

MODELADO Y PREDICCIÓN DEL ÍNDICE DE INTENSIDAD ENERGÉTICA POR PROVINCIAS DEL ECUADOR UTILIZANDO REDES NEURONALES Y ANÁLISIS DE CONGLOMERADOS CON ALGORITMOS DE APRENDIZAJE AUTOMÁTICO

Xavier Rodriguez-Cruz¹ y Julio Litardo²

Universidad de Guayaquil – Facultad de Ingeniería Industrial
Guayaquil, Ecuador

Información

Recibido:
10 de marzo de 2025

Aceptado:
4 de junio de 2025

Palabras clave:
Intensidad energética
Redes neuronales
artificiales
Aprendizaje automático
K-means
DBSCAN.

JEL:
C45, C53, Q41, O13

DOI:
<https://doi.org/10.47550/RCE/35.1.8>

Resumen

Mediante un diseño experimental comparativo, se evalúan cuatro arquitecturas de redes neuronales feedforward, LSTM, feedforward secuencial y LSTM robusto para predecir el índice de intensidad energética provincial, utilizando datos de panel de 2018 a 2023 con variables claves relacionadas con el consumo energético, la actividad económica y la dinámica demográfica. El modelo feedforward secuencial mostró mayor precisión predictiva ($R^2 = 0,724$), superando el rendimiento de los enfoques recurrentes en la capacidad para capturar patrones temporales, precisión y eficiencia computacional.

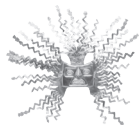
Los métodos de agrupamiento permiten caracterizar el comportamiento de las provincias en términos de intensidad energética, identificando grupos con patrones de consumo energético similares en relación con su producción económica. Orellana, Pastaza, Santa Elena y Zamora Chinchipe destacan por presentar perfiles energéticos diferenciados o atípicos, resaltando la utilidad de algoritmos como DBSCAN para detectar dinámicas regionales singulares.

Los resultados permiten identificar asimetrías territoriales en eficiencia energética, proporcionando una base empírica para el diseño e implementación de políticas energéticas diferenciadas. Estos conocimientos son particularmente valiosos en contextos geográficos y socioeconómicos diversos, donde promover el uso sostenible de la energía y reducir las desigualdades en la distribución y el consumo son objetivos fundamentales.

¹ORCID: 0000-0002-2931-0702. CRediT: investigación, análisis formal, metodología, redacción – borrador original. Correo electrónico: lester.rodriquezc@ug.edu.ec.

²ORCID: 0009-0008-3931-8151. CRediT: conceptualización, supervisión, validación, redacción – revisión y edición. Correo electrónico: julio.litardou@ug.edu.ec.

Copyright © 2025 Rodriguez-Cruz y Litardo. Los autores conservan los derechos de autor del artículo. El artículo se distribuye bajo la licencia Creative Commons Attribution 4.0 License.



MODELING AND PREDICTION OF THE ENERGY INTENSITY INDEX BY PROVINCES OF ECUADOR USING NEURAL NETWORKS AND CLUSTER ANALYSIS WITH MACHINE LEARNING ALGORITHMS

Xavier Rodriguez-Cruz¹ and Julio Litardo²

University de Guayaquil– Faculty of Industrial Engineering
Guayaquil, Ecuador

Article Info

Received:

10th March 2025

Accepted:

4th June 2025

Keywords:

Energy Intensity
Artificial Neural
Networks
Machine Learning
K-means
DBSCAN

JEL:

C45, C53, Q41, O13

DOI:

<https://doi.org/10.47550/RCE/35.1.8>

Abstract

Using a comparative experimental design, four neural network architectures—feedforward, LSTM, sequential feedforward, and robust LSTM—are evaluated to predict the provincial energy intensity index, based on panel data from 2018 to 2023 and key variables related to energy consumption, economic activity, and demographic dynamics. The sequential feedforward model demonstrated the highest predictive accuracy ($R^2 = 0.724$), outperforming recurrent approaches in capturing temporal patterns, precision, and computational efficiency.

Clustering methods allow for the characterization of provinces' behavior in terms of energy intensity, identifying groups with similar energy consumption patterns relative to their economic output. Orellana, Pastaza, Santa Elena, and Zamora Chinchipe stand out for exhibiting distinct or atypical energy profiles, highlighting the usefulness of algorithms such as DBSCAN in detecting unique regional dynamics.

The results help identify regional differences in energy efficiency, providing an empirical basis for designing and implementing tailored energy policies. These insights are especially useful in varied geographic and socio-economic settings, where encouraging sustainable energy use and reducing disparities in distribution and consumption are key objectives.

¹ORCID: 0000-0002-2931-0702. CRediT: research, formal analysis, methodology, writing – original draft. Email: lester.rodriguezc@ug.edu.ec.

²ORCID: 0009-0008-3931-8151. CRediT: conceptualization, supervision, validation, writing – review and editing. Email: julio.litardou@ug.edu.ec.

Copyright © 2025 Rodríguez-Cruz and Litardo. The authors retain copyright of the article. The article is distributed under the Creative Commons Attribution 4.0 License.

1. INTRODUCCIÓN

La transición hacia economías sostenibles y bajas en carbono requiere un enfoque riguroso en la eficiencia energética como un eje central de las políticas públicas. La previsión precisa y eficiente del consumo de energía es un requisito previo crucial para una planificación y formulación de políticas energéticas sustentables y eficaces dentro de una economía.

En ese contexto, la intensidad energética (IE), definida como la cantidad de energía consumida (CE) por unidad de producto interno bruto (PIB) en un determinado periodo, es un indicador crítico que no solo mide la eficiencia en el uso de recursos energéticos, sino que también refleja las interacciones entre el crecimiento económico, el progreso tecnológico y las estrategias de sostenibilidad en una economía.

Una menor intensidad energética indica un uso más eficiente de la energía para generar actividad económica; en otras palabras, se necesita menos energía para generar la misma cantidad de PIB, reflejando una mayor eficiencia energética. En cambio, si la intensidad energética es alta, significa que se requiere más energía para generar cada unidad de PIB, lo que indica una menor eficiencia energética.

De acuerdo con la International Energy Agency (2023), reducir la intensidad energética se ha establecido como una prioridad global para optimizar el uso de recursos y minimizar emisiones de gases de efecto invernadero. Por esta razón, los estudios subnacionales tienen relevancia debido a que permiten identificar dinámicas locales que podrían ser ignoradas en análisis a nivel nacional.

Se debe de considerar que la disminución de emisiones de gases de efecto invernadero se alinea con compromisos internacionales de mitigación del cambio climático, especificados en los Objetivos de Desarrollo Sostenible (ODS), de manera particular con el ODS13, y que una mayor eficiencia energética puede liberar recursos que podrían ser destinados a otros sectores como el de educación, salud e infraestructura.

Otro aspecto relevante del indicador se relaciona con el impacto sobre los recursos naturales. Un menor consumo energético por unidad de PIB puede implicar una menor presión sobre recursos como el petróleo, el gas, el carbón o la biomasa, promoviendo un desarrollo más sostenible. En este sentido, constituye una señal de avance hacia una economía baja en carbono, reflejada en la evolución de sectores económicos de mayor valor agregado y menor intensidad energética, como los servicios y las industrias basadas en el conocimiento.

Ecuador enfrenta desafíos significativos debido al crecimiento sostenido de la demanda energética, su histórica dependencia de fuentes no renovables y una alta participación de la generación hídrica en la capacidad instalada. Esta combinación ha dejado al país vulnerable a las variaciones climáticas y a la disponibilidad de agua, como se evidenció en la reciente crisis eléctrica de 2024, agravada por los efectos del cambio climático.

Las dinámicas anteriormente descritas, en conjunto con las disparidades en el consumo regional de energía, las limitaciones tecnológicas y de infraestructura, plantean desafíos importantes para la planificación y gestión eficiente de los recursos energéticos.

De acuerdo con Rodríguez-Cruz & Balda-Cruz (2022), la economía ecuatoriana es energéticamente intensiva en el desarrollo de su actividad productiva. Bajo esta perspectiva, comprender la distribución territorial y su evolución a nivel provincial es clave para el diseño de políticas públicas que promuevan una transición hacia un uso más eficiente y sostenible de la energía en el Ecuador.

La relación entre intensidad energética, crisis energética y cambio climático es intrínseca. La crisis energética impone desafíos adicionales a la eficiencia del uso de la energía, mientras que el cambio climático altera las condiciones bajo las cuales se produce y consume la energía. Este vínculo resalta la necesidad de un enfoque integral que considere no solo la predicción de la intensidad energética, sino también la adaptación a los efectos del cambio climático y la crisis energética en Ecuador.

Esta investigación propone un enfoque mediante la integración de redes neuronales para la predicción del índice de intensidad energética y el análisis de conglomerados para clasificar las provincias ecuatorianas en función de sus patrones de intensidad energética.

Implementando algoritmos de redes neuronales y de aprendizaje automático, se busca proporcionar un marco metodológico que permita a los formuladores de políticas identificar regiones de alta intensidad e implementar estrategias que fomenten la eficiencia energética, reduzcan la dependencia de fuentes no renovables y contribuyan a los compromisos nacionales e internacionales en materia de sostenibilidad y mitigación del cambio climático.

2. REVISIÓN DE LA LITERATURA

El sector energético es fundamental para el desarrollo sostenible de cualquier país. Según la OIEA (2008), la energía es esencial para potenciar el desarrollo sostenible, la competitividad y la mejora del bienestar social, y, en la mayoría de los casos, es fundamental para generar riqueza industrial y comercial. Por esta razón, la planificación y gestión del consumo de energía es un factor clave para la toma de decisiones en políticas energéticas y de sostenibilidad.

A continuación, se presenta una revisión de la literatura sobre los principales avances en la aplicación de redes neuronales y *clustering* en la predicción de consumo de energía.

Benidis et al. (2022) proporciona una revisión integral sobre el uso del aprendizaje profundo en la predicción de series de tiempo, destacando la capacidad para superar modelos tradicionales en escenarios complejos. Este trabajo

es relevante al presentar un material práctico sobre arquitecturas claves como *long short-term memory* (LSTM), redes neuronales convolucionales (CNN) y transformadores, facilitando su aplicación a diversas industrias.

En consecuencia, podemos afirmar que las redes neuronales y las técnicas de aprendizaje automático representan enfoques sofisticados para la modelación y predicción de fenómenos complejos, como son el consumo de energía y la intensidad energética.

Owda et al. (2014) exploran el uso de redes neuronales artificiales (ANN) para predecir el consumo de energía eléctrica. El estudio presenta un modelo basado en ANN que incorpora múltiples variables, incluyendo factores económicos, climáticos y patrones históricos de consumo. Se destaca la importancia de una adecuada selección de variables y de la optimización de los parámetros de la red para mejorar la precisión de los resultados. El estudio concluye que el uso de ANN puede mejorar significativamente la planificación y gestión energética, y señala que su efectividad podría incrementarse aún más mediante la incorporación de modelos más avanzados y datos en tiempo real.

Nugaliyadde et al. (2019) emplean redes neuronales recurrentes profundas (*deep recurrent neural networks*, RNNs) para predecir el consumo de electricidad. Destacan su capacidad capturar patrones temporales complejos en series de tiempo energéticas y mejorar la precisión de las predicciones en comparación con modelos tradicionales. Además, resaltan que el uso de arquitecturas avanzadas como LSTM permite mejorar la eficiencia de la modelización del consumo energético en distintos escenarios.

En esa misma línea, el artículo de Bai et al. (2024) presenta un algoritmo mejorado de redes neuronales para la predicción del consumo energético, presentan un modelo híbrido avanzado que combina CNN, LSTM y bidireccional (Bi-LSTM), destacando la relevancia en la optimización del uso de energía y la sostenibilidad. Los autores comparan el rendimiento de su modelo con enfoques tradicionales, como son las regresiones estadísticas y redes neuronales estándar, evidenciando mejoras significativas en precisión y eficiencia computacional.

Uno de los principales aportes de este trabajo es la integración de mecanismos avanzados de aprendizaje automático, como la optimización de hiperparámetros y técnicas de regularización, que permiten reducir el sobreajuste y mejora la capacidad de generalización del modelo. Otro aspecto relevante es que los autores enfatizan la importancia de incorporar factores externos —como variables económicas y climáticas— para lograr predicciones más precisas.

Por otro lado, Cui (2016) señala que el crecimiento económico suele ir acompañado de un elevado consumo de energía fósil, lo que conlleva mayores emisiones de gases de efecto invernadero y contribuye al calentamiento global. El estudio destaca que, pese a los esfuerzos en eficiencia energética, la economía no ha logrado una mejora sustancial en su intensidad energética,

lo que ha limitado los avances en la reducción de emisiones y en el cumplimiento de objetivos de sostenibilidad.

Como se indicó anteriormente, el ahorro energético y la protección del medio ambiente se han convertido en desafíos de alcance global. En este contexto, a nivel local, se han referenciado los siguientes trabajos por su relación con la presente investigación.

El estudio realizado por Vallejo (2017) analiza la intensidad energética en los sectores de uso final de la economía ecuatoriana, aplicando el método de descomposición basado en índices Divisia de media logarítmica (LMDI) para evaluar la evolución del consumo energético. Aunque el estudio no emplea redes neuronales ni técnicas de *clustering*, se cita en esta investigación debido a sus resultados relevantes. En particular, el artículo concluye que la eficiencia energética es un factor clave en la reducción del consumo de energía, y demuestra que las mejoras tecnológicas, junto con políticas implementadas en determinados periodos, han contribuido significativamente a disminuir la intensidad energética en el país. Por lo tanto, la capacidad de pronosticar con precisión la intensidad energética e identificar patrones complejos en los datos se vuelve fundamental para una planificación e implementación efectiva de políticas de eficiencia energética en Ecuador.

Estudios recientes han demostrado que las redes neuronales artificiales son eficaces para predecir el consumo mensual de electricidad en Ecuador, utilizando datos específicos de cantones, lo que permite mejorar la planificación energética, como lo especifica Guachimboza-Davalos et al. (2021). Este estudio emplea variables predictoras como el año, el mes, el cantón, el número de clientes y el monto facturado por consumo eléctrico, indicadores que pueden reflejar características socioeconómicas relevantes.

Guamán et al. (2024) presentan un modelo basado en aprendizaje automático para proyectar el consumo eléctrico a largo plazo en Ecuador. El modelo, entrenado con datos históricos y variables socioeconómicas, logra una capacidad predictiva superior a la de los métodos tradicionales. Para ello, se emplean redes neuronales multicapa (MLP), máquinas de soporte vectorial (SVM) y regresión lineal robusta (RLR) con el objetivo de modelar y predecir la demanda eléctrica en Ecuador desde 2023 hasta 2050. El estudio utiliza como variables explicativas indicadores socioeconómicos históricos, tales como población, PIB, acceso a la electricidad y porcentaje de población urbana.

Por otra parte, Hinojosa Bassantes (2022) aborda la implementación de un modelo de predicción de la demanda eléctrica mediante redes neuronales artificiales (RNA). El estudio analiza datos históricos de consumo energético y variables asociadas con el objetivo de anticipar el comportamiento de la demanda futura. El modelo propuesto mejora la precisión de las predicciones en comparación con los métodos convencionales. Sin embargo, se limita al uso de registros de demanda diaria horaria de energía eléctrica correspondientes al barrio San José de Cutuglahua, en el cantón Quito.

En el marco de la planificación energética, la aplicación de técnicas avanzadas de análisis de datos, como las redes neuronales artificiales y los métodos

de *clustering*, ha demostrado ser fundamental para mejorar la precisión de las predicciones. Estas herramientas permiten modelar relaciones complejas entre variables económicas y energéticas, capturando patrones no lineales y dinámicos que los enfoques tradicionales, como los modelos econométricos y de series de tiempo, no logran representar adecuadamente debido a sus supuestos de estabilidad y linealidad en las relaciones económicas. Esta limitación explica, en gran medida, su menor eficacia en proyecciones de largo plazo.

Esta característica permite que las redes neuronales sean herramientas altamente adaptables para la predicción en entornos con datos desordenados, incompletos o afectados por ruido, donde los métodos tradicionales pueden presentar limitaciones como refiere Zaneck (2023). Es decir, su capacidad de aprendizaje automático permite identificar patrones y realizar inferencias sin importar supuestos previos sobre la forma funcional o la relación exacta entre las variables del sistema, siendo particularmente útiles en contextos reales con alta incertidumbre (Haykin, 2009, pp. 18-21, 156-158)

En la revisión de la literatura no se identificaron estudios específicos orientados a la predicción del índice de intensidad energética, sino investigaciones centradas en la estimación de la demanda de energía. En este sentido, el presente trabajo contribuye a la literatura existente al enfocarse en la predicción del índice de intensidad energética mediante el uso de una arquitectura específica de red neuronal y técnicas de análisis de *clustering*.

Es importante destacar que, en economías en desarrollo como la de Ecuador, donde la demanda energética está estrechamente vinculada al desarrollo económico y a factores estructurales, la implementación de redes neuronales para la predicción resulta particularmente relevante. A su vez, los algoritmos de agrupamiento permiten segmentar regiones y poblaciones según patrones de intensidad energética, lo que facilita el diseño de políticas públicas más focalizadas y efectivas.

3. MATERIALES Y MÉTODOS

Para la elaboración del presente trabajo de investigación, se utilizaron los registros del consumo de energía (GWh) y de los clientes regulados a nivel provincial de la Estadística Anual y Multianual del Sector Eléctrico Ecuatoriano¹ de la Agencia de Regulación y Control de Electricidad (ARCONEL), y los registros de las Cuentas Nacionales Anuales Regionales² del Banco Central del Ecuador (BCE) para el valor agregado bruto (VAB) a nivel provincial a valores corrientes.

¹ <https://controlelectrico.gob.ec/publicaciones-estadistica-del-sector-electrico-2/>

² https://contenido.bce.fin.ec/documentos/informacioneconomica/cuentasnacionales/ix_cuentasnacionalesanuales.html#

La población por cada provincia se obtuvo del portal web del Instituto Ecuatoriano de Estadísticas y Censos (INEC)³, y el número de empresas por provincias del sector manufacturero en el sitio web de la Superintendencia de Compañías del Ecuador⁴.

A continuación, en la tabla 1 se presentan las variables a utilizar:

Tabla 1 Variables económicas, energéticas y sociodemográficas

Variables	Descripción
Económicas	
Valor agregado bruto VAB	Valor de la producción menos el valor del consumo intermedio. Es una medida de la contribución al PIB hecha por unidad de producción, industria o sector, expresada en millones de USD.
Empresas sectores económicos	Registra el número de empresas de cada sector bajo el estado de compañías activas para la Superintendencia de Compañías.
Energéticas	
Consumo de energía	Se refiere a la energía facturada a consumidores de las empresas distribuidoras que se encuentran sujetos al pliego tarifario. La unidad de medida para el presente trabajo es el GWh.
Clientes residenciales	Persona natural o jurídica, pública o privada, que utiliza la energía eléctrica exclusivamente para uso doméstico. Se incluyen los consumidores de escasos recursos económicos y de bajo consumo que tienen integrada a su residencia una pequeña actividad comercial o artesanal.
Clientes comerciales	Persona natural o jurídica, pública o privada, que utiliza la energía eléctrica para fines de negocio, actividades profesionales o cualquier actividad con fines de lucro.
Clientes industriales	Persona natural o jurídica, pública o privada, que utiliza la energía eléctrica para la elaboración o transformación de productos. Se consideran a los agroindustriales, que transforman productos de la agricultura, ganadería, riqueza forestal y pesca.
Clientes otros	Consumidores como entidades oficiales, asistencia social, servicios comunitarios, otros.
Sociales	
Población	Número de personas por cada una de las provincias de acuerdo con el INEC

Fuente: Registros estadísticos ARCONEL, BCE, Superintendencia de Compañías e INEC.

Elaboración: autores

³ <https://www.ecuadorencifras.gob.ec/proyecciones-poblacionales/>

⁴ <https://appscvsmovil.supercias.gob.ec/ranking/reporte.html>

Debido a la disponibilidad de los datos, el periodo de cada una de las series abarca desde 2018 hasta 2023. En el apartado de anexos se incluyen los siguientes: energía facturada por provincia 2018-2023 (GWh) (anexo 1); valor agregado bruto provincial 2018-2023 (miles de USD) (anexo 2); unidades de medida eléctrica (anexo 3), y variación anual de la demanda de energía (GWh) y del Valor Agregado Bruto (USD) por provincias (anexo 4).

Los registros incluyen los grupos de consumo establecidos en el pliego tarifario, a excepción del Servicio de Alumbrado Público General (SAPG). Así, para el grupo de consumo *residencial* se consideran las tarifas: residencial, residencial para el programa PEC y residencial temporal. En el grupo de consumo *comercial* se incluyen las tarifas: comercial sin demanda, comercial con demanda y comercial con demanda horaria. Para el grupo de consumo *industrial*, se contemplan las tarifas: industrial con demanda, industrial con demanda horaria, industrial con demanda horaria diferenciada e industrial artesanal. Finalmente, el grupo de consumo *otros* abarca a consumidores como entidades oficiales, organizaciones de asistencia social, servicios comunitarios, sistemas de bombeo de agua, escenarios deportivos y estaciones de carga rápida.

El consumidor regulado es la persona natural o jurídica que, mediante la suscripción de un contrato de suministro, se beneficia con la prestación del servicio público de energía eléctrica y del servicio de alumbrado público general, bien como propietario del inmueble o como receptor directo del servicio, y paga un valor como tarifa, por la energía que consume y la demanda de potencia eléctrica que requiere para satisfacer sus diferentes necesidades, según sus modalidades de consumo y nivel de voltaje.

En la tabla 2 se presentan las principales estadísticas descriptivas de cada una de las variables. Se ha incrementado la dimensionalidad del conjunto de datos en comparación con una serie temporal agregada, al incorporar variables adicionales como la población, número de clientes residenciales, comerciales, industriales y de otros sectores, lo que proporciona una mayor capacidad para capturar patrones heterogéneos entre provincias.

No obstante, se reconoce que el tamaño del conjunto de datos puede ser relativamente reducido para la implementación de modelos complejos como las redes neuronales, lo cual podría aumentar el riesgo de sobreajuste (*overfitting*). Por esta razón, en el presente estudio se han aplicado medidas metodológicas adecuadas para mitigar dicho riesgo en el diseño y entrenamiento de los modelos de redes neuronales.

Tabla 2. Principales estadísticos de las variables de estudio

Variables	Número	Media	Desviación estándar	Mínimo	25 %	50 %	75 %	Máximo
Valor agregado bruto provincial (VABp)	144	4.288.632	7.859.265	222.136	778.326	1.667.876	3.372.844	35.912.077
Empresas manufactureras (NE)	144	4.377	9.467	165	520	866	1.854	45.815
Consumo de energía (DE)	144	826	1.537	48	95	409	660	8.236
Clientes residenciales (Clr)	144	199.944	273.855	90.729	47.829	128.393	202.934	1.127.179
Clientes comerciales (Clc)	144	20.791	32.721	20.015	4.881	10.330	19.702	150.529
Clientes industriales (Clin)	144	1.696	2.860	119	261	520	1.612	13.973
Clientes Otros (Clo)	144	3.397	3.807	512	1.440	1.810	3.396	18.846
Población (Pob)	144	730.040	1.034.088	28.205	188.593	477.250	628.492	4.690.613

Fuente: Banco Central del Ecuador, ARCONEL y Superintendencia de Compañías

Elaboración: autores

Según García et al (2011) la intensidad energética (IE) es el cociente entre el consumo de energía (CE) y el producto interno bruto (PIB), en un determinado periodo. Se calcula a valores corrientes o valores constantes con un año base determinado.

Como el cálculo es a nivel de provincia se utiliza el valor agregado bruto por provincia para aproximar el valor total de la producción de la economía, considerando que el rubro de *otros elementos* del PIB⁵ no se distribuye por provincia. La fórmula de cálculo es la siguiente:

$$IE_{ij} = \frac{CE_{ij}}{VAB_{ij}} \quad (1)$$

Donde:

- IE_{ij} es la intensidad energética en el año i de la provincia j (bep/10³ USD).
- CE_{ij} es el consumo energético del año i de la provincia j expresado en unidades calóricas (10³ bep).
- VAB_{ij} corresponde al valor total agregado de la producción en el año i de la provincia j (10⁶ USD).

Como se indicó anteriormente, la predicción del índice de intensidad energética por provincias se realizará a través de redes neuronales. Se debe destacar que, en el desarrollo de esta investigación, se realizó un proceso iterativo de exploración y evaluación de arquitecturas de redes para modelar el fenómeno de interés con la disponibilidad de datos especificados.

A partir de las pruebas experimentales y comparativas, se seleccionaron cuatro modelos representativos que mostraron un balance adecuado entre capacidad predictiva, estabilidad durante el entrenamiento y complejidad computacional. Esta práctica es coherente con los lineamientos metodológicos establecidos en la literatura sobre aprendizaje profundo, donde se recomienda realizar comparaciones empíricas entre distintas arquitecturas con el fin de identificar estructuras que se adapten a la naturaleza de los datos y a los objetivos de análisis. Como propone Chollet (2021, pp. 295-328), las redes LSTM permiten capturar dependencias temporales en series, siendo útiles para modelado secuencial, mientras que el uso de la normalización Dropout y Batch mejora la generalización de modelos *feedforward* (Chollet, 2021, pp.259-293).

⁵ Otros elementos del PIB: Incluye impuestos indirectos netos + derechos arancelarios + impuestos netos sobre importaciones + impuesto al valor agregado (IVA).

Con base en lo anterior, se mitiga el sesgo asociado a la selección de una única arquitectura y se garantiza una comparación robusta entre modelos. En este estudio se evaluaron cuatro arquitecturas representativas: dos redes neuronales *feedforward* (una básica y otra adaptada a secuencias mediante *flattening*), y dos modelos LSTM (uno simple y otro optimizado mediante normalización por lotes, dropout y validación cruzada temporal). Esta selección permite diferenciar entre arquitecturas estáticas y secuenciales, así como entre configuraciones simples y avanzadas.

Para garantizar la comparabilidad objetiva, cada modelo se entrenó bajo condiciones experimentales homogéneas, incluyendo el uso del mismo conjunto preprocesado, número de épocas, tasas de aprendizaje, función de pérdida y optimizador.

3.1. Red neuronal *feedforward*

Se trata de una red neuronal artificial (ANN) clásica con flujo unidireccional de datos desde la capa de entrada, pasando por capas ocultas, hasta la salida, sin ciclos ni retroalimentación. La arquitectura transforma la entrada, compuesta por seis dimensiones, a un espacio de 64 dimensiones. Luego aplica una función de activación ReLU para introducir no linealidad, seguida de una capa de *dropout* que apaga aleatoriamente el 30 % de las neuronas, con el fin de reducir el riesgo de sobreajuste (*overfitting*). Posteriormente, la red reduce progresivamente la dimensionalidad a 32, luego a 16, hasta llegar a una única salida. Su principal ventaja radica en su simplicidad y velocidad de entrenamiento; sin embargo, no está diseñada para procesar datos secuenciales. La tabla 1 presenta en detalle su arquitectura.

Tabla 3. Feedforward clásico

Capa	Notación matemática	Dimensión de pesos	Activación
Entrada	$x \in \mathbb{R}^6$	-	-
Linear(6,64)	$h_1 = W_1 x + b_1$	$W_1 \in \mathbb{R}^{6 \times 6}$	ReLU
Dropout(0,3)	$h'_1 = \text{Dropout}(h_1)$		-
Linear(64,32)	$h_2 = W_2 h'_1 + b_2$	$W_2 \in \mathbb{R}^{32 \times 64}$	ReLU
BatchNorm1d(32)	$h'_2 = \text{BatchNorm}(h_2)$	-	-
Linear(32,16)	$h_3 = W_3 h'_2 + b_2$	$W_3 \in \mathbb{R}^{16 \times 32}$	ReLU
Linear(16,1)	$\hat{y} = W_4 h_3 + b_4$	$W_4 \in \mathbb{R}^{1 \times 16}$	-

Elaboración: autores

El modelo matemático para la red feedforward clásico es:

$$\hat{y} = W_4 \cdot \text{ReLU}(W_3 \cdot \text{BatchNorm}(\text{ReLU}(W_2 \cdot \text{Dropout}(\text{ReLU}(W_1 x + b_1)))) + b_4) \quad (2)$$

3.2. Red LSTM básica

Es una red recurrente (RNN) especializada en capturar dependencias temporales mediante “puertas” que regulan el flujo de información. En este caso en particular, procesa secuencia de 3 pasos (lookback = 3) con 6 características cada una. Cada capa LSTM calcula:

- Puertas *forget/input/output*: deciden la información a recordar/descartar.
- Estado de memoria (C_t): almacena información relevante a largo plazo.

Luego, toma el último estado oculto h_t y lo mapea a la salida. Una ventaja de esta arquitectura de red es que modela patrones temporales complejos, pero el coste computacional es muy alto con respecto a la red *feedforward*. A continuación, se presenta la arquitectura en la tabla 4.

Tabla 4. Red LSTM (Long Short-Term Memory)

Capa	Notación matemática	Dimensión de pesos	Activación
Entrada	$X \in \mathbb{R}^{3 \times 6}$ (secuencia)	-	-
LSTM(6,64,2)	$(h_t, c_t) = \text{LSTM}(x_t, h_{t-1}, c_{t-1})$	$W_{\text{LSTM}} \in \mathbb{R}^{4 \times 64 \times (6+64)}$	Tanh/ σ
Dropout(0,3)	$h'_T = \text{Dropout}(h_T)$		
Linear(64,1)	$\hat{y} = W_h h'_T + b_h$	$W_h \in \mathbb{R}^{1 \times 64}$	

Elaboración: autores

Ecuación LSTM para cada paso t :

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (3)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (4)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (5)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (6)$$

$$o_t = \sigma(W_0 \cdot [h_{t-1}, x_t] + b_0) \quad (7)$$

$$h_t = o_t \odot \tanh(C_t) \quad (8)$$

Donde:

f_t es la puerta de olvido;
 i_t es la puerta de entrada;
 $\tilde{C}_t$ es la memoria candidata;
 C_t es la memoria actualizada;
 o_t es la puerta de salida, y
 h_t es la salida de la celda en el tiempo t .

La salida final de la red o el último estado oculto se especifica de la siguiente manera:

$$\hat{y} = W_h \cdot h_t + b_h \quad (9)$$

3.3. Red neuronal *feedforward* secuencial (FFNN secuencial)

El tercer modelo corresponde a una red neuronal *feedforward* con estructura secuencial. Conserva la arquitectura del modelo *feedforward* básico, pero adapta su entrada mediante una expansión que convierte las secuencias en vectores planos. En este estudio, se utilizaron secuencias de tres pasos con seis variables por paso, lo que genera un total de 18 entradas aplanadas. Su principal ventaja es que combina la simplicidad de una FFNN con la capacidad de incorporar información temporal en la entrada. La tabla 3 presenta en detalle la arquitectura de esta red.

Tabla 5. Feedforward Secuencial

Capa	Notación matemática	Dimensión de pesos	Activación
Entrada	$X_{\text{flat}} \in \mathbb{R}^{18}$ (3 pasos x 6 features)	-	-
Linear(18,64)	$h_1 = W_1 X_{\text{flat}} + b_1$	$W_1 \in \mathbb{R}^{64 \times 18}$	ReLU
Dropout(0,3)	$h'_1 = \text{Dropout}(h_1)$	-	-
Linear(64,32)	$h_2 = W_2 h'_1 + b_2$	$W_2 \in \mathbb{R}^{32 \times 64}$	ReLU
BatchNorm1d(32)	$h'_2 = \text{BatchNorm}(h_2)$	-	-
Linear(32,16)	$h_3 = W_3 h'_2 + b_3$	$W_3 \in \mathbb{R}^{16 \times 32}$	ReLU
Linear(16,1)	$\hat{y} = W_4 h_3 + b_4$	$W_4 \in \mathbb{R}^{1 \times 16}$	-

Elaboración: autores

A continuación se presenta la notación matemática del modelo FFNN Secuencial

$$y\hat{y} = W_4 \cdot \text{ReLU}(W_3 \cdot \text{BatchNorm}(\text{ReLU}(W_2 \cdot \text{Dropout}(\text{ReLU}(W_1 X_{\text{flat}} + b_1)))) + b_4) \quad (10)$$

3.4. Modelo LSTM robusto

Por último, se implementó una versión mejorada del modelo LSTM, incorporando funciones de activación ReLU y normalización mediante BatchNorm, con el objetivo principal de lograr un modelado avanzado de secuencias y una mayor capacidad de generalización. Este modelo incluye una red LSTM con *dropout* interno del 30 %, normalización por lotes y capas densas adicionales de 64, 32 y 1 neuronas, respectivamente. Para la normalización de las predicciones, se utiliza una función de activación *sigmoid* en la salida. Estas características fortalecen su capacidad predictiva frente a secuencias complejas y permiten integrar mecanismos de prevención del sobreajuste como *dropout* y *early stopping*, así como una regularización más robusta mediante la combinación de *dropout* y BatchNorm.

A continuación, se presenta la arquitectura de la red LSTM robusta:

Tabla 6. Red LSTM robusto

Capa	Notación matemática	Dimensión de pesos	Activación
Entrada	$X \in \mathbb{R}^{3 \times 6}$	-	-
LSTM(6,64,2) + Dropout(0,3)	$h_T = \text{LSTM}(X)[-1]$	$W_{\text{LSTM}} \in \mathbb{R}^{4 \times 64 \times (6+64)}$	Tanh/ σ
BatchNorm1d(64)	$h'_T = \text{BatchNorm}(h_T)$	-	-
Linear(64,32)	$h_1 = W_1 h'_T + b_1$	$W_1 \in \mathbb{R}^{32 \times 64}$	ReLU
Linear(32,1)	$\hat{y} = \sigma(W_2 h_1 + b_2)$	$W_1 \in \mathbb{R}^{1 \times 32}$	Sigmoid

Elaboración: autores

La notación matemática para la arquitectura de red LSTM robusto se describe a continuación:

$$\hat{y} = \sigma(W_2 \cdot \text{ReLU}(W_1 \cdot \text{BatchNorm}(\text{LSTM}(X)[-1] + b_1) + b_2) \quad (11)$$

Se proponen las siguientes arquitecturas distintas con mecanismos explícitos de deregularización: *dropout* (30 %), normalización *batch* y la inclusión de una estrategia de validación con particiones diferenciadas según el modelo. De forma particular, el modelo LSTM robusto incorpora *early stopping* con validación anidada, lo que permite una mejor detección de sobreajuste durante el entrenamiento.

En ese mismo orden de ideas, se ha utilizado el optimizador Adam o AdamW con tasa de aprendizaje de 0,001, y se ha mantenido una semilla fija (SEED = 42) para garantizar la reproducibilidad. De igual forma, la división del conjunto de datos (80/20) respeta el orden temporal, evitando el uso de aleatorización que podría inducir pérdida de información.

Como se indicó al comienzo de este apartado, estas prácticas metodológicas están alineadas con las recomendaciones de la literatura para el entrenamiento confiable de redes neuronales con conjunto de datos pequeños como lo destacan Goodfellow et al. (2016, pp. 111-135).

En esta investigación se seguirá un enfoque de análisis de conglomerados similar al propuesto por Rodríguez-Cruz (2024), con la diferenciación de que se busca la utilidad del análisis de agrupamiento para identificar patrones de intensidad energética diferenciados entre provincias del Ecuador.

El algoritmo de K-means se lo puede expresar de la siguiente manera:

$$J = \sum_{i=1}^K \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (12)$$

Donde:

K es el número de cluster;
 C_i es el conjunto de puntos de datos en el i -ésimo clúster;
 x representa un punto de datos;
 μ_i es el centroide del i -ésimo clúster, y
 $\|x - \mu_i\|^2$ es la distancia euclidiana cuadrada entre el punto de datos x y el centroide μ_i .

Cada punto de datos x se asigna al clúster cuyo contenido μ_i minimice la distancia euclidiana:

$$C_i = \{x : \|x - \mu_i\|^2 \leq \|x - \mu_j\|^2 \forall j, 1 \leq j \leq K\} \quad (13)$$

$$\Delta J = \left(\frac{n_i}{n_i - 1} \right) \|x - \mu_i\|^2 - \left(\frac{n_j}{n_j + 1} \right) \|x - \mu_j\|^2 \quad (14)$$

Donde:

n_i y n_j son las cantidades de puntos en los clusteres C_i y C_j respectivamente.

μ_i y μ_j son los centroides de los clusteres C_i y C_j respectivamente.

Para encontrar los clústeres en K-means se utilizará el Elbow Method, más conocido como el método del codo, técnica para determinar el número óptimo de grupos o clústeres.

Para el algoritmo DBSCAN, el modelo se expresa como:

$$N_\varepsilon(P) = \{q \in D \mid \text{dist}(P, q) \leq \varepsilon\} \quad (15)$$

Donde:

$N_\varepsilon(P)$ es el conjunto de puntos en la vecindad ε del punto P .

q es el punto de datos en el conjunto de datos D .

$\text{dist}(P, q)$ es la distancia (típicamente euclidiana) entre los puntos P y q .

ε es el radio de vecindad.

Se define el punto central como:

$$|N_\varepsilon(P)| \geq \text{MinPts} \quad (16)$$

Donde:

$|N_\varepsilon(P)|$ es el número de puntos en la vecindad ε del punto P .

MinPts es el número mínimo de puntos requeridos para formar un clúster.

Finalmente, se utiliza el algoritmo de aglomeramiento jerárquico (*hierarchical clustering*). Este algoritmo no exige especificar previamente un número de clústeres, ya que genera un dendrograma que representa los niveles de similitud entre los puntos de datos. En este caso, se utiliza el enfoque aglomerativo junto con una medida de distancia basada en el enlace promedio, que se detalla a continuación:

$$C = \{\{x_1\}, \{x_2\}, \dots, \{x_n\}\} \quad (17)$$

Donde C se considera el conjunto de clústers iniciales y x_i es un punto de datos. La distancia es:

$$d(C_i, C_j) = \frac{1}{|C_i| \cdot |C_j|} \sum_{x \in C_i} \sum_{y \in C_j} dist(x, y) \quad (18)$$

La metodología de K-means se emplea para identificar provincias con patrones similares de intensidad energética, mientras que DBSCAN se utiliza para detectar provincias con comportamientos atípicos. Por su parte, el clúster jerárquico aglomerativo resulta particularmente útil para análisis espaciales y temporales, como lo señalan Hastie et al. (2005).

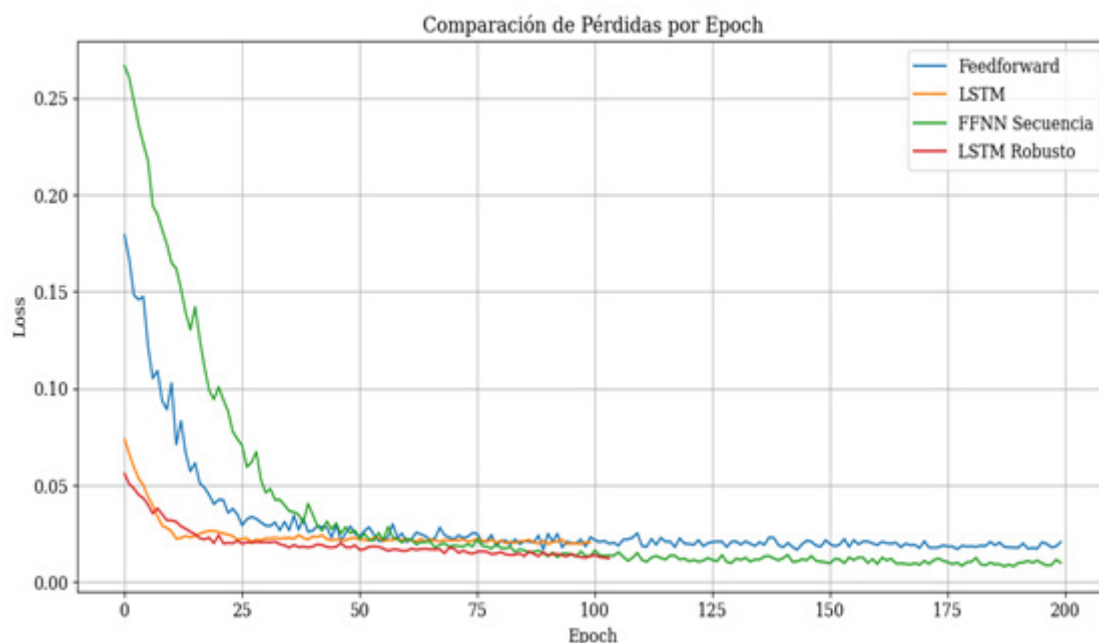
Cabe destacar que, en el presente trabajo, se utilizan las series a precios corrientes del valor agregado bruto provincial, y se realiza la conversión de medidas calóricas para transformar la demanda de energía de GWh a kBEP.

4. RESULTADOS Y LIMITACIONES

En esta sección se presentan y analizan los resultados obtenidos mediante la aplicación de las diferentes arquitecturas de red propuestas para la predicción del índice de intensidad energética por provincias. En el anexo 5, se presenta una matriz comparativa de las características principales de las arquitecturas de redes neuronales.

La figura 1 muestra la evolución de la función de pérdida (MSE) durante las épocas de entrenamiento y validación para cada una de las arquitecturas evaluadas.

Figura 1. Evolución de la función de pérdida durante el entrenamiento



Elaboración: autores

Se observa que el modelo *feedforward* clásico presenta oscilaciones significativas y una alta pérdida persistente: señal de inestabilidad en el entrenamiento y una limitada capacidad de capturar patrones generales en los datos de validación. Sugiere una posible tendencia al sobreajuste, con una menor capacidad de extrapolación hacia datos no observados.

En contraste, el modelo *feedforward* secuencial exhibe una curva de pérdida suave y decreciente, logrando una convergencia eficiente y una menor divergencia entre las fases de entrenamiento y validación en comparación con otras arquitecturas. Esto evidencia una mayor capacidad predictiva fuera de la muestra; en otras palabras, refleja una buena estabilidad de aprendizaje y una capacidad robusta para extrapolar a dinámicas provinciales no observadas en los datos de entrenamiento.

Por otro lado, las arquitecturas LSTM muestran una evolución estable, aunque con una tasa de convergencia menor respecto al modelo secuencial, como se muestra en las tasas de convergencia del error de entrenamiento por modelo, indicada en la tabla 7.

Tabla 7. Tasas de convergencia del error de entrenamiento

Epoch	Feedforward	LSTM	Feedforward secuencial	LSTM robusto
1	0,179313	0,073928	0,266544	0,056109
20	0,040270	0,026430	0,094411	0,020113
40	0,026022	0,022036	0,040594	0,017944
60	0,025187	0,021746	0,020481	0,016229
80	0,018849	0,020335	0,019511	0,014783
100	0,023335	0,020487	0,013419	0,013423
120	0,019986	-	0,012235	-
140	0,019084	-	0,011359	-
160	0,020052	-	0,012018	-
180	0,018425	-	0,010234	-
200	0,020791	-	0,009992	-

Elaboración: autores

Los modelos LSTM muestran un aprendizaje consistente, sin evidencias de sobreajuste ni inestabilidad numérica. No obstante, aunque estos modelos logran aprender, lo hacen a un ritmo más lento en comparación con el modelo *feedforward* secuencial. Esto implica que, a lo largo de las épocas de entrenamiento, el LSTM requiere más tiempo para reducir el error hasta niveles óptimos, lo cual representa una desventaja en términos de mayor duración del entrenamiento, mayor costo computacional y un riesgo incrementado de quedar atrapado en mínimos locales.

En la tabla 8 se presentan las métricas de desempeño utilizadas para comparar las diferentes redes neuronales. En este caso, se destaca que el modelo *feedforward* secuencial presenta el mejor ajuste predictivo: comete pocos errores al tener un RMSE de 0,0344 y explica aproximadamente el 72 % de la variabilidad del índice de intensidad energética entre provincias.

Tabla 8. Comparación de desempeño de las arquitecturas de redes neuronales

Modelo de red neuronal	Error cuadrático medio (MSE)	Raíz del error cuadrático medio (RMSE)	Coefficiente de determinación (R2)
Feedforward clásico	0,00434	0,06586	-0,011
Feedforward secuencial	0,00119	0,03444	0,724
LSTM	0,00401	0,06335	0,065
LSTM robusto	0,00244	0,04936	0,432

Elaboración: autores

El modelo LSTM, si bien logra desempeños aceptables, no supera al secuencial. El modelo clásico evidencia bajo ajuste y potencial sobreajuste, confirmado por su R2 negativo.

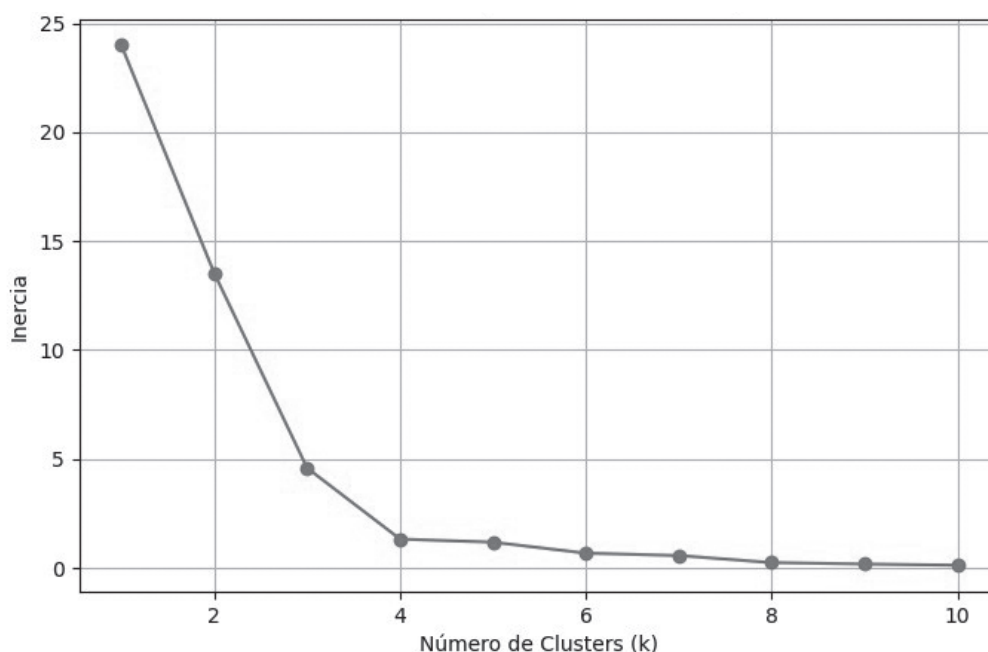
Las métricas obtenidas sobre el conjunto de prueba permiten una evaluación objetiva de la capacidad de generalización de los modelos. Los resultados muestran que la arquitectura de la red *feedforward* secuencial es la más adecuada para la predicción del índice de intensidad energética.

Aunque el tamaño de muestra puede parecer reducido desde una perspectiva de aprendizaje profundo, la evidencia demuestra que la red *feedforward* secuencial ofrece resultados robustos incluso con un conjunto de datos moderado, siempre que se apliquen adecuadamente el preprocesamiento, la selección de arquitectura y las técnicas de regularización para mitigar el sobreajuste, en línea con lo planteado por Chollet (2021).

Por otro lado, la red mitiga problemas de multicolinealidad sin necesidad de emplear técnicas de reducción de dimensiones, debido a la combinación de métodos de regularización como *dropout* o BatchNorm, lo cual coincide con lo señalado por Goodfellow et al. (2016), quienes destacan que las redes profundas pueden aprender representaciones latentes no lineales, evitando redundancias en los datos y mejorando su capacidad de generalización.

En el anexo 6 se presenta el resultado de la predicción del índice de intensidad energética en kBEP/103 USD para el año 2024 y 2025, y en el anexo 7 se presenta la evolución del índice de intensidad energética por provincias para el periodo 2018 a 2025.

Conforme a la aplicación de los algoritmos de agrupamiento, el análisis de conglomerados basado en los resultados de K-means permitió identificar cuatro clústeres diferenciados, cada uno de los cuales refleja variaciones en el promedio del índice de intensidad energética durante el periodo analizado.

Figura 2. Clústeres óptimos para K-means

Fuente: Banco Central del Ecuador y ARCONEL

Elaboración: autores

El análisis de clústeres realizado revela una clara diferenciación en la intensidad energética de las provincias agrupadas en cuatro clústeres. El clúster 1 se concentra en Orellana y Pastaza, provincias amazónicas con baja participación de consumidores regulados y una baja densidad poblacional y dispersión territorial. Aunque son provincias que están ligadas a la industria petrolera, no necesariamente tienen un alto consumo energético en sectores residenciales y comerciales, debido a una menor industrialización y economía más orientada a actividades de bajo consumo energético. En este sentido, este grupo representa un nivel muy bajo de intensidad energética.

El clúster 2, con un promedio de 0,112073, incluye provincias como Azuay y Napo, donde predomina la mayor concentración de centrales de generación renovable. Cabe destacar que, al cierre de 2023, de los 1.596,16 MW de potencia instalados en las centrales ubicadas en la provincia de Napo, la mayor parte corresponde a la central Coca Codo Sinclair. Este grupo, el más numeroso entre los conglomerados identificados, está conformado por provincias andinas y amazónicas cuyo uso energético resulta relativamente eficiente en relación con el valor económico que generan.

El clúster 0 está conformado por provincias como Guayas y Manabí, con un promedio de intensidad energética de 0,150153, lo que sugiere una mayor concentración de actividades productivas o un consumo elevado en sectores específicos. Este grupo incluye economías con presencia de industrias medianas, actividades agroindustriales y manufactureras, así como sectores petroquímicos —como la refinería ubicada en Esmeraldas— y economías

orientadas a la agroexportación, el comercio y los servicios. También forman parte de este clúster provincias con una base agrícola diversificada, como Cotopaxi, Chimborazo e Imbabura, lo que lo convierte en un conglomerado relevante, caracterizado por una intensidad energética media-alta.

En cuanto al clúster 3, este registra el promedio más alto (0,203154), conformado por provincias de Santa Elena y Zamora Chinchipe, que tienen una estructura basada en sectores extractivos con alto consumo energético y bajo procesamiento, lo que da señales de menor eficiencia energética.

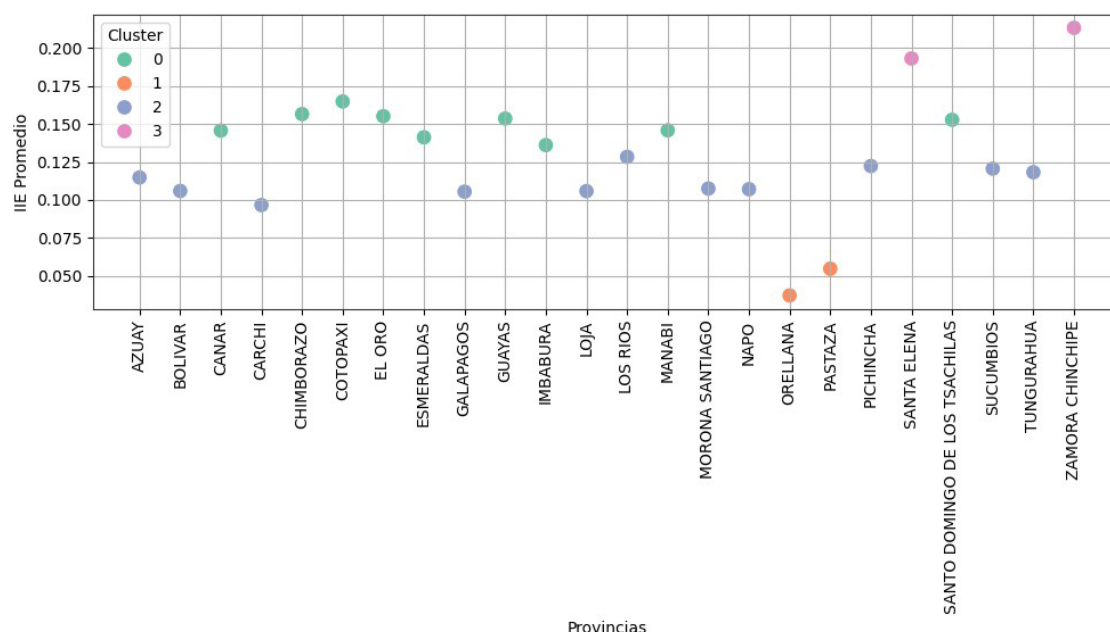
En la tabla 9 se muestran las provincias para cada uno de los clústeres identificados con la aplicación del algoritmo K-means.

Tabla 9. Análisis por clúster K-means-*feedforward* secuencial

Clúster	Provincias	IEE promedio	Comentarios
1	Orellana y Pastaza	0,045961	Provincias con una baja intensidad energética, explicada en medida por ser economías basadas en actividades como la agricultura tradicional, ganadería extensiva, turismo ecológico.
2	Azuay, Bolívar, Carchi, Galápagos, Loja, Los Ríos, Morona Santiago, Napo, Pichincha, Sucumbíos, Tungurahua	0,112073	En Sucumbíos se registra un alto consumo eléctrico debido a que se atiende la carga a EP Petroecuador, lo que eleva el promedio, considerando que su densidad poblacional es relativamente baja. Está conformada por economías más diversificadas o con mayor orientación a servicios y producción con alto valor agregado. Baja intensidad energética.
0	Cañar, Chimborazo, Cotopaxi, El Oro, Esmeraldas, Guayas, Imbabura, Manabí, Santo Domingo de los Tsáchilas	0,150153	Integrado por economías mixtas, con actividad agroindustrial relevante y oportunidades de mejora en eficiencia energética. Clúster de media intensidad energética
3	Santa Elena y Zamora Chinchipe	0,203154	Integrados por provincias con alta intensidad energética, debido a que sus economías dependen de sectores extractivos intensivos en energía (minería, pesca, petroquímica) con bajo valor agregado por unidad de energía consumida. Grupo con alta intensidad energética.

Fuente: Banco Central del Ecuador y ARCONEL

Elaboración: autores

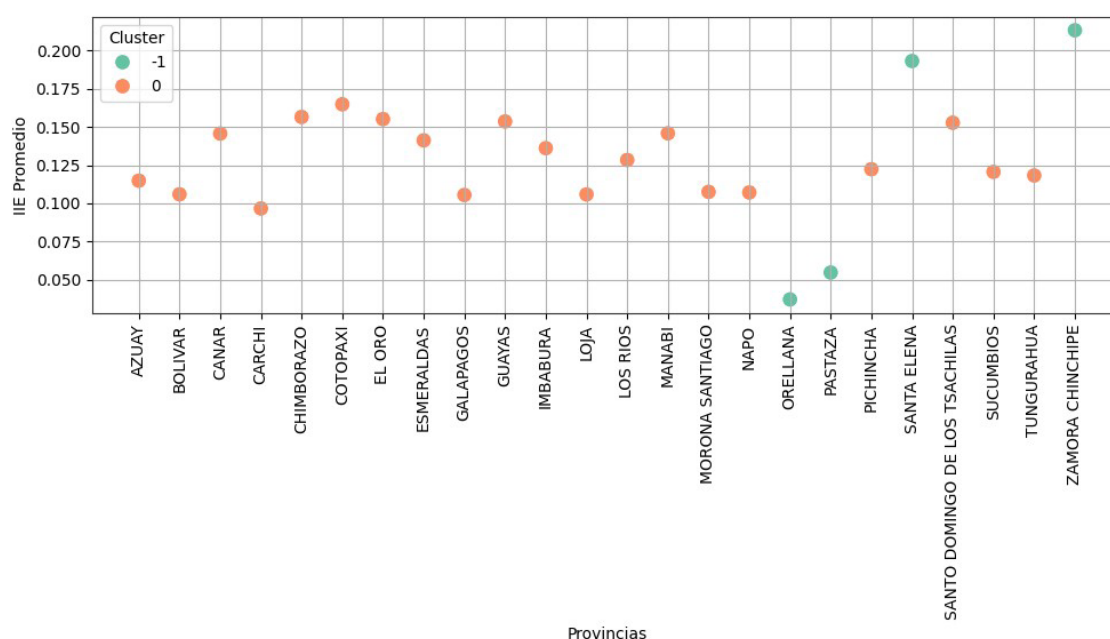
Figura 3. Clústeres K-means

Fuente: Banco Central del Ecuador y ARCONEL

Elaboración: autores

Con respecto a los resultados obtenidos mediante el algoritmo DBSCAN, se identificó un grupo de provincias atípicas conformado por Orellana, Pastaza —economías con baja transformación de bienes y servicios, como se indicó previamente—, así como Santa Elena y Zamora Chinchipe. Este grupo fue clasificado como clúster -1, lo cual el algoritmo interpreta como *ruido*, es decir, provincias cuyos valores del índice de intensidad energética presentan una alta heterogeneidad o no se ajustan a los patrones predominantes en el resto del país. En contraste, el clúster 0 agrupa a todas las demás provincias, lo que sugiere una mayor homogeneidad en la relación entre el PIB y el consumo energético, representando estructuras económicas más estables y típicas del contexto ecuatoriano.

Figura 4. Clúster DBSCAN

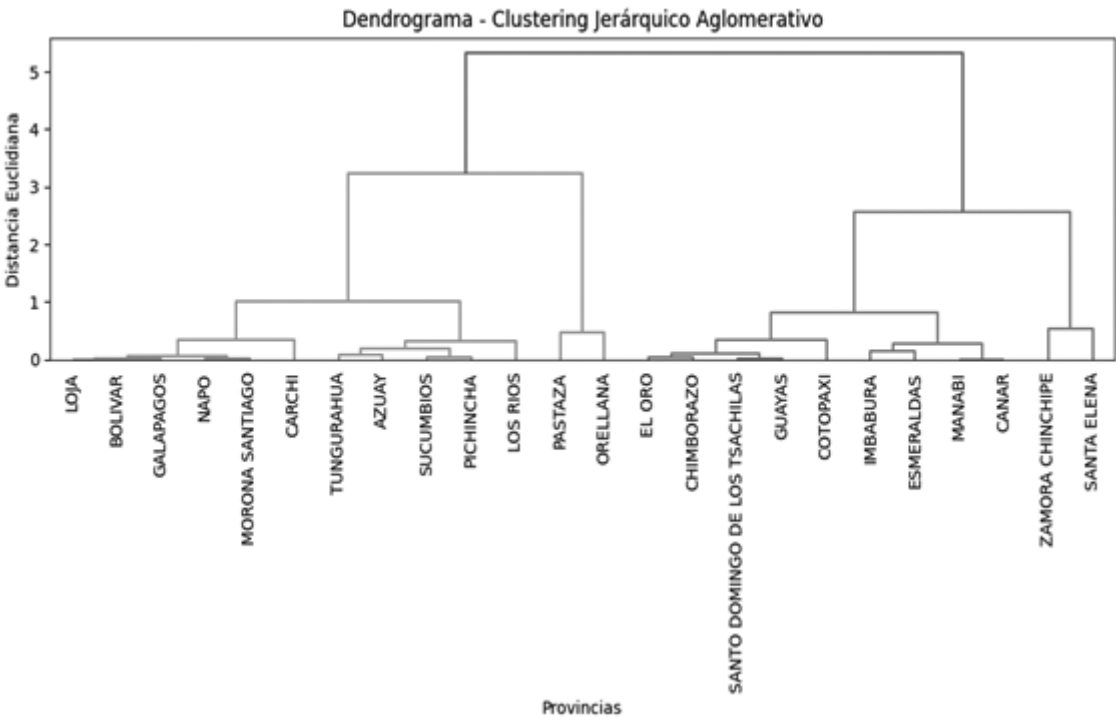


Fuente: Banco Central del Ecuador y ARCONEL

Elaboración: autores

En cuanto a la aplicación del algoritmo de aglomeramiento jerárquico (*hierachical clustering*), el dendograma permite identificar a las provincias que se agrupan en función de su similitud con respecto al índice de intensidad energética. Se observan tres grupos principales: el primero se forma a una distancia menor a 1,5 e incluye provincias como Loja, Bolívar, Galápagos y Napo, entre otras del centro y la Amazonía, reflejando patrones energéticos similares. El segundo grupo se consolida a una distancia aproximada de 2,5 y está conformado por provincias como Guayas, Manabí, Esmeraldas e Imbabura, las cuales comparten un perfil energético posiblemente asociado a un mayor consumo o desarrollo económico. Finalmente, el tercer grupo lo integran Zamora Chinchipe y Santa Elena, que se agrupan recién a una distancia superior a 5, lo que sugiere que presentan características energéticas atípicas o significativamente distintas del resto. Este último hallazgo es coherente con los resultados del algoritmo DBSCAN, que también las identificó como *outliers*.

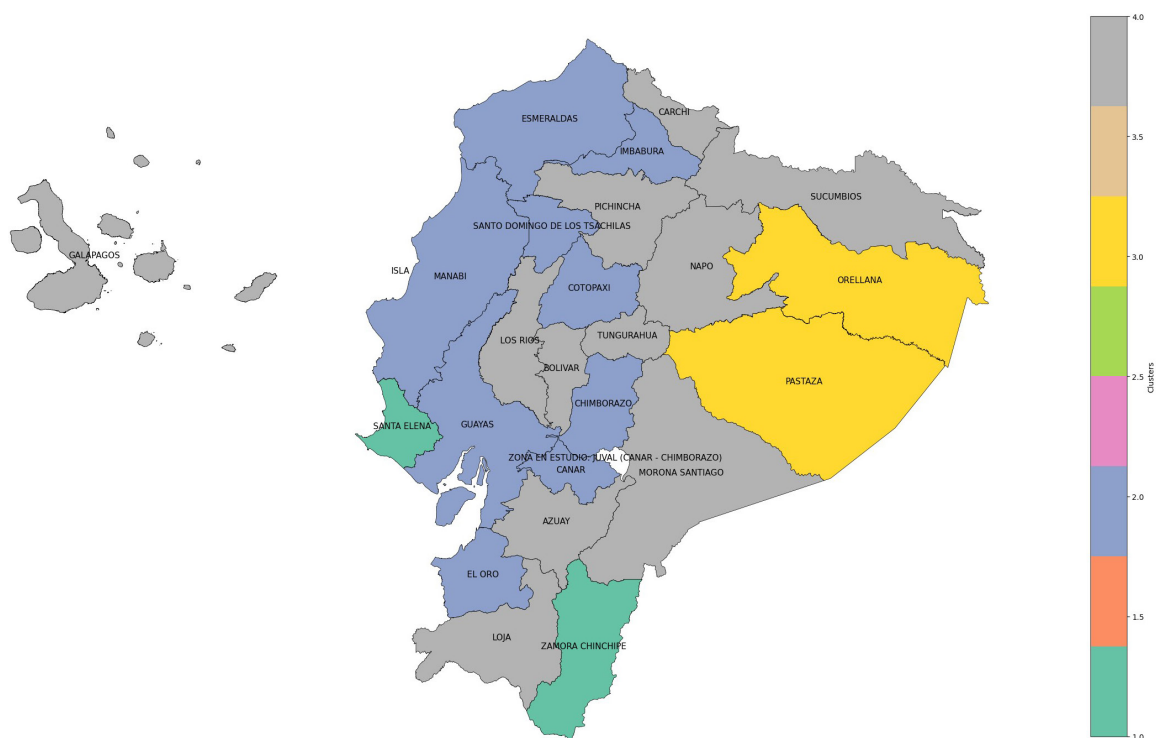
Figura 5. Dendrograma-Clúster jerárquico aglomerativo



Fuente: Banco Central del Ecuador y ARCONEL

Elaboración: autores

Figura 6. Geografía de los clústeres jerárquicos por provincias



Fuente: Banco Central del Ecuador y ARCONEL

Elaboración: autores

5. CONCLUSIONES Y DISCUSIÓN

En esta investigación se evaluaron cuatro arquitecturas de redes neuronales para la predicción del índice de intensidad energética a nivel provincial: *feedforward* tradicional, LSTM básico, *feedforward* secuencial y LSTM robusto. Los resultados demostraron que el modelo *feedforward* secuencial obtuvo mejor rendimiento, con un R^2 de 0,724, superando a las demás arquitecturas en precisión y estabilidad. Esta elección se fundamentó en su capacidad para capturar dependencias temporales mediante el aplanamiento de secuencias, combinado con una estructura eficiente de capas densas y técnicas de regularización (*dropout* y BatchNorm), lo que permitió una mejor generalización sin caer en sobreajuste.

Utilizando el *feedforward* secuencial, se realizaron predicciones del índice de intensidad energética para el año 2024 y 2025, por cada una de las provincias. Estas predicciones se integraron posteriormente para ser analizadas mediante técnicas de clustering como K-means, DBSCAN y jerárquico aglomerativo para identificar patrones geográficos en el comportamiento energético.

La aplicación de los algoritmos para aglomeramiento, para el caso particular de K-means, permite identificar regiones de manera clara y eficiente en cuatro grupos. Se puede evidenciar que provincias como Orellana y Pastaza registran en promedio un índice de intensidad energética de 0,046 kBEP/10³ USD; este valor indica que, para generar 1.000 dólares de valor agregado bruto provincial, se consumen 0,046 kilogramos de barriles equivalentes de petróleo (BEP) o 46 gramos de BEP, lo que supone una muy alta eficiencia relativa, aunque puede estar relacionadas con su limitada estructura productiva.

Provincias como Pichincha y Guayas integran grupos de clúster con baja y media intensidad energética, corresponden a regiones con más altos niveles de industrialización. Por otro lado, las provincias de Santa Elena y Zamora Chinchipe conforman un clúster con un alto índice promedio de intensidad energética, de 0,203 kBEP/10³ USD explicada por economías que dependen de sectores extractivos intensivos en energía, como la minería, la pesca y la petroquímica, que generan un bajo valor agregado por unidad de energía consumida.

Los resultados del análisis de densidad DBSCAN identifican un pequeño grupo de provincias atípicas integradas por Orellana, Pastaza, Santa Elena y Zamora Chinchipe. Estas provincias tienen características energéticas que difieren notablemente del resto, consistente con la agrupación generada por el algoritmo K-means y resaltando el valor de la aplicabilidad del algoritmo DBSCAN para identificar regiones con comportamientos energéticos diferenciados; sin embargo, ambos clústeres tienen valores promedio del índice de intensidad energética muy cercanos, una diferencia de 0,00465 kBEP/10³ USD entre los grupos, lo que sugiere que no hay una separación drástica entre los grupos.

En cuanto a la aplicación del algoritmo de clústeres jerárquicos, este permite visualizar progresivamente cómo se forman los grupos, identificando tres grandes conglomerados a distintos niveles de similitud. Se observa que varias provincias del centro y de la Amazonía ecuatoriana comparten características energéticas similares, mientras que otras, como Santa Elena y Zamora Chinchipe, destacan por presentar un comportamiento energético notablemente divergente.

Cabe señalar que no se incorporó un análisis de sensibilidad, ya que este no fue contemplado en el diseño metodológico original. La presente investigación se enfocó en evaluar el rendimiento predictivo de diversas arquitecturas de redes neuronales a partir de un conjunto específico de variables y parámetros. No obstante, la inclusión de un análisis de sensibilidad podría constituir una extensión relevante para investigaciones futuras, especialmente para valorar la estabilidad de los modelos frente a variaciones en los insumos, en los hiperparámetros o con datos a nivel cantonal.

Finalmente, se recomienda realizar un contraste sistemático entre los resultados obtenidos mediante la arquitectura de red neuronal y los generados por metodologías econométricas para datos de panel, con el objetivo de

comparar el desempeño de ambas y corroborar los hallazgos reportados en la literatura, ejercicio que permitirá validar si las redes neuronales mantienen ventajas predictivas consistentes, particularmente en contextos característicos de relacionales no lineales y estructuras temporales complejas.

BIBLIOGRAFÍA

- Bai, J., Wang, J., Ran, J., Li, X., & Tu, C. (2024). An Improved Neural Network Algorithm for Energy Consumption Forecasting. *Sustainability*, 16(21), 9332.
- Benidis, K., Rangapuram, S. S., Flunkert, V., Wang, Y., Maddix, D., Turkmen, C., ... & Januschowski, T. (2022). Deep Learning for Time Series Forecasting: Tutorial and Literature Survey. *ACM Computing Surveys*, 55(6), 1–36.
- Chollet, F. (2021). *Deep Learning With Python*. Simon and Schuster.
- Cui, H. (2016). China's Economic Growth and Energy Consumption. *International Journal of Energy Economics and Policy*, 6(2), 349–355.
- García, F., Hernández, G., & Luna, N. (2011). *Manual de estadísticas energéticas OLADE*.
- Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep Learning* (Vol. 1, No. 2). Cambridge: MIT Press.
- Guachimboza-Davalos, J. I., Llanes-Cedeño, E. A., Rubio-Aguilar, R. J., Peralta-Zurita, D. B., & Núñez-Barrionuevo, O. F. (2021). Prediction of Monthly Electricity Consumption by Cantons in Ecuador Through Neural Networks: A Case Study. En *Systems and Information Sciences: Proceedings of ICCIS 2020* (pp. 28–37). Springer International Publishing.
- Guamán, W., Benalcazar, P., Cordova-Garcia, J., & Torres, M. (2024, October). Machine Learning-Based Projections of Long-Term Electricity Consumption: The Case Study of Ecuador. En *International Conference on Advanced Research in Technologies, Information, Innovation and Sustainability* (pp. 174–187). Cham: Springer Nature Switzerland.
- Hastie, T., Tibshirani, R., Friedman, J., & Franklin, J. (2005). The Elements of Statistical Learning: Data Mining, Inference and Prediction. *The Mathematical Intelligencer*, 27(2), 83–85.
- Haykin, S. (2009). *Neural Networks and Learning Machines* (3rd ed.). Pearson Education India.
- Hinojosa Bassantes, J. S. (2022). *Implementación de un modelo de predicción de la demanda eléctrica mediante redes neuronales artificiales*.
- International Energy Agency. (2023). *Energy Efficiency 2023*. <https://www.iea.org/reports/energy-efficiency-2023>
- Nugaliyadde, A., Somaratne, U., & Wong, K. W. (2019). Predicting Electricity Consumption Using Deep Recurrent Neural Networks. *arXiv preprint arXiv:1909.08182*.
- OIEA. (2008). *Indicadores energéticos del desarrollo sostenible: directrices y metodologías*.

- Owda, H. M., Omoniwa, B., Shahid, A. R., & Ziauddin, S. (2014). Using Artificial Neural Network Techniques for Prediction of Electric Energy Consumption. *arXiv preprint arXiv:1412.2186*.
- Rodríguez-Cruz y Balda Cruz. (2022). Descomposición de la variación del índice de intensidad energética y desacople de la economía en Ecuador. *Compendium: Cuadernos de Economía y Administración*, 9(2), 115–130. <https://doi.org/10.46677/compendium.v9i2.1002>
- Rodríguez-Cruz, L. X. (2024). Análisis de la eficiencia y rentabilidad de las empresas manufactureras ecuatorianas: una aplicación de algoritmos no supervisados. *Cuestiones Económicas*, 34(2). <https://doi.org/10.47550/RCE/34.2.5>
- Vallejo, C. (2017). Análisis de intensidad energética para los sectores de uso final de la economía ecuatoriana. *Cuestiones Económicas*, 27(2), 17–17.
- Zanek, F. (2023). *Pronóstico de demanda de electricidad con redes neuronales artificiales para Salta, Argentina*. *Revista Científica de Informática ENCRIPAR*, 6(12), 1–18.

ANEXOS

Anexo A. Energía facturada por provincia 2018-2023 (GWh)

Provincias	2018	2019	2020	2021	2022	2023
Azuay	918,48	929,83	855,35	916,71	947,27	955,15
Bolívar	68,48	69,07	70,41	71,79	72,32	76,38
Carchi	83,51	85,37	84,62	88,13	87,69	92,62
Cañar	182,87	194,96	177,19	195,54	186,37	199,62
Chimborazo	324,38	308,63	288,47	378,38	400,16	342,01
Cotopaxi	486,77	449,38	418,79	466,43	493,84	460,74
El Oro	826,45	892,64	881,38	946,69	988,79	1.018,93
Esmeraldas	445,42	438,95	422,01	445,25	474,44	500,50
Galápagos	49,25	53,94	47,53	48,66	48,28	70,98
Guayas	6.740,21	6.954,41	6.890,15	7.289,97	7.435,29	8.235,84
Imbabura	293,41	302,45	289,17	314,98	328,57	347,61
Loja	243,75	248,80	244,38	246,24	261,52	274,32
Los Ríos	635,62	663,96	658,15	683,16	676,40	756,21
Manabí	1.368,14	1.426,03	1.455,90	1.464,39	1.549,09	1.797,17
Morona Santiago	71,22	72,95	72,57	78,91	82,61	86,34
Napo	81,92	85,84	82,50	87,91	87,31	95,64
Orellana	123,49	130,74	132,00	134,40	140,39	154,23
Pastaza	58,17	59,80	59,10	61,89	61,42	68,05
Pichincha	3.910,24	3.743,40	3.424,28	3.505,90	3.601,82	3.730,60
Santa Elena	356,73	393,02	396,25	427,56	445,76	444,35
Santo Domingo de los Tsáchilas	420,29	429,05	422,84	445,36	460,32	503,82
Sucumbíos	478,53	555,74	561,73	579,41	601,91	601,95
Tungurahua	462,85	463,05	439,68	465,20	451,58	485,75
Zamora Chinchipe	53,01	145,50	278,31	449,94	715,16	780,06

Fuente: ARCONEL

Elaboración: autores

Anexo B. Valor agregado bruto provincial 2018-2023 (miles de USD)

Provincias	2018	2019	2020	2021	2022	2023
Azuay	4.651.017	4.655.332	4.190.464	4.606.093	5.084.289	5.003.715
Bolívar	416.238	413.974	407.563	427.146	474.118	474.612
Carchi	599.037	604.673	532.326	582.816	600.053	592.440
Cañar	802.705	788.835	705.026	766.147	815.845	807.579
Chimborazo	1.339.981	1.378.326	1.179.658	1.251.765	1.387.496	1.356.964
Cotopaxi	1.772.718	1.733.860	1.561.525	1.665.710	1.680.184	1.693.471
El Oro	3.359.398	3.383.287	3.206.664	3.601.883	3.821.418	4.011.472
Esmeraldas	2.473.491	2.151.602	1.932.036	2.017.708	1.825.719	1.805.635
Galápagos	339.378	361.892	241.890	222.136	248.505	282.115
Guayas	29.484.069	30.192.388	28.743.798	30.802.249	34.108.591	35.912.077
Imbabura	1.567.277	1.579.691	1.315.622	1.347.340	1.530.066	1.655.475
Loja	1.526.839	1.622.200	1.407.102	1.458.391	1.778.577	1.890.553
Los Ríos	3.140.980	3.142.045	3.220.731	3.369.363	3.490.584	3.905.489
Manabí	4.998.039	4.898.244	4.515.587	4.906.461	5.496.687	5.903.548
Morona Santiago	453.101	446.051	421.692	434.820	484.079	523.641
Napo	532.271	497.831	420.198	451.705	512.414	494.701
Orellana	5.376.224	5.400.857	3.335.301	5.284.770	7.226.709	5.822.719
Pastaza	875.868	782.386	609.717	889.916	1.115.514	929.851
Pichincha	28.267.714	28.054.282	24.411.601	27.201.525	28.019.826	30.310.540
Santa Elena	1.577.586	1.351.889	958.249	1.121.571	1.196.778	1.217.470
Santo Domingo de los Tsáchilas	1.806.044	1.769.035	1.670.042	1.743.385	1.902.616	2.003.079
Sucumbíos	3.268.888	2.848.515	1.999.518	3.008.797	3.833.738	3.104.639
Tungurahua	2.524.904	2.600.733	2.151.575	2.406.918	2.588.576	2.752.890
Zamora Chinchipe	314.384	346.946	676.613	1.060.200	1.312.899	1.656.755

Fuente: Banco Central del Ecuador BCE

Elaboración: autores

Anexo C. Unidades de medida eléctrica

Prefijo	Símbolo	Múltiplo	Potencia de diez
Tera	T	1.000.000.000.000	10 ¹²
Giga	G	1.000.000.000	10 ⁹
Mega	M	1.000.000	10 ⁶
Kilo	k	1.000	10 ³
		1	10 ⁰
Mili	m	1/1.000	10 ⁻³
Micro	μ	1/1.000.000	10 ⁻⁶
Nano	n	1/1.000.000.000	10 ⁻⁹
Pico	p	1/1.000.000.000.000	10 ⁻¹²

Fuente: *Manual de estadísticas energéticas* OLADE (2017)

Elaboración: autores

Nota: Adicional de las unidades de medida eléctricas “estándar”, se utilizan otras unidades para indicar otros calores y cantidades como:

Wh(Watts por hora): cantidad de energía eléctrica consumida por un circuito durante un periodo de tiempo;

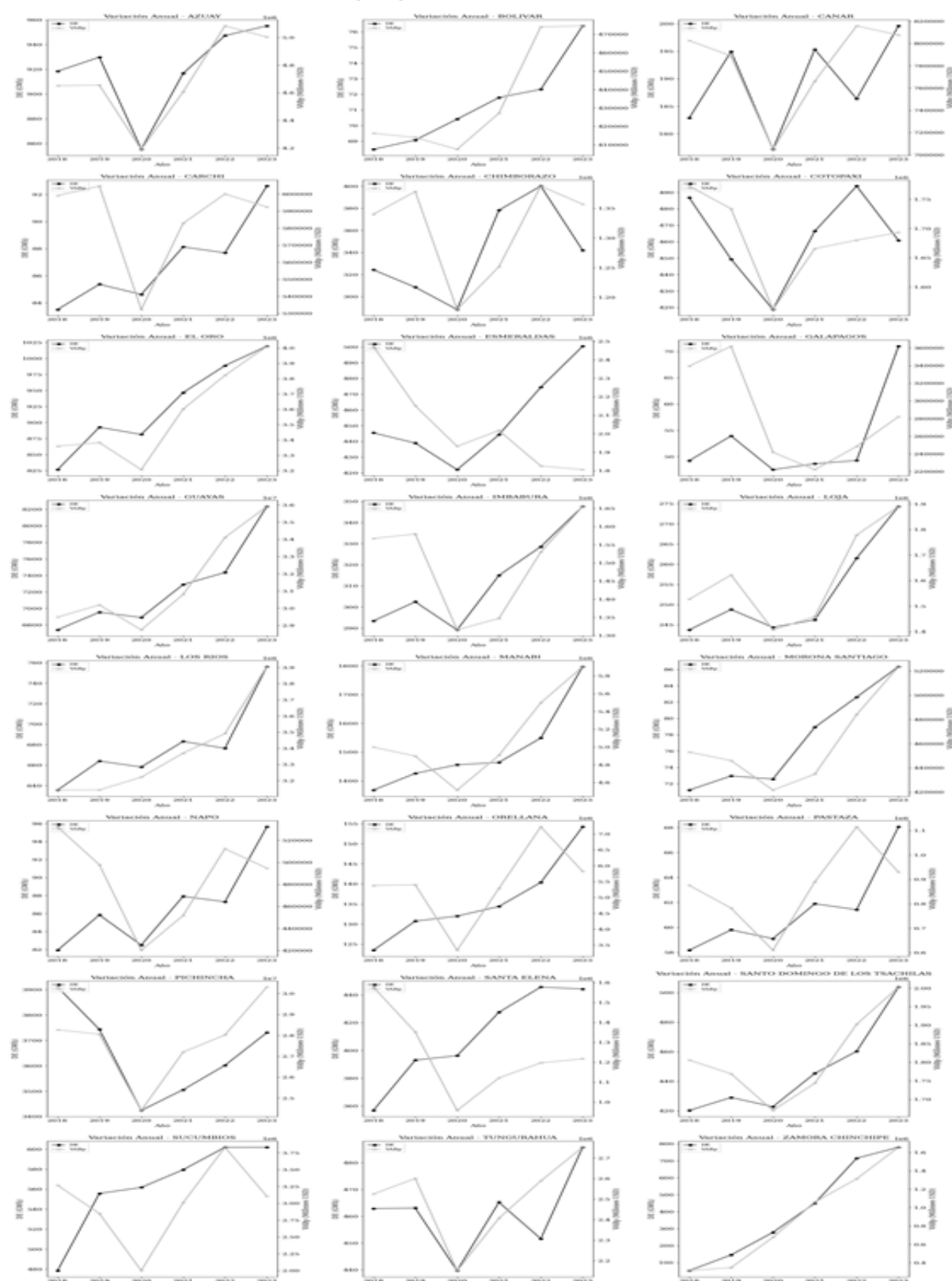
kWh (kilo watts-hora) = 1.000 watts-hora;

MWh (mega watts-hora) = 1.000.000 watts-hora, y

GWh(giga watts-hora) = 1.000.000.000 watts-hora.

Nota: De acuerdo al *Manual de estadísticas energética 2017* de OLADE, el factor de conversión calórico aplicado para Ecuador, en su calidad de país miembro, es de 0,620 para transformar unidades de GWh (unidad original) a kBEP.

Anexo D. Variación anual de demanda de energía (GWh) y valor agregado bruto (USD)



Fuente: Banco Central del Ecuador BCE y ARCONEL

Elaboración: autores

Anexo E. Matriz comparativa de arquitecturas de redes neuronales

Características	Feedforward	LSTM	Feedforward secuencial	LSTM robusto
Estructura	4 capas densas (64-32-16-1)	2 capas LSTM (64 unidades)+1 densa	4 capas densas (64-32-16-1) con entrada vectorizada	2 capas LSTM(64 unidades) + 2 capas densas (32-1)
Entrada	6 <i>features</i> (sin contexto temporal)	Secuencia 3D (ventana=3,6 <i>features</i> / paso)	18 <i>features</i> (ventana = 3 aplanada)	Secuencia 3D (ventana=3,6 <i>features</i> / paso) <i>Dropout</i> (30 %)
Mecanismo	<i>Dropout</i> (30 %), BatchNorm	<i>Dropout</i> (30 %)	<i>Dropout</i> (30 %) BatchNorm	BatchNorm sigmoide final
Optimizador	AdamW (lr=0,001)	Adam (lr=0,001)	AdamW (lr=0,001)	Adam (lr=0,001)
Épocas de entrenamiento	200 (sin <i>early stopping</i>)	100 (sobreajuste detectado)	200 (convergencia estable)	127 8 <i>early stopping</i> con paciencia =10)
Tamaño de batch	32	16 (limitado por memoria)	64 (eficiencia por datos 2D)	16
Datos de validación	20 % del conjunto de entrenamiento (<i>split</i> inicial)	20 % del conjunto de entrenamiento (<i>split</i> inicial)	20 % del conjunto de entrenamiento (<i>split</i> inicial)	10 % del conjunto de entrenamiento (<i>split</i> anidado para <i>early stopping</i>)
Tiempo época(GPU T4)	~ 42 segundos	~ 1 min 15 seg	~ 38 segundos	~ 1 min 50 seg
Curva de aprendizaje	Alta variabilidad post-epoch 50	Sobreajuste después de epoch 85	Estable después de epoch 35	Mejor val_loss en epoch 97
Preprocesamiento temporal	Ninguno	Ventana de 3 pasos	Secuencias aplanadas (3 pasos →18 <i>features</i>)	Ventana de 3 pasos

Características	Feedforward	LSTM	Feedforward secuencial	LSTM robusto
Hiperparámetros clave	-	hidden_size=64, num_layers=2	-	hidden_size=64, capas FC adicionales

Fuente: Arquitecturas de redes neuronales propuestas en la investigación

Elaboración: autores

Nota:

Partición de Datos:

Trainging/test: 80 %/20 % para todos los modelos (sin *shuffle*, preservando orden temporal)

Validación:

Feedforward, LSTM, *Feedforward* secuencial: 20 % del entrenamiento inicial

LSTM robusto: 10 % de entrenamiento (*split* anidado) para monitoreo de *early stopping*.

Hiperparámetros comunes: compartido para todas las arquitecturas

Para reproducibilidad, SEED = 42

Para regularización, Dropout_rate = 0,3

Tasa de aprendizaje, Learning_rate = 0,001

Función de pérdida, Loss_function = nn.MSELoss

Anexo F. Índice intensidad energética en kBEP/103 USD - *feedforward* secuencial

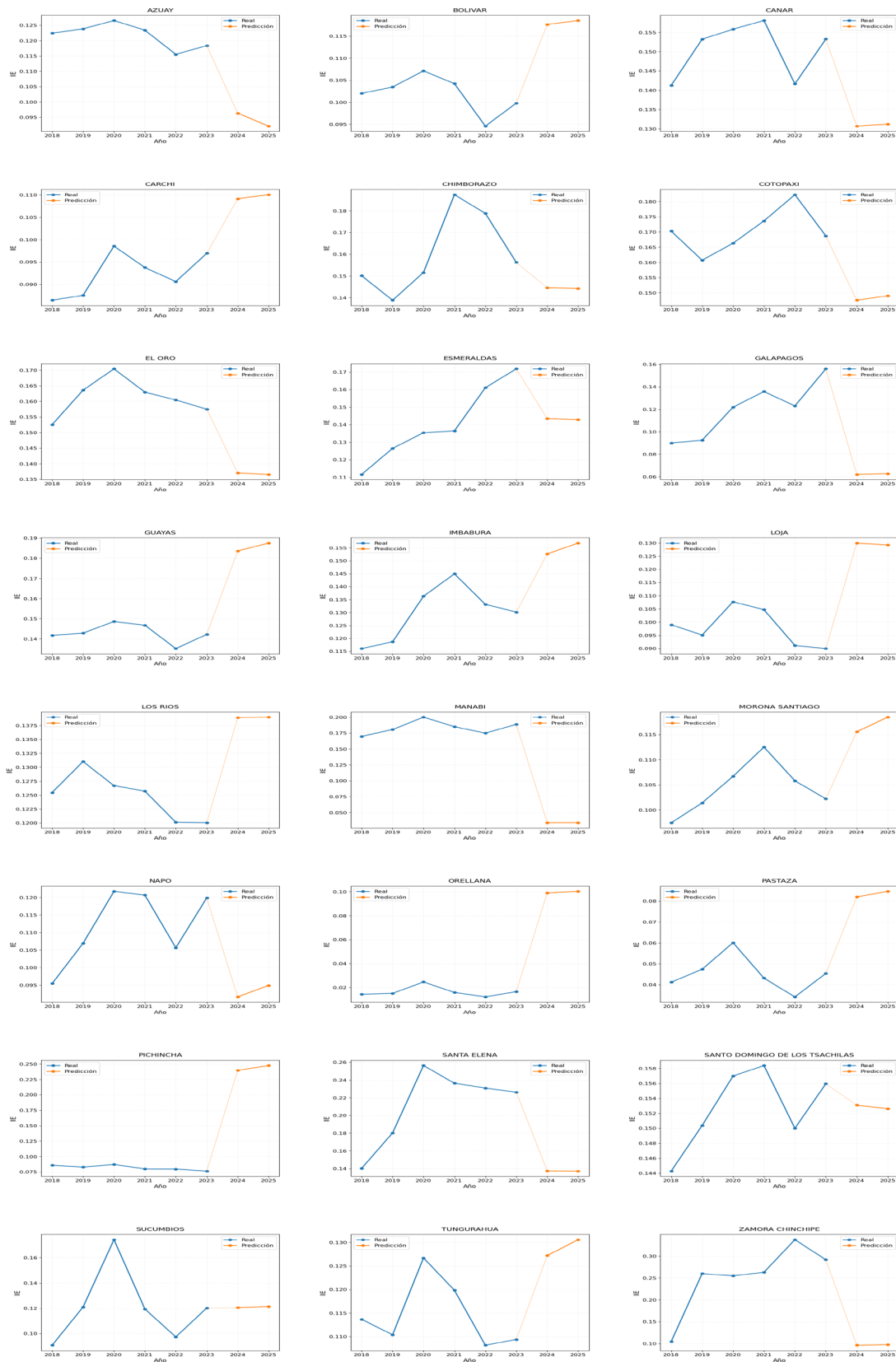
Provincias	2018	2019	2020	2021	2022	2023	2024p	2025p
Azuay	0,122437	0,123835	0,126553	0,123393	0,155140	0,118351	0,096323	0,092118
Bolívar	0,102023	0,103445	0,107110	0,104203	0,094572	0,099778	0,117653	0,118550
Carchi	0,086432	0,087534	0,098557	0,093753	0,090605	0,096929	0,109158	0,110062
Cañar	0,141247	0,153233	0,155821	0,158078	0,141632	0,153254	0,130687	0,131216
Chimborazo	0,150088	0,138828	0,151613	0,187412	0,178811	0,156265	0,144574	0,144216
Cotopaxi	0,170246	0,160691	0,166280	0,173612	0,182231	0,168682	0,147533	0,148998
El Oro	0,152527	0,163580	0,170413	0,162956	0,160425	0,157492	0,137011	0,136552
Esmeraldas	0,111648	0,126487	0,135425	0,136509	0,161116	0,171856	0,143473	0,142927
Galápagos	0,089973	0,092411	0,121827	0,135814	0,122950	0,155992	0,061960	0,062705
Guayas	0,141735	0,142809	0,148620	0,146735	0,135153	0,142187	0,183582	0,187560
Imbabura	0,116070	0,118706	0,136274	0,144943	0,133140	0,130185	0,152613	0,156846
Loja	0,098979	0,095091	0,107679	0,104683	0,091164	0,089962	0,129931	0,129199
Los Ríos	0,125465	0,131015	0,126696	0,125709	0,120143	0,120049	0,138894	0,138967
Manabí	0,169716	0,180501	0,199898	0,185046	0,174730	0,188742	0,033921	0,033992
Morona Santiago	0,097454	0,101399	0,106697	0,112516	0,105805	0,102228	0,115570	0,118467
Napo	0,095422	0,106905	0,121728	0,120663	0,105641	0,119864	0,091611	0,094891
Orellana	0,014241	0,015009	0,024538	0,015768	0,012044	0,016422	0,099018	0,100340
Pastaza	0,041177	0,047388	0,060097	0,043118	0,034137	0,045374	0,082029	0,084675
Pichincha	0,085764	0,082729	0,086969	0,079909	0,079698	0,076309	0,239808	0,247511
Santa Elena	0,140197	0,180246	0,256379	0,236354	0,230929	0,226286	0,137282	0,137094
Santo Domingo de los Tsáchilas	0,144282	0,150371	0,156979	0,158383	0,150003	0,155944	0,153082	0,152607

Provincias	2018	2019	2020	2021	2022	2023	2024p	2025p
Sucumbíos	0,090761	0,120961	0,174178	0,119395	0,097342	0,120210	0,120560	0,121436
Tungurahua	0,113550	0,110388	0,126699	0,119831	0,108160	0,109400	0,127199	0,130567
Zamora Chinchipe	0,104542	0,260012	0,255024	0,263123	0,337725	0,291918	0,096025	0,097329

Fuente: Banco Central del Ecuador BCE y ARCONEL

Elaboración: autores

Anexo G. Evolución del índice de intensidad energética por provincias



Fuente: Banco Central del Ecuador BCE y ARCONEL

Elaboración: autores