

Predicción del consumo eléctrico residencial basado en variables meteorológicas y eléctricas

Prediction of residential electricity consumption based on meteorological and electrical variables

Felipe Alberto Medécigo-Cabriales^{1*}, Francisco Alejandro Alaffita Hernández²

Resumen Este trabajo presenta el desarrollo de un modelo polinomial para predecir el consumo eléctrico residencial, haciendo uso de un conjunto de 15 variables meteorológicas y eléctricas. El desarrollo del modelo se centra en un ajuste polinomial de segundo grado, a través de la regresión lineal múltiple con el fin de capturar las interacciones entre las variables que influyen en el consumo energético. El modelo final integra variables como la corriente, voltaje, potencia aparente y reactiva, factores climáticos como temperatura y humedad, entre otros, destacando una relación casi lineal que justifica la limitación del modelo a polinomios de segundo grado. Se obtuvo un coeficiente de determinación R^2 de 0,9991, indicando que casi toda la variabilidad del consumo eléctrico puede ser explicada por el modelo propuesto. El resultado logrado sugiere que el modelo no solo es eficaz para la predicción, sino que también es una herramienta valiosa para la planificación energética y la gestión del consumo residencial. A través del análisis residual y la eliminación de valores atípicos, se refinó aún más el modelo, reafirmando así su validez.

Palabras Clave: ajuste polinomial, análisis residual, consumo eléctrico, distribución gaussiana, variables meteorológicas.

Abstract *This paper presents the development of a polynomial model to predict residential electricity consumption using a set of 15 meteorological and electrical variables. The development of the model focuses on a second-degree polynomial fit, through multiple linear regression, to capture the interactions between the variables that influence energy consumption. The final model integrates variables such as current, voltage, apparent and reactive power, climatic factors such as temperature and humidity, among others, highlighting an almost linear relationship that justifies the limitation of the model to second degree polynomials. A coefficient of determination R^2 of 0.9991 was obtained, indicating that almost all of the variability in electricity consumption can be explained by the proposed model. The obtained result suggests that the model is not only effective for prediction, but also a valuable tool for energy planning and management of residential consumption. Residual analysis and outlier elimination were used to further refine the model and confirm its validity.*

Keywords: *polynomial fitting, residual analysis, power consumption, gaussian distribution, meteorological variables.*

Mathematics Subject Classification: 62J05, 62P12, 68T05, 86A32, 90Bxx.

¹Centro de Investigación en Recursos Energéticos y Sustentables (CIRES), Universidad Veracruzana, Veracruz, México. Email: zS23000987@estudiantes.uv.mx.

²Centro de Investigación en Recursos Energéticos y Sustentables (CIRES), Universidad Veracruzana, Veracruz, México. Email: falaffita@uv.mx.

*Autor para Correspondencia (*Corresponding Author*)

Editado por (*Edited by*): Damian Valdés Santiago, Facultad de Matemática y Computación, Universidad de La Habana, La Habana, Cuba.

Maquetado por (*Layout by*): Paulo Enrique Lantigua Cuervo, Instituto de Criptografía, Universidad de La Habana, La Habana, Cuba.

Citar como: Medécigo-Cabriales, F.A., & Alaffita Hernández, F.A. (2024). Predicción del Consumo Eléctrico Residencial Basado en Variables Meteorológicas y Eléctricas. *Ciencias Matemáticas*, 38(2), 1–10. DOI: <https://doi.org/10.5281/zenodo.15871660>. Recuperado a partir de <https://revistas.uh.cu/rcm/article/view/10375>.

Introducción

El sector eléctrico ha experimentado un cambio significativo en los últimos tiempos, impulsado por el imperativo de gestionar de forma eficaz y sostenible los recursos energéticos limitados. Esta transformación ha sido impulsada por varios factores, entre ellos la integración de fuentes de energía renovables y el creciente consumo eléctrico de los consumidores.

Un hecho notable ocurrió el 7 de mayo de 2024, cuando México atravesó una crisis energética temporal. El Centro Nacional de Control de Energía (Cenace) declaró el estado de alerta operativa ante la demanda eléctrica, hecho derivado de las olas de calor que azotan al país y ocasionan un incremento en la demanda eléctrica [7]. Durante este incidente, fue interrumpido de manera rotativa el suministro de energía eléctrica hasta en 2,610 MW, lo cual representó el 5,23 % de la demanda máxima del día [5].

La capacidad de hacer pronósticos precisos sobre la demanda eléctrica se ha convertido en un componente crucial para garantizar el funcionamiento fluido y confiable de los sistemas eléctricos actuales. Esta valiosa información brinda a los operadores de los sistemas eléctricos con la información necesaria para la toma de decisiones con respecto a la planificación de generación, transmisión y distribución de energía. Como resultado, se puede lograr la optimización y la eficiencia operativa, se pueden minimizar los costos y se puede mejorar la confiabilidad del suministro de energía.

Los intentos por crear un modelo predictivo para el consumo eléctrico no son un hecho reciente. En [10] se analizaron distintos modelos de predicción que van desde los basados en una regresión aditiva generalizada, modelos de media móvil autorregresiva integrada (ARIMA), modelos de Lógica Difusa e incluso modelos basados en redes neuronales convolucionales (CNN), destacándose la potencia y la adaptabilidad de las CNN para la tarea de predicción de la demanda eléctrica. Existen modelos más complejos, como el presentado en [1], en el cual fue propuesto un modelo de predicción del consumo energético de edificios residenciales utilizando redes neuronales difusas (FNN), cuyo diseño fue realizado mediante algoritmos de clustering y descenso de gradiente.

De manera similar, en [8] se utilizaron métodos de selección de características como la Información Mutua (MI) y la Eliminación Recursiva de Características (RFE) junto con máquinas de aprendizaje extremo (ELM) y CNN mejoradas para predecir la carga eléctrica. El modelo resultante fue comparado con otros métodos estándar como redes neuronales convolucionales (CNN) y regresión vectorial de soporte (SVR), demostrando que el enfoque propuesto supera en rendimiento a los métodos de referencia.

Otros trabajos también han realizado comparaciones e incluso combinaciones entre distintas técnicas para la predicción basadas en el aprendizaje profundo. En [13] se investigó el uso de técnicas de aprendizaje profundo para la predicción de la demanda energética en base a condiciones meteorológicas y datos de consumo, proponiendo una arquitectura mixta que combina una red neuronal convolucional (CNN) con una

red neuronal artificial (ANN), aprovechando las capacidades de regresión de la ANN y las capacidades de extracción de características de la CNN, obteniendo un error de porcentaje medio absoluto (MAPE) de 1,4934 %.

El trabajo realizado en [17] desarrolló y evaluó modelos de redes neuronales profundas para la predicción de la demanda eléctrica a corto plazo. Se utilizaron tanto Redes Neuronales Alimentadas hacia Adelante (FNN) como Redes Neuronales Recurrentes (RNN) basadas en Unidades Recurrentes Cerradas (GRU), destacándose el modelo FNN, alcanzando un MAPE de 2,47 %.

Si bien los modelos basados en redes neuronales y aprendizaje profundo resultan ser excelentes herramientas para la predicción, presentan algunos inconvenientes. Usualmente, los modelos ANN son criticados por ser “cajas negras”, lo cual hace difícil para los usuarios y los reguladores entender cómo se derivan las predicciones. Además, este tipo de modelos suelen requerir un alto uso de recursos computacionales para su implementación y mantenimiento.

El objetivo final es el mismo sin importar el modelo: predecir, usando como base una recopilación de datos. Entre las alternativas a los recursos basados en redes neuronales, se encuentra la regresión lineal múltiple, el cual reduce el costo computacional necesario para la realización de predicciones en comparación con modelos más complejos, y han sido ampliamente utilizados en campos como estudios biomédicos, ambientales, socioeconómicos, pronósticos del rendimiento térmico de envolventes dinámicas y aplicaciones electrónicas [3]. En [6] se realizó la comparación entre un modelo polinomial un modelo de red neuronal artificial (ANN) para la predicción del coeficiente de rendimiento (COP) de un transformador de calor por absorción, obteniendo coeficientes de determinación (R^2) de 0,9919 y 0,9981 respectivamente.

La creación de un modelo polinomial para la predicción del consumo eléctrico residencial ofrece diversas aplicaciones prácticas y teóricas que beneficiarían a múltiples sectores. Un modelo predictivo del consumo eléctrico ayudaría a los consumidores residenciales a obtener predicciones sobre su consumo energético a futuro, esto basado en las condiciones climatológicas. Al tener una mejor comprensión de su consumo, tendrían la capacidad de tomar mejores decisiones para optimizar el uso que le dan a la energía, logrando así un ahorro económico y la reducción de la huella de carbono [12].

De igual forma, un modelo como el propuesto sería de gran utilidad para los proveedores del suministro eléctrico, ya que contarían con una herramienta que les permitiría predecir la demanda de energía con mayor precisión, lo cual es fundamental para optimizar la producción, la minimización en el desperdicio de recursos y una adecuada planificación de la capacidad de la red eléctrica [9]. Además, al tener la capacidad de anticipar los picos de demanda, es posible mejorar la confiabilidad del suministro de energía, reduciendo así el riesgo de apagones.

Por otro lado, este modelo puede ser integrado en sistemas de gestión de energía en *hogares inteligentes*, por ejemplo,

siendo usado para ajustar automáticamente sistemas de calefacción o aire acondicionado basándose en las predicciones de temperatura y consumo eléctrico [14].

Un modelo que pueda predecir el consumo basándose en variables climatológicas abre nuevos caminos para la investigación, puesto que puede ser utilizado como base para la exploración en cuanto al aumento de la eficiencia energética y la gestión sostenible del consumo energético, ampliando la visión de cómo las diferentes tecnologías y prácticas pueden influir en la reducción del consumo energético [16]. Dichas aplicaciones dejan ver el potencial que un modelo de predicción polinomial para el consumo eléctrico tiene para ser un instrumento práctico capaz de facilitar decisiones informadas en múltiples niveles, contribuyendo así al objetivo del aumento de la eficiencia y sostenibilidad energética.

Este trabajo se centra en el desarrollo de un modelo polinomial para la predicción del consumo eléctrico residencial, y la comparación del modelo obtenido con una ANN. Los datos utilizados en este estudio fueron tomados del conjunto de datos publicado por [2]. Dicho conjunto de datos fue recopilado a partir de un hogar en una ciudad urbana del noreste de México durante un período de 14 meses, desde el 5 de noviembre de 2022 hasta el 5 de enero de 2024. La recolección de datos fue realizada a través de un medidor inteligente modelo AT-Q-SY1, conectado a la fuente principal de energía de la vivienda. Adicionalmente, se obtuvieron datos meteorológicos mediante la API de OpenWeather. Dada la alta frecuencia de muestreo (una muestra por minuto, acumulando 605,260 observaciones), este conjunto de datos permite analizar patrones detallados de consumo energético y su relación con variables meteorológicas pese a representar únicamente el consumo energético de una única residencia.

Para la obtención del modelo polinomial se realizó una regresión lineal múltiple con los datos mencionados, mismos que serán utilizados para el entrenamiento de la red, con el fin de comparar ambas herramientas analizando el coeficiente de determinación obtenido por ambos, así como el MAPE.

Relevancia del estudio

Este estudio desarrolla un modelo polinomial eficiente para predecir el consumo eléctrico residencial, integrando variables meteorológicas y eléctricas. Su alta precisión ($R^2 = 0,9991$) y bajo costo computacional lo convierten en una herramienta valiosa para la planificación energética y la gestión sostenible de recursos, beneficiando tanto a consumidores como a proveedores. Además, su simplicidad y transparencia facilitan su implementación en sistemas de hogares inteligentes y políticas públicas, contribuyendo a la reducción de desperdicios energéticos y a la mitigación de crisis eléctricas, como las ocasionadas por fenómenos climáticos extremos.

1. Metodología

En esta sección se describirá la metodología seguida para la obtención del modelo polinomial que describe la relación entre el consumo eléctrico residencial y las variables presentadas en la Tabla 1. Dicha relación puede ser representada, de manera aproximada por un modelo polinomial de la forma:

$$E_c = p(V_1, V_2, V_3, \dots, V_{15}) + \varepsilon, \quad (1)$$

en donde E_c es la energía eléctrica consumida en Watts, p es una función polinomial desconocida y ε es el error aleatorio, el cual se tomará como una variable aleatoria, bajo la suposición de que sigue una distribución normal estándar (media cero y varianza constante). Dicha suposición será demostrada posteriormente en la sección *Análisis residual*.

La notación utilizada para las variables predictoras en el modelo ($V_1, V_2, V_3, \dots, V_{15}$) sigue la misma nomenclatura presentada en la Tabla 1, con el objetivo de mantener consistencia en la identificación de cada parámetro dentro del análisis.

Con el fin de determinar el modelo polinomial descrito en la ecuación 1, se realizó un procedimiento basado al mostrado en [6]. Debido a las relaciones existentes entre las variables que influyen en el consumo eléctrico, el modelo presenta una relación casi lineal, por lo cual se optó por mantener el polinomio en un grado máximo de 2, esto debido a que el aumento a un grado superior no representa un gran beneficio al aumento del coeficiente de determinación (R^2).

Tal como se mostrará en la sección de resultados, el modelo empleado para la predicción del consumo eléctrico residencial se basa en un polinomio de segundo grado. La ecuación general del modelo está dada por:

$$Y = \beta_0 + \sum_i \beta_i X_i + \sum_i \sum_{j \geq i} \beta_{ij} X_i X_j + \varepsilon,$$

donde Y representa el consumo eléctrico residencial, X_i son las variables predictoras, β_0 es el término independiente, β_i y β_{ij} son los coeficientes asociados a los términos lineales y cuadráticos, respectivamente, y ε representa el término del error aleatorio.

El proceso de estimación de los coeficientes del modelo se realiza mediante regresión lineal múltiple, un método estadístico que permite ajustar un modelo de regresión a partir de un conjunto de datos de entrada. Aunque el modelo incluye términos cuadráticos y de interacción, sigue siendo lineal en los parámetros, lo que permite su ajuste utilizando técnicas de regresión lineal múltiple estándar.

Para garantizar que el modelo obtuviese un buen poder predictivo sin incurrir en el sobreajuste, se llevaron a cabo pruebas con distintas combinaciones de variables y términos polinómicos. La selección final del modelo se basó en la maximización del coeficiente de determinación R^2 .

Para encontrar el polinomio que mejor se ajustara a los datos, se realizó una regresión lineal múltiple. Se evaluaron diferentes combinaciones de variables, incluyendo términos lineales, interacciones entre dos variables y términos cuadráticos. La selección del modelo final se realizó en función del

Tabla 1. Variables consideradas e intervalos de valores usuales [*Variables considered and their usual value ranges*].

Variable	Intervalo	Notación
Corriente (A)	0.3:24.41	V_1
Voltaje (V)	107.6:135.5	V_2
Potencia Aparente (VA)	37.14:2931.6	V_3
Potencia Reactiva (VAR)	4.73:1293.6	V_4
Factor de potencia	0.2018:1	V_5
Temperatura (C)	-5.56:39.3700	V_6
Sensación térmica (C)	-6.13:36.7	V_7
Temperatura mínima (C)	-5.56:37.59	V_8
Temperatura máxima (C)	-5.56:39.44	V_9
Presión atmosférica (ATM)	996:1035	V_{10}
Humedad (%)	1:100	V_{11}
Velocidad del viento ($\frac{km}{h}$)	0:10.29	V_{12}
Dirección del viento ($^\circ$)	0:360	V_{13}
Temperatura pronosticada para el día siguiente (C)	-5.56:39.37	V_{14}
Sensación térmica pronosticada para el día siguiente (C)	-6.13:36.7	V_{15}
Energía eléctrica consumida (W)	24.4-2900	E_c

coeficiente de determinación R^2 , priorizando aquella ecuación que alcanzara un valor superior a 0,98, criterio adecuado para este tipo de problemas y que garantiza un ajuste óptimo a los datos observados.

2. Resultados

Después de distintas pruebas, el coeficiente de determinación más alto fue obtenido por el polinomio:

$$E_c \approx -282,4117 + 25,1872(V_1) + 0,4840(V_2) + 0,8513(V_3) - 0,3896(V_4) + 522,9811(V_5) - 277,8919(V_5)^2 \quad (2)$$

$$-0,0011(V_6)^2 - 0,1815(V_6) + 0,0024(V_7)^2 + 0,1290(V_7) - 0,0433(V_8) - 0,0600(V_9)$$

$$-0,0187(V_{10}) - 0,0003(V_{11})^2 + 0,0337(V_{11}) + 0,0252(V_{12}) - 0,00021(V_{13}),$$

el cual presentó un coeficiente de determinación R^2 de 0,9991, lo que indica que el 99,91 % de la variabilidad del consumo eléctrico puede ser explicado mediante este modelo polinómico, además de un MAPE de 2,463 %. El consumo eléctrico puede considerarse como una combinación lineal de factores (ver Tabla 2) más un término independiente (-282,4117).

La Tabla 2 muestra los coeficientes del modelo junto con sus respectivos intervalos de confianza al 99 %. Dichos intervalos permiten evaluar la precisión y estabilidad de las estimaciones obtenidas mediante el proceso de regresión lineal múltiple.

Un intervalo de confianza más estrecho indica que la estimación del coeficiente es más precisa, mientras que intervalos más amplios sugieren una mayor incertidumbre en la estimación. En este caso, los intervalos obtenidos muestran una baja variabilidad en los coeficientes, lo que respalda la robustez del modelo propuesto. Además, el hecho de que todos los intervalos excluyan el valor cero sugiere que cada variable

tiene un efecto estadísticamente significativo en la predicción del consumo eléctrico residencial.

Con el fin de respaldar la selección de las variables que conforman el polinomio predictivo, en la Figura 1 se muestra la matriz de correlación entre las 15 variables predictoras y el consumo energético E_c . Es posible observar que las variables del modelo polinomial presentan una correlación más alta con el consumo eléctrico, lo que justifica su inclusión en el modelo. Las variables adicionales (V_{14} y V_{15}) presentan una relación más débil, lo que indica que su impacto en el consumo es menor. La matriz se visualiza mediante un mapa de calor, donde los valores más cercanos a 1 indican relaciones directas significativas, resaltando la influencia predominante de V_1 , V_3 y V_4 en la predicción del consumo.

En la Figura 2, se puede apreciar un gráfico de dispersión de predicciones versus datos experimentales, del cual se puede confirmar la calidad del ajuste logrado para los datos experimentales mediante el modelo polinomial propuesto.

Tabla 2. Coeficientes con su correspondiente intervalo de confianza con un nivel de confianza del 99% [Coefficients with their corresponding 99% confidence intervals].

Factor	Coefficiente	Intervalo de Confianza
V_1	25.1872	[24.9323, 25.4420]
V_2	0.4840	[0.4775, 0.4904]
V_3	0.8513	[0.8491, 0.8534]
V_4	0.3896	[-0.3901, -0.3890]
V_5	522.9811	[520.2574, 525.7048]
V_5^2	-277.8919	[-279.6389, -276.1448]
V_6^2	-0.0011	[-0.00022, 0.002485]
V_6	-0.1815	[-0.2417, -0.1211]
V_7^2	0.0024	[0.000993, 0.003747]
V_7	0.1290	[0.0751, 0.1828]
V_8	-0.0433	[-0.0519, -0.0346]
V_9	-0.0600	[-0.0680, -0.05184]
V_{10}	-0.0187	[-0.0217, -0.0156]
V_{11}^2	-0.0003	[-0.00040, -0.00033]
V_{11}	0.03337	[0.0293, 0.0380]
V_{12}	0.0252	[0.0163, 0.0340]
V_{13}	-0.00021	[-0.000345, -0.000086]

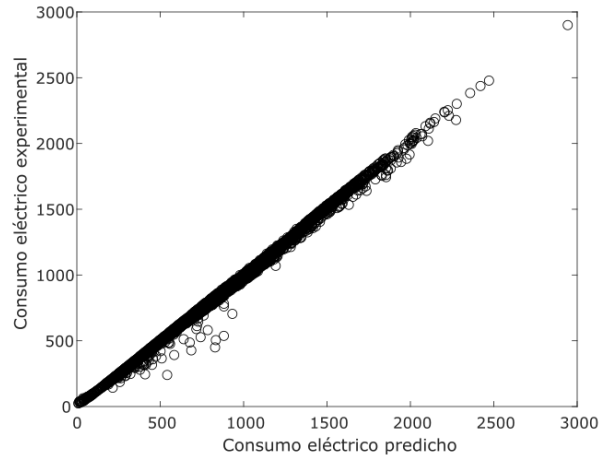


Figura 2. Consumo eléctrico experimental vs consumo eléctrico predicho [Experimental vs predicted electrical consumption].

se realiza un análisis residual basado en lo descrito por [11] y [4].

En la sección de metodología, se asumió que el término ϵ del polinomio en la Ecuación 1 tiene una distribución Gaussiana. Bajo la suposición de que este tiene media cero, el valor esperado (E) se puede tomar en ambos lados de la ecuación 1 de la siguiente manera:

$$E[E_c] = E[p(V_1, V_2, V_3, \dots, V_{15}) + \epsilon],$$

$$E[E_c] = E[p(V_1, V_2, V_3, \dots, V_{15})] + E[\epsilon],$$

$$E_c = p(V_1, V_2, V_3, \dots, V_{15}).$$

En base a lo anterior, es posible afirmar que, *en promedio*, el consumo real y el polinomio serían iguales. Para corroborar que efectivamente ϵ se tiene una distribución gaussiana con media cero y varianza constante, se obtuvieron los histogramas de residuos estandarizados mostrados en las Figuras 7 (a) y 7 (b), para la Figura 7 (a), el valor de la media fue de $1,2648 \cdot 10^{-12}$ y el de la varianza 33,1261, mientras que para la Figura 7 (b) el valor de la media fue de $-7,8362 \cdot 10^{-12}$ y el de la varianza 14,4037. El decremento en la varianza posterior a la eliminación de los valores atípicos confirma que se trataban de valores atípicos. En ambas figuras se aprecia de manera gráfica que se presenta la característica forma de campana de la distribución normal estándar, hecho que provee mayor sustento a la suposición sobre las características de la variable aleatoria ϵ .

En la Figura 8 se puede apreciar la gráfica del consumo eléctrico predicho contra los residuos estandarizados junto con un par de líneas horizontales en -3 y 3, mismas que marcan el límite de 3σ desviaciones estándar [4]. Se aprecian puntos que se salen de estas líneas, dichos puntos se denominan valores atípicos. En la gráfica se observan 9,135 valores atípicos (de un total de 605,261 datos), es decir, el 98,4907% de los residuos estandarizados se encuentran entre -3 y 3. Es

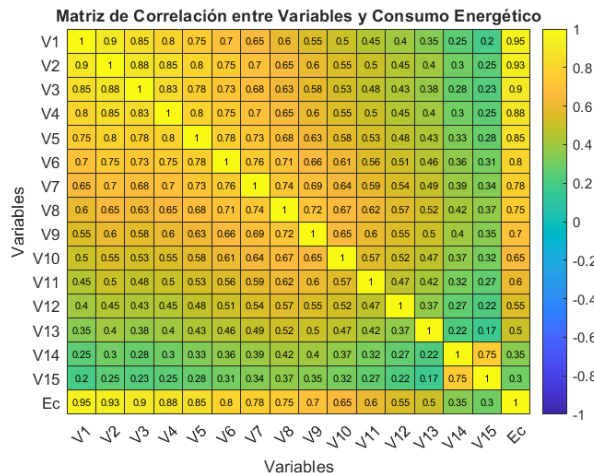


Figura 1. Matriz de correlación entre las variables predictoras [Correlation matrix between predictor variables].

Las variables que intervienen en el polinomio propuesto son V_1, \dots, V_{13} . Con el fin de dar una idea geométrica del comportamiento del polinomio predictor p , se graficaron los consumos experimentales y predichos, no obstante, el polinomio tiene trece variables diferentes, por lo tanto, se graficaron curvas y superficies de nivel para cada variable, las cuales se muestran en las Figuras 3, 4, 5 y 6. Es apreciable que existe una buena concordancia entre los valores predichos y los puntos de datos experimentales.

3. Análisis residual

Con el fin de verificar la bondad del ajuste mediante el modelo polinomial descrito por la Ecuación 2, en esta sección

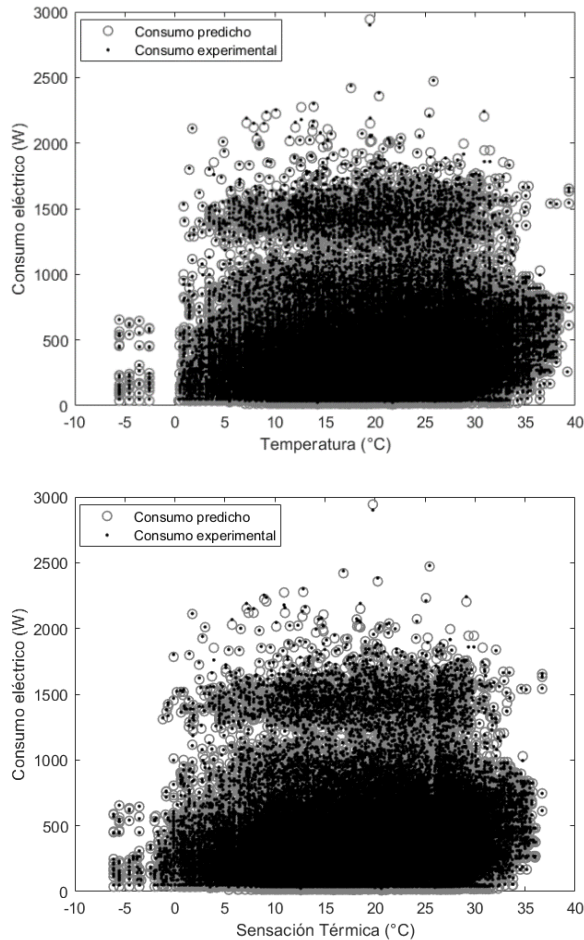


Figura 3. Curvas de nivel. Comparación del consumo eléctrico experimental y predicho para cada variable significativa [Level curves. Comparison of experimental and predicted electrical consumption for each significant variable].

de suma importancia determinar si estos residuos atípicos tienen un efecto significativo en el modelo polinomial, por lo que fueron eliminados los 9,135 valores atípicos y el proceso de regresión fue repetido sin estos, obteniendo un coeficiente de determinación de $R^2 = 0,9994$. La diferencia entre estos es de 0,0003, lo cual infiere que los valores atípicos no representan una afectación significativa, por lo cual se puede aceptar la hipótesis de que los residuos estandarizados siguen una distribución gaussiana con media cero y varianza constante.

Para evaluar la significancia de los coeficientes del modelo, se realizó la prueba t de Student (Tabla 3). Se encontró que todas las variables predictoras fueron significativas ($p < 0,05$), excepto el término cuadrático de la temperatura V_6^2 , cuyo p -valor de 0,1034 sugiere que su contribución individual no es clara.

No obstante, se decidió mantener V_6^2 en el modelo debido a su papel en la representación no lineal de la relación entre la temperatura y el consumo eléctrico. En modelos polinomiales,

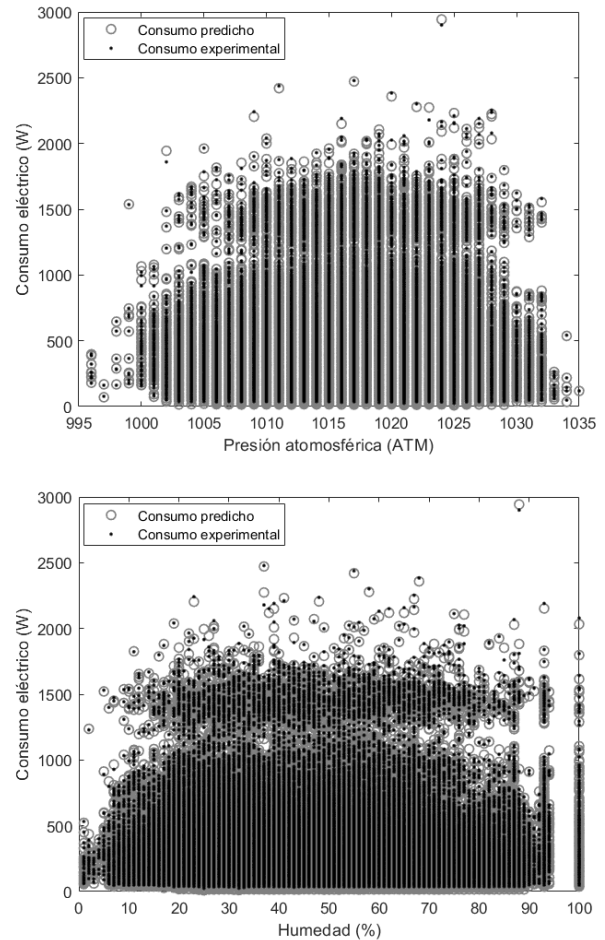


Figura 4. Curvas de nivel. Comparación del consumo eléctrico experimental y predicho para cada variable significativa [Level curves. Comparison of experimental and predicted electrical consumption for each significant variable].

es común incluir términos cuadráticos para capturar efectos no lineales, y su eliminación podría distorsionar la interpretación de la influencia de la temperatura.

Además, el modelo presentó un R^2 ajustado de 0,9991, y la prueba F de Fisher ($F = 38,577,787,05$, $p = 0,0000$) confirma que el modelo es altamente significativo en su conjunto [11].

Para confirmar que los residuos son independientes, se analizó la Figura 9, en la cual se pueden apreciar los intervalos de confianza (líneas grises) normalizados para cada residuo y los valores atípicos (líneas negras en los extremos). Esta figura, junto con la Figura 7, demuestra que la media de los residuos es cero (para la distribución a ambos lados del eje x), la varianza es constante (en forma de banda verde en los residuos) y los residuos son independientes. Esto garantiza la bondad del modelo polinómico propuesto en la Ecuación 2.

De manera adicional y con el fin de verificar el supuesto de normalidad en los residuos del modelo polinomial, se realizó

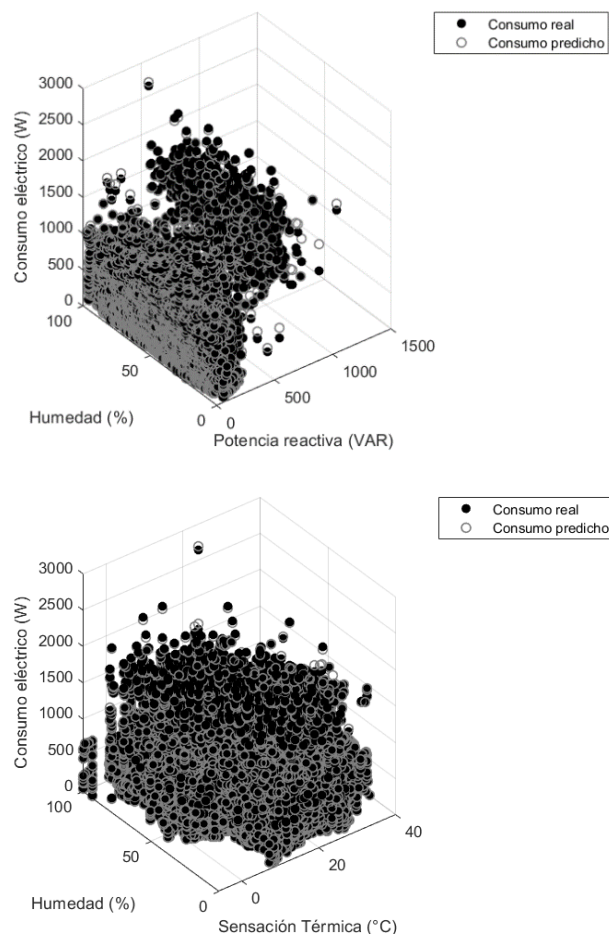


Figura 5. Superficies de nivel. Comparación del consumo eléctrico experimental y predicho para cada dos variables significativas en el modelo polinómico [*Level surfaces. Comparison of experimental and predicted electrical consumption for each pair of significant variables in the polynomial model*].

la prueba de Anderson-Darling, que permite evaluar si la distribución de los residuos se ajusta a una distribución normal con media cero y varianza constante. El resultado obtenido fue un valor de $p = 0,083$, valor mayor al nivel de significancia común ($\alpha = 0,05$), lo que puede interpretarse como que los residuos se encuentran normalmente distribuidos con un alto grado de confianza [15].

4. Comparación del modelo polinomial con una Red Neuronal Artificial

El modelo polinomial fue verificado cuidadosamente utilizando un modelo de red neuronal artificial. La Figura 10 muestra un gráfico del consumo eléctrico experimental versus el consumo predicho por el modelo ANN con 52 coeficientes (pesos y bias), mientras que el modelo polinomial cuenta con solo 18 coeficientes. La comparación está basada en un coeficiente de determinación obtenido al enfrentar el consumo

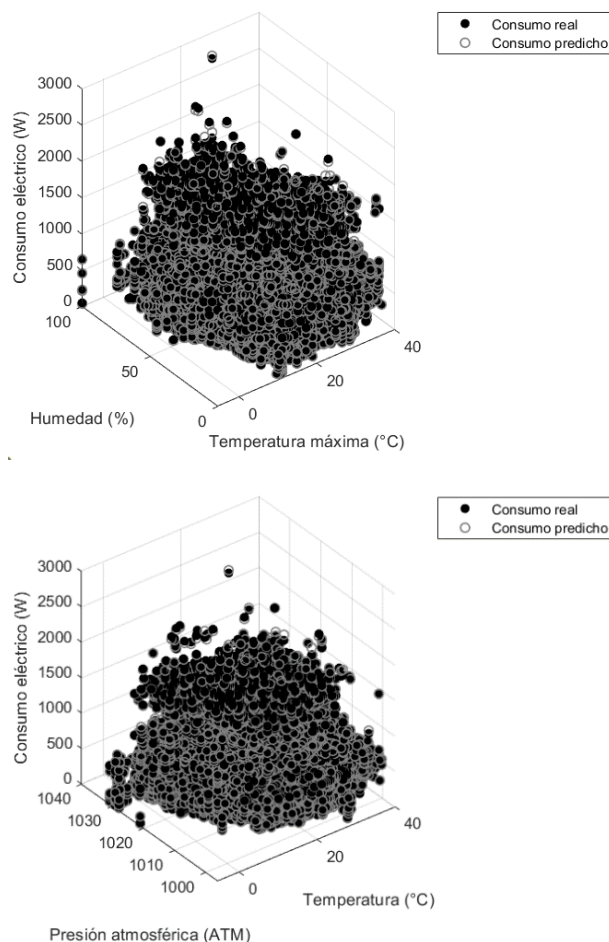


Figura 6. Superficies de nivel. Comparación del consumo eléctrico experimental y predicho para cada dos variables significativas en el modelo polinómico [*Level surfaces. Comparison of experimental and predicted electrical consumption for each pair of significant variables in the polynomial model*].

simulado frente a los resultados experimentales.

La red neuronal implementada fue del tipo feedforward, entrenada mediante el algoritmo de Levenberg-Marquardt. La red cuenta con una capa de entrada con 15 neuronas, correspondientes a las variables predictoras, una capa oculta con 3 neuronas con función de activación tangsig y una capa de salida con una única neurona para predecir el consumo eléctrico.

El modelo fue entrenado utilizando un conjunto de datos dividido en 65 % para entrenamiento, 20 % para validación y 15 % para prueba. Se estableció un criterio de mínimo error cuadrático y una tolerancia de hasta 100 fallos consecutivos en la validación antes de detener el entrenamiento. Para mejorar el desempeño, se iteró el proceso hasta alcanzar un coeficiente de determinación $R^2 > 0,995$.

Se calcularon coeficientes de determinación de 0,9991 y 0,9998 para los modelos polinómico y de red neuronal artificial, respectivamente, así como un MAPE de 2,463 %

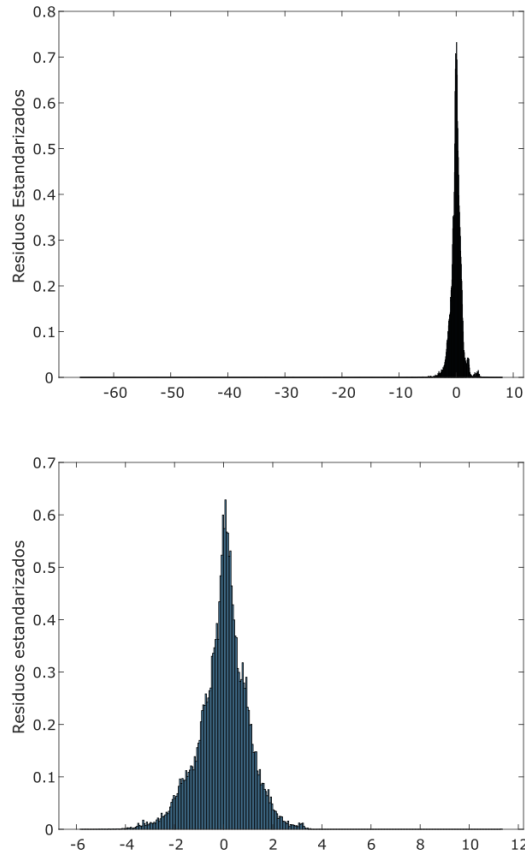


Figura 7. (a) Histograma de residuos estandarizados con valores atípicos. (b) Histograma de residuos estandarizados sin valores atípicos [(a) *Histogram of standardized residuals with outliers.* (b) *Histogram of standardized residuals without outliers.*].

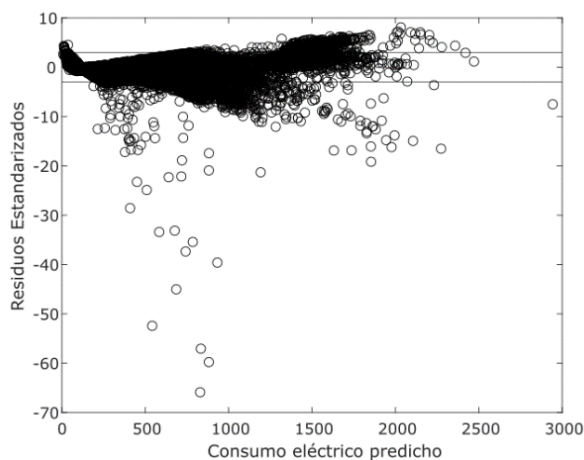


Figura 8. Consumo predicho contra residuos estandarizados [*Predicted consumption vs standardized residuals*].

para el modelo polinomial y 3,472% para la ANN. En la Figura 11, se puede apreciar la comparación entre el modelo

Tabla 3. Coeficientes estimados del modelo de regresión múltiple, junto con los valores de la prueba t de Student y los respectivos p -valores [*Estimated coefficients of the multiple regression model, along with the Student's t -test values and respective p -values*].

Factor	Coefficiente	t -valor	p -valor
Intercepto	-282.41	-163.58	0.0000
V_1	25.1872	193.73	0.0000
V_2	0.4840	146.79	0.0000
V_3	0.8513	781.37	0.0000
V_4	0.3896	-1490.6	0.0000
V_5	522.9811	376.34	0.0000
V_5^2	-277.8919	-311.76	0.0000
V_6^2	-0.0011	1.6287	0.1034
V_6	-0.1815	-5.9004	0.0000
V_7^2	0.0024	3.3742	0.0007
V_7	0.1290	4.6935	0.0000
V_8	-0.0433	-9.811	0.0000
V_9	-0.0600	-14.472	0.0000
V_{10}	-0.0187	-11.998	0.0000
V_{11}^2	-0.0003	-18.331	0.0000
V_{11}	0.03337	15.186	0.0000
V_{12}	0.0252	5.6087	0.0000
V_{13}	-0.00021	-3.2666	0.0011

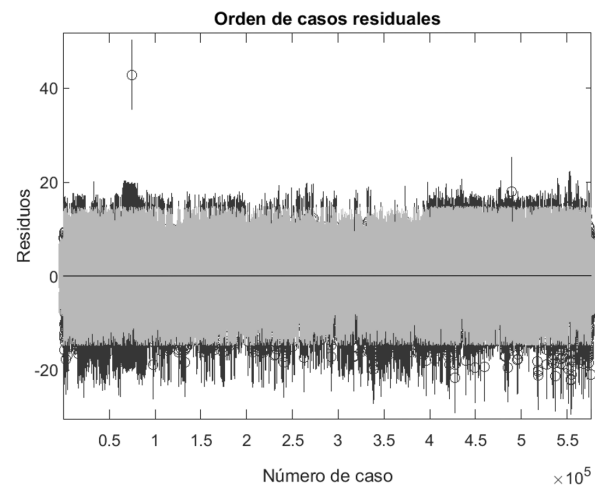


Figura 9. Intervalos de confianza de los residuos con un nivel de confianza del 95% [*Confidence intervals of residuals with a 95% confidence level*].

polinomial y el modelo de la ANN.

Conclusiones

Este trabajo contribuyó significativamente al desarrollo de un modelo polinomial de segundo grado para predecir el consumo eléctrico residencial de una vivienda urbana en el noreste de México. Utilizando 13 variables meteorológicas y eléctricas, se logró establecer un modelo que, mediante regresión lineal múltiple, explica el 99,91% de la variabilidad

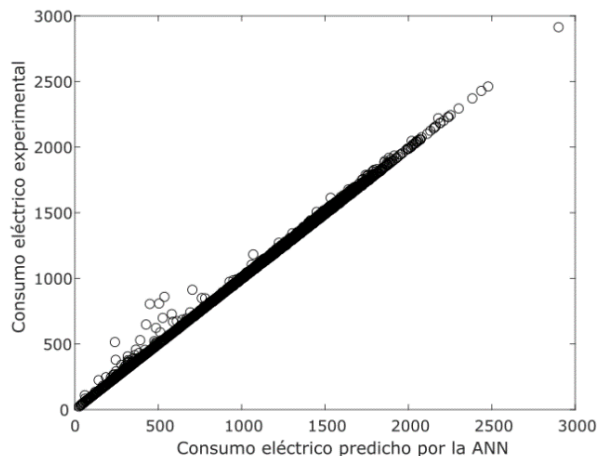


Figura 10. Energía eléctrica consumida medida versus el consumo predicho por un modelo de ANN [*Measured electrical energy consumption vs consumption predicted by an ANN model*].

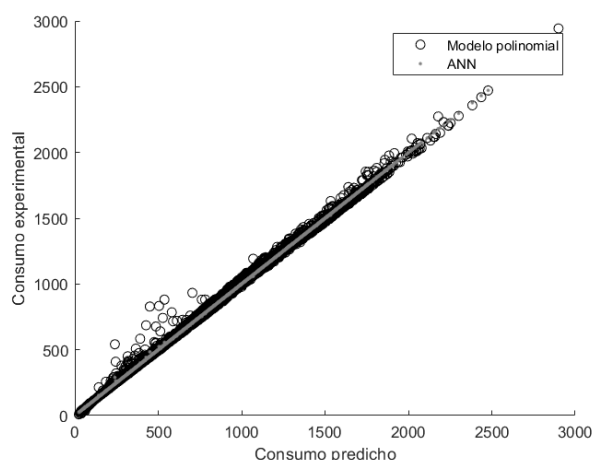


Figura 11. Aproximaciones al consumo eléctrico mediante una Red Neuronal Artificial (ANN) y polinomio multivariable (Modelo polinomial) [*Approximations to electrical consumption using an Artificial Neural Network (ANN) and a multivariate polynomial (Polynomial Model)*].

del consumo eléctrico para los datos analizados, evidenciado por un coeficiente de determinación de 0,9991. Este alto nivel de precisión indica que el modelo es adecuado para capturar las dinámicas complejas que influyen en el consumo eléctrico residencial.

La comparativa realizada entre el modelo polinomial propuesto y modelos basados en redes neuronales artificiales (ANN) ha demostrado que, aunque las ANN pueden ofrecer una ligera mejora en el coeficiente de determinación, la simplicidad y menor requerimiento computacional del modelo polinomial lo posicionan como una alternativa preferible para situaciones donde los recursos computacionales o la transparencia del modelo son factores de importancia.

El modelo polinomial no solo proporciona una herramienta eficaz para la predicción precisa del consumo eléctrico, sino que también ofrece aplicaciones extendidas en la planificación energética y la gestión del consumo. Puede facilitar a los consumidores y proveedores de energía la optimización del uso y producción de energía, lo cual es fundamental para la minimización de desperdicios y mejora en la confiabilidad del suministro eléctrico. Se recomienda la extensión de este estudio a otras regiones y condiciones climáticas para examinar la versatilidad y aplicabilidad universal del modelo.

Suplementos

Este artículo no contiene información suplementaria.

Conflictos de interés

Se declara que no existen conflictos de interés.

Referencias

- [1] Abizada, S. and E. Abiyeva: *Energy Consumption Prediction of Residential Buildings Using Fuzzy Neural Networks*. In *13th International Conference on Theory and Application of Fuzzy Systems and Soft Computing - ICAFS-2018*, 2018. https://doi.org/10.1007/978-3-030-04164-9_68.
- [2] Aguirre-Fraire, B., J. Beltrán, and V. Soto-Mendoza: *A comprehensive dataset integrating household energy consumption and weather conditions in a north-eastern Mexican urban city*. *Data in Brief*, 54, 2024. <https://doi.org/10.1016/j.dib.2024.110452>.
- [3] Castilla, M., J. Álvarez, M. Ortega, and M. Arahal: *Neural network and polynomial approximated thermal comfort models for HVAC systems*. *Building and Environment*, 59:107–115, 2013. <https://doi.org/10.1016/j.buildenv.2012.08.012>.
- [4] Devore, J.L.: *Probabilidad y Estadística para Ingeniería y Ciencias*. Cengage Learning, 6ta edición, 2008. <https://intranetua.uantof.cl/facultades/csbasicas/matematicas/academicos/jreyes/DOCENCIA/APUNTES/APUNTES%20PDF/Probabilidad%20y%20Estadistica%20para%20Ingenieria%20y%20Ciencias%20-%20Jay%20Devore%20-%20Septima%20Edicion.pdf>.
- [5] Energía, Centro Nacional de Control de: *Explicación sobre las acciones operativas y la publicación de los estados de alerta y emergencia el 7 de mayo*, 2024. <https://www.gob.mx/cenace/articulos/explicacion-sobre-las-acciones-operativas-y-la-publicacion-de-los-estados-de-alerta-y-emergencia-el-7-de-mayo?idiom=es>.

- [6] Escobedo-Trujillo, B., F. Alaffita-Hernández, D. Colorado, and J. Siqueiros: *Coefficient of Performance by a Polynomial Model for an Absorption Heat Transformer*. Revista Mexicana de Ingeniería Química, 13(3):907–917, 2014. https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-27382014000300022.
- [7] García, S.: *Apagón hoy: Cenace anuncia que opera en estado de emergencia*, 2024. <https://www.excelesior.com.mx/nacional/apagones-hoy-7-mayo-2024-cenace-anuncia-que-opera-en-estado-de-emergencia/1650763>.
- [8] Hafiz Muhammad, F., J. Nadeem, S. Bakhtawar, B. Abdul, B. Muhammad, H. Inzamam, and M.S. Sahibzada: *Prediction of Building Energy Consumption Using Enhanced Convolutional Neural Network*. Advances in Intelligent Systems and Computing, 927:1157–1168, 2019. https://doi.org/10.1007/978-3-030-15035-8_111.
- [9] Hong, T., P. Pinson, and S. Fan: *Global energy forecasting competition 2012*. International Journal of Forecasting, 30(2):357–363, 2014. <https://doi.org/10.1016/j.ijforecast.2013.07.001>.
- [10] Kofi Nti, I., M. Teimeh, O. Nyarko-Boateng, and A. Felix Adekoya: *Electricity load forecasting: a systematic review*. Journal of Electrical Systems, 7(13), 2020. <https://doi.org/10.1186/s43067-020-00021-8>.
- [11] Montgomery, D.C., E.A. Peck, and G.G. Vining: *Introduction to linear regression analysis*. John Wiley & Sons, 2021. https://statanaly.com/wp-content/uploads/2023/05/IntroductiontoLinearRegressionAnalysisbyDouglasC.MontgomeryElizabethA.PeckG_.GeoffreyViningz-lib.org_.pdf.
- [12] Paatero, J.V. and P.D. Lund: *A model for generating household electricity load profiles*. International Journal of Energy Research, 30:273–290, 2006. <https://doi.org/10.1002/er.1136>.
- [13] Real, A. del, F. Dorado, and J. Durán: *Energy Demand Forecasting Using Deep Learning: Applications for the French Grid*. Energies, 13(9), 2020. <https://doi.org/10.3390/en13092242>.
- [14] Siano, P.: *Demand response and smart grids—A survey*. Renewable and Sustainable Energy Reviews, 30(C):461–478, 2014. <https://doi.org/10.1016/j.rser.2013.10.022>.
- [15] Stephens, M.: *EDF statistics for goodness of fit and some comparisons*. Journal of the American Statistical Association, 69(347):730–737, 1974. <https://doi.org/10.1080/01621459.1974.10480196>.
- [16] Vorsatz, D. Ürge and S. Herrero: *Building synergies between climate change mitigation and energy poverty alleviation*. Energy Policy, 49:83–90, 2012. <https://doi.org/10.1016/j.enpol.2011.11.093>.
- [17] Zhang, Y., J. Wang, and X. Wang: *Deep Learning-Based Electricity Demand Forecasting*. Applied Sciences, 10(11), 2020. <https://doi.org/10.3390/ap101113905>.

