ISBN: 978-628-96613-1-6. ISSN: 2414-6390. Digital Object Identifier: https://dx.doi.org/10.18687/LACCEI2025.1.1.1996

Long-Term Electric Demand Forecasting Using a Multiple Linear Regression Model

Wiliam Lorenzo Tolentino Lezama, Student[1], Luis Angel Saravia Velasquez, Student[2],
Angel E. Obispo, Dr.[3], y Joel Villavicencio Gastelu, Dr.[4]
[1,4] Universidad Tecnológica del Perú, Perú,
aobispo@utp.edu.pe

Abstract*– The continuous increase in energy demand highlights the need for its forecasting to ensure energy supply. Various methods have been used to predict this demand; however, this study proposes an approach based on a multiple linear regression model to perform energy demand forecasting. Demand energy is considered the dependent variable, while the independent variables include socioeconomic factors such as population, gross domestic product (GDP), energy production and temperature. To validate the accuracy of the model, statistical indicators such as the coefficient of determination (R^2) and percentage error metrics, including the mean absolute deviation (MAD) and the mean absolute percentage error (MAPE), are applied. The proposed method is applied to the energy demand of the Peruvian national case, using historical data collected from 1995 to 2021. The results show an error of 1.7% between the historical demand used for model training (70% of the data, corresponding to 1995–2013) and the forecasted demand for the validation period (30% of the historical data, corresponding to 2014–2021), indicating a good level of accuracy in the estimates. Based on these results, long-term forecasting is performed, projecting electric demand from 2022 to 2038. The proposed methodology is useful for energy demand forecasting studies carried out by operators of national electrical power systems, such as Economic System Operation Committee (COES) in the case of Peru.

Keywords*-- Load forecasting, Electric demand, Electrical energy.

Previsión de la Demanda Eléctrica a Largo Plazo Basado en el Modelo de Regresión Múltiple

Wiliam Lorenzo Tolentino Lezama, Estudiante[1], Luis Angel Saravia Velasquez, Estudiante[2], Angel E. Obispo, Dr.[3], y Joel Villavicencio Gastelu, Dr.[4]
[1,4] Universidad Tecnológica del Perú, Perú,
aobispo@utp.edu.pe

Resumen*– El aumento continuo en la demanda de energía resalta la necesidad de su predicción a fin de asegurar su suministro en el futuro. Diversos métodos han sido utilizados para predecir esta demanda; sin embargo, este trabajo se utiliza un modelo de regresión lineal múltiple para realizar la previsión de la demanda. La demanda eléctrica se considera como la variable dependiente, mientras que las variables independientes incluyen factores socioeconómicos como la población, el producto bruto interno (PBI), la producción de energía y la temperatura. Para validar la precisión del modelo, se aplican indicadores estadísticos como el coeficiente de determinación (R^2) y métricas de error porcentual, incluyendo el error absoluto medio (MAD) y el error porcentual absoluto medio (MAPE). El modelo propuesto se aplica al caso de la demanda de energía nacional del Perú, utilizando datos históricos recopilados entre 1995 y 2021. Los resultados muestran un error del 1.7% entre la demanda histórica utilizada para el entrenamiento del modelo (70% de los datos, correspondientes a 1995-2013) y la demanda prevista para el período de validación (30% de los datos históricos, correspondientes a 2014-2021), lo que indica un buen nivel de precisión en las estimaciones realizadas. En base a estos resultados, se realiza un pronóstico a largo plazo, proyectando la demanda eléctrica desde 2022 hasta 2038. La metodología presentada es útil para estudios de previsión de demanda que llevan a cabo los operadores de los sistemas eléctricos nacionales, como es el Comité de Operación Económica del Sistema (COES) en el caso del Perú.

Palabras clave*-- Factores socioeconómicos, Previsión de demanda, Regresión lineal múltiple.

I. Introducción

A. Contextualización y Planteamiento del Problema

El crecimiento constante de la demanda de energía eléctrica, impulsado por factores como el aumento poblacional, la expansión industrial y el desarrollo tecnológico, plantea la necesidad de una planificación eficiente de las fuentes de generación y de las redes eléctrica. Tradicionalmente, esta planificación se ha basado en proyecciones de demanda a nivel global dentro de las zonas de operación de las concesionarias eléctricas. Sin embargo, depender exclusivamente de estas estimaciones globales puede llevar a decisiones subóptimas, generando costos adicionales al actualizar o expandir la infraestructura, debido a la falta de consideración de dinámicas locales específicas.

Para optimizar el uso de los recursos y mejorar la precisión en la toma de decisiones, se propone el uso de proyecciones basadas en series de datos históricos, un enfoque conocido como "previsión de la demanda a largo plazo". Este método permite identificar patrones de consumo y anticipar escenarios futuros considerando variables socioeconómicas y ambientales. Además, permite mejorar la gestión de la infraestructura existente y es clave para planificar nuevas inversiones, integrar energías renovables y desarrollar redes inteligentes, contribuyendo así a un sistema eléctrico más eficiente y sostenible.

B. Revisión Literaria

La predicción de la demanda eléctrica es fundamental para la expansión eficiente y sostenible de la red eléctrica, ya que permite obtener estimaciones precisas del consumo futuro, facilitando una mejor planificación de los recursos energéticos. Esta precisión es esencial para evitar tanto la sobrecarga de la infraestructura como el desperdicio de recursos. Para mejorar la exactitud de estas proyecciones y minimizar los márgenes de error, se han desarrollado diversos métodos, programas y algoritmos avanzados que combinan enfoques estadísticos tradicionales con técnicas modernas de inteligencia artificial.

Estos métodos de pronóstico requieren la integración de recursos técnicos, humanos, económicos, tecnológicos e informáticos específicos, así como un profundo conocimiento de los requerimientos de cada enfoque y técnica para garantizar resultados confiables. En este contexto, el pronóstico de la demanda eléctrica emplea modelos matemáticos y métricas de error específicas, como el error absoluto medio (MAD) y el error absoluto porcentual medio (MAPE), para evaluar la precisión de las estimaciones de carga futura.

Por ejemplo, en [1] se integran de manera estratégica dos modelos de aprendizaje automático, CatBoost y XGBoost, junto con algoritmos de optimización como el Algoritmo de Búsqueda de Gorriones (SSA), la Optimización por Enjambre de Partículas Fásicas (PPSO) y la Optimización Híbrida del Lobo Gris (GWO). Los resultados obtenidos destacan que el modelo XGBoost-SSA ofrece el mejor rendimiento, mostrando una alta precisión y estabilidad en la predicción del consumo eléctrico en distintos escenarios. Este enfoque innovador no solo mejora la gestión de la demanda energética, sino que también contribuye a la optimización de recursos y a la promoción del desarrollo sostenible en el sector eléctrico, con potenciales aplicaciones en diferentes regiones y horizontes temporales más amplios.

Asimismo, en [2] se presentan tres métodos robustos para la predicción de carga a corto plazo: RME multivariado, MHDE multivariado y DEXR univariado. Estos modelos han sido desarrollados a partir de un exhaustivo análisis de la literatura existente y han sido validados mediante series temporales de demanda eléctrica en Francia. Los resultados evidencian la

capacidad de estos métodos para identificar cambios abruptos en el comportamiento estocástico de la serie, conocidos como "rupturas", que suelen estar relacionados con eventos especiales, como días festivos públicos o transiciones entre días festivos y días laborales. En particular, el método RME basado en modelos SARIMA multivariados ha demostrado ser el más eficaz, siendo recomendado para la modelización y predicción en tiempo real de la demanda eléctrica. Además, se destaca la efectividad de un enfoque de modelado multivariado simple y práctico, capaz de abordar la heterocedasticidad de la serie y mejorar la robustez de las proyecciones en entornos dinámicos y cambiantes.

En otro estudio [3], se propone un modelo basado en la teoría de análisis de agrupamiento para la previsión de la demanda eléctrica. Este método incluye varias técnicas como ISFM, GFM, LRFM, entre otras, que dividen las estimaciones en regiones celulares y utilizan datos históricos para desarrollar modelos de predicción. Se comparan métricas como AAE, AARE y RMS en cada método, concluyendo que el enfoque de análisis de agrupamiento mejora la precisión de las predicciones y reduce los errores de pronóstico, siendo útil para la gestión de la demanda eléctrica en áreas específicas.

En [4] aborda la importancia de la planificación precisa del suministro de energía eléctrica mediante el desarrollo de un método híbrido de pronóstico de carga eléctrica, denominado VMD-RVFL. Este enfoque combina la descomposición precisa de la carga eléctrica en funciones de modo intrínseco (IMF) mediante la técnica VMD, con la capacidad de predicción rápida y precisa de la red RVFL. El estudio demuestra que el VMD-RVFL supera a los métodos existentes con relación a la precisión en la previsión de la carga eléctrica, ofreciendo una solución efectiva para contrarrestar la no linealidad y la naturaleza no estacionaria de los datos de carga eléctrica. Además, se identifica como trabajo futuro la exploración de variantes computacionalmente más eficientes del VMD y su combinación con RVFL para mejorar aún más la eficacia del pronóstico a corto plazo de la carga eléctrica.

En [5] se desarrolla un modelo eficiente para la predicción de carga eléctrica a corto plazo utilizando el algoritmo de Bosques Aleatorios (RF). La complejidad del pronóstico de debe a diversos factores como tendencias no lineales, estacionalidades múltiples y variabilidad variable. El RF aborda estas dificultades mediante la combinación de árboles de regresión a través de técnicas de bagging y subespacios aleatorios, lo que mejora la precisión de la predicción y reduce la variabilidad del modelo. El modelo propuesto se evaluó utilizando datos históricos de carga horaria de cuatro subestaciones transformadoras que abastecen diferentes áreas del campus de la Universidad de Beira Interior, Portugal. Los resultados evidencian la eficacia del modelo RF al prever la carga horaria en el corto plazo, incluso con solo un día de anticipación. Este descubrimiento tiene el potencial de optimizar el análisis en las operaciones de la red inteligente.

En [6] se presenta un modelo de regresión de vectores de soporte (SVR) con algoritmo inmune (IA) para determinar parámetros para la predicción de la demanda. Se compara con métodos tradicionales como redes neuronales artificiales (ANN) y modelos de regresión, demostrando una mayor precisión, especialmente en regiones específicas. La combinación de SVR e IA proporciona resultados superiores al abordar patrones de datos no lineales y complejos, destacando su eficacia en la predicción precisa de la carga eléctrica, lo que sugiere su potencial utilidad en la industria eléctrica. Además, se señala la posibilidad de integrar factores sociales y estacionales para mejorar aún más la capacidad explicativa del modelo y se plantea la exploración de otras técnicas de optimización avanzada para refinar el proceso de determinación de parámetros en las investigaciones futuras en el espacio de la teoría SVR.

Además, en [7] se presenta un método, donde para un pronóstico propuesto aplica una regresión ponderada geográficamente, que contempla la carga instalada con variables socioeconómicas independientes o dependientes del lugar de estudio, donde la variable dependiente es la sub-área y la variable independiente es el área total. Los resultados muestran una alta probabilidad de desarrollo de la demanda eléctrica, lo cual es consistente con el crecimiento de dicha área y se presentan en forma de mapas que representan las preferencias futuras de crecimiento eléctrico.

En [8], se presenta un método de predicción de carga anual en sistemas eléctricos mediante el uso de la técnica de optimización por enjambre de partículas (PSO). Al formular el problema como una estimación en un espacio de estados, se aplica PSO para reducir el error en los parámetros del modelo propuesto. Para validar los resultados, se comparan con datos verdaderos de las redes eléctricas de Kuwait y Egipto utilizando la técnica de mínimos cuadrados (LES). Los hallazgos muestran que PSO proporciona estimaciones más precisas, destacando su potencial como una herramienta valiosa para la estimación de parámetros. Además, el estudio sugiere que el modelo cuadrático se ajusta mejor a los datos disponibles y señala la necesidad de investigaciones futuras que consideren variables adicionales, como el precio del petróleo, para mejorar aún más la precisión de las predicciones.

Por otro lado, en [9] se enfoca en el modelo de aprendizaje incremental de regresión de vectores de soporte (SVR) para el pronóstico de carga eléctrica, optimizando el manejo de datos que llegan por lotes y conjuntos de datos grandes. El modelo propuesto, denominado SVR incremental (I-SVR), emplea la reconstrucción del espacio de fases (PSR) para integrar series temporales y selecciona un subconjunto de entrenamiento óptimo, reduciendo así la complejidad temporal y espacial del entrenamiento. Cuando se añaden nuevos datos, se utiliza un método de reconstrucción de conjuntos de datos representativos para reentrenar rápidamente el modelo, y un marco de optimización de enjambre de partículas anidadas (NPSO) para seleccionar los parámetros del modelo.

Los experimentos demostraron que el I-SVR supera a los modelos tradicionales en precisión y eficiencia computacional. Este enfoque evita la necesidad de reentrenar el modelo completo cada vez que se reciben nuevos datos, manteniendo una alta capacidad de generalización y mejorando la interpretabilidad y rendimiento del pronóstico de carga eléctrica. Las futuras investigaciones se centrarán en mejorar la selección

de candidatos a vectores de soporte y en desarrollar reglas difusas para la configuración de parámetros, así como en integrar variables exógenas en el modelo.

Por último, en [10] se aborda el modelo regresión lineal múltiple (MLR). Este método se emplea para prever la demanda eléctrica de un día a otro, considerando variables como la temperatura, el punto de rocío, la carga del día anterior, las horas y la carga de la semana anterior. La precisión del modelo se evalúa mediante el error porcentual absoluto medio y el coeficiente de determinación (R^2). La correcta previsión de la demanda eléctrica es crucial para asegurar un suministro continuo y económico de energía, dado que factores como las condiciones climáticas, los eventos festivos y la estructura tarifaria influyen significativamente en el consumo energético.

C. Contribución

Este trabajo propone un modelo de regresión lineal múltiple que logra un error inferior al 2%, lo que representa una mejora respecto a otros estudios que presentan errores mayores. Esta mejora se debe a la inclusión de variables adicionales, como la producción de energía y la temperatura, que contribuyen a un ajuste más preciso del modelo. A diferencia de enfoques que se centran únicamente en factores socioeconómicos tradicionales, la incorporación de estas variables permite capturar de manera más adecuada la influencia de factores ambientales y operativos en la demanda eléctrica. En este sentido, el modelo no solo contribuye a reducir el margen de error en la estimación de la demanda, sino que también facilita una mejor toma de decisiones en el ámbito de la gestión energética a largo plazo.

II. METODOLOGÍA

A. Datos de Entrada

- **Históricos de Demanda de Energía:** Representa el consumo energético por parte de los habitantes a lo largo de los años. Esta información será utilizada para ajustar nuestro modelo de predicción y para realizar pronósticos futuros de demanda de energía. La información sobre la demanda de energía se obtuvo de [11].
- **Producto Bruto Interno (PBI):** El PBI es un reflejo directo del nivel de actividad económica de una nación y, por ende, de su consumo energético. Cuando el PBI experimenta un crecimiento, suele ir acompañado de un aumento en la demanda eléctrica. Los datos relativos al PBI fueron obtenidos en [12].
- **Población:** La población influye directamente en la previsión de la demanda, debido a que el aumento del número de habitantes va acompañado con el aumento del consumo de energía. Los datos de la población fueron adquiridos en [13].
- **Producción de Energía Eléctrica:** La producción de la energía eléctrica está directamente vinculada con la demanda, y, debe satisfacer la creciente demanda. La información sobre la producción de energía eléctrica se obtuvo en [11].
- **Temperatura:** Debido al cambio climático, se reflejan variaciones en la temperatura ambiente. Por ejemplo, en verano, el aumento de la temperatura conlleva en muchos casos al uso de equipos de aire acondicionados, mientras que, en invierno, el frío hace necesario el uso de calefactores eléctricos. Por tanto, la variación de la temperatura está directamente relacionado el consumo de energía eléctrica y, por ende, con la predicción de la demanda. La información sobre la temperatura se obtuvo en [14].

B. Previsión de la Demanda usando Regresión Lineal Múltiple

La regresión lineal múltiple es una técnica estadística que se utiliza para analizar y modelar la relación entre una variable dependiente y dos o más variables independientes. Su propósito es comprender cómo las variables independientes influyen en la variable dependiente y, a partir de esa relación, realizar predicciones sobre su comportamiento futuro [15]. Esta metodología es ampliamente utilizada en diversas disciplinas debido a su capacidad para identificar patrones, evaluar el impacto de factores específicos y generar modelos predictivos robustos.

En el contexto de este estudio, se aplica la regresión lineal múltiple para pronosticar la demanda eléctrica (variable dependiente), considerando como variables independientes diversos factores socioeconómicos relevantes. El modelo adoptado se expresa mediante la siguiente ecuación:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \quad (1)$$

y : Demanda eléctrica (MW)
β_0 : Intercepción (MW)
β_1 : Coeficiente de PBI (MW por millón de USD)
β_2 : Coeficiente de población (MW por persona)
β_3 : Coeficiente de producción de energía
β_1 : Coeficiente de temperatura
X_1 : Variable de PBI (millones de USD)
X_2 : Variable de población (personas)
X_3 : Variable de producción de energía eléctrica (MW)
X_4 : Variable de temperature (°C)

Los coeficientes de cada variable independiente (β_1, β_2 , β_3 y β_4) se determinaron a partir de datos históricos de las respectivas variables, recopilados entre 1995 y 2021. Para el análisis estadístico, se empleó la herramienta de análisis de datos de Excel. Los coeficientes se obtuvieron a través del comando de análisis de regresión, el cual proporciona no solo los valores de los coeficientes, sino también estadísticas complementarias, como el coeficiente de determinación (R^2) y el error estándar, que permiten evaluar la precisión del modelo [16].

A diferencia de estudios previos que consideran únicamente dos variables, como el PBI y la población [17], este trabajo amplía el análisis al incorporar cuatro variables

independientes. La inclusión de la producción de energía y la temperatura tiene un impacto significativo en la mejora de la capacidad explicativa del modelo. Estos factores adicionales permiten captar de manera más precisa las fluctuaciones en la demanda eléctrica, ya que reflejan tanto el crecimiento económico y demográfico como las variaciones estacionales y operativas del sistema eléctrico. No obstante, es importante resaltar que la producción de energía se considera como una variable independiente únicamente en el sentido estadístico, al haber mostrado una correlación con la demanda eléctrica en los datos históricos. No se interpreta como una causa directa, sino como una variable que refleja el comportamiento conjunto del sistema energético durante el período analizado.

Además, la incorporación de estas variables contribuye a una reducción en el error promedio del modelo, lo que se traduce en pronósticos más confiables para la planificación del sistema eléctrico a largo plazo. En particular, la producción de energía actúa como un indicador de la capacidad operativa del sistema, mientras que la temperatura permite captar la influencia de factores climáticos en el consumo de energía, especialmente en contextos donde la climatización tiene un peso relevante en la demanda total.

C. Precisión del Pronóstico

Para la interpretación de la precisión de los pronósticos de la demanda se aplican diferentes métricas o estadísticas de regresión.

Para validar la capacidad de pronóstico se usan los valores de la demanda eléctrica de los años 1995-2013 que representan el 70% de todos los datos históricos. El resto de los datos (30%) se usó para evaluar la precisión del pronóstico desde el 2014 al 2021. Adicionalmente, se harán pronósticos a futuro a partir del año 2022.

Según la obtención de los datos que representan el 70%, se presenta la estadística de regresión R^2 y el R^2 ajustado [18]. Lo que permite evaluar el grado de probabilidad del pronóstico de la demanda. Un R^2 que se acerca a 1, indica un modelo altamente preciso, para una predicción adecuada se considera un R^2 mayor al 0.85 (85%). Lo cual, es considerado de alta precisión en muchos contextos.

Por otro lado, el análisis de errores relacionados con las métricas de rendimiento incluye el Índice de Error Absoluto Medio (MAD), es la media de las diferencias absolutas entre los valores reales y los valores predichos. El Índice de Error Porcentual Absoluto Medio (MAPE), refleja la capacidad de pronóstico de un modelo en porcentaje, donde el valor menor al 10% se considera con buen resultado [19].

- **Error absoluto de la media (MAD)**

$$MAD = \frac{\sum\left|\frac{D_r - D_p}{D_r}\right|}{N} \tag{2}$$

- **Error absoluto de la media (MAD)**

$$MAPE = \frac{\sum\left|\frac{D_r - D_p}{D_r}\right| \cdot 100}{N} \tag{3}$$

D_r : Demanda real

D_p : Demanda pronosticada

N : Número de periodos

III. Resultados y discusiones

La metodología propuesta se aplicó para realizar la prever la demanda de energía eléctrica en el Perú, utilizando datos históricos que abarcan el período de 1995 a 2021 [11-13].

A. Resultados

En la Tabla I se presentan las estadísticas de regresión obtenidas.

TABLA I
ESTADÍSTICA DE REGRESIÓN

Coeficiente de determinación R^2	0.996533813
R^2 ajustado	0.995543474

Se observa que los valores del coeficiente de determinación (R^2) y del R^2 ajustado son cercanos a 1, lo que indica que el modelo explica de manera adecuada la variabilidad de la demanda eléctrica en función de las variables independientes consideradas. En términos de precisión, esto se refleja en errores porcentuales bajos, calculados en 0.35% para el R^2 y 0.45% para el R^2 ajustado. Estos márgenes de error sugieren que el modelo tiene un buen nivel de ajuste y es capaz de proporcionar estimaciones confiables para pronósticos futuros.

La Tabla II muestra los coeficientes de regresión obtenidos para cada variable independiente. El término de intercepción (β_0) presenta un valor negativo (-6455.62985), lo cual no implica un fallo en el modelo, ya que este valor representa el punto de corte cuando todas las variables independientes son cero, lo que no necesariamente tiene una interpretación directa en este contexto, siempre que las variables independientes tomen valores típicos dentro del rango observado en los datos.

TABLA II
COEFICIENTES DE LAS VARIABLES INDEPENDIENTES

Coeficientes	β
Intercepción (β_0)	-6455.62985
Población (P)	0.000102687
PBI	2.002292114
Producción de energía (PE)	0.086853172
Temperatura (T)	200.9563118

Por otro lado, los coeficientes asociados a las variables independientes indican el cambio esperado en la demanda eléctrica por cada unidad de incremento en dichas variables,

manteniendo las demás constantes. Por ejemplo, un aumento de una unidad en el PBI está asociado con un incremento de aproximadamente 2.0023 unidades en la demanda eléctrica, mientras que la temperatura muestra un impacto significativo, con un coeficiente de 200.9563, lo que sugiere una fuerte relación entre el clima y el consumo de energía.

Considerando estos coeficientes, se puede formular la ecuación de regresión múltiple que modela la demanda eléctrica (y) de la siguiente manera:

$$y = -6455.6299 + 0.00010269 \cdot P + 2.002292114 \cdot PBI + 0.086853172 \cdot PE + 200.9563118 \cdot T \quad (4)$$

Es importante mencionar que, en el contexto peruano, el consumo eléctrico por calefacción es prácticamente nulo, lo que justifica una aproximación lineal entre temperatura y demanda para temperaturas superiores a 18 °C, donde predomina el uso de refrigeración.

Con la ecuación ajustada para cada variable independiente, se realiza una prueba del modelo considerando los datos históricos y los valores pronosticados para el período 2014-2021, como se muestra en la Fig. 1.

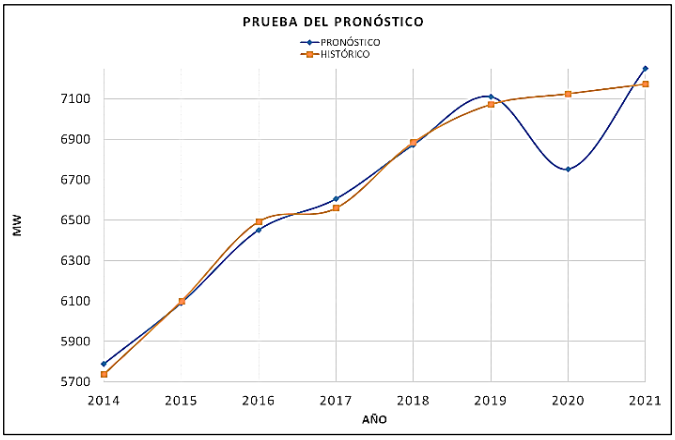

Fig. 1. Curva de Datos históricos vs Pronóstico.

En la Fig. 1 se presenta la comparación entre la curva de los datos históricos de demanda eléctrica (línea naranja) y los valores pronosticados por el modelo (línea azul). El objetivo de esta comparación es evaluar la precisión del modelo al verificar su capacidad para replicar el comportamiento en el intervalo temporal observado. Se puede apreciar que la línea de pronóstico se mantiene cercana a la línea de los datos históricos, lo cual sugiere una buena capacidad predictiva del modelo y refuerza su fiabilidad para proyecciones futuras. Para esta validación, se ha utilizado el 30% de los datos históricos, correspondiente al período 2014-2021, comparándolos con los valores calculados.

Además, se observa una disminución notable en la demanda eléctrica en la línea de pronóstico correspondiente al año 2020. Esta anomalía podría explicarse por el impacto de la pandemia de COVID-19, que afectó a varios países, provocando una caída significativa en la actividad económica y, por ende, en el consumo de energía eléctrica. A partir de 2021, se identifica una tendencia de recuperación en la demanda, lo que refleja una estabilización progresiva de la economía. Esta recuperación sugiere que el modelo es sensible a eventos atípicos y puede adaptarse a cambios abruptos en el comportamiento de la demanda, proporcionando así un análisis robusto para escenarios futuros.

Para evaluar la precisión del modelo, se analizaron las siguientes métricas de error, definidas según (2) y (3).

En la Tabla III se presentan las métricas de error, calculadas utilizando el 30% de los datos del modelo correspondiente al período 2014-2021. El refleja la diferencia promedio entre los valores históricos y los valores pronosticados de la demanda eléctrica.

TABLA III
MÉTRICAS DE ERROR

MAD	80.8998 MW
MAPE	1.7447%

En este caso, el valor de 80.89 MW indica que, en promedio, existe una desviación de aproximadamente esa magnitud entre la demanda observada y la estimada, ya sea por exceso o por defecto, en relación con los datos históricos. Este valor podría considerarse elevado si la previsión realizada hubiera sido en una pequeña extensión de territorio, como un distrito. Sin embargo, la previsión fue realizada a nivel nacional para el cual dicho valor del MAD es ínfimo en términos porcentuales.

Por otro lado, el MAPE representa el porcentaje promedio de las diferencias absolutas entre los valores reales y los pronosticados, en relación con los valores observados. Con un valor de 1.74%, este resultado sugiere que el modelo presenta un bajo margen de error relativo, lo que indica un buen desempeño en la capacidad predictiva. Este porcentaje, cercano al 2%, es consistente con resultados aceptables en estudios de pronóstico similares, lo que refuerza la viabilidad del modelo de regresión múltiple para realizar estimaciones futuras de la demanda eléctrica.

B. Discusión

Para un pronóstico a futuro se considera una tasa de crecimiento anual para cada variable independiente que son: población 1.3%, PBI 3.7%, producción de energía 2.33% y la temperatura promedio 0.2%. Posterior a ello, se realiza una ampliación de datos de cada variable a partir del 2021 hasta 2038, tomando en cuenta la tasa de crecimiento. Por último, se vuele a realizar el cálculo ecuación (1) considerando los coeficientes de regresión, para el obtener los resultados del pronóstico de la demanda que se muestra en la Fig. 2.

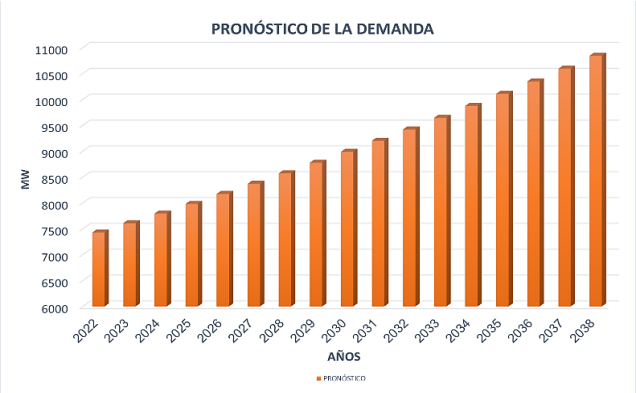

Fig. 2. Curva pronóstico a futuro.

En el gráfico de barras se observa una tendencia continua de crecimiento en la demanda pronosticada año tras año desde 2022 hasta 2038, sin ningún año de decremento, alcanzando un valor aproximado de 10,841 MW. La tasa de crecimiento anual de la demanda entre 2022 y 2038 es aproximadamente del 2.39%. Esto indica que, en promedio, la demanda está creciendo un 2.39% cada año durante este período. Según resultados obtenidos se hacen comparación en [20].

IV. Conclusiones

Este trabajo presentó un análisis de la previsión de la demanda eléctrica utilizando un modelo de regresión lineal múltiple. El modelo propuesto se aplicó al caso peruano, obteniendo previsiones anuales. Los resultados fueron obtenidos considerando el promedio de datos diarios, semanales y mensuales, lo que permitió obtener proyecciones más consistentes sobre el crecimiento futuro de la demanda eléctrica.

Los resultados muestran que el modelo ofrece un buen nivel de ajuste, con un coeficiente de determinación (R^2) del 99%, lo que sugiere que puede representar de manera adecuada la relación entre las variables consideradas. Además, se obtuvieron métricas de error satisfactorias, con un error absoluto medio (MAD) de 80,89 y un error absoluto porcentual medio (MAPE) de 1,74%. Estos valores indican un buen desempeño del modelo en la predicción de la demanda eléctrica, especialmente al compararlos con estudios similares donde los errores suelen ser mayores.

El uso de cuatro variables independientes —población, producto bruto interno (PBI), producción de energía y temperatura— ha contribuido a mejorar la capacidad explicativa del modelo, permitiendo captar de forma más precisa algunos de los factores que influyen en la demanda eléctrica.

Los resultados alcanzados indican que el modelo de regresión lineal múltiple utilizado en este estudio es una herramienta adecuada para la previsión de la demanda eléctrica a largo plazo. No obstante, se sugiere que, en futuros trabajos, se considere la inclusión de nuevas variables o la aplicación de otros métodos estadísticos y de aprendizaje automático para evaluar posibles mejoras en la capacidad predictiva del modelo.

References

[1] X. Li, Z. Wang, C. Yang, and A. Bozkurt, "An advanced framework for net electricity consumption prediction: Incorporating novel machine learning models and optimization algorithms," Energy, vol. 296, pp. 131-259, 2024.

[2] Y. Chakhchoukh, P. Panciatici, and L. Mili, "Electric load forecasting based on statistical robust methods," IEEE Trans. Power Syst., vol. 26, no. 3, pp. 982-991, Aug. 2011.

[3] B. Xiao, G. Peng-wei, M. Gang, Y. Gan-gui, L. Ping, C. Hong-wei, L. Jie-fu, and B. Yang, "A spatial load forecasting method based on the theory of clustering analysis," Energy, vol. 24, pp. 176-183, Jan. 2011.

[4] S. Shing, A. Rasheed, P. Kil-Heum, and K. Veluvolu, "Fast and accurate short-term load forecasting with a hybrid model," Electronics, vol. 13, no. 6, pp. 1079, 2024.

[5] B. Magalhães, P. Bento, J. Pombo, M. Calado, and S. Mariano, "Short-term load forecasting based on optimized random forest and optimal feature selection," Energies, vol. 17, no. 8, pp. 19-26, 2024.

[6] Y. Yang, J. Che, Y. Li, Y. Zhao, and S. Zhu, "An incremental electric load forecasting model based on support vector regression," Energy, vol. 113, pp. 796-808, 2016.

[7] M. Joel David, A. Padilha-Feltrin, and E. M. Carreno, "Spatial pattern recognition of urban sprawl using a geographically weighted regression for spatial electric load forecasting," IEEE, vol. 15, pp. 978-5090, Feb. 2015.

[8] M. AlRashidi and K. El-Naggar, "Long-term electric load forecasting based on particle swarm optimization," Appl. Energy, vol. 87, no. 1, pp. 320-326, 2010.

[9] S. Fan and R. J. Hyndman, "Forecasting electricity demand in Australian National Electricity Market," in Proc. IEEE Power Energy Soc. Gen. Meeting, San Diego, CA, USA, pp. 1-4, 2012.

[10] R. Bareth, M. Kochar, and A. Yadav, "Electrical load forecasting based on multiple regression using Excel tools," in Proc. 2023 3rd Int. Conf. Adv. Electr., Comput., Commun. Sustain. Technol. (ICAECT), Bhilai, India, pp. 1-4, 2023.

[11] Ministerio de Energía y Minas, "Anuario estadístico de electricidad 2021," Oct. 18, 2022. [Online]. Available: https://www.gob.pe/institucion/minem/informes-publicaciones/3600118-anuario-estadistico-de-electricidad-2021. [Accessed: May 24, 2024].

[12] Banco Central de Reserva del Perú, "PBI (millones US$)," 2024. [Online]. Available: https://estadisticas.bcrp.gob.pe/estadisticas/series/trimestrales/resultados/PN02635BQ/html. [Accessed: May 24, 2024].

[13] Instituto Nacional de Estadística e Informática, "Anuario estadístico de electricidad 2021," 2021. [Online]. Available: https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1722/. [Accessed: May 24, 2024].

[14] D. Wu, B. Wang, D. Precup, and B. Boulet, "Multiple kernel learning-based transfer regression for electric load forecasting," IEEE Trans. Smart Grid, vol. 11, no. 2, pp. 1183-1192, Mar. 2020.

[15] T. Hong, M. Gui, M. E. Baran, and H. L. Willis, "Modeling and forecasting hourly electric load by multiple linear regression with interactions," in Proc. IEEE PES Gen. Meeting, Minneapolis, MN, USA, pp. 1-8, 2010.

[16] N. Amral, C. S. Ozveren, and D. King, "Short-term load forecasting using multiple linear regression," in Proc. 2007 42nd Int. Univ. Power Eng. Conf., Brighton, UK, pp. 1192-1198, 2007.

[17] F. M. Tuaimah and H. M. A. Abass, "Short-term electrical load forecasting for Iraqi power system based on multiple linear regression method," Int. J. Comput. Appl., vol. 100, pp. 975-8887, Aug. 2014.

[18] S. M. Badran and O. B. Abouelatta, "Forecasting electrical load using ANN combined with multiple regression method," Res. Bull. Jordan ACM, vol. 2, no. 2, pp. 152-158, 2011.

[19] N. Amral, C. S. Ozveren, and D. King, "Short-term load forecasting using multiple linear regression," in Proc. 2007 42nd Int. Univ. Power Eng. Conf., Brighton, UK, pp. 1192-1198, 2007.

[20] F. M. Tuaimah and H. M. A. Abass, "Short-term electrical load forecasting for Iraqi power system based on multiple linear regression method," Int. J. Comput. Appl., vol. 100, pp. 975-8887, Aug. 2014.