



**ESCUELA DE NEGOCIOS**

**MAESTRÍA EN INTELIGENCIA DE NEGOCIOS Y CIENCA DE DATOS**

**PRONÓSTICO DEL SECTOR ELÉCTRICO EN ECUADOR: EVOLUCIÓN  
MENSUAL POR TIPO DE FUENTE DESDE 2013 HASTA 2023**

**Profesor  
PhD Manuel Eugenio Morocho Cayamcela**

**Autor  
Claudio David Arias Piedra**

**2024**

## RESUMEN

El análisis de la generación eléctrica a nivel nacional cobra importancia debido al importante incremento de la demanda, por lo que es esencial contar con métodos predictivos que permitan estimar con precisión los comportamientos futuros para una toma de decisiones más acertada. Por tal motivo, el presente trabajo analiza y desarrolla modelos predictivos en series temporales mediante SARIMA (*seasonal autoregressive integrated moving average*) y Prophet aplicado a la generación eléctrica por tipo de fuente durante el período 2013-2023 en Ecuador. Los resultados han sido evaluados mediante las métricas o funciones de pérdida RMSE (*root mean square error*) y MAPE (*mean absolute percentage error*) para los sets de entrenamiento y prueba. Se concluye que Prophet presenta un nivel de precisión elevado y bueno para los sets de entrenamiento y prueba respectivamente en las series estudiadas, su facilidad de implementación, uso y ajuste lo convierten en una herramienta altamente precisa, sin embargo, ambos modelos comparten disminuciones importantes en el rendimiento para el set de prueba, evidenciando posibles mejoras futuras.

## ABSTRACT

The analysis of electricity generation at a nationwide level becomes important due to the significant increase of the demand, so it is essential to have predictive methods that allow accurate estimation of future behaviors for better decision making. Therefore, this thesis analyzes and develops predictive models in time series using SARIMA (*seasonal autoregressive integrated moving average*) and Prophet applied to electricity generation by type of source during the period 2013-2023 in Ecuador. The results have been tested using the RMSE (*root mean square error*) and MAPE (*mean absolute percentage error*) metrics. It is concluded that Prophet presents a high and good level of accuracy for the training and test sets respectively, its ease of implementation, use and adjustment make it a highly accurate tool, however, both models share significant decreases in performance for the test set, showing possible future improvements.

## ÍNDICE DEL CONTENIDO

RESUMEN .....	2
ABSTRACT .....	3
ÍNDICE DEL CONTENIDO.....	4
ÍNDICE DE TABLAS .....	5
ÍNDICE DE FIGURAS .....	6
INTRODUCCIÓN .....	1
REVISIÓN DE LITERATURA.....	2
IDENTIFICACIÓN DEL OBJETO DE ESTUDIO .....	10
PLANTEAMIENTO DEL PROBLEMA .....	12
OBJETIVO GENERAL .....	13
OBJETIVOS ESPECÍFICOS .....	13
JUSTIFICACIÓN Y APLICACIÓN DE LA METODOLOGÍA .....	14
SELECCIÓN DE LA BASE DE DATOS .....	14
PREPROCESAMIENTO DE DATOS .....	15
IDENTIFICACIÓN, DESCRIPCIÓN Y VISUALIZACIÓN DE VARIABLES .....	16
SELECCIÓN DEL MODELO ESTADÍSTICO .....	19
RESULTADOS .....	22
ANÁLISIS DESCRIPTIVO DE LA SERIE DE TIEMPO .....	22
ANÁLISIS PREDICTIVO MEDIANTE SARIMA .....	25
ANÁLISIS PREDICTIVO MEDIANTE PROPHET .....	34
DISCUSIÓN DE LOS RESULTADOS Y PROPUESTA DE SOLUCIÓN .....	39
CONCLUSIONES Y RECOMENDACIONES .....	41
REFERENCIAS.....	42
ANEXOS .....	48

## ÍNDICE DE TABLAS

Tabla 1 Matriz comparativa de investigaciones académicas similares.....	7
Tabla 2 Detalle de fuentes primarias y secundarias .....	9
Tabla 3 Estructura de los datasets originales.....	14
Tabla 4 Estadísticas descriptivas de las series temporales .....	16
Tabla 5 Selección de variables.....	17
Tabla 6 Resultados del test ADF para las series de tiempo .....	25
Tabla 7 Estructura del modelo para la serie Hidráulica .....	28
Tabla 8 Resultados de la prueba Ljung-Box para el modelo SARIMA de “Hidráulica” .....	30
Tabla 9 Categorización de MAPE para pronósticos .....	31
Tabla 10 Resultados de las métricas de error para el modelo SARIMA en Hidráulica .....	31
Tabla 11 Estructura del modelo para la serie Térmica MCI .....	32
Tabla 12 Resultados de la prueba Ljung-Box para el modelo SARIMA de “Térmica MCI” .....	33
Tabla 13 Resultados de las métricas de error para el modelo SARIMA en Térmica MCI.....	34
Tabla 14 Hiperparámetros del modelo Prophet para la serie Hidráulica .....	35
Tabla 15 Resultados de las métricas de error para el modelo Prophet en Hidráulica .....	36
Tabla 16 Hiperparámetros del modelo Prophet para la serie Térmica MCI.....	37
Tabla 17 Resultados de las métricas de error para el modelo Prophet en Térmica MCI.....	37
Tabla 18 Resumen de las métricas de rendimiento para los modelos estudiados .....	38

## ÍNDICE DE FIGURAS

Figura 1 Visualización de las series de tiempo.....	18
Figura 2 Visualización de las series de tiempo seleccionadas .....	19
Figura 3 Serie total de la producción eléctrica neta por mes.....	22
Figura 4 Tasa de variación mensual de la producción eléctrica .....	23
Figura 5 Composición de la producción eléctrica mensual por tipo de fuente..	24
Figura 6 Porcentaje de aporte de las series Hidráulica y Térmica MCI.....	25
Figura 7 Gráfico de autocorrelación y autocorrelación parcial para la serie Hidráulica .....	26
Figura 8 Gráfico de autocorrelación y autocorrelación parcial para la serie Térmica MCI.....	27
Figura 9 Division del conjunto de datos en train y test set para la series .....	27
Figura 10 Residuos del modelo SARIMA aplicado a la serie Hidráulica .....	29
Figura 11 Serie original y predicciones de la serie Hidráulica con SARIMA.....	30
Figura 12 Residuos del modelo SARIMA aplicado a la serie Térmica MCI.....	33
Figura 13 Serie original y predicciones de la serie Térmica MCI con SARIMA	34
Figura 14 Serie original y predicciones de la serie Hidráulica con Prophet.....	36
Figura 15 Serie original y predicciones de la serie Térmica MCI con Prophet .	38

## INTRODUCCIÓN

En un contexto de transformación digital y evolución tecnológica, la generación de electricidad se convierte en un aspecto de elevada importancia para mantener la continuidad del servicio en un entorno donde el consumo eléctrico promedio se ha visto incrementado (Lange et al., 2020) y la oferta no es lo suficiente alta para satisfacer todos los requerimientos (Qin et al., 2022).

Por tal motivo, el estudio de las series temporales y el modelamiento de la generación eléctrica constituyen elementos de alto valor para los actores tanto público como privados en la sociedad debido a su correlación en el desarrollo económico de los países (Stern et al., 2019).

El presente trabajo de investigación buscar estudiar el comportamiento de las series temporales para la generación o producción eléctrica neta del Ecuador por tipo de fuente entre los años 2013 a 2023 con una frecuencia mensual y utilizando datos provenientes de los organismos de control.

En un primer momento, se realiza la revisión bibliográfica con respecto a publicaciones similares y evaluar los posibles modelos a ser aplicados. Posteriormente, se presenta el objeto de estudio, problema y establecimiento de los objetivos (general y específicos) para brindar el contexto metodológico de los modelos estadísticos.

En un segundo momento, se indican las estadísticas descriptivas y el modelado de las series “Hidráulica” y “Térmica” utilizando técnicas como *seasonal autoregressive integrated moving average* (SARIMA) y Prophet, así como también las respectivas métricas de evaluación en términos de *root mean square error* (RMSE) y *mean absolute percentage error* (MAPE). Finalmente, se debaten los hallazgos estadísticos, implicaciones, conclusiones y recomendaciones.

## REVISIÓN DE LITERATURA

La relación entre el modelado estadístico-matemático y el sector energético ha sido ampliamente discutida en la academia con respecto a la utilización de técnicas tradicionales y la validación de nuevas metodologías derivadas de los importantes avances en machine learning e inteligencia artificial (M. K. Islam et al., 2020). La necesidad real de obtener pronósticos precisos deriva del importante impacto que posee este sector en la macroeconomía del país a través de indicadores como las tasas de crecimiento y empleo (Sievers et al., 2019), considerando, además, la existencia de un periodo de transición vigente que busca una generación de electricidad más sustentable mediante el uso energías renovables (Maji & Saari, 2021).

Según Klyuev et al. (2022), la literatura científica reciente ha evaluado aproximadamente un total de 200 metodologías individuales diferentes y múltiples variaciones de estos procedimientos en estudios relacionados con el pronósticos o *“forecasting”* de los niveles de demanda energética en las industrias, concluyendo que no existe un método general que sea superior en todos las situaciones, cada caso de estudio requiere de un proceso analítico adecuado al igual que una correcta configuración de los parámetros en cada modelo.

En este sentido, el modelado de series temporales aplicados a la demanda y producción de ha permitido establecer una correcta formulación de políticas estratégicas entorno al desarrollo del sector energético a través de los modelos EMP (por sus siglas en ingles), destacando a la complejidad, la naturaleza y el objetivo de estudio como elementos clave para una correcta elección de los métodos predictivos (Debnath & Mourshed, 2018).

En Ecuador, el estudio de la generación eléctrica ha sido evaluado desde la capacidad hidráulica según lo presentado por Barzola-Monteses et al. (2019),



quienes utilizaron una serie histórica desde el año 2000 hasta el 2009 con la metodología Box-Jenkins *autoregressive integrated moving average* (ARIMA) y Box-Tiao *autoregressive integrated moving average with exogenous variables* (ARIMAX). Los resultados evidencian una superioridad del modelo ARIMAX (1,1,1)(1,0,0) para la predicción mensual de hasta un año, considerando una variable exógena al momento de las estimaciones.

Otros enfoques incorporan, como se mencionó anteriormente, nuevas metodologías basadas en inteligencia artificial como: redes neuronales (Kouvara & Vogiatzis 2023) y árboles de decisión (Pinto et al. 2021). En consecuencia, Ünlü (2019) establece un estudio comparativo sobre las capacidades predictivas de la generación eléctrica en Turquía utilizando *support vector regression* (SVR), redes neuronales profundas mediante *long short term memory* (LSTM) y *multi-layer perceptron* (MLP), demostrando que LSTM presenta los mejores resultados en las métricas de evaluación y cuyo impacto real puede verse aplicado en una mejor planificación en materia de sostenibilidad.

Bitirgen y Filik (2020), Pinto et al. (2021) y Ledmaoui et al. (2023) incorporan técnicas como *random forest*, *adaboost*, *decision trees*, *xgboost* y *gradient boosted regression* para la estimación de series temporales con respecto al consumo y generación de energía eléctrica a través de múltiples periodos de tiempo. Los resultados muestran un excelente desempeño con respecto a modelos como ARIMA y marcan una tendencia al ser altamente precisos con bajos niveles de errores.

Esta estructura analítica para el estudio de múltiples metodologías y sus capacidades predictivas en series temporales han derivado en la obtención de importantes investigaciones comparativas sobre la viabilidad de usar la estadística tradicional y las nuevas tendencias de modelado. Los estudios presentados por Pontoh et al. (2021) con respecto al modelamiento de pasajeros en tren y Toharudin et al. (2023) para la estimación de temperaturas mínimas en máximas en Indonesia, los cuales evaluaron e implementaron modelos de redes

neuronales LSTM y un novedoso enfoque mediante el modelo Prophet desarrollado inicialmente por Facebook ahora Meta (Facebook, 2017; Taylor & Letham, 2017) aplicado a series temporales. Las conclusiones resaltan que Prophet posee una importante capacidad predictiva frente a metodologías más complejas, con métricas de rendimiento muy similares y competitivos (Feng et al., 2022).

Similares resultados se pueden observar en Kouvara y Vogiatzis (2023), quienes compararon las capacidades de pronóstico en cinco modelos: ARIMA, *seasonal autoregressive integrated moving average* (SARIMA), *seasonal autoregressive integrated moving average with exogenous variables* (SARIMAX), Prophet y LSTM para una serie de tiempo sobre el consumo eléctrico por hora y datos climáticos durante un período de 3 años en Grecia, con la configuración de un set de entrenamiento para los años 2019 a 2022 y un set de prueba utilizando el año 2022. Los autores concluyen que los modelos SARIMA y Prophet reportan las mejores métricas de rendimiento frente a otros modelos en términos de *mean square error* (MSE), *root mean square error* (RMSE), *mean absolute percentage error* (MAPE) y coeficiente de determinación ( $R^2$ ).

Los investigadores también han realizado importantes esfuerzos para profundizar y evaluar estos dos modelos (SARIMA y Prophet) en diferentes conjuntos de datos sobre la generación de estimaciones para el sector energético, esto debido a sus propiedades para el manejo de estacionalidad y descomposición de componentes.

Samal et al. (2019) evalúan estos modelos en una serie temporal para un conjunto de diez indicadores de contaminación en India durante un período de 10 años. Durante este proceso se consideró la aplicación de un logaritmo a las series evaluadas con el objetivo de estabilizar las series y estudiar su impacto en las métricas finales. Los resultados y conclusiones del modelamiento permitieron determinar que Prophet supera a SARIMA y se desempeña mejor al utilizar las series logarítmicas con respecto al modelo base.

Atasever et al. (2022) utilizaron un esquema similar de comparación, pero aplicado al pronóstico de generación eléctrica mediante un sistema eólico durante 2017 y 2018, utilizando como métricas a RMSE y MAE (*mean absolute error*). Los resultados respaldan las evidencias académicas y presentan a Prophet como un modelo altamente preciso, con una alta capacidad de adaptabilidad sobre los datos y una menor cantidad de ajuste o *fine-tuning* en los parámetros. El comportamiento del modelo puede ser controlado de forma eficiente mediante parámetros específicos relacionados a los puntos de cambios en la serie temporal, tipo de crecimiento, tipo de estacionalidad, entre otros. Entre los hiperparámetros utilizados encontramos: `changepoint_prior_scale`, `changepoint_range`, `seasonality_prior_scale`, `seasonality_mode`, `growth`, `weekly_seasonality` y `daily_seasonality`.

El reciente estudio de M. K. Islam et al. (2023), sobre los pronósticos de recursos solares y eólicos para la generación de electricidad, reitera nuevamente las conclusiones presentadas en la literatura científica sobre el uso de Prophet como un instrumento altamente eficiente y con valores de RMSE y MAE inferiores a los presentados por SARIMA en una serie temporal compuesta por un registro mensual desde 1990 hasta 2018.

Considerando los antecedes antes presentados, se observa que el uso del modelo Prophet para la estimación y pronóstico de series temporales destaca sobre otras metodologías comúnmente utilizadas debido a su alto nivel de precisión y bajo error. Es así como ciertos investigadores han aplicado este concepto a la generación eléctrica y a múltiples áreas de estudio que incluye, pero no se limitan a la predicción de contaminantes en el aire (Hasnain et al., 2022) e incluso el pronóstico de contagios (Aditya Satrio et al., 2021).

Chadalavada et al. (2020) realizan un estudio aplicado a la predicción de consumo de energía con el dataset “Appliance” de UCI Machine Learning y cuya metodología se centra exclusivamente en el uso de un único modelo (Prophet)

para el cálculo de los valores futuros y sus correspondientes métricas. Los resultados concuerdan y avalan los estudios antes presentados al reportar evidencia sobre el elevado rendimiento del modelo.

Finalmente, Huang y Dong (2022), modelan una serie temporal para la demanda de material eléctrico en China mediante Prophet, reportando un alto nivel de precisión y aplicabilidad de las predicciones generadas, al permitir el establecimiento de políticas previsorias al momento de requerir y preparar los materiales de las empresas de energía eléctrica en el futuro.

En la Tabla 1 se encuentran resumidas las diferentes de investigaciones realizadas en torno a la diferentes metodologías utilizadas, las bases de datos, los resultados obtenidos y las principales implicaciones derivadas de estos procesos analíticos. Además, en la Tabla 2 se ha presentado el detalle de las fuentes utilizadas para el presente documento de investigación así como también una breve descripción sobre su contenido y origen.

**Tabla 1***Matriz comparativa de investigaciones académicas similares*

<b>Autor(es)</b>	<b>Datos</b>	<b>Metodología</b>	<b>Resultados</b>	<b>Implicaciones</b>
(Ünlü, 2019)	Serie mensual univariada de 1975 a 2017 en Turquía (TEIAS).	SVR, MLP, LSTM.	LSTM supera a los otros métodos de regresión.	Planificación futura en materia de sostenibilidad.
(Kouvara & Vogiatzis, 2023)	Serie desde 2019 hasta 2021 de consumo por hora en Grecia.	ARIMA, SARIMA, SARIMAX, Prophet, LSTM.	SARIMA y Prophet obtuvieron los mejores resultados.	Mejor comprensión, monitorización y pronóstico.
(M. K. Islam et al., 2023)	Serie mensual de la radiación y velocidad en Australia.	SARIMA y Prophet.	El modelo de Prophet es más eficiente.	Planificación y desarrollo de un sistema sostenible.
(Pinto et al., 2021)	Consumo de energía promedio por hora durante 10 días.	Random Forest, Gradient Boosted Regression, AdaBoost.	AdaBoost presenta un desempeño superior.	Los pronósticos permiten analizar entre el consumo y generación.
(Barzola-Monteses et al., 2019)	Serie mensual de producción mensual bruta en Ecuador 2014 2015.	ARIMA y ARIMAX.	El modelo ARIMAX desempeña mejor que el ARIMA univariado.	Desarrollar estrategias para la planificación del sector energético.
(Pontoh et al., 2021)	Serie mensual desde 2006 hasta 2019 en la región de Java Island.	Feed Foward Neural Network (FFNN) y Prophet.	Diferencias entre RMSE no es significativa entre los modelos.	Políticas para la gestión eficaz de la capacidad.

Autor(es)	Datos	Metodología	Resultados	Implicaciones
(Atasever et al., 2022)	Serie de la velocidad del viento, Francia 2017 y 2018.	SARIMA y Prophet.	El modelo Prophet presenta mejores resultados.	Estimaciones precisas sobre los aerogeneradores y su disponibilidad.
(Ledmaoui et al., 2023)	Serie diaria de generación eléctrica solar, Marruecos 2022.	SVR, ANN, DT, RF, GAM, XGBoost.	El modelo de ANN reporta los mejores resultados en las métricas utilizadas.	Programar actividades de mantenimiento preventivo y eficiencia.
(Chadalavada et al., 2020)	UCI Machine Learning Repository “Appliance” Energy Prediction.	Prophet	Prophet presenta mejores resultados que métodos como ARIMA.	Analizar los niveles de ahorro para asegurar la disponibilidad.
(Huang & Dong 2022)	Serie de 12 meses sobre la demanda de materiales eléctricos en China en 2021	Prophet	Resultados precisos y un error controlado menor al 10%.	Estimaciones precisas para el establecimiento de políticas previsorias.
(Bitirgen & Filik, 2020)	Serie histórica de 1 año con registros sobre el precio obtenido de NYISO.	XGBoost y ARIMA.	XGBoost tiene mejores capacidades de pronóstico y los errores son bajos.	Cuantificar el nivel de incertidumbre, mejorar la gestión de riesgos y operaciones del mercado eléctrico.

Fuente: Elaboración propia.

**Tabla 2***Detalle de fuentes primarias y secundarias*

<b>Autor(es)</b>	<b>Tipo</b>	<b>Descripción</b>	<b>Fuente</b>	<b>Detalle</b>
Agencia de Regulación y Control Eléctrico	Base de datos	Base de datos anual consolidada desde 2013 hasta 2023.	Sistema de Reportes - SISDAT	Primaria
(M. K. Islam et al., 2023)	Artículo Científico	Documento científico para la metodología (modelo).	Energies	Primaria
(Pontoh et al., 2021)	Artículo Científico	Documento científico para la metodología (modelo).	Journal of Physics: Conference Series	Primaria
(Chadalavada et al., 2020)	Artículo Científico	Documento científico para la metodología (serie).	Indian Journal of Science and Technology	Primaria
(Samal et al., 2019)	Artículo Científico	Documento científico para la metodología (modelo).	ITCC'19	Primaria
(Barzola-Monteses et al., 2019)	Artículo Científico	Documento científico de estudios en el país.	Sustainability (Switzerland)	Secundaria
Klyuev et al. (2022)	Artículo Científico	Documento científico de revisión bibliográfica.	Energies	Secundaria
(Debnath & Mourshed, 2018)	Artículo Científico	Documento científico de revisión bibliográfica.	Renewable and Sustainable Energy Reviews	Secundaria

Fuente: Elaboración propia.

## IDENTIFICACIÓN DEL OBJETO DE ESTUDIO

El modelamiento de series temporales constituye un elemento de gran relevancia en el contexto de planificación y toma de decisiones estratégicas para múltiples áreas de aplicación a nivel sectorial, a saber: operaciones, finanzas, ambiente, energía, entre otros (Petropoulos et al., 2022). Específicamente, el sector energético puede beneficiarse de este tipo de metodologías para el desarrollo de políticas adecuadas e impactando directamente en la eficiencia y desempeño de los sistemas (Ahmad et al., 2020).

Para el caso ecuatoriano, la generación de electricidad constituye actualmente un elemento de gran interés y discusión debido a la existente dificultad para cubrir la demanda de forma interna, recurriendo a otros métodos como la importación de energía, para brindar soluciones temporales a un problema que requiere importantes cambios estructurales. Según Párraga Palacios et al. (2019), el Ecuador podría utilizar fuentes de recursos renovables, como la energía eólica, para aumentar la capacidad generativa de estas fuentes en la matriz energética.

En consecuencia, la presente investigación hace uso de un conjunto de datos en formato de serie temporal con información perteneciente a la generación eléctrica neta por megavatio hora (MWh) con una segmentación por tipo de subcentral y una frecuencia mensual durante el periodo 2013 - 2023. El desarrollo individualizado de estas series permitirá la generación de estimaciones para periodos futuros mediante la construcción de un modelo robusto.

El proyecto propuesto se justifica debido a la relevancia de los modelos para pronósticos en la planificación de estrategias y medidas de contención que se podrían implementar en casos donde la generación esperada de electricidad llegue a niveles críticos. Además, el uso de visualizaciones permite una mejor



comprensión de la información, en un entorno donde el tiempo apremia y la toma de decisiones debe ser precisa para salvaguardar la disponibilidad de energía a nivel nacional.

## PLANTEAMIENTO DEL PROBLEMA

*Data-driven decisión-making (D3M)* establece patrones claros de análisis para brindar un soporte numérico en pro de una mejora en cuanto a la calidad y fiabilidad de las estrategias a ser aplicadas (Lu et al., 2019). Todo esto bajo un contexto de incertidumbre y dinamismo sobre las variables que modifican el comportamiento del objeto de estudio, requiriendo de algoritmos (o modelos) que potencialmente mejoren la objetividad y reduzcan los sesgos al momento de establecer los procedimientos a ser implementados (Oliver, 2019).

Para el caso gubernamental y la administración recursos públicos, el uso de analítica de datos permite una mejor comprensión de la comunidad (Washington, 2014). Manikam et al. (2019) señalan que esta metodología genera un impacto positivo en productividad, eficiencia, efectividad, reducción de costos y beneficios compartidos que permiten fomentar el desarrollo a través del tiempo (Ongena & Davids 2023).

Es entonces imperativo el uso de metodologías adecuadas para expandir las fuentes de información hacia los tomadores de decisiones y brindar, en consecuencia, un análisis exhaustivo que permita comprender íntegramente todos los posibles resultados y los planes de acción correspondientes para evitar obstáculos futuros.

En relación con la problemática expuesta, la presente investigación utiliza un enfoque analítico con el objetivo de establecer, estadísticamente, la construcción de un modelo predictivo robusto mediante las metodologías de Prophet y SARIMA para series temporales de generación eléctrica en Ecuador, que derive en una toma de decisiones con un mayor nivel de certeza, así como también una mejora en la comprensión sobre el comportamiento de las variables y su importancia dentro del análisis sectorial para la elaboración de políticas adecuadas.

Para este efecto, se ha planteado el uso de la generación eléctrica neta en Ecuador a nivel mensual cuya unidad de medida es megavatios por hora (MWh) para un período de análisis entre 2013 y 2023 con respecto a los diferentes tipos de fuente u origen. Las estimaciones generadas de los modelos serán utilizadas para mejorar las capacidades predictivas de los agentes públicos y privados sobre el comportamiento futuro esperado de la generación eléctrica en Ecuador y planificar estratégicamente los procedimientos a seguir en casos requeridos.

### **OBJETIVO GENERAL**

Desarrollar un modelo de predicción para series temporales mediante las metodologías SARIMA y Prophet para la generación eléctrica neta a nivel nacional por tipo de origen.

### **OBJETIVOS ESPECÍFICOS**

- Investigar las series temporales mediante el uso de estadística descriptiva para comprender su evolución y desarrollo a través del tiempo.
- Calcular los niveles de errores mediante las funciones mediante el uso de métricas específicas como RMSE y MAPE.
- Comparar las métricas en los sets de entrenamiento y prueba para establecer su nivel de precisión en las predicciones realizadas.
- Interpretar los resultados obtenidos con respecto al contexto de la generación eléctrica y su influencia en el sector público y privado.

## JUSTIFICACIÓN Y APLICACIÓN DE LA METODOLOGÍA

Metodológicamente esta investigación hace uso de herramientas como la estadística descriptiva para un análisis exploratorio inicial sobre el comportamiento de las series a través del tiempo, técnicas de modelamiento matemático-estadístico para la construcción de los modelos predictivos y finalmente, comparar sus resultados para determinar aquel modelo que mejor rendimiento reporta sobre los datos.

La selección de los modelos estadísticos responde a su frecuencia de uso y capacidad predictiva reportada en investigaciones similares realizadas a un objeto de estudio similar en el campo de la generación eléctrica. De este modo, los modelos SARIMA y Prophet serán evaluados mediante métricas o funciones de pérdidas como RMSE y MAPE para cada una de las series temporales seleccionadas.

### Selección de la base de datos

Los registros han sido obtenidos directamente de la Agencia de Regulación y Control de Electricidad mediante SISDAT para la generación de reportes automáticos con un detalle anual desde el año 2013 hasta el año 2023 utilizando la tipología “Energía Producida” y cuyo formato de descarga corresponde a un documento Excel (.xlsx) (Agencia de Regulación y Control de Electricidad, 2024).

**Tabla 3**

*Estructura de los datasets originales*

<b>Dataset</b>	<b>Número de observaciones</b>	<b>Número de columnas</b>
Año 2013	9039	28
Año 2014	9387	28
Año 2015	10721	28
Año 2016	11796	28
Año 2017	11975	28

Año 2018	11281	28
Año 2019	11398	28
Año 2020	11312	28
Año 2021	11868	28
Año 2022	11944	28
Año 2023	12077	28

---

Fuente: Elaboración propia.

La Tabla 3 presenta el conteo de registros originales conformados por el conjunto de documentos Excel individuales para cada uno de los años en estudio, la estructura interna de estos documentos se define mediante un diseño tabular donde cada fila corresponde una observación específica y cada columna representa una variable de medición o identificación.

### **Preprocesamiento de datos**

Para la construcción de las series temporales se procedió a utilizar una tabla pivotante cuyo índice principal corresponde a la variable “MES” (cardinalidad de 12), el eje columnar se define mediante las categorías presentes en la variable “Subtipo Central” (cardinalidad de 7) y cuyos valores corresponde a la sumatoria de la generación eléctrica de la variable “Energía Neta (MWh)” para cada tipo específico de origen en cada uno de los meses, proceso que fue aplicado sobre cada uno de los periodos estudiados. Finalmente, se combinaron cada una de las tablas individuales en un solo dataframe, resultando en una serie temporal con un nivel de detalle mensual para un intervalo desde el año 2013 hasta 2023 a nivel nacional para cada fuente de origen.

De acuerdo con lo presentado en la Tabla 4, la estructura de las series muestra conjunto de 132 registros sin datos faltantes o N/A, las estadísticas de media y desviación estándar sugieren que la generación eléctrica neta “Energía Neta (MWh)” se da en mayor nivel para las categorías: “Hidráulica”, “Térmica MCI (motor de combustión interna)” y “Térmica Turbovapor”, con una dispersión

importante al implicar una considerable fluctuación de los datos alrededor de la media.

**Tabla 4**

*Estadísticas descriptivas de las series temporales*

<b>Fuente</b>	<b>Media</b>	<b>Mediana</b>	<b>Desv. Est</b>	<b>Conteo</b>	<b>N/A</b>
Eólica	7.308,897	6.432,332	4.532,588	132	0
Fotovoltaica	2.708,318	3.053,140	970,212	132	0
Hidráulica	1.637.118,315	1.709.248,258	506.373,662	132	0
Interconexión	34.844,273	4.082,856	59.738,134	132	0
Térmica MCI	444.219,592	408.976,600	102.116,272	132	0
Térmica Turbogas	155.025,773	123.195,628	88.774,518	132	0
Térmica Turbovapor	163.899,041	151.453,472	63.959,400	132	0

Fuente: Elaboración propia.

Esta investigación empleó la herramienta de Python (versión 3.10.11) para la generación de gráficos y estimaciones de los modelos predictivos. El acceso a los documentos de datos, visualizaciones y el código en Python utilizado para la presente investigación puede ser accedido desde un navegador web mediante el repositorio de GitHub (Arias, 2024) disponible de forma gratuita.

### **Identificación, descripción y visualización de variables**

La selección de las series analizadas (Tabla 5) se justifica, en primer lugar, debido a la extensión y objetivo del presente trabajo investigativo, el cual buscar explorar las capacidades predictivas de los modelos utilizados en el ámbito de la predicción de generación eléctrica con un enfoque hacia la realidad ecuatoriana previo a una implementación completa para las series restantes.

**Tabla 5***Selección de variables*

<b>Variable</b>	<b>Descripción</b>	<b>Tipo</b>	<b>Medición</b>
Hidráulica	Serie mensual desde el año 2013 hasta 2023 para la energía eléctrica neta generada hidráulicamente.	Numérica	MWh
Térmica MCI	Serie mensual desde el año 2013 hasta 2023 para la energía eléctrica neta generada térmicamente mediante MCI.	Numérica	MWh

Fuente: Elaboración propia.

En segundo lugar, se ha considerado el nivel de aporte de generación eléctrica para cada una de las fuentes, donde la contribución conjunta de las centrales hidroeléctricas “Hidroeléctrica” y centrales termoeléctricas por motor de combustión interna “Térmica MCI” aportan con aproximadamente el 80% del total de energía neta a nivel nacional durante los años estudiados, convirtiéndose en elementos fundamentales para el sistema eléctrico del país y en series altamente relevantes sobre las cuales aplicar modelos predictivos.

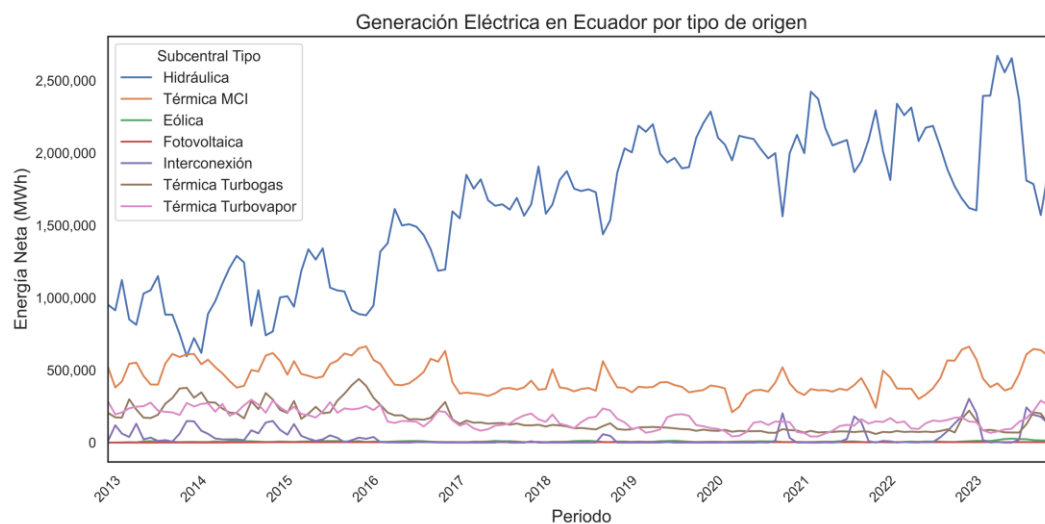
En tercer lugar, el uso de la variable “Energía Neta” permite obtener un análisis real de la cantidad real de energía que se pone a disposición para los consumidores a través de los sistemas de transmisión como se evidencia en su aplicación para los estudios de Wang et al. (2019), Porse et al. (2020) y Samsó et al. (2023).

De acuerdo con lo presentado en la Figura 1, el comportamiento de la generación eléctrica muestra una fuerte dependencia de las centrales hidroeléctricas y térmica MCI, destacando su importancia para un análisis estadístico detallado dado el contexto situacional del país. Estas fuentes forman parte del Sistema Nacional Interconectado (SIN), el cual corresponde a todo el conjunto de

elementos que facilitan tanto la producción como la transferencia de energía desde los diferentes centros generadores hasta los consumidores (Ministerio de Energías y Minas 2022).

**Figura 1**

*Visualización de las series de tiempo*



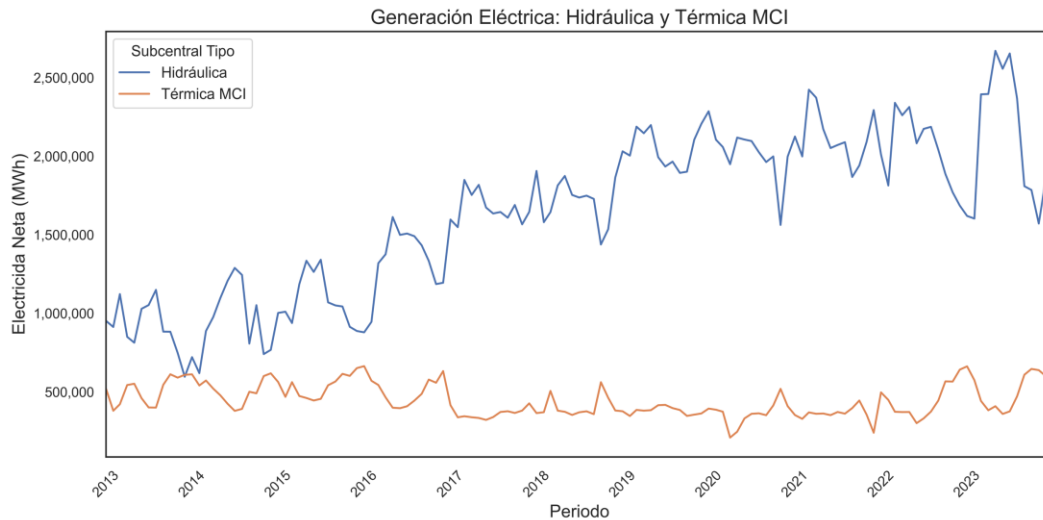
Fuente: Elaboración propia.

Al extraer exclusivamente las dos series analizadas en la Figura 2, se evidencia en mayor detalle que el comportamiento de la generación “Hidráulica” es creciente a través del tiempo y con un cierto nivel cíclico durante los últimos meses de cada año. La serie “Térmica MCI” por su parte, presenta una disminución en los valores de energía neta generada desde aproximadamente el año 2017 hasta finales de 2022, donde recupera inicialmente sus valores históricos de producción a los presentados en el año 2016 o anteriores.



**Figura 2**

*Visualización de las series de tiempo seleccionadas*



Fuente: Elaboración propia.

### Selección del modelo estadístico

El modelo Prophet considera un esquema basado en la descomposición temporal en tres segmentos bien definidos: tendencia de la serie, estacionalidad y feriados o días especiales, con un enfoque hacia el diseño intuitivo para un ajuste del modelo de una forma sencilla sin requerir que los analistas conozcan en profundidad el funcionamiento interno del mismo (Taylor & Letham, 2017).

Prophet se especifica matemáticamente mediante la Ecuación 1 donde  $g(t)$  representa el componente tendencial para el modelamiento de cambios no periódicos,  $s(t)$  representa los cambios que ocurren de forma periódica a nivel semana - anual y  $h(t)$  para el control de efectos potenciales irregulares debido a eventos específicos durante un cierto intervalo de tiempo.  $\epsilon_t$  agrupa los otros cambios particulares no considerados en las otros componentes.

### Ecuación 1

*Modelo general de Prophet*

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

Para el correcto modelamiento de las diferentes características de las series temporales y aplicaciones en el entorno empresarial de Facebook (ahora Meta), Prophet utiliza dos estructuras para el modelamiento del componente tendencial: crecimiento no lineal saturado (Ecuación 2) y lineal por partes (Ecuación 3) donde:  $C$  representa la capacidad,  $k$  es la medida de crecimiento,  $m$  es el medida offset,  $\delta$  ;  $\gamma$  representan vectores de ajuste para la serie (Pontoh et al., 2021).

### **Ecuación 2**

*Modelo de crecimiento saturado*

$$g(t) = \frac{C}{1 + \exp(-k(t - m))'}$$

### **Ecuación 3**

*Modelo de crecimiento lineal*

$$g(t) = \frac{C(t)}{1 + \exp(-(k + \alpha(t)^T \delta)(t - (m + \alpha(t)^T \gamma)))'}$$

El componente estacional por su parte considera el efecto el efecto cíclico de los eventos, los cuales son modelados utilizando las series de Fourier y la ecuación (4) para funciones periódicas de  $t$ , proporcionando un ajuste adecuado y pronósticos precisos mediante el parámetro  $P$  (periodo regular) (Pontoh et al., 2021; Taylor & Letham, 2017).

### **Ecuación 4**

*Componente estacional de Prophet*

$$s(t) = \sum_{n=1}^N (a_n \cos\left(\frac{2\pi nt}{P}\right) + b_n \sin\left(\frac{2\pi nt}{P}\right))$$

El modelo SARIMA se define como una variante de ARIMA que incorpora capacidades adicionales para series de tiempo con fuertes patrones estacionales (Al-Shaikh et al., 2019), su configuración se presenta en la estructura de  $ARIMA(p, d, q)(P, D, Q)s$  donde  $p$  se define como el orden o ventana autorregresiva,  $d$  representa el número de diferenciación en la serie,  $q$  establece el intervalo del promedio móvil,  $P$  define la ventana autorregresiva estacional,  $D$  especifica la diferenciación de serie estacional,  $Q$  detalla el intervalo de la media móvil estacional y  $s$  determina el tiempo o ciclo estacional (Noor et al., 2022).

Según lo presentado por Guo et al. (2011) y M. K. Islam et al (2023), SARIMA puede expresarse matemáticamente mediante la Ecuación 5, la cual comparte características del modelo autorregresivo (AR), integrado (I) y media móvil (MA). Por tal motivo,  $\alpha$  representa la parte AR no estacional,  $B$  agrupa la parte AR estacional,  $\theta$  simboliza MA no estacional,  $\phi$  denota MA estacional y  $L$  actúa como operador de retraso o *backward shift operator* que controla el nivel retraso de la serie.

### Ecuación 5

*Estructura del modelo SARIMA*

$$\alpha_p(L)B_p(L^s)\nabla^d\nabla_m^D X_s = \theta_q(L)\phi_q(L^s)\varepsilon_s + a$$

Box y Jenkins (1976) mencionan que existen tres etapas para un correcto modelamiento de series temporales para modelos ARIMA y similares, a saber: identificación, estimación y diagnóstico. La primera etapa busca establecer, de forma preliminar, los valores correspondientes de la estructura  $(p, d, q)(P, D, Q)$  mediante el use de la autocorrelación (*auto correlation function* – ACF) y autocorrelación parcial (*partial auto correlation function* – PACF). La segunda y tercera etapa se fundamentan en el modelamiento, cálculo de los parámetros y evaluación del nivel de ajuste final sobre los datos (Eze et al., 2020).

El uso de los gráficos ACF y PACF permiten determinar los valores correspondientes a los parámetros  $q$  y  $p$  respectivamente debido a que ACF mide la dependencia lineal entre observaciones dado un desfase  $q$  y PACF permite establecer la cantidad necesaria de términos autorregresivos (Chang et al., 2012). La prueba estadística Augmented Dickey Fuller (ADF) (Dickey & Fuller, 1979), evalúa la estacionariedad mediante una hipótesis nula ( $H_0$ ) de no estacionariedad de los datos y un criterio de análisis para rechazar ( $H_0$ ) en el caso donde p-value es inferior a un nivel de significancia ( $\alpha$ ) (Eze et al., 2020), y cumplir con el requerimiento de estacionariedad (Milenković et al., 2016).

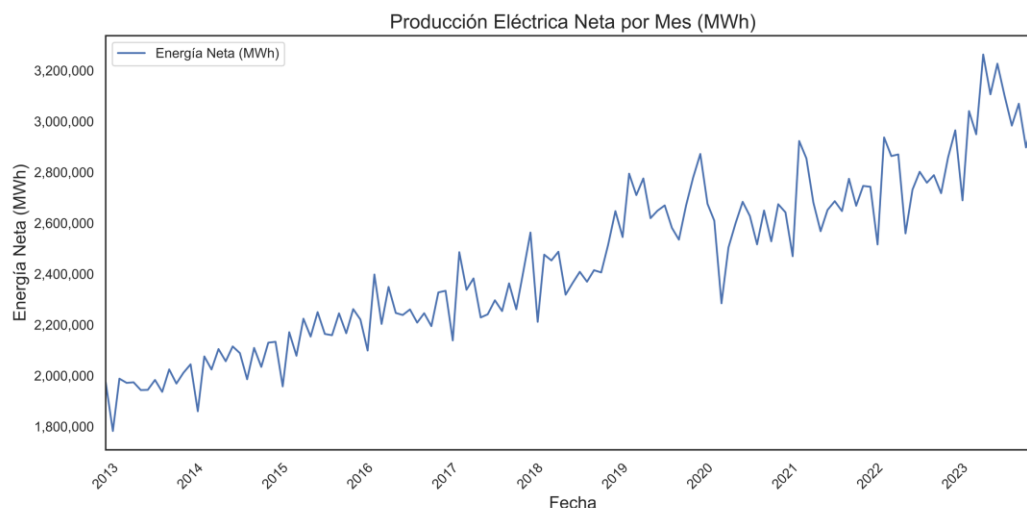
## RESULTADOS

### Análisis mediante estadística descriptiva de las series temporales

En la Figura 3 se evidencia una tendencia de crecimiento sostenido en el intervalo de tiempo estudiado con una importante estructura estacional caracterizada en los últimos meses, donde la producción total disminuye con respecto a los valores presentados en meses anteriores, en los años más recientes se evidencian fluctuaciones más pronunciadas y posibles cambios en la infraestructura.

**Figura 3**

*Serie total de la producción eléctrica neta por mes*



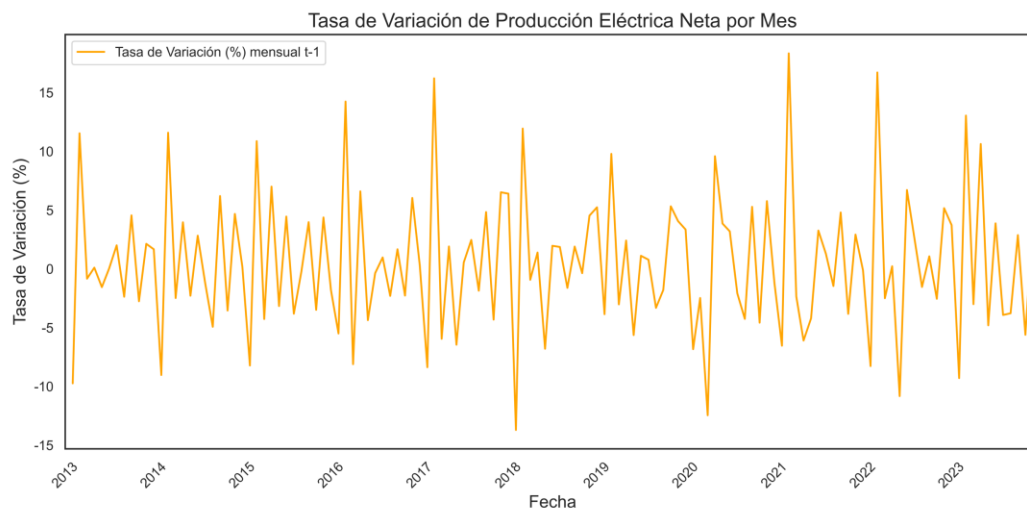
Fuente: Elaboración propia.

Las estadísticas descriptivas indica que en el promedio de producción eléctrica neta es de aproximadamente 2,445,124.21 MWh con una desviación estándar de 332,670.59 MWh para el período 2013 – 2023. Además, se ha evidenciado que el máximo histórico se presenta en mayo del año de 2023 con una producción total de 3,262,864.17 MWh y un mínimo histórico de 1,782,461.24 MWh en febrero del 2013.

Para profundizar el análisis se procedió a realizar el cálculo de la variación mensual de la producción total (t-1) para determinar el comportamiento de la generación eléctrica de una forma más detallada como se presenta en la Figura 4. De manera preliminar se observa un patrón cíclico, pero no la presencia de una tendencia clara que indique el aumento o disminución de las variaciones.

**Figura 4**

*Tasa de variación mensual de la producción eléctrica*



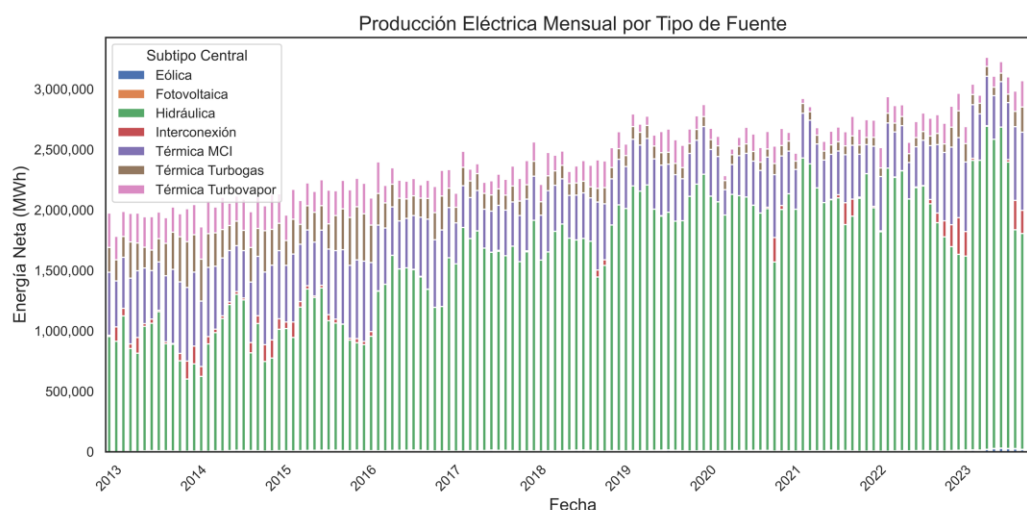
Fuente: Elaboración propia.

Estructuralmente la serie muestra una variación promedio de 0.4923% con una desviación estándar 5.9349%, es decir, la generación eléctrica entre meses no fluctúa de manera considerable, pero cuando lo hace, presenta un importante valor de aproximadamente el 6% de fluctuación sobre la producción del mes anterior. Históricamente se observa que el mínimo histórico de variación se da entre los meses febrero y enero del año 2018 con un valor de -13.7026% y un máximo entre los meses marzo y febrero del 2021 con 18.3450%. Según Salameh y Safari (1992), las variaciones de producción eléctrica se ven afectadas por factores que modifican directamente los resultados obtenidos de cada fuente como la velocidad media del viento, las características de la fuente, precipitaciones o lluvias (Barzola-Monteses et al., 2019), entre otros.

Desde otra perspectiva, la evolución mensual de generación eléctrica en Ecuador puede ser abordada en torno a la composición o aporte de cada fuente al Sistema Nacional Interconectado (Figura 5). La generación eléctrica mediante centrales hidráulicas pose un rol muy importante para la sostenibilidad energética del país frente a otras fuentes como la generación térmica (MCI, Turbogas y Turbovapor) que poco a poco han ido reduciendo su participación en el esquema nacional de producción y oferta de energía.

**Figura 5**

*Composición de la producción eléctrica mensual por tipo de fuente*

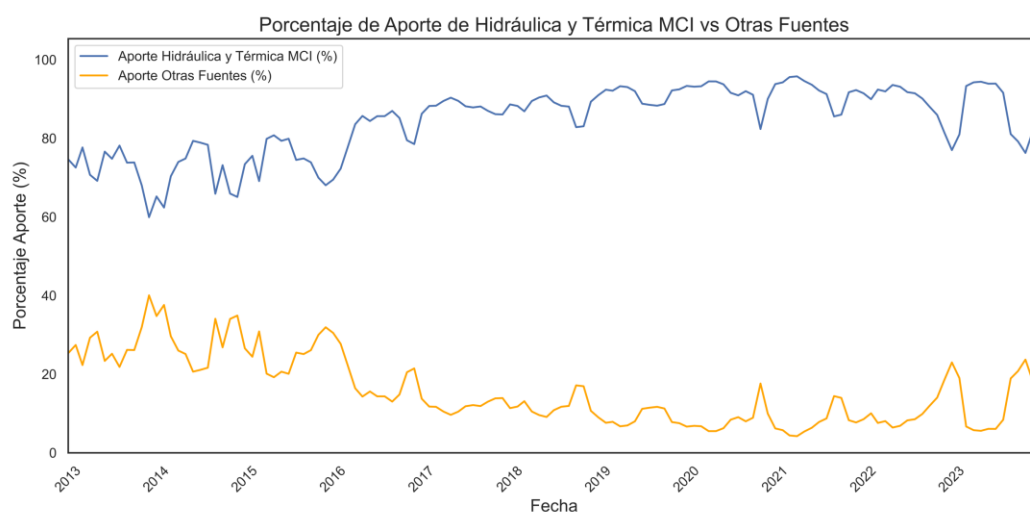


Fuente: Elaboración propia.

De forma complementaria, la Figura 6 expone que el porcentaje de aporte de estas dos series en conjunto aportan con aproximadamente el 80% del total de la energía neta producida. Esto se sustenta gracias a las estadísticas descriptivas generadas de serie, la cual reporta valores de aporte promedio de 84,2969% con una desviación con respecto a la media de 8,7637%, el máximo histórico se da en marzo del 2021 con un porcentaje de aporte registro un 95,7698% y un mínimo de 59,9389% en diciembre del 2013 para el tiempo de análisis seleccionado.

**Figura 6**

*Porcentaje de aporte de las series Hidráulica y Térmica MCI*



Fuente: Elaboración propia.

### **Análisis predictivo mediante SARIMA**

Para empezar con el análisis y modelado estadístico se procedió a realizar una prueba de estacionariedad para las series seleccionadas utilizando el test ADF para validar la existencia de este requerimiento. Como se presenta en la Tabla 6 y bajo una significancia del 5%, permitió determinar que las series en su estado original no poseen estacionariedad, por lo tanto, se procedió a diferenciarlas en un período permitió la estabilización de las series y se validó su relevancia estadística.

**Tabla 6**

*Resultados del test ADF para las series de tiempo*

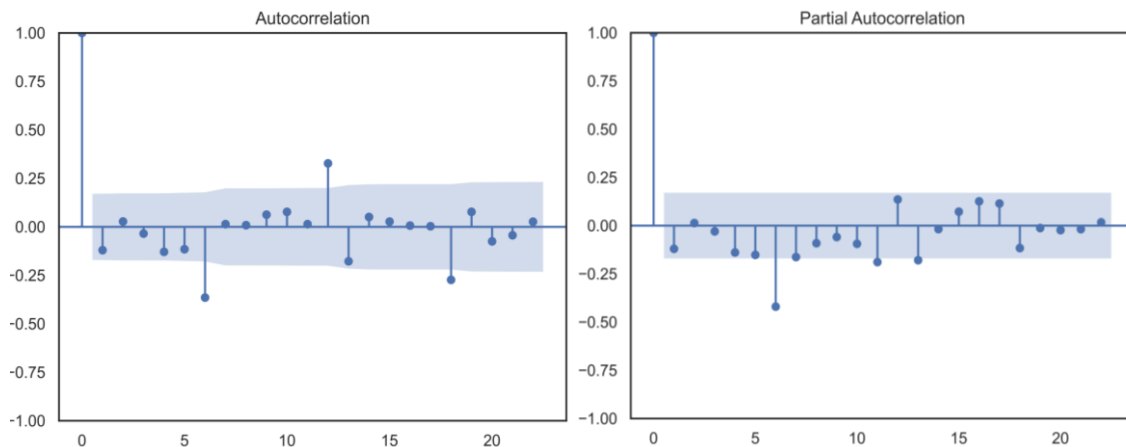
<b>Serie</b>	<b>Estadístico</b>	<b>p-valor</b>	<b>Resultado</b>
Hidráulica (d=0)	-1.7892	0.3857	No estacionario
Hidráulica (d=1)	-5.4622	2.50341e-06	Estacionario
Térmica MCI (d=0)	-1.5555	0.5058	No estacionario
Térmica MCI (d=1)	-7.4883	4.5754e-11	Estacionario

Fuente: Elaboración propia.

Una vez validado el requerimiento de estacionariedad se procedió a realizar el análisis de ADF y PACF. Para el caso de la serie “Hidráulica” (Figura 7), se observa un alto nivel de correlación en el primer rezago y la presencia de rezagos estadísticamente relevantes en los puntos (6,11,17), convirtiéndose en posibles candidatos para el parámetro  $q$ . Para el caso de PACF se reporta un comportamiento similar, alta correlación en el primer rezago y posibles candidatos para el parámetro  $p$  en los puntos (6,11,13).

**Figura 7**

*Gráfico de autocorrelación y autocorrelación parcial para la serie Hidráulica*



Fuente: Elaboración propia.

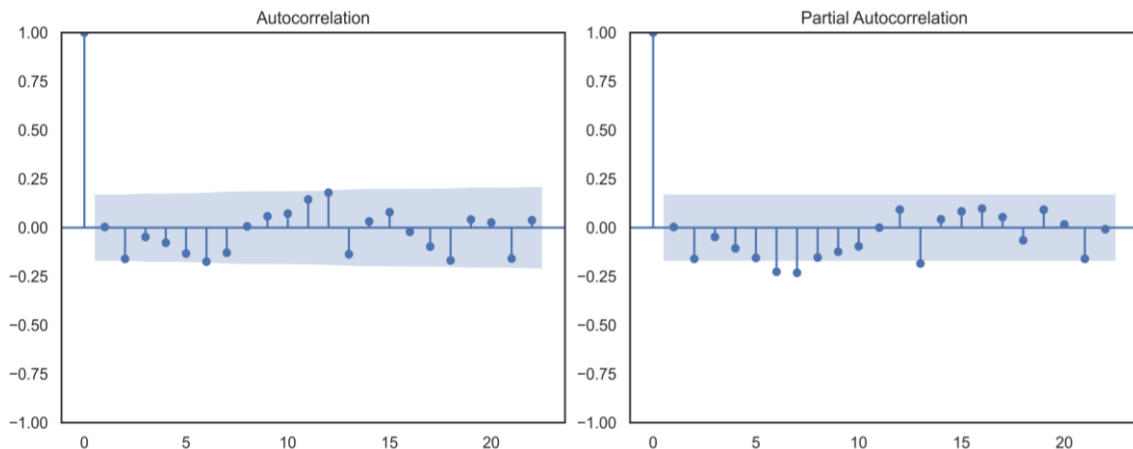
De forma análoga, se analizó la serie temporal “Térmica MCI” utilizando los gráficos ACF y PACF (Figura 8) para determinar las correlaciones de interés, donde se determinó que los posibles valores para el parámetro  $q$  se presentan posiblemente en (2,6,11) para autocorrelación y (2,5,6,7,13) para autocorrelación parcial (parámetro  $p$ ).

El uso de las gráficas ACF y PACF se justifica no solo a través de su aplicación rutinaria en el modelado de series temporales, el aspecto empírico y la distribución teórica los han convertido en un elemento fundamental en la inferencia estadística para un correcto análisis de series temporales (Hassani et al., 2024).



**Figura 8**

*Gráfico de autocorrelación y autocorrelación parcial para la serie Térmica MCI*

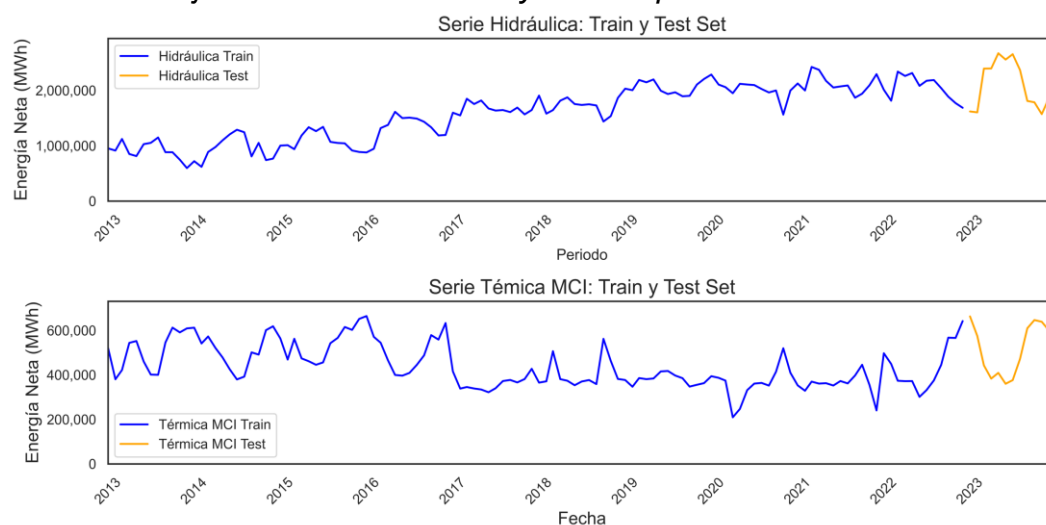


Fuente: Elaboración propia.

Posteriormente se aplicó una división o *split* de las series temporales de acuerdo con lo implementado por Feng et al. (2022), M. K. Islam et al. (2023; Kouvara & Vogiatzis, 2023), al excluir específicamente el último año como conjunto de pruebas (*test*) y utilizar los años restantes como conjunto de entrenamiento (*train*). Como se observa en la Figura 9, el train set se compone de los registros mensuales desde 2013 hasta 2022 y el train set de los 12 meses del año 2023.

**Figura 9**

*Division del conjunto de datos en train y test set para la series*



Fuente: Elaboración propia.

Para el modelado se utilizó la librería *pmdarima* (versión 2.0.4) y el objeto *auto\_arima* para realizar un proceso de búsqueda exhaustiva de los parámetros del modelo y retornar aquella combinación que reporta el mejor ajuste de los datos medido a través del valor más bajo de AIC (*Akaike Information Criterion*). Ulteriormente, se modeló la combinación final mediante la librería *statsmodels* (versión 0.14.1) para ambas series estudiadas y generar las tablas de resumen y gráficos de resumen.

Para la serie “Hidráulica” se ha determinado que el mejor modelo (menor AIC) se encuentra compuesto por ARIMA (1,1,1)(2,0,1)<sub>12</sub>, que contempla un ajuste de orden no estacional (AR(1), I(1), MA(1)) y estacional (AR(2), I(0), MA(1)) con una frecuencia mensual (12). Al analizar la Tabla 7, se reporta que el ajuste del modelo posee un AIC de 3191.765 y un total de observaciones de 120, la significancia de los coeficientes *ar.L1*, *ma.L1*, *ar.S.L12* y *ma.S.L12* indican que el modelo se encuentra bien ajustado y que la relación estacional reportada es notable y muestra patrones anuales claros. Al analizar *ar.S.L24* se observa que la relación con el segundo año anterior no es fuerte y la serie posee una mayor relación con los valores del periodo directamente anterior y con el mismo periodo de un año atrás.

**Tabla 7**

*Estructura del modelo para la serie Hidráulica*

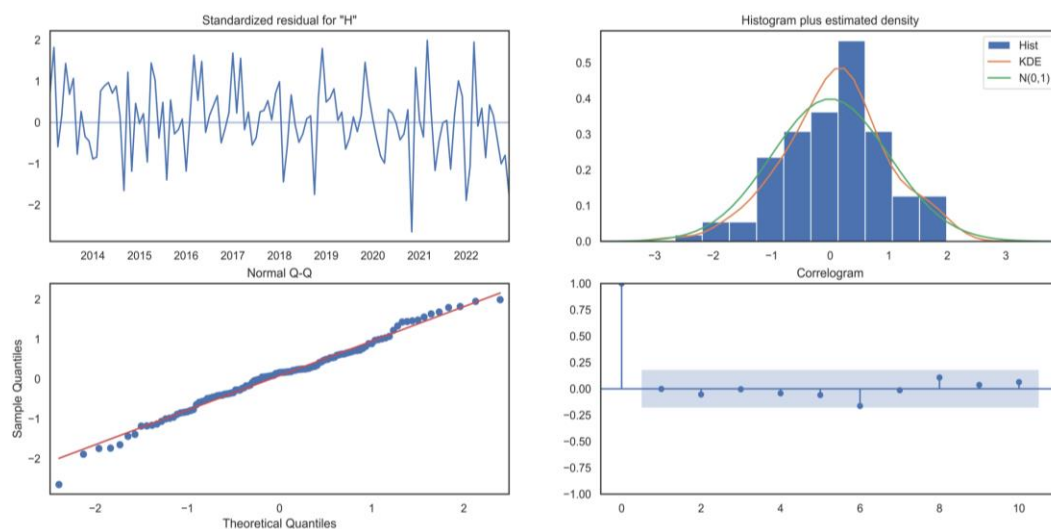
Dep. Variable	Hidráulica	No. Observaciones	120	
Modelo	(1,1,1)(2,0,1) <sub>12</sub>	AIC	3191.765	
	Coefficiente	Error Estándar	z	P> z
ar.L1	0.5903	0.181	3.269	0.001
ma.L1	-0.8711	0.119	-7.346	0.000
ar.S.L12	0.7467	0.305	2.448	0.014
ar.S.L24	0.0787	0.184	0.427	0.670
ma.S.L12	-0.5778	0.288	-2.007	0.045
sigma	2.923e+10	9.03e-12	3.24e+21	0.000

Fuente: Elaboración propia.

El diagnóstico del modelo y sus residuos se presentan en la Figura 10, donde se evidencia que su comportamiento se maneja entre -2 y +2 desviaciones estándar sin la evidencia de una tendencia visible. El histograma muestra que los residuos se aproximan a un comportamiento relativamente normal y la curva KDE (*Kernel Density Estimator*) se asemeja bastante a la normal teórica y el gráfico Q-Q se ajustan correctamente a los cuantiles teóricos normales. El correlograma evidencia que no existe autocorrelaciones significativas y los valores se encuentran entre las bandas de confianza.

**Figura 10**

*Residuos del modelo SARIMA aplicado a la serie Hidráulica*



Fuente: Elaboración propia.

Adicionalmente, se procedió a realizar la prueba Ljung-Box (1978) de autocorrelación de residuos mediante la hipótesis nula ( $H_0$ ) sobre ruido blanco (independencia de residuos), un criterio de análisis regla de decisión para rechazar ( $H_0$ ) si los resultados de p-value son inferiores a un nivel de significancia ( $\alpha$ ). Los resultados del test aplicado a los residuos (Tabla 8), evidencian que no existe un comportamiento de autocorrelación en los residuos en un total de 10 lags analizados.

**Tabla 8**

*Resultados de la prueba Ljung-Box para el modelo SARIMA de “Hidráulica”*

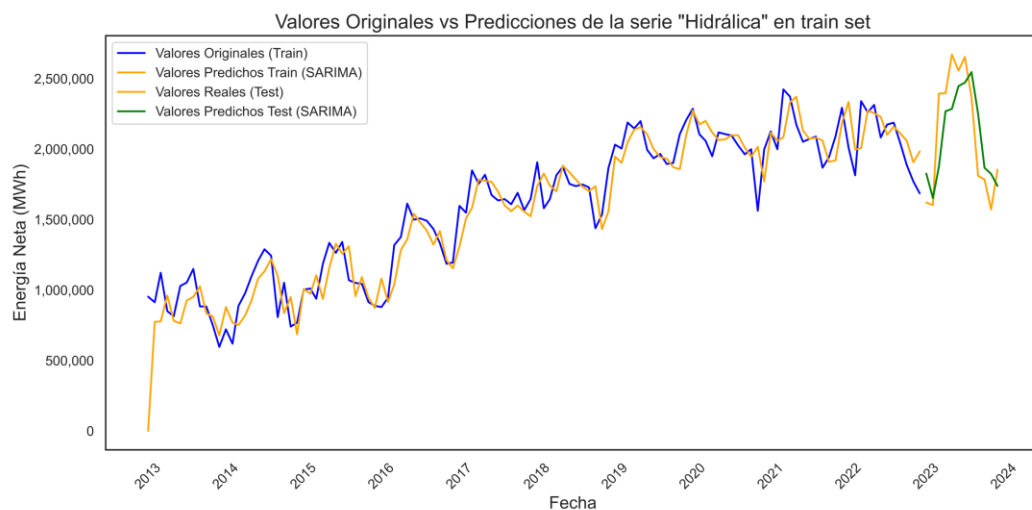
Test Ljung-Box (1978)			
Estadístico	Lags	p-valor	Resultado
2.5316	10	0.990	No autocorrelación

Fuente: Elaboración propia.

Al graficar los valores originales de la generación eléctrica en MWh para la serie “Hidráulica” (train y test set) frente a los valores predichos por el modelo SARIMA (Figura 11), se observa que su comportamiento es bastante similar durante todos los periodos analizados y es capaz de identificar los complejos patrones estacionales. El proceso iterativo de pronóstico utiliza los datos mensuales disponibles del año 2023 y realiza las estimaciones correspondientes a los nuevos valores ajustados a los movimientos más recientes de acuerdo con la configuración establecida anteriormente del modelo.

**Figura 11**

*Serie original y predicciones de la serie Hidráulica con SARIMA*



Fuente: Elaboración propia.

Para cuantificar estos errores se hará uso de las funciones de pérdida RMSE y MAPE considerando lo presentado por Lewis (1982) y Pontoh et al. (2021) con respecto a la categorización estándar de MAPE para evaluación de pronósticos

(Tabla 9). De esta manera se puede utilizar, de forma simultánea, una métrica numérica más comparable entre modelos predictivos en forma de categoría y una métrica en las mismas unidades que la serie analizada.

**Tabla 9**

*Categorización de MAPE para pronósticos*

<b>MAPE</b>	<b>Criterio de Pronóstico</b>
<10%	La precisión del modelo es muy elevada
10-20%	La precisión del modelo es buena
20-50%	La precisión del modelo es bastante buena
>50%	La precisión del modelo no es buena

Fuente: Pontoh et al. (2021).

Los hallazgos se muestran en la Tabla 10, donde la capacidad predictiva del modelo planteado posee una precisión muy elevada en cuanto al set de entrenamiento y disminuye ligeramente hasta alcanzar el umbral de la categoría buena para el set de prueba. Estos valores complementan los indicadores anteriores sobre la calidad del modelo y confirman un buen desempeño general, aunque podría existir ciertos indicios de *overfitting* o sobreajuste a los datos.

**Tabla 10**

*Resultados de las métricas de error para el modelo SARIMA en Hidráulica*

	<b>RMSE</b>	<b>Valor</b>	<b>MAPE</b> <b>Categoría</b>
Hidráulica (Train)	173867.70	9.27%	Elevada precisión
Hidráulica (Test)	263709.84	10.60%	Buena precisión

Fuente: Elaboración propia.

El modelado para la segunda serie “Térmica MCI” siguió un proceso similar al aplicado para el caso “Hidráulica”. Luego de haber determinado los posibles candidatos para los parámetros del modelo se utilizó `auto_arima` para la búsqueda exhaustiva entre combinaciones y encontrar aquella con el menor AIC.

De forma análoga, la estructura del modelo se presenta en la Tabla 11, donde el proceso automático determinó que la mejor combinación para esta serie se da en un modelo ARIMA (1,1,1)(0,0,0)12, donde se identificó un orden no estacional (AR(1), I(1), MA(1)) y estacional (AR(0), I(0), MA(0)) con una frecuencia mensual (12). Esto implica que la mejor combinación de los parámetros deriva en un modelo que no considera términos estacionales y se rige exclusivamente por patrones autorregresivos y de media móvil. La significancia de los coeficientes ar.L1 y ma.L1 indican que existe una relación positiva con respecto al mes anterior y el error de predicción es corregido correctamente e influye en el valor actual.

**Tabla 11**

*Estructura del modelo para la serie Térmica MCI*

<b>Dep. Variable</b>	Térmica MCI	<b>No. Observaciones</b>	120	
<b>Modelo</b>	(1,1,1)	<b>AIC</b>	2985.704	
	<b>Coefficiente</b>	<b>Error Estándar</b>	<b>z</b>	<b>P&gt; z </b>
ar.L1	0.6966	0.121	5.776	0.000
ma.L1	-0.8951	0.101	-8.875	0.000
sigma	4.479e+09	5.47e-12	8.19e+20	0.000

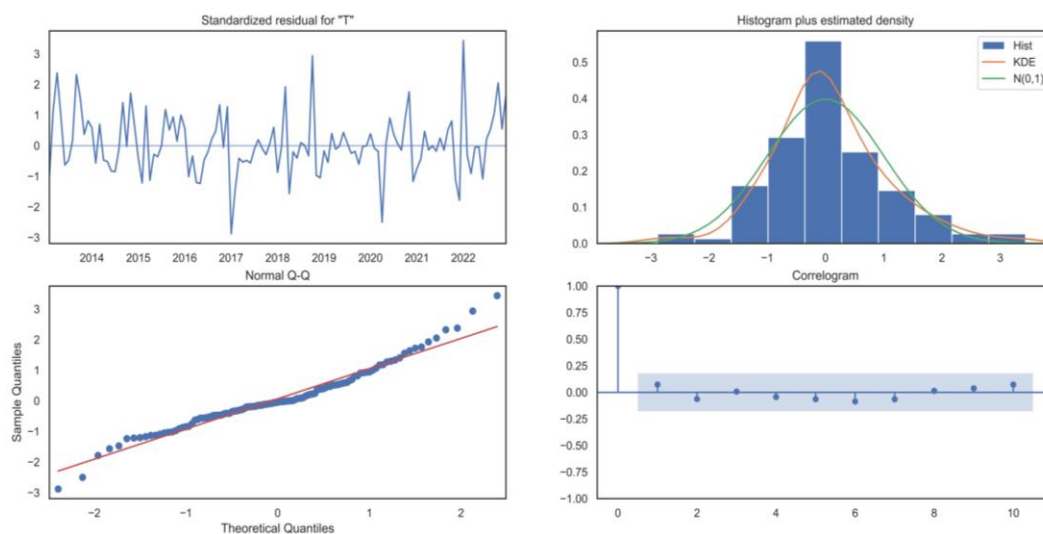
Fuente: Elaboración propia.

El análisis de residuos se presenta en la Figura 12, donde su comportamiento se mantiene entre -3 y +3 desviaciones estándar sin un patrón visible. El histograma y la proyección de KDE muestran una distribución relativamente normal, aunque existe la presencia de ciertas desviaciones o ligeras asimetrías. El gráfico Q-Q muestra un ajuste regular con los cuantiles teóricos normales y el correlograma indica la presencia de no autocorrelación significativa.

Los resultados del test Ljung-Box aplicado a los residuos del modelo para la serie “Térmica MCI” se presentan en la Tabla 12 y se concluye los residuos no presentan autocorrelación al rechazar rechazar la hipótesis nula  $H_0$ .

**Figura 12**

*Residuos del modelo SARIMA aplicado a la serie Térmica MCI*



Fuente: Elaboración propia.

**Tabla 12**

*Resultados de la prueba Ljung-Box para el modelo SARIMA de "Térmica MCI"*

Test Ljung-Box (1978)			
Estadístico	Lags	p-valor	Resultado
6.3741	10	0.782	No autocorrelación

Fuente: Elaboración propia.

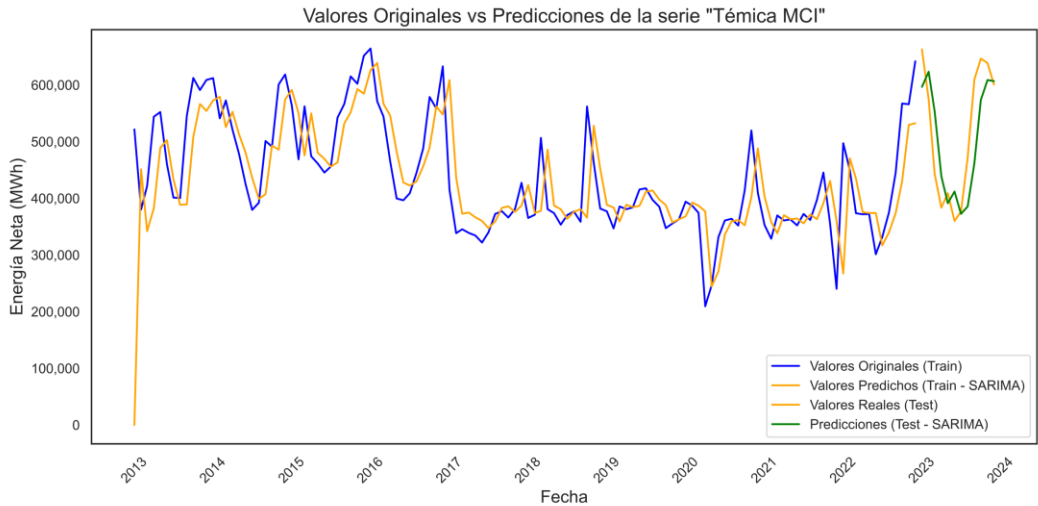
La grafica comparativa de los originales y las predicciones realizadas con SARIMA en la serie "Térmica MCI" (Figura 13) muestran que el modelo reporta un alto nivel de precisión, no obstante, se evidencia ciertas diferencias importantes en los meses iniciales de cada periodo, esto presenta posible evidencia de que el modelo capturó el patrón general de la serie, pero existe cierto nivel de dificultad en la magnitud de los valores predichos.

De acuerdo con las métricas de error reportadas en la Tabla 13, el modelo generación eléctrica por "Térmica MCI" ha generado predicciones con una calidad inferior a los reportados para el caso "Hidráulica" en términos de MAPE dentro del rango aceptable ("Buena Precisión"). Además, se destaca una

disminución en la métrica RMSE, indicando que no existe, preliminarmente, un sobreajuste del modelo en entrenamiento.

Figura 13

Serie original y predicciones de la serie Térmica MCI con SARIMA



Fuente: Elaboración propia.

Tabla 13

Resultados de las métricas de error para el modelo SARIMA en Térmica MCI

	RMSE	MAPE	
		Valor	Categoría
Térmica MCI (Train)	81372.86	11.81%	Buena precisión
Térmica MCI (Test)	71026.93	11.43%	Buena precisión

Fuente: Elaboración propia.

Análisis predictivo mediante Prophet

En cuanto al análisis utilizando la librería Prophet, se diseñó una función de búsqueda exhaustiva o grid search para optimizar los tiempos de búsqueda en la selección de los mejores hiperparámetros. En este sentido, se planteó un total de 4 hiperparámetros principales que tienen una relación directa con respecto al comportamiento general del modelo y la forma de análisis para el tratamiento de estacionalidades, estos son: 'changepoint\_prior\_scale', 'changepoint\_range',



'seasonality\_prior\_scale' y 'seasonality\_mode'. El proceso de optimización utiliza RMSE como función de pérdida y retorna aquella combinación menor valor de RMSE, es decir, el modelo que mejor se adaptó al comportamiento de los datos. Los resultados optimizados para la serie “Hidráulica” se presentan en la Tabla 14, los hiperparámetros indican que esta modelo tiene una elevada sensibilidad a los cambios y permite un ajuste en los cambios de tendencias con un enfoque más hacia la parte inicial de la serie. Además, se evidencia que se espera una variabilidad considerable para la captura de patrones estacionales y una estacionalidad caracterizada por un efecto multiplicativo.

**Tabla 14**

*Hiperparámetros del modelo Prophet para la serie Hidráulica*

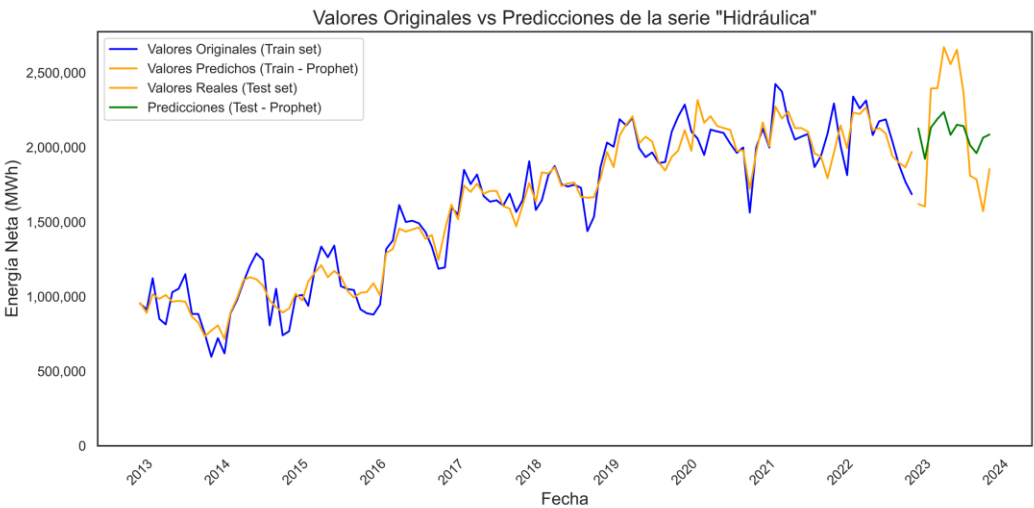
<b>Parámetro</b>	<b>Valor</b>
changepoint_prior_scale	1.9999999999999998
changepoint_range	0.7000000000000001
seasonality_prior_scale	1.7999999999999998
seasonality_mode	'multiplicative'

Fuente: Elaboración propia.

La representación visual del ajuste para el modelo Prophet en la serie “Hidráulica” se presenta en la Figura 14. Inicialmente se observa un muy buen ajuste en los datos de entrenamiento con predicciones bastante cercanas a los datos reales, luego el rendimiento predictivo del sistema disminuye y se complica con el importante cambio tendencial en los periodos del año 2013.

Esto se evidencia en las métricas de error RMSE y MAPE (Tabla 15), donde se observa un buen rendimiento para la serie de entrenamiento y las predicciones permiten asignar la categoría de “Elevada precisión”. Para el caso del set de prueba, el rendimiento cae sustancialmente a un valor de MAPE de 16,58% y desciende a la categoría de “Buena precisión”, esto se refleja también un aumento sustancial del valor RMSE.

**Figura 14**  
*Serie original y predicciones de la serie Hidráulica con Prophet*



Fuente: Elaboración propia.

**Tabla 15**  
*Resultados de las métricas de error para el modelo Prophet en Hidráulica*

	RMSE	MAPE	
		Valor	Categoría
Hidráulica (Train)	115216.24	6.48%	Elevada precisión
Hidráulica (Test)	360062.63	16,58%	Buena precisión

Fuente: Elaboración propia.

Del mismo modo, los resultados para la serie “Térmica MCI” se presentan en la Tabla 16, donde la configuración del modelo y sus hiperparámetros indican una elevada sensibilidad a los cambios y habilitan un ajuste relativamente estable de tendencias para la mayor parte de los datos. Los hiperparámetros ‘seasonality\_prior\_scale’ y ‘seasonality\_mode’ indican que existe una flexibilidad moderada para la captura de patrones estacionales y una magnitud de efectos estacionales que varían con el nivel de la serie.

**Tabla 16***Hiperparámetros del modelo Prophet para la serie Térmica MCI*

<b>Parámetro</b>	<b>Valor</b>
changepoint_prior_scale	1.9999999999999998
changepoint_range	0.9
seasonality_prior_scale	0.8999999999999999
seasonality_mode	'multiplicative'

Fuente: Elaboración propia.

Utilizando las métricas de rendimiento (Tabla 17) se observa que el modelo ha reportado un nivel de precisión muy elevado para la serie “Térmica MCI” en términos de MAPE y comparte lo evidenciado en el caso de la serie “Hidráulica” al reportar una disminución en la capacidad predictiva al presentar un RSME muy elevado y un cambio en la categoría de precisión.

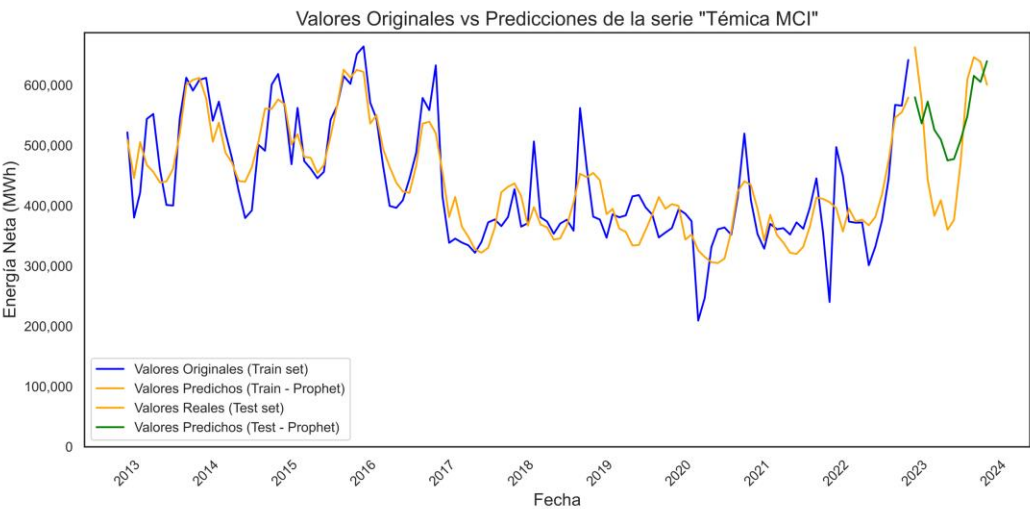
**Tabla 17***Resultados de las métricas de error para el modelo Prophet en Térmica MCI*

	<b>RMSE</b>	<b>MAPE</b>
	<b>Valor</b>	<b>Categoría</b>
Térmica MCI (Train)	47911.59	9.35% Elevada precisión
Térmica MCI (Test)	85568.87	16.96% Buena precisión

Fuente: Elaboración propia.

Finalmente, se presentan las series originales y predicciones realizadas por el modelo Prophet para la serie “Térmica MCI” en la Figura 15. El ajuste inicial en los datos de entrenamiento indica un nivel adecuado en la identificación de tendencias a lo largo de la serie, pero su comportamiento disminuye notablemente en la transición para el período 2023 debido al abrupto cambio que existe en el mes de enero.

**Figura 15**  
*Serie original y predicciones de la serie Térmica MCI con Prophet*



Fuente: Elaboración propia.

A modo de resumen, se presenta los resultados combinados de las métricas de evaluación para los dos modelos y cada uno de los conjuntos de datos (Tabla 18). La comparación entre los resultados muestra una importante mejora en las capacidades predictivas de Prophet sobre SARIMA en entramiento, evidenciando además un comportamiento similar entorno a la reducción en la precisión para ambos modelos en el set de prueba.

**Tabla 18**  
*Resumen de las métricas de rendimiento para los modelos estudiados*

	SARIMA		Prophet	
	RMSE	MAPE	RMSE	MAPE
Hidráulica (Train)	173867.70	9.27%	115216.24	6.48%
Hidráulica (Test)	263709.84	10.60%	360062.63	16,58%
Térmica MCI (Train)	81372.86	11.81%	47911.59	9.35%
Térmica MCI (Test)	71026.93	11.43%	85568.87	16.96%

Fuente: Elaboración propia.

## DISCUSIÓN DE LOS RESULTADOS Y PROPUESTA DE SOLUCIÓN

Considerando los resultados presentados en el apartado anterior, es evidente que el sector energético del Ecuador en términos de producción neta ha experimentado un incremento sostenido desde el 2013 hasta el 2023. Sin embargo, la creciente dependencia de la generación hidráulica ha derivado en un mayor nivel sensibilidad ante cambios o fluctuaciones en las fuentes que alimentan las centrales hidroeléctricas y cuyo comportamiento se refleja en el agregado nacional de producción, con patrones cíclicos cada vez más evidentes.

En cuanto a la modelización de series temporales mediante las metodologías SARIMA y Prophet, estas presentan importantes capacidades predictivas al reportar niveles de errores que, de acuerdo con la escala categórica de MAPE, permiten ubicar la calidad de los modelos entre “elevada precisión” y “buena precisión” en los respectivos conjuntos de entrenamiento y prueba a pesar de una compleja transición en los meses del año 2023.

Por una parte, el análisis estadístico tradicional de SARIMA requirió la diferenciación de un periodo para ambas series con el objetivo de estabilizarlas y lograr un nivel de aceptable de estacionariedad. El análisis de los gráficos ACF y PACF permitió identificar los posibles valores para los parámetros necesarios en la configuración del modelo en ambos casos para las series “Hidráulica” y “Térmica MCI”. Utilizando las métricas de rendimiento RMSE y MAPE se estableció que existe un comportamiento más preciso sobre las predicciones realizadas en la serie “Hidráulica” para el train set y una ligera disminución en el test set. De forma similar, el rendimiento para las predicciones del modelo en “Térmica MCI” se caracterizó por una “buena precisión” con valores MAPE aceptables.

Por otra parte, el modelado con la metodología “Prophet” se caracterizó por un proceso más simplificado con la consideración de que es computacionalmente más exigente, al requerir la optimización de múltiples hiperparámetros de forma simultánea. Los resultados de rendimiento evidenciaron una disminución

considerable en los valores MAPE para los sets de entrenamiento y se vieron incrementados al momento de utilizar los datos de prueba tanto en términos de RMSE como de MAPE, aunque su precisión todavía se encuentra dentro los intervalos permitidos para ser calificada como “Buena precisión”.

Consecuentemente, la propuesta de solución se ha encuentra formada por el uso de un modelo Prophet para la estimación de series temporales que, a pesar de presentar ciertas dificultades para generalizar en nuevos datos, la capacidad de “personalización” y ajuste mediante los hiperparámetros la convierten una opción altamente viable para predecir los comportamientos de la producción eléctrica por tipo de fuente a nivel nacional.

Es importante considerar también que el rendimiento de la metodología estadística tradicional no es despreciable, pues a pesar de los constantes avances tecnológicos, el modelado estadístico con supuestos o restricciones ha permanecido vigente y con resultados igual de válidos que los sistemas actuales, ofreciendo una fuente de comparación sólida para inferencia estadística en series de tiempo.

Las implicaciones de estos resultados hacia una aplicabilidad real se presentan a través desde una perspectiva pública y privada. En primer lugar, la utilización del modelo Prophet puede mejorar la planificación energética en torno a infraestructura y política energética a largo plazo, los operadores energéticos pueden optimizar la administración de recursos eléctricos e influenciar positivamente para la construcción de políticas de transición energética.

En segundo lugar, las predicciones de Prophet puede aportar en el sector privado al permitir establecer estrategias de inversión para la compra de plantas eléctricas, planificación de la producción para periodos donde existe restricción energética y una optimización del consumo energético general para el ahorro de recursos.

## CONCLUSIONES Y RECOMENDACIONES

El presente documento de investigación realizó un análisis y modelamiento de series temporales para la generación eléctrica neta del Ecuador por tipo origen en base a dos metodologías ampliamente utilizadas en la literatura académica. El análisis descriptivo permitió presentar la situación actual de sector eléctrico y su importante nivel de dependencia en los sistemas hidroeléctricos y la poca relevancia de otras fuentes.

La construcción de los modelos predictivos y la posterior evaluación de su rendimiento ha permitido observar dos opciones altamente válidas para la generación de pronósticos con alta precisión. Tanto Prophet como SARIMA reportaron rendimientos muy elevados para la calidad de predicciones en los sets de entrenamiento y prueba para ambas series analizadas mediante RMSE y MAPE.

Sin embargo, el nivel de ajuste sobre los datos y extensa manipulación sobre la configuración del modelo mediante los hiperparámetros, convierten a Prophet en una herramienta altamente robusto frente a otras opciones similares, esto permite combinar el conocimiento del analista y facilitar su adopción a diferentes casos de aplicación en con diferentes tendencias, estacionalidades y efectos.

Como recomendaciones para estudios posteriores es necesario mencionar la introducción de otros hiperparámetros con el objetivo de optimizar el modelo a la serie de tiempo a través de una búsqueda exhaustiva mediante combinatorias. Además, se podrían explorar otras opciones en cuanto al tratamiento previo de la serie temporal al implementar transformaciones logarítmicas o similares y evaluar el desempeño de las metodologías predictivas. Finalmente, se podrían explorar otras opciones de modelación cuya complejidad es más elevada, aunque podrían presentar interesantes resultados sobre las series de generación eléctrica como: redes neuronales, árboles de decisión y algoritmos con descenso de gradiente (xgboost, catboost, lightgbm, entre otros).

## REFERENCIAS

- Aditya Satrio, C. B., Darmawan, W., Nadia, B. U., & Hanafiah, N. (2021). Time series analysis and forecasting of coronavirus disease in Indonesia using ARIMA model and PROPHET. *Procedia Computer Science*, 179, 524–532. <https://doi.org/10.1016/j.procs.2021.01.036>
- Agencia de Regulación y Control de Electricidad. (2024). Reportes de Información Estadística del Sector Eléctrico. <http://reportes.controlrecursosyenergia.gob.ec/>
- Ahmad, T., Zhang, H., & Yan, B. (2020). A review on renewable energy and electricity requirement forecasting models for smart grid and buildings. In *Sustainable Cities and Society* (Vol. 55). Elsevier Ltd. <https://doi.org/10.1016/j.scs.2020.102052>
- Al-Shaikh, H., Rahman, Md. A., & Zubair, A. (2019). Short-Term Electric Demand Forecasting for Power Systems using Similar Months Approach based SARIMA. 2019 IEEE International Conference on Power, Electrical, and Electronics and Industrial Applications (PEEIACON), 122–126. <https://doi.org/10.1109/PEEIACON48840.2019.9071939>
- Arias, C. (2024). Repositorio de Github. [https://github.com/1-echo/Maestria\\_UDLA](https://github.com/1-echo/Maestria_UDLA)
- Atasever, S., Öztürk, B., & Bilgiç, G. (2022). A new approach to short-term wind speed prediction: the prophet model. *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*, 44(4), 8826–8841. <https://doi.org/10.1080/15567036.2022.2126035>
- Barzola-Monteses, J., Mite-León, M., Espinoza-Andaluz, M., Gómez-Romero, J., & Fajardo, W. (2019). Time series analysis for predicting hydroelectric power production: The ecuador case. *Sustainability (Switzerland)*, 11(23). <https://doi.org/10.3390/su11236539>
- Bitirgen, K., & Filik, Ü. B. (2020). Electricity Price Forecasting based on XGBooST and ARIMA Algorithms. *BSEU JOURNAL OF ENGINEERING RESEARCH AND TECHNOLOGY*, 1, 1–7. <https://orcid.org/0000-0002-4468-4905>



- Box, G. E. P., & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control* (3rd ed.). Holden-Day.
- Chadalavada, R. J., Raghavendra, S., & Rekha, V. (2020). Electricity requirement prediction using time series and Facebooks PROPHET. *Indian Journal of Science and Technology*, 13(47), 4631–4645. <https://doi.org/10.17485/IJST/v13i47.1847>
- Chang, X., Gao, M., Wang, Y., & Hou, X. (2012). Seasonal autoregressive integrated moving average model for precipitation time series. *Journal of Mathematics and Statistics*, 8(4), 500–505. <https://doi.org/10.3844/jmssp.2012.500.505>
- Debnath, K. B., & Mourshed, M. (2018). Forecasting methods in energy planning models. *Renewable and Sustainable Energy Reviews*, 88, 297–325. <https://doi.org/10.1016/j.rser.2018.02.002>
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. In *Source: Journal of the American Statistical Association* (Vol. 74, Issue 366).
- Eze, N. M., Asogwa, O. C., Obetta, A. O., Ojide, K. C., & Okonkwo, C. I. (2020). A Time Series Analysis of Federal Budgetary Allocations to Education Sector in Nigeria (1970-2018). *American Journal of Applied Mathematics and Statistics*, 8(1), 1–8. <https://doi.org/10.12691/ajams-8-1-1>
- Facebook. (2017). Prophet - Forecasting at scale. <https://facebook.github.io/prophet/>
- Feng, T., Zheng, Z., Xu, J., Liu, M., Li, M., Jia, H., & Yu, X. (2022). The comparative analysis of SARIMA, Facebook Prophet, and LSTM for road traffic injury prediction in Northeast China. *Frontiers in Public Health*, 10. <https://doi.org/10.3389/fpubh.2022.946563>
- Guo, Z., Zhao, J., Zhang, W., & Wang, J. (2011). A corrected hybrid approach for wind speed prediction in Hexi Corridor of China. *Energy*, 36(3), 1668–1679. <https://doi.org/10.1016/j.energy.2010.12.063>
- Hasnain, A., Sheng, Y., Hashmi, M. Z., Bhatti, U. A., Hussain, A., Hameed, M., Marjan, S., Bazai, S. U., Hossain, M. A., Sahabuddin, M., Wagan, R. A., & Zha, Y. (2022). Time Series Analysis and Forecasting of Air Pollutants Based

- on Prophet Forecasting Model in Jiangsu Province, China. *Frontiers in Environmental Science*, 10. <https://doi.org/10.3389/fenvs.2022.945628>
- Hassani, H., Marvian, L., Yarmohammadi, M., & Yeganegi, M. R. (2024). Unraveling Time Series Dynamics: Evaluating Partial Autocorrelation Function Distribution and Its Implications. *Mathematical and Computational Applications*, 29(4). <https://doi.org/10.3390/mca29040058>
- Huang, X., & Dong, L. (2022). Research on Power Material Demand Forecasting Strategy Based on Prophet Algorithm. *Journal of Physics: Conference Series*, 2404(1), 012043. <https://doi.org/10.1088/1742-6596/2404/1/012043>
- Islam, M. A., Che, H. S., Hasanuzzaman, M., & Rahim, N. A. (2020). Energy demand forecasting. In *Energy for Sustainable Development* (pp. 105–123). Elsevier. <https://doi.org/10.1016/B978-0-12-814645-3.00005-5>
- Islam, M. K., Hassan, N. M. S., Rasul, M. G., Emami, K., & Chowdhury, A. A. (2023). Forecasting of Solar and Wind Resources for Power Generation. *Energies*, 16(17), 6247. <https://doi.org/10.3390/en16176247>
- Klyuev, R. V., Morgoev, I. D., Morgoeva, A. D., Gavrina, O. A., Martyushev, N. V., Efremenkova, E. A., & Mengxu, Q. (2022). Methods of Forecasting Electric Energy Consumption: A Literature Review. *Energies*, 15(23), 8919. <https://doi.org/10.3390/en15238919>
- Kouvara, D., & Vogiatzis, D. (2023). Forecasting Residential Energy Consumption: A Case Study for Greece. *International Conference on Enterprise Information Systems, ICEIS - Proceedings*, 1, 484–492. <https://doi.org/10.5220/0011854500003467>
- Lange, S., Pohl, J., & Santarius, T. (2020). Digitalization and energy consumption. Does ICT reduce energy demand? *Ecological Economics*, 176. <https://doi.org/10.1016/j.ecolecon.2020.106760>
- Ledmaoui, Y., El Maghraoui, A., El Aroussi, M., Saadane, R., Chebak, A., & Chehri, A. (2023). Forecasting solar energy production: A comparative study of machine learning algorithms. *Energy Reports*, 10, 1004–1012. <https://doi.org/10.1016/j.egyr.2023.07.042>
- Lewis, C. (1982). *Industrial and Business Forecasting Methods*. Butterworths Publishing.

- Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303. <https://doi.org/10.1093/biomet/65.2.297>
- Lu, J., Yan, Z., Han, J., & Zhang, G. (2019). Data-Driven Decision-Making (D3M): Framework, Methodology, and Directions. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 3(4), 286–296. <https://doi.org/10.1109/TETCI.2019.2915813>
- Maji, I. K., & Saari, M. Y. (2021). Renewable Energy Consumption and Sectoral Based Output. *International Journal of Finance Research*, 2(3), 143–153. <https://doi.org/10.47747/ijfr.v2i3.339>
- Manikam, S., Sahibudin, S., & Kasinathan, V. (2019). Business intelligence addressing service quality for big data analytics in public sector. *Indonesian Journal of Electrical Engineering and Computer Science*, 16(1), 491–499. <https://doi.org/10.11591/ijeecs.v16.i1.pp491-499>
- Milenković, M., Švadlenka, L., Melichar, V., Bojović, N., & Avramović, Z. (2016). SARIMA modelling approach for railway passenger flow forecasting. *Transport*, 33(5), 1–8. <https://doi.org/10.3846/16484142.2016.1139623>
- Ministerio de Energías y Minas. (2022). Factor de emisión de CO2 del Sistema Nacional Interconectado de Ecuador. <https://www.rekursosyenergia.gob.ec/wp-content/uploads/2023/08/wp-1692720103183.pdf>
- Noor, T. H., Almars, A. M., Alwateer, M., Almaliki, M., Gad, I., & Atlam, E. S. (2022). SARIMA: A Seasonal Autoregressive Integrated Moving Average Model for Crime Analysis in Saudi Arabia. *Electronics (Switzerland)*, 11(23). <https://doi.org/10.3390/electronics11233986>
- Oliver, N. (2019). Governance in the Era of Data-driven Decision-making Algorithms. *Women Shaping Global Economic Governance*, 171–180. <https://doi.org/https://doi.org/10.18356/7f891b82-en>
- Ongena, G., & Davids, A. (2023). Big Data Analytics Capability and Governmental Performance: An Empirical Examination. *International Journal of Electronic Government Research*, 19(1). <https://doi.org/10.4018/IJEGR.321638>

- Párraga Palacios, Á. G., Intriago Rodríguez, S. A., Velasco Fuentes., E. D., Cedeño Quinto, V. M., Cedeño Quinto, V. M., Murillo Párraga, N. L., & Zambrano Gavilanes, F. E. (2019). Producción de energía eólica en Ecuador. *Ciencia Digital*, 3(3), 22–32. <https://doi.org/10.33262/cienciadigital.v3i3.610>
- Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Ben Taieb, S., Bergmeir, C., Bessa, R. J., Bijak, J., Boylan, J. E., Browell, J., Carnevale, C., Castle, J. L., Cirillo, P., Clements, M. P., Cordeiro, C., Cyrino Oliveira, F. L., De Baets, S., Dokumentov, A., ... Ziel, F. (2022). Forecasting: theory and practice. *International Journal of Forecasting*, 38(3), 705–871. <https://doi.org/10.1016/j.ijforecast.2021.11.001>
- Pinto, T., Praça, I., Vale, Z., & Silva, J. (2021). Ensemble learning for electricity consumption forecasting in office buildings. *Neurocomputing*, 423, 747–755. <https://doi.org/10.1016/j.neucom.2020.02.124>
- Pontoh, R. S., Zahroh, S., Nurahman, H. R., Aprillion, R. I., Ramdani, A., & Akmal, D. I. (2021). Applied of feed-forward neural network and facebook prophet model for train passengers forecasting. *Journal of Physics: Conference Series*, 1776(1). <https://doi.org/10.1088/1742-6596/1776/1/012057>
- Porse, E., Fournier, E., Cheng, D., Hirashiki, C., Gustafson, H., Federico, F., & Pincetl, S. (2020). Net solar generation potential from urban rooftops in Los Angeles. *Energy Policy*, 142. <https://doi.org/10.1016/j.enpol.2020.111461>
- Qin, P., Liu, M., Su, L., Fei, Y., & Tan-Soo, J.-S. (2022). Electricity consumption in the digital era: Micro evidence from Chinese households. *Resources, Conservation and Recycling*, 182, 106297. <https://doi.org/10.1016/j.resconrec.2022.106297>
- Salameh, Z. M., & Safari, I. (1992). Optimum windmill-site matching. *IEEE Transactions on Energy Conversion*, 7(4), 669–676. <https://doi.org/10.1109/60.182649>
- Samal, K. K. R., Babu, K. S., Das, S. K., & Acharaya, A. (2019). Time Series based Air Pollution Forecasting using SARIMA and Prophet Model. *Proceedings of the 2019 International Conference on Information*

- Technology and Computer Communications, 80–85.  
<https://doi.org/10.1145/3355402.3355417>
- Samsó, R., Crespin, J., García-Olivares, A., & Solé, J. (2023). Examining the Potential of Marine Renewable Energy: A Net Energy Perspective. *Sustainability (Switzerland)*, 15(10). <https://doi.org/10.3390/su15108050>
- Sievers, L., Breitschopf, B., Pfaff, M., & Schaffer, A. (2019). Macroeconomic impact of the German energy transition and its distribution by sectors and regions. *Ecological Economics*, 160, 191–204.  
<https://doi.org/10.1016/j.ecolecon.2019.02.017>
- Stern, D. I., Burke, P. J., & Bruns, S. B. (2019). Energy and Economic Growth  
 Title The Impact of Electricity on Economic Development: A Macroeconomic Perspective. <https://escholarship.org/uc/item/7jb0015q>
- Taylor, S. J., & Letham, B. (2017). Forecasting at Scale.  
<https://doi.org/10.7287/peerj.preprints.3190v2>
- Toharudin, T., Pontoh, R. S., Caraka, R. E., Zahroh, S., Lee, Y., & Chen, R. C. (2023). Employing long short-term memory and Facebook prophet model in air temperature forecasting. *Communications in Statistics: Simulation and Computation*, 52(2), 279–290.  
<https://doi.org/10.1080/03610918.2020.1854302>
- Ünlü, R. (2019). A Comparative Study of Machine Learning and Deep Learning for Time Series Forecasting: A Case Study of Choosing the Best Prediction Model for Turkey Electricity Production. *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 23(2), 635–646.  
<https://doi.org/10.19113/sdufenbed.494396>
- Wang, C., Cao, X., Mao, J., & Qin, P. (2019). The changes in coal intensity of electricity generation in Chinese coal-fired power plants. *Energy Economics*, 80, 491–501. <https://doi.org/10.1016/j.eneco.2019.01.032>
- Washington, A. (2014). Big Data and Public Sector Information.  
<http://ssrn.com/abstract=2386150>

## ANEXOS

