



**Universidad
Europea**

UNIVERSIDAD EUROPEA DE MADRID

ESCUELA DE ARQUITECTURA, INGENIERÍA Y DISEÑO

MÁSTER UNIVERSITARIO EN ANÁLISIS DE DATOS MASIVOS (BIG DATA)

TRABAJO FIN DE MÁSTER

**Predicción del Precio de Bolsa Nacional day-ahead en el
mercado colombiano: comparación de
arquitecturas de aprendizaje profundo bajo
variabilidad climática**

ANDRÉS GRIMALDOS ECHAVARRÍA

Dirigido por

DR. WOLFRAM ROZAS RODRÍGUEZ

CURSO 2025-2026

TÍTULO: Predicción del Precio de Bolsa Nacional day-ahead en el mercado colombiano: comparación de arquitecturas de aprendizaje profundo bajo variabilidad climática

AUTOR: ANDRÉS GRIMALDOS ECHAVARRÍA

TITULACIÓN: MÁSTER UNIVERSITARIO EN ANÁLISIS DE DATOS MASIVOS (BIG DATA)

DIRECTOR/ES DEL PROYECTO: DR. WOLFRAM ROZAS RODRÍGUEZ

FECHA: abril de 2026

RESUMEN

Colombia es un país que se caracteriza por tener un como fuente dominante de energía el agua, lo que lo hace fuertemente sensible a cambios climáticos como el fenómeno de El Niño o La Niña, lo que agrega componentes de gran incertidumbre a la dinámica del precio de bolsa nacional (PBN) y provoca que diferentes inversionistas no tengan interés en asumir riesgos en este mercado y los agentes económicos tengan una mayor dificultad en la toma de decisiones. Este trabajo utiliza 6 arquitecturas de aprendizaje profundo que son comparadas de manera experimental por medio de métricas puntuales y probabilísticas con el propósito de crear estimaciones puntuales e intervalos de confianza de calidad para la toma de decisiones en este mercado a través de una mitigación del riesgo. Como resultado se obtuvo que la mejor arquitectura en todos los aspectos para un periodo de tiempo hasta enero de 2025 es el *Temporal Fusion Transformer*, lo que contribuye al campo de estimación de precios de energía en Colombia para que continúe siendo estudiado y mercados de energías renovables como Colombia tengan una barrera de entrada más sencilla para su inversión.

Palabras clave: Predicción precios de energía, Temporal Fusión Transformer, Aprendizaje Profundo, XM.

ABSTRACT

Colombia is a country with a high dependency of water in their energy matrix which cause a big sensibility to weather phenomena such as El Niño or La Niña, therefore the energy price in Colombia is characterized by uncertainty and investors do not see this market as an option because of their high risk and the economic stakeholders have difficulties in decision making. This work employees 6 deep learning architectures which are compared by means of point and probabilistic metrics with the purpose of making good points and confidence interval estimations to improve the decision-making by mitigating the risk. As a result, we got that the best model in both point and probabilistic estimation was Temporal Fusion Transformer in a period until January 2025, which contributed to the field of energy prices forecasting in Colombia in order to continue studying this market and making more easily available to invest and take decisions and improve engagement in renewable energies.

Keywords: Energy Price Forecasting, Temporal Fusion Transformer, Deep Learning, XM

AGRADECIMIENTOS

En primera instancia me gustaría agradecer a mi tutor Dr. Wolfram Rozas Rodríguez por estar siempre disponible cuando se le ha necesitado, estar pendiente del proceso de este trabajo y enseñarme sobre los temas relacionados a este TFM, son enseñanzas tanto académicas como personales para toda la vida.

También me gustaría darle las gracias a la Universidad Europea por brindarme educación de calidad durante este periodo de máster. Este aprendizaje me ha sido útil para el crecimiento profesional.

Por otro lado, me encantaría agradecerle a mi familia, porque es mi pilar de vida y son las personas más importantes para mí que siempre han estado dándome apoyo y animo en todo lo que hago.

Finalmente, agradecerle a cada uno de mis amigos a los que les he conversado sobre el máster y siempre han estado empujándome a continuar adelante y motivándome incluso en los momentos más complejos.

"The goal of forecasting is not to predict the future but to tell you what you need to know to take meaningful action in the present."

— Paul Saffo, *Six Rules for Effective Forecasting*, Harvard Business Review (2007)

TABLA RESUMEN

	DATOS
Nombre y apellidos:	Andrés Grimaldos Echavarría
Título del proyecto:	Predicción del Precio de Bolsa Nacional day-ahead en el mercado colombiano: comparación de arquitecturas de aprendizaje profundo bajo variabilidad climática
Directores del proyecto:	Dr. Wolfram Rozas Rodríguez
El proyecto se ha realizado en colaboración de una empresa o a petición de una empresa:	NO
El proyecto ha implementado un producto:	NO
El proyecto ha consistido en el desarrollo de una investigación o innovación:	SI
Objetivo general del proyecto:	Predecir el precio del mercado spot de energía colombiano para el día siguiente comparando arquitecturas de aprendizaje profundo

Índice

RESUMEN	3
ABSTRACT	4
TABLA RESUMEN	7
Capítulo 1. RESUMEN DEL PROYECTO	13
1.1 Contexto y justificación.....	13
1.2 Planteamiento del problema	13
1.3 Objetivos del proyecto.....	13
1.4 Resultados obtenidos	14
1.5 Estructura de la memoria	14
Capítulo 2. ANTECEDENTES / ESTADO DEL ARTE	15
2.1 Estado del arte	15
2.2 Contexto y justificación.....	18
2.3 Planteamiento del problema	23
Capítulo 3. OBJETIVOS.....	25
3.1 Objetivos generales	25
3.2 Objetivos específicos	25
3.3 Beneficios del proyecto	26
Capítulo 4. DESARROLLO DEL PROYECTO	27
4.1 Planificación del proyecto.....	27
4.2 Descripción de la solución, metodologías y herramientas empleadas.....	29
4.3 Recursos requeridos	67
4.4 Presupuesto	67
4.5 Viabilidad	68
4.6 Resultados del proyecto	69

Andrés Grimaldos Echavarría

Capítulo 5.	DISCUSIÓN.....	78
Capítulo 6.	CONCLUSIONES	80
6.1	Conclusiones del trabajo.....	80
6.2	Conclusiones personales.....	80
Capítulo 7.	FUTURAS LÍNEAS DE TRABAJO	82
Capítulo 8.	REFERENCIAS.....	83

Índice de Figuras

Figura 1. Generación eléctrica real anual del SIN en 2025. Participación porcentual de la generación por fuente: hidráulica (80.7%), combustible fósil (12.7%), solar (5.3%), biomasa (1.1%) y eólica (0.2%). Fuente: elaboración propia con datos de XM/Sinergox (XM, 2026).	19
Figura 2. Relación inversa entre aporte de caudal y precio de la energía. Fuente: elaboración propia con datos del equipo de analítica de XM (XM, 2026b).....	20
Figura 3. Comportamiento del cálculo mensual del ONI entre 2014 y 2025. Fuente: elaboración propia con datos de NOAA (NOAA, 2026).....	21
Figura 4. Precio spot promedio mensual y su relación con episodios de fenómeno de El Niño y La Niña. Fuente: elaboración propia con datos del equipo de analítica de XM y NOAA (NOAA, 2026; XM, 2026b).....	22
Figura 5. Metodología CRISP-DM. Fuente: elaboración propia.	27
Figura 6. Disponibilidad de los datos. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026).....	34
Figura 7. Distribuciones de las variables utilizando KDE*. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026). *KDE: Kernel Density Estimation es un método no paramétrico para estimar la densidad de una variable.	36
Figura 8. Diagrama ACF y PACF de PBN. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.).....	37
Figura 9. Series temporales de las variables de estudio. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026)	39
Figura 10. Descomposición STL* del PBN. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.). STL: Seasonal and Trend decomposition using Loess: es un método muy usado para dividir una serie temporal en tendencia, estacionalidad y ruido.....	40
Figura 11. Relación de valores extremos del PBN y su co-ocurrencia con el fenómeno de El Niño y La Niña. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026) .	41
Figura 12. Matrices de correlación de Pearson y Spearman y diferencia entre ambas. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026)	43
Figura 13. Correlaciones cruzadas con de PAC, AES, PVUD y ONI con PBN. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026)	44
Figura 14. VIF de las variables del trabajo. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026).....	45

Figura 15. Boxplot interanual del PBN. Elaboración propia con datos de (Equipo Analítica XM, s. f.).....	46
Figura 16. Flujo de preparación de los datos. Fuente: elaboración propia	47
Figura 17. Mecanismo RNN. Fuente: elaboración propia	53
Figura 18. Mecanismo LSTM	54
Figura 19. Arquitectura CNN-LSTM. Fuente: elaboración propia.	55
Figura 20. Arquitectura TCN. Fuente: elaboración propia	56
Figura 21. Arquitectura DeepAR. Fuente: elaboración propia basada en Salinas et al. (2020)..	58
Figura 22. Arquitectura TFT. Fuente: extraído de Lim et al. (2021).	59
Figura 23. Comparación de predicciones y valor real. Fuente: elaboración propia con datos de Equipo Analítica XM (s. f.)	70
Figura 24. Regresión de predicción y valor real. Fuente: elaboración propia	71
Figura 25. Distribución de los errores de estimación. Fuente: elaboración propia	71
Figura 26. Ratio MAE spike / MAE normal. Fuente: elaboración propia.	73
Figura 27. Comportamiento de atención. Fuente: librería pytorch-forecasting donde en el eje x tenemos el tiempo y en el eje y el valor numérico de atención	74
Figura 28. Importancia de las variables estáticas. Fuente: librería pytorch-forecasting donde el eje x muestra el porcentaje de atención y el eje y muestra las variables estáticas.	74
Figura 29. Importancia de las variables en la etapa encoder. Fuente: librería pytorch-forecasting donde el eje x muestra el porcentaje de importancia en la fase de encoder y el eje y las variables	75
Figura 30. Importancia de las variables en la fase decoder. Fuente: librería pytorch-forecasting donde el eje x muestra el porcentaje de importancia y el eje y las variables.	75

Índice de Tablas

Tabla 1. Resumen de la literatura destacada..... 17

Tabla 2. Cronograma de trabajo. Fuente: elaboración propia..... 29

Tabla 3. Métricas de API XM elegidas para los modelos. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.) 32

Tabla 4. Medidas de tendencia central y dispersión. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026) 35

Tabla 5. Medidas de posición y forma. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026)..... 35

Tabla 6. Resultados pruebas ADF y KPSS. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026)..... 38

Tabla 7. Resumen comparativo modelos del estudio. Fuente: elaboración propia. 60

Tabla 8. Resultado métricas de evaluación sobre el conjunto de pruebas. Fuente: elaboración propia. 69

Capítulo 1. RESUMEN DEL PROYECTO

1.1 Contexto y justificación

La matriz energética colombiana está compuesta en su mayoría por fuentes hídricas, lo cual convierte a Colombia en un país que le apuesta a la generación de energía por medio de fuentes renovables, pero fuertemente dependiente de un tipo de fuente, por lo que la afectación de esta puede afectar notablemente los precios de la electricidad y dificultar el ingreso de inversores o entidades en este mercado. Es por ello, que surge la necesidad de crear estrategias que fomenten estos clientes y así permitir seguir usando orígenes renovables para la energía, lo cual va en pro de los objetivos de desarrollo sostenible.

1.2 Planteamiento del problema

A causa de la fuerte dependencia de las afluencias de agua en el proceso de generación de energía en Colombia, fenómenos climáticos como El Niño o La Niña pueden traer comportamientos inesperados dentro de este mercado debido a sequías o fuertes lluvias respectivamente, por tanto, la relación entre situaciones meteorológicas y el precio de la energía en Colombia, trae la problemática de reducir la incertidumbre en los procesos de gestión de riesgos por parte de inversionistas y entidades. Es por ello, que han surgido múltiples investigaciones que han intentado atacar el problema utilizando diferentes aproximaciones para la predicción del precio de la electricidad en Colombia, pero debido a la complejidad de variables que pueden afectar la dinámica, se requiere de estrategias avanzadas como el *Deep learning* para ser probadas con múltiples fuentes de información que puedan ser útil para esta tarea.

1.3 Objetivos del proyecto

El proyecto está orientado a un objetivo principal el cual es la predicción de los precios de energía en el mercado spot colombiano de electricidad del día siguiente mediante un modelo predictivo basado en aprendizaje profundo o *deep learning* contrastado contra otros modelos estadísticos y de aprendizaje automático utilizando métricas suministrada por la entidad administradora del Mercado de Energía Mayorista de Colombia denominada XM por medio de API pública y la inclusión del Índice Oceánico del Niño para efectos de posibles periodos del

fenómeno de El Niño o La Niña. Esto con el propósito de estimular el mercado de energía para los diferentes agentes e incentivar el uso de energías renovables.

1.4 Resultados obtenidos

El análisis brindado por este trabajo dio como resultado que el mejor modelo para la predicción del precio de energía en el day-ahead en el mercado colombiano es el *Temporal Fusion Transformer*, el cual tuvo el mejor rendimiento en métricas puntuales y probabilísticas, lo que hace que los agentes del Mercado de Energía Mayorista de Colombia tengan la posibilidad de tener valores confiables que le permitan una mejor toma de decisiones mitigando su riesgo.

Adicionalmente, se obtuvo que hay una estrecha relación entre las variables de los aportes de caudal, los aportes de energía, el porcentaje de volumen útil y el fenómeno de El Niño en la predicción del precio de la energía que fue detectada por medio de análisis exploratorio de los datos y por medio del mecanismo de atención del modelo.

1.5 Estructura de la memoria

Esta memoria tiene una estructura que comienza por un primer capítulo donde se menciona de manera resumida el contexto, el planteamiento del problema, los objetivos, los resultados y la estructura del documento. Posteriormente, se tiene un capítulo orientado a presentar el estado actual de la literatura asociada al campo de predicción de precios de energía, dar un contexto y un planteamiento del problema con mayor profundidad. Luego, se tiene el capítulo en donde se plantean los objetivos generales y específicos de este TFM que nos sirven de guía para lo que se quiere lograr. En el capítulo 4 se expone la planificación de este trabajo y cómo se aplica la metodología CRISP-DM a este trabajo y los resultados obtenidos. Por otro lado, el capítulo 5 corresponde a establecer un análisis y discusión alrededor de los resultados que se obtuvieron. En el capítulo 6 se presentan las conclusiones de este trabajo durante su implementación. Finalmente, en el capítulo 7 se tienen las futuras líneas de trabajo que surgen de este proyecto y en el capítulo 8 tenemos las referencias utilizadas para este TFM.

Capítulo 2. ANTECEDENTES / ESTADO DEL ARTE

En este capítulo se expone la literatura previa consultada asociada a métodos utilizados para la predicción del precio de la energía en Colombia y en otros países para tener un panorama bibliográfico. Posteriormente, se procede a dar contextualización del Mercado de Energía Mayorista de Colombia junto con su principal problemática de incertidumbre y la importancia de reducirla. Finalmente, se menciona el objetivo con este trabajo asociado a este problema en el contexto colombiano.

2.1 Estado del arte

La predicción del precio de energía en el mercado spot del Mercado de Energía Mayorista es un problema complejo debido a la incertidumbre de la generación de energía en Colombia a causa de la fuerte dependencia de las fuentes hídricas las cuales pueden ser sensibles a fenómenos climáticos cuya naturaleza inherente es estocástica. Es por ello, que en primer lugar se revisa literatura asociada a la comprobación de influencia climática en el precio de la energía. Luego, se pasa a mostrar investigaciones de aplicación de métodos estadísticos. Posteriormente, se estudia literatura que indica estrategias que se han implementado para identificar las variables más influyentes para modelos de predicción de aprendizaje automático (*machine learning*) y de aprendizaje profundo (*deep learning*) aplicado al mercado de energía en Colombia y casos internacionales para esta última aproximación.

La **¡Error! No se encuentra el origen de la referencia.** ilustra un resumen de la literatura que se consultó en el orden expuesto anteriormente, donde se busca ilustrar la referencia, su objetivo, el dataset que se utilizó y el método usado para lograr el objetivo que tenían propuesto.

Referencia	Objetivo	Dataset	Método
(Villa-Loaiza et al., 2023)	Analizar si el fenómeno de El Niño afecta la relación entre los precios de la energía y los mercados financieros en el contexto colombiano	Para el rango de 1 de julio a 31 de diciembre de 2021: Índice Oceánico El Niño proporcionado por el <i>Climate Prediction</i>	Prueba de contagio Bootstrap de correlación Gaussiana local

		<i>Centre de National Oceanic and Atmospheric Administration.</i> • Precio de Bolsa Nacional de XM. • Índice General de la Bolsa de Valores de Colombia.	
(Monica et al., 2018)	Identificación de anomalías financieras como “efecto calendario”, análisis de volatilidad y evaluar influencia climática a través del modelamiento y pronóstico del precio de energía en el mercado spot.	• Precios de energía del mercado spot tanto diarios como mensuales. • Aportes hídricos.	• ARMA • SARIMAX • GARCH • IGARCH
(Villarreal et al., 2023)	Predecir el precio de la energía en el mercado colombiano con el propósito de que inversionistas puedan calcular ingresos en los modelos de valoración de proyectos de generación de energía renovable en un horizonte de 3 años	Para el rango de 1 de enero de 2000 a 31 de julio de 2023: • Precio de energía en bolsa nacional, aportes hídricos y en caudal de XM. • Índice Oceánico El Niño proporcionado por el <i>Climate Prediction Centre</i> de <i>National Oceanic and Atmospheric Administration</i>.	• Regresión lineal simple y múltiple. • VAR. • ARIMA • ARIMAX • SARIMAX
(Ramírez Londoño & Giraldo Zuluaga, 2023)	Permitir a los agentes de mercado ajustar ofertas para mitigar riesgos y maximizar	Para el rango de 2013 a 2018: • Precio de bolsa, demanda de	• DNN • LSTM • GRU • CNN-LSTM

Andrés Grimaldos Echavarría

		beneficios por medio de la implementación de diferentes modelos de <i>Deep learning</i> .	energía, disponibilidad de plantas térmicas y no térmicas, volumen útil de embalses, aportes hidrológicos y precio de oferta de plantas de API pública de XM.	
(Rozas 2024)	Rodríguez,	Mejorar el rendimiento de los Mercados Locales de Energía (LEM) mediante una solución digital integral denominada Smart Local Energy Market (SLEM).	<i>Cornwall Local Energy Market (LEM) Residential Electricity Dataset</i> (2018-2020), facilitado por <i>Trilemma Consulting</i> y <i>Centrica PLC</i> .	• K-Means. • Two-Step. • Redes de Kohonen. • SARIMAX. • ESA. • TCM. • Regresión lineal (OLS, Ridge, Lasso). • Árboles de decisión (CART, CHAID, C5.0). • Random forest. • XGB. • MLP. • CNN + RNN (LSTM, GRU, Bi-LSTM y Bi-GRU).
(Failing et al., 2026)		Explorar y evaluar diferentes arquitecturas de <i>Deep learning</i> para predecir los precios del mercado eléctrico spot español en un horizonte de 1 día.	Para el rango de 1 de enero de 2020 a 29 de junio de 2023: • API del sistema de información del operador del sistema español (eSIOS). • Precios del gas de Mibgas.	• ANN. • CNN. • LSTM.

Tabla 1. Resumen de la literatura destacada.

2.2 Contexto y justificación

El sector eléctrico colombiano se organiza bajo un enfoque de mercado regulado en el que conviven (i) la operación física del sistema eléctrico (que debe mantener el equilibrio instantáneo entre oferta y demanda) y (ii) un conjunto de transacciones comerciales que determinan precios, pagos e incentivos para inversión y prestación del servicio. Este arreglo se enmarca en el régimen definido por la Ley 143 de 1994 que establece las actividades del sector: generación, interconexión, transmisión, distribución y comercialización (*Ley 143 de 1994 - Gestor Normativo - Función Pública*, s. f.).

Dentro de este marco se encuentra el Mercado de Energía Mayorista (MEM) en el cual participan generadores, transmisores, distribuidores, comercializadores y consumidores no regulados y está administrado por XM compañía de expertos en mercados (XM) y es regulado por la Comisión de Regulación de Energía y Gas (CREG) (*Ley 143 de 1994 - Gestor Normativo - Función Pública*, s. f.; XM, s. f.-a). Adicionalmente, este está segmentado en 2:

- Mercado de contratos bilaterales: en este se realizan operaciones de compra y venta entre generadores y comercializadores principalmente que les permite gestionar sus riesgos con el propósito de generar cobertura y estabilidad en sus ingresos y costos (XM, s. f.-b).
- Bolsa de energía (mercado spot): es aquel donde se transa energía en el corto plazo (con precios para las 24 horas del día siguiente) con participación obligatoria de todos los generadores registrados de precios de oferta y sin intervención financiera de los contratos bilaterales (XM, s. f.-b).

En la dinámica del mercado energético colombiano XM tiene un papel importante, porque no solo es el administrador del MEM, sino también el que se encarga de operar el Sistema Interconectado Nacional colombiano (SIN) a través del Centro Nacional de Despacho (CND) el cual se encarga de la planificación y supervisión de la operatividad de los recursos del SIN para que cubran la demanda diaria a partir de la disponibilidad ofertada por los generadores (XM, s. f.-a, s. f.-b). Por consiguiente, XM desde el CND tiene la responsabilidad de fijar el precio de la energía horaria en el mercado spot de acuerdo con la demanda proyectada de la manera más segura, económica y confiable posible e independiente de los contratos bilaterales.

Para la formación del precio en bolsa, XM exige a los generadores presentar ofertas de precio para las siguientes 24 horas del día de acuerdo con sus capacidades técnicas y de generación

durante el día y con esa información, el tipo de planta (e.g. hidroeléctrica o térmica) y aspectos técnicos de la misma, el CND ejecuta un algoritmo de optimización que da como resultado el despacho ideal. Posteriormente, según CREG (2023) se ordena de menor a mayor precio ofertado las plantas del despacho y se halla el máximo precio ofertado (MPO) como el asociado a la última planta requerida para atender la demanda y finalmente se le adiciona un valor ΔI que corresponde al costo de las plantas térmicas que salen en el despacho ideal. Por tanto, se observa que el proceso del cálculo del precio de bolsa nacional de la energía depende de múltiples factores diarios que pueden desencadenar que una planta entre en el despacho ideal o no, por ejemplo, circunstancias climáticas.

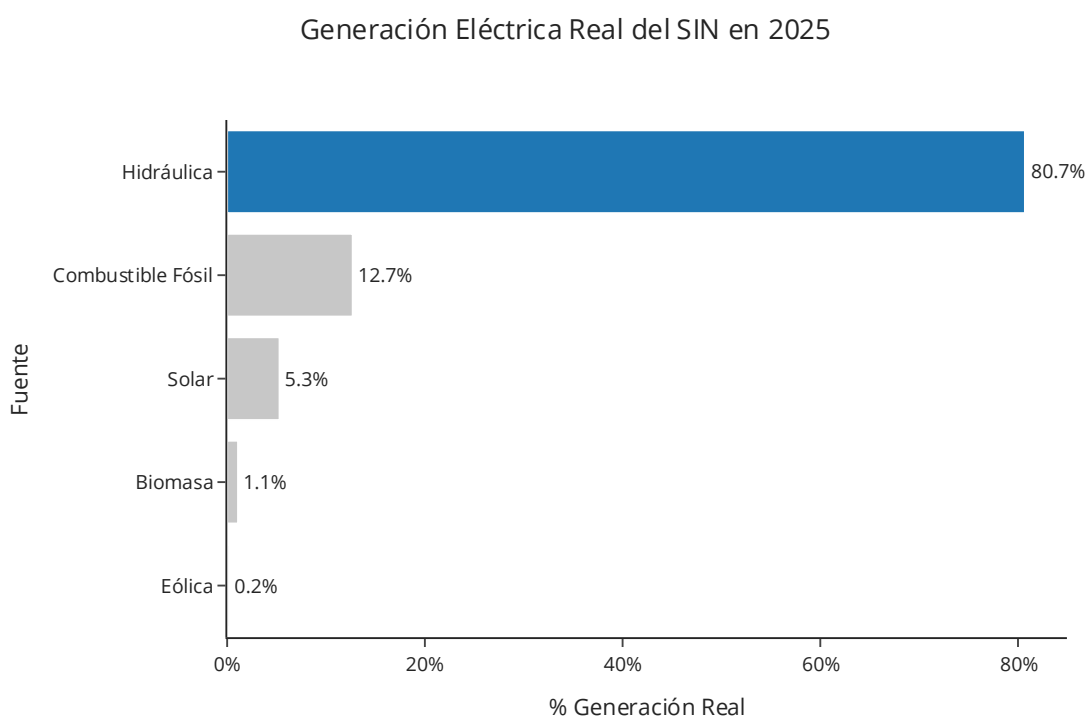


Figura 1. Generación eléctrica real anual del SIN en 2025. Participación porcentual de la generación por fuente: hidráulica (80.7%), combustible fósil (12.7%), solar (5.3%), biomasa (1.1%) y eólica (0.2%). Fuente: elaboración propia con datos de XM/Sinergox (XM, 2026).

En sintonía con lo anteriormente mencionado, en la Figura 1 se observa la composición de la matriz eléctrica real anual del SIN en 2025 en la cual hay una alta dominancia hidráulica lo que confirma la fuerte dependencia del sistema eléctrico colombiano de las afluentes de agua. En segundo lugar, aparece la generación con combustibles fósiles que suele actuar como respaldo y aporta flexibilidad para atender condiciones de mayor estrés del sistema. La participación de la energía solar evidencia el avance reciente de fuentes renovables no convencionales, mientras que la biomasa y la eólica mantienen una contribución marginal en el total anual, lo que sugiere

un potencial de crecimiento futuro y/o limitaciones de expansión, integración y disponibilidad de recursos.

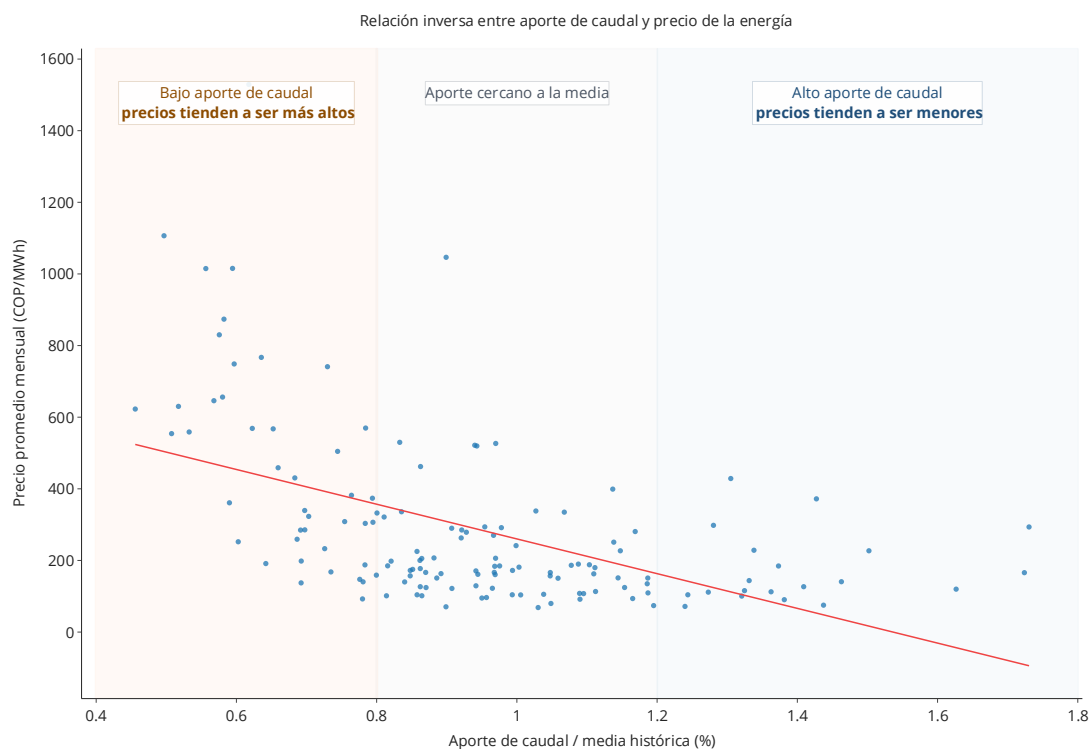


Figura 2. Relación inversa entre aporte de caudal y precio de la energía. Fuente: elaboración propia con datos del equipo de analítica de XM (XM, 2026b).

Adicionalmente, en la Figura 2 se observa la relación entre el precio promedio mensual de la energía y el aporte de caudal con respecto a la media histórica, donde se vuelve a evidenciar la fuerte dependencia que tiene Colombia con las afluencias hídricas, puesto que a menor sea el aporte del caudal, el precio es mayor y viceversa, debido a que una reducción de las fuentes de agua, provoca que las plantas hidroeléctricas protejan sus recursos y suban sus precios debido a la escasez de esta materia prima y, por consiguiente los precios del Despacho Ideal sean más altos y haya más probabilidad de entrada a fuentes de combustibles fósiles.

El Fenómeno de El Niño y La Niña son las dos fases opuestas del ENSO (El Niño–Southern Oscillation), un patrón climático natural asociado a cambios en la temperatura superficial del mar y en la circulación atmosférica del océano Pacífico tropical. El Niño ocurre cuando las aguas del Pacífico ecuatorial central y oriental se calientan por encima de lo normal, lo que altera los vientos y la distribución de lluvias a escala global; en Colombia suele asociarse, en términos generales, con reducción de precipitaciones, aumento de temperaturas y mayor probabilidad

de sequías e incendios en varias regiones. En contraste, La Niña se presenta cuando esas aguas se enfrían más de lo normal, favoreciendo condiciones opuestas: incremento de la nubosidad y la precipitación, con mayor riesgo de inundaciones y deslizamientos en zonas vulnerables.

Ahora bien, dado que el Despacho Ideal lo más común es esperarse que la mayoría de las fuentes generadoras del día sean hidráulicas, se muestra una de las características del MEM en Colombia: la incertidumbre, la cual está estrechamente relacionada con la aleatoriedad climática que caracteriza a Colombia y que puede llegar a ser fuertemente afectada por fenómenos como El Niño o La Niña que influyen en la disponibilidad del recurso hídrico en embalses y caudales, provocando fluctuaciones complejas de predecir en los precios de la energía. Este aspecto es causante de que las decisiones de inversión o financiación en este sector tiendan a no ser tan atractivas debido a que se está exponiendo a un riesgo que puede llegar a ser difícil de controlar.

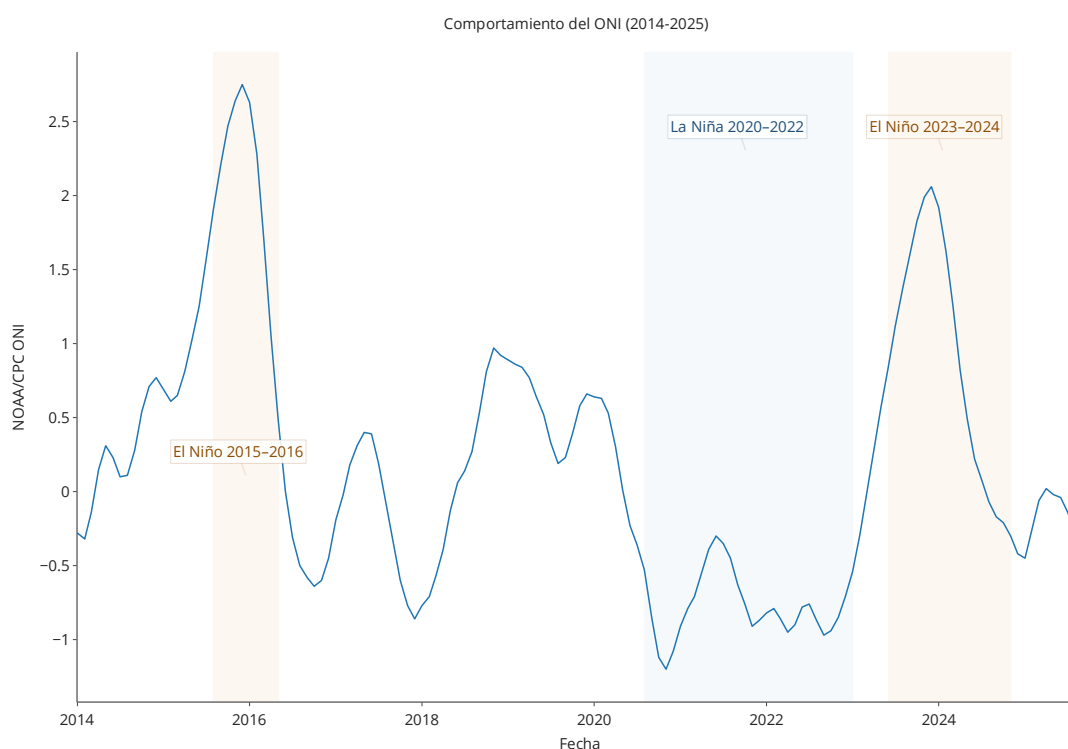


Figura 3. Comportamiento del cálculo mensual del ONI entre 2014 y 2025. Fuente: elaboración propia con datos de NOAA (NOAA, 2026).

Reforzando lo anterior, la institución *National Oceanic and Atmospheric Administration* (NOAA) que está encargada del estudio y predicción de cambios en el océano, atmósfera, espacio y el sol reporta el denominado *Oceanic Niño Index* (ONI) que es un indicador que representa la

diferencia entre la temperatura promedio móvil de los últimos 3 meses de la superficie del mar Pacífico en la región Niño 3.4 y el promedio de los últimos 30 años sirviendo como medio para monitorear el fenómeno de El Niño y La Niña presentados cuando el índice presenta un valor de encima de 0.5 y por debajo -0.5 respectivamente (NOAA, 2009). En la Figura 3 se observa el comportamiento de este indicador entre 2014 y 2025 señalando 3 momentos de ejemplo en los que valores consecutivos superiores a 0.5 o inferiores a -0.5 se presentaron, indicando presencia del fenómeno de El Niño y La Niña respectivamente.

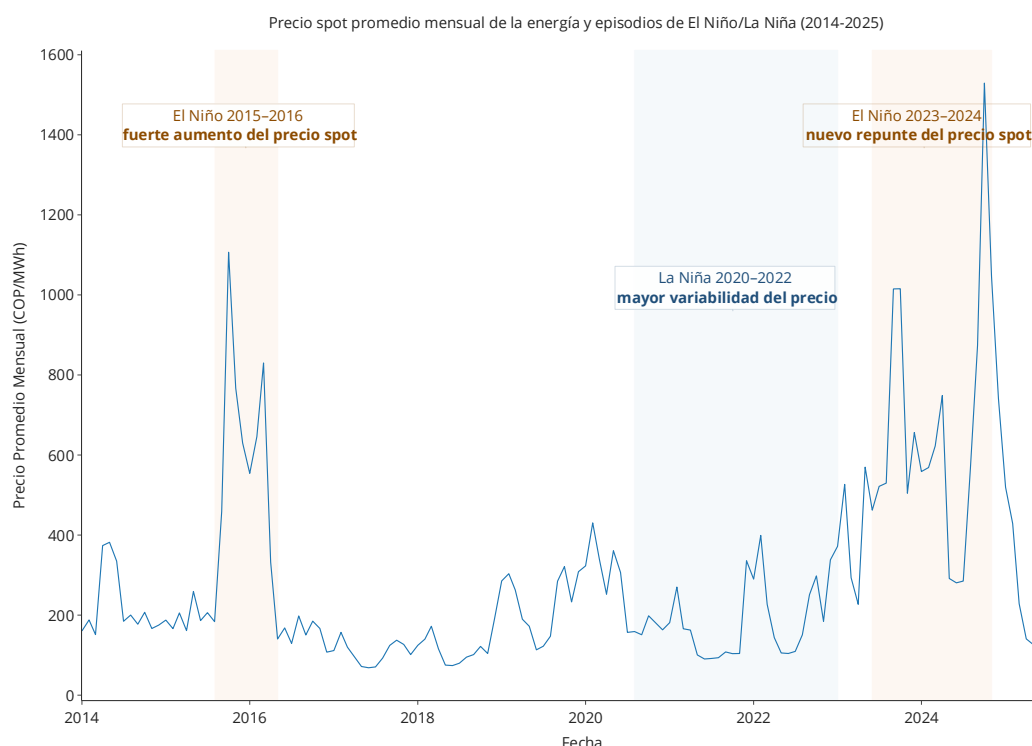


Figura 4. Precio spot promedio mensual y su relación con episodios de fenómeno de El Niño y La Niña. Fuente: elaboración propia con datos del equipo de analítica de XM y NOAA (NOAA, 2026; XM, 2026b).

En la Figura 4 se presenta el comportamiento del precio spot promedio mensual de la energía en Colombia entre 2014 y 2025 señalando episodios del fenómeno de El Niño y La Niña, mostrando que para el primero se tiene la tendencia a aumentar el valor de los precios, puesto que en esta fase suele haber una reducción de los aportes de los caudales, y como se observó anteriormente, esto provoca un aumento en el precio de la energía en Colombia. De manera análoga, ocurre cuando está en la fase de La Niña, donde el aumento de las lluvias puede dar pie a una volatilidad mayor del precio, pero en su mayoría manteniendo periodos donde baja el precio de la energía a causa del aumento de los aportes hídricos cuando pueden ser aprovechados.

A partir de esto, surge la necesidad de buscar diferentes formas de incentivar a que inversionistas y diferentes entidades se interesen por este sector sin la necesidad de eliminar de las fuentes principales las energías renovables. Es por ello, que técnicas avanzadas como el aprendizaje profundo que tienen el factor diferencial de encontrar patrones que otros métodos estadísticos convencionales o métodos de aprendizaje automático no podrían identificar, lo cual permite una reducción de la incertidumbre y, por tanto, el riesgo de ingreso para los inversionistas puede ser gestionado de una manera más sencilla y tomar decisiones de contratación u operación mejor. Este detalle puede facilitar la expansión de energías renovables en Colombia a futuro con consecuencias como el ingreso de la energía solar con mayor contribución que podría compensar en situaciones como el fenómeno de El Niño por sus sequías.

2.3 Planteamiento del problema

El mercado spot de energía en Colombia cumple un papel central en la asignación de señales de corto plazo para la operación y comercialización del Sistema Interconectado Nacional; sin embargo, su dinámica de precios presenta alta volatilidad y episodios de comportamiento extremo que complican la toma de decisiones de generadores, comercializadores, grandes consumidores e inversionistas. Una causa estructural de esta incertidumbre es la fuerte dependencia de la matriz de generación respecto al recurso hídrico, lo que vuelve al precio especialmente sensible a variaciones hidrológicas y a choques climáticos asociados a los fenómenos de El Niño y La Niña, además de condiciones de oferta, demanda y restricciones operativas. En este contexto, los agentes deben formular estrategias de contratación, cobertura e inversión bajo un entorno donde el riesgo climático se traduce en riesgo financiero, y donde la complejidad de variables y su interacción no lineal incrementa los costos de análisis, elevando barreras de entrada para participantes no especializados.

A partir del estado del arte, se evidencia que, aunque existen múltiples enfoques para el pronóstico de precios eléctricos, persisten vacíos relevantes cuando se busca modelar mercados con alta sensibilidad climática y no estacionariedad, como el colombiano:

- Limitaciones para capturar relaciones no lineales y dependencias temporales complejas entre variables hidrometeorológicas y el precio spot.
- Desempeño reducido ante escenarios extremos o cambios de régimen (e.g., eventos Niño/Niña)
- Ausencia de marcos predictivos que traduzcan el pronóstico en reducción de incertidumbre útil para la toma de decisiones (e.g., mediante escenarios o distribución

de probabilidad del precio) con potencial impacto en la financiación y estructuración de proyectos.

Esta falta de una solución predictiva robusta y adaptada al contexto local restringe la capacidad de gestionar el riesgo de precio, lo que puede desincentivar la entrada de nuevos inversionistas y limitar la expansión de tecnologías renovables que requieren mayor certidumbre para su bancabilidad. Ante estas limitaciones, aproximaciones más modernas como el *deep learning* surgen como una opción viable para reducir la incertidumbre inherente al proceso de los precios debido a su gran nivel de rendimiento en la captura de patrones.

Es por lo anterior que observamos en la literatura más reciente asociada a la resolución de este problema que los métodos utilizados son de *deep learning* tratando de integrar variables que permitan tener en cuenta los aspectos que caracterizan esta dinámica compleja como se vio en el estado del arte. No obstante, dentro de los intentos de usar *Deep learning* se evidencia que aún hay posibilidades de utilizar otro tipo de variables como el Índice Oceánico El Niño proporcionado por el *Climate Prediction Centre* de *National Oceanic and Atmospheric Administration* que permite tener una alerta directa al modelo en caso de posible acontecimiento del fenómeno de El Niño o La Niña.

Por tanto, surge la necesidad de desarrollar y evaluar un modelo de predicción del precio spot colombiano basado en métodos de *deep learning* integrando variables utilizadas en otras aproximaciones y el Índice Oceánico El Niño, orientado a mejorar la anticipación de variaciones inducidas por fenómenos climáticos y a proveer insumos cuantitativos que faciliten decisiones de contratación e inversión, contribuyendo así a un entorno más favorable para atraer capital y acelerar la transición hacia una matriz más diversificada y renovable.

Capítulo 3. OBJETIVOS

En el presente capítulo se mencionan el objetivo general que se quiere lograr con este trabajo y sus respectivos objetivos específicos que nos permiten tener los pasos concretos y verificables de la metodología CRISP-DM que guían a alcanzar esta meta final.

3.1 Objetivos generales

Se establece un objetivo general para este trabajo el cual es:

OG.1. Desarrollar y evaluar modelos de deep learning para el pronóstico day-ahead del precio spot de la electricidad en Colombia, incorporando variables hidrológicas (aportes y volumen útil de embalses) e indicadores climáticos (Índice Oceánico de El Niño), con el fin de reducir la incertidumbre asociada a la variabilidad climática y generar insumos cuantitativos para la toma de decisiones de contratación e inversión en el mercado eléctrico, favoreciendo la participación de nuevos inversionistas y la integración de energías renovables.

3.2 Objetivos específicos

La implementación del siguiente trabajo se realiza bajo la metodología *Cross-Industry Standard Process for Data Mining* (CRISP-DM), por tanto, los objetivos específicos están orientados hacia el cumplimiento de las fases de este marco de trabajo.

OE1. Comprensión del negocio (Business Understanding). Delimitar el problema de pronóstico day-ahead del precio spot en el SIN y establecer los criterios de éxito del modelo, definiendo el caso de uso (toma de decisiones de contratación e inversión) y las métricas de evaluación que permitan cuantificar la reducción de incertidumbre bajo variabilidad climática.

OE2. Comprensión de los datos (Data Understanding). Identificar, recopilar y caracterizar las fuentes de datos necesarias para el pronóstico (API XM e índice Oceánico El Niño dado por NOAA), incluyendo series históricas de precio spot, variables hidrológicas (aportes y volumen útil de embalses) e indicadores climáticos (e.g., Índice Oceánico de El Niño), evaluando su calidad, consistencia temporal, disponibilidad y posibles rezagos.

OE3. Preparación de los datos (Data Preparation). Construir el conjunto de datos modelable mediante limpieza, alineación temporal, imputación de faltantes y transformación de variables, incorporando ingeniería de características (rezagos, ventanas móviles, estacionalidad, variables calendario) y definiendo esquemas de partición y validación temporal adecuados para series de tiempo.

OE4. Modelado (Modeling). Diseñar e implementar modelos de deep learning para el pronóstico day-ahead del precio spot (e.g., arquitecturas recurrentes, convolucionales o basadas en atención), ajustando hiperparámetros y comparando variantes con y sin variables climáticas/hidrológicas para medir su contribución al desempeño predictivo.

OE5. Evaluación (Evaluation). Evaluar el desempeño del modelo frente a líneas base (e.g., modelos naive/estadísticos o ML tradicionales) usando métricas de error y, cuando aplique, métricas orientadas a incertidumbre/riesgo (intervalos, cuantiles, cobertura), analizando robustez en periodos asociados a El Niño/La Niña y escenarios de alta volatilidad.

OE6. Despliegue/uso (Deployment). Proponer un esquema reproducible de uso del modelo (pipeline) para generar pronósticos day-ahead y reportes interpretables (escenarios, bandas de confianza y sensibilidad de variables), de modo que sirva como insumo para decisiones de contratación e inversión y para discutir su potencial contribución a la integración de energías renovables.

3.3 Beneficios del proyecto

Este proyecto aporta el beneficio de reducir la incertidumbre asociada a la variabilidad climática en el precio spot day-ahead del mercado eléctrico colombiano mediante el desarrollo y evaluación de un modelo de *deep learning* que integra variables hidrológicas (aportes y volumen útil de embalses) e indicadores climáticos (e.g., el Índice Oceánico de El Niño).

En relación con los objetivos, el modelo busca mejorar la capacidad de anticipación de cambios de precio y entregar insumos cuantitativos (pronósticos) para apoyar decisiones de contratación y gestión de riesgo de los agentes del MEM. Como resultado, se espera disminuir costos de análisis y barreras técnicas para participantes no especializados, facilitando un entorno más propicio para la entrada de nuevos inversionistas y para la estructuración de proyectos y contratos que incrementen la participación de energías renovables en la matriz, al contar con una evaluación más informada del riesgo de mercado.

Capítulo 4. DESARROLLO DEL PROYECTO

4.1 Planificación del proyecto

La metodología de este trabajo es de tipo experimental y comparativa, ya que pretende evaluar el rendimiento de múltiples modelos de *deep learning* con un conjunto de variables predictoras que den el contexto necesario para la predicción del precio de las 24 horas siguientes de la energía en el mercado spot colombiano con la intención de identificar el mejor modelo que podría favorecer la gestión del riesgo en las decisiones comerciales de agentes económicos implicados.

Como se mencionó en la sección de objetivos, el marco de trabajo utilizado será la metodología CRISP-DM que constituye un marco metodológico de referencia para el desarrollo de proyectos de analítica de datos, minería de datos y aprendizaje automático. Su principal fortaleza radica en que articula, de manera sistemática e iterativa, la comprensión del problema, el análisis de la información disponible, la construcción de modelos predictivos, la evaluación de su desempeño y la formulación de lineamientos para su eventual implementación. Esta metodología está ilustrada en la Figura 5 y consiste en las siguientes fases iterativas:

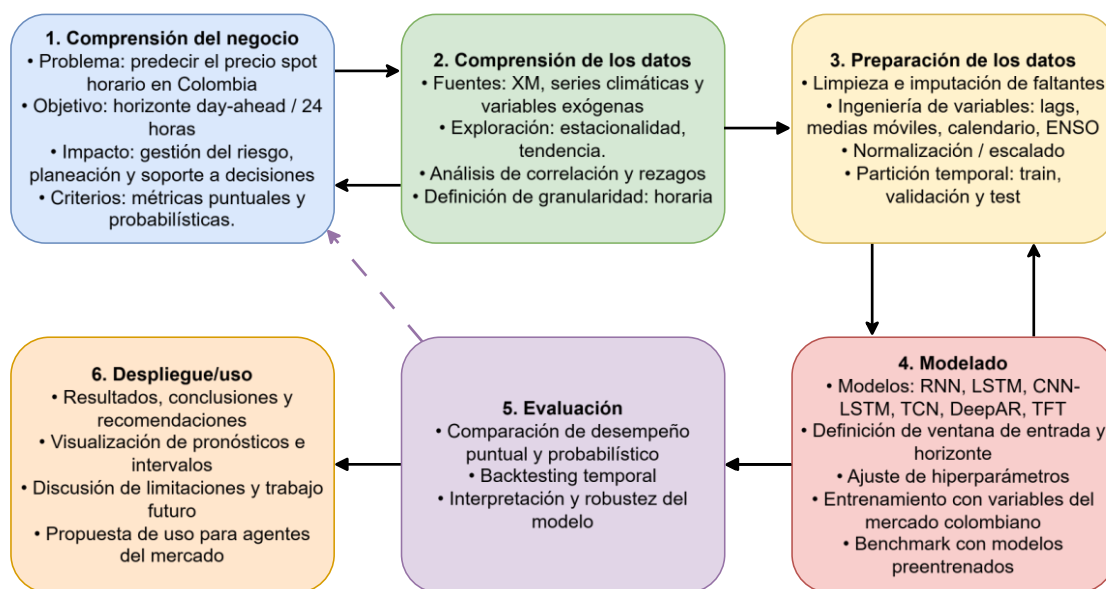


Figura 5. Metodología CRISP-DM. Fuente: elaboración propia.

- **Comprensión del negocio (Business Understanding):** en esta fase se busca tener un entendimiento profundo de la necesidad de negocio para establecer un alcance y resultados esperados al final del trabajo.
- **Comprensión de los datos (Data Understanding):** para esta fase se pretende realizar la recolección de datos de las fuentes disponibles y se procede a realizar un análisis exploratorio de los datos.
- **Preparación de los datos (Data Preparation):** con respecto a esta fase, se busca implementar un procesamiento de los datos para limpiarlos y transformarlos a una estructura adecuada para la siguiente fase.
- **Modelado (Modeling):** el objetivo de esta fase es definir cada uno de los modelos que se comparan entre sí para encontrar la mejor solución al problema a través de su entrenamiento y continuo ajuste de hiperparámetros.
- **Evaluación (Evaluation):** el propósito de esta fase es medir el rendimiento de los modelos planteados en la fase anterior para sopesar cuál de las soluciones anteriores es la mejor.
- **Despliegue/uso (Deployment):** en esta etapa final se busca establecer un marco de trabajo para la implementación de la solución propuesta en un entorno real, es decir, se explica cómo podría utilizarse en la práctica.

Así, esta metodología nos permite tener un enfoque mixto: cualitativo para poder comprender el contexto del mercado spot de energía colombiano, entender sus retos y complejidades, así como sus necesidades y la interacción de múltiples factores que marcan la dinámica; cuantitativo para tener un criterio de decisión al momento de inclinarse por una solución u otra que esté alineada a la gestión del riesgo en este sistema. Adicionalmente, nos permite tener un flujo iterativo que a través de retroalimentación nos permite llegar a una solución confiable.

Teniendo en cuenta lo anterior, el cronograma de tareas segregado por cada una de las fases de la metodología CRISP-DM se presenta en Tabla 2.

Fase	Tarea	Mes 1	Mes 2	Mes 3	Mes 4	Mes 5	Mes 6	Mes 7
Comprensión del negocio	Entendimiento del mercado de energía colombiano	X	X					
	Lectura de literatura para estado del arte		X	X				
	Búsqueda de fuentes de datos			X				

Andrés Grimaldos Echavarría

	Definición de alcance y objetivos	X			
	Redacción secciones 1, 2 y 3 de la memoria	X	X		
Comprensión de los datos	Análisis exploratorio de los datos		X		
Preparación de los datos	Definir estructura del dataset y su procesamiento		X		
Modelado	Definición de modelos y métricas de rendimiento		X		
	Redacción de la sección 4 de la memoria		X	X	
Evaluación	Implementación y resultados			X	X
	Análisis de los resultados y conclusiones				X
Despliegue/uso	Redacción secciones 5, 6 y 7 de la memoria			X	X

Tabla 2. Cronograma de trabajo. Fuente: elaboración propia.

4.2 Descripción de la solución, metodologías y herramientas empleadas

En esta subsección se expone, con un mayor nivel de profundidad analítica, la manera en que la metodología CRISP-DM será implementada en el marco del presente trabajo, cuyo propósito es la predicción day-ahead de los precios de la energía en el mercado colombiano. En particular, se desarrolla una caracterización sistemática de cada una de sus fases, precisando los fundamentos conceptuales que las orientan, las decisiones de diseño que estructuran su aplicación, los procedimientos metodológicos previstos y los entregables asociados, en coherencia con los objetivos, el alcance y las exigencias técnicas de esta investigación.

4.2.1 Comprensión del negocio

La primera fase de CRISP-DM consiste en delimitar el problema desde el contexto de negocio. En este trabajo, dicho problema se centra en la necesidad de anticipar el precio de la energía para el día siguiente en el mercado spot colombiano para una gestión del riesgo y toma de decisiones mejor e incentivar inversionistas a entrar a un mercado dominado por una fuente de energía renovable, caracterizado por una alta sensibilidad a variables operativas, hidrológicas, climáticas y de demanda. La relevancia de esta predicción se explica por su potencial utilidad para agentes generadores, comercializadores, grandes consumidores, analistas del sector e inversionistas, en la medida en que el precio spot resume, en tiempo real, el equilibrio dinámico entre oferta, demanda, disponibilidad de generación, restricciones del sistema y condiciones hidrometeorológicas.

Desde la perspectiva metodológica, esta fase exige responder preguntas tales como: ¿cuál es la relevancia económica y técnica de predecir el precio day-ahead en Colombia? ¿Qué implicaciones tiene la volatilidad del precio sobre la toma de decisiones de los agentes? ¿Cuál es el horizonte exacto de predicción que se pretende modelar —por ejemplo, una hora adelante, las 24 horas del día siguiente o el perfil horario completo—? Y ¿cuál será la contribución específica de este trabajo? Ya sea en términos de comparación de arquitecturas de aprendizaje profundo, mejora del desempeño predictivo, análisis de variables explicativas o incorporación de componentes probabilísticos en el pronóstico.

Adicionalmente, se formulan igualmente los objetivos generales y específicos del trabajo presentados en la sección anterior de objetivos, de donde se deduce que el horizonte de tiempo de predicción será las 24 horas siguientes, es decir el day-ahead. El resultado principal de esta fase es el entendimiento de la necesidad de la predicción del precio de la energía enmarcado en el contexto colombiano.

4.2.2 Comprensión de los datos

La segunda fase se orienta a identificar, recopilar y caracterizar las fuentes de información necesarias para este trabajo. En el caso del mercado de energía colombiano, implica reconocer qué series históricas se encuentran disponibles, cuál es su frecuencia temporal, qué nivel de calidad presentan y en qué medida pueden contribuir a explicar la dinámica del precio spot.

A partir de lo anterior se recopilaron 2 fuentes principales de datos útiles para el problema de este trabajo: API XM y serie histórica de ONI tomand como horizonte desde 2014-01-01 hasta 2025-01-01 de información.

API XM

Esta fuente de información consiste en un repositorio construido por el equipo de analítica de XM en el cual se implementa con el propósito de que las personas puedan construir servicios que consuman la información del Mercado de Energía Mayorista colombiano, ya sea por medio de objetos en Python, Excel con VBA o directamente el consumo por API (Equipo Analítica XM, s. f.).

En esta API es posible consultar listas o métricas de acuerdo con su frecuencia (horaria, diaria o mensual) con los siguientes *endpoints* (no requieren de una API key para su consumo):

- <https://servapibi.xm.com.co/hourly>
- <https://servapibi.xm.com.co/daily>
- <https://servapibi.xm.com.co/monthly>
- <https://servapibi.xm.com.co/lists>

En primera instancia, se consultó el *endpoint* correspondiente al listado de métricas, ya que este retorna las métricas disponibles asociadas al SIN, junto con su respectivo significado y frecuencia. El cuerpo de la solicitud empleado en esta petición se presenta a continuación:

```
{
  "MetricId": "ListadoMetricas"
}
```

A partir de la lista de métricas obtenida, las variables con su respectiva frecuencia se presentan en la Tabla 3.

Métrica	Frecuencia	Id de la métrica	Descripción
Precio de bolsa nacional histórico (PBN)	Horaria	PrecBolsNaci	Precio de bolsa horario nacional

Porcentaje de aportes de caudal del SIN (PAC)	Diaria	PorcApor	Aportes % por sistema, medido como [Aportes Caudal / Media Histórica]*100
Aportes de energía en el sistema (AES)	Diaria	AporEner	Caudales en energía de los ríos que aportan agua a algún embalse del SIN.
Porcentaje de volumen útil diario del sistema (PVUD)	Diaria	PorcVoluUtilDiar	Porcentaje de volumen almacenado en el embalse por encima del Nivel Mínimo Técnico.

Tabla 3. Métricas de API XM elegidas para los modelos. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.)

Para la consulta de las métricas presentadas en la Tabla 3 se utiliza el siguiente cuerpo de la solicitud, la cual debe realizarse múltiples veces para cada métrica puesto que el tamaño máximo de ventana de tiempo de consulta es 30 días (consultando el *endpoint* respectivo de acuerdo con su frecuencia):

```
{
  "MetricId": "metric_id",
  "StartDate": "Fecha inicial en formato ISO",
  "EndDate": "Fecha final en formato ISO",
  "Entity": "Sistema"
}
```

Oceanic Niño Index (ONI)

El Índice Oceánico de El Niño (*Oceanic Niño Index, ONI*) constituye uno de los indicadores de referencia más ampliamente empleados para la caracterización de la variabilidad climática asociada al fenómeno *ENSO* (*El Niño–Southern Oscillation*). Su relevancia radica en que permite representar, de manera sintética, el comportamiento de las anomalías de la temperatura superficial del mar en la región Niño 3.4 del Pacífico tropical, comprendida entre 5°N–5°S y 120°–170°W, zona reconocida por su papel central en la dinámica océano-atmósfera de escala interanual (NOAA, 2009).

De acuerdo con la NOAA, el ONI se estima a partir de una media móvil trimestral de las anomalías de la temperatura superficial del mar, lo que favorece la identificación de patrones persistentes por encima de fluctuaciones de corto plazo (NOAA, 2009). En este sentido, el índice se ha consolidado como uno de los criterios operativos más utilizados para la delimitación de las fases cálida y fría del ENSO, correspondientes, respectivamente, a los eventos El Niño y La Niña.

Desde una perspectiva interpretativa, 5 valores sostenidos iguales o superiores a $+0,5$ °C se asocian con condiciones de El Niño, mientras que registros iguales o inferiores a $-0,5$ °C se vinculan con condiciones de La Niña (NOAA, 2026). Esta convención no solo facilita la clasificación del estado térmico del Pacífico ecuatorial central, sino que también proporciona una base analítica robusta para examinar la propagación de sus efectos sobre distintos sistemas climáticos regionales.

En el ámbito de la investigación, el ONI resulta particularmente pertinente en estudios que exploran las interrelaciones entre la variabilidad oceánico-atmosférica y el comportamiento de variables hidroclimáticas, energéticas y económicas. En el caso específico de este trabajo, su incorporación responde a la necesidad de considerar un descriptor climático con capacidad explicativa sobre la dinámica hidrológica del sistema colombiano y, por esta vía, sobre las variaciones observadas en el precio de la energía en el mercado spot day-ahead, puesto que es un componente muy importante en la incertidumbre que caracteriza de este. Para este trabajo se extraen sus valores históricos de (NOAA, 2026).

Ahora, a partir de estas variables elegidas, se realiza un análisis exploratorio de los datos con el propósito de caracterizar medidas estadísticas básicas de las variables como medidas de tendencia central, de dispersión, asimetría, curtosis. Adicionalmente, se busca encontrar relaciones lineales y no lineales y multicolinealidad entre las variables por medio de coeficientes de correlación y tests estadísticos, así como propiedades de series de tiempo como estacionariedad, estacionalidad, autocorrelación, entre otras. Finalmente, se buscan diferentes conclusiones a partir de visualizaciones que permitan mostrar conclusiones relevantes para la fase de modelado.

Integralidad de los datos

Al realizar un sondeo de datos faltantes para el periodo de estudio, se evidenció que para las variables utilizadas en este trabajo no había necesidad de aplicar ningún tipo de estrategia para esta situación, esto se reporta en la Figura 6. Adicionalmente, se identificó que el número de datos es $N = 105,192$ que es el número de horas que hay en el intervalo de tiempo estudiado.

Asimismo, no se detectaron datos duplicados, por tanto, se llega a la conclusión de que no es necesario aplicar ninguna estrategia para manejo de duplicidad o falta de datos.

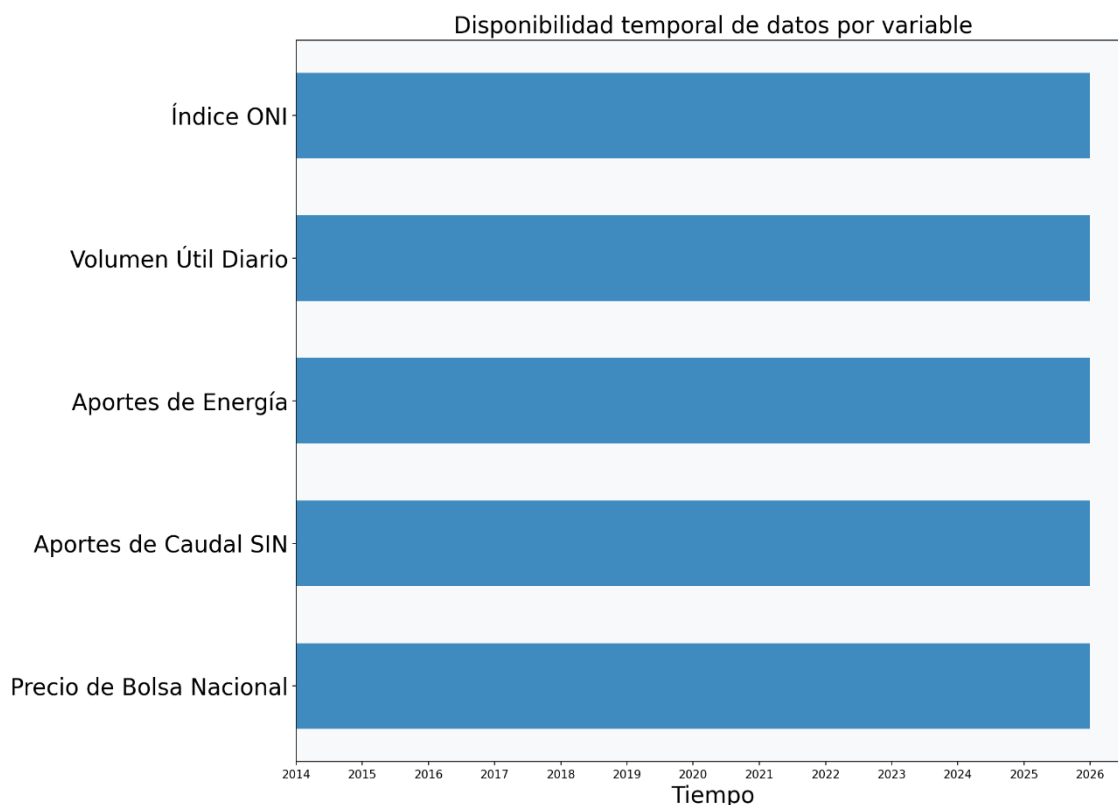


Figura 6. Disponibilidad de los datos. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026)

Estadísticas de las variables

En este apartado, se presentan las estadísticas de medidas de tendencia central, dispersión, posición y forma en las Tablas Tabla 4 Tabla 5, así como la estimación de la distribución de las variables a utilizar en este trabajo en la Figura 7. En primera instancia, se evidencia claramente que hay una diferencia de escala entre cada una de las variables que debe tenerse en cuenta y nos indica que es necesario realizar un proceso que permita mitigar este aspecto que podría afectar el aprendizaje en los modelos.

Esto último es reforzado por la heterogeneidad en que existe a nivel de dispersión y forma distribucional de las variables. En particular, el precio de la energía (PBN) presenta un valor de la media muy por encima de su mediana que junto al hecho de tener una desviación alta y un

coeficiente de variación de 96,4% nos indica que hay una volatilidad notable, lo cual es lo que se espera debido a la incertidumbre inherente a la dinámica de esta variable financiera. Adicionalmente, un valor alto positivo de asimetría y de curtosis representa que la distribución de PBN tiene colas pesadas y es asimétrico a la derecha, lo cual es debido a episodios donde el precio de la energía se elevó en una magnitud considerable para cambiar la forma de su distribución, ilustrando la complejidad esperada que tiene el entendimiento de esta variable.

Variable	Media	Mediana	Desviación Estándar	CV (%)	IQR
PBN	287.2096	186.3343	276.8780	96.40	221.9301
PAC	0.9435000	0.8686000	0.3654000	38.73	0.4425000
AES	173691200	158815700	86260070	49.66	110367800
PVUD	0.6495000	0.6642000	0.1391000	21.42	0.2061000
ONI	0.1708000	0.01	0.8860000	518.63	1.180000

Tabla 4. Medidas de tendencia central y dispersión. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026)

Variable	Mínimo	P25	P75	Máximo	Asimetría	Curtosis
PBN	38.94190	119.4082	341.3384	2821.521	2.9623	13.4640
PAC	0.2819	0.6887	1.1312	3.441800	1.2771	2.7884
AES	36780700	110043200	220411000	643404200	1.1064	1.6093
PVUD	0.2856	0.5537	0.7598	0.8824000	-0.4025	-0.6724
ONI	-1.2	-0.53	0.65	2.75	0.9032	0.3766

Tabla 5. Medidas de posición y forma. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026)

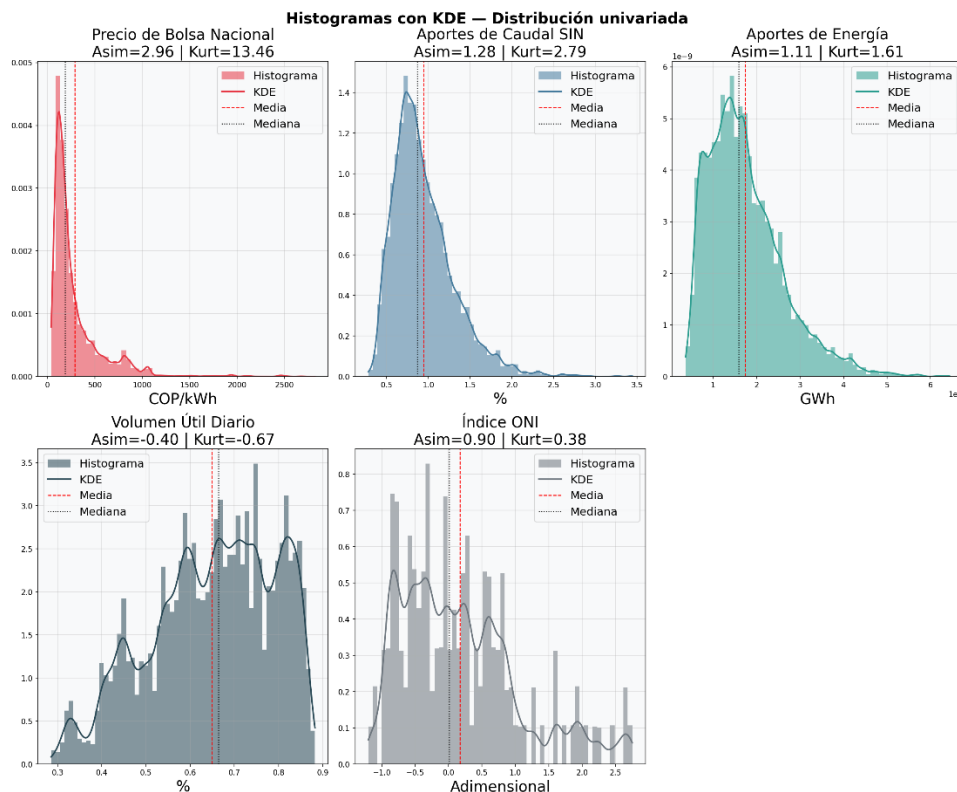


Figura 7. Distribuciones de las variables utilizando KDE*. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026). *KDE: Kernel Density Estimation es un método no paramétrico para estimar la densidad de una variable.

Las variables PAC y AES, aunque menos extremas que PBN, también exhiben rasgos distribucionales que merecen atención. En ambos casos, la superioridad de la media respecto de la mediana, junto con la asimetría positiva y la distancia apreciable entre los percentiles superiores y los valores máximos observados, sugiere la existencia de episodios de intensificación o de observaciones extremas que alteran su distribución. Este comportamiento permite inferir que dichas variables podrían estar reflejando condiciones operativas o contextuales de especial relevancia para la formación del precio, lo cual les otorga potencial capacidad explicativa en el marco del modelo predictivo.

Análisis temporal y relación entre variables

Ahora, se realiza un análisis a nivel temporal asociado a las variables de este trabajo. En primera instancia se revisan aspectos de estacionariedad, luego haremos una descomposición STL a la serie de precios de energía del mercado colombiano para identificar tendencias, estacionalidad

y ruido, posteriormente, observaremos relaciones temporales, lineales y no lineales entre las variables.

Comenzando un análisis de estacionariedad, se realiza inicialmente un análisis visual por medio de las gráficas de ACF y el PACF de la serie de precios de energía presentada en la Figura X. Donde se observa que la serie de precios de energía del mercado colombiano tiene un decaimiento en ACF lento y estadísticamente significativo hasta más de 100 días y el PACF tiene una caída abrupta después del rezago 1, estos 2 aspectos nos dan indicios de que la serie no es estacionaria y adicionalmente, su naturaleza es de memoria a largo plazo con un patrón de corto plazo que se extiende hasta pocos rezagos, de acuerdo con el PACF, 1 rezago.

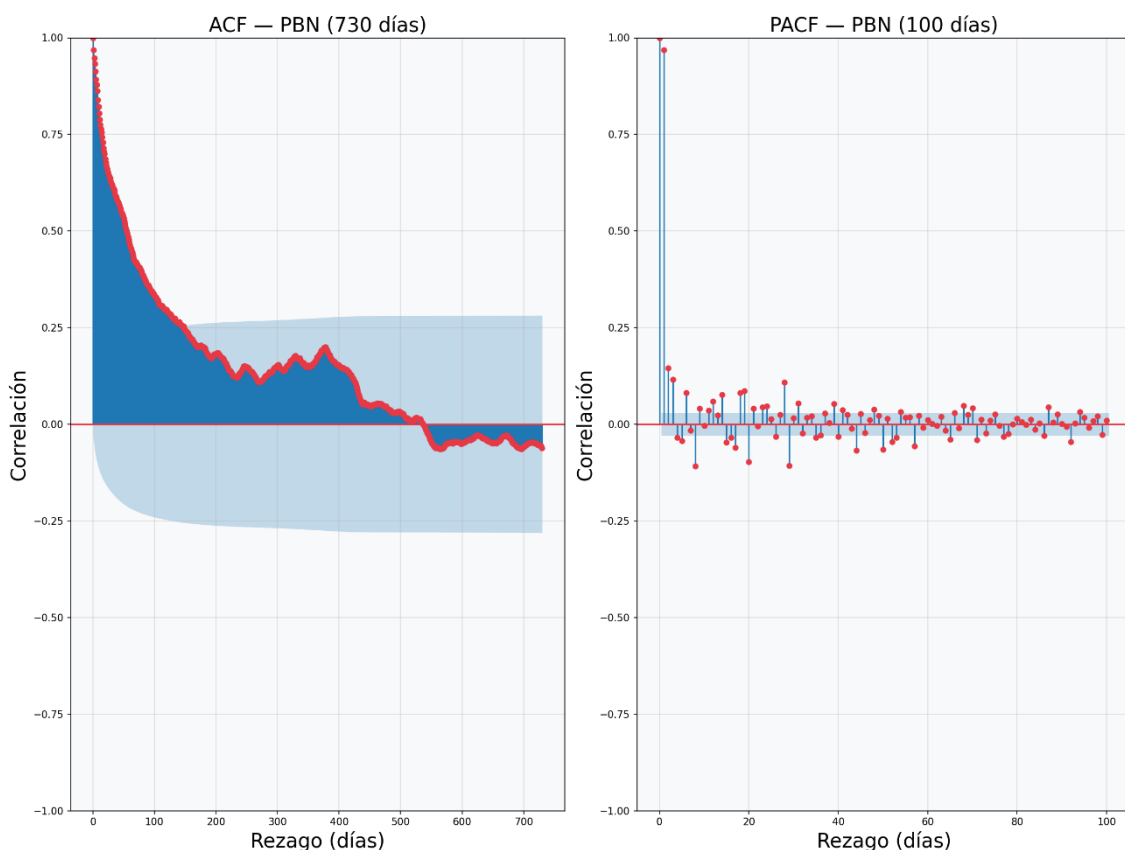


Figura 8. Diagrama ACF y PACF de PBN. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.)

En adición a lo anterior, se realizan 2 pruebas estadísticas para comprobar la significancia estadística de que la serie PBN es estacionaria o no. En este trabajo, se aplica la prueba de Dickey-Fuller Aumentado (ADF) que determina si la media y la varianza de una serie es constante en el tiempo y sus hipótesis son las siguientes:

H_0 : la serie tiene raíz unitaria (no estacionaria)

H_1 : la serie no tiene raíz unitaria (es estacionaria)

Y también utilizaremos la prueba de Kwiatkowski-Phillips-Schmidt-Shin (KPSS) que evalúa si la serie de tiempo es estacionaria alrededor de una media o una tendencia lineal a través de raíz unitaria como ADF, sus hipótesis son las siguientes:

H_0 : la serie no tiene raíz unitaria (es estacionaria)

H_1 : la serie tiene raíz unitaria (no es estacionaria)

En la Tabla 6 se presentan los resultados de estas pruebas aplicados a las variables, donde observamos que hay casos donde ADF indica que hay evidencia estadística significativa del 5% de que las series son estacionarias, pero la prueba KPSS señala lo contrario, es decir en ambas pruebas se rechaza las hipótesis nulas, por lo tanto, esto nos permite concluir que las series son estacionarias según ADF, pero fraccionalmente integrado y con memoria larga por parte de KPSS, comprobando lo mencionado anteriormente.

Variable	p-valor ADF	p-valor KPSS
PBN	0.0002	0.0100
PAC	0.000	0.0290
AES	0.000	0.0100
PVUD	0.000	0.0242
ONI	0.000	0.0100

Tabla 6. Resultados pruebas ADF y KPSS. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026)

Adicionalmente, en la Figura 9 se observan las series temporales de las variables de este trabajo, en donde podemos notar que la tendencia entre PAC, AES, PVUD tiende a ser muy similar, y contrario a PBN y ONI, dándonos los primeros indicios de una relación inversa entre PBN y PAC, AES y PVUD y directa entre PBN y ONI, aspectos que tienen sentido, ya que como se explicó en la sección de contexto el precio de energía en el mercado colombiano se ve fuertemente afectado ante periodos de baja capacidad hídrica y presencia del fenómeno de El Niño (valores de ONI altos). En adición, se observa que el comportamiento cíclico de PVUD es más suave,

puesto que corresponde a valores de niveles acumulados, mientras que PAC y AES al ser flujos de entrada y salida constante tienen una volatilidad más marcada, aspecto que debe ser importante considerar al momento de modelar, ya que PAC y AES podrían generar ruido en algunos periodos.

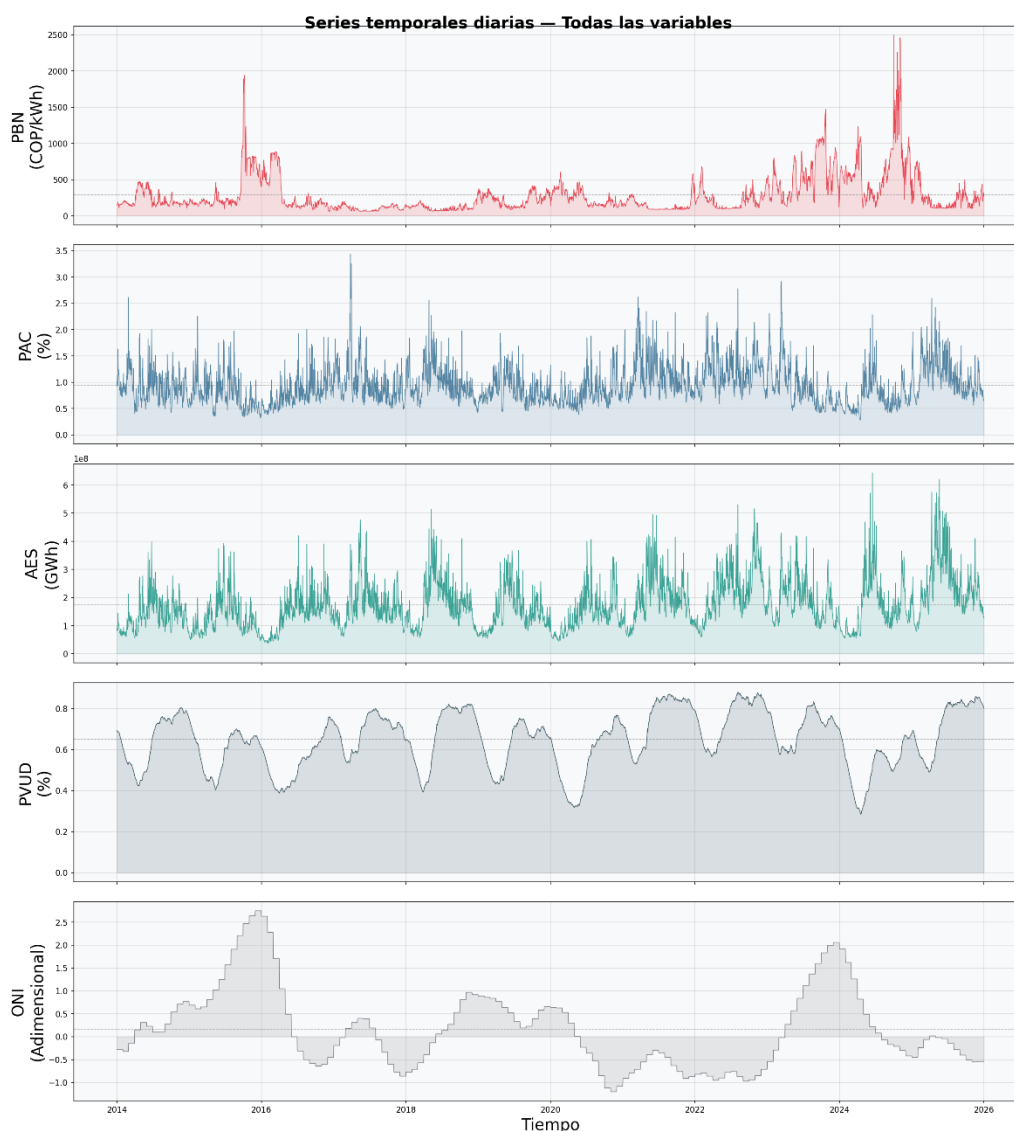


Figura 9. Series temporales de las variables de estudio. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026)

Por otro lado, en la Figura 10 tenemos la descomposición STL de la serie temporal de precios de energía en el mercado colombiano (PBN), de donde podemos observar una tendencia creciente marcada con algunas subidas o bajadas coincidentes con periodos climáticos o estructurales. Adicionalmente, se observa que el patrón estacional ha ido creciendo con el tiempo, lo que

puede indicar una estacionalidad multiplicativa. Sin embargo, el aspecto más importante es que se evidencia un residuo alto principalmente en épocas coincidentes con el fenómeno de El Niño, lo que nos indica que para poder modelar el comportamiento de esta serie temporal es necesario utilizar variables exógenas que permitan representar estos patrones que la descomposición no es capaz de identificar.

Estos patrones desconocidos para la descomposición se planean modelar con las variables propuestas, ya que en comportamientos como el ilustrado en la Figura 11 se identifica que los valores extremos del PBN son coincidentes o tienen un rezago con respecto a épocas del fenómeno de El Niño y la Niña, por ejemplo, los valores más altos con épocas con El Niño, por lo que el ONI parece ser una variable que debe estar en nuestros modelos.

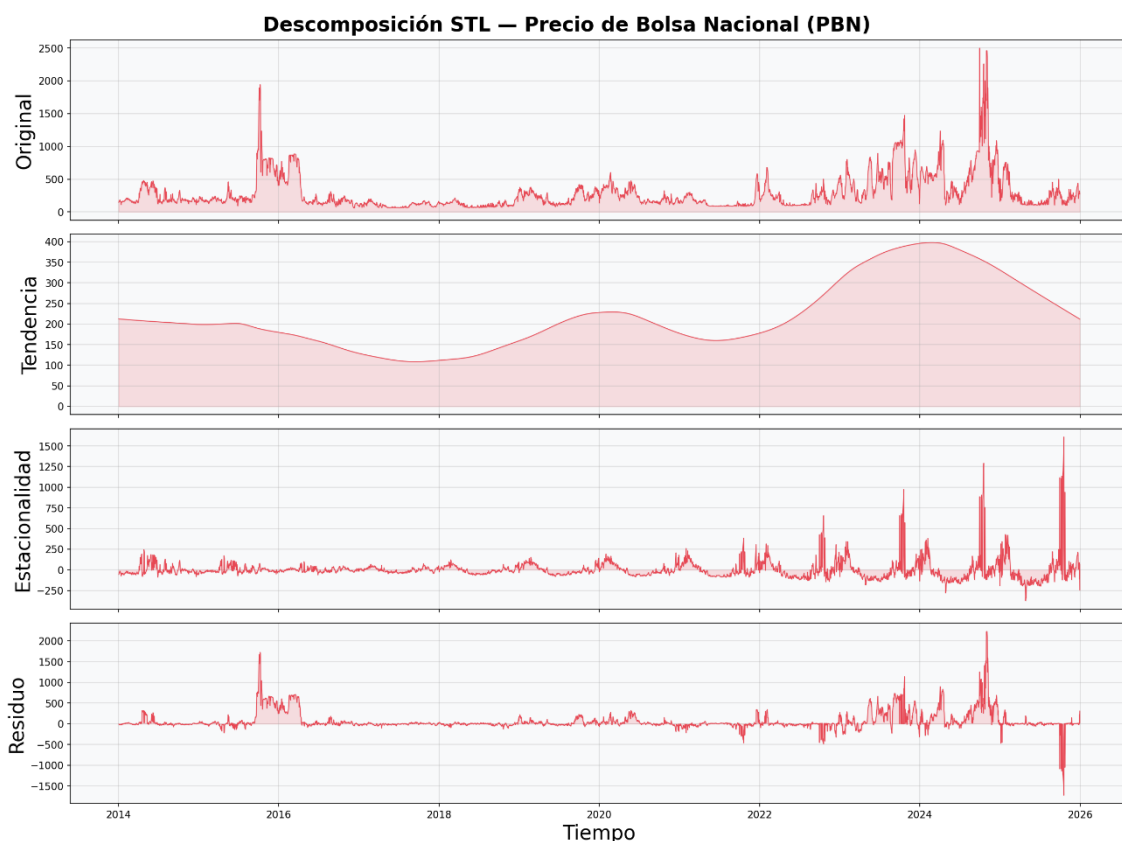


Figura 10. Descomposición STL del PBN. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.). STL: Seasonal and Trend decomposition using Loess: es un método muy usado para dividir una serie temporal en tendencia, estacionalidad y ruido.*

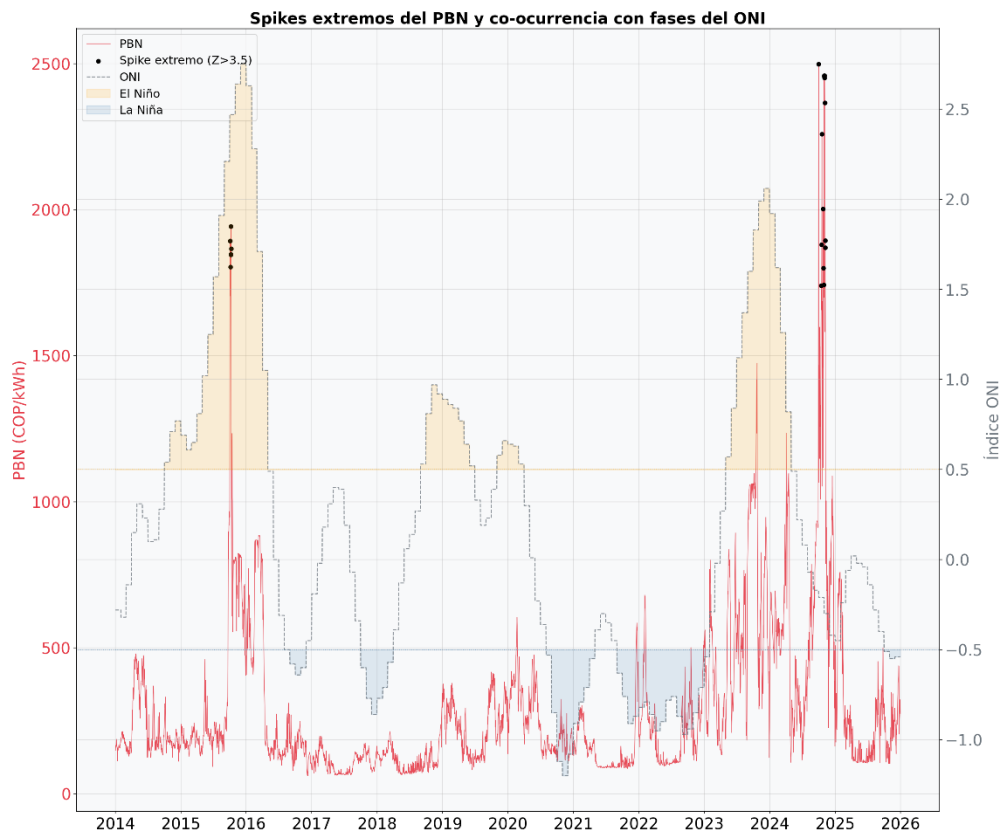


Figura 11. Relación de valores extremos del PBN y su co-ocurrencia con el fenómeno de El Niño y La Niña.
Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026)

Ahora, con respecto a la medición de relaciones entre variables utilizaremos el coeficiente de Pearson que permite medir la relación lineal entre 2 variables en un rango entre -1 y 1, siendo valores positivos indicio de relaciones directas y negativos relaciones inversas, es decir mide la fuerza lineal a través de la desviación de los valores de cada variable con respecto a su media con la siguiente fórmula:

$$r_{pearson} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 (Y_i - \bar{Y})^2}}$$

Adicionalmente, se calcula el coeficiente de Spearman que permite la medición de relaciones monótonas crecientes o decrecientes entre 2 variables no necesariamente linealmente, por lo que se usa como métrica para capturar relaciones no lineales entre las variables, esto lo hace calculando el coeficiente de Pearson sobre el rango de cada variable lo que permite centrarse en el orden relativo y no en la magnitud, de la siguiente manera:

$$r_{spearman} = r_{pearson}(rango(X), rango(Y))$$

Asimismo, también se calcula la diferencia absoluta entre estos coeficientes para identificar qué tipo de relación (lineal o no lineal) es predominante. En la Figura 12 se observa las matrices de correlación y la diferencia entre ellas. Donde se observa que, en términos generales, todas las variables exógenas presentan correlaciones estadísticamente significativas con el Precio de Bolsa Nacional (PBN), aunque con magnitudes moderadas que reflejan la complejidad multifactorial del mercado eléctrico colombiano.

La variable con mayor correlación con el PBN es el Índice ONI, con un coeficiente de Pearson de 0.41 y de Spearman de 0.35. La diferencia absoluta entre ambos coeficientes es de apenas 0.06, lo que indica que la relación ONI–PBN es predominantemente lineal dentro del rango de variación histórica del índice, sugiriendo que esta variable puede ser incluida directamente en los modelos sin requerir transformaciones no lineales de sus valores como covariable exógena.

Por otro lado, las variables hídricas PAC y AES presentan correlaciones negativas con el PBN, con coeficientes de Pearson de -0.38 y -0.28 respectivamente, que se intensifican al medirse con Spearman hasta -0.46 y -0.41 . Este incremento sistemático del coeficiente de Spearman respecto al de Pearson —con diferencias absolutas de 0.08 para PAC y 0.13 para AES— constituye evidencia estadística de que ambas relaciones tienen una componente no lineal significativa de tipo convexo: cuando los aportes de caudal y energía son muy bajos, el PBN no sólo aumenta, sino que lo hace de forma desproporcionada, generando los picos de precio característicos de los periodos de estrés hídrico. Este comportamiento no lineal justifica el uso de arquitecturas de aprendizaje profundo con funciones de activación no lineales. Adicionalmente, la elevada correlación entre PAC y AES —de 0.77 en Pearson y 0.76 en Spearman, con una diferencia de apenas 0.01— confirma que ambas variables son cuasi-redundantes, dado que describen el mismo fenómeno físico (el ingreso de agua al sistema) desde perspectivas de flujo de caudal y energía potencial respectivamente.

El Porcentaje de Volumen Útil Diario (PVUD) presenta una correlación negativa con el PBN de -0.22 en Pearson y -0.34 en Spearman, con una diferencia de 0.12 que sitúa a esta variable como la que exhibe mayor no linealidad relativa en su relación con el precio de bolsa. Aunque la magnitud del coeficiente de Pearson podría sugerir una relación débil, el análisis conjunto con Spearman y la diferencia entre ambos revela que la asociación real es considerablemente más fuerte de lo que el coeficiente lineal indica, concentrándose en los extremos de la distribución: cuando el PVUD desciende por debajo de niveles críticos, el impacto sobre el PBN es exponencialmente mayor que en condiciones normales de almacenamiento. Este hallazgo es

coherente con la naturaleza física del sistema de embalses colombiano, donde la pérdida de capacidad de regulación hídrica en condiciones de sequía extrema fuerza el ingreso masivo de generación térmica de alto costo al despacho, produciendo saltos de precio que no pueden ser capturados por modelos que asuman linealidad en esta relación.

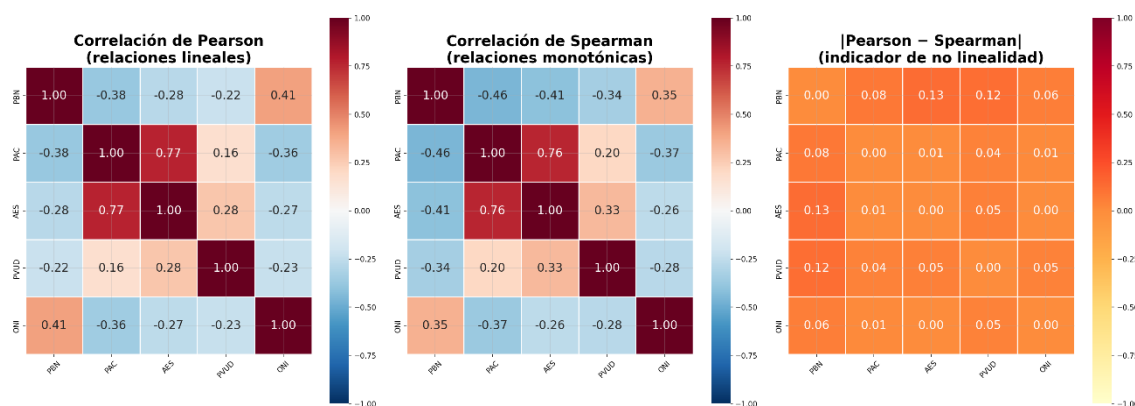


Figura 12. Matrices de correlación de Pearson y Spearman y diferencia entre ambas. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026)

Complementando la búsqueda de relaciones útiles, se realiza un análisis de correlación cruzada (CCF) entre cada variable exógena y el PBN con rezagos de 0 a 90 días proporciona información fundamental sobre la estructura temporal de la predictibilidad del precio de bolsa, determinando no solo si una variable es relevante para la predicción sino también con cuántos días de anticipación su información es estadísticamente significativa.

En la Figura 13 se observa el resultado de este análisis donde el Porcentaje de Aportes de Caudal (PAC) y los Aportes de Energía al Sistema (AES) exhiben su correlación cruzada más intensa en el rezago 0 ($r = -0.38$ y $r = -0.28$ respectivamente), con una curva que decrece gradualmente en valor absoluto a medida que aumenta el rezago, pero se mantiene estadísticamente significativa —por encima del intervalo de confianza al 95%— durante los 90 días analizados. Este patrón indica que la información del caudal aportante tiene una utilidad predictiva sobre el PBN que es inmediata pero también persistente a lo largo del tiempo, reflejando la capacidad del sistema de embalses para amortiguar las variaciones de corto plazo en los aportes: un día de caudal bajo no impacta de inmediato el precio si los embalses están bien cargados, pero la acumulación sostenida de déficit hídrico termina por transmitirse al mercado de precios con un retardo que puede extenderse varios meses. Para el diseño de los modelos de predicción, esta estructura sugiere incluir el PAC con múltiples rezagos en el vector de entrada, considerando ventanas de hasta 30 días. Adicionalmente, la similitud entre PAC y AES está en concordancia con la colinealidad identificada en la matriz de correlación de Pearson.

Con respecto Porcentaje de Volumen Útil Diario del Sistema (PVUD), a diferencia del PAC y el AES, el PVUD alcanza su correlación cruzada más intensa con el PBN en el rezago 89 días ($r = -0.29$), con una curva que crece progresivamente en valor absoluto desde el rezago 0 hasta el extremo del horizonte analizado. Este comportamiento revela que el nivel de almacenamiento hídrico de los embalses del SIN predice mejor el precio de bolsa dentro de tres meses que el precio de mañana, lo que es físicamente interpretable: el estado actual de los embalses determina la capacidad de respuesta del sistema ante futuros periodos de estiaje, y su efecto sobre los precios se materializa con el vaciado gradual de los embalses en condiciones de déficit hídrico. Esta propiedad convierte al PVUD en una variable de valor predictivo diferido excepcional, y justifica su inclusión con rezagos explícitos de 30, 60 y 90 días en el vector de covariables de entrada para todos los modelos considerados

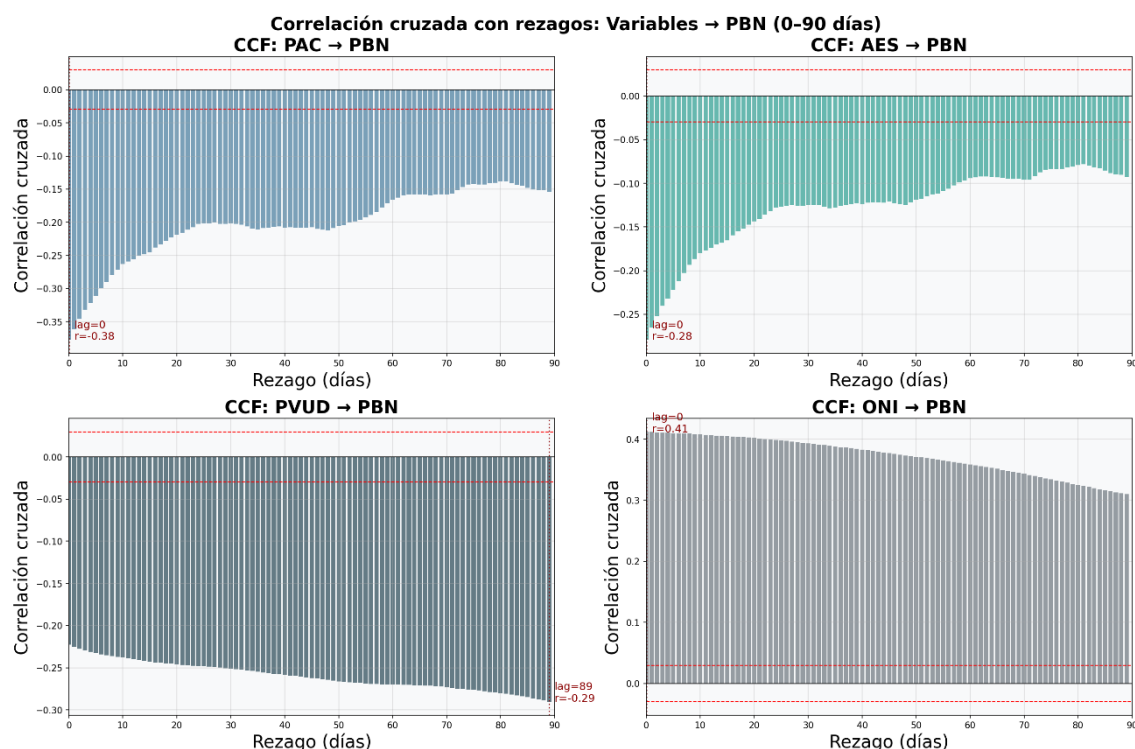


Figura 13. Correlaciones cruzadas con de PAC, AES, PVUD y ONI con PBN. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026)

También se realiza el análisis del Factor de Inflación de Varianza (VIF) para todas las variables de este trabajo calculado como:

$$VIF_j = \frac{1}{1 - R_j^2} \text{ para una variable } j$$

Donde R_j^2 es el coeficiente de determinación al incluir la variable x_j contra el resto de las variables, esta medida es muy útil para detectar multicolinealidad. En la Figura X se presenta el valor asociado a cada una de las variables usando como umbral moderado y alto 5 y 10 respectivamente. De donde se observa que no hay indicios de multicolinealidad, lo que nos indica que es posible utilizar cada una de las variables en nuestro vector de características para los modelos, ya que las variables son informativos y sin redundancia cuando son utilizadas en un mismo sistema.

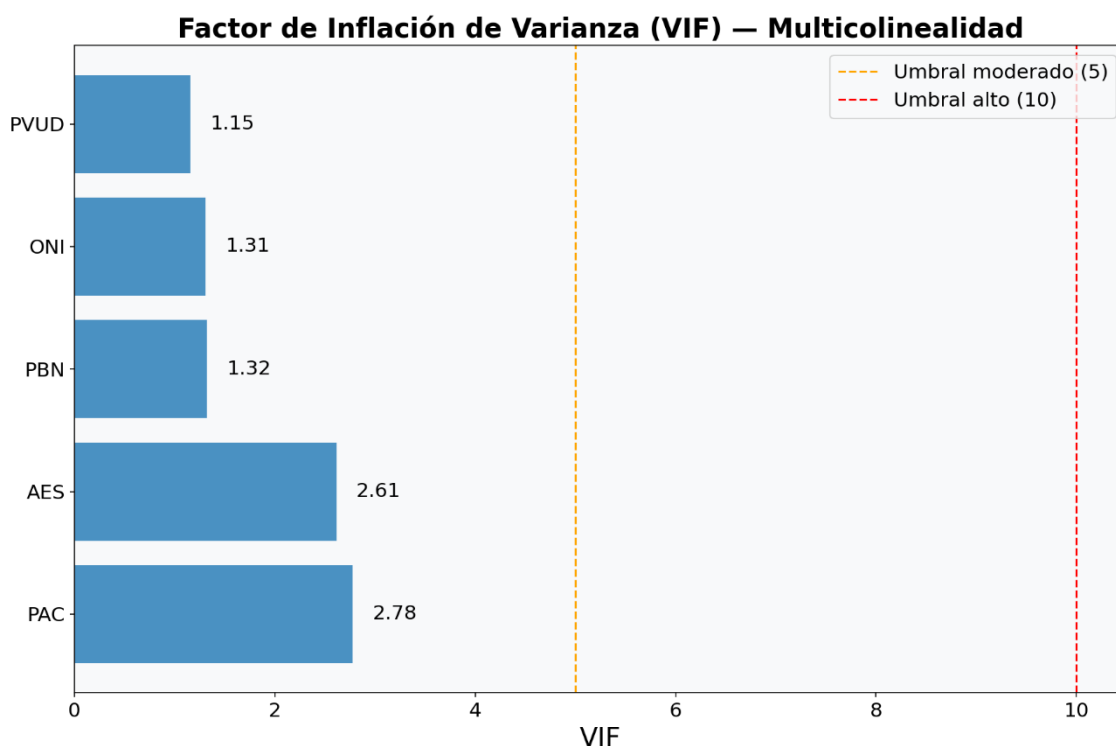


Figura 14. VIF de las variables del trabajo. Fuente: elaboración propia con datos de (Equipo Analítica XM, s. f.; NOAA, 2026)

Finalmente, se realiza un gráfico del boxplot de los precios del mercado de energía colombiano por año ilustrado en la Figura 15, de donde se observa notables diferencias distribucionales de un periodo a otro, lo que indica que la varianza no es constante en el tiempo y esta serie es heterocedástica y, por consiguiente el comportamiento del PBN está compuesto por diferentes regímenes caracterizados por diversos factores, por ejemplo el climático, ya que en épocas del fenómeno de El Niño como en 2015-2016 y 2023-2024, la amplitud de la caja es mucho más grande, evidenciando una varianza mayor con respecto a otros periodos.

Este último aspecto es muy importante, porque debe ser tenido en cuenta en la estrategia de partición en los conjuntos de entrenamiento, validación y test para la evaluación de nuestros modelos bajo escenarios reales, implicando por ejemplo en tener conjuntos de prueba de hasta 1 año o más para que el modelo esté inmerso en la realidad y ver su rendimiento real.

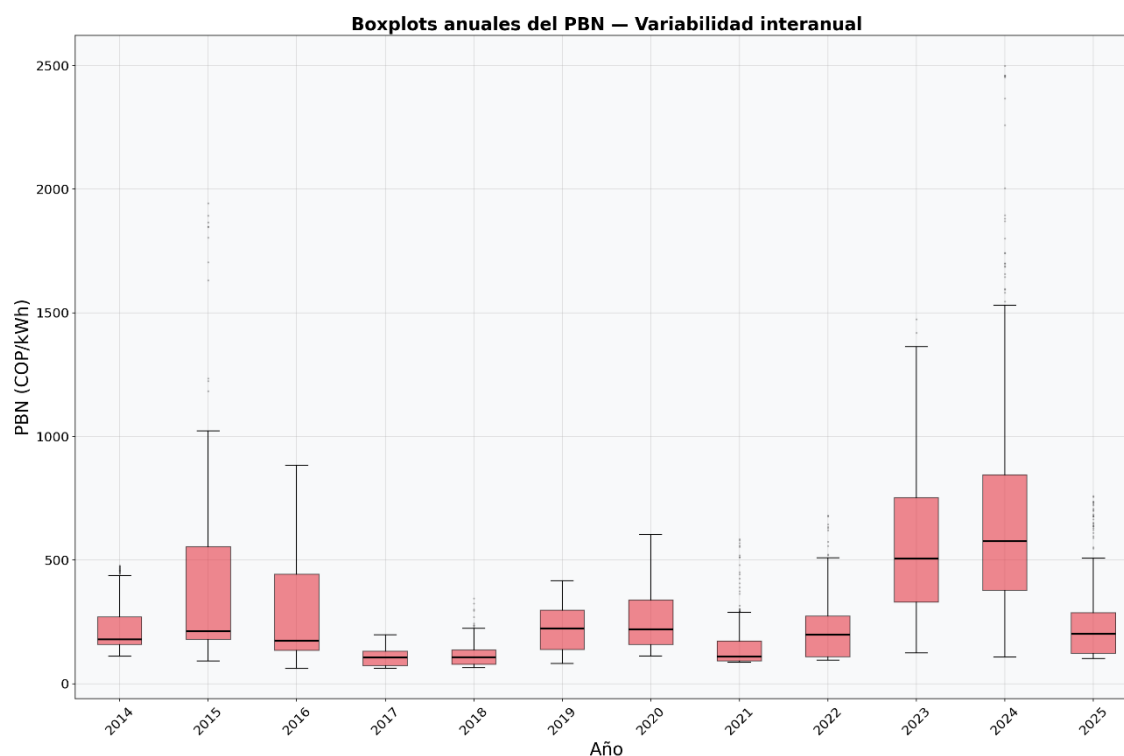


Figura 15. Boxplot interanual del PBN. Elaboración propia con datos de (Equipo Analítica XM, s. f.)

4.2.3 Preparación de los datos

A partir del análisis exploratorio de los datos se obtienen conclusiones útiles que sirven para diseñar un flujo de transformaciones alrededor de las variables de estudio que sean adecuadas y faciliten el proceso de aprendizaje de los modelos seleccionados. Para explicar el flujo de procesamiento de los datos antes de pasar a los modelos, se ilustra el flujo en la Figura 16.

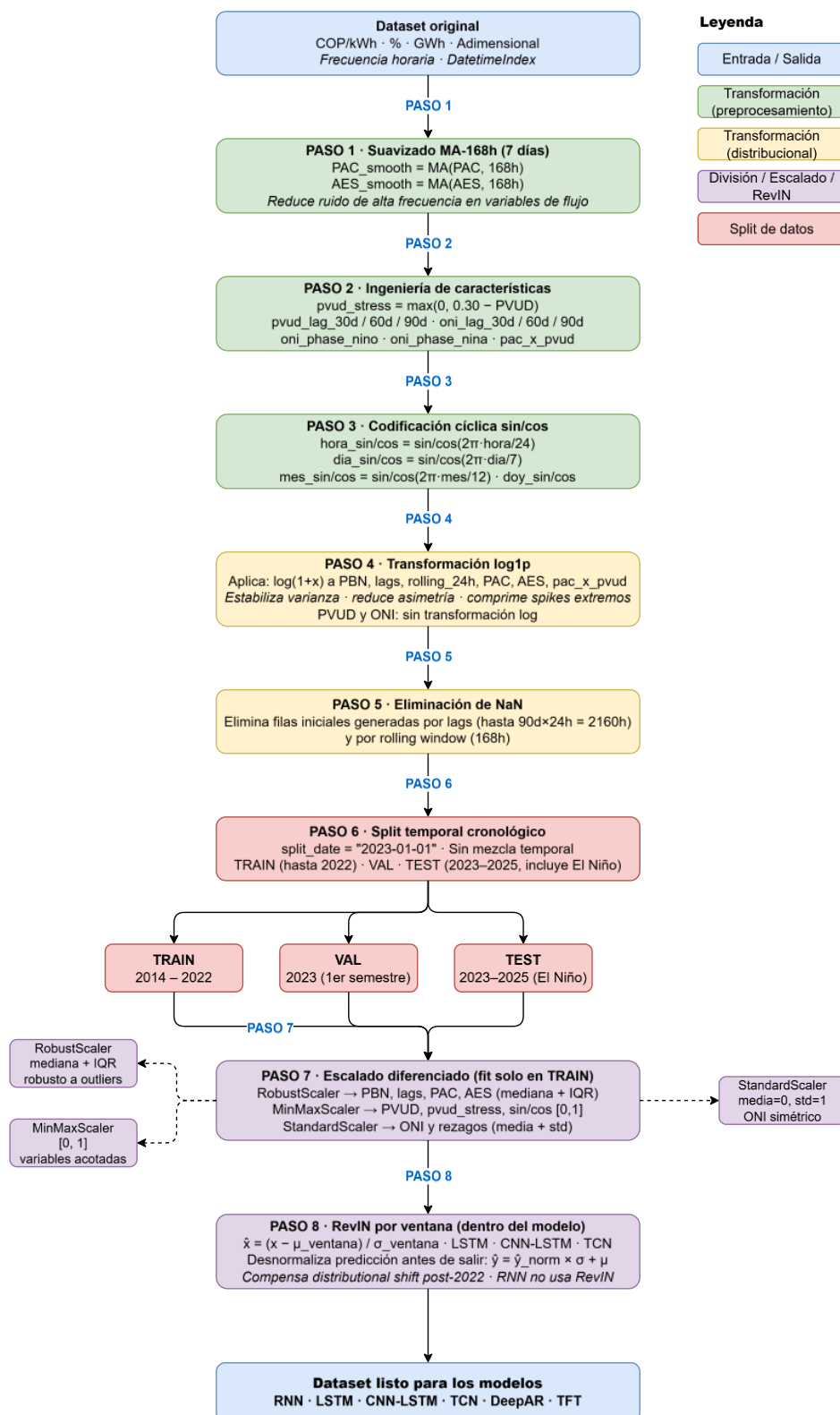


Figura 16. Flujo de preparación de los datos. Fuente: elaboración propia

Paso 1

Para este primer paso, se comienza con un suavizado de las variables PAC y AES, ya que por su naturaleza son flujos de entrada y salida sensibles a factores externos que tienen mucha incertidumbre como los climáticos, esto implica que sus valores tengan una volatilidad muy marcada de picos que podrían provocar ruido para los modelos, por lo que se procede a realizar un suavizado de media móvil de 7 días (168 horas), ya que es el periodo de tiempo útil para que esta señal contribuya en el PBN. También se dejan los valores originales para capturar información del corto plazo.

Paso 2

En este paso, procederemos a aplicar ingeniería de características sobre las variables de estudio para crear otras que permitan dar más información. En primer lugar, se identificó que cuando el PVUD está en niveles por debajo de 0.3 (situación de estrés hídrico importante) los precios tienden a dispararse, por lo que creamos una variable para representar esta situación como $\max(0, 0.3 - PVUD)$.

Adicionalmente, basándonos en los diagramas de correlación cruzada creamos variables de rezagos de 30, 60 y 90 días para PVUD y ONI, ya que se evidenció que los valores de estas hoy proveen información útil para predecir el precio dentro de hasta 3 meses. Asimismo, se crean 2 variables binarias que indican si en el instante t hay ocurrencia del fenómeno de El Niño o de La Niña, utilizando como umbral 0.5 y -0.5 respectivamente. Finalmente, se crea una variable que represente el sistema hídrico en cada instante del tiempo como el producto $PVUD * PAC$, ya que esto representa la interacción entre el almacenamiento y flujo del agua.

Paso 3

Para este paso, se realiza el último preprocesamiento de los datos antes de comenzar con transformaciones distribucionales. Se procede a agregar información básica a nivel de calendario al modelo: si el día es fin de semana, el mes, el día y la hora. Sin embargo, es importante tener en cuenta que debemos proyectar un comportamiento cíclico, por ejemplo para las horas, el valor 23 es cercano a 0, pero numéricamente esto no es así; es por ello que realizamos una proyección circular utilizando funciones de seno y coseno de la siguiente manera:

$$hora_{proyectada_sin} = \sin\left(\frac{2\pi}{24} * hora\right)$$

$$hora_{proyectada_cos} = \cos\left(\frac{2\pi}{24} * hora\right)$$

$$dia_{proyectado_sin} = \sin\left(\frac{2\pi}{7} * dia\right)$$

$$dia_{proyectado_cos} = \cos\left(\frac{2\pi}{7} * dia\right)$$

$$mes_{proyectado_sin} = \sin\left(\frac{2\pi}{12} * mes\right)$$

$$mes_{proyectado_cos} = \cos\left(\frac{2\pi}{12} * mes\right)$$

Paso 4

En este paso, se realizan las transformaciones distribucionales. Se empieza aplicando la transformación $\log(1 + x)$ a las variables PBN y sus rezagos, PAC y AES junto con sus suavizados y a $PVUD * PAC$ con el propósito de estabilizar su varianza, forma y posición distribucional, ya que como se vio el PBN se caracteriza por su heterocedasticidad, por su asimétrica y curtosis leptocúrtica con colas pesadas. Adicionalmente, se observó que PAC y AES tienen más fuerza como correlación no lineal con PBN, y esta transformación permite linealizarla un poco. Finalmente, se usa $\log(1 + x)$ en lugar de $\log(x)$ para mitigar eventos donde haya flujo 0 en variables como PAC y AES.

Paso 5

En esta etapa se termina de dejar el conjunto de datos listo antes de comenzar a hacer la partición en conjunto de entrenamiento, validación y prueba. Esto es, eliminar los valores NaN que surgen al agregar los valores de diferentes rezagos.

Paso 6

En este paso, se realiza la partición del conjunto de datos en entrenamiento, validación y pruebas antes de pasar al procesamiento de escalamiento. Para este trabajo se opta por tomar la serie temporal desde 2014-01-01 hasta 2022-12-31 como conjunto de entrenamiento de tal manera que los modelos puedan aprender patrones de diferentes regímenes como episodios del fenómeno de El Niño. Se usa el periodo del primer semestre de 2023 como periodo de

validación y finalmente desde el segundo semestre de 2023 hasta 2025-12-31 como conjunto de pruebas de tal manera que tengamos un horizonte considerable y con episodios de El Niño para evaluar los modelos en situaciones extremas, siguiendo las sugerencias de Lago et al. (2021) de incluir un tiempo considerable de 1 a 2 años y capture el comportamiento de energías renovables, que en Colombia es el tipo de energía dominante, en particular la hídrica como se ha visto durante este trabajo.

Paso 7

Para este paso se aplican las transformaciones de escala de escalamiento estándar, robusto y minmax, que se encuentran disponibles en librerías como scikit-learn de Python. Las variables cuya varianza, forma y posición distribucional y presencia de valores extremos sea más evidente se les aplica un escalamiento robusto basado en la mediana y el rango Inter cuartil (IQR) las cuales son de acuerdo con el análisis exploratorio de los datos: PAC y sus rezagos y media móvil de 24 horas, PAC y PVUD y sus versiones suavizadas y $PAC * PVUD$. Con respecto a la transformación minmax, se realiza sobre las variables calendario que al aplicar funciones de senos y cosenos están en un rango $[-1, 1]$, por lo que para mantener estabilidad numérica y convergencia de los modelos se escala a $[0, 1]$. Finalmente, el escalamiento estándar solo se utiliza sobre el ONI y sus rezagos, puesto que es una variable que está centrada en 0 y no posee valores extremos.

Por otro lado, es importante mencionar que estas transformaciones se aplican por separado en los conjuntos de entrenamiento, validación y pruebas, ya que realizarlo sobre todo el conjunto de datos implicaría problemas de cambio distribucional o *distributional shift* puesto que los datos y los modelos estarían observando información que no debería y habría una menor caracterización de cada régimen, se realiza el ajuste de los escaladores con el conjunto de entrenamiento y luego con esos parámetros locales, se utiliza en el conjunto de validación y pruebas.

Paso 8

En este paso se incluye una capa que se usa en los modelos presentados en la sección siguiente LSTM, CNN-LSTM y TCN y permite ser otra forma de mitigar el problema de cambio distribucional o *distribution shift* el cual es un problema en nuestro contexto por el cambio de regímenes marcado por fenómenos climáticos y por el cambio del valor del dinero en el tiempo que provocan que la serie no sea estacionaria y sea heterocedástica presentando cambios considerables en los valores en algunos periodos de tiempo. Por tanto, lo que se realiza es

normalizar cada ventana con las estadísticas propias de ella con una estrategia propuesta por Kim et al. (s. f.) denominada *Reversible Instance Normalization (RevIN)* definida como:

$$\begin{aligned}\mu_i &= \frac{1}{V} \sum_{t=1}^V x_{it} \\ \sigma_i &= \sqrt{\frac{1}{V} \sum_{t=1}^V (x_{it} - \mu_i)^2 + \epsilon} \\ u_{it} &= \frac{x_{it} - \mu_i}{\sigma_i}\end{aligned}$$

Para un vector $x \in \mathbb{R}^{V \times d}$ con V y d el tamaño de la ventana y número de variables respectivamente, μ_i, σ_i, u_{it} son la media, la desviación estándar y el valor escalado de la característica i . De esta manera u_{it} representa el valor que recibirá el modelo, pero al final se desnormaliza la predicción, ya que el modelo también la genera con esa escala, de la siguiente manera:

$$\hat{y}_{pred} = \hat{y}_t \sigma_y + \mu_y$$

Siendo μ_y y σ_y la media y desviación estándar de la variable de precios en la ventana y $\hat{y}_t$ la estimación del modelo.

4.2.4 Modelado

Como se ha mencionado en los objetivos de este trabajo, la línea de solución mediante experimentación y el análisis previo realizado es a través de modelos de *deep learning* los cuales se caracterizan por su alta capacidad de abstracción de patrones que pueden llegar a ser muy complejos para otros modelos o incluso un ser humano. Sin embargo, es claro que su planteamiento y forma de enseñarles debe llevarse a cabo de una manera cuidadosa buscando un *trade-off* entre el sesgo y la varianza del modelo, aspectos que son mitigados por medio del análisis y procesamiento de datos de las secciones anteriores que permiten sopesar sobre cuál es la información más útil de acuerdo con el contexto y como representarlo cuantitativamente para que los modelos puedan entender.

Adicionalmente, cada modelo tiene un conjunto de hiper-parámetros que debemos calibrar que son propio de cada modelo y nos puede traer beneficios en el entrenamiento o desventajas si no sabemos manejarlos. Es por ello por lo que a continuación se presentan los modelos que se utilizan para este trabajo y cómo se aplican al contexto del problema estudiado de la predicción de precios de energía en el mercado spot colombiano.

Para todos los problemas consideremos $\mathbf{D} \in \mathbb{R}^{N \times d}$ la matriz que contiene las N observaciones del conjunto de datos para cada instante del tiempo $t \geq 0$ para cada una de las d variables utilizadas en el estudio, es decir $\mathbf{x}_t \in \mathbb{R}^{1 \times d}$ es un vector de una observación en el instante t con los valores de cada variable (PBN y sus rezagos, PAC, AES, PVUD, ONI, etc). Sin embargo, para los modelos, la entrada se trabaja con una ventana de tamaño V , por lo que nuestra entrada será una matriz $\mathbf{X}_t = [\mathbf{x}_{t-V+1}, \dots, \mathbf{x}_t] \in \mathbb{R}^{V \times d}$. Y sea $\mathbf{y}_t = [y_{t+1}, \dots, y_{t+H}] \in \mathbb{R}^H$ el vector de precios de energía del día siguiente (es decir $H = 24$) que se quiere predecir, es decir queremos encontrar por medio de los modelos una función:

$$f: \mathbb{R}^{N \times d} \rightarrow \mathbb{R}^H$$

Tal que

$$\hat{\mathbf{y}}_t = f(\mathbf{X}_t) = [\hat{y}_{t+1}, \dots, \hat{y}_{t+H}]$$

Siendo $\hat{\mathbf{y}}_t$ el vector de precios estimados por el modelo.

Recurrent Neural Network (RNN)

Este modelo es uno de los más reconocidos en el campo de *deep learning* para trabajar con datos secuenciales como en este caso, una serie temporal. Su mecanismo se basa en crear una “memoria” a corto plazo a partir de transferir conocimiento por medio de la integración de la información de la capa oculta anterior $\mathbf{H}_{t-1} \in \mathbb{R}^{V \times h}$ (siendo h el número de neuronas en la capa oculta) y el conjunto de datos del instante a predecir $\mathbf{X}_t$ para estimar la salida $\hat{\mathbf{y}}_t$. En la Figura 17 se representa visualmente su funcionamiento.

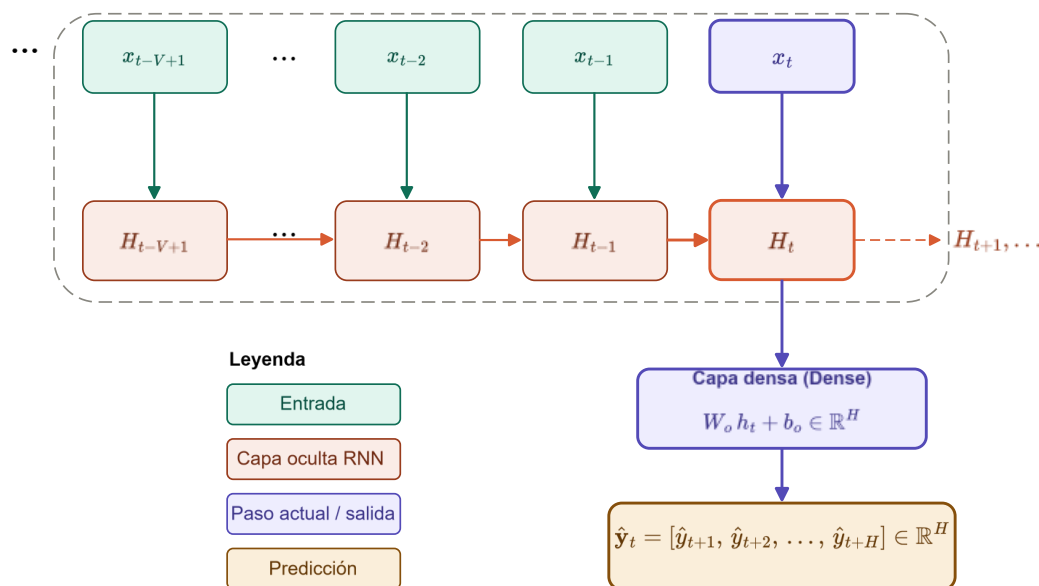


Figura 17. Mecanismo RNN. Fuente: elaboración propia

A partir de lo explicado, se observa que al momento de aplicar *backpropagation* y actualizar los pesos del modelo, hay posibilidad de que para pasos del tiempo muy hacia atrás ocurra una explosión o desvanecimiento de los gradientes, lo que dificulta que este modelo pueda tener una ventana de contexto o “memoria” a largo plazo y, por tanto, no resulte ser el modelo adecuado en ocasiones. Es por estas razones que RNN funciona como una referencia base para comparar con los demás modelos.

Long Short-Term Memory (LSTM)

Este tipo de arquitectura se utiliza para mitigar el problema de memoria a largo plazo de RNN, en particular la explosión y desvanecimiento del gradiente por medio de la incorporación de las denominadas “células de memoria” que en lugar de tener nodos recurrentes que evitan multiplicaciones recurrentes directas y gradientes más estables, tienen un estado interno que se asegura de recordar solo la información necesaria apoyándose en 3 compuertas: entrada, olvido y salida que se encargan de indicar qué tanto debe impactar la entrada al estado interno, si el estado interno debería olvidar los pasos anteriores y qué tanto debería impactar la salida al estado interno respectivamente. Matemáticamente, lo podemos expresar: sea $I_t \in \mathbb{R}^{V \times h}$, $F_t \in \mathbb{R}^{V \times h}$ y $O_t \in \mathbb{R}^{V \times h}$ las compuertas de entrada, olvido y salida, sea el nodo de entrada $\tilde{C}_t \in \mathbb{R}^{V \times h}$ y $H_t \in \mathbb{R}^h$ la capa del estado oculto con h el número de neuronas en la capa oculta y sea $\sigma(\cdot)$ la función sigmoide que permite manejar cuándo se debe olvidar y cuándo recordar. En la Figura se observa el mecanismo de este modelo.

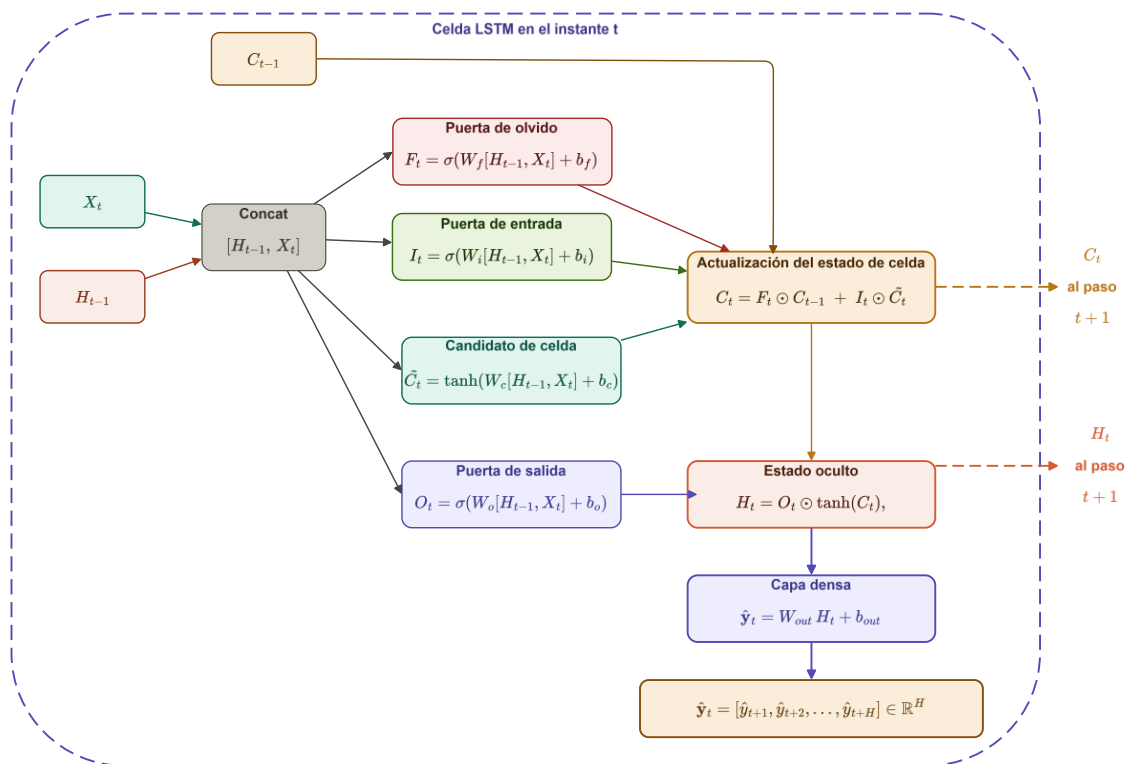


Figura 18. Mecanismo LSTM

Convolutional Neural Networks and Long Short-Term Memory (CNN-LSTM)

Otra aproximación que se propone evaluar en este trabajo es una arquitectura *CNN-LSTM*, la cual nos permite combinar una red neuronal convolucional para extraer patrones locales que puede ser muy útil para este problema, pues permite mitigar uno de los problemas que tiene la LSTM de capturar dependencias a largo plazo y lidiar con el ruido en el corto plazo debido a los picos. En este caso la CNN se encarga de resolver el problema de obtener contexto y manejar el ruido para luego codificarlo y enviarlo a una LSTM, de esta manera los gradientes del ruido y de las dependencias largas no compiten entre sí como ocurre en una LSTM pura como la explicada anteriormente. En la Figura se ve la arquitectura utilizada en este trabajo, donde cabe destacar que el argumento “causal” es necesario para evitar *data leakage*, es decir que el modelo vea datos futuros, lo cual no es deseable.

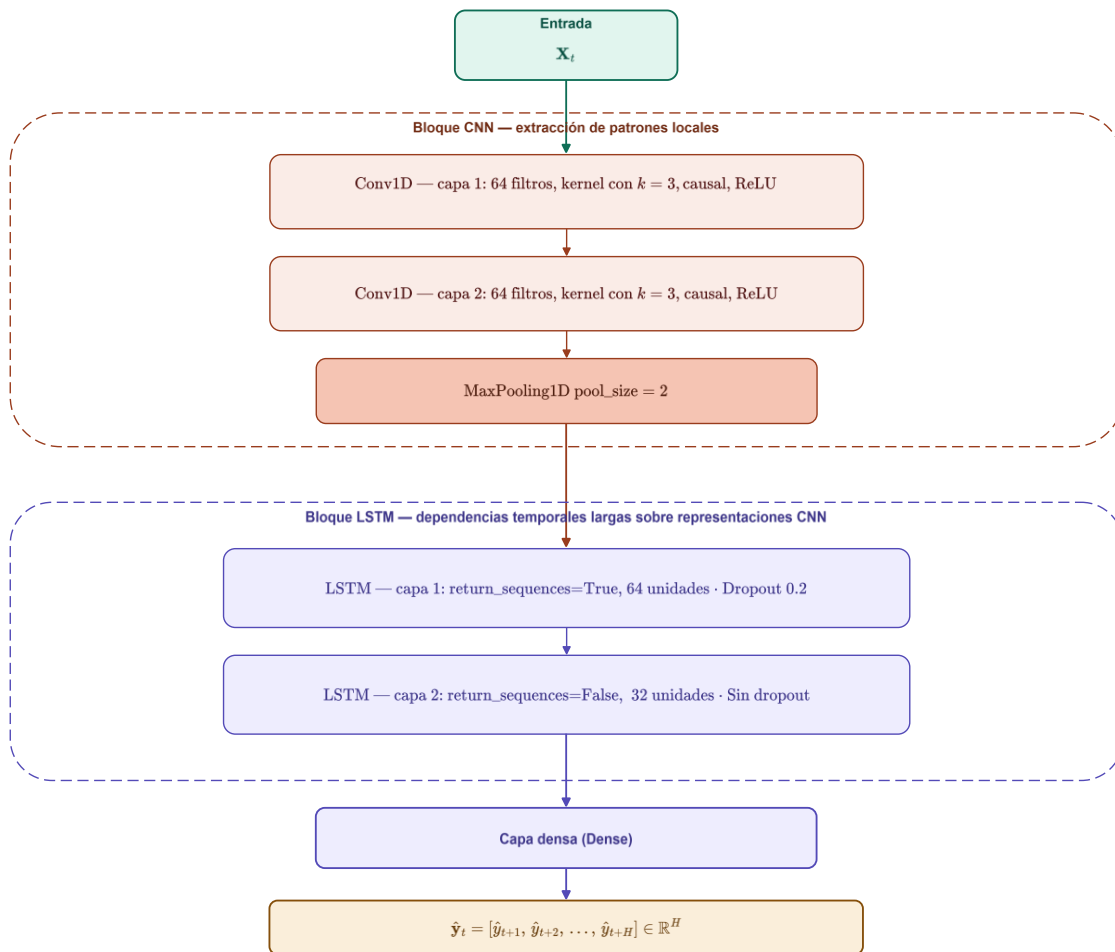


Figura 19. Arquitectura CNN-LSTM. Fuente: elaboración propia.

Temporal Convolutional Networks (TCN)

Este modelo propone una arquitectura haciendo un cambio de paradigma con respecto a los otros modelos presentados, ya que propone una estrategia para capturar dependencias en el largo plazo sin tener que utilizar propagación de memoria por medio de recurrencia, lo que permite que su cómputo no sea secuencial, sino que por medio de su propuesta se pueda realizar el proceso de manera paralela y mejorar los tiempos de entrenamiento del modelo. Para lograr esto, TCN se compone de bloques residuales compuestos por una capa de convolución dilatada, normalización, una función de activación (normalmente ReLU) y una capa *Dropout*, en cada uno de esos bloques también se aplica una estrategia causal que permite que el tamaño del kernel de las convoluciones observe información en el tiempo con pasos hacia atrás y nunca hacia adelante para ver información que el modelo no debe.

Adicionalmente, la parte principal de su mecanismo es la mencionada capa de dilatación, porque lo que permite es que el modelo no tenga que limitarse a ver pasos uno a uno hacia atrás, sino que pueda verlo en diferentes tamaños, por ejemplo, una dilatación $d = 2$, implica ver en los tiempos $[t - 4, t - 2, t]$ para un kernel de tamaño 3. Por otro lado, el hecho de no utilizar una MaxPool nos puede permitir mitigar ese riesgo en CNN-LSTM donde se baja la dimensional a la mitad por medio de una MaxPool. Estas características convierten a este modelo como una propuesta interesante para el problema de predicción, debido a que nos permite capturar contexto en el largo plazo de la serie temporal por medio de un cálculo más eficiente y con una estrategia diferente. En la Figura 20 se observa la arquitectura utilizada para este trabajo.

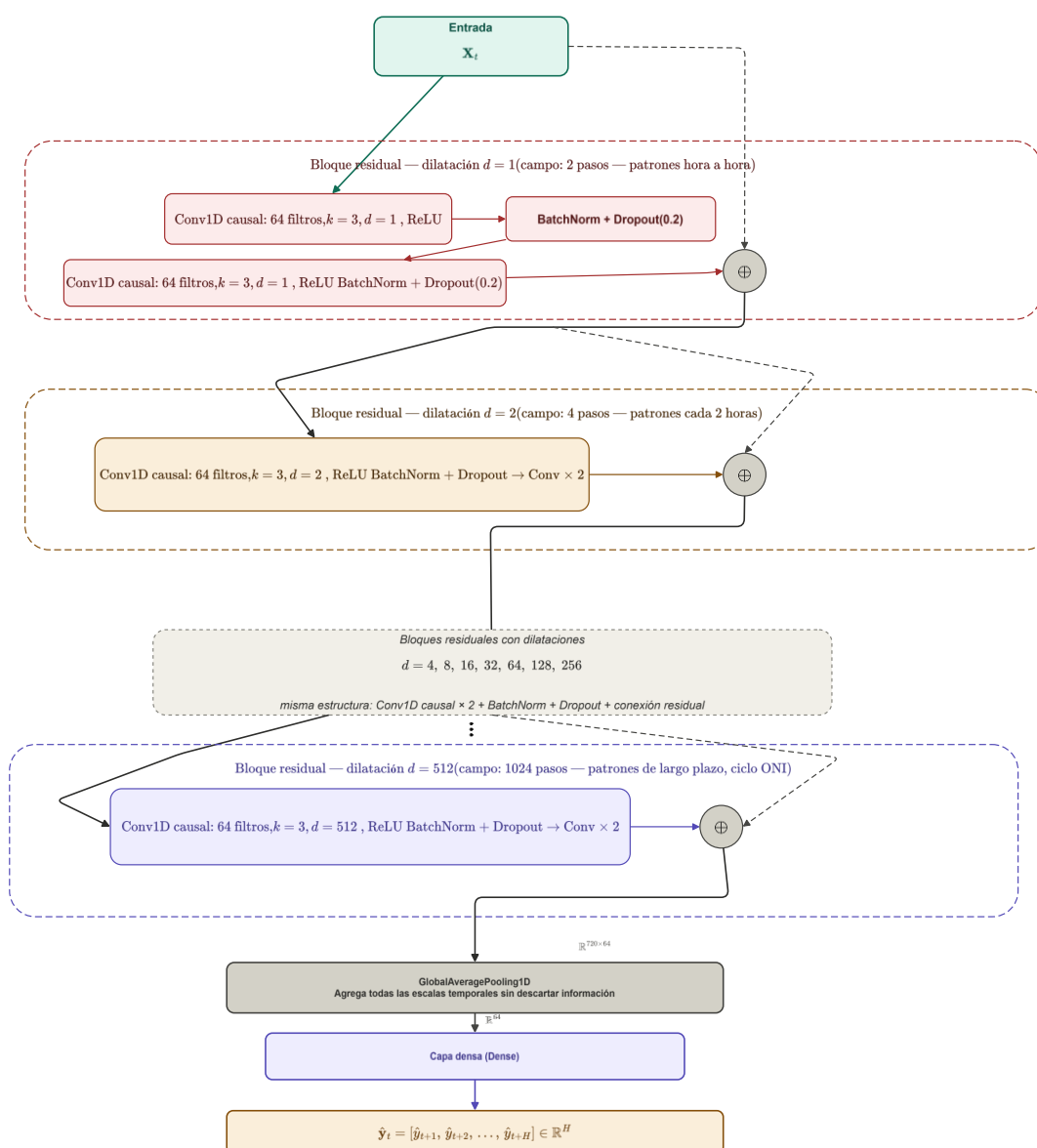


Figura 20. Arquitectura TCN. Fuente: elaboración propia

DeepAR

Este modelo propuesto por Salinas et al. (2020) en *Amazon Research* surge con la intención de resolver problemas de predicción de demanda de productos en e-commerce donde la distribución de las ventas es normalmente asimétrica, tiene colas pesadas y puede variar de acuerdo con la temporada; estas características son las mismas que tenemos en el problema de la predicción del PBN, por lo que resulta una buena opción para este trabajo. Adicionalmente, este modelo tiene la particularidad que no tiene como objetivo una estimación puntual, sino la distribución de una variable objetivo para un horizonte de predicción (en nuestro caso 24 horas) partir de la estimación de sus parámetros de media y varianza; estas predicciones probabilísticas tienen ventajas interpretables para los agentes del mercado de energía colombiano como se explicará en la sección de evaluación. Asimismo, otra característica de este modelo es que es posible utilizar covariables con valores conocidos hasta el instante t (como las variables del SIN y el ONI y sus respectivas transformaciones) así como covariables conocidas en el futuro, por ejemplo, las variables asociadas a hora, día de la semana y mes.

Matemáticamente, lo podríamos definir como:

$$\begin{aligned} H_t &= LSTM(H_{t-1}, [y_{t-1}, X_t, c]) \\ (\mu_t, \sigma_t) &= g_\theta(H_t) \\ y_t &\sim N(\mu_t, \sigma_t) \end{aligned}$$

Donde y_{t-1} es el valor del precio anterior (real para etapa de entrenamiento y el anterior inferido en etapa de inferencia), X_t son las covariables conocidas en t , c es el embedding estático del grupo “colombia”, puesto que solamente estamos trabajando con una serie temporal objetivo, H_t es el estado oculto de la LSTM y $g_\theta(\cdot)$ es una red que proyecta H_t a los parámetros de la distribución. En la Figura 21 se observa la arquitectura de esta modelo utilizada en este trabajo.

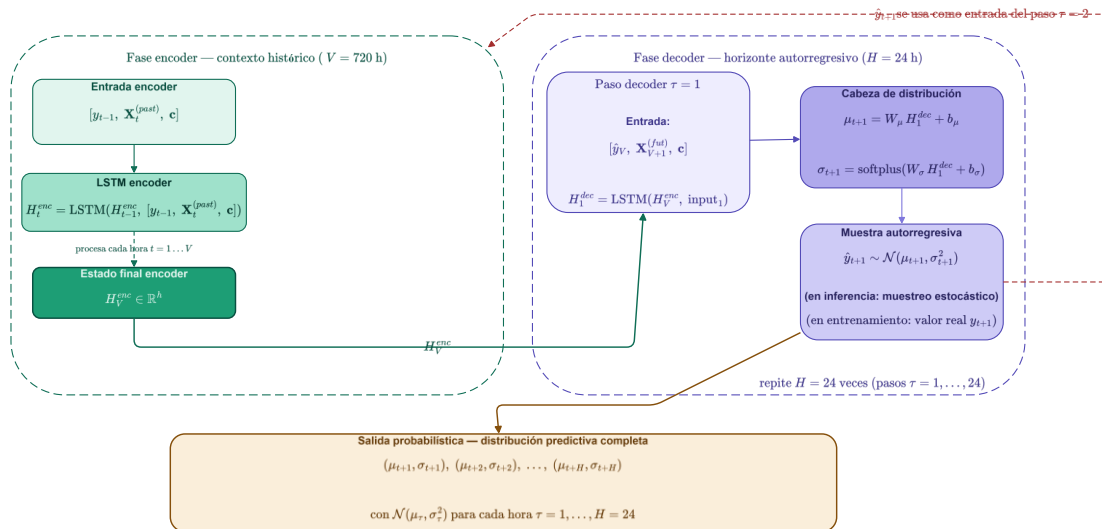


Figura 21. Arquitectura DeepAR. Fuente: elaboración propia basada en Salinas et al. (2020).

Temporal Fusion Transformers (TFT)

Esta arquitectura fue propuesta por Lim et al. (2021) en *Google Cloud AI*, su motivación parte de diseñar un modelo que resuelva el problema de poder distinguir qué covariables son más relevantes en un horizonte de predicción con el propósito de darle más interpretabilidad al modelo, el cual es un aspecto que no se ve en los otros modelos presentados, puesto que las tratan como igualmente informativas. Por otro lado, también plantea una estrategia para ampliar el uso de la arquitectura *Transformer* basada en el mecanismo de autoatención propuesto por Vaswani et al. (2017), que plantea una estrategia para brindar un mayor contexto en problemas especialmente de texto; sin embargo, en ese tipo de problemas es posible observar las palabras como un solo conjunto de datos, sin realizar una partición en la cual especificar qué datos se pueden observar en el pasado y cuáles en el futuro, este problema también se busca resolver con *TFT*.

En la Figura 22 se ilustra la arquitectura presentada en la publicación original, donde observamos que hay 4 mecanismos claves en el funcionamiento de este modelo. En primer lugar, para poder elegir qué variables debemos mantener y comprimir o incluso eliminarlas de acuerdo con su nivel informativo, se utilizan *Gated Residual Network (GRN)* la cual es un bloque cuyo componente principal es la compuerta *Gated Linear Unit* que decide cuánta información de la transformación pasa hacia adelante para posteriormente calcular la importancia de cada variable mediante una función *softmax* que da como resultado *Variable Selection Network (VSN)*. En segundo lugar, el modelo permite realizar un procesamiento de las variables según su tipo como en *DeepAR*, es decir en si son variables estáticas y conocidas o no conocidas en el

59

Predicción probabilística	Monte Carlo - Dropout	Monte Carlo - Dropout	Monte Carlo - Dropout	Monte Carlo - Dropout	Por medio de distribución con parámetros μ, σ	Cuantiles
Supuesto distribucional	Gaussiano (Monte Carlo)	Gaussiano (Monte Carlo)	Gaussiano (Monte Carlo)	Gaussiano (Monte Carlo)	Normal	Ninguno
Covariables futuras conocidas	No	No	No	No	Sí	Sí
Selección de variables interpretable	No	No	No	No	No	Sí (VSN)
Paralelizable en entrenamiento	No	No	Parcial	Sí	No	Sí
Interpretabilidad	Baja	Baja	Baja	Baja	Baja	Alta (VSN + atención)
Complejidad computacional	Baja	Media	Media	Media-alta	Media	Alta

Tabla 7. Resumen comparativo modelos del estudio. Fuente: elaboración propia.

4.2.5 Evaluación

Para este trabajo la elección de la evaluación del rendimiento de los modelos se realizó basado en la posición de los 3 agentes principales del Mercado de Energía Mayorista colombianos: operador (XM), comercializadoras y plantas generadoras. Para este fin, se seleccionan métricas puntuales y probabilísticas que permitan tener una noción de la estimación promedio del precio de la energía y un intervalo de confianza para la gestión del riesgo, la operatividad del mercado y la toma de decisiones principalmente orientada a costos y utilidades, por ejemplo, para elegir comprar un modelo sobre otro, subir o bajar el precio, reservar agua, entre otras opciones. Por ello, se explica cada una de las métricas y su significado y posible uso.

Métricas de estimación puntual

Error Absoluto Medio (Mean Average Error, MAE):

Esta métrica mide el promedio en magnitud de los errores entre el valor puntual estimado por el modelo y el valor verdadero y se calcula de la siguiente manera:

$$MAE = \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t|$$

Siendo T la cantidad de periodos de tiempo que se están prediciendo y y_t e $\hat{y}_t$ el valor real y el estimado por el modelo respectivamente. Esta métrica es especialmente útil, porque es fácilmente interpretable, ya que se encuentra en las mismas unidades de medida que la variable objetivo, en este caso COP/kWh. Adicionalmente, esta métrica no es tan sensible a los valores extremos como el MSE, sino que tiene una representación más equilibrada, lo cual es muy útil para este problema, donde se pueden presentar épocas de El Niño que elevan el precio considerablemente.

Raíz del Error Cuadrático Medio (Root Mean Squared Error, RMSE):

Este indicador mide la raíz cuadrada del promedio de los errores entre el valor puntual estimado por el modelo y el valor real, calculado de la siguiente forma:

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2}$$

Donde $T, y_t, \hat{y}_t$ tienen el mismo significado que en MAE. Con respecto a esta métrica, se observa que es un complemento al MAE, ya que castiga cuadráticamente los errores más grandes, que en el contexto de este problema pueden estar asociado a épocas de fenómeno de El Niño, sin embargo, al sacar raíz cuadrada reduce un poco ese castigo promedio, ya que no deseamos que sea tan grande. Adicionalmente, se espera que en este problema $RMSE > MAE$, debido a la incertidumbre climática que tiene gran repercusión en el mercado de energía colombiano que genera regímenes con valores extremos.

Error Porcentual Absoluto Medio (Mean Average Percentage Error, MAPE):

Esta métrica mide el error porcentual con respecto al valor real promedio. Es calculado de la siguiente manera:

$$MAPE = \frac{100\%}{T} \sum_{t=1}^T \left| \frac{y_t - \hat{y}_t}{y_t} \right|$$

Donde $T, y_t, \hat{y}_t$ tienen el mismo significado que en MAE. Adicionalmente, este valor es muy útil para el contexto del precio de energía en Colombia, ya que como se observó en el análisis exploratorio de los datos, la dinámica está dividida en regímenes que han sido marcados principalmente por fenómenos climáticos, lo que puede indicar que una diferencia de estimación de 100 COP/kWh puede no significar la misma magnitud de error en 2015 que en 2024 debido también a la valorización del precio en el tiempo; y esta métrica al medir los valores porcentualmente es adecuado para tener una perspectiva comparable entre periodos.

Error Porcentual Absoluto Medio Simétrico (Symmetric Mean Absolute Percentage Error, SMAPE):

Este indicador es el promedio de los errores porcentuales con respecto al valor real y al estimado. Se calcula de la siguiente manera:

$$SMAPE = \frac{100\%}{T} \sum_{t=1}^T \frac{|y_t - \hat{y}_t|}{\frac{|y_t| + |\hat{y}_t|}{2}}$$

Donde $T, y_t, \hat{y}_t$ tienen el mismo significado que en MAE. Esta métrica tiene la misma utilidad que el MAPE, sin embargo, busca resolver un problema que tiene el MAPE y es la sensibilidad a grandes contrastes entre el valor real y el estimado, ya que un valor real pequeño y un valor estimado grande puede inflar la cifra y viceversa, por tanto, lo que hace el SMAPE es resolver esa asimetría tomando como denominador el promedio del valor real y el estimado.

Métricas de estimación probabilísticas

Pérdida de Quantil o Pérdida de Pinball (Quantile Loss or Pinball Loss):

Es función que permite evaluar el pronóstico probabilístico a partir de cuantiles por medio de la siguiente regla para un cuantil q :

$$PB_q(y_t, \hat{q}_t) = \begin{cases} q(y_t - \hat{q}_t) & \text{si } y_t \geq \hat{q}_t \\ (1 - q)(y_t - \hat{q}_t) & \text{si } y_t < \hat{q}_t \end{cases}$$

Donde q es el cuantil que estamos calculando su estimación, $\hat{q}_t$ es la estimación del cuantil del instante t e y_t es el valor real en el instante t . Esta función tiene propiedades que son muy útiles dentro del contexto de este trabajo, ya que cuando estamos calculando cuantiles por encima de la mediana, es decir $q > 0.5$ la función castiga las subestimaciones más que las sobreestimaciones, lo cual es muy útil para una comercializadora al gestionar su riesgo, puesto que nos indica que es más caro que un riesgo de exposición a un valor extremo no cubierto se materialice a que quede con más cobertura de la necesaria, aspecto que suele ser común en los análisis de estos agentes. De manera análoga, cuando se tiene cuantiles por debajo de la mediana $q < 0.5$ se castiga con mayor severidad las sobreestimaciones que las subestimaciones, cuyo resultado también es útil para una planta hidroeléctrica, ya que le sirve para decidir cuándo proteger su agua a partir del costo de oportunidad.

Cobertura de Intervalo:

Esta métrica mide la proporción de datos en un periodo de tiempo $t \in [1, T]$ están dentro de un intervalo de confianza, es decir su cálculo es para una confianza del 90% es:

$$Coverage_{90\%} = \frac{1}{T} \sum_{t=1}^T 1(\hat{q}_{0.1,t} \leq y_t \leq \hat{q}_{0.9,t})$$

Donde $\hat{q}_{0.1,t}$, $\hat{q}_{0.9,t}$ son la estimación el cuantil 0.1 y 0.9 respectivamente. Esta métrica es muy útil porque nos permite identificar una banda de confianza en la cual el precio se encuentra y la cantidad de predicciones que se encuentran dentro de este intervalo. Esto es útil, porque si se tiene un buen coverage, XM podría utilizar el valor del intervalo dado por los cuantiles, específicamente el superior, para predecir si el precio va a superar el valor de precio de escasez, aspecto que es importante para ellos como operador y administrador.

Puntaje de interval o Puntaje de Winkler (Interval Score or Winkler Score):

Es un indicador que permite medir la calidad de un intervalo de confianza $1 - \alpha$, de la siguiente manera:

$$IS_{1-\alpha} = (U - L) + \frac{2}{\alpha} \max(L - y_t, 0) + \frac{2}{\alpha} \max(y_t - U, 0)$$

Donde L, U es el valor inferior y superior del intervalo respectivamente. Esta medida es muy útil porque castiga el ancho del intervalo y al mismo tiempo penaliza datos que no se encuentran en el intervalo. Esta medida es muy útil para el mercado de energía colombiano, ya que un intervalo estrecho y preciso puede representar una cobertura adecuada para una comercializadora, mientras que uno ancho, puede representar costos de cobertura innecesarios.

Continuous Ranked Probability Score (Gaussian):

Esta es una métrica que es una generalización del MAE a nivel probabilístico y se describe como:

$$CRPS(F, y_t) = \int_{-\infty}^{\infty} [F(x) - 1(x \geq y_t)]^2 dx$$

Donde F es la distribución de probabilidad, y_t es el valor del precio en el instante t y $1(\cdot)$ es la función indicadora. Esta medida nos indica 2 aspectos importantes de nuestra estimación probabilística: sharpness (qué tan estrecha es la distribución predictiva) y calibración (qué tanto se corresponde la distribución predictiva con la real), es decir permite medir el ancho y la precisión del intervalo de estimación. Dado que la definición es compleja de calcular, se utiliza el cálculo siguiente para distribuciones no paramétricas (Arnold et al., 2023; Gneiting & Raftery, 2007; Laio & Tamea, 2007):

$$CRPS(F, y_t) = \int_0^1 PB_q(F^{-1}(q), y_t) dq$$

Es decir, es la integral sobre todos los quantiles de la función de pérdida pinball, que para este trabajo estimaremos como se presenta en Bracher et al. (2021) donde se deduce que es posible calcularlo como:

$$CRPS \approx WIS$$

Donde:

$$WIS = \frac{1}{K + \frac{1}{2}} \left(w_0 |y_t - \hat{q}_{0.5}| + \sum_{k=1}^K w_k IS_{1-\alpha_k} \right)$$

Con $w_0 = 1/2$, $IS_{1-\alpha_k}$ es el puntaje de intervalo definido anteriormente con un nivel de significancia de α_k y K es el número de intervalos de predicción centrales, para este trabajo utilizaremos los cuantiles $\hat{q}_{0.1}, \hat{q}_{0.25}, \hat{q}_{0.5}, \hat{q}_{0.75}, \hat{q}_{0.9}$, por lo que $K = 2$, esta aproximación es útil para el modelo TFT, debido a su estimación de cuantiles.

Por otro lado, para modelos como el DeepAR, donde la distribución es gaussiana, podemos aplicar la fórmula mencionada por Gneiting et al. (2005) de la siguiente manera:

$$CRPS(N(\mu, \sigma^2), y_t) = \sigma \left[z(2\Phi(z) - 1) + 2\phi(z) - \frac{1}{\sqrt{\pi}} \right]$$

Donde $z = \frac{y_t - \mu}{\sigma}$ y μ, σ es la media y desviación estándar respectivamente de las predicciones realizadas con simulaciones Monte Carlo, $\Phi(\cdot)$ es la distribución de probabilidad acumulada de una normal estándar y $\phi(\cdot)$ su función de densidad de probabilidad.

Otra métrica que vale la pena calcular es una que permita medir la efectividad de los modelos en episodios extremos del fenómeno de El Niño, ya que tiene gran relevancia para los grupos de interés del mercado de energía colombiano, para ello, lo que se realiza es calcular un ratio de valores atípicos de la siguiente manera, primero calculamos el percentil 90 de los precios reales para tener un umbral que nos permita definir valores extremos, para nuestro caso fue de 297 COP/kWh, luego separamos los datos en 2 grupos y calculamos el MAE sobre cada uno y realizamos la siguiente división:

$$spikes\ ratio = \frac{MAE\ spikes}{MAE\ normal}$$

4.2.6 Despliegue / uso

En esta última etapa se plantea posibles estrategias para realizar el despliegue del mejor modelo y su mantenimiento y uso a lo largo del tiempo, para ello es importante abordarlo desde 3 perspectivas: ¿Cómo desplegar este modelo e integrarlo?, medición constante del rendimiento de negocio y rendimiento y recalibración del modelo.

¿Cómo desplegar este modelo e integrarlo?

Para este caso es posible tener una arquitectura de datos en nube basada en grandes volúmenes de datos que permita tener como punto de partida la ingesta de datos de XM y ONI por medio de servicios como Kinesis Data Streams o Kinesis Firehouse en *Amazon Web Services (AWS)* y almacene la información en un componente de persistencia como puede ser una base de datos relacional o no relacional que luego permita acceder a la información para recalibrar. Adicionalmente, tener en cuenta que la predicción debe realizarse antes de las 8 am si se quisiera utilizar en XM, puesto que es el momento donde se publican los precios.

Medición constante del rendimiento de negocio

Esto está asociado a que el modelo siempre tendrá asociado unos costos por su uso y alojamiento en caso de que esté en nube, por lo que es importante estar monitoreando que el ROI esté de esta herramienta tenga sentido y esté dentro de los objetivos estratégicos de la empresa.

Medición de rendimiento y recalibración del modelo

Es importante para aquellos agentes que quieran utilizar el modelo, que ante episodios críticos marcados por métricas deterioradas notablemente por fenómenos o acontecimientos atípicos el modelo se busque dejar de usar mientras se equilibra la situación respectiva. También es posible buscar modificar el modelo para adaptarse a ese periodo, pero será una opción más costosa.

Adicionalmente, es importante que haya una periodicidad fija de reajuste del modelo o reentrenamiento, podría ser de 2 a 3 meses debido a las relaciones identificadas en el análisis

exploratorio de los datos para que el modelo pueda estar actualizado a los patrones más recientes.

4.3 Recursos requeridos

Para la ejecución de este proyecto se requirieron los siguientes recursos:

Recursos técnicos

Fuentes de datos: API XM e índice mensual ONI.

Entorno de Google Colab más reciente a la fecha 13 de abril de 2026 con GPU A100 y memoria RAM ampliada (Google Colab pro + 200 unidades de procesamiento).

Dispositivos

Este trabajo fue realizado en un computador ACER modelo Nitro AN515-54 con procesador Intel Core i5-9300H CPU @ 2.40GHz con 1TB de disco HDD, 256 GB de SSD y 16 GB RAM, el cual fue usado para la investigación de papers, uso de Google Colab y redacción del paper.

Asistencia de expertos

Reuniones de asesoramiento con el tutor en las cuales se recibía retroalimentación de este a la vez que orientaba y enseñaba sobre el tema. Este acompañamiento fue muy importante para la elaboración de este trabajo.

4.4 Presupuesto

Tipo de coste	Valor	Comentarios
Horas de trabajo en el proyecto	117 horas * 10 €/hora = 1170 €	Tiempo dedicado al desarrollo del trabajo: lectura de papers, planificación de solución, implementación de la solución, redacción de documento y reuniones con el tutor.

Equipo técnico utilizado	800 €	Valor aproximado de computador ACER modelo Nitro AN515-54 con procesador Intel Core i5-9300H CPU @ 2.40GHz con 1TB de disco HDD, 256 GB de SSD y 16 GB RAM
Software utilizado	29,97 €	Google Colab pro + 200 unidades de procesamiento
Estudios e informes	0 €	No se ha tenido que adquirir estudios ni informes
Materiales empleados	0 €	No se han utilizado materiales para este trabajo más allá de lo especificado de Google Colab en Software utilizado
Total	2000 €	Costo total del trabajo

4.5 Viabilidad

La viabilidad de este proyecto puede analizarse desde 3 perspectivas: coste, beneficios y sostenibilidad del proyecto en el futuro.

En primera instancia, este proyecto no representa un gran coste más allá de la investigación y la implementación y su entrenamiento inicial en alguna máquina en la nube que permita tener hardware adecuado. Sin embargo, si se quiere plantear todo un sistema integrado de extracción de la información, persistencia y recalibración diaria del modelo seleccionado, los costes suben, pero aun siendo manejables.

En segundo lugar, los beneficios que puede traer un proyecto como este son muy altos y se multiplican si es integrado a un sistema que utilice algún agente del Mercado de Energía Mayorista colombiano, puesto que puede ser una herramienta muy útil para la toma de decisiones de XM, comercializadoras para temas de cobertura y para generadoras para proteger y usar sus recursos eficientemente. De esta manera se logra incentivar inversionistas en este mercado con energía renovable dominante.

Finalmente, este proyecto debido a que es implementado con tecnologías de vanguardia como el lenguaje de programación Python y librerías que son muy utilizadas y fáciles de actualizar en caso de ser necesario, permiten que esta implementación sea mantenible en el tiempo y fácilmente escalable.

4.6 Resultados del proyecto

Para este apartado, se presentan los resultados de la metodología CRISP-DM propuesta en secciones anteriores, donde observaremos el rendimiento de los modelos desde los diferentes puntos de vista que han sido considerados importantes a lo largo de este proyecto: *tradeoff* entre la predicción puntual y la probabilística, puesto que esta última es realmente útil para los agentes del Mercado de Energía Mayorista colombiano, esto implica tener un modelo preciso, pero con intervalos lo suficientemente bien calibrados en cuanto a cobertura y ancho.

En la Tabla 8 se presentan los resultados de las métricas de evaluación sobre el conjunto de pruebas planteadas en la fase de evaluación.

Modelo	MAE	R^2	CRPS	Cov90%	Sesgo
TFT	7.141	0.975	5.122	0.736	0.2
DeepAR	10.148	0.951	13.579	0.951	-1.2
RNN	12.916	0.915	11.569	0.630	-8.2
TCN	30.556	0.635	37.404	0.028	29.4
LSTM	55.155	0.068	50.720	0.003	56.7
CNN-LSTM	56.824	-0.052	55.515	0.004	55.1

Tabla 8. Resultado métricas de evaluación sobre el conjunto de pruebas. Fuente: elaboración propia.

En primera instancia, se puede observar que a nivel de métricas puntuales y probabilísticas el *TFT* y el *DeepAR* fueron los que tuvieron el mejor rendimiento mostrando que arquitecturas modernas con una estructura especializada en series de tiempo tienden a tener un mejor rendimiento para este problema. Sin embargo, es interesante hacer hincapié en los resultados de *RNN* que fue el que tuvo el mejor rendimiento de los modelos keras, porque muestra que la implementación realizada en este trabajo de tener una ventana de 168 horas para *RNN* lo convierte en un modelo robusto en la predicción de precios de energía en el mercado colombiano.

Andrés Grimaldos Echavarría

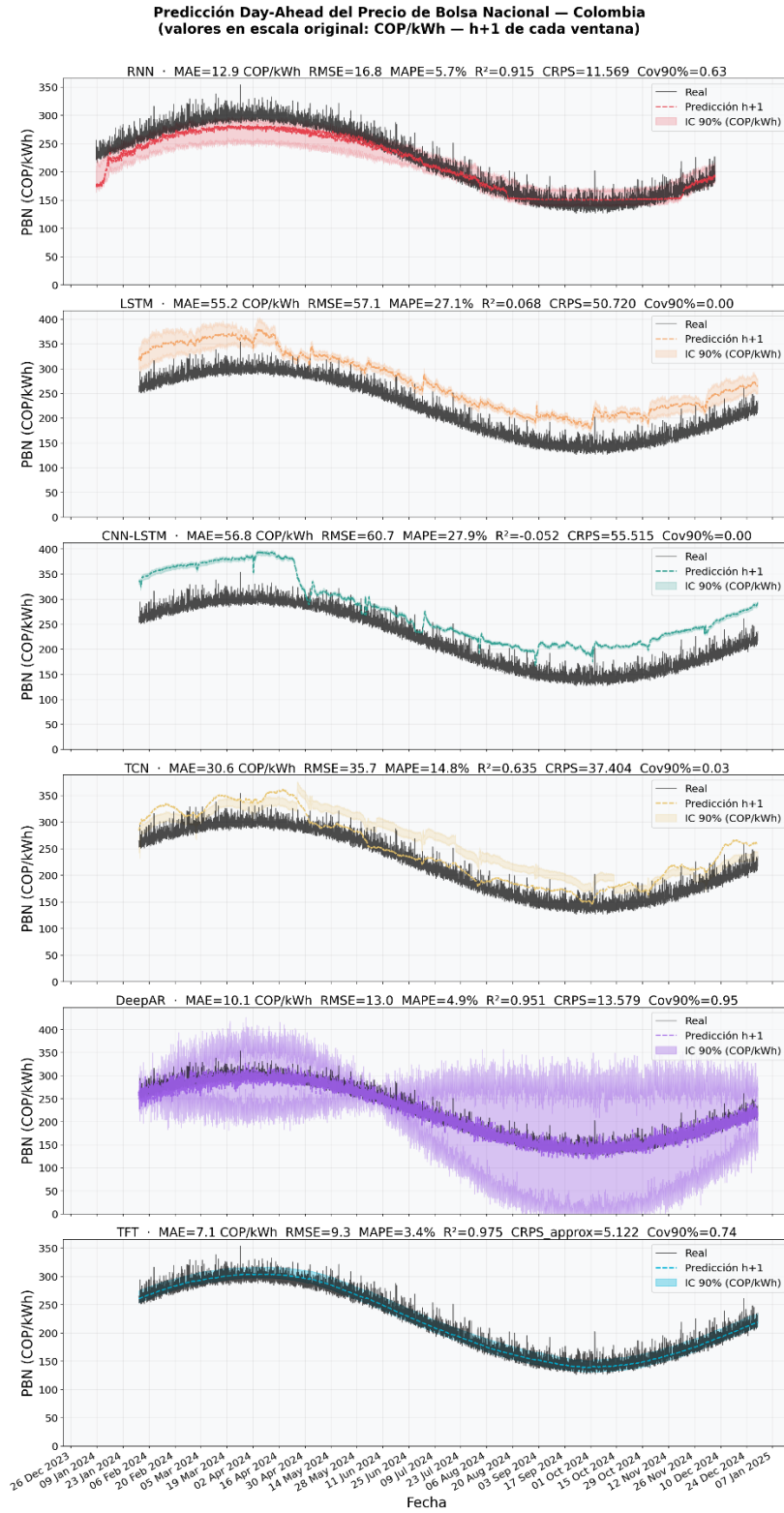


Figura 23. Comparación de predicciones y valor real. Fuente: elaboración propia con datos de Equipo Analítica XM (s. f.)

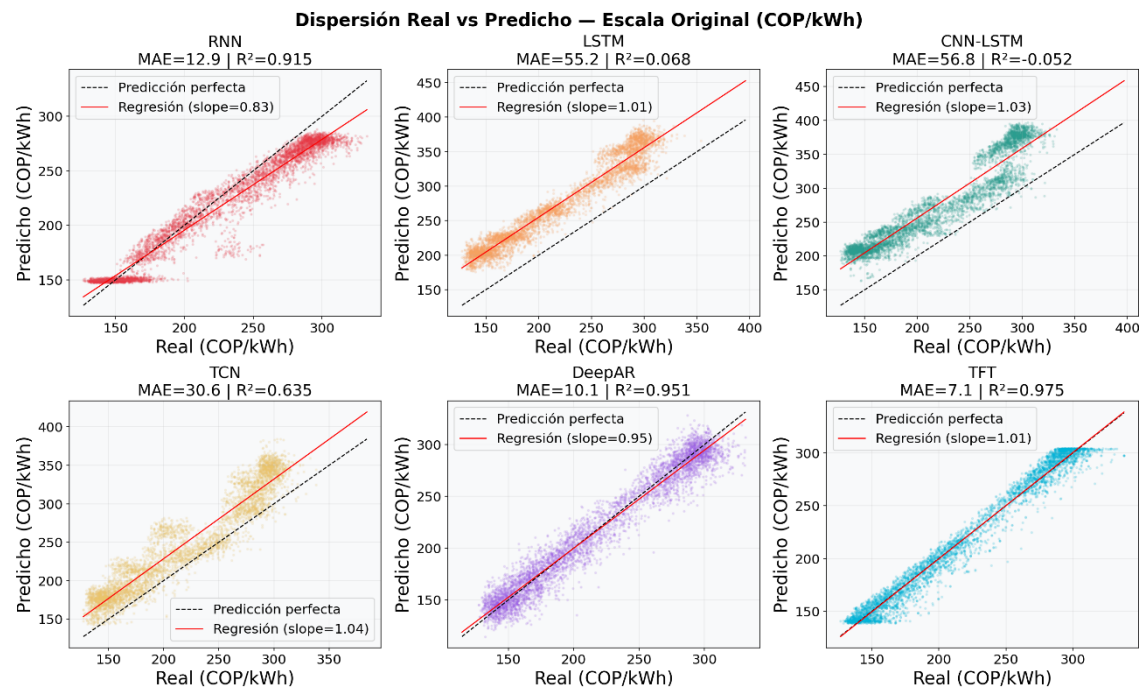


Figura 24. Regresión de predicción y valor real. Fuente: elaboración propia

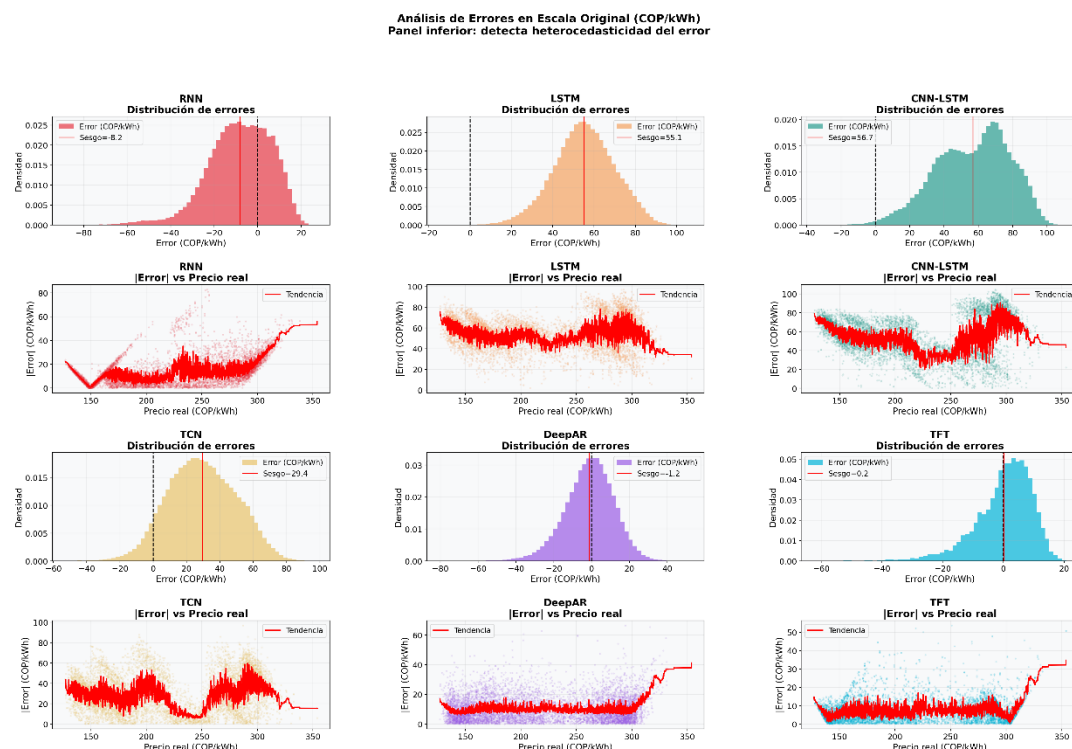


Figura 25. Distribución de los errores de estimación. Fuente: elaboración propia

Adicionalmente, en las Figuras Figura 23, Figura 24 y Figura 25 se presentan las gráficas de los resultados asociados al ajuste de la serie de precios de energía reales en el conjunto de pruebas donde se tiene la comparación de serie real con la serie estimada y los intervalos de confianza, la regresión lineal entre el valor estimado y el real y la distribución asociada a los errores de estimación para cada uno de los modelos.

A partir de lo anterior, se observa que los modelos *LSTM* y *CNN-LSTM* presentan un sesgo positivo sistemático de +55,1 y +56,7 COP/kWh respectivamente, con un coeficiente de determinación R^2 de 0,068 y -0,052 y una cobertura del intervalo de confianza al 90% cercana a cero. Este comportamiento no refleja una incapacidad del modelo para aprender la dinámica del precio, sino una falla en la etapa de desnormalización de la capa RevIN. Como evidencia de ello, la pendiente de regresión en el diagrama de dispersión es cercana a 1,0 (1,01 y 1,03 respectivamente), lo que indica que ambos modelos capturan correctamente la variación relativa del Precio de Bolsa Nacional, pero introducen un desplazamiento constante en la escala original. Por esta razón, los resultados de LSTM y CNN-LSTM deben interpretarse con cautela y no son directamente comparables con los demás modelos en términos de métricas absolutas.

Con respecto al *TCN* se observa que en periodos con valores altos de precios tiende a tener un rendimiento más bajo que en los periodos sin tanta volatilidad, lo que indica que el modelo puede ser bueno para temporadas normales en términos de variabilidad, pero puede no capturar el comportamiento correcto para estas situaciones extremas como el fenómeno de El Niño, esto también se evidencia en la Figura donde se muestra una ratio mayor a 1. Asimismo, también se observa que los intervalos construidos por este modelo no son de buena calidad, puesto que su coverage es muy bajo, puesto que solo captura un 3% de las veces, muy lejos del deseado 90%.

Por otro lado, el modelo *RNN*, como se mencionó fue el mejor modelo de los implementados con Keras, puesto que se observa que, sin embargo, subestima valores en épocas extremas, lo cual se vio en su ratio de spikes, capturó el comportamiento mejor que los demás modelos de Keras y sus intervalos dieron como resultado un coverage de 0.63, por lo que los valores obtenidos a partir de la capa Montecarlo Dropout dan una calibración decente de intervalos.

En adición, el modelo *DeepAR* que posee una arquitectura más moderna, logra obtener un resultado bastante bueno en valores puntuales y también en probabilísticos, no obstante, esto último hay que revisarlo, porque si bien su coverage es de 0.95, se observa en la Figura 23 que sus intervalos son muy anchos, por lo que tienen una buena precisión, pero una amplitud que podría ser alta para algunos agentes del mercado, puesto que puede dificultar la gestión de riesgos con sobrecoberturas, por ejemplo. Además, este modelo tuvo un buen rendimiento en

Andrés Grimaldos Echavarría

la ratio de spikes con un valor de 1.3x, el cual es un valor tolerable a comparación de por ejemplo *RNN*, pero que aún indica que falta capturar más patrones para modelar épocas con climatología extrema.

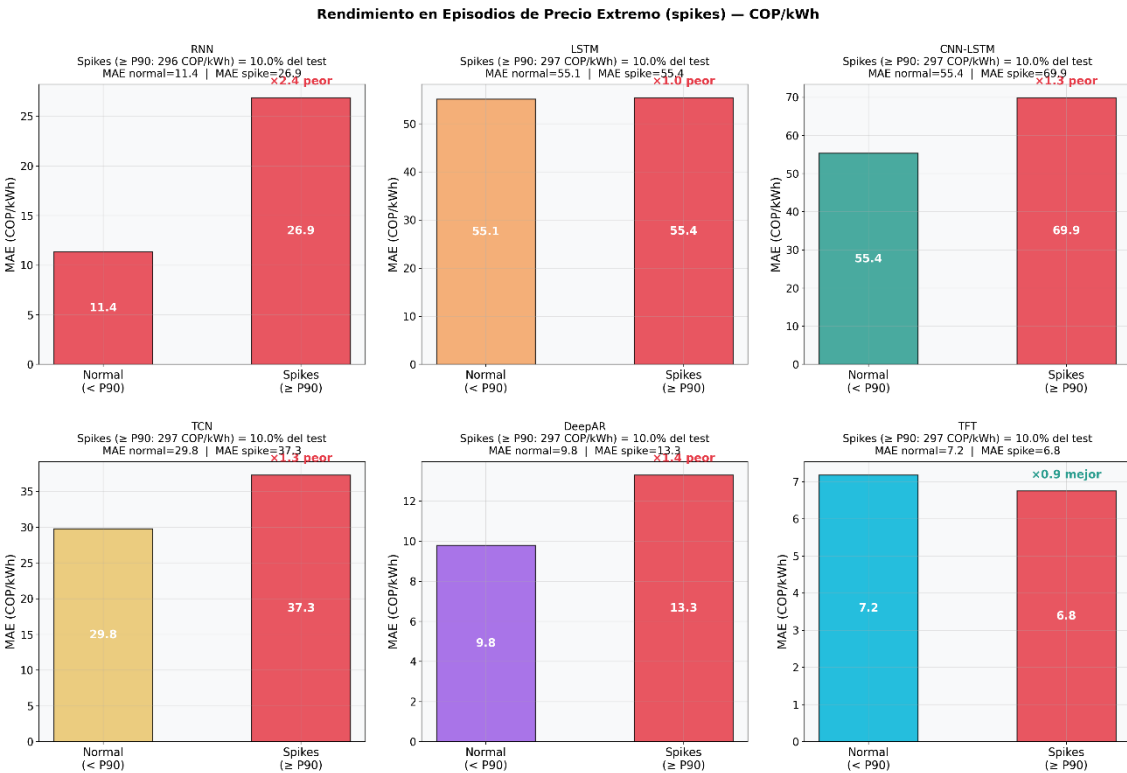


Figura 26. Ratio MAE spike / MAE normal. Fuente: elaboración propia.

Finalmente, tenemos los resultados del modelo *TFT*, el cual sin lugar a duda es el que tuvo el mejor rendimiento en todas las métricas y tiene el valor agregado de brindar interpretabilidad para el periodo de 2023 a 2025 en el mercado de energía colombiano. En términos probabilísticos, el TFT alcanza un CRPS aproximado de 5.1 y una cobertura del intervalo de confianza al 90% de 0.74, que representa el mejor equilibrio entre precisión y calibración del conjunto de modelos evaluados.

El análisis de la serie temporal de predicciones en escala original (COP/kWh) revela que el TFT logra seguir con alta fidelidad la tendencia de largo plazo del PBN colombiano durante todo el período de evaluación. La gráfica muestra cómo el modelo captura correctamente el ciclo hidrológico anual: los precios elevados de principios de 2024 (aproximadamente 300 COP/kWh durante la temporada seca de enero a abril), la pronunciada caída hacia mediados de año

Andrés Grimaldos Echavarría

(mínimos cercanos a 150 COP/kWh en junio–julio 2024 durante la temporada lluviosa del primer semestre) y la recuperación gradual hacia finales del período. Adicionalmente, es el modelo con el resultado más diferenciador de este trabajo con respecto al ratio de spikes, lo que quiere decir que su mecanismo sofisticado de selección de variables junto con el aprendizaje del contexto mediante atención permite capturar de una manera positiva el comportamiento en épocas extremas climáticas como el fenómeno de El Niño o La Niña.

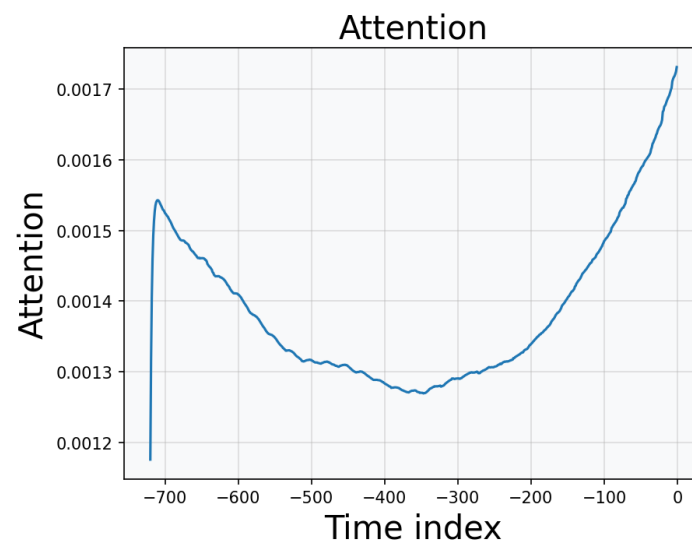


Figura 27. Comportamiento de atención. Fuente: librería *pytorch-forecasting* donde en el eje x tenemos el tiempo y en el eje y el valor numérico de atención

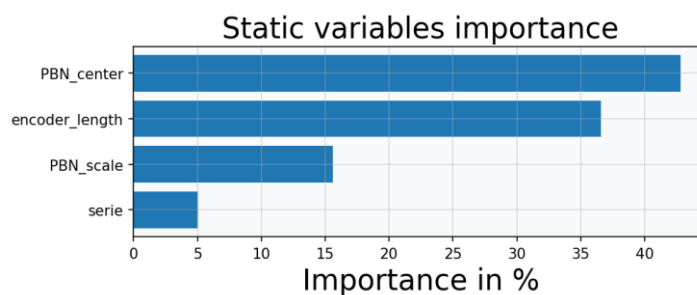


Figura 28. Importancia de las variables estáticas. Fuente: librería *pytorch-forecasting* donde el eje x muestra el porcentaje de atención y el eje y muestra las variables estáticas.

Andrés Grimaldos Echavarría

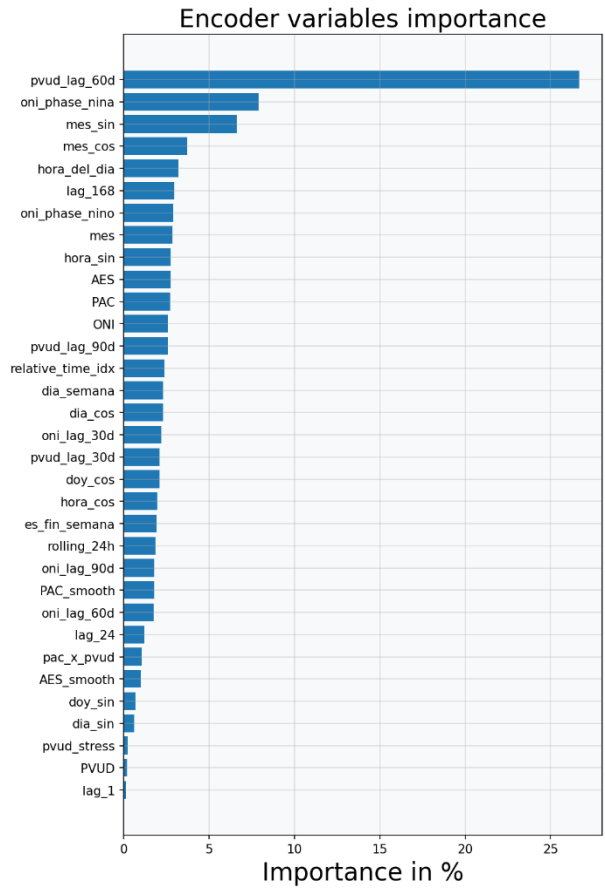


Figura 29. Importancia de las variables en la etapa encoder. Fuente: librería pytorch-forecasting donde el eje x muestra el porcentaje de importancia en la fase de encoder y el eje y las variables

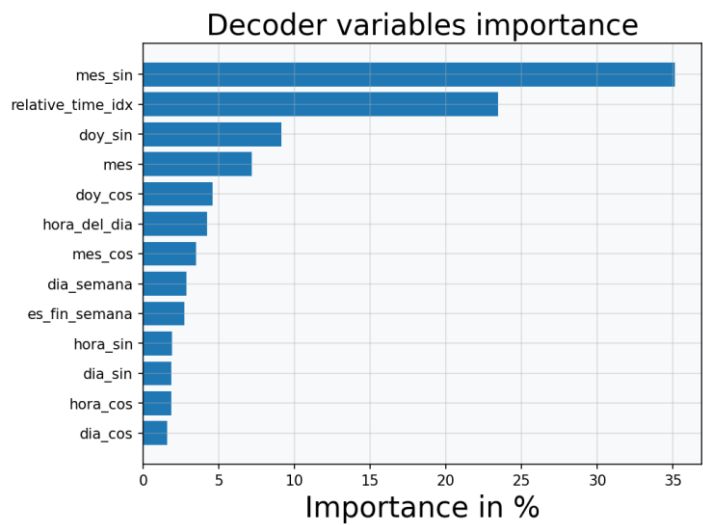


Figura 30. Importancia de las variables en la fase decoder. Fuente: librería pytorch-forecasting donde el eje x muestra el porcentaje de importancia y el eje y las variables.

Una de las contribuciones más valiosas del TFT para el análisis del mercado eléctrico colombiano es la interpretabilidad de su mecanismo de selección de variables mediante la Variable Selection Network (VSN) que se observa en las Figuras 28, 29, 30 y 30. El análisis de importancia de las variables del encoder en la Figura 29 revela que el rezago de 60 días del Porcentaje de Volumen Útil Diario (`pvud_lag_60d`) es la variable más informativa para la predicción del PBN, con un peso relativo del 26%. Este hallazgo valida empíricamente uno de los resultados más significativos del Análisis Exploratorio de Datos: la Función de Correlación Cruzada entre PVUD y PBN mostró que la correlación no es máxima en el rezago inmediato, sino que crece continuamente hasta alcanzar su valor máximo alrededor de los 89 días ($r = -0,29$), reflejando la inercia física del vaciado gradual del sistema de embalses colombiano. El hecho de que el modelo aprenda automáticamente, sin intervención manual, que el estado de los embalses hace dos meses es más informativo que el nivel actual constituye una validación cuantitativa de la hipótesis hidrológica central de ese trabajo. La segunda variable más importante es la fase climática La Niña (`oni_phase_nina`, 8%), seguida de las variables de estacionalidad mensual (`mes_sin`, 6%, y `mes_cos`, 4%). La presencia de `oni_phase_nina` como segunda variable más relevante confirma que el modelo identifica el régimen hídrico favorable como el principal moderador de los precios extremos en el mercado colombiano, coherente con el análisis de co-ocurrencia de spikes y fases del Índice Oceánico de El Niño realizado en el EDA.

El análisis de importancia de las variables del decoder de la Figura 30 revela que la componente sinusoidal del mes (`mes_sin`) es la covariable futura más relevante con un peso del 35%, seguida del índice temporal relativo (`relative_time_idx`, 23%) y la componente sinusoidal del día del año (`doy_sin`, 9%). La dominancia de las variables de estacionalidad mensual en el decoder es coherente con la estructura del ciclo hidrológico colombiano, que alterna entre temporadas secas (diciembre–marzo y julio–agosto) y lluviosas (abril–junio y septiembre–noviembre) con patrones bien definidos. El modelo aprende que conocer en qué punto del ciclo anual se encuentra el horizonte de predicción es el insumo más valioso para estructurar la distribución del precio futuro, lo que justifica la inclusión de covariables de calendario como variables futuras conocidas en el diseño del experimento.

El patrón de atención temporal de la Figura 27 exhibe una forma en U característica con dos regiones de alta atención: el pasado muy reciente (últimas horas, con peso máximo de aproximadamente 0,00173) y el pasado de largo alcance (25–30 días atrás, con pesos superiores a 0,00150), y una región de menor atención en el período intermedio de 2–3 semanas (pesos mínimos de aproximadamente 0,00127 alrededor de los 300–400 pasos hacia atrás). Este patrón es físicamente interpretable: el precio de mañana depende del precio de hoy a través de la fuerte autocorrelación de corto plazo identificada en el ACF del EDA, y del estado de los embalses hace aproximadamente un mes a través del retardo hidrológico del ciclo de llenado y vaciado. El período intermedio de 2–3 semanas es menos informativo porque corresponde a

una escala temporal donde ni la autocorrelación inmediata del precio ni los efectos diferidos del ciclo hídrico tienen máxima expresión.

Finalmente, las variables estáticas en la Figura 28 muestran que PBN_center (42%) y encoder_length (36%) dominan sobre el embedding del grupo serie (5%). Estos dos primeros elementos son metavariables automáticas generadas por pytorch-forecasting cuando se activan las opciones `add_target_scales=True` y `add_encoder_length=True` respectivamente: PBN_center representa la media del PBN en la ventana de escalado, y encoder_length codifica la longitud efectiva del encoder. Su alta importancia relativa indica que el modelo usa activamente el nivel medio del precio en el contexto histórico para condicionar sus predicciones, lo que es equivalente funcionalmente al mecanismo RevIN implementado en los modelos Keras. El bajo peso del embedding serie (5%) es esperado dado que el experimento modela el mercado colombiano como una única serie temporal, situación en la que el embedding estático no puede aportar información diferenciadora entre grupos. Esta limitación podría superarse en trabajo futuro modelando múltiples series paralelas (por ejemplo, por zona del SIN o por tipo de recurso de generación), escenario en el que el embedding estático del TFT capturaría las características estructurales de cada subsistema.

Capítulo 5. DISCUSIÓN

Una vez realizada el análisis del problema de estimación del PBN en el day-ahead, la experimentación con múltiples modelos de aprendizaje profundo y la evaluación del rendimiento de estos es oportuno comentar algunos aspectos interesantes alrededor de este trabajo en sus resultados.

En primer lugar, es interesante mencionar el rendimiento superior de los modelos probabilísticos *TFT* y *DeepAR* frente a los modelos deterministas con estimación por Monte Carlo Dropout (*RNN*, *LSTM*, *CNN-LSTM* y *TCN*) en el problema abordado en este trabajo, lo cual también es un beneficio aprovechable por el negocio, ya que tener predicciones con intervalos de confianza bien calibrados y de calidad permite a XM identificar periodos con alta probabilidad de alcanzar precios de escasez, lo cual implica un golpe al mercado que debe administrar y lo podría hacer con mayor antelación. De manera similar, estos intervalos son muy útiles para las comercializadoras al momento de realizar contratos bilaterales y compromisos, puesto que con estas bandas pueden establecer márgenes necesarios de cobertura o proyecciones de rendimiento sobre los negocios pactados. Finalmente, las plantas generadoras también pueden verse muy beneficiadas de esto, debido a que tener una banda inferior muy bien estimada le permite tomar decisiones críticas operativas de cuándo ofrecer su agua y cuánto cobrar por ella.

En segundo lugar, otro aspecto que vale la pena hacer hincapié es en la concordancia alta de resultados encontrados en el análisis exploratorio de los datos y en la interpretabilidad resultante del modelo *TFT*, porque se observó en la gráfica del comportamiento de atención de la Figura 27 que la forma en “U” representaba que la atención desarrollada por el modelo era más alta en rezagos cortos o en rezagos largos, mientras que en el intermedio no era necesario, su atención era más baja. Esto es muy interesante, ya que sin tener que realizar una intervención manual, el modelo fue capaz de capturar ese comportamiento, esto también nos indica otra utilidad de este modelo y son sus pesos, puesto que la coherencia que logra puede extrapolarse a un método de selección de características para otras aproximaciones. Adicionalmente, es necesario mencionar que si bien el coverage del modelo es alto, aún podría trabajar para calibrarlo más y lograr acercarse al 90%.

En tercer lugar, el modelo *DeepAR* obtuvo un resultado paradójico que se debe discutir, ya que fue el modelo con el segundo mejor MAE, pero al momento de evaluar los intervalos de confianza fueron muy sobrestimados, lo que se ve en las marcadas bandas púrpuras de la Figura 23, la causa más plausible de este comportamiento es que este modelo debe estar limitado a la distribución del precio y no tiene la capacidad de incorporar las variables exógenas de una

manera segregada como lo hace *TFT*, lo que provoca que esas bandas de confianza más anchas sea debido a esa información que no es capaz de capturar el modelo.

En cuarto lugar, uno de los aspectos más relevantes en este trabajo era el análisis de rendimiento en épocas de climas extremos como periodo de el fenómeno de El Niño y La Niña. Este aspecto tuvo un resultado diferenciador, puesto que esta métrica de la ratio sobre los valores extremos busca capturar este rendimiento que el MAE global no es capaz de capturar. En este caso, el único modelo que tuvo un comportamiento mejor en episodios extremos fue el *TFT* lo que nos indica que el mecanismo de atención temporal es fundamental para predecir el precio en esas circunstancias; esto es muy importante porque, por ejemplo, para un comercializador épocas de El Niño o La Niña son los momentos con mayor riesgo financiero, por lo que subestimar puede acarrear en pérdidas notables.

En cuarto lugar, es necesario comentar lo relacionado con el rendimiento superior de *RNN* dentro de los modelos implementados con la librería Keras de Python, indicando que una ventana de horizonte de 168 horas es un buen parámetro para la construcción de este modelo para la solución del problema objetivo de este trabajo. Sin embargo, este resultado debe tener un análisis más profundo, porque los resultados de los modelos *LSTM* y *CNN-LSTM* no necesariamente son inadecuados para este contexto, sino que su rendimiento inferior podría estar atribuido a la metodología de implementación de estos, en particular a la adición de la capa extra *RevIN* que no produjo las ventajas ni el comportamiento que se esperaba, también podría deberse a como se realizan los cálculos computacionales en Keras a diferencia de otras librerías de Python como PyTorch. Esto representa una limitación comparativa de estos modelos contra los demás, una posible aproximación podría ser migrar la implementación a PyTorch y repetir la experimentación de estos modelos.

Finalmente, este trabajo realiza una contribución a la literatura asociada a las aproximaciones para estimar el precio de la energía en Colombia en el día siguiente por medio de una nueva comparativa utilizando modelos de aprendizaje profundo y arquitecturas modernas como *DeepAR* y *TFT*, lo cual amplía el panorama de estudio, ya que surgen múltiples líneas futuras que se mencionan en el siguiente capítulo. Además, hay que mencionar que este estudio está limitado a un conjunto de pruebas que va hasta enero de 2025, sería bastante bueno ampliar este periodo del tiempo.

Capítulo 6. CONCLUSIONES

6.1 Conclusiones del trabajo

Este trabajo tiene como objetivo la experimentación en la estimación del day-ahead del precio de la energía en Colombia por medio de modelos de *deep learning* para mejorar la toma de decisiones de los agentes económicos de interés en este mercado para fomentarlo. Esto se considera que fue un objetivo cumplido, ya que se implementaron 6 arquitecturas de aprendizaje profundo que fueron comparadas entre sí con rigurosas métricas en la estimación del precio de la energía tanto de manera puntual como probabilística.

Estos resultados obtenidos permiten tener estadísticas y métricas que son interpretables para el contexto del Mercado de Energía Mayorista colombiano, porque brindan información útil que como se mencionaba en el capítulo de discusión, pueden ser utilizadas por XM, comercializadoras y plantas generadoras para su toma de decisión. En adición, lo pueden hacer por medio de modelos que también brindan información interpretable como el *TFT*.

Finalmente, a partir del análisis de resultados realizado se observa que hay evidencia significativa de que las relaciones climáticas realmente tienen un impacto notable sobre el precio de la energía debido a la fuerte dominancia hídrica como fuente, lo que también evidencia lo retador que es para los grupos de interés de este mercado operar y obtener rentabilidad esperada; sin embargo, con la posibilidad de tener métodos predictivos que le den confianza para la toma de decisiones, se fomenta que mercados basados en energías renovables como este, sean más llamativos y menos riesgosos.

6.2 Conclusiones personales

La implementación de este proyecto fue una experiencia enriquecedora y retadora, porque tiene como propósito aplicar métodos no vistos durante el máster de aprendizaje profundo a un problema con una gran complejidad debido a la incertidumbre provocada por fenómenos climáticos en los precios de energía del mercado colombiano. Adicionalmente, fue también desafiante entender la contextualización y los conceptos de negocio asociados a los mercados de energía, en particular al colombiano, pero fue una situación que se facilitó gracias al apoyo de mi tutor.

Asimismo, todos los aprendizajes de planificación, técnicos y de negocio me los llevo para mi crecimiento profesional y personal que desde este momento ya están siendo aplicados en la práctica, siendo trascendentales para mí. Este proyecto tiene una gran importancia para mí, ya que profundicé en temas que siempre había querido estudiar y se pudo dar la oportunidad en este trabajo.

Finalmente, fue retador pero interesante leer sobre las publicaciones de otras personas para entender los métodos de vanguardia que se utilizan hoy en día en problemas complejos de predicción en series temporales.

Capítulo 7. FUTURAS LÍNEAS DE TRABAJO

Este trabajo tuvo como foco principal la comparación de modelos de aprendizaje profundo en la predicción del day-ahead de los precios de energía en el mercado spot colombiano, sin embargo, hay que tener en cuenta que esos modelos son solo un subconjunto de un conjunto con un gran número de modelos que van creciendo día a día debido al auge de la inteligencia artificial, por lo que sería muy interesante continuar atacando este problema con las nuevas arquitecturas que se vayan proponiendo y contrastarlas contra estos modelos utilizados en este trabajo.

Adicionalmente, este trabajo no tenía en su alcance modelos pre-entrenados, aspecto que podría ser interesante de aplicar en este problema, utilizando modelos como TimesFM o Chronos y utilizar aproximaciones de *zero shot vs zero shot + fine tuning* para comprobar el rendimiento de *transfer learning* en esta problemática y comparar con las mismas métricas de evaluación utilizadas en este trabajo.

Por otro lado, este trabajo se limitó a un conjunto de variables cuya disponibilidad era del 100% en la API de XM, no obstante, esta fuente de datos contiene muchas otras métricas que vale la pena analizar si pudieran brindar información para la abstracción de más patrones asociados al mercado de energía colombiano. Asimismo, vale la pena investigar otras fuentes de información no convencionales que podrían aportar conocimiento útil a los modelos, por ejemplo, noticias.

Finalmente, este trabajo se limitó a realizarse en Google Colab, pero también vale la pena implementar y mejorar el diseño de una arquitectura completa en nube que permita la extracción de las fuentes de información y recalibrar constantemente el modelo utilizado para la predicción de precios de energía, lo cual podría ser muy útil en un sistema para alguno de los agentes involucrados en el Mercado de Energía Mayorista colombiano.

Capítulo 8. REFERENCIAS

- Arnold, S., Walz, E. M., Ziegel, J., & Gneiting, T. (2023). Decompositions of the mean continuous ranked probability score. *Electronic Journal of Statistics*, 18(2), 4992-5044. <https://doi.org/10.1214/24-EJS2316>
- Bracher, J., Ray, E. L., Gneiting, T., & Reich, N. G. (2021). Evaluating epidemic forecasts in an interval format. *PLOS Computational Biology*, 17(2), e1008618. <https://doi.org/10.1371/JOURNAL.PCBI.1008618>
- CREG. (2023). *METODOLOGÍA PARA LA DETERMINACIÓN DEL PRECIO DE BOLSA PROPUESTAS DE AJUSTE TEMPORAL*. https://gestornormativo.creg.gov.co/gestor/entorno/docs/originales/Proyecto_Resoluci%C3%B3n_CREG_701_028_2023/Documento%20Soporte%20CREG%20901%20044%20de%202023.pdf
- Equipo Analítica XM. (s. f.). *GitHub - EquipoAnaliticaXM/API_XM: Este repositorio se crea con el fin de compartir herramientas de consulta para extraer información relevante del Mercado de Energía Mayorista colombiano* · GitHub. Recuperado 11 de abril de 2026, de https://github.com/EquipoAnaliticaXM/API_XM
- Failing, J. M., Segarra-Tamarit, J., Cardo-Miota, J., & Beltran, H. (2026). Deep learning-based prediction models for spot electricity market prices in the Spanish market. *Mathematics and Computers in Simulation*, 240, 96-104. <https://doi.org/10.1016/J.MATCOM.2025.07.010>
- Gneiting, T., & Raftery, A. E. (2007). Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association*, 102, 359-378. <https://doi.org/10.1198/016214506000001437>
- Gneiting, T., Raftery, A. E., Westveld, A. H., Goldman, T., Gneiting, T., Raftery, A. E., III, A. H. W., & Goldman, T. (2005). Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation. *Monthly Weather Review*, 133(5), 1098-1118. <https://doi.org/10.1175/MWR2904.1>
- Kim, T., Kim, J., Vuno, Y. T., Park, C., Choi, J.-H., & Choo, J. (s. f.). *Reversible Instance Normalization for Accurate Time-Series Forecasting against Distribution Shift*.
- Lago, J., Marcjasz, G., De Schutter, B., & Weron, R. (2021). Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark. *Applied Energy*, 293, 116983. <https://doi.org/10.1016/J.APENERGY.2021.116983>
- Laio, F., & Tamea, S. (2007). Verification tools for probabilistic forecasts of continuous hydrological variables. *Hydrology and Earth System Sciences*, 11(4), 1267-1277. <https://doi.org/10.5194/HESS-11-1267-2007>

- Ley 143 de 1994 - Gestor Normativo - Función Pública. (s. f.). Recuperado 5 de enero de 2026, de <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=4631>
- Lim, B., Arik, S., Loeff, N., & Pfister, T. (2021). Temporal Fusion Transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4), 1748-1764. <https://doi.org/10.1016/j.ijforecast.2021.03.012>
- Monica, A. A. A., Botero, S. B., & Jaime, H. H. B. (2018). Anomalías Financieras en el Mercado de Electricidad: Análisis empririco de los precios spot. *Iberian Conference on Information Systems and Technologies, CISTI, 2018-June*, 1-7. <https://doi.org/10.23919/CISTI.2018.8398640>
- NOAA. (2009). *Climate Variability: Oceanic Niño Index* | NOAA Climate.gov. <https://www.climate.gov/news-features/understanding-climate/climate-variability-oceanic-nino-index>
- NOAA. (2026). *ONI Timeseries*. <https://psl.noaa.gov/data/timeseries/month/DS/ONI/>
- Ramírez Londoño, E. A., & Giraldo Zuluaga, S. (2023). *Predicción del precio de la energía eléctrica en Colombia con horizonte de 24 horas utilizando modelos de Deep Learning*. Universidad de Antioquia. <https://hdl.handle.net/10495/37596>
- Rozas Rodríguez, C. W. (2024). *Artificial intelligence predictive models for local renewable energy markets*. 1. <https://dialnet.unirioja.es/servlet/tesis?codigo=387065&info=resumen&idioma=ENG>
- Salinas, D., Flunkert, V., Gasthaus, J., & Januschowski, T. (2020). DeepAR: Probabilistic forecasting with autoregressive recurrent networks. *International Journal of Forecasting*, 36(3), 1181-1191. <https://doi.org/10.1016/j.ijforecast.2019.07.001>
- Vaswani, A., Brain, G., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention Is All You Need*. 1. <https://arxiv.org/pdf/1706.03762>
- Villa-Loaiza, C., Taype-Huaman, I., Benavides-Franco, J., Buenaventura-Vera, G., & Carabalí-Mosquera, J. (2023). Does climate impact the relationship between the energy price and the stock market? The Colombian case. *Applied Energy*, 336, 120800. <https://doi.org/10.1016/J.APENERGY.2023.120800>
- Villarreal, Y. J., Luis, M., Flores, A., Martín, S., María, P., Hurtado, A., Eafit, U., De Economía, E., & Finanzas, Y. (2023). *Predicción del precio de la energía eléctrica en Colombia mediante un enfoque de machine learning*. Universidad EAFIT. <https://hdl.handle.net/10784/33253>
- XM. (s. f.-a). *Operación del SIN y Administración del MEM en Colombia* | XM. Recuperado 5 de enero de 2026, de <https://www.xm.com.co/nuestra-empresa/nosotros/que-hacemos>
- XM. (s. f.-b). *Preguntas Frecuentes* – XM. Recuperado 5 de enero de 2026, de <https://www.xm.com.co/herramientas/preguntas-frecuentes>
- XM. (2026a). *Páginas - Generación Real del SIN*. <https://sinergox.xm.com.co/oferta/Paginas/Informes/GeneracionSIN.aspx>

XM. (2026b). *Repositorio Equipo Analítica XM*. https://github.com/EquipoAnaliticaXM/API_XM

[PÁGINA INTENCIONADAMENTE EN BLANCO]