

## Desarrollo de un modelo inteligente para predecir el precio marginal local (PML) en un nodo de la red eléctrica nacional

Marcos Fidel Guzmán Escobar<sup>1</sup>, Alberto Alfonso Aguilar Lasserre<sup>1</sup>,  
Marco Julio Argumedo del Moral<sup>1</sup>, Nicasio Hernández Flores<sup>2</sup>,  
Gustavo Arroyo Figueroa<sup>2</sup>

<sup>1</sup> Tecnológico Nacional de México, Instituto Tecnológico de Orizaba,  
División de Estudios de Posgrado e Investigación,  
México

<sup>2</sup> Instituto Nacional de Electricidad y Energías Limpias,  
México

{fidel.mfge, marcojulioarg}@gmail.com,  
albertoaal@hotmail.com,  
{nicasio.hernandez, garroyo}@ineel.mx

**Resumen.** Esta investigación se centró en analizar 28 nodos representativos de los tres sistemas de interconexión de México para desarrollar pronósticos precisos del Precio Marginal Local (PML). Se utilizaron cinco modelos: Árboles de Decisión para Regresión, Regresión Polinomial de Segundo Grado, SARIMA, LSTM y una combinación de Prophet con LSTM Secuencial. Los resultados mostraron errores porcentuales medios absolutos (MAPE) en un rango de 42% a 23%. Además, se incluyeron métricas como el Error Absoluto Medio (MAE), el Error Cuadrático Medio (MSE) y la Raíz del Error Cuadrático Medio (RMSE) para evaluar la eficacia de los algoritmos en el pronóstico del PML. La metodología desarrollada, combina la integración y análisis de datos climáticos y el uso de algoritmos avanzados de Machine Learning (ML), para erigir un modelo replicable y robusto. Esta investigación proporciona un referente valioso para futuras investigaciones y aplicaciones prácticas en el ámbito del pronóstico energético en México. El proyecto culmina con el desarrollo de una Interfaz de Usuario Gráfica (GUI) interactiva, llamada “PML Explorer”, que permite a los usuarios visualizar de manera intuitiva los datos del Precio Marginal Local (PML). Además, incorpora una funcionalidad clave: la capacidad de ejecutar un código desarrollado en Python que integra cinco algoritmos avanzados. Esta herramienta está diseñada para predecir el PML hasta 48 horas en el futuro, mejorando significativamente la precisión de los pronósticos y facilitando la toma de decisiones basada en datos concretos y análisis predictivo.

**Palabras clave:** Machine learning, precio marginal local, árboles de decisión para regresión, LSTM, regresión polinomial, SARIMA, prophet, interfaz de usuario gráfica (GUI).

## Development of an Intelligent Model to Predict the Local Marginal Price (LMP) at a Node of the National Electric Grid

**Abstract.** This research focused on analyzing 28 representative nodes of the three interconnection systems in Mexico to develop accurate forecasts of the Locational Marginal Price (LMP). Five models were used: Regression Trees, Second Degree Polynomial Regression, SARIMA, LSTM, and a combination of Prophet with Sequential LSTM. The results showed mean absolute percentage errors (MAPE) ranging from 42% to 23%. Additionally, metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) were included to evaluate the effectiveness of the algorithms in forecasting LMP. The developed methodology combines the integration and analysis of climatic data and the use of advanced Machine Learning (ML) algorithms to build a replicable and robust model. This research provides a valuable reference for future research and practical applications in the field of energy forecasting in Mexico. The project culminates with the development of an interactive Graphical User Interface (GUI) called "PML Explorer," allowing users to intuitively visualize Locational Marginal Price (LMP) data. It also incorporates a key functionality: the ability to execute code developed in Python that integrates five advanced algorithms. This tool is designed to predict LMP up to 48 hours into the future, significantly improving forecast accuracy and facilitating decision-making based on concrete data and predictive analysis.

**Keywords:** Machine learning, local marginal price, decision trees for regression, LSTM, polynomial regression, SARIMA, prophet, graphical user interface (GUI).

### 1. Introducción

El Precio Marginal Local (PML) refleja el valor de la energía en un tiempo y lugar determinado. Este se calcula por medio de un modelo de optimización que minimiza los costos de generación mientras se satisface la demanda y se respetan las restricciones físicas de la red eléctrica (Alonso and Gabriel, 2017). Por lo tanto, la importancia de un precio justo mejorará los indicadores de disponibilidad, confiabilidad y eficiencia que se maneja en todos los procesos de potencia que se manejan en Empresa Generadora de Electricidad, que, a su vez, afectarán de manera positiva a los clientes consumidores de electricidad (Andrés Ramos, Gonzalo Cortés, Jesus Maria Latorre, and Cerisola, 2006).

Se presenta un modelo inteligente de pronóstico con redes neuronales para el Precio Marginal Local (PML) considerando el clima como factor a integrar y que influye en el precio de la energía de un nodo en específico. En la realización del modelo se llevó a cabo con el Instituto Nacional de Electricidad y Energías Limpias (INEEL), el cual es un centro de investigación dedicado a dar solución a las necesidades del sector eléctrico y energético. El artículo detalla trabajos relacionados en la materia y continúa describiendo la metodología empleada, detallando las etapas seguidas para desarrollar el modelo inteligente. Se destaca la optimización del modelo para mejorar su

rendimiento, así como la validación mediante métricas que cuantifican los errores en los pronósticos. En la sección dedicada a los resultados, se presentan los hallazgos de los cinco modelos de pronóstico evaluados, resaltando el rendimiento superior del modelo LSTM secuencial. Se ofrecen métricas de errores para comprender la eficiencia del modelo de Machine Learning al implementarlo en 28 nodos de interconexión. Finalmente, el proyecto culmina con el desarrollo de una Interfaz de Usuario Gráfica (GUI) denominada "PML Explorer". Esta herramienta interactiva no solo permite visualizar de manera intuitiva los datos del Precio Marginal Local (PML), sino que también integra la capacidad de ejecutar un código desarrollado en Python, el cual incorpora cinco algoritmos avanzados de aprendizaje automático, basados en datos concretos y análisis predictivo.

## **2. Trabajos relacionados**

La predicción de series temporales se ha destacado por el uso de redes neuronales recurrentes (RNN), las cuales poseen la capacidad inherente de retener una "memoria" que las hace idóneas para la predicción de secuencias extensas. Diversas variantes de RNN han demostrado su eficacia en este ámbito, permitiendo capturar patrones complejos en los datos temporales (SatvikKhuntia, AtharHanif, QadeerAhmed, JohnLahti, and MaartenMeijer, 2022). Por ejemplo, El enfoque de aprendizaje automático bayesiano (BML) se ha aplicado para la predicción de ondas con incertidumbre cuantificada, utilizando métodos de aproximación como la deserción Monte Carlo (MC) y hardware de ML moderno (Jincheng Zhang, Xiaowei Zhao, Siya Jin, and Greaves, 2022). El uso de redes neuronales en la optimización de la producción de energía eólica y la maximización de ingresos ha sido investigado, aplicando técnicas como el algoritmo genético cuántico y la coordinación de parques eólicos con centrales eléctricas de almacenamiento de energía (Xiyun Yang, Liwei Fan, Xiangjun Li, and Meng, 2022). El pronóstico del PML ha sido abordado utilizando diversos enfoques, como redes neuronales artificiales (RNA), el método de Mínimos Cuadrados Generalizados por División de Grupo (GMDH) y sistemas de inferencia difusa basados en redes neuronales adaptativas (ANFIS); donde el día del año fue la variable con mayor impacto en los componentes de pérdida y congestión.

Lo anterior está asociado al déficit de capacidad de transmisión eléctrica, como también los precios de los combustibles impactan significativamente en el componente energético, que está relacionado con el perfil de mezcla de generación de la región peninsular [6]. En el artículo (Jeovani E. Santiago López and García, 2018), se describe y caracteriza la evolución temporal de la serie de tiempo de los PML's horarios – diarios resultantes del MECP en México, analizando simultáneamente su volatilidad a lo largo de los periodos para cada mercado de energía (MDA y MTR) y para cada sistema interconectado (SIN, BCA y BCS) analizando la serie de tiempo no es estacionaria debido a patrones asociados a la frecuencia semanal donde el análisis se hizo enfocándose a datos históricos analizados estadísticamente.

El estudio del artículo (Andrés Ramos, Gonzalo Cortés, Jesus Maria Latorre, and Cerisola, 2006) analiza la relación entre el precio marginal del mercado de electricidad y la demanda mediante el uso de conglomerados analizados matemáticamente, donde se definió que el precio de la electricidad depende fundamentalmente de la demanda y

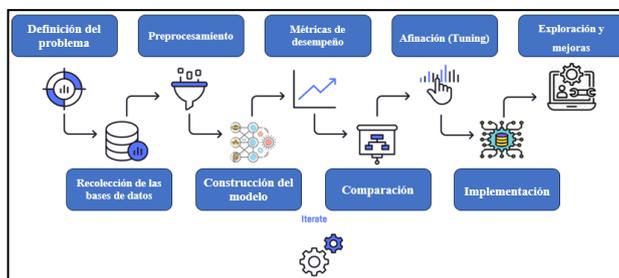


Fig. 1. Proceso de construcción de un modelo de Machine Learning.

la potencia disponible que es cambiante a lo largo del tiempo. Por otra parte, métodos como las redes neuronales de propagación inversa (BP) se han empleado para pronosticar los precios de liquidación del mercado energético diario (Deepak Singhal and Swarup, 2011). Estudios recientes han explorado funciones de pérdida para optimizar la predicción de precios al contado de la electricidad, utilizando redes neuronales con técnicas como K-vecino más cercano (KNN) y redes LSTM con aprendizaje reforzado por retro propagación (Ahmad Amine Loutfi, Mengtao Sun, Ijlal Loutfi, and Solibakke, 2022).

### 3. Metodología

A continuación, se muestra la metodología en la Fig. 1 que simplifica el proceso de construcción del modelo de ML:

El proceso a nivel general, estos son los pasos para seguir:

**Definición del problema:** El objetivo es predecir con precisión el Precio Marginal Local (PML) de nodos interconectados de la red nacional; lo que es un problema de regresión en un contexto de series temporales. Se necesita un enfoque de aprendizaje supervisado para una predicción precisa, con un horizonte de predicción de hasta 48 horas. Esto implica comprender los datos históricos y modelar sus variables y fluctuaciones.

**Recolección de las bases de datos:** Se recopilan datos de dos fuentes principales: registros horarios del PML del Centro Nacional de Control de Energía (CENACE) y variables climatológicas de la Comisión Nacional del Agua (CONAGUA), lo que plantea un desafío debido a la variabilidad en la operatividad de las estaciones climatológicas y la dificultad para obtener datos consistentes y completos.

**Preprocesamiento:** El preprocesamiento de datos es una etapa crítica para garantizar la calidad y utilidad de los datos en la modelización. Se utiliza “Python” y la biblioteca “Pandas” para preparar e identificar las variables en los datos. Esta fase incluye la unificación de archivos, la detección y manejo de valores faltantes, la normalización de datos entre 0 y 1 con “MinMaxScaler” que ayude a que todas las características contribuyan de manera equitativa al modelo sin que una característica domine debido a su escala más amplia, además de una ingeniería de características aplicando una conversión senoidal para que se tenga el patrón temporal de día y noche en los datos;

Por ejemplo, en una escala de tiempo de 24 horas, el salto de 23 a 0 crea una discontinuidad que no refleja la realidad cíclica del tiempo. Las funciones seno y coseno suavizan este salto, representando las 23:00 y 00:00 como puntos cercanos en el ciclo.

**Construcción del Modelo:** Se desarrollan una variedad de modelos de aprendizaje automático para abordar la complejidad de los datos del PML. Estos incluyen árboles de decisión, regresión polinomial, modelo en SARIMA y redes neuronales LSTM, estableciendo los subconjuntos 80% entrenamiento, 10% validación y 10% de prueba. Los pasos clave dentro del modelo LSTM + Prophet requiere:

1. **Preparación de los datos en formato seq2seq:** Los datos se organizan en pares de entrada y salida para permitir al modelo predecir secuencias futuras en lugar de valores individuales. Esta estructura de datos es fundamental para el enfoque seq2seq.
2. **Utilización de una capa LSTM seguida de una capa Time Distributed:** El modelo LSTM se emplea para capturar dependencias temporales complejas en los datos. La capa Time Distributed permite aplicar una capa densa a cada punto de tiempo en la secuencia de salida, lo que facilita la generación de predicciones secuenciales.
3. **Uso de Time Series Split para dividir los datos temporalmente:** La división temporal de los datos con Time Series Split garantiza que el modelo se entrene y valide de manera adecuada, evitando filtrar información del futuro hacia el pasado durante el proceso de entrenamiento.
4. **Preparación de Datos para Prophet:** Esto implicó indicar las columnas “ds” para las fechas y “y” para la variable objetivo (precios de energía) que son los pronósticos del modelo LSTM. Prophet requiere este formato específico para realizar sus predicciones de series temporales. Prophet está diseñado para modelar series temporales univariadas, lo que significa que solo puede manejar una variable objetivo (en este caso, 'y') y una variable de tiempo (en este caso, 'ds').
5. **Modelado y Entrenamiento con Prophet:** Se configuró y entrenó un modelo Prophet, excluyendo la estacionalidad anual (yearly\_seasonality=False) pero incluyendo la diaria (daily\_seasonality=True). Esta configuración se eligió considerando la naturaleza horaria de los datos de precios de energía.
6. **Generación de Predicciones Futuras:** Se generaron predicciones para un horizonte de 48 horas utilizando Prophet. Para esto, se creó un DataFrame “future\_lstm” que contenía las fechas futuras para las predicciones. Prophet produjo predicciones para este período, abordando las tendencias y estacionalidades identificadas.

**Métricas de desempeño:** Se evalúa el rendimiento de los modelos desarrollados utilizando métricas estándar como el Error Absoluto Medio (MAE), el Error Cuadrático Medio (MSE), la Raíz del Error Cuadrático Medio (RMSE) y el Error Porcentual Absoluto Medio (MAPE), así como técnicas de validación temporal con “timeseriesplit” y también con “walkforward-validation” para una evaluación rigurosa apropiada para series de tiempo.



Fig. 2. Localización geográfica de los nodos para la implementación del modelo.



Fig. 3. Menú principal de la interfaz gráfica de usuario "PML Explorer".

**Comparación de los modelos:** Para lograr una comparación objetiva y cuantitativa, se propone una tabla comparativa donde se presentan las métricas de rendimiento clave, incluyendo los errores MAE, MSE, RMSE y MAPE para cada modelo. Esta tabla permite no solo visualizar el rendimiento de cada modelo en términos de precisión de predicción, sino también entender sus fortalezas y limitaciones en el contexto específico del pronóstico de PML.

**Afinación (Tuning):** El objetivo es refinar el modelo LSTM para maximizar su precisión y eficiencia en la predicción de los PML.

- **Análisis de Sensibilidad en Redes LSTM:** Se llevó a cabo en redes LSTM, probando dos arquitecturas diferentes y ajustando manualmente sus hiperparámetros en 10 iteraciones, registrando los cambios en MAPE y MAE para identificar las configuraciones más efectivas.
- **Búsqueda Grid en LSTM Secuencial + PROPHET:** Para el modelo híbrido de LSTM con Prophet, se utilizó una búsqueda grid para explorar sistemáticamente una amplia gama de combinaciones de hiperparámetros y encontrar la configuración óptima para una mayor precisión predictiva.

**Implementación:** Esta fase realiza la aplicación de los cinco algoritmos en 28 nodos de interconexión en México, y se lleva a cabo un análisis utilizando los árboles de decisión para regresión y matrices de correlación.

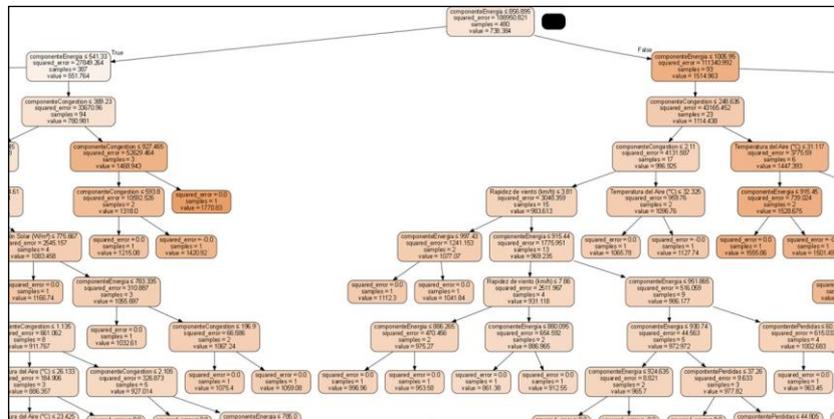
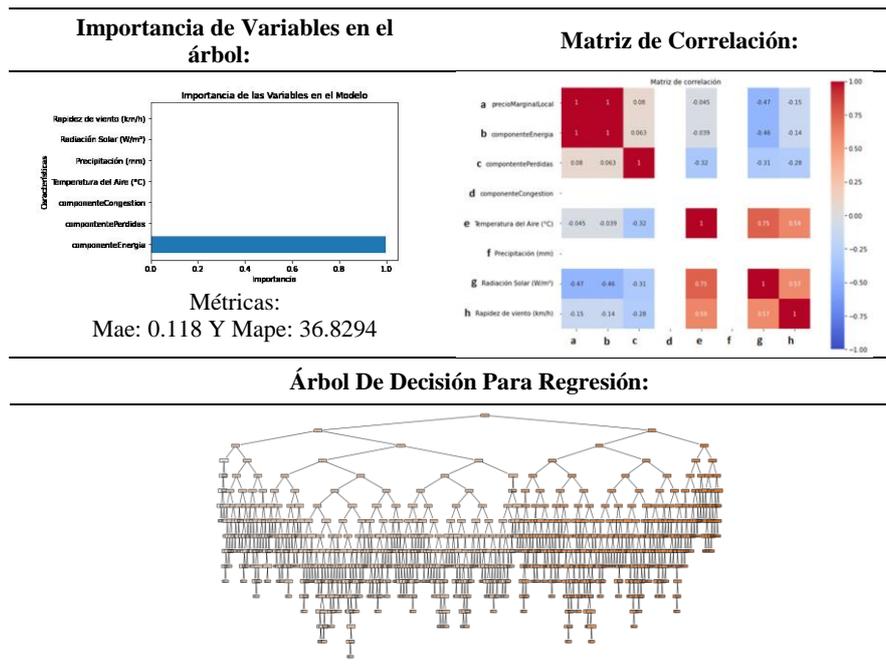


Fig. 4. Diagrama de árbol del nodo 1 de la implementación del modelo.

Tabla 1. Análisis de variables del nodo 14 (07LPZ-115), con árbol de decisión y matriz de correlación.



**Exploración y mejoras:** En esta etapa complementaria, se crea una Interfaz Gráfica de Usuario (GUI); usando “Tkinter” se propone una herramienta para visualizar los datos en una gráfica que le permite seleccionar diferentes periodos y guardarlos, si así se desea.

Una vez visualizados se regresa al menú para correr el código Python al oprimir el botón que ejecuta el código en entorno “Jupyter” notebook.

Incorporar una interfaz de usuario para que los analistas puedan ajustar manualmente la visualización gráfica y dinámica de los datos del PML es una buena herramienta.

#### 4. Resultados

Los árboles de decisión para regresión son modelos predictivos de valores numéricos por lo que el Precio Marginal Local (PML) puede ser tratado como un problema de regresión con funcionalidad de los árboles de decisión; donde cada nodo en el árbol representa una pregunta o una condición sobre las características de los datos, y cada rama representa una posible respuesta a esa pregunta.

Las hojas del árbol contienen las predicciones finales y no se limitó la altura del árbol; en la práctica, se utilizaron 720 registros (30 días de datos horarios), y la profundidad del árbol oscilaba entre 16 y 17 niveles para los diferentes nodos de la etapa de implementación. Al evaluar el modelo mediante gráficos de desempeño y validaciones cruzadas para series de tiempo, se determinó que esta profundidad proporcionaba un buen desempeño.

Esta configuración permitió al árbol capturar adecuadamente la complejidad de los datos sin necesidad de imponer una restricción explícita en su altura. Para la visualización más detallada se muestran en la Fig. y Tabla 1, y se pone a disposición un enlace a una carpeta en Google Drive para acceder al árbol en alta resolución<sup>1</sup>.

Tabla 1 muestra el nodo 07LPZ-115 de La Paz, Baja California Sur, donde el componente de energía es crucial para el pronóstico según el algoritmo de árbol de decisión y la matriz de correlación. Con una base de datos de PML de 28 nodos, el modelo LSTM secuencial puede aprender de las variables y ofrecer pronósticos confiables. Mejoramos el enfoque LSTM con una neurona "Dense" para predecir 48 horas a futuro. Aunque las métricas de prueba fueron menos eficientes que las de entrenamiento, el MSE, MAE y RMSE fueron bajos. A pesar de un MAPE elevado, el modelo generaliza bien, como se muestra en la Fig.

En la Fig. 6, se puede observar que cada división (split) se aproxima a cero, alcanzando un error cuadrático medio (MSE) muy cercano a cero antes de la época 80. Esto indica que el modelo aprendió eficazmente de la base de datos y generalizó bien los pronósticos con datos reales.

La Fig muestra una comparación entre los datos reales y los predichos por el modelo en el mismo período de tiempo. Sin embargo, la Fig. 1 va más allá al demostrar la capacidad del modelo para generalizar tanto con datos reales como para los próximos 48 PML imprevistos.

Otra mejora aplicada se logró al complementarla con Prophet; para ello es necesario que disponga de una neurona de salida llamada "*TimeDistributed*" fundamental para predecir una secuencia que pueda Prophet reconocer y lograr el ensamble de ambas herramientas útiles en series temporales. La razón además fue obtener un intervalo de confianza útil para nuestros pronósticos en LSTM fue Prophet. La proyección a futuro como se muestran en la Fig. 1 permite saber que tan bien logra generalizar los valores pronosticados y resultan ser bastante similares:

<sup>1</sup> [drive.google.com/drive/folders/1pMgxpHxrnHC8b68Ke9jTOLHo7fvLGeA?usp=drive\\_link](https://drive.google.com/drive/folders/1pMgxpHxrnHC8b68Ke9jTOLHo7fvLGeA?usp=drive_link)

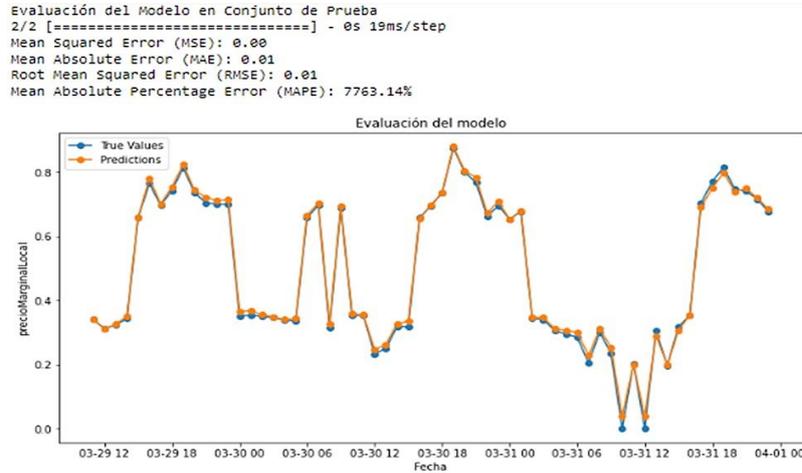


Fig. 5. Validación del modelo en el conjunto de prueba.

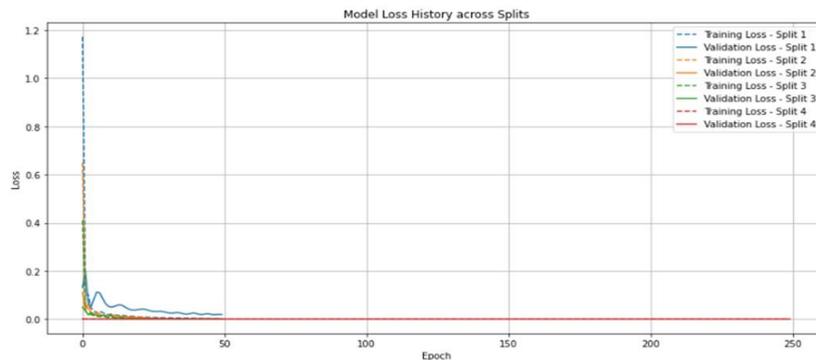
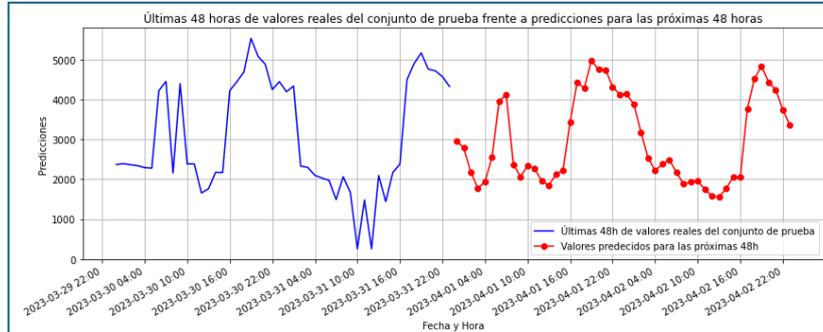


Fig. 6. Validación del modelo LSTM.

Además, se hizo una búsqueda en cuadrícula para asignar días festivos y parámetros de relevancia para la tendencia, que busquen automáticamente el pronóstico eficiente por considerar, por ejemplo, días de alta demanda entre semana, al modelo LSTM + Prophet. Dicha búsqueda abarca 3 parámetros con 4 valores, por lo tanto, el número total de combinaciones es  $4 \times 4 \times 4 = 64$  combinaciones a probar, para encontrar los mejores hiperparámetros para el modelo Prophet en función del MAPE.

Este itera a través de todas las combinaciones de parámetros y entrena un modelo Prophet para cada combinación. A continuación, los resultados del ensamble de dos poderosos algoritmos de pronóstico, véase Tabla 1 y Fig. :

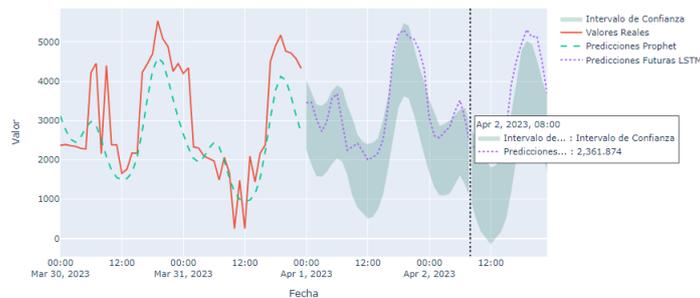
Las predicciones incluyen tres columnas relevantes para los intervalos de confianza, donde se muestran las predicciones generadas por el modelo LSTM (Long Short-Term Memory) en conjunto con el algoritmo Prophet para el Precio Marginal Local (PML) y cada fila representa una predicción para una hora específica, y las columnas proporcionan la fecha correspondiente, la predicción del PML, así como los límites inferior y superior de la predicción. Véase Fig. :



**Fig. 1.** Gráfica de los pronósticos a futuro del modelo LSTM.

**Tabla 1.** Lista de predicciones del PML para el nodo 07LPZ-115 La Paz, Baja California Sur, BCS.

	Fecha	Predicción	Limite Inferior	Limite Superior
720	2023-04-01 00:00:00	3112.005631	2193.283494	4038.913784
721	2023-04-01 01:00:00	2770.515998	1844.436235	3736.092496
722	2023-04-01 02:00:00	2517.998256	1660.765271	3393.738719
723	2023-04-01 03:00:00	2440.362533	1537.300040	3325.764496
724	2023-04-01 04:00:00	2572.406561	1680.079793	3476.827607
725	2023-04-01 05:00:00	2813.649044	1853.277718	3698.457952
726	2023-04-01 06:00:00	2963.441176	2103.976816	3896.577917
727	2023-04-01 07:00:00	2861.869027	1912.291430	3788.422816
728	2023-04-01 08:00:00	2512.488324	1604.671978	3448.388620
729	2023-04-01 09:00:00	2073.066336	1205.472347	3027.894809
730	2023-04-01 10:00:00	1724.196337	812.573193	2594.136177
731	2023-04-01 11:00:00	1541.021564	617.754523	2419.545514
732	2023-04-01 12:00:00	1485.305152	514.042457	2402.229779
733	2023-04-01 13:00:00	1511.350722	576.648782	2516.085676



**Fig. 8.** Gráfica de predicciones del PML del nodo 07LPZ-115 La Paz, Baja California Sur, BCS.

La Tabla 3 muestra los nodos de la implementación del modelo, donde resaltan en colores azules los nodos de los sistemas BCA y BCS, mientras que los que no tienen colores son del sistema SIN. Lo más destacable de la tabla es que resume en orden las correlaciones de las variables que se identificaron de relevancia en el proyecto, están ordenadas de acuerdo con el coeficiente de correlación “r” y además respaldadas por

**Tabla 3.** Resultado de los nodos interconectados y ordenados de acuerdo con el nivel de correlación.

NODOP				PRODESEN				Coeficiente de correlación
CLAVE	NOMBRE	NIVEL DE TENSION (kV)	SISTEMA	ZONA DE CARGA	REGION DE TRANSMISION	COMPONENTE		
15 07SNT-15	Santiago	115	BCS	LOS CABOS	LOS CABOS	energía	0.9393	
9 07TCT-69	Tecate	69	BCA	TUJANA	TUJANA	energía	0.9398	
12 07CAD-115	Cabo San Lucas Dos	115	BCS	LOS CABOS	LOS CABOS	energía	0.9398	
14 07LPCZ-115	La Paz	115	BCS	LA PAZ	LA PAZ	energía	0.9398	
7 07SAF-115	San Felipe	115	BCA	ENSENADA	ENSENADA	energía	0.9391	
18 07FLUM-69	Flumorsa	69	BCA	MEXICALI	MEXICALI	energía	0.9355	
8 07PCD-230	Cerro Prieto Dos	230	BCA	MEXICALI	MEXICALI	energía	0.9338	
27 03AMX-69	Almex	69	SIN	GUADALAJARA	GUADALAJARA	energía	0.9323	
13 07GAD-115	Central Termica Gral. Agustín Olachea	115	BCS	CONSTITUCION	CONSTITUCION	energía	0.9316	
19 03LVT-69	La Venta	69	SIN	MINAS	GUADALAJARA	energía	0.9311	
11 07LPC-115	Loreto	115	BCS	CONSTITUCION	CONSTITUCION	energía	0.9302	
10 07VPA-230	Valle de Puebla	230	BCA	MEXICALI	MEXICALI	energía	0.9302	
5 03AGH-400	Aguaumbia	400	SIN	TEPC VALLARTA	TEPC	energía	0.9297	
24 03AM-95	Alorandoc	95	SIN	CENTRO ORIENTE	CENTRAL	energía	0.9544	
23 03AMF-230	Antequero Internacional Felipe Angeli	230	SIN	CENTRO ORIENTE	CENTRAL	energía	0.9521	
3 03ITH-230	Tecohuacan	230	SIN	VDNORTE	CENTRAL	energía	0.9505	
1 02MMT-400	Manuel Moreno Torres	400	SIN	TUXTLA	GRUVALVA	energía	0.8667	
20 05MVU-400	Cuadro de Maniobras Villa Nueva	400	SIN	LAGUNA	LAGUNA	energía	0.8527	
21 05PAR-115	Parral	115	SIN	CAMARGO	CHIRAHUA	energía	0.8486	
2 06MON-115	Monclova	115	SIN	MONCLOVA	MONTREFFY	energía	0.8449	
17 06CBD-400	Carbon Dos	400	SIN	PIEDRAS NEGRAS	RIO ESCONDIDO	energía	0.8447	
26 06ECN-138	Central	138	SIN	MATAMOROS	MATAMOROS	energía	0.8067	
22 04GSV-115	Guasave	115	SIN	GUASAVE	LOS MOCHIS	energía	0.5565	
25 04PLD-230	Puerto Libertad	230	SIN	CAOBRA	HERMOSELLLO	energía	0.5442	
16 08TUM-115	Tulum	115	SIN	RIVERA MAYA	CANCLUN	congestión	0.8395	
6 06PTE-115	Poniente	115	SIN	MERIDA	MERIDA	congestión	0.8391	
4 08SLC-230	Santa Lucía	230	SIN	CAMPECHE	LERMA	congestión	0.7358	
28 02CBE-400	Ciclo Combinado Centro	400	SIN	MIDRELOS	PUEBLA	pérdidas	0.7883	

**Tabla 4.** Métricas de errores registradas en la implementación a 28 nodos del país.

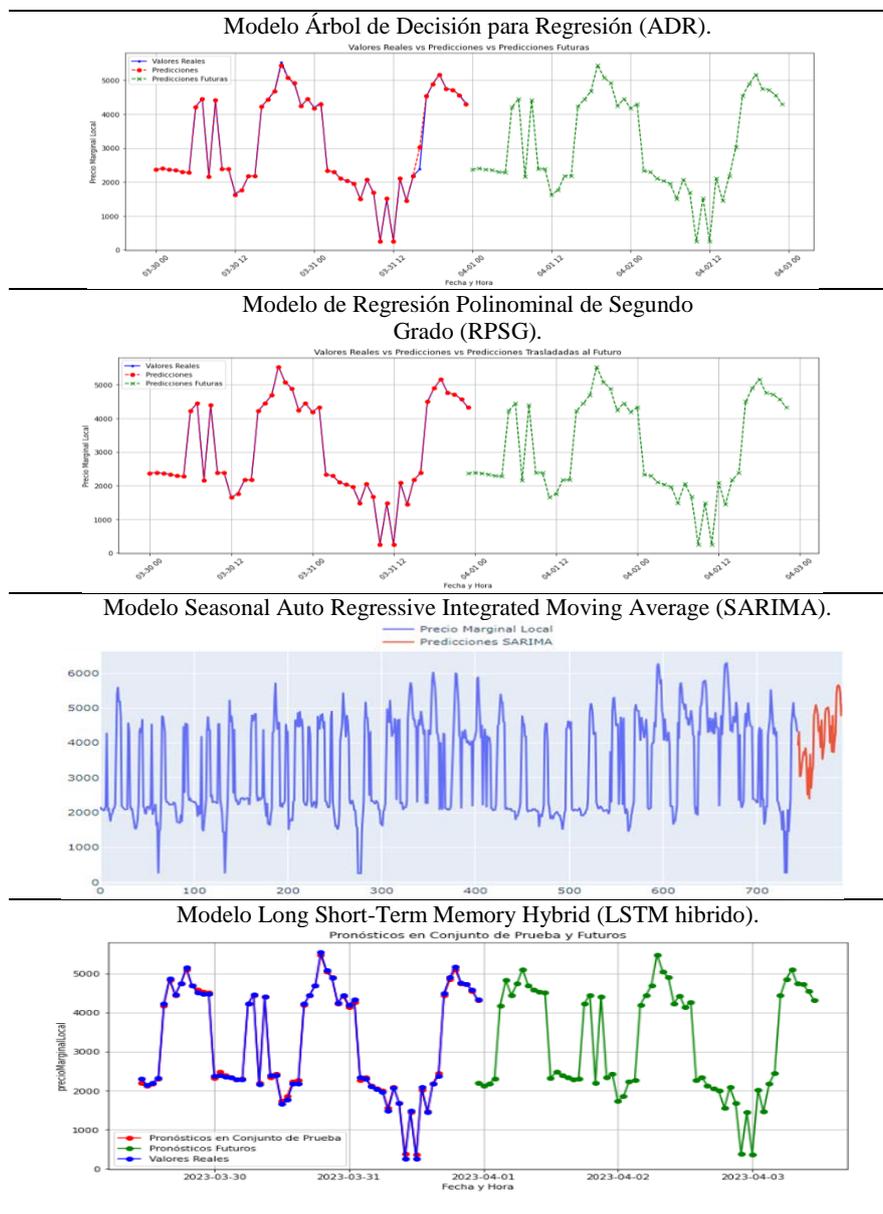
SISTEMA	CLAVE	MAE					MSE					RMSE					SARIMA				
		Actul D.R.	Reg. Pol.	SARIMA	LSTM	PROPHET	Actul D.R.	Reg. Pol.	SARIMA	LSTM	PROPHET	Actul D.R.	Reg. Pol.	SARIMA	LSTM	PROPHET	Actul D.R.	Reg. Pol.	SARIMA	LSTM	PROPHET
SIN	03AMT-400	11.4405	0.0043	537.0088	0.0069	0.1725	6316.894	0.000003	941655.1608	0.00007	0.0508	83.1682	0.006	795.3723	0.0086	0.2254	3.626	0.0006	44.4545	6.4789	328.2149
SIN	06AMN-115	71.5268	0.0054	1447.649	0.0406	0.123	33136.87	0.000006	5121280.048	0.0045	0.0314	182.0393	0.008	1585.0489	0.0671	0.1773	17.0954	0.0012	355.0385	166.3571	518.2491
SIN	07TTC-230	59.9408	0.0043	750.3296	0.0332	0.1226	28021.94	0.000006	847216.4049	0.0023	0.0264	167.3974	0.0063	920.4435	0.048	0.1627	4.2772	0.0007	81.7216	22.9482	199.1243
SIN	08SCC-230	50.8738	0.0048	524.2447	0.007	0.1693	10794.28	0.000005	529913.6658	0.0007	0.0474	101.4629	0.0062	737.9516	0.0088	0.2178	4.4949	0.0007	43.7878	4.9753	345.8007
SIN	09AMM-400	51.1106	0.0051	675.1131	0.0151	0.0988	5450.342	0.000004	635331.2236	0.0004	0.0167	73.8264	0.0067	797.0768	0.0203	0.1294	1.1867	0.0008	84.8911	5.5892	65.7652
SIN	09PFC-115	47.3855	0.0048	503.5288	0.0059	0.1348	7018.459	0.000004	449155.0947	0.0006	0.0319	83.7762	0.0065	846.9494	0.0078	0.2439	5.3199	0.0008	40.2207	19.2353	310.1529
BCA	05AAT-115	20.2111	0.004	807.5241	0.028	0.0761	757.0735	0.000003	962394.4338	0.0012	0.0091	27.5149	0.0058	981.017	0.0353	0.0955	1.8332	0.0004	73.3821	6.2028	23.1534
BCA	07CFA-230	11.0956	0.004	620.3987	0.0474	0.0855	9634.895	7659.906	712104.8159	0.0031	0.0105	98.1574	87.5208	849.2377	0.0564	0.1029	6.6876	6.7974	56.2545	13.7006	24.4397
BCA	07CFA-230	11.4024	0.0027	750.5297	0.0211	0.0768	946.5218	0.000004	855381.7229	0.0018	0.0145	23.7897	0.0046	924.8684	0.028	0.1028	0.9603	0.0003	70.7357	6.3302	21.9658
BCA	07VPA-230	51.9494	0.0038	621.8733	0.0432	0.079	9944.47	7659.967	742491.9878	0.0032	0.01	99.7219	87.5212	851.1122	0.0567	0.1	6.6446	6.7723	56.2474	12.2253	23.161
BCS	07RUC-115	47.6165	0.0043	4845.34	0.0081	0.1467	5092.122	0.000004	2900992.08	0.0001	0.0311	71.2602	0.0064	5385.232	0.0114	0.1822	1.77502	0.0004	233.5007	51.4104	49.4714
BCS	07CAG-115	17.2161	0.0038	4609.394	0.014	0.1782	459.2146	0.000005	2468970.5	0.0002	0.048	21.4226	0.0056	5448.0079	0.0162	0.2191	0.5298	0.0001	214.2511	88.9771	65.4005
BCS	07GAS-115	45.5725	0.0029	4401.108	0.0092	0.1605	4111.436	0.000002	27147794	0.0001	0.0389	44.1204	0.0048	5210.3544	0.0111	0.1974	1.9036	0.0002	234.8511	28.9372	54.6034
BCS	07RUC-115	23.7731	0.0037	6461.757	0.0045	0.118	6247.49	0.000003	5985404.48	0.0003	0.0218	79.041	0.0007	7671.9918	0.0068	0.1479	0.8284	0.0001	292.1631	80.814	66.8294
BCS	07SNA-115	11.2319	0.0037	4600.364	0.0119	0.1893	210.2421	0.000003	2961768.138	0.0002	0.0528	14.4997	0.0007	3442.2071	0.0157	0.2289	0.4294	0.0001	218.1658	332.0457	58.8845
SIN	07RUC-115	43.7178	0.0029	894.0714	0.0015	0.1773	8084.944	0.000004	3242147.934	0.0005	0.0148	89.9272	0.0044	1806.1417	0.0074	0.1219	4.0024	0.0004	42.222	31.81	571.4953
SIN	08CFA-400	71.8741	0.0063	1397.145	0.0556	0.1252	30367.19	0.000007	2331418.314	0.0071	0.0231	174.2641	0.0056	1507.7821	0.0463	0.176	16.2759	0.0015	348.4454	107.0837	2314.218
BCA	07VPA-230	23.4918	0.0037	5.8807	0.0575	0.061	1461.336	473.701	842194.0418	0.0048	0.0059	40.0166	21.7846	937.7313	0.0677	0.0771	2.0219	0.7514	70.5532	15.7968	17.8318
SIN	07VPA-230	48.7424	0.0037	359.1993	0.0177	0.0867	10791.91	0.000003	262878.9032	0.0009	0.0146	103.8819	0.0034	331.6914	0.0096	0.1209	4.7389	0.0001	30.42	4.051	89.4614
SIN	09AMV-400	61.4294	0.0052	1178.337	0.0426	0.1132	20700.4	4.9273	1464773.792	0.0034	0.025	143.8763	0.007	1283.2668	0.0589	0.1582	11.0606	0.0015	332.2425	13.5999	85.2353
SIN	09AMV-400	83.1033	0.0057	605.0316	0.0328	0.0819	25616.39	0.000007	545331.827	0.0019	0.0101	160.0312	0.0084	738.3983	0.0442	0.1007	11.342	0.0012	312.4586	12.1817	63.5174
SIN	08CAG-115	65.1607	0.004	440.0458	0.0319	0.0644	12106.22	0.000004	365793.3668	0.0014	0.0063	110.0273	0.0064	604.8081	0.0401	0.0784	8.1348	0.0006	50.5459	8.8899	15.7533
SIN	09AMV-230	60.2828	0.0046	2280.12	0.0228	0.1101	14200.18	0.000004	4667487.794	0.0009	0.0303	119.1653	0.0063	2582.1866	0.0301	0.1741	2.9683	0.0008	484.7839	19.6202	157.7258
SIN	09AMV-230	61.1824	0.005	753.6432	0.0309	0.1263	30275.01	0.000009	863052.5363	0.0023	0.0282	173.9971	0.0094	909.0052	0.0506	0.1681	4.4264	0.0003	80.6032	15.3646	201.2359
SIN	09MPC-230	124.255	0.0034	344.3625	0.0331	0.0683	69786.55	0.000002	197559.5895	0.0017	0.0083	264.1714	0.0052	444.4767	0.0421	0.0914	44.8935	0.0047	1109.71	71.8674	70.8578
SIN	08CFA-138	78.1978	0.0054	753.8242	0.0401	0.0567	43618.48	0.000007	734480.1616	0.003	0.0063	209.3291	0.0084	858.1843	0.0553	0.0799	25.8139	0.0016	162.6107	49.796	50.0497
SIN	09AMV-69	46.9664	0.0036	373.049	0.0137	0.1107	9050.403	0.000003	285336.7824	0.0002	0.023	96.1802	0.0056	338.3364	0.0173	0.1518	4.3261	0.0005	38.7913	71.3294	106.4486
SIN	08CFA-400	39.9155	0.0049	676.2495	0.0128	0.1347	10037.39	0.000004	612856.0327	0.0002	0.0317	100.1897	0.0066	782.8512	0.0149	0.1782	3.3314	0.0007	114.4917	12.3911	231.8203

los árboles de decisión para regresión; ambas técnicas representan la influencia medida de las variables con el PML, en sentido e intensidad.

Los resultados aplicados a los nodos de energía en la implementación se muestran en la Tabla 4, que concentra las métricas obtenidas para determinar la capacidad de predicción en los nodos de la república, donde en general, los modelos tuvieron el siguiente orden de eficacia de acuerdo con sus métricas: Regresión polinomial, árbol de decisión para regresión, LSTM, LSTM + Prophet y finalmente SARIMA, véase Tabla 4.

Además, se representan los diferentes modelos en la siguiente tabla para una conclusión directa del modelo y una comparativa para determinar sus cualidades frente los modelos de regresión tradicionales y de medias móviles. véase Tabla . Los resultados de la Tabla 5 indican que el modelo de árbol de decisión para regresión muestra un desempeño excepcional en los conjuntos de validación y prueba, destacándose por sus métricas sobresalientes.

**Tabla 5.** Gráficas de los pronósticos obtenidos en cada modelo de Machine Learning para el nodo 07LPZ-115, La Paz, Baja Cal., BCS.



Sin embargo, su capacidad predictiva se limita a un solo paso adelante debido a su naturaleza no secuencial, lo que resulta en pronósticos estancados a largo plazo. La regresión polinómica de segundo grado, por su parte, presenta un ajuste excelente a los datos históricos, superando al árbol de decisión en ciertos aspectos. No obstante, su

incapacidad para capturar la temporalidad y generalizar valores volátiles se traduce en predicciones que son meras extrapolaciones de datos pasados. En contraste, el modelo SARIMA muestra un rendimiento inferior en los pronósticos, con errores significativos en comparación con los valores reales. En cuanto a la LSTM, aunque generalmente tiene un buen rendimiento, su elevado MAPE sugiere dificultades para predecir con precisión, posiblemente debido a la volatilidad de los datos o a la presencia de valores cero en variables clave como la precipitación y la congestión. Estas consideraciones son cruciales para interpretar la capacidad predictiva de cada modelo en escenarios dinámicos y realistas.

## **5. Conclusiones y trabajo a futuro**

Los árboles de decisión, utilizados en un contexto de regresión, demuestran ser una herramienta valiosa no solo para predecir precios sino también para comprender las relaciones subyacentes entre las variables y permite identificar las variables más relevantes para la predicción, lo que respalda la elección de árboles de decisión en el análisis de modelos de aprendizaje automático en este proyecto, convirtiéndolo en un referente en la construcción de modelos de pronósticos eficientes. La aplicación de técnicas de Machine Learning fue esencial para abordar la complejidad inherente a los datos de series temporales y mejorar sustancialmente la precisión de los pronósticos de PML en los nodos de energía. El modelo más destacado es la red LSTM "dense", con Prophet el cual demostró ser muy efectivo, especialmente por sus pronósticos razonables a futuro y la provisión de intervalos de confianza. Los resultados obtenidos establecen una base sólida para futuros desarrollos y aplicaciones prácticas en el sector energético de México. Para futuros desarrollos, se propone una mejora continua. Un desafío clave es la adaptación de las redes neuronales para prever “n” datos futuros, lo que requiere flexibilidad en la manipulación de “arrays” desiguales o incompatibles, lo que podría requerir un proyecto especializado para garantizar flexibilidad y realizar pruebas exhaustivas.

## **Referencias**

1. Alonso, G.: Reporte anual del mercado eléctrico mayorista 2016. Monitor independiente del Mercado (2017)
2. Ramos, A., Cortés, G., Latorre, J.M., Cerisola, S.: Análisis de la relación precio marginal y demanda de electricidad mediante conglomerados. In: X Congreso de Ingeniería de Organización (2006)
3. Loutfi, A.A., Sun, M., Loutfi, I., Solibakke, P.B.: Empirical Study of Day-ahead Electricity Spot-price Forecasting: Insights into a Novel Loss Function for Training Neural Networks. Elsevier, vol. 319 (2022). DOI: 10.1016/j.apenergy.2022.119182.
4. Singhal, D., Swarup, K.: Electricity Price Forecasting using Artificial Neural Networks. In: International Journal of Electrical Power and Energy Systems, vol. 33, no. 3, pp. 550–555 (2011). DOI: 10.1016/j.ijepes.2010.12.009.
5. Santiago-López, J.E., García, F.G.: Precios marginales locales: Análisis de volatilidad del MEM. In: IEEE, RVP-AI/2018, pp. 18 (2018)
6. Livas-García, A., May-Tzuc, O., Cruz-May, E., Tariq, R., Jimenez-Torres, M., Bassam, A.: Forecasting of Locational Marginal Price Components with Artificial Intelligence and

- Sensitivity Analysis: A Study under Tropical Weather and Renewable Power for the Mexican Southeast. Elsevier, vol. 206, (2022). DOI: 10.1016/j.epsr.2022.107793.
7. Yang, X., Fan, L., Li, X., Meng, L.: Day-ahead and Real-Time Market Bidding and Scheduling Strategy for Wind Power Participation based on Shared Energy Storage. *Electric Power Systems Research*, vol. 214 (2022). DOI: 10.1016/j.epsr.2022.108903.
  8. Zhang, J., Zhao, X., Jin, S., Greaves, D.: Phase-resolved Real-Time Ocean Wave Prediction with Quantified Uncertainty based on Variational Bayesian Machine Learning. Elsevier, vol. 234 (2022)
  9. Khuntia, S., Hanif, A., Ahmed, Q., Lahti, J., Meijer, M.: Cabin Load Prediction Using Time Series Forecasting in Long-haul Trucks for Optimal Energy Management. Elsevier, pp. 342–347 (2022)