https://www.dof.gob.mx/nota_detalle.php?codigo=5505475&fecha=23/11/2017#gsc.tab=0).
- Grupo IOA (2024). Ioa group of icaycc unam. <http://grupo-ioa.atmosfera.unam.mx/>. Accessed: 2025-03-11.
- Guo, Z., Zhao, J., Zhang, W., & Wang, J. (2011). A corrected hybrid approach for wind speed prediction in hexi corridor of china. *Energy*, 36(3), 1668–1679. <https://doi.org/10.1016/j.energy.2010.12.063>.
- GWEC (2020). Gwec global wind report 2019. *Global Wind Energy Council: Bonn, Germany*.
- Hagens, N. J. (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.
- 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>.
- ICAYCC (2024). Instituto de ciencias de la atmósfera y cambio climático (icaycc) de la unam. <https://www.atmosfera.unam.mx/>. Accessed: 2025-03-11.

## BIBLIOGRAFÍA

---

- INEEL (2025). Atlas de energía y medio ambiente - lista de proyectos meteorológicos del instituto nacional de electricidad y energías limpias (ineel). Accedido el 3 de mayo de 2025.  
URL <https://aems.ineel.mx/ProjectListText.aspx?&Rnd=788679>
- Jaramillo, O. A., & Borja, M. A. (2004). Wind speed analysis in la ventosa, mexico: a bimodal probability distribution case. *Renewable energy*, 29(10), 1613–1630. <https://doi.org/10.1016/j.renene.2004.02.001>.
- 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, Mexico.
- 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.
- Liu, M.-D., Ding, L., & Bai, Y.-L. (2021a). Application of 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, 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. (2021b). 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.2019.114137>.
- 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 mexico. *Atmosphere*, 11(11), 1242. <https://doi.org/10.3390/atmos11111242>.

- Lopez-Villalobos, C. A., Rodriguez-Hernandez, O., Campos-Amezcuca, 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. [https://doi.org/10.1016/0169-2070\(89\)90017-4](https://doi.org/10.1016/0169-2070(89)90017-4).
- 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.
- 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, Mexico.
- Romero-Centeno, R., Zavala-Hidalgo, J., Gallegos, A., & O'Brien, J. J. (2003). Isthmus 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](https://doi.org/10.1175/1520-0442(2003)016<2628:IOTWCA>2.0.CO;2).
- San Miguel, G., Regueiro Ferreira, R., Gomez-Catusus, J., & Gomez Villarino, M. (2023). *Tecnología y sostenibilidad de la energía eólica. Preguntas y respuestas*. Ediciones Mundi-Prensa.
- SENER (2024). Informe pormenorizado sobre el desempeño y las tendencias de la industria eléctrica nacional 2023. Tech. rep., Secretaria de Energia (SENER).
- Singh, S., Mohapatra, A., et al. (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>.

- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., et al. (2019). A description of the advanced research wrf version 4. *NCAR tech. note*, 145(10.5065).
- Sun, F., & Jin, T. (2022). A hybrid approach to multi-step, short-term wind speed forecasting using correlated features. *renewable energy*, 186, 742–754. <https://doi.org/10.1016/j.renene.2022.01.041>.
- Talayero, A., & Telmo, E. (2011). *Energía eólica (Serie Energías renovables) (2ª ed.)*. Textos Docentes. Prensas de la Universidad de Zaragoza.
- 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.
- 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>.
- Villarubia, M. (2007). *Energía eólica*. Energías Alternativas y Medioambiente. Ceac.
- Villarubia, M. (2013). *Ingeniería de la energía eólica*. Alpha Editorial.
- 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>.
- 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. (2023a). 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. (2023b). 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>.


## BIBLIOGRAFÍA

---

- Zhang, D., Hu, G., Song, J., Gao, H., Ren, H., & Chen, W. (2024). A novel spatio-temporal 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, J., Wei, Y., & Tan, Z. (2020). An adaptive hybrid model for short term wind speed forecasting. *Energy*, 190, 115615. <https://doi.org/10.1016/j.energy.2019.06.132>.
- 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). Gwec global wind report 2024. *Global Wind Energy Council: Bonn, Germany*.
- 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 modeling. *Energy*, (p. 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/10.1016/j.energy.2020.119509>.

# Maritza Talhia Bernabé Morales

## Modelo híbrido WRF-SARIMA para mejorar los pronósticos de la velocidad del viento.pdf

 Universidad Michoacana de San Nicolás de Hidalgo

### Detalles del documento

Identificador de la entrega

trn:oid:::3117:578926466

Fecha de entrega

15 abr 2026, 11:06 a.m. GMT-6

Fecha de descarga

15 abr 2026, 11:10 a.m. GMT-6

Nombre del archivo

Modelo híbrido WRF-SARIMA para mejorar los pronósticos de la velocidad del viento.pdf

Tamaño del archivo

15.4 MB

82 páginas

21.818 palabras

112.676 caracteres

# Formato de Declaración de Originalidad y Uso de Inteligencia Artificial

Coordinación General de Estudios de Posgrado  
Universidad Michoacana de San Nicolás de Hidalgo



A quien corresponda,

Por este medio, quien abajo firma, bajo protesta de decir verdad, declara lo siguiente:

- Que presenta para revisión de originalidad el manuscrito cuyos detalles se especifican abajo.
- Que todas las fuentes consultadas para la elaboración del manuscrito están debidamente identificadas dentro del cuerpo del texto, e incluidas en la lista de referencias.
- Que, en caso de haber usado un sistema de inteligencia artificial, en cualquier etapa del desarrollo de su trabajo, lo ha especificado en la tabla que se encuentra en este documento.
- Que conoce la normativa de la Universidad Michoacana de San Nicolás de Hidalgo, en particular los Incisos IX y XII del artículo 85, y los artículos 88 y 101 del Estatuto Universitario de la UMSNH, además del transitorio tercero del Reglamento General para los Estudios de Posgrado de la UMSNH.

| <b>Datos del manuscrito que se presenta a revisión</b> |   |                                     |
|--|---|-------------------------------------|
| <b>Programa educativo</b>                              | Doctorado en Ciencias en Ingeniería Mecánica                                      |                                     |
| <b>Título del trabajo</b>                              | Modelo híbrido WRF-SARIMA para mejorar los pronósticos de la velocidad del viento |                                     |
|  | <b>Nombre</b>   | <b>Correo electrónico</b>           |
| <b>Autor/es</b>  | Maritza Talhia Bernabé Morales  | 0915588d@umich.mx                   |
| <b>Director</b>  | Erasmus Cadenas Calderón  | ecadenas@umich.mx                   |
| <b>Codirector</b>                                      | Víctor López Garza  | victor.garza@umich.mx               |
| <b>Coordinador del programa</b>                        | Gilberto Gonzalez Avalos  | doc.cs.ingenieria.mecanica@umich.mx |


| <b>Uso de Inteligencia Artificial</b> |                    |                    |
|---------------------------------------|--------------------|--------------------|
| <b>Rubro</b>                          | <b>Uso (sí/no)</b> | <b>Descripción</b> |
| <b>Asistencia en la redacción</b>     | No                 |                    |

# Formato de Declaración de Originalidad y Uso de Inteligencia Artificial

Coordinación General de Estudios de Posgrado  
Universidad Michoacana de San Nicolás de Hidalgo



| Uso de Inteligencia Artificial             |             |                              |
|--|-------------|------------------------------|
| Rubro                                      | Uso (sí/no) | Descripción                  |
| Traducción al español                      | No          |                              |
| Traducción a otra lengua                   | Si          | Uso del traductor de Google. |
| Revisión y corrección de estilo            | Si          | Uso de Grammarly y Deepseek  |
| Análisis de datos                          | No          |                              |
| Búsqueda y organización de información     | No          |                              |
| Formateo de las referencias bibliográficas | No          |                              |
| Generación de contenido multimedia         | No          |                              |
| Otro                                       | No          |                              |

| Datos del solicitante |   |
|-----------------------|---|
| Nombre y firma        | Maritza Talhia Bernabé Morales  |
| Lugar y fecha         | Morelia, Michoacán, a 6 de mayo de 2026.  |