

Multivariate Statistical Analysis Based Methodology for Long-Term Demand Forecasting

J. Jiménez, A. Pertuz, C. G. Quintero, and J. Montaña

Abstract—Forecasting models are necessary in electrical utilities to set the energy cover for several years in order to minimize the operational cost. In this sense, a low error is required to avoid high levels energy exchange with market and hence an increase in the operation costs. However, demand energy has monthly a behavior correlated with some economics variables (e.g., gross domestic product, gold price), type and amount of days, population, historical demand, inter alia. For this reason, it is necessary to design a statistical methodology that allows selecting suitable variables and characterize them to establish which factors are significant for understanding. This paper proposes a methodology to identify those significant factors in the monthly forecasting model for energy demand in the power purchases area. The climatic scenario to increase the sensibility of the model, when the probability of the “Niño” or “Niña” phenomena increase month to month, was proposed.

Index Terms—Demand forecasting, Artificial neural networks, Factorial analysis, Time series analysis.

I. INTRODUCCIÓN

EL mercado de la energía eléctrica en el mundo funciona bajo la premisa de la oferta y la demanda entre los generadores y distribuidores, los cuales son regulados a través del valor establecido en la bolsa de energía. El valor del kWh dentro del mercado eléctrico depende de la antelación con la que se asegure la compra de la energía requerida por parte del comercializador. Por tal motivo, se recurre a predicciones de energía eléctrica de hasta 5 años para garantizar el cubrimiento de la demanda con el mayor margen de rentabilidad posible [1].

El mercado de la energía eléctrica en Colombia ha experimentado diversos inconvenientes en la prestación de su servicio de distribución de energía. Lo anterior debido al aumento en el número de los clientes a los cuales suministrarle energía acarreando consigo un incremento en la demanda de energía eléctrica nacional [2]. De igual manera el crecimiento de la incertidumbre al momento de gestionar la distribución como resultado de la configuración del sistema eléctrico del país. Asimismo, no es posible dejar de mencionar los retos que representan para el suministro fenómenos climáticos intensos como el fenómeno del Niño y de la Niña [2], [3]. Ahora bien, este aumento de la demanda de energía y el reto de garantizar

el suministro, reflejan con claridad la necesidad de tener modelos de predicción de consumo de energía eléctrica confiables. Dichos modelos deben brindar información suficiente y acertada para estimar los cambios en la demanda de energía eléctrica vistos desde la perspectiva del distribuidor y a su vez, permitan identificar los factores que inciden dentro la demanda de dicha energía [4].

Las compras de energía (comercializadores comprándole a generadores) se pueden dar en dos escenarios. Uno de largo plazo el cual consiste en contratos, resultado de una convocatoria pública, y otro de corto plazo conocido como bolsa de energía, en donde generadores y comercializadores compran y venden energía según el precio al que se liquide el kilovatio hora (kWh) en un periodo de tiempo [5]. En el primer escenario se basa el enfoque de pronóstico aquí analizado, debido a que se plantea una propuesta de pronóstico mensual para los siguientes dos años con el fin de suministrar dicha información a la convocatoria pública que finalmente es quien define qué generadores se adjudican parte de la demanda de energía a cubrir [6].

En la actualidad, la mayoría de unidades de compra de energía establecen su presupuesto en el pronóstico de la demanda anual a través de una metodología basada en las políticas de crecimiento de la infraestructura eléctrica establecidas por la empresa. Este proceso de proyección como apoyo al presupuesto y valores base de demanda para la inscripción de la convocatoria abierta se lleva a cabo una vez al año [6]. Por tal motivo, el interés hacia modelos de previsión no es nuevo en el mundo, y se han desarrollado modelos para diferentes países y ciudades que permiten una gestión inteligente del recurso eléctrico. Sin embargo, en Colombia los desarrollos entorno a entender la demanda nacional y el consumo en ciudades costeras ha sido muy escaso y poco aplicado. Por tal razón, este estudio tiene como finalidad el análisis estadístico de las series temporales de demanda nacional y consumo de energía eléctrica en la Región Caribe, basado en técnicas de análisis estadístico multivariado como el análisis factorial [7]. De acuerdo con lo anterior, se busca a través del análisis factorial encontrar los factores que inciden en la demanda de energía, usando la relación con variables macroeconómicas como el precio de la bolsa de energía, el Índice Mensual de Actividad Económica (IMACO), la población, el valor del dólar, el precio del petróleo y el índice de precios de combustibles, variables exógenas propias del clima y coyunturales como la incidencia de los fenómenos del Niño y la Niña, con la intención de establecer un conjunto de factores significativos, que sean nombrados y establecidos propiamente, que posteriormente permitan hacer proyecciones para el consumo de energía.

J. Jiménez, Universidad del Norte, Km 5 vía Puerto Colombia, Barranquilla, Colombia (e-mail: jmares@uninorte.edu.co).

A. Pertuz, Universidad del Norte, Km 5 vía Puerto Colombia, Barranquilla, Colombia (e-mail: abimaelp@uninorte.edu.co).

C. G. Quintero, Universidad del Norte, Km 5 vía Puerto Colombia, Barranquilla, Colombia (e-mail: christianq@uninorte.edu.co).

J. Montaña, Universidad Técnica Federico Santa María, Avenida España 1680, Valparaíso, Chile (e-mail: johny.montana@usm.cl).

El trabajo aquí propuesto se encuentra organizado de la siguiente manera. En la sección II se realiza una presentación de los trabajos relacionados. En la sección III se presenta una descripción de la metodología propuesta. Finalmente, en las secciones IV y V se muestra un análisis de los resultados obtenidos en donde se destaca la adaptabilidad de la metodología propuesta para la predicción de la demanda de energía eléctrica mensual en la Región Caribe Colombiana.

II. TRABAJOS RELACIONADOS

Para las empresas encargadas de la distribución de energía eléctrica es importante llevar a cabo un proceso de estimación de la demanda de energía eléctrica para los próximos 1, 3 y hasta 5 años con bajos niveles de error. Lo anterior debido a la necesidad de establecer niveles de cobertura que garanticen la provisión del servicio por lo menos en un 90% de la demanda esperada.

En primera instancia es notable como en los trabajos [5], [17], [21], [23], [27] se vinculan las variables económicas tales como el crecimiento poblacional, precio del crudo, precio del platino, precio del dólar, entre otras para llevar a cabo los modelos de predicción mensual y anual. Los resultados alcanzados por los modelos planteados presentaron un desempeño MAPE (*Mean Absolute Percentage Error*) por debajo del 3%. En segunda instancia se destaca la utilización amplia de redes neuronales ANN (*Artificial Neural Networks*) en comparación con técnicas basadas en la metodología ARIMA (*AutoRegressive Integrated Mobile Average*). Asimismo, es importante resaltar como el diseño de modelos de predicción de la demanda de energía anual difiere en gran medida con respecto a lo planteado en los de alcance horario presentados en [12]–[15], en donde se utiliza la información climática como principal variable exógena.

De acuerdo a lo mostrado en la Tabla I, las redes neuronales se presentan como la técnica de modelado usualmente utilizada para la predicción de la demanda de energía. Las series de tiempo analizadas mostraron características no estacionarias lo cual motiva el uso de técnicas como máquinas de SVM (*Support Vector Machine*) [13], [16], FL (*Fuzzy Logic*) [10], [17], [18], MC (*Markov Chain*) [19] y ANN [7], [11], [20]. A pesar de los resultados alcanzados por algunos autores, en la mayoría de los casos no se describe una metodología estadística que justifique la selección de las variables exógenas utilizadas por los modelos propuestos. La propuesta aquí planteada presenta una metodología para evaluar la interacción de variables exógenas como el crecimiento poblacional, el producto interno bruto (PIB), el precio del oro, el precio del crudo, entre otras, con la demanda eléctrica mensual de la Región Caribe Colombiana.

III. METODOLOGÍA PROPUESTA

El desarrollo de la metodología propuesta en este trabajo consta de las siguientes etapas fundamentales:

1. Determinación y selección de las variables de entrada.
2. Identificación de la estructura de las redes neuronales para los datos de demanda de energía eléctrica.

3. Determinación de los parámetros de la red neuronal.

TABLA I
MODELOS UTILIZADOS PARA LA PREDICCIÓN DE LA DEMANDA DE ENERGÍA ELÉCTRICA

Autor	Variable de entrada	Alcance	Técnica	MAPE
[8]	L,VE	HR	ARIMA	2%
[21]	L	D	PCA, ES, ANN, ARIMA	1.4%,1.35%,1.5%,2.1%
[22]	L,T	D	ANN (dinámica)	1.63%
[23]	L,T,M,e	M	SVM	3.9%
[24]	L	D	ANN,ANFIS,ARIMA	1.25%,18.72%,5.07%
[7]	L,T,H	HR,D	ANN	3.81%
[12]	L	HR	ARIMA,ANN,PM	1.67%,2.2%,1.62%
[13]	L,T,H	HR	SVM	2.04%
[17]	L	D	FL	3.57%
[16]	L	A	SVM,GA	1.84%
[20]	L,T	M,D	ANN	10% (M), 3,12% (D)
[25]	L,T,W,TC	D	ANN,ANFIS,HANFIS	1.79%, 1.21%, 0.83%
[26]	L,T	D	ANN,ARMAX	4%
[9]	L,VE	A	ANN	
[14]	L	HR	ANN,SVR	
[27], [28]	L, H	D	ANN (Dynamic)	
[10]	L, VE	A	FL	
[29]	L	M	ANN, ARIMA	0.018%, 0.04%
[30]	L,VE	A	ANN	2.44%
[19]	L	A	GM	0.54%
[31]	L	M	ANN	5%
[32]	L, T, H	D	ANN	7.46%
[11]	L,VE	A	ANN	0.95%
[15]	L,T	HR	ANN	1.94%
[18]	L	D	FL,ES	1.12%

Variable de entrada – L = Carga, T = Temperatura; H = Humedad, VE = Variables económicas, M = Mes, e = Error, TC = Tipo de cliente, W = Clima (variables diferentes a T y H).

Alcance – HR = Horario, M = Mensual, A = Anual, D = Diario.

Técnica – ARIMA = Autoregressive Integrated Moving Average, PCA = Principal Components Analysis, ES = Exponential Smoothing, ANN = Artificial Neural Networks, SVM = Support Vector Machines, FL = Fuzzy Logic, GA = Genetic Algorithms, ANFIS = Adaptive Neuro Fuzzy Inference System, PM = ARIMA modified, HANFIS = Hybrid ANFIS, ARMAX = Autoregressive Moving Average Model, GM = Grey Markov.

La propuesta aquí planteada busca evaluar la interacción de algunas variables exógenas con la demanda de energía eléctrica mensual de la Región Caribe Colombiana. Para ello, se hace necesario recurrir a una metodología estadística y sistemática que permita establecer qué variables resultan significativas y contribuyen dentro del grado de variabilidad de la demanda de energía. En primera instancia se evalúa el grado de fiabilidad de las variables con el objetivo de verificar si éstas de manera preliminar explican el comportamiento de

la demanda. En segunda instancia se hace uso del análisis factorial para reducir la cantidad de variables de tal forma que se asocien aquellas que presenten un comportamiento altamente correlacionado entre sí. Acto seguido, se evalúa el grado de sensibilidad producido por fenómenos externos tales como el fenómeno del Niño o de la Niña. Finalmente, se plantea la selección de la configuración de la red neuronal a partir del diseño planteado en [33].

A. Determinación y Selección de las Variables de Entrada

En esta sección se lleva a cabo el análisis de las variables explicativas con el fin de identificar aquellas con un mayor grado de significación dentro de las variaciones de la demanda de energía eléctrica. Para ello se acudirá al análisis factorial con el fin de identificar aquellas variables exógenas con un mayor grado de incidencia dentro de la demanda de energía. Para el análisis aquí planteado se tomaron como referencia los datos de la demanda de energía eléctrica de la Región Caribe comprendidos entre los años 2005 y 2016.

En la Fig. 1 se presenta una muestra de la demanda de energía eléctrica de la Región Caribe de los años 2011 a 2016. Se resalta la presencia de la componente estacional y de tendencia, lo cual evidencia la naturaleza no estacionaria de la serie de tiempo. El test de *Dickey-Fuller* aplicado a la serie de tiempo demostró la aseveración antes mencionada, dado que la hipótesis nula (existencia de una raíz unitaria) no fue rechazada.

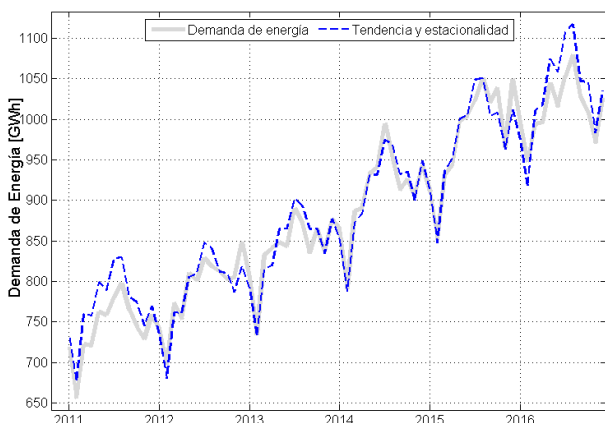


Fig. 1. Demanda de energía eléctrica mensual para el periodo 2011 – 2016.

En la Fig. 2 se muestra el resultado del test de multicomparación de *Tukey-Kramer* en donde se evidencia que la demanda de energía de todos los meses del periodo seleccionado no presentó diferencia significativa. Lo anterior permite establecer un único modelo de predicción dentro de la metodología aquí planteada.

La Fig. 3 permite observar que la demanda de energía mensual de la Región Caribe de los últimos cuatro meses, proporciona información significativa debido a que mantiene un nivel de correlación por encima del umbral mínimo dado por la relación $\frac{2}{\sqrt{n}}$.

Por otro lado, con el objetivo de evaluar las variables climáticas dentro del análisis propuesto, se recurrió a la información de la página <http://www.wunder-ground.com/>.

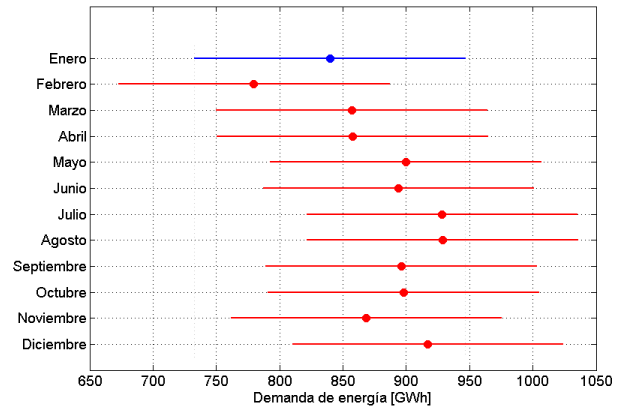


Fig. 2. Multicomparación de medias de la demanda de energía eléctrica diaria por mes.

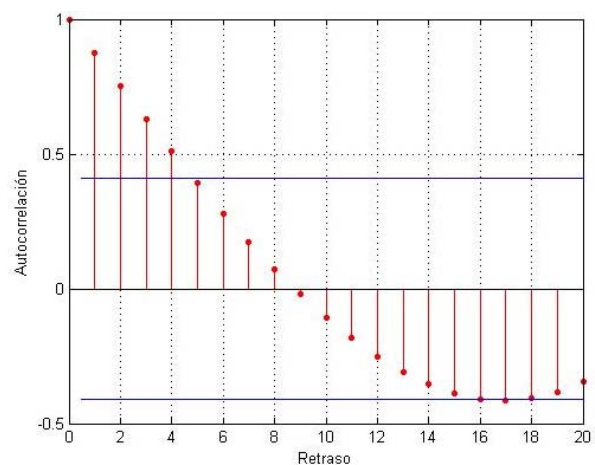


Fig. 3. Autocorrelación total de la demanda de energía eléctrica.

En la Tabla II se presenta la correlación a nivel mensual de las diferentes variables que fueron consideradas, en donde se muestra una leve o nula correlación de la demanda mensual de energía con la temperatura, humedad y velocidad del viento para la ventana de tiempo escogida (2005 a 2016). De igual manera, al incluir estas variables dentro de los datos a utilizar previo al análisis factorial, proporcionaron un *Alfa de Cronbach* igual a 0.2, lo cual desestima el uso de estas variables dentro del análisis factorial.

TABLA II
ANÁLISIS DE CORRELACIÓN DE LAS VARIABLES CLIMÁTICAS CON LA DEMANDA DE ENERGÍA ELÉCTRICA

Variable	Correlación Mensual
Temperatura Mínima	-0.0368
Temperatura Media	0.3128
Temperatura Máxima	0.0921
Humedad Mínima	0.0119
Humedad Media	0.1840
Humedad Máxima	0.2206
Velocidad Mínima del Viento	-0.2009
Velocidad Media del Viento	-0.1963
Velocidad Máxima del Viento	-0.1134

A pesar de los resultados mostrados en la Tabla II, es necesario seguir con la caracterización de los fenómenos climáticos como el Niño y la Niña para medir su efecto dentro

de la demanda de energía eléctrica de la Región Caribe. En la Fig. 4 se muestra el grado de afectación de los fenómenos del Niño y de la Niña dentro del comportamiento de la demanda eléctrica para los años 2011 a 2016 [34]. Es de resaltar como la aparición del fenómeno del Niño en el segundo semestre del 2015 produjo un aumento de la demanda de energía. Mientras que el fenómeno de la Niña durante el segundo semestre del año 2016 ocasionó una disminución de la demanda para el mismo año. Por tal motivo, se hace necesaria la inclusión de escenarios climáticos (Niño, Niña o Normal) que permitan integrar el efecto climático antes mencionado.

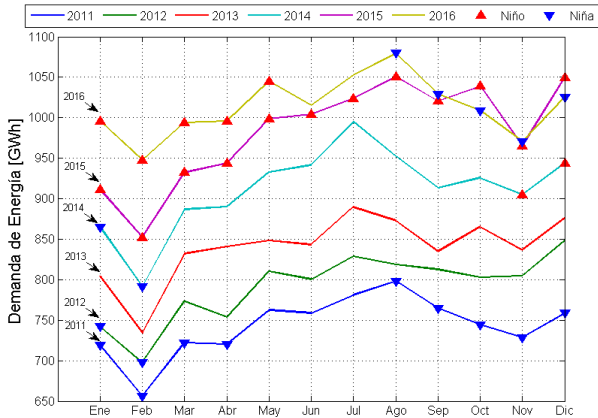


Fig. 4. Demanda de energía mensual y caracterización climática de la Región Caribe para el periodo 2011 – 2016.

Por otro lado, se hace necesario incorporar variables que brinden indicios claros sobre la tendencia anual de la serie, es de ahí que se haya optado por incorporar algunas variables relacionadas con el crecimiento demográfico de cada uno de los departamentos de la Región Caribe (Atlántico, Bolívar, Cesar, Córdoba, La Guajira, Magdalena, Sucre). De igual manera, la demanda de energía se ve afectada por algunas variables económicas que inciden directamente en la adquisición de bienes y servicios, tal es el caso del PIB.

En la Fig. 5 se muestra el factor de contribución de cada una de las variables consideradas dentro del análisis utilizado para la pre-selección de cada una de las variables macroeconómicas. Los resultados de esta etapa brindan información previa al análisis factorial aquí propuesto.

B. Comparación con Otras Técnicas: Una Mirada Desde el Análisis Factorial

El análisis factorial es una técnica estadística de reducción de dimensiones que describe cada variable en términos de una combinación lineal de un pequeño número de factores comunes no observables y un factor único para cada variable. De esta manera, el objetivo es encontrar los factores comunes que recojan el máximo de información de las variables originales [35].

Para el análisis propuesto, se incluye un total de 29 variables, que basado en lo establecido en el estado del arte, se considera y espera tengan influencia sobre la demanda de energía en la Región Caribe. A continuación se listan dichas

variables:

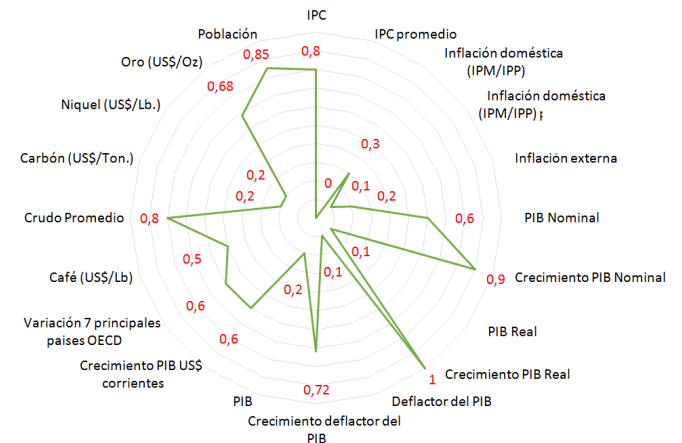


Fig. 5. Factor de contribución de las variables económicas dentro de la variabilidad de la demanda de energía eléctrica.

1. Indicador número del mes a pronosticar.
2. Demanda nacional mensual de energía.
3. Promedio mensual de demanda de energía nacional en los últimos cuatro (4) meses.
4. Demanda mensual de energía en la Región Caribe.
5. Promedio mensual de demanda de energía en la Región Caribe en los últimos cuatro (4) meses.
6. Demanda Nacional no atendida.
7. Exportaciones de Energía.
8. Importaciones de Energía.
9. Generación Hidráulica.
10. Generación Térmica.
11. Precio promedio del dólar.
12. Precio dólar último día del mes.
13. PIB.
14. IMACO.
15. Índice de precios de combustibles.
16. Precio del petróleo tipo Brent.
17. Precio del crudo.
18. Población de la Región Caribe Colombiana.
19. Precio del oro.
20. Precio de la plata.
21. Precio del platino.
22. Precio de la bolsa de energía.
23. Fenómeno del Niño
24. Fenómeno de la Niña.
25. Número de días en el mes a pronosticar.
26. Días laborales del mes a pronosticar.
27. Días sábados del mes a pronosticar.
28. Días domingos del mes a pronosticar.
29. Días festivos del mes a pronosticar.

Como puede evidenciarse en las variables incluidas, se tienen en cuenta no solo variables propias del sector energético, sino también variables que describen la producción industrial del país y el crecimiento económico, además de variables climáticas como el fenómeno del Niño y la Niña en el intervalo de Enero de 2005 a Diciembre de 2016, las medidas de las variables son mensuales y las variables (1), (16) y (17) son de tipo categórico.

Es importante anotar que no se incluye la temperatura, la humedad y la velocidad del viento en este análisis, debido a que el *alfa de Cronbach* mostrado en la Tabla III, el cual mide la fiabilidad de escala en ese segmento de datos, no generaba un valor aceptable para el análisis. Acorde con lo anterior y la poca correlación mostrada en la Tabla II, se decide realizar el análisis sin las variables temperatura, humedad y velocidad del

viento.

Por otro lado, el análisis inicial de escala muestra una covarianza promedio negativa entre los elementos, incumpliendo los supuestos del modelo de fiabilidad; lo anterior era de esperarse, teniendo en cuenta las grandes diferencias entre las unidades de las variables evaluadas. En adelante el análisis se realiza sobre las variables estandarizadas.

TABLA III
ESTADÍSTICAS DE FIABILIDAD

Alfa de Cronbach	Alfa de Cronbach basada en elementos estandarizados	Número de elementos
0.252	0.785	29

Continuando con el análisis factorial, es importante revisar la significancia de la correlación de las variables del estudio con respecto a la demanda de energía en la Región Caribe. Se encuentran valores muy pequeños de la *Demanda nacional no atendida*, el *PIB*, *días laborales del mes*, *días sábados del mes* y *días festivos del mes*. Estas últimas tres variables no solo no tenían correlaciones significativas con la demanda de la Región Caribe, sino con la mayoría de las variables en estudio. Para los casos de la *Demanda nacional no atendida* y el *PIB*, solo se encontraron correlaciones con las variables económicas y no con otro tipo de variables; de esta manera, estas variables son excluidas del análisis.

Ahora bien, es necesario determinar si las variables en estudio poseen las condiciones mínimas para un análisis de tipo factorial. Para ello, se extraen aquellos valores con alta correlación (valores por encima de 0.3) con el fin de calcular los valores de *KMO* (*Medida Kaise-Meyer-Olkin de adecuación de muestreo*) y *Bartlett*. En la Tabla IV se muestra el *KMO* calculado para las variables seleccionadas, y dado que es mayor que 0.5, entonces hay correlación y por lo tanto el análisis factorial es aplicable.

TABLA IV
PRUEBA DE KMO Y BARLETT

Medida Kaise-Meyer-Olkin de adecuación de muestreo	0.793
Aprox. Chi-cuadrado	3698.895
Prueba de esfericidad de Barlett	gl 231
	Sig. .000

Por otro lado, el estadístico de *Bartlett* al rechazarse la hipótesis nula que presume independencia entre las variables, brinda la posibilidad de la utilización del análisis factorial.

Se ejecuta un análisis factorial basado en covarianzas, debido a que los datos ya están estandarizados, con opción de rotación *varimax*. Como resultado se demuestra que son necesarios cinco (5) factores para explicar en un 84.49% de la varianza. El gráfico de sedimentación mostrado en la Fig. 6 soporta estas conclusiones. Ahora bien, antes de revisar las cargas factoriales y comunalidades, es importante validar la adecuación muestral del análisis (*MSA*).

La *MSA* utilizada se soporta en la matriz de correlación anti-imagen. Este índice constituye la diagonal principal de la matriz de correlaciones anti-imagen y se utiliza como diagnóstico de la adecuación de cada variable a un modelo

factorial. Es similar al *KMO*, con la particularidad que el análisis se realiza para probar independencia por variables. Un índice *MSA* menor que 0.5 indica que la variable deber ser excluida del análisis.

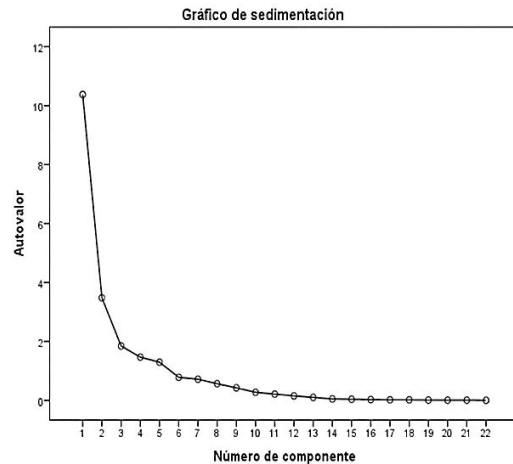


Fig. 6. Gráfico de sedimentación de las componentes.

De acuerdo con el criterio de *MSA* las variables *indicador de mes a pronosticar* y *número de días en el mes a pronosticar* deben ser excluidas del análisis. De igual manera, se revisan los valores de la correlación residual, resultante de restar a la correlación observada a la correlación reproducida. Esta medida es importante, porque valores muy pequeños cercanos a cero, indican que el modelo es correcto.

Las Tablas V y VI muestran las comunalidades y las cargas factoriales, respectivamente. Es importante destacar, que al excluir las variables antes mencionadas (*indicador de mes a pronosticar* y *número de días en el mes a pronosticar*) se generan tan solo cuatro (4) factores y no cinco (5) como resulta al incluirlas, debido a que estas dos corresponden a un único (1) factor. Teniendo en cuenta las cargas factoriales se agrupan de la siguiente manera:

1. Factor económico externo: Precio promedio del dólar, precio dólar del último día, índice de precios de combustibles, precio del petróleo Brent, precio del crudo, precio del platino.
2. Factor demanda de energía y producción industrial interna: Demanda nacional de energía, promedio de demanda nacional de los últimos cuatro (4) meses, demanda de energía de la Región Caribe, promedio de demanda Región Caribe de los últimos cuatro (4) meses, exportaciones de energía, IMACO, población, precio del oro y precio de la plata.
3. Factor de producción: Generación hidráulica, generación térmica y precio de bolsa de energía.
4. Indicador estacional: Indicador de mes y días del mes a pronosticar (factor excluido con base en *MSA*).
5. Impacto climático: Importaciones de energía y fenómenos del Niño y de la Niña.

TABLA V
COMUNALIDADES POR EL MÉTODO DE EXTRACCIÓN DE COMPONENTES PRINCIPALES

Variables	Puro		Reescalado	
	Inicial	Extracción	Inicial	Extracción
Indicador número del mes a pronosticar.	1.015	.644	1.000	.634
Demanda nacional mensual de energía.	.988	.938	1.000	.950
Promedio mensual de demanda de energía nacional en los últimos cuatro (4) meses.	1.001	.912	1.000	.911
Demanda mensual de energía en la Región Caribe.	.983	.910	1.000	.925
Promedio mensual de demanda de energía en la Región Caribe en los últimos cuatro (4) meses.	.996	.929	1.000	.932
Exportaciones Energía	.901	.656	1.000	.728
Importaciones Energía	1.000	.803	1.000	.803
Generación Hidráulica	1.000	.939	1.000	.939
Generación Térmica	1.031	.962	1.000	.933
Precio promedio del dólar.	.998	.846	1.000	.848
Precio dólar último día del mes.	.995	.811	1.000	.815
IMACO	1.026	.685	1.000	.668
Índice precios combustibles	.990	.940	1.000	.950
Precio petróleo Brent	.992	.904	1.000	.911
Precio petróleo crudo	.992	.880	1.000	.887
Población de la Región Caribe Colombiana.	.993	.969	1.000	.976
Precio del oro.	1.005	.949	1.000	.945
Precio de la plata.	1.002	.884	1.000	.882
Precio del platino.	.946	.652	1.000	.689
Precio de la bolsa de energía.	1.028	.876	1.000	.853
Fenómeno del Niño	.947	.614	1.000	.649
Número de días en el mes a pronosticar.	1.019	.757	1.000	.743

TABLA VI
CARGAS FACTORIALES DE LAS VARIABLES ECONÓMICAS

Variables	Puro				
	1	2	3	4	5
Indicador número del mes a pronosticar.	.014	.141	-.001	.734	.293
Demanda nacional mensual de energía.	.462	.664	.132	.492	-.154
Promedio mensual de demanda de energía nacional en los últimos cuatro (4) meses.	.502	.789	.153	.109	-.038
Demanda mensual de energía en la Región Caribe.	.477	.698	.203	.372	-.126
Promedio mensual de demanda de energía en la Región Caribe en los últimos cuatro (4) meses.	.477	.793	.227	.143	-.010
Exportaciones Energía	-.030	.795	-.069	-.035	.127
Importaciones Energía	.417	.118	.102	.052	-.776
Generación Hidráulica	.230	.315	-.854	.229	-.068
Generación Térmica	.045	.179	.960	.073	.035
Precio promedio del dólar.	.770	-.412	-.123	-.076	.249
Precio dólar último día del mes.	.759	-.401	-.127	-.067	.233
IMACO	.249	-.613	-.397	.046	.297
Índice precios combustibles	.926	.249	-.102	.058	-.082
Precio petróleo Brent	.919	.148	-.169	.072	.067
Precio petróleo crudo	.918	.093	-.145	.072	.042
Población de la Región Caribe Colombiana.	.525	.812	.114	-.029	-.142
Precio del oro.	.504	.833	.032	.026	-.004
Precio de la plata.	.651	.658	.092	.021	.133
Precio del platino.	.793	.126	.048	-.041	-.057
Precio de la bolsa de energía.	-.052	.353	.856	.046	-.122
Fenómeno del Niño	.202	.105	.502	.057	.555
Número de días en el mes a pronosticar.	.029	-.056	-.058	.832	-.242

C. Identificación de la Estructura de las Redes Neuronales Para los Datos de Demanda de Energía Eléctrica.

De acuerdo con los resultados de las subsecciones A y B con respecto a la demanda de energía eléctrica y las variables exógenas analizadas, es posible plantear la estructura del modelo de predicción presentado en la Fig. 7.

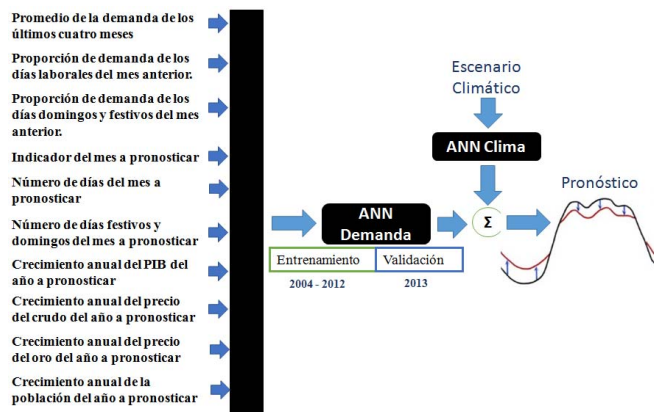


Fig. 7. Modelo de predicción de la demanda de energía eléctrica mensual.

Basándose en los resultados del análisis expuesto en el inciso anterior se plantea un segundo modelo que contempla las entradas mostradas en la Fig. 7. A continuación se exponen las razones de la selección de algunas de las variables anteriormente mostradas como entradas al modelo propuesto.

- Entrada 1: Promedio mensual de demanda de energía en la Región Caribe en los últimos cuatro (4) meses. A través del análisis de autocorrelación de la demanda es posible determinar que el mayor grado de correlación se presenta en los cuatro (4) meses inmediatamente anteriores al actual.
- Entrada 2: Proporción de demanda de los días laborales. Métrica referida a la demanda de energía presente de acuerdo al tipo de día por mes. Lo anterior brinda información de la cantidad de energía consumida durante los días hábiles del mes anterior.
- Entrada 3: Proporción de demanda de los días domingos y festivos. Brinda información de la cantidad de energía consumida durante los días domingos y festivos del mes anterior.
- Entrada 4: Indicador del mes a pronosticar. Referencia numérica del mes a pronosticar.
- Entrada 5: Número de días del mes a pronosticar.
- Entrada 6: Número de días festivos y domingos del mes a

pronosticar.

- Entrada 7: Crecimiento anual del PIB. Índice de crecimiento en la economía del país, este valor también es utilizado como un indicio de la tendencia de la economía.
- Entrada 8: Crecimiento anual del precio del crudo. El precio del crudo ejerce influencia sobre el costo de generación de la energía eléctrica.
- Entrada 9: Crecimiento anual del precio del oro. Como un metal precioso, se utiliza para indicar el poder adquisitivo de la moneda nacional.
- Entrada 10: Crecimiento anual de la población. Se basa en la premisa de correlación de la población y un mayor consumo de energía eléctrica debido al incremento de unidades habitacionales.

Tal y como se muestra en la Fig. 7, la propuesta aquí planteada contempla la incorporación de una red neuronal entrenada para integrar la sensibilidad climática generada por la afectación directa de los fenómenos climáticos como el Niño o la Niña. La Fig. 7 muestra el ajuste proporcionado por la red neuronal en consideración del efecto generado por cada tipo de fenómeno dentro de la demanda de energía eléctrica. La red neuronal utilizada tiene como entrada la desviación de demanda normalizada previamente tipificada de acuerdo con las transiciones de un escenario climático a otro, i.e., Normal a Niño, Normal a Niña, Niño a Niña, Niña a Niño, Niño a Normal y Niña a Normal.

D. Estimación de los Parámetros de la Red Neuronal para la Predicción de la Demanda de Energía

Dado que no existe una teoría formal que permita garantizar la selección de parámetros y configuración óptimos de las redes neuronales que brinden un máximo nivel de generalización para fines de pronóstico, se propone una metodología de diseño experimental. La metodología planteada en [33] evidenció como factor significativo dentro de la configuración de las redes neuronales el número de neuronas en la capa oculta.

Por lo tanto, para llevar a cabo la selección de la red neuronal adecuada, se aplicó el algoritmo de entrenamiento propuesto en [33]; en donde se utilizó un *script* desarrollado en MATLAB que busca modificar parámetros de la red neuronal recorriendo una amplia zona de experimentación para evaluar el MAPE y seleccionar aquella configuración con mejor desempeño. La red neuronal seleccionada para la predicción de la demanda cuenta con 10 entradas y 29 neuronas en la capa oculta, mientras que la red neuronal utilizada para el ajuste de la curva de demanda en función del clima tiene 1 entrada y 26 neuronas en la capa oculta.

IV. RESULTADOS

Para llevar a cabo la validación de la metodología aquí propuesta se utilizó la información de la demanda de energía eléctrica en la Región Caribe Colombiana durante los años 2015, 2016 y 2017. Con el objetivo de evaluar el desempeño del modelo propuesto se utilizó el MAPE como métrica de comparación, debido a que en la literatura es de común uso

para medir el nivel de ajuste y generalización [33]. Los resultados aquí analizados corresponden a la implementación del modelo durante los años 2015, 2016 y 2017.

La Fig. 8 y 9 muestran respectivamente las curvas de demanda de energía eléctrica durante el año 2015 y 2016 en dos escenarios.

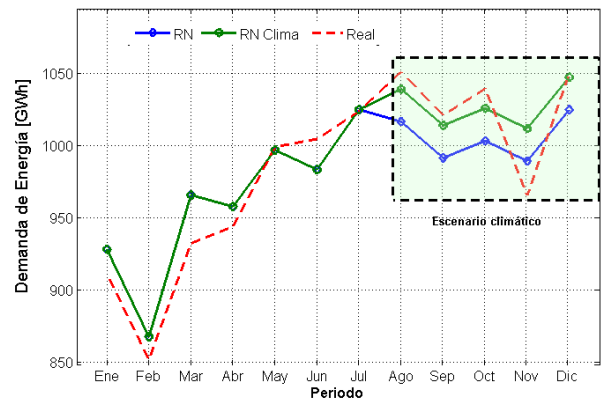


Fig. 8. Demanda de energía prevista para el año 2015.

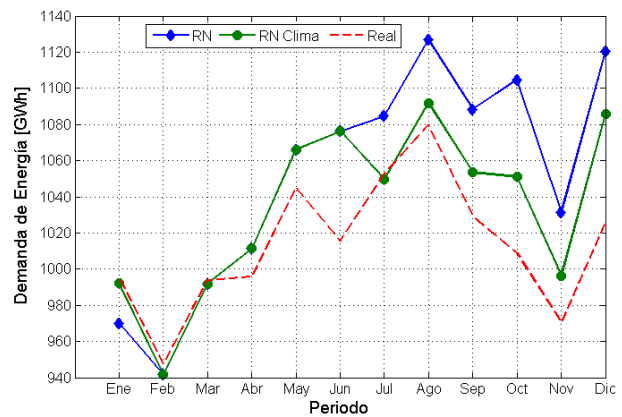


Fig. 9. Demanda de energía prevista para el año 2016.

En primera instancia se presenta la predicción obtenida por el modelo basado únicamente en la red neuronal (*RN*) sin la sensibilidad climática (línea azul). En segunda instancia se muestran los resultados del modelo completo (línea verde), el cual incorpora una red neuronal (*RN Clima*) entrenada para ajustar la predicción de la demanda de energía con el fin de incorporar sensibilidad acorde con algunos fenómenos como el Niño o la Niña. De acuerdo a lo descrito en [34], durante el año 2015 el fenómeno del niño se presentó con mayor fuerza a partir del mes de julio hasta abril de 2016. Bajo esta consideración se observa en la Tabla VII una mejoría considerable cuando se incluye el aporte de la red mostrada en la Fig. 7.

En la Tabla VII se muestra como la inclusión de un escenario climático permite ajustar la sensibilidad del modelo ocasionando una reducción del error de pronóstico. Para el año 2015 y 2016 se logró una mejoría en el MAPE igual a 0.3% y 2%, respectivamente. En términos generales el planteamiento de escenarios climáticos (fenómenos del Niño y de la Niña) permite alcanzar un rendimiento promedio igual a 1.8%.

TABLA VII
DESEMPEÑO DE LA METODOLOGÍA PROPUESTA

Año	MAPE	
	RN	RN Clima
2015	2,1%	1,8%
2016	4,3%	2,3%
2017	2,2%	1,6%
Total	3,1%	1,8%

En la Fig. 10 se muestra el desempeño de la metodología propuesta durante lo transcurrido del año 2017. El desempeño del modelo basado en el escenario climático presentó un MAPE igual a 1.6% mientras que el modelo que no contempló ningún fenómeno climático tuvo un desempeño igual a 2.2%.

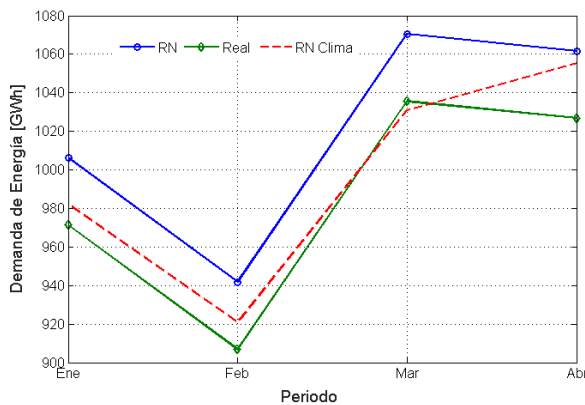


Fig. 10. Demanda de energía prevista para el año 2017.

V. CONCLUSIONES

En este trabajo se ha propuesto una metodología para el desarrollo de un modelo de predicción de la demanda de energía mensual acorde con las características climáticas de la región. A diferencia de lo presentado en el estado del arte, la propuesta aquí planteada integra técnicas estadísticas de análisis multivariado tales como el análisis factorial con el fin de caracterizar la serie de demanda de energía y su correlación con variables exógenas como el clima e indicadores económicos. Como resultado se obtuvieron dos modelos basados en redes neuronales para modelar las características socio-económicas de la serie y otro para ajustar la sensibilidad del pronóstico basado en los escenarios climáticos esperados (Niño o Niña). Por otro lado se llevó a cabo un diseño de experimentos para realizar el entrenamiento de las redes neuronales utilizadas para el pronóstico de la demanda de energía y el ajuste climático [33].

La metodología propuesta permitió establecer como variables significativas aquellas que permitieron obtener un nivel de fiabilidad por encima de 0.8. Adicionalmente, el análisis factorial resultó adecuado para integrar aquellas variables con alto grado de correlación entre ellas, permitiendo pasar de 29 variables a 10 variables a considerar dentro del modelo aquí propuesto.

La inclusión de la sensibilidad al clima permitió un ajuste dentro de la curva de demanda de energía propuesta por la red neuronal presentada en la Fig. 7. Los resultados mostrados en la Tabla VII muestran un mejor desempeño para la red

neuronal con sensibilidad climática durante los años 2015, 2016 y 2017.

El MAPE mostrado por ambos modelos evidenció una mejoría del 1.3% de la red neuronal basada en sensibilidad climática con respecto a la red neuronal sin este ajuste.

REFERENCIAS

- [1] Comisión de Regulación de Energía y Gas, *Reglamento de Comercialización de Energía Eléctrica*. Colombia, 2014.
- [2] R. Ferreira, P. H. Corredor, H. Rudnick, X. Cifuentes, and L. Barroso, "Electrical expansion in South America: Centralized or distributed generation for Brazil and Colombia," *IEEE Power Energy Mag.*, vol. 17, no. 2, pp. 50–60, 2019.
- [3] J. D. Velasquez, I. Dyrer, and C. J. Franco, "Modeling the Effect of Macroeconomic Events on River Inflows in the Colombian Electricity Market," *IEEE Lat. Am. Trans.*, vol. 14, no. 10, pp. 4287–4292, 2016.
- [4] E. E. El-Attar, J. Y. Goulermas, and Q. H. Wu, "Forecasting electric daily peak load based on local prediction," in *Power & Energy Society General Meeting, 2009. PES'09. IEEE*, 2009, pp. 1–6.
- [5] T. E. Del Carpio Huayllas and D. S. Ramos, "Electric power forecasting methodologies of some South American countries: A comparative analysis," *IEEE Lat. Am. Trans.*, vol. 8, no. 5, pp. 519–525, 2010.
- [6] Comisión de Regulación de Energía y Gas, *Resolución 020 de 1996*. Colombia, 1996.
- [7] M. Beccali, M. Cellura, V. Lo Brano, and A. Marvuglia, "Forecasting daily urban electric load profiles using artificial neural networks," *Energy Convers. Manag.*, vol. 45, pp. 2879–2900, 2004.
- [8] J. Olaya, "Un modelo spline para el pronóstico de la demanda de energía eléctrica," *Rev. Colomb. Estadística*, no. 2, pp. 187–202, 2007.
- [9] OLANREWAJU, "Energy Efficiency Assessment Using Artificial Neural Network Combined with Data Envelopment Analysis," 2011.
- [10] H. S. G. M. Rastad, and J. Nazarzadeh, "A Hybrid Nonlinear Model for the Annual Maximum Simultaneous Electric Power Demand," vol. 21, no. 3, pp. 1069–1078, 2006.
- [11] M. M. Dalvand, S. Bahram, Z. Azami, and H. Tarimoradi, "Long-term Load Forecasting of Iranian Power Grid Using Fuzzy and Artificial Neural Networks," in *Universities Power Engineering Conference, 2008.*, 2011.
- [12] S. Huang, S. Member, and K. Shih, "Short-Term Load Forecasting Via ARMA Model Identification Including Non-Gaussian," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 673–679, 2003.
- [13] S. Fan and L. Chen, "Short-Term Load Forecasting Based on an Adaptive Hybrid Method," *IEEE Trans. Power Syst.*, vol. 21, no. 1, pp. 392–401, 2006.
- [14] Gabriel I. S. Ruas, "Electrical Energy Demand Prediction Using Artificial Neural Networks and Support Vector Regression," 2008.
- [15] R.-H. Liang and C.-C. Cheng, "Combined regression-fuzzy approach for short-term load forecasting," *IEE Proc. - Gener. Transm. Distrib.*, vol. 147, no. 4, p. 261, 2000.
- [16] P. Pai and W. Hong, "Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms," *Electr. Power Syst. Res.*, vol. 74, pp. 417–425, 2005.
- [17] K. Song, Y. Baek, D. H. Hong, and G. Jang, "Short-Term Load Forecasting for the Holidays Using," vol. 20, no. 1, pp. 96–101, 2005.
- [18] K.-B. Song, S.-K. Ha, J.-W. Park, D.-J. Kweon, and K.-H. Kim, "Hybrid Load Forecasting Method With Analysis of Temperature Sensitivities," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 869–876, 2006.
- [19] X. Wang and M. Meng, "Forecasting Electricity Demand Grey-Markov Model," no. July, pp. 12–15, 2008.
- [20] T. Yalcinoz and U. Eminoglu, "Short term and medium term power distribution load forecasting by neural networks," *Energy Convers. Manag.*, vol. 46, pp. 1393–1405, 2005.
- [21] J. W. Taylor, L. M. de Menezes, and P. E. McSharry, "A comparison of univariate methods for forecasting electricity demand up to a day ahead," *Int. J. Forecast.*, vol. 22, no. 1, pp. 1–16, Jan. 2006.
- [22] T. Senjyu, H. Takara, K. Uezato, and T. Funabashi, "One-Hour-Ahead Load Forecasting," *Trans. Eng. Comput.*, vol. 17, no. 1, pp. 113–118, 2002.
- [23] G. Oğcu, O. F. Demirel, and S. Zaim, "Forecasting Electricity Consumption with Neural Networks and Support Vector Regression,"

- Procedia - Soc. Behav. Sci.*, vol. 58, pp. 1576–1585, Oct. 2012.
- [24] L. C. Moreira de Andrade and I. Nunes da Silva, "Very Short-Term Load Forecasting Based on ARIMA Model and Intelligent Systems," *2009 15th Int. Conf. Intell. Syst. Appl. to Power Syst.*, pp. 1–6, Nov. 2009.
- [25] C. Pacheco et al., "Electric load forecasting: evaluating the novel hierarchical neuro-fuzzy BSP model," vol. 26, pp. 131–142, 2004.
- [26] G. A. Darbellay and M. Slama, "Forecasting the short-term demand for electricity Do neural networks stand a better chance?," vol. 16, pp. 71–83, 2000.
- [27] D. C. Park, M. A. El-Sharkawi, R. J. Marks, L. E. Atlas, and M. J. Damborg, "Electric Load Forecasting Using an Artificial Neural Network," *IEEE Trans. Power Syst.*, 1991.
- [28] T. G. Barbounis, J. B. Theocharis, M. C. Alexiadis, and P. S. Dokopoulos, "Long-term Wind Speed and Power Forecasting Using Local Recurrent Neural Network Models," in *IEEE Transactions on Energy Conversion*, 2006, pp. 273–284.
- [29] A. A. Azadeh, S. F. Ghaderi, and S. Sohrabkhani, "A simulated-based neural network algorithm for forecasting electrical energy consumption in Iran," *Energy Policy*, vol. 36, pp. 2637–2644, 2008.
- [30] K. Nagasaka and M. Al Mamun, "Long-term Peak Demand Prediction of 9 Japanese Power Utilities Using Radial Basis Function Networks," pp. 1–8, 2004.
- [31] A. Miguel, E. Gonza, and D. Carmona-ferna, "Forecasting of the electric energy demand trend and monthly fluctuation with neural networks," vol. 52, pp. 336–343, 2007.
- [32] A. Mizutani, T. Yukawa, K. Numa, and Y. Kuze, "Improvement of Input-Output Correlations of Electric Power Load Forecasting By Scatter Search," pp. 429–433, 2005.
- [33] J. Jiménez, K. Donado, and C. G. Quintero, "A Methodology for Short-Term Load Forecasting," vol. 15, no. 3, pp. 400–407, 2017.
- [34] Climate Prediction Center Internet Team, "NOAA," *Climate Prediction Center Internet Team*, 2017. .
- [35] L. G. D. Monroy and L. Guillermo, *Estadística multivariada: Inferencia y métodos*. Universidad Nacional de Colombia, 2002.



Jamer Jimenez Mares es actualmente estudiante de Doctorado de la Universidad del Norte (UN) en Barranquilla, Colombia. Recibió su título en Ingeniería Electrónica en el 2009 y su Maestría en Ingeniería Electrónica (Cum Laude) con énfasis en eficiencia energética y sistemas inteligentes en UN en el 2013. Sus intereses en investigación incluyen;

1) Análisis de datos – metodologías y técnicas que permiten extraer información útil de bases de datos con ruido. – 2) Aprendizaje de máquina –Diseño de algoritmos meta heurísticos. Motores de inferencia para propósitos generales (redes neuronales, algoritmos genéticos, lógica difusa). 3) Eficiencia energética – Técnicas de gestión de la demanda para alcanzar ahorros de energía en edificaciones. – 4) Modelado de series de tiempo – Modelos paramétricos y no paramétricos (modelos ARIMA, k-vecinos, técnicas de soft-computing).



Abimael J. Pertuz Beltrán es actualmente estudiante de Maestría en Ingeniería Industrial de la Universidad del Norte (UN) en Barranquilla, Colombia. Recibió su título en Ingeniería Industrial en el 2014 en la misma Universidad. Sus intereses en investigación incluyen;

1) Diseño de sistemas de información inteligente para la apoyo a tratamientos en salud 2) Análisis de datos – metodologías y técnicas de diseño y análisis estadístico factorial. – 3) Modelado de series de

tiempo – Modelos paramétricos y no paramétricos (modelos ARIMA, k-vecinos). 4) Análisis financiero y proyecciones económicas basadas en datos históricos.



Christian G. Quintero M. es profesor de tiempo completo del Departamento de Ingeniería Eléctrica y Electrónica de la Universidad del Norte (UN) en Barranquilla, Colombia. Recibió su título de Doctorado (Cum Laude) en el programa doctoral de Tecnologías de la Información de la Universitat de Girona (UdG), España en el 2007. Recibió su título en Ingeniería Electrónica (Cum Laude) de la Universidad Industrial de Santander, Colombia en el 2001 y su título de Maestría en Tecnologías de la Información de la UdG en el 2005. Actualmente, es miembro del Grupo de Investigación en Robótica y Sistemas Inteligentes y profesor de Control Automático y Diseño de Sistemas Inteligentes en UN. Consejero y evaluador del programa nacional de Electrónica, Telecomunicaciones e Informática de Colciencias. Sus intereses en investigación y docencia incluyen el desarrollo de planteamientos de inteligencia computacional y control automático aplicados conjuntamente en varios dominios.



Johny Montaña es profesor de tiempo completo del Departamento de Ingeniería Eléctrica de la Universidad Técnica Federico Santa María en Valparaíso, Chile desde 2014. Recibió su título en Ingeniería Eléctrica en el año 1999, su título de Maestría en Alta Tensión en 2002 y su título de Doctorado en Ingeniería Eléctrica

en 2006 en la Universidad Nacional de Colombia. Fue asistente de investigación en el grupo PAAS de la Universidad Nacional de Colombia entre el 2000 y el 2005, fue Ingeniero de Diseño Jr. en Siemens SA entre el 2006 y el 2009 y fue profesor tiempo completo en la Universidad del Norte en Barranquilla, Colombia entre el 2010 y el 2013. Sus intereses en investigación y docencia incluyen sistemas de protección contra rayos, sistemas de localización de rayos, sistemas de puesta a tierra y redes inteligentes.