Evaluación del Efecto de la Variación de la Temperatura Regional en la Demanda de Energía Eléctrica Desagregada por Áreas Operativas

Cristian Alberto Gómez Olarte

Universidad Nacional de Colombia
Facultad Nacional de Minas
Departamento de Ciencias de la Computación y la Decisión
Programa de Maestría en Ingeniería – Sistemas Energéticos
Medellín, Colombia
Noviembre de 2016

Evaluación del Efecto de la Variación de la Temperatura Regional en la Demanda de Energía Eléctrica Desagregada por Áreas Operativas

Cristian Alberto Gómez Olarte

Trabajo final de maestría presentado como requisito parcial para optar al título de:
Magíster en Ingeniería – Sistemas Energéticos

Director de Tesis
Yris Olaya Morales

Línea de Investigación:
Mercados de Energía
Grupo de Investigación:
Ciencias de la Decisión

Universidad Nacional de Colombia
Facultad Nacional de Minas
Departamento de Ciencias de la Computación y la Decisión
Medellín, Colombia
Noviembre de 2016

A Dios, mi esposa, mi hijo, mi madre y mis hermanos.

AGRADECIMIENTOS

Agradezco a la profesora Yris Olaya Morales por su orientación, paciencia y apoyo en el desarrollo del presente trabajo.

RESUMEN

La evaluación del efecto de la variación de la temperatura media del aire sobre la demanda de electricidad es un tema trascendental e importante en el sector eléctrico, ya que la temperatura podría convertirse en un subyacente sobre activos estructurados que permitan la cobertura de los agentes del mercado, en especial en un contexto hidrodependiente. La relación entre ambas variables ha sido muy investigada en particular en países con estaciones climáticas marcadas en los cuales la demanda de energía para satisfacer necesidades de enfriamiento y calefacción tiene una relación evidente con la temperatura media del aire. En países tropicales ha sido menos estudiado el vínculo entre dichas variables por las menores variaciones estacionales de la temperatura, ya que se ha demostrado, como en el caso de Colombia, que la demanda regulada de energía eléctrica está más estrechamente relacionada con el Producto Interno Bruto (PIB) y con el efecto calendario anual (número de días del mes y de días festivos, periodos vacaciones, entre otros) (Velásquez, 2008). Sin embargo, el crecimiento económico y otras variables de comportamiento pueden influir en que haya una mayor respuesta de la demanda a cambios en la temperatura, ya que al aumentar el ingreso de los hogares puede cambiar el comportamiento respecto a las temperaturas de confort térmico. Por esta razón en el presente trabajo final de maestría se pretende determinar si existe una relación entre dichas variables. Para ello, se revisan las técnicas de completamiento de series temporales microclimáticas, se construyen series de indicador de temperatura media del aire y de demanda de energía eléctrica para cada una de las áreas operativas que conforman el Sistema Interconectado Nacional (SIN) y se determina si éstas tienen alguna relación temporal. Así mismo, se propone un modelo econométrico, que no tiene propiedades predictivas, pero a través del cual se analiza la importancia de incluir las variaciones de la temperatura media del aire en el pronóstico de la demanda de electricidad en el país. Los resultados muestran que, las variaciones en las condiciones climáticas, particularmente en la temperatura, inciden en el comportamiento de la demanda de energía eléctrica. Finalmente, se presentan recomendaciones sobre posibles trabajos futuros a realizar a partir del presente trabajo final de maestría.

Palabras clave: Demanda de energía eléctrica, temperatura del aire, causalidad de Granger.

ABSTRACT

The evaluation of the effect of the variation of the average air temperature on the electricity demand is a key aspect in electricity sector, because the temperature could become an underlying structured asset which allows coverage of market players, especially in a hydro-dependent context. The relationship between the two variables has been investigated particularly in countries with marked seasons in which energy demand has a clear relationship with the average air temperature to satisfy heating and cooling needs. In tropical countries, the link between these variables has been less studied due to lower seasonal variations in temperature, and it has been shown that in the case of Colombia, regulated demand of electricity is more closely related to the Gross Domestic Product (GDP) and the annual calendar effect (number of days in the month and holidays, vacation periods and so on) (Velásquez, 2008). However, the economic growth and other behavior variables can influence if there was a greater demand response to changes in temperature, since increasing the household income can change the behavior regarding thermal comfort temperatures. For this reason, this master's degree final project aims to determine whether there is a relationship between these variables. To do this, the techniques of completion of microclimatic time series are revised and indicator series for average air temperature and electricity demand indicators are calculated for each of the operational areas that make up the National Interconnected System (NIS) and it determines if they have any temporal relationship. Also, an econometric model is proposed, which has no predictive properties, but through which the importance of including variations in the average air temperature is analyzed to build a forecast for electricity demand in the country. The results show that changes in weather conditions, particularly in temperature, affect the behavior of the electricity demand. Finally, recommendations on possible future to be made are presented, taking into account this master's degree final project.

Palabras clave: Electricity demand, air temperature, Granger causality.

CONTENIDO

1. Introducción 1
1.1 Necesidad del estudio de la demanda de electricidad y la temperatura 1
1.2 Experiencias en el relleno de datos faltantes en series climáticas 1
1.3 Experiencias en el análisis de relación entre temperatura y demanda de electricidad 2
1.4 Limitantes del análisis de relación entre demanda de electricidad y temperatura 5
1.4.1 La transformación de las series de demanda por áreas operativas 5
1.4.2 Las falencias de las series de temperatura en Colombia 6
1.5 Objetivo general y específicos 6
1.5.1 Objetivo general 7
1.5.2 Objetivos específicos 7
1.6 Alcance de la investigación 7
1.7 Aportes y contribuciones 8
1.8 Organización del documento 8
2. Materiales y métodos 9
2.1 Materiales 9
2.1.1 Temperatura media del aire 9
2.1.2 Demanda de energía eléctrica 11
2.2 Métodos 12
2.2.1 Protocolos de completamiento y transformación de las series 12
2.2.1.1 Temperatura media del aire 12
2.2.1.2 Demanda de energía eléctrica 18
2.2.2 Propiedades estadísticas de las series 19
2.2.3 Propiedades de las series de tiempo 20
2.2.4 Análisis de estacionariedad de las series 20
2.2.4.1 Correlogramas de las series 20
2.2.4.2 Test Durbin-Watson 21
2.2.4.3 Test Dickey-Fuller Aumentado 22
2.2.4.4 Test Phillips-Perrón 23
2.2.5 Prueba de causalidad de Granger 23
2.2.6 Prueba de variables irrelevantes 23
2.2.7 Prueba de correlación serial de los residuos 24
2.2.8 Análisis de los residuos 24

2.2.9 Estimación de la Volatilidad 24
2.2.9.1 Modelo GARCH 24
2.2.9.2 Modelo EGARCH 24
3. Pre procesamiento: series de temperatura 25
3.1 La temperatura media del aire 25
3.1.1 Control de calidad 25
3.1.2 Pruebas de homogeneidad 25
3.1.3 Análisis de correlación lineal 27
3.1.4 Significancia estadística 28
3.1.5 Completamiento de las series 28
3.1.6 Indicador de temperatura por área operativa 30
4. Pre procesamiento: demanda de electricidad 33
4.1 La demanda de energía eléctrica 33
4.1.1 Demanda comercial y demanda regulada 35
4.1.2 Demanda comercial, regulada y no regulada por área operativa 35
4.2 Eliminación de efectos económicos en las series de demanda 37
5. Resultados 39
5.1 Estadística descriptiva de las series 39
5.2 Análisis de estacionariedad 43
5.2.1 Correlogramas 43
5.2.2 Test Dickey-Fuller Aumentado 44
5.2.3 Test Phillips-Perron 45
5.2.4 Discusión sobre las pruebas de raíz unitaria 46
5.3 Test de causalidad de Granger 46
5.4 Modelos de regresión 48
5.4.1 Análisis estructural de los modelos 51
5.5 Conclusiones 64
5.5.1 Objetivo 1 65
5.5.2 Objetivo 2 65
5.5.3 Objetivo 3 66
5.6 Trabajo futuro 66
6. Referencias 68

LISTA DE TABLAS

Tabla 1.1 Investigaciones sobre completamiento de series microclimáticas. 2
Tabla 1.2 Investigaciones sobre la relación entre temperatura y la demanda de electricidad. 3
Tabla 2.1 Número de estaciones según rango de completamiento. 9
Tabla 2.2 Número de estaciones según rango de completamiento y por área operativa. 10
Tabla 3.1 Número de estaciones según coeficiente de determinación. 27
Tabla 3.2 Número de estaciones según rango de completamiento. 29
Tabla 3.3 Número de estaciones según rango de completamiento por área operativa. 29
Tabla 5.1 Estadísticos descriptivos de las series de indicador de temperatura media del aire. 39
Tabla 5.2 Estadísticos descriptivos de las series de demanda de energía eléctrica. 41
Tabla 5.3 Matriz de coeficientes de correlación entre las series en estudio. 46
Tabla 5.4 Resultados de la prueba de causalidad de Granger. 47
Tabla 5.5 Modelos de demanda de energía eléctrica de cada una de las áreas operativas. 49
Tabla 5.6 Variables utilizadas en los modelos de regresión. 49
Tabla 5.7 Modelos de demanda de energía eléctrica modificados. 50
Tabla 5.8 Coeficientes de regresión de los modelos modificados. 52
Tabla 5.9 Modelos de volatilidad de las áreas operativas. Excepto Centro. 52
Tabla 5.10 Modelos de volatilidad del área operativa Centro. 53
Tabla 5.11 Resumen de los modelos de volatilidad. 53
Tabla 5.12 Parámetros estadísticos de los modelos de cada área operativa. 54
Tabla 5.13 Correlograma de los residuos. Antioquia. 55
Tabla 5.14 Correlograma de los residuos. Centro. 56
Tabla 5.15 Correlograma de los residuos. Costa Atlántica. 56
Tabla 5.16 Correlograma de los residuos. CQR. 56
Tabla 5.17 Correlograma de los residuos. Oriente. 56
Tabla 5.18 Correlograma de los residuos. Sur. 57
Tabla 5.19 Correlograma de los residuos. THC. 57
Tabla 5.20 Correlograma de los residuos. Valle. 57
Tabla 5.21 Correlograma de los residuos al cuadrado. Antioquia. 58
Tabla 5.22 Correlograma de los residuos al cuadrado. Centro. 58
Tabla 5.23 Correlograma de los residuos al cuadrado. Costa Atlántica. 58
Tabla 5.24 Correlograma de los residuos al cuadrado. CQR. 58
Tabla 5.25 Correlograma de los residuos al cuadrado. Oriente 59
Tabla 5.26 Correlograma de los residuos al cuadrado. Sur. 59
Tabla 5.27 Correlograma de los residuos al cuadrado. THC 59
Tabla 5.28 Correlograma de los residuos al cuadrado. Valle. 59

LISTA DE FIGURAS

Figura 2.1 Estaciones climáticas utilizadas en el completamiento de datos de temperatura. 10
Figura 2.2 Áreas operativas del Sistema Interconectado Nacional. 11
Figura 3.1 Serie de tiempo de referencia de temperatura en Colombia. 26
Figura 3.2 Test de homogeneidad de la estación San Miguel de Sema. 26
Figura 3.3 Test de homogeneidad de la estación Apto José M Córdova. 27
Figura 3.4 Observatorios climatológicos colombianos seleccionados. 30
Figura 3.5 Serie de Indicador de Temperatura para las ocho áreas operativas (a). 31
Figura 3.6 Serie de Indicador de Temperatura para las ocho áreas operativas (b). 32
Figura 4.1 Series de demanda de energía eléctrica del SIN. 33
Figura 4.2 Series de pronóstico de demanda de energía eléctrica del SIN en cada área operativa. . 34
Figura 4.3 Series de demanda comercial y regulada de energía eléctrica en el SIN. 35
Figura 4.4 Series de demanda regulada de cada una de las áreas operativas. 35
Figura 4.5 Demanda comercial regulada corregida de cada una de las áreas operativas (a). 37
Figura 4.6 Demanda comercial regulada corregida de cada una de las áreas operativas (b). 38
Figura 5.1 Histogramas de indicador de temperatura media del aire (a). 40
Figura 5.2 Histogramas de indicador de temperatura media del aire (b). 41
Figura 5.3 Histogramas de demanda de energía eléctrica. 42
Figura 5.4 Correlograma del indicador de temperatura media del aire de Antioquia. 43
Figura 5.5 Correlograma de la demanda de energía eléctrica de Antioquia. 44
Figura 5.6 Correlograma de la demanda de energía eléctrica de la Costa Atlántica. 44
Figura 5.7 Modelo de regresión y residuales de Antioquia. 60
Figura 5.8 Modelo de regresión y residuales de Centro. 61
Figura 5.9 Modelo de regresión y residuales de Costa Atlántica. 61
Figura 5.10 Modelo de regresión y residuales de CQR. 62
Figura 5.11 Modelo de regresión y residuales de Oriente. 62
Figura 5.12 Modelo de regresión y residuales de Sur. 63
Figura 5.13 Modelo de regresión y residuales de THC. 63
Figura 5.14 Modelo de regresión y residuales de Valle. 64

1. INTRODUCCIÓN

1.1 Necesidad del estudio de la demanda de electricidad y la temperatura

La caracterización de la demanda de energía eléctrica es un proceso en el cual variables microclimáticas como la precipitación, la humedad, la nubosidad o la temperatura, son fundamentales. En el caso de la temperatura, algunos investigadores han determinado que ésta variable reúne todos los efectos del cambio en las condiciones climáticas sobre el comportamiento de la demanda de energía eléctrica.(Catarlis, Synodinou, Proedrou, & Tsanfrassoulis, 2001; Considine, 2000; Engle, Granger, Rice, & Weiss, 1986; Filippini, 1995; Henley & Peirson, 1998; Moral & Vicéns, 2003). Sin embargo, en Colombia, al no contar con estaciones climáticas marcadas, la temperatura no ha sido considerada dentro de la explicación del comportamiento de la demanda de electricidad. No obstante, con la urbanización acelerada del país han aumentado las temperaturas de las ciudades (León Aristizábal, 2000) y surge la pregunta de si podemos observar una relación entre temperatura y demanda de energía eléctrica.

En el estudio de la relación entre los cambios de la temperatura media del aire y la demanda de energía eléctrica, para el caso colombiano, se debe tener en cuenta que el territorio nacional es heterogéneo y por tanto es necesario desagregar el Sistema Interconectado Nacional en áreas que se encuentren geográficamente cercanas, y que en la medida de lo posible, posean características culturales similares. Por ello, en el presente trabajo final de maestría se analizan los efectos del cambio en la temperatura media del aire sobre la demanda desagregada de energía eléctrica por área operativa. Vale la pena resaltar que en vista de que cada área, en su interior, tiene a su vez condiciones climáticas distintas, se propone utilizar la metodología propuesta por Moral & Vicéns (2003) para la estimación de un indicador de temperatura que describa las variaciones de ésta en cada una de las áreas consideradas. Finalmente debe tenerse en cuenta que para la estimación del indicador enunciado ha de estudiarse en primer lugar la calidad de las series de temperatura, por tal motivo en la sección 1.2 se presentan las experiencias en el relleno de datos faltantes en series climáticas.

1.2 Experiencias en el relleno de datos faltantes en series climáticas

Como se indicó antes, para el estudio de series climáticas es necesario no sólo analizar la calidad de los datos de dichas series sino también determinar, en caso de ser necesario, cuáles son las técnicas de completamiento más apropiadas para rellenar lagunas y sustituir datos incoherentes. En este tópico, (Barrera, 2004), se destaca por analizar diversas técnicas de relleno de datos faltantes en series de variables microclimáticas y determinar a partir de estimaciones estadísticas cuáles de los métodos

disponibles en la literatura que mejor se ajustan a la naturaleza de cada variable meteorológica. En la **Tabla 1.1** se presentan los estudios más significativos en ésta área de investigación.

Tabla 1.1 Investigaciones sobre completamiento de series microclimáticas.

Estudio	Autores	Revista / Institución	Tipo de Investigación (Método/Modelo)	Número de Citas
SueMulador: Herramienta para la Simulación de Datos Faltantes en Series Climáticas Diarias de Zonas Ecuatoriales	(Chica, Peña, Giraldo, Obando, & Riaño, 2014)	Revista Facultad Nacional de Agronomía Medellín	Aplicación (Cadena de Markov)	4
Técnicas de Completado de Series Mensuales y Aplicación al Estudio de la Influencia de la NAO en la Distribución de la Precipitación en España	(Barrera, 2004)	Universidad de Barcelona	Metodológico	3
Creating a Serially Complete, National Daily Time Series of Temperature and Precipitation for the Western United States,	(Eischeid, Pasteris, Diaz, Plantico, & Lott, 2000)	Journal of Applied Meteorology	Aplicación (Interpolación Espacial)	133
Notas Metodológicas para la Cumplimentación de Series Climáticas y Extrapolación de Datos. Su Aplicación al Mapa de Temperaturas de Andalucía	(Justicia & Dominguez, 1992)	Baetica. Estudios de Arte, Geografía e Historia.	Aplicación (Regresión Lineal)	1

Fuente: Elaboración Propia.

1.3 Experiencias en el análisis de relación entre temperatura y demanda de electricidad

El estudio de los efectos del cambio de la temperatura media del aire sobre la demanda de energía eléctrica ha sido abordado explícita o implícitamente por investigadores en Europa, Brasil y Estados Unidos. Para ello, los autores han utilizado dos técnicas de evaluación: modelos estadísticos e inteligencia artificial. El primer grupo incluye modelos basados en regresión, series de tiempo o técnicas econométricas tales como la regresión lineal simple y múltiple (Ahmed, Muttaqi, & Agalgaonkar, 2012; Chandramowli & Felder, 2014; Mirasgedis et al., 2007; Oliveira, Marzec, Bordin, Bretas, & Bernardon, 2012; Sathaye et al., 2013), la regresión multivariada (Pilli-Sihvola, Aatola, Ollikainen, & Tuomenvirta, 2010; Schaeffer et al., 2012), los modelos autorregresivos (Zachariadis, 2010) y los modelos de autocorrelación (Climent Diranzo, Valor, Torró, & Caselles Miralles, 2003). El segundo grupo está conformado por los modelos basados en programación dinámica lineal (Jaglom et al., 2014), regresión difusa y redes neuronales artificiales (Guégan, Uvo, & Madani, 2012).

En la **Tabla 1.2** se listan algunos estudios representativos publicados en torno al tópico de investigación.

Tabla 1.2 Investigaciones sobre la relación entre temperatura y la demanda de electricidad.

Estudio	Autores	Revista / Institución	Tipo de Investigación (Método/Modelo)	Número de Citas
Impact of climate change heating and cooling energy use in buildings in the United States	(Wang & Chen, 2014)	Energy and Buildings	Aplicación (GCM)	1
Impact of climate change on electricity systems and markets – A review of models and forecasts	(Chandramowli & Felder, 2014)	Sustainable Energy Technologies and Assessments	Revisión Bibliográfica	5
Impact of climate change on Taiwanese power market determined using linear complementarity model	(Tung et al., 2013)	Applied Energy	Aplicación (LCM)	2
Estimating impacts of warming temperatures on California's electricity system	(Sathaye et al., 2013)	Global Environmental Change	Metodológico	2
The impact of climate change on the European energy system	(Dowling, 2013)	Energy Policy	Aplicación (POLES)	5
Efecto del cambio climático sobre la previsión de demanda de energía eléctrica	(P. M. O. Oliveira et al., 2012)	Ingeniería Eléctrica	Aplicación (Regresión Lineal Múltiple)	0
Climate change impacts on electricity demand in the State of New South Wales, Australia	(Ahmed et al., 2012)	Applied Energy	Aplicación (Regresión Lineal Múltiple)	4
Global warming and electricity demand in the rapidly growing city of Delhi: A semiparametric variable coefficient approach	(Gupta, 2012)	Energy Economics	Aplicación (Regresión de Mínimos Cuadrados)	4
Developing a module for estimating climate warming effects on hydropower pricing in California	(Guégan et al., 2012)	Energy Policy	Reseña Bibliográfica	5
Modelling the effects of climate change on the energy system—A case study of Norway	(Seljom et al., 2011)	Energy Policy	Aplicación (MARKAL)	8
Climate change effect on very short-term electric load forecasting	(M. Oliveira, Marzec, Bordin, Bretas, & Bernardon, 2011)	IEEE	Aplicación (Regresión Lineal Múltiple)	8
Electricity demand elasticities and temperature: Evidence from panel smooth transition regression with instrumental variable approach	(Lee & Chiu, 2011)	Energy Economics	Aplicación (PSTR)	31
Forecast of electricity consumption in Cyprus up to the year 2030: The potential impact of climate change	(Zachariadis, 2010)	Energy Policy	Aplicación (Econométrico: ARDL)	9
Climate change and electricity consumption—Witnessing increasing or decreasing use and costs?	(Pilli-Sihvola et al., 2010)	Energy Policy	Aplicación (Regresión Econométrica Multivariable)	16

Estudio	Autores	Revista / Institución	Tipo de Investigación (Método/Modelo)	Número de Citas
Dynamic temperature dependence patterns in future energy demand models in the context of climate change	(Hekkenberg, Moll, & Uiterkamp, 2009)	Energy	Aplicación (Regresión Lineal Múltiple)	32
The non-linear link between electricity consumption and temperature in Europe: A threshold panel approach	(Bessec & Fouquau, 2008)	Energy Economics	Aplicación (Regresión no Lineal)	96
Modeling framework for estimating impacts of climate change on electricity demand at regional level - Case of Greece	(Mirasgedis et al., 2007)	Energy Conversion and Management	Aplicación (Regresión Múltiple)	31
Análisis de la influencia de la temperatura en la demanda residencial de energía eléctrica en Andalucía	(Gutiérrez & Arévalo, 2006)	Revista de Estudios Andaluces	Aplicación - Metodológico	0
Regional Energy Demand and Adaptations to Climate Change: Methodology and application to the state of Maryland, USA	(Ruth & Lin, 2006)	Energy Policy	Aplicación (Regresión Lineal)	66
Modelos STR para el análisis de la respuesta no lineal de la demanda de electricidad a la temperatura	(Moral & Vicéns, 2005)	Instituto L.R. Klein - Centro Gauss	Aplicación (Modelos LSTR)	0
Modelización del impacto de la temperatura en la demanda residencial de energía eléctrica	(Moya, 2004)	Congreso de Ingeniería de la Organización	Aplicación (Modelos de Regresión Dinámica)	0
Climate change impacts on electricity demand	(Parkpoom, Harrison, & Bialek, 2004)	IEEE	Aplicación (Regresión Lineal Múltiple)	31
Un modelo de previsión de la demanda de energía eléctrica_ Thor II	(Moral & Vicéns, 2003)	Instituto L.R. Klein - Centro Gauss	Aplicación (Regresión Lineal Múltiple)	0
Incidencia de la climatología en el consumo de gas y electricidad en España	(Climent Diranzo et al., 2003)	Infraestructuras: Transportes e Industrias de Red	Aplicación (Modelo autorregresivo y de Medias Móviles)	4
El modelo español de consumo sectorial de electricidad	(Valor, Climent, Meneu, & Caselles, 2002)	Universidad de Valencia	Aplicación (Regresión Lineal Múltiple)	3
Temperature and seasonality influences on Spanish electricity load	(Pardo, Meneu, & Valor, 2002)	Energy Economics	Aplicación (Regresión Lineal Múltiple)	196
Daily Air Temperature and Electricity Load in Spain	(Valor, Meneu, & Caselles, 2001)	American Meteorogical Society	Aplicación (Gráfico)	169
Temperature effects on the winter daily electric load	(Bolzern & Fronza, 1982)	Journal of Applied Meteorology	Aplicación (Regresión Lineal)	10

Fuente: Elaboración Propia.

De la revisión sistemática de la literatura, ha sido posible identificar cuáles son las principales técnicas implementadas en la evaluación de la relación entre la temperatura y la demanda de electricidad, así como también cuál es la información primaria y secundaria que se requiere para llevar a cabo el análisis. Dentro de los estudios más representativos se encuentra el realizado por Climent, Valor, Torró & Caselles (2003), quienes determinaron a partir de la aplicación de técnicas econométricas

que existe una relación no lineal entre la demanda de energía eléctrica y la temperatura media del aire respecto a un valor medio de referencia, a partir del cual a medida que la temperatura disminuye o aumenta, la demanda de electricidad incrementa. Estos autores concluyeron además que la demanda doméstica o regulada es la única que devela relación con la temperatura ambiente debido a que la no regulada o industrial solo depende de la laboralidad o dinámica operativa de las empresas industriales. Por otra parte, Moral & Vicéns (2003) analizan el modelo econométrico de la demanda de energía eléctrica español y determinan que la inclusión de variables como la laboralidad, la actividad económica y la temperatura en el pronóstico de la demanda de electricidad, permitirían realizar una predicción más aproximada de la demanda real de energía eléctrica en España. Así mismo, Moya (2004) propone un modelo de regresión dinámica en el cual demuestra que el consumo de energía eléctrica español se ve influenciado por variaciones en la temperatura. Es importante aclarar que en el presente trabajo final de maestría no se desea estimar demanda de energía eléctrica, sino determinar si existe o no una relación entre temperatura media del aire y demanda de electricidad en el sector residencial. No obstante, dado que en caso de tomar la temperatura y la demanda con rezagos como únicas variables de análisis podría generarse un desequilibrio en la estimación, se incorpora el efecto día o laboralidad a través de variables dummy que permiten acotar el comportamiento de la demanda de electricidad en las distintas áreas operativas.

En resumen, el estudio de los efectos del cambio de las condiciones climáticas sobre la demanda de energía eléctrica ha dejado en evidencia que la temperatura es una variable clave para explicar el comportamiento de demanda de electricidad y a su vez para el pronóstico de la misma. Sin embargo, existen limitantes que dificultan dicho análisis. En la sección 1.4 se presentan las limitaciones identificadas para analizar la relación entre las variables objeto de estudio para el caso colombiano.

1.4 Limitantes del análisis de relación entre demanda de electricidad y temperatura

Existen barreras que dificultan el análisis de relación entre las variaciones de la temperatura media del aire y la demanda de energía eléctrica. Estos limitantes se listan a continuación:

1.4.1 La transformación de las series de demanda por áreas operativas

Para determinar si la variación de la temperatura media del aire genera cambios sobre la demanda de energía eléctrica es conveniente limitar el objeto de estudio a ‘pequeños’ espacios geográficos y por tanto, efectuar un análisis desagregado de la demanda de electricidad como lo indican (Moral & Vicéns, 2005) en su investigación. En este contexto, resulta oportuno y necesario utilizar series temporales que reflejen la dinámica de la demanda de carga en cada una de las áreas operativas que integran el Sistema Interconectado Nacional – SIN. Sin embargo, la información disponible en el portal del operador del mercado XM Compañía de Expertos en Mercados S.A. E.S.P. no se publica desagregada y por tanto surgen los siguientes problemas:

- ¿Qué información puede ser utilizada para transformar la serie agregada de demanda de energía eléctrica?

- ¿Cómo transformar la serie agregada de demanda de energía eléctrica?
- ¿Qué supuestos deben formularse para transformar la serie agregada de demanda de energía eléctrica?
- ¿Cuáles críticas pueden originarse respecto a los supuestos formulados para transformar la serie agregada de demanda de energía eléctrica?

El protocolo para dar solución a las limitaciones anteriores se presenta en la sección 2.2.1.2.

1.4.2 Las falencias de las series de temperatura en Colombia

De acuerdo con los lineamientos de la Ley 1712 del 6 de marzo de 2014, por medio de la cual se regula el derecho de acceso a la información pública de carácter nacional, el Instituto de Hidrología, Meteorología y Estudios Ambientales IDEAM proporcionó lo registros diarios de la temperatura media del aire en el periodo comprendido entre el 01:01:2013 y el 31:12:2013 de todas las estaciones climatológicas del territorio nacional (IDEAM, 2014). Sin embargo, los registros no son continuos, es decir, tienen lagunas de información y además, en algunos casos, se observan errores de transcripción. Por lo anterior, surgen las siguientes dificultades:

- Las series temporales de temperatura no son continuas y homogéneas, ya que en el 97% de las estaciones climáticas falta entre el 1 y el 98% de los registros diarios de temperatura media del aire. (Véase 2.1.1).
- Es posible que existan errores de transcripción en los registros de temperatura media del aire.
- No hay evidencias reportadas en la literatura que indiquen cómo seleccionar un único modelo de completamiento de datos en series microclimáticas.
- Las lagunas de información en las series dificultan la aplicación de un único método de completamiento.
- Se requiere un método de completado para que los valores estimados sean realistas y así no formular conclusiones erróneas.

Teniendo en cuenta la necesidad del análisis de relación entre las variaciones en la temperatura media del aire y la demanda de energía eléctrica en Colombia descrita en la sección 1.1 y las limitantes anteriores, en la siguiente sección se definen los objetivos y el alcance de este trabajo final de maestría.

1.5 Objetivo general y específicos

A partir de la problemática presentada en la sección 1.1 y las limitaciones identificadas en la sección 1.4, se establecen el objetivo general y los objetivos específicos en el presente trabajo final de maestría:

1.5.1 Objetivo general

Evaluar la relación histórica entre la temperatura media del aire regional y la demanda desagregada de energía eléctrica por áreas operativas del SIN en Colombia, en el periodo comprendido entre 01:01:2007 y 31:12:2013.

1.5.2 Objetivos específicos

Los objetivos específicos son:

I. Identificar las fuentes de información y construir, para cada área operativa, las series de demanda de energía eléctrica desagregada y el indicador diario de temperatura media del aire.
II. Determinar el grado de relación, en cada área operativa, entre las series históricas de demanda de energía eléctrica y el indicador diario de temperatura media del aire.
III. Evaluar el impacto de las variaciones de la temperatura en la demanda de energía eléctrica en cada una de las áreas operativas que integran el Sistema Interconectado Nacional – SIN.

1.6 Alcance de la investigación

La ciencia se define como el "conjunto de conocimientos obtenidos mediante la observación y el razonamiento, sistemáticamente estructurados y de los que se deducen principios y leyes generales", y la tecnología por su parte es el conjunto de teorías y técnicas que permiten el aprovechamiento práctico del conocimiento científico (Valencia, 2012). De acuerdo a estas definiciones, este trabajo final de maestría se encuentra en la categoría de investigación tecnológica ya que aprovecha los conocimientos científicos y herramientas teóricas y tecnológicas disponibles en la literatura con objeto de dar respuesta a la pregunta emergente de investigación descrita en la sección 1.1.

Por otra parte, el alcance de una investigación indica el resultado que se espera obtener de ella. Según Hernández, Fernández & Baptista (2003), los resultados de un trabajo investigativo pueden ser exploratorios, correlacionales, descriptivos o explicativos. Los estudios exploratorios tienen como objetivo investigar un tópico poco estudiado, los correlacionales identificar si existe relación entre dos o más variables a partir de modelos empíricos o estadísticos, los descriptivos detallar las características más relevantes de un fenómeno o problema de investigación, y finalmente en los estudios explicativos se pretende explicar las causas que dan origen al fenómeno o problema, por qué ocurre y bajo qué condiciones éste tiene lugar (Hernández, Fernández, & Baptista, 2003). En este sentido, la pregunta emergente de investigación se encuentra en fase correlacional, ya que si bien en Colombia aún no se han desarrollado investigaciones dirigidas directa o indirectamente a la evaluación de la relación entre la temperatura media del aire y la demanda de energía eléctrica; en países como Alemania, Brasil, España y Estados Unidos ya se han publicado estudios metodológicos y/o casos de aplicación en dicho tópico.

Finalmente, con el desarrollo del presente trabajo de investigación se espera caracterizar la relación entre las variaciones de la temperatura media del aire y la demanda de energía eléctrica en las áreas operativas que integran el Sistema Interconectado Nacional (SIN), y por tanto, no se estudiará todo

el territorio nacional ya que no se tendrán en cuenta las zonas no interconectadas. Así mismo, como se presentó en la sección 1.4.2, por la disponibilidad de información, el periodo de tiempo de evaluación es de siete años y corresponde a los registros diarios de temperatura tomados desde el 01:01:2007 hasta el 31:12:2013.

1.7 Aportes y contribuciones

En este trabajo final de maestría se realizarán contribuciones en los siguientes temas:

I. Se evalúan las técnicas de completamiento de series microclimáticas, y de acuerdo con las características de la información disponible, se determina cuáles son los métodos más apropiados para completar las series de temperatura media del aire en Colombia.
II. Se ofrece una mayor comprensión de la influencia de las variaciones en la temperatura sobre la demanda de energía eléctrica en Colombia, a partir de la aplicación de herramientas estadísticas.

Así mismo, se espera que los resultados obtenidos sean la base de trabajos futuros en el área, por ejemplo en el pronóstico de la demanda de electricidad, ya que las conclusiones del presente trabajo final de maestría podrían ser utilizadas por investigadores y agentes del mercado como insumo para complementar sus desarrollos científicos.

1.8 Organización del documento

Este trabajo final de maestría está organizado de la siguiente manera: en el Capítulo 2 se presentan los materiales y los métodos utilizados en la investigación. En el Capítulo 3, muestran los resultados del pre procesamiento de las series de temperatura media del aire y en el Capítulo 4 de las series de demanda de energía eléctrica. Finalmente, en el Capítulo 5 se presentan los resultados y la discusión de los mismos y se exponen las conclusiones respectos a los objetivos propuestos en el presente trabajo final de maestría.

2. MATERIALES Y MÉTODOS

2.1 Materiales

Los materiales utilizados para la elaboración del presente trabajo final de maestría corresponden a la información diaria de temperatura media del aire, la cual fue proporcionada por el Instituto de Hidrología, Meteorología y Estudios Ambientales de Colombia – IDEAM entre 01:01:2007 y 31:12:2013, y a las series diarias de demanda de energía eléctrica de Colombia agregada, comercial no regulada y al pronóstico de carga eléctrica de cada una de las áreas operativas que integran el Sistema Interconectado Nacional – SIN, las cuales fueron descargadas del portal corporativo del operador del mercado XM Compañía de Expertos en Mercados S.A. E.S.P., para el mismo periodo de evaluación. Los materiales utilizados son descritos en las siguientes dos sub secciones.

2.1.1 Temperatura media del aire

Los datos diarios de temperatura media del aire utilizados para la elaboración de este trabajo corresponden a 448 series temporales de ésta variable climática registrada en igual número de observatorios meteorológicos, los cuales están localizados en los departamentos donde opera el Sistema Interconectado Nacional entre el 01:01:2007 y el 31:12:2013. Es importante notar que algunas series presentan lagunas puntuales y/o continuas de información.

Con el objetivo de caracterizar las series de temperatura media del aire, se definieron cuatro rangos de porcentaje de completamiento. En la **Tabla 2.1** se presenta el número de estaciones según el porcentaje de datos registrados en el periodo de análisis.

Tabla 2.1 Número de estaciones según rango de completamiento.

Clasificación	Rango de Completamiento Valores faltantes	No. Estaciones	% de Estaciones según Rango de Completamiento
Grupo 1	$0 < x \leq 25$	12	2,7
Grupo 2	$25 < x \leq 50$	38	8,5
Grupo 3	$50 < x \leq 75$	106	23,7
Grupo 4	$75 < x \leq 100$	292	65,2

Fuente: Elaboración propia.

Así mismo, en la **Tabla 2.2** se muestra el número de estaciones según porcentaje de completamiento y por departamento.

Tabla 2.2 Número de estaciones según rango de completamiento y por área operativa.

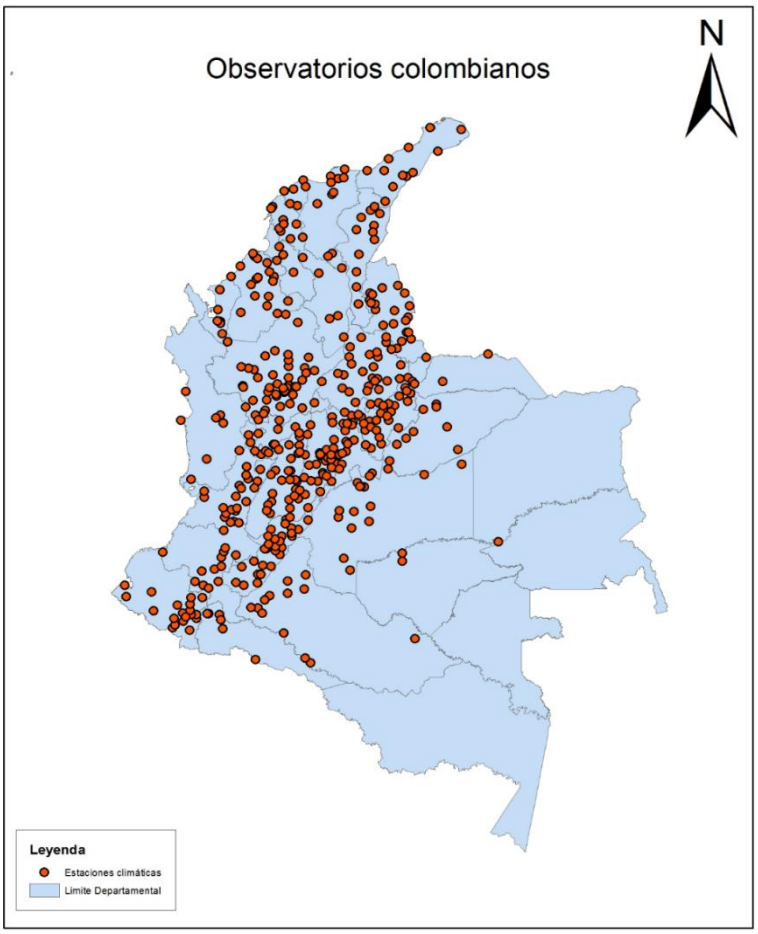
Departamento	Rango de Completamiento				Total
	$0 < x \leq 25$	$25 < x \leq 50$	$50 < x \leq 75$	$75 < x \leq 100$	
Antioquia	1	3	16	39	59
Centro	6	13	20	27	66
Costa Atlántica	1	6	24	45	76
CQR	0	1	4	9	14
Oriente	2	6	11	80	99
Sur	0	0	9	36	45
THC	2	9	17	43	71
Valle	0	0	5	13	18
Total de Estaciones	**12**	**38**	**106**	**292**	**448**

Fuente: Elaboración propia.

El porcentaje de completamiento inicial de cada una de las series de temperatura de las estaciones climáticas utilizadas se muestra en el **Anexo 8.1**.

La localización de cada una de las estaciones climáticas utilizadas en el presente trabajo final de maestría en la etapa de completamiento de datos se presenta en la **Figura 2.1.**

Figura 2.1 Estaciones climáticas utilizadas en el completamiento de datos de temperatura.



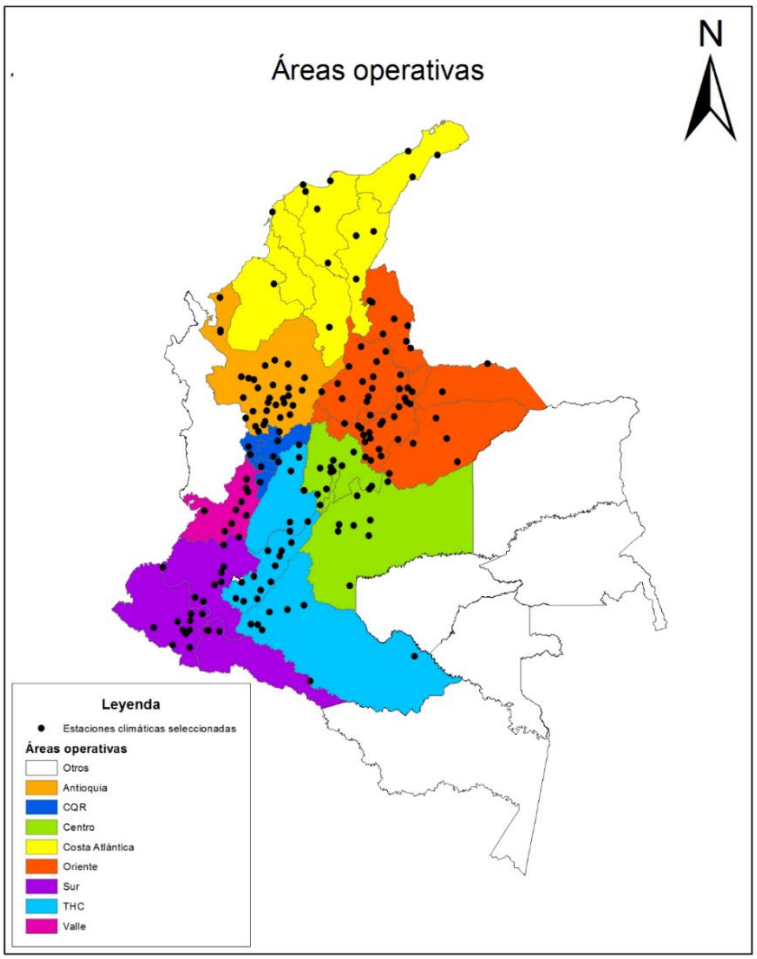
Fuente: Elaboración propia.

2.1.2 Demanda de energía eléctrica

La información base tomada de demanda de carga está compuesta por las series de demanda total del SIN, generación, demanda no atendida, exportaciones e importaciones de electricidad (XM Compañía de Expertos en Mercados S.A. E.S.P., 2015b). Sin embargo, la serie histórica de demanda agregada de energía eléctrica debe ser transformada, y por tanto, a partir de la información disponible en el portal del operador del mercado (XM Compañía de Expertos en Mercados S.A. E.S.P., 2015e), se estableció que era necesario utilizar la serie de demanda comercial no regulada (XM Compañía de Expertos en Mercados S.A. E.S.P., 2015c) y la serie del pronóstico oficial de demanda de electricidad por área operativa del SIN (XM Compañía de Expertos en Mercados S.A. E.S.P., 2015d). La distribución de las áreas operativas del SIN en el territorio nacional se muestra en la **Figura 2.2.**

Los datos de demanda mensual fueron tomados entre el 01:01:2007 y el 31:12:2013, es decir, para el mismo periodo de tiempo de la información proporcionada por el Instituto de Hidrología, Meteorología y Estudios Ambientales de Colombia – IDEAM. La resolución de la información es diaria y está completa.

Figura 2.2 Áreas operativas del Sistema Interconectado Nacional.



Fuente: Elaboración propia.

Nota: Los criterios de selección de las estaciones climáticas se presentan en el siguiente capítulo.

Los datos de demanda diaria mensual fueron tomados entre el 01:01:2007 y el 31:12:2013, es decir, para el mismo periodo de tiempo de la información proporcionada por el Instituto de Hidrología, Meteorología y Estudios Ambientales de Colombia – IDEAM. La resolución de la información es diaria y está completa.

2.2 Métodos

En esta sección se definen los protocolos o procedimientos para el completamiento y transformación de las series de temperatura media del aire y para la desagregación de la serie total de demanda de energía eléctrica del Sistema Interconectado Nacional que se van a utilizar en el presente trabajo. Los materiales descritos en la sección anterior son los elementos de trabajo en este sub capítulo. Así mismo, se presentan las herramientas teóricas a emplear para realizar la caracterización de las series de cada una de las áreas operativas.

2.2.1 Protocolos de completamiento y transformación de las series

Debido a que la información requerida para determinar si las variaciones en la temperatura media del aire generan cambios en la demanda de energía eléctrica en Colombia debe ser completada y/o transformada, a continuación se presentan los protocolos de completamiento y/o transformación de las series temporales de las variables objeto de investigación.

2.2.1.1 Temperatura media del aire

Como se indicó en la sección 1.4.2, el 97% de las series de temperatura media poseen lagunas de información en el periodo comprendido entre el 01:01:2007 y el 31:12:2013 (véase sección 2.1.1 y **Anexo 8.1**). Por tal motivo, a continuación se describe el protocolo para realizar el completamiento de las series de temperatura y de ésta manera dar solución a las dificultades planteadas en dicha sección. Dicho protocolo se basa en la tesis doctoral de Barrera (2004) titulada "Técnicas de completado de series mensuales y aplicación al estudio de la influencia de la NAO en la distribución de la precipitación en España".

a. Control de calidad

Etapa de verificación de la calidad de los registros de temperatura, en la cual se debe determinar si las series temporales tienen valores incorrectos o dudosos que deban ser corregidos o eliminados.

b. Prueba de homogeneidad

Con el objetivo de establecer si los cambios en las series de temperatura media del aire en cada una de las estaciones climáticas obedecen únicamente a cambios en las condiciones meteorológicas, se debe realizar una prueba o test de homogeneidad para descartar las series heterogéneas y de esta manera evitar que en el completamiento se utilicen series que conduzcan a conclusiones erróneas

(Barrera, 2004). Se propone el método de las dobles acumulaciones o dobles masas para verificar la homogeneidad de las series. En este método, se construye una serie de tiempo de referencia a partir de la media aritmética de cada uno de los registros diarios de las series en evaluación y posteriormente se calcula la serie acumulada para la serie de referencia y para la serie i a la cual se le quiere probar su homogeneidad (Searcy & Hardison, 1960). Se dice que una serie no es homogénea si presenta una tendencia no lineal (Barrera, 2004).

c. Análisis de correlación lineal

De acuerdo con Barrera (2004), la correlación entre datos indica si dos series temporales guardan algún tipo de relación estadística y/o dependencia entre sí. Para ello, se debe calcular el coeficiente de correlación lineal de Pearson, el cual varía entre -1 y 1 dependiendo si la correlación es completa positiva o completa negativa. Cabe resaltar que un valor del coeficiente cercano a cero indica que las series temporales no están correlacionadas linealmente. La expresión matemática utilizada para el cálculo del coeficiente de correlación de Pearson se presenta a continuación:

$$\gamma = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2}\sqrt{\sum(y_i - \bar{y})^2}}$$

Donde, γ es corresponde al coeficiente de correlación lineal de Pearson, las x_i y las y_i son los datos correspondientes, $\bar{x}$ e $\bar{y}$ son los valores medios de las series temporales.

d. Significancia estadística de la correlación de Pearson

Debido a que el coeficiente de correlación de Pearson no indica que la correlación observada sea estadísticamente significativa o que una correlación observada sea más fuerte que otra y tampoco indica que exista una relación de causalidad entre dos variables; se debe calcular la significancia estadística de la correlación. Una correlación de Pearson es significativa si se puede afirmar que esta es diferente de cero o si se puede rechazar la siguiente hipótesis nula "El coeficiente de correlación procede de una población cuya correlación es cero" (Press, Flannery, Teukolsky, & Vetterling, 1992).

Para determinar la significancia estadística de la correlación de Pearson, se puede aplicar la prueba de contraste de hipótesis basada en la distribución de la **t** de student con N-2 grados de libertad. Donde γ es el coeficiente de correlación de Pearson y N corresponde al número de pares de datos considerados en el cálculo del coeficiente de correlación de Pearson para cada pareja de series de temperatura (Press et al., 1992). Según la siguiente expresión:

$$t = \frac{\gamma}{\sqrt{\frac{1 - \gamma^2}{N - 2}}}$$

Una vez calculado el estadístico **t**, se rechaza la hipótesis nula si el valor de dicho estadístico estimado es mayor al estadístico teórico de la distribución para un nivel de confianza del 95%. En este caso la correlación obtenida no procede de una población con correlación cero, y por tanto las dos series

están correlacionadas. En caso de que el estadístico **t** calculado sea menor a su valor teórico, se acepta la hipótesis nula y las dos series temporales no están correlacionadas (Press et al., 1992).

e. Análisis de los métodos de completamiento

En la literatura se reportan diferentes métodos de completamiento de series temporales. Sin embargo, su aplicabilidad no es universal dado que no todos describen de forma adecuada la evolución o el comportamiento de una variable. En el caso de las series de temperatura, éstas como otras series de variables microclimáticas tienen características particulares propias de las condiciones bajo las cuales son registradas (Barrera, 2004). A continuación se presentan los métodos de completado de datos en series de tiempo en dos grupos: métodos univariados y métodos multivariantes de una única variable.

- **Métodos univariados**

Según Barrera (2004), los métodos univariados son aquellos en los cuales el completamiento de los vacíos de información en una serie puede realizarse con los datos de la misma serie. Estos métodos se basan en el estudio de la dependencia de una serie de datos entre sí y lo que buscan es determinar si un valor depende del registro anterior o del siguiente en una escala de tiempo dada (Barrera, 2004). Para ello, se pueden utilizar pruebas de tipo cuantitativo como la prueba de Durbin-Watson, el Contraste d_4 de Wallis, el Contraste h de Durbin, el Contrate de Breusch-Godfrey, el Contraste de Sargan y el Contraste de Box-Pierce-Ljung, entre otros (Arranz & Zamora, 2010). Barrera (2004) propone que los métodos univariados pueden ser utilizados si se determina que las series temporales presentan autocorrelación entre sí. Sin embargo, cuando las series de tiempo corresponden a valores registrados en escalas de tiempo sucesivas y a escala diaria y horaria y si la serie no sigue una caminata aleatoria, es posible suponer que un dato en un instante dado está influenciado por las observaciones en instantes anteriores (Juan, Kizys, & Manzanedo, 2006). Por lo tanto, en el presente trabajo final de maestría se supone que la autocorrelación es alta entre las series de temperatura media del aire, y por tanto, no se propone realizar pruebas de contraste de autocorrelación. Los métodos univariados utilizados en el presente trabajo se presentan a continuación:

- Criterio del dato anterior o posterior: En este método se reemplaza un dato faltante por el valor registrado antes o después del mismo. Para el caso de estudio del presente trabajo, un dato faltante se debería reemplazar por el registro de temperatura tomado el día anterior o posterior al día que carece de información. No obstante, según Barrera (2004), éste método no es apropiado para la escala mensual y diaria de la temperatura ya que el error relativo medio obtenido en su investigación en ambos casos fue superior al 10%.

- Criterio del valor medio: En este método se sustituye un registro faltante por el valor medio de la serie para un periodo de tiempo determinado. En caso de que la serie a completar sea estacionaria, genera una perturbación mínima (Rodríguez, Llasat, & Martín Vide, 1999). Sin embargo, para Barrera (2004), este método sólo es aplicable cuando el coeficiente de variación estadístico es inferior al 5%; y es útil, para el caso de la temperatura, cuando el completado de lagunas se realiza en series diarias anuales.

- **Métodos multivariantes de una única variable**

Barrera (2004) presenta algunos métodos que se utilizan para el completamiento de lagunas de información en una serie a través de los registros de otras. En el caso de la presente investigación, para el relleno de datos faltantes en una serie de temperatura a partir de las mediciones tomadas en otros observatorios climáticos. Estos métodos se basan en el estudio de la correlación lineal entre las series y la estimación de su significancia estadística. Barrera (2004) propone que una vez se determine si existe una fuerte correlación lineal y si dicha estimación es estadísticamente significativa, los datos faltantes pueden ser completados mediante uno de los métodos que se presentan a continuación:

¬ Criterio de Karl: En este método se sustituye un dato faltante por el valor registrado en otra estación que posea un alto coeficiente de correlación lineal de Pearson (Karl & Williams, 1987). Para aplicar el método las series temporales deben estar normalizadas de acuerdo con la expresión que se presenta a continuación:

$$Z_i = \frac{x_i - \overline{x_j}}{\sigma_j}$$

Donde Z_i es el valor i-ésimo normalizado; x_i es el valor i-ésimo del mes j-ésimo; $\overline{x_j}$ es el valor medio del mes j-ésimo y σ_j es la desviación estándar del mes j-ésimo (Barrera, 2004).

Barrera (2004) propone que este método es útil para completar series de temperatura en cualquier escala de tiempo, ya que la mayoría de los coeficientes de correlación estimados para los distintos pares de series fueron apreciablemente altos.

¬ Razón normal: En este método el dato faltante se calcula a partir de los registros de tres estaciones vecinas que presenten un alto coeficiente lineal de correlación de Pearson (Paulus & Kohler, 1952). La expresión matemática utilizada se muestra a continuación:

$$x(t) = \frac{1}{3}\left(\frac{\bar{x}}{\overline{x_1}}x_1(t) + \frac{\bar{x}}{\overline{x_2}}x_2(t) + \frac{\bar{x}}{\overline{x_3}}x_3(t)\right)$$

Donde $\bar{x}, \overline{x_1}, \overline{x_2}$ y $\overline{x_3}$ son las medias de las variables de la serie incompleta y de las tres series vecinas que poseen registro y que cumplen con las condiciones del método para realizar el completamiento; y $x_1(t)$, $x_2(t)$ y $x_3(t)$ son los registros de las tres series vecinas a utilizar en el completamiento. Paulus et al. (1952) plantean que este método suaviza la influencia que podría generarse por la presencia de un error en una de las tres series usadas para llevar a cabo el completamiento. Por su parte, Barrera (2004) determinó que éste método es apropiado para completar series de temperatura por la alta correlación lineal entre las series usadas en el desarrollo de su investigación.

¬ Combinación lineal ponderada: En este método el dato faltante es sustituido por los registros de series vecinas con un alto índice de correlación lineal de Pearson. En este sentido, un dato inexistente es estimado a partir de la combinación lineal ponderada de los datos de las series

que se van a utilizar en el completado con un peso proporcional al coeficiente de correlación de Pearson entre cada serie auxiliar y la serie incompleta (Peterson & Easterling, 1994). Rodríguez et al. (1999) sugieren que en el caso de la temperatura, se deben seleccionar para el completamiento series que presenten coeficientes mayores a 0,8. La expresión matemática que describe el método de la combinación lineal ponderada se presenta a continuación:

$$x(t) = \frac{\gamma_1 x_1(t) + \gamma_2 x_2(t) + \gamma_2 x_2(t) + \cdots}{\gamma_1 + \gamma_2 + \gamma_3 + \cdots}$$

Donde $x(t)$ es el dato incompleto, γ_i es el coeficiente de correlación de Pearson de la i-ésima serie usada para completar el registro inexistente y $x_i(t)$ es el valor registrado de la i-ésima serie en el instante t. Peterson et al. (1994) sugieren que el número de series que pueden ser usadas en el completamiento es arbitraria, sin embargo éste nunca debe ser inferior a dos. Para aplicar el método las series temporales deben estar normalizadas de acuerdo con la expresión utilizada en el criterio de Karl (Barrera, 2004).

Al igual que en los métodos del criterio de Karl y de la razón normal, Barrera (2004) determinó que el completamiento a través de la combinación lineal ponderada es apropiado para completar series de temperatura que posean alta correlación lineal.

- Criterio de la correlación lineal: En este método se sustituye un dato incompleto a través de la aplicación del método de mínimos cuadrados entre dos series: una incompleta que corresponde a la variable dependiente y otra completa que hace las veces de variable independiente. Este método implica que entre ambas series debe existir una alta correlación lineal, preferiblemente cercano a uno, y además han de tener una misma escala de tiempo y proximidad geográfica (Fernández, 1995). Para aplicar el método las series temporales deben estar normalizadas de acuerdo con la expresión utilizada en el criterio de Karl y en el método de combinación lineal ponderada (Barrera, 2004), y la serie completada debe presentar menor varianza que la serie original y por tanto el coeficiente de determinación de la regresión debe ser mayor o igual a uno (Fernández, 1995). Esto último se prueba mediante la siguiente expresión:

$$r^2 \leq \frac{1}{n_1 - 2} = R^2$$

Donde r^2 es el coeficiente de determinación de la regresión, n_1 es la longitud de la serie incompleta y R^2 es el coeficiente de ajuste de la regresión lineal.

Barrera (2004), determinó que este método, al igual que el criterio de Karl y la razón normal, es apropiado para completar series de temperatura que posean alta correlación lineal.

f. Criterios de selección del método de completamiento

Con el objetivo de completar las series de temperatura media del aire, a continuación se muestran los criterios que se deben tener en cuenta para el completado de datos en dichas series. Todos los criterios

parten de la base de que se debe completar la mayor cantidad de lagunas de información procurando que los valores estimados sean representativos y/o coherentes respecto a las observaciones registradas en cada una de las series de temperatura. Estos criterios se detallan a continuación:

- La sustitución de un dato incompleto por otro en una serie con un alto coeficiente de correlación lineal no garantiza que el dato sustituido sea representativo en la serie incompleta. Esto se debe a que la temperatura es una variable sensible a la altura sobre el nivel del mar y por tanto, se pueden presentar casos en los cuales se de una alta correlación lineal entre dos series de estaciones meteorológicas con condiciones climáticas diferentes.

- Los datos faltantes en una serie son más representativos si se tiene en cuenta la media histórica de la serie incompleta.

- A mayor cantidad de series utilizadas en el completamiento, se puede suavizar más la influencia de errores presentes en cada una de las series utilizadas para estimar un dato sobre la serie incompleta.

- Si dos estaciones climáticas poseen igual índice de correlación de Pearson respecto a la serie incompleta, la serie más representativa a utilizar será aquella que se encuentre más cerca de la serie con datos faltantes.

- El valor mínimo del coeficiente de correlación de Pearson debe ser superior a 0,7 (Barrera, 2004).

- Ante la presencia de grandes lagunas de información en algunas series, el completamiento debe realizarse a través de un método que utilice información de otras series temporales. En caso de que en esta fase no se completen todas las series, Barrera (2004) recomienda utilizar métodos univariados para estimar los datos que aún sigan como faltantes.

g. Cálculo del indicador de temperatura por área operativa

Debido a que en la presente investigación se pretende evaluar el efecto de las variaciones de la temperatura media del aire sobre la demanda de energía eléctrica en cada una de las áreas que integran el Sistema Interconectado Nacional – SIN. A continuación se muestra la metodología para el cálculo del indicador de temperatura en cada área operativa a partir de la propuesta de Moral et al. (2005).

- Una vez realizado el completamiento de las series de temperatura se agrupan las series que tengan el 100 % de l-s datos.

- Selección de las series de temperatura representativas de cada área operativa, según el coeficiente de variación estadístico (Departamento Administrativo Nacional de Estadística, 2008). De acuerdo con la guía del DANE (2008) se supone un valor máximo de 10 % para considerar que la serie de temperatura es representativa.

- En caso de que en un municipio dos o más estaciones cumplan con los criterios anteriores, se debe seleccionar la serie con menor coeficiente de variación estadístico.
- Asignación de las series de proyección de población a cada una de las series de temperatura seleccionadas teniendo en cuenta el municipio donde está localizada.

- Estimación del indicador de temperatura de cada área operativa como una media ponderada de las temperaturas de cada serie de acuerdo con la población de los municipios donde éstas están localizadas.

2.2.1.2 Demanda de energía eléctrica

En la sección 1.4.1 se identificaron las limitaciones para llevar a cabo la transformación de la serie total de demanda de energía eléctrica. En adelante, se describe el protocolo para transformar dicha serie en otras series que reflejen el comportamiento de la demanda de carga en cada una de las áreas operativas que integran el Sistema Interconectado Nacional – SIN.

a. Estimación de la demanda comercial y de la demanda regulada

La demanda comercial corresponde a la diferencia entre la demanda total y la demanda no atendida o a la suma entre la demanda regulada y no regulada (XM Compañía de Expertos en Mercados S.A. E.S.P., 2015a), y la demanda regulada corresponde a la carga requerida por usuarios industriales, comerciales y residenciales cuando estos demandan menos 55 MWh/mes de energía eléctrica o menos a 0.1 MW de potencia (Compañía Energética del Tolima S.A. E.S.P., 2015). A continuación se presentan las expresiones matemáticas utilizadas para estimar el valor de la demanda comercial y de la demanda regulada de energía eléctrica en el SIN.

$$D_c = D_{SIN} - D_{na}$$

$$D_r = D_c - D_{nr}$$

Donde D_c es la demanda comercial, D_{SIN} es la demanda real del SIN, D_{na} es la demanda no atendida, D_r es la demanda regulada y D_{nr} es la demanda no regulada.

b. Estimación de la demanda comercial, regulada y no regulada por área operativa

La demanda comercial, la demanda regulada y la demanda no regulada por área operativa, no se pueden estimar de forma directa con la información disponible en el portal del operador del mercado. Por tanto, es necesario suponer que la demanda comercial, la demanda regulada y demanda no regulada en cada área operativa del SIN, son proporcionales al pronóstico de demanda en cada área operativa. En este contexto, la demanda comercial diaria de un área operativa será igual al producto entre la demanda comercial del SIN y la relación entre el pronóstico de demanda de dicha área y el pronóstico para la demanda total o agregada. De forma equivalente, la demanda regulada y la no regulada de un área operativa será igual al producto entre la demanda regulada y no regulada del SIN

respectivamente, y la relación entre el pronóstico de demanda de dicha área y el pronóstico de la demanda agregada del Sistema Interconectado Nacional - SIN.

A partir de los supuestos formulados, es posible transformar la demanda agregada de energía eléctrica en series de demanda para cada una de las áreas operativas del Sistema Interconectado Nacional. En este sentido, se presenta una solución al segundo y tercer limitantes identificados en la sección 1.4.1 para llevar a cabo el proceso de transformación de la serie de demanda de electricidad en Colombia.

Respecto a las críticas que pueden originarse en relación con los supuestos formulados para la transformación de la serie agregada de demanda de energía eléctrica, es de aclarar que estos no fueron propuestos para obtener series reales de demanda en cada área de operación sino para construir series que reflejen el comportamiento o la dinámica en la demanda de carga en dichas áreas operativas.

c. Eliminación de efectos económicos en las series de demanda

Con el objetivo de eliminar el efecto generado en la demanda por el crecimiento económico del país, las series de demanda comercial, regulada y no regulada por área operativa serán corregidas teniendo en cuenta el factor de corrección (Sailor & Muñoz, 1997) adaptado para una serie con resolución diaria. Las expresiones matemáticas a utilizar se presentan a continuación:

$$f_i = \frac{\overline{E}_i}{\overline{\overline{E}}}$$

Donde $\overline{E}_i$ es la demanda de electricidad media del año i y $\overline{\overline{E}}$ es la demanda diaria promedio de toda la serie. Finalmente, la demanda de electricidad corregida será:

$$EC_{ij} = \frac{E_{ij}}{f_i}$$

Donde EC_{ij} es la demanda corregida de energía eléctrica del día j del año i, E_{ij} es la demanda original y f_i es el factor de corrección del año i.

2.2.2 Propiedades estadísticas de las series

Con el objetivo de realizar una caracterización básica de las series de indicador de temperatura media del aire y de demanda de energía eléctrica de cada una de las áreas operativas, se deben estimar los siguientes estadísticos básicos (Startz, 2015).

- **Media:** media aritmética.
- **Mediana:** Valor que se encuentra en la posición central de la serie.
- **Máximo**
- **Mínimo**
- **Desviación estándar:** Medida de la dispersión de los datos.

- **Coeficiente de Asimetría (Skewness):** Parámetro que permite determinar si los datos se distribuyen de forma uniforme alrededor de la media.
- **Curtosis:** Medida del grado de concentración de los datos alrededor de la media.
- **Estadístico Jarque-Bera:** Estadístico que evalúa si una serie se distribuye normalmente. Mide la diferencia entre la asimetría de la distribución y la curtosis. Una serie tiene una distribución normal si el estadístico Jarque-Bera es inferior a 5,99.

2.2.3 Propiedades de las series de tiempo

En el proceso de caracterización de una serie de tiempo es importante analizar sus componentes, ya que a partir de estos se define el procedimiento idóneo para utilizar estudiar dicha serie. A continuación se describen los principales componentes de una serie de tiempo (Murillos, Trejos, & Carvajal, 2003).

- **Autocorrelación:** Medida de la relación o independencia entre dos variables que se encuentran separadas o rezagadas k periodos.
- **Estacionariedad:** Se dice que una serie es estacionaria cuando la media y varianza no cambian en el tiempo.
- **Estacionalidad:** Una serie de tiempo es estacional cuando en ella se observa un patrón de cambio semana tras semana, mes a mes, año tras año, entre otros.
- **Heterocedasticidad:** Medida de la variabilidad de la serie, es decir, de la varianza de las observaciones. Se dice que una serie es heterocedástica cuando la varianza de dicha serie no es constante en el tiempo. En caso contrario, la serie se considera homocedástica.
- **Raíz unitaria:** Indica si una serie no retorna a su tendencia. Se dice que una serie tiene raíz unitaria cuando no es estacionaria o cuando no recupera su tendencia, es el caso de series que presentan crecimiento o declinación a largo plazo.

2.2.4 Análisis de estacionariedad de las series

En la presente sección se describen los análisis de autocorrelación, estacionariedad y estacionalidad realizadas a las series de indicador de temperatura media del aire y de demanda de energía eléctrica de cada una de las áreas operativas que integran el Sistema Interconectado Nacional.

2.2.4.1 Correlogramas de las series

Los correlogramas son representaciones gráficas de la función de autocorrelación (AC) y de la función de autocorrelación parcial (PAC), con los cuales se puede establecer si una serie es un proceso autorregresivo. En otras palabras, permiten identificar si los valores que toma una variable en el tiempo dependen o no de valores anteriores de la misma variable (McCandless, Gabrielli, & Murphy, 2001).

Un proceso es autorregresivo AR(1), cuando el valor de la función de autocorrelación disminuye exponencialmente a cero o lo hace de forma sinusoidal convergente a cero. Por otra parte, se dice que un proceso es autorregresivo AR(2) cuando los valores de la función de autocorrelación decrecen

exponencialmente hacia cero a partir del segundo valor o cuando lo hacen de forma sinusoidal no convergente a cero (McCandless et al., 2001).

Los correlogramas permiten además identificar cuáles son los rezagos que deben considerarse en un modelo, en especial, a partir de los valores de la función de autocorrelación parcial (McCandless et al., 2001). Para efectos del presente trabajo final de maestría, se considera que los rezagos a tener en cuenta en la construcción de modelos de análisis de variables son aquellos cuyo valor de PAC es superior a 0.5.

A continuación se presenta la descripción de las dos formas que existen para medir la dependencia de las variables en una serie de tiempo.

a. Autocorrelación

La función de autocorrelación (AC) corresponde a la medida de correlación entre las observaciones ordenadas de una serie de tiempo desfasadas diferentes órdenes o periodos, es decir, proporciona información sobre la relación lineal entre observaciones de una serie que están separadas por k rezagos (McCandless et al., 2001). La función de autocorrelación de un proceso AR(p) se presenta a continuación (Sánchez González, 1999):

$$\rho_k = \Phi_1 \rho_{k-1} + \Phi_2 \rho_{k-2} + \cdots + \Phi_p \rho_{k-p}$$

La demostración de la función de autocorrelación se presenta en el **Anexo 8.2.**

b. Autocorrelación Parcial

La función de autocorrelación parcial (PAC) corresponde, al igual que en la AC, a la medida de correlación entre las observaciones ordenadas de una serie de tiempo desfasadas diferentes órdenes o periodos, pero con la diferencia de que en la PAC se aíslan los efectos lineales con las demás observaciones de la serie de tiempo. La función de autocorrelación parcial de un proceso AR(p) se estima mediante el siguiente procedimiento (Sánchez González, 1999):

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \cdots + \Phi_p Y_{t-p} + \mathrm{a}_t$$

2.2.4.2 Test Durbin-Watson

La prueba Durbin-Watson es una medida de la correlación serial en los residuales en el cual se supone que el término del error estocástico a_t se genera a través de un proceso autorregresivo de primer orden (Durbin & Watson, 1950, 1951):

$$a_t = \mu_t + \theta a_{t-1} \quad , t \geq 2$$

Donde μ_t es el error aleatorio no correlacionado con otro componente o ruido blanco, a_{t-1} es el término de error retardado un periodo y θ una constante cuyo valor absoluto es inferior a la unidad. La expresión de contraste o estadístico de prueba en la prueba Durbin Watson es:

$$DW = \frac{\sum_{t=2}^{n}(a_t - a_{t-1})^2}{\sum_{t=1}^{n} {a_t}^2}$$

Donde a_t es el residual del periodo y a_{t-1} el residual del periodo anterior y las hipótesis que desea probar son:

$H_0: \rho = 0$ → No existe autocorrelación

$H_1: |\rho| > 0$ → Existe autocorrelación

Donde el estadístico experimental se puede expresar como:

$$DW = 2(1 - \rho)$$

Debido a que $-1 \leq \rho \leq 1$, el estadístico experimental DW puede variar entre cero y cuatro así:

$$DW = 0, \quad sí\ \rho = 1 \ \rightarrow \ Autocorrelación\ positiva$$

$$DW = 2, \quad sí\ \rho = 0 \ \rightarrow \ No\ hay\ autocorrelación$$

$$DW = 4, \quad sí\ \rho = -1 \ \rightarrow \ Autocorrelación\ negativa$$

La demostración del estadístico experimental de DW se presenta en el **Anexo 8.3**.

2.2.4.3 Test Dickey-Fuller Aumentado

La prueba Dickey-Fuller Aumentado (ADF) es una prueba de estacionariedad en el cual se contrasta la presencia de una o más raíces unitarias en un proceso autorregresivo de orden AR(p) (Dickey & Fuller, 1979), a partir del siguiente modelo:

$$\Delta Y_t = \omega_0 + \gamma Y_{t-1} + \sum_{i=2}^{p} \beta_i \, \Delta Y_{t-i+1} + a_t$$

Donde ω_0 es una constante, γ es el estadístico de prueba para Y_{t-1}, que en caso de ser igual a cero indica la presencia de una raíz unitaria. Además:

$$\gamma = -\left(1 - \sum_{i=1}^{p} C_i\right)$$

y:

$$\beta_i = \sum_{j=i}^{p} \omega_j$$

La demostración del contraste ADF se presenta en el **Anexo 8.4**.

2.2.4.4 Test Phillips-Perrón

La prueba Phillips-Perron (PP), es una prueba de estacionariedad en la cual se contrasta, al igual que en el ADF, la presencia de una o más raíces unitarias en un proceso autorregresivo de orden AR(p). No obstante, este se diferencia del ADF porque no considera los términos de diferencias retardados y porque utiliza métodos no paramétricos para evitar la correlación serial en los términos del error, es decir, tiene en cuenta la heterocedasticidad en el proceso del error ya que este no necesariamente es ruido blanco (Phillips & Perron, 1988). El estadístico de prueba se obtiene considerando que la serie de tiempo es generada a partir de un proceso como el que se describe a continuación:

$$Y_t = \alpha Y_{t-1} + a_t$$

Modelo para el cual Phillips & Perron (1988) consideraron la siguiente ecuación de regresión:

$$Y_t = \tilde{\mu} + \tilde{\beta}\left(t - \frac{T}{2}\right) + \tilde{\gamma} Y_{t-1} + \tilde{a}_t$$

Donde $\tilde{\mu}$, $\tilde{\beta}$, $\tilde{\gamma}$ son coeficientes de regresión. Phillips & Perron (1988), definieron X como la matriz $T \times 3$ a partir de la cual se determinan los coeficientes de regresión, y los estadísticos de prueba $t_{\tilde{\mu}}$. $t_{\tilde{\beta}}$ y $t_{\tilde{\gamma}}$ con los cuales se establece si la serie tiene una raíz unitaria bajo la siguiente hipótesis:

H_0: $\tilde{\mu} = 0, \tilde{\beta} = 0, \tilde{\gamma} = 1$→ La serie tiene una raíz unitaria.

H_1:→ La serie es estacionaria.

Para ver la demostración completa la prueba de la prueba Phillips-Perrón véase (Phillips & Perron, 1988).

2.2.5 Prueba de causalidad de Granger

La prueba de causalidad de Granger es un test que estudia la relación entre dos variables, como lo dicen Gutiérrez Moya & Arévalo Quijada (2003), la prueba de causalidad de Granger más que indicar causalidad sugiere una relación de precedencia de la variable explicativa sobre la variable respuesta, en el caso, del indicador de temperatura media del aire sobre la demanda de energía eléctrica de cada una de las áreas operativas que integran el Sistema Interconectado Nacional.

2.2.6 Prueba de variables irrelevantes

La prueba de variables irrelevantes o redundantes permite determinar la significancia estadística de una variable o un conjunto de variables en un modelo econométrico, es decir, ayuda a establecer cuáles son las variables explicativas que deben ser incluidas en la construcción de dicho modelo. Se considera que una variable no debe ser incluida en un modelo cuando la probabilidad del estadístico F y de la relación log likelihood es inferior a 5% (Gutiérrez Moya & Arévalo Quijada, 2003).

2.2.7 Prueba de correlación serial de los residuos

La prueba de correlación serial de los residuos es un test que diagnostica la existencia de autocorrelación serial de segundo orden entre estos. Se considera que los residuos de un modelo econométrico no están autocorrelacionados cuando las probabilidades de los estadístico F y LM son superiores al nivel de significancia del 5% (Gutiérrez Moya & Arévalo Quijada, 2003).

2.2.8 Análisis de los residuos

El análisis de los residuos consiste en analizar los correlogramas de los residuos en niveles y el correlograma de los residuos al cuadrado. En los primeros se determina si los residuos se encuentran autocorrelacionados; mientras que a partir de los segundos es posible establecer si los residuos presentan heterocedasticidad condicional autorregresiva, es decir, si se requiere modelar volatilidad en el modelo (Gutiérrez Moya & Arévalo Quijada, 2003).

2.2.9 Estimación de la Volatilidad

Cuando los correlogramas de los residuos al cuadrado indican que éstos presentan autocorrelación, es decir, heterocedasticidad condicional autorregresiva, es necesario estimar la volatilidad, ya que los cambios en ésta pueden generar distorsiones en la variable en niveles (Granados Castro, 2015). Para la modelación de dichos efectos, en la literatura se encuentran referenciados distintos modelos. No obstante, en el presente trabajo final de maestría sólo se describen los que se presentan a continuación:

2.2.9.1 Modelo GARCH

El modelo GARCH (p, q) surge como una generalización del modelo ARCH (q), el cual fue desarrollado, como lo indica Granados Castro (2015), con el objetivo de replicar regularidades de activos financieros, agrupamientos de volatilidad por periodos, colas en la distribución de los errores y la no independencia de los errores. La diferencia entre el modelo GARCH (p, q) y el modelo ARCH (q) radica en que el segundo generaliza la dependencia de la varianza condicional de los residuos al incorporar valores pasados de la varianza condicional, por lo que se requiere de un menor número de cuadrados de residuos (Tecano Molano, 2010). Para ampliar la descripción de los modelos véase (Bollerslev, 1986).

2.2.9.2 Modelo EGARCH

El modelo EGARCH o GARCH exponencial es una variación del modelo GARCH, en el cual se modela el efecto de apalancamiento o de cambios asimétricos de los incrementos y decrementos de una serie de rendimiento, que no puede ser modelado mediante un modelo GARCH (Sosa et al., 2014). Para ampliar la descripción del modelo véase (Nelson, 1991).

3. PRE PROCESAMIENTO: SERIES DE TEMPERATURA

3.1 La temperatura media del aire

Como se indicó en la sección 2.1.1 para la elaboración de este trabajo se emplearon los registros diarios de temperatura de 448 estaciones climáticas distribuidas en el territorio nacional, en particular en los departamentos donde opera el Sistema Interconectado Nacional, entre el 01:01:2007 y el 31:12:2013. Sin embargo, algunas de las series presentan lagunas puntuales o continuas de información que deben ser eliminadas a través de la aplicación de técnicas de completamiento. Para ello se utilizó la metodología propuesta en la sección 2.2.1.1, en la cual se presentan los pasos y los criterios que deben tenerse en cuenta para completar y transformar las series de temperatura media del aire.

A continuación se presentan los resultados del pre procesamiento de las series de temperatura media del aire para la obtención de la serie de indicador de temperatura para cada una de las áreas operativas que integran el Sistema Interconectado Nacional.

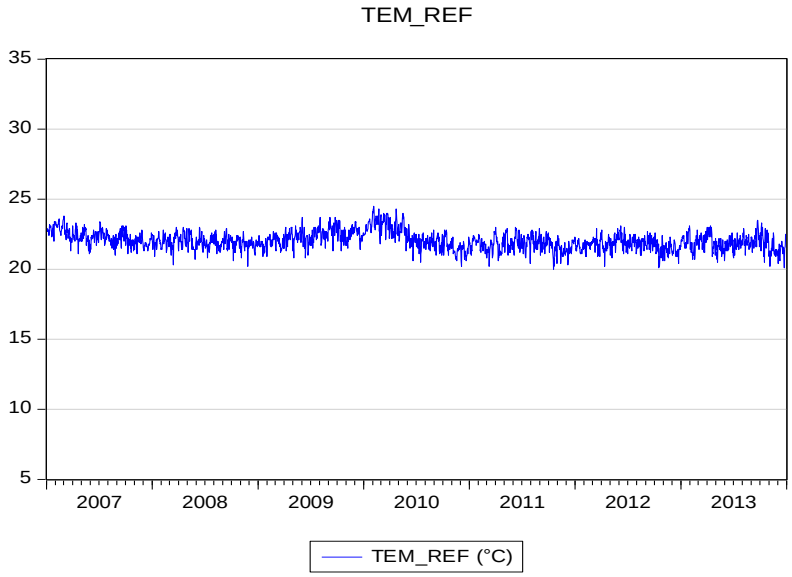
3.1.1 Control de calidad

Las 448 series de temperatura media del aire fueron sometidas a un control de calidad en el cual se determinó que la calidad de los registros tomados era muy diversa no sólo por el porcentaje de completamiento de datos sino también porque algunas series poseían valores incorrectos o dudosos, los cuales fueron corregidos o eliminados. Cabe notar, que es probable que dichos valores estuvieran asociados con errores en el registro o en la digitación de los mismos. Por ejemplo, algunas estaciones poseían dígitos adicionales o espacios entre los dígitos del registro.

3.1.2 Pruebas de homogeneidad

Una vez finalizado el control de calidad, se procedió a establecer si los cambios en las series de temperatura media del aire de cada estación climática obedecían únicamente a cambios en las condiciones meteorológicas. Para ello, se realizó la prueba de dobles masas o dobles acumulaciones a través de la cual se determina si una serie es homogénea o heterogénea. De acuerdo con el planteamiento presentado en el numeral b de la sección 2.2.1.1, a continuación se muestra la gráfica de la serie de tiempo de referencia.

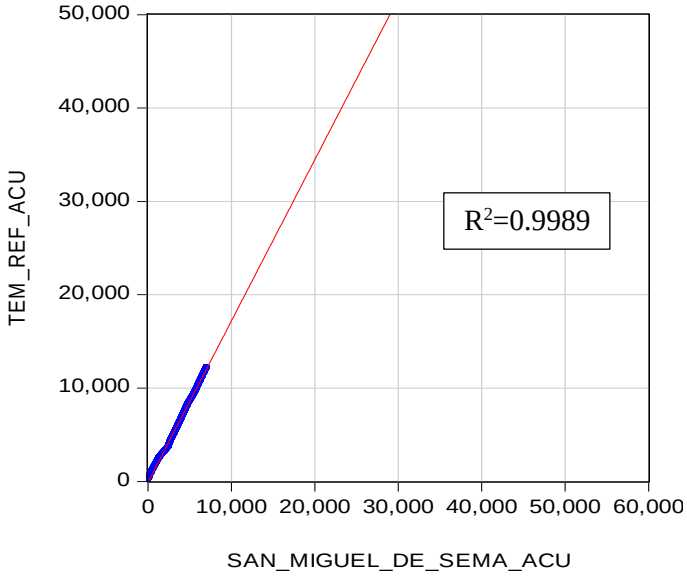
Figura 3.1 Serie de tiempo de referencia de temperatura en Colombia.



Fuente: Elaboración propia a partir de Eviews.

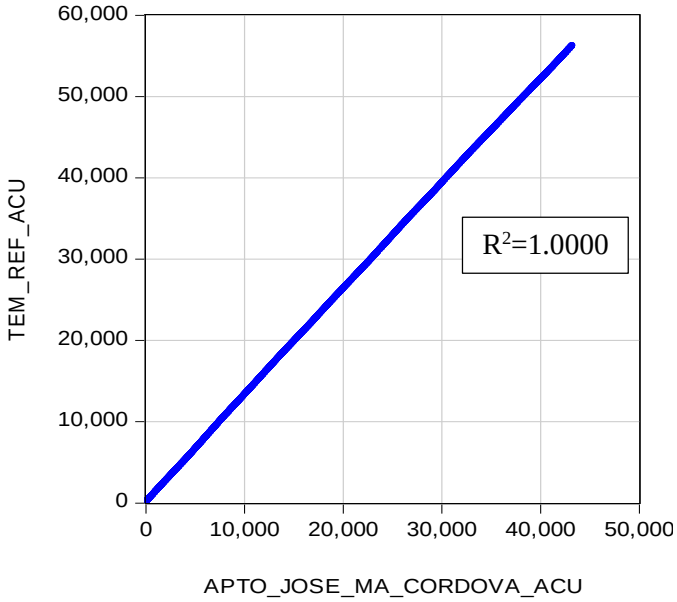
A partir de la serie presentada en la **Figura 3.1** se verificó la homogeneidad de las 448 estaciones climáticas localizadas en los departamentos donde opera el Sistema Interconectado Nacional. Los resultados obtenidos indican que el 100 % de las series analizadas son homogéneas ya que las series acumuladas de cada una en función de la serie de referencia acumulada presentaron un comportamiento lineal y no se evidenciaron saltos que indicaran que estas fueran heterogéneas (Barrera, 2004). Las gráficas de dobles masas o acumulaciones de las 448 estaciones meteorológicas evaluadas se presentan en el **Anexo 8.5**. A continuación se muestran las gráficas de la estaciones con menor y con uno de los mayores coeficientes de determinación o R Cuadrado.

Figura 3.2 Test de homogeneidad de la estación San Miguel de Sema.



Fuente: Elaboración propia.

Figura 3.3 Test de homogeneidad de la estación Apto José M Córdova.



Fuente: Elaboración propia.

En la **Tabla 3.1** se presentan el número de estaciones de acuerdo con el valor del coeficiente de determinación o R Cuadrado obtenido en la prueba de homogeneidad realizada a las 448 estaciones climáticas utilizadas en la fase de completamiento.

Tabla 3.1 Número de estaciones según coeficiente de determinación.

R^2	No. de Estaciones
0,9989	1
0,9990	1
0,9994	2
0,9995	2
0,9996	1
0,9997	8
0,9998	14
0,9999	108
1,0000	311
Total	**448**

Fuente: Elaboración propia.

3.1.3 Análisis de correlación lineal

Debido a que en el completamiento de datos en una serie puede utilizarse la información de otra serie temporal registrada en el mismo periodo de tiempo, en el presente trabajo se calculó el coeficiente de correlación lineal de Pearson para los 200.704 emparejamientos de series de temperatura.

El objetivo de realizar el análisis de la correlación lineal entre las serie fue determinar cuáles eran las series candidatas a ser elegidas para completar datos o lagunas de información en otras series. Para ello, se tuvo en cuenta que Barrera (2004), en su trabajo de investigación, estableció que el umbral

mínimo aceptable de coeficiente de correlación lineal entre series es de 0,7. Cabe resaltar que en caso de que dos series presentaran igual coeficiente de correlación lineal de Pearson respecto a otra serie, se consideró con mayor representatividad aquella que estuviera más cerca geográficamente y por tanto, en el análisis de correlación lineal también se calcularon las distancias relativas de una estación dada respecto a los otros 447 observatorios meteorológicos.

3.1.4 Significancia estadística

Con el objetivo de verificar si los coeficientes de correlación de Pearson eran estadísticamente significativos, se estimó la significancia estadística de estos a través de la aplicación de la prueba de contraste de hipótesis basada en la distribución de la **t** de student con N-2 grados de libertad. Los resultados obtenidos indican que el 100 % de los coeficientes de correlación de Pearson son estadísticamente significativos, ya que en todos los casos se rechazó la hipótesis nula de que la correlación procediera de una población con correlación cero.

3.1.5 Completamiento de las series

En el completamiento de las series se tuvieron en cuenta los criterios descritos en el numeral f de la sección 2.2.1.1 y las recomendaciones de Barrera (2004). En primer lugar se definió que se utilizaría uno de los cuatro métodos multivariantes presentados en el numeral e de la sección 2.2.1.1 para realizar el completamiento de las series de temperatura. El método seleccionado fue el de la razón normal propuesto por Paulus et al. (1952). Sin embargo, todas las series no contaban con al menos tres series a partir de las cuales realizar la estimación de datos incompletos y otras, en cambio, poseían aún más series que este valor umbral. Por tal motivo, se definió que si bien el método de la razón normal era el indicado para realizar el completamiento de las lagunas de información, porque suaviza la influencia de posibles errores y evita que se asignen datos no representativos de una serie a otra, no podría ser utilizado conforme a la propuesta elaborada por Paulus et al. (1952). En este sentido, se propuso una modificación metodológica en la cual se planteó que se emplearía la información de hasta máximo cinco estaciones climáticas que presentaran un coeficiente de correlación lineal de Pearson mayor a 0,7; respecto a una serie dada, para estimar un dato incompleto en ésta.

En la primera fase del completamiento, las series de temperatura base fueron separadas en 84 meses entre 01:2007 y 12:2013, con el objetivo de utilizar para la estimación de los datos incompletos, en un mes determinado, la media mensual de cada serie y de esta forma evitar incluir en el completamiento de los datos los efectos de las variaciones climáticas propias de otros periodos climáticos interanuales e históricos.

En la segunda fase de completado, se recalculó el porcentaje de completamiento de los datos y se determinó que era necesario rellenar lagunas que no habían podido ser eliminadas. Para ello, se repitió el procedimiento realizado en dicha fase. Por lo anterior, se reestimaron los coeficientes de correlación lineal de Pearson de los emparejamientos para las nuevas series, se verificó que éstos cumplieran con el criterio de la significancia estadística y de nuevo, fueron calculados los porcentajes de completamiento. No obstante, aún eran pocas las estaciones que estaban completas al 100 % respecto al número inicial de observatorios y por tanto, se dio lugar a una tercera fase de

completamiento en la cual se estimaron los datos faltantes puntuales que podían ser calculados, a partir de los valores registrados o estimados en las etapas previas, como el promedio de la medición de temperatura del día anterior y del día posterior. Luego, se recalculó el porcentaje de completamiento y para aquellas series que poseían un valor superior a 90 % e inferior a 100 %, se aplicó el criterio del valor medio. Cabe notar que según Barrera (2004) este método es aplicable cuando el coeficiente de variación estadístico es inferior al 5%, sin embargo el criterio del valor medio fue aplicado sólo a series que tenían como máximo tres datos faltantes, razón por la cual se supone que este método genera una perturbación mínima como la descrita por Rodríguez et al. (1999).

A pesar de realizar cuatro completamientos a las series base de temperatura, éstas aún no presentaban completamientos del 100 % y por tanto fue necesario efectuar un último completado. En esta etapa, se empleó de nuevo el método de la razón normal propuesto por Paulus et al. (1952), pero para los siete periodos anuales de estudio. Los porcentajes de completamiento de las series se presentan en el **Anexo 8.6**.

Con el objetivo de caracterizar las series de temperatura media del aire completadas, se utilizaron los mismos cuatro rangos de porcentaje de completamiento definidos en la sección 2.1.1. En la **Tabla 3.2** se presenta el número de estaciones según el porcentaje de datos registrados en el periodo de análisis.

Tabla 3.2 Número de estaciones según rango de completamiento.

Clasificación	Rango de Completamiento	No. Estaciones	% de Estaciones según Rango de Completamiento
Grupo 1	$0 < x \leq 25$	3	0,6
Grupo 2	$25 < x \leq 50$	21	4,7
Grupo 3	$50 < x \leq 75$	63	14,1
Grupo 4	$75 < x \leq 100$	361	80,6

Fuente: Elaboración propia.

Así mismo, en la **Tabla 3.3** se muestra el número de estaciones según porcentaje de completamiento y por departamento.

Tabla 3.3 Número de estaciones según rango de completamiento por área operativa.

Departamento	Rango				Total
	$0 < x \leq 25$	$25 < x \leq 50$	$50 < x \leq 75$	$75 < x \leq 100$	
Antioquia	0	3	12	44	59
Centro	2	8	9	47	66
Costa Atlántica	0	1	17	58	76
CQR	0	0	2	12	14
Oriente	0	5	10	84	99
Sur	0	0	1	44	45
THC	1	4	9	57	71
Valle	0	0	3	15	18
Total de Estaciones	3	21	63	361	**448**

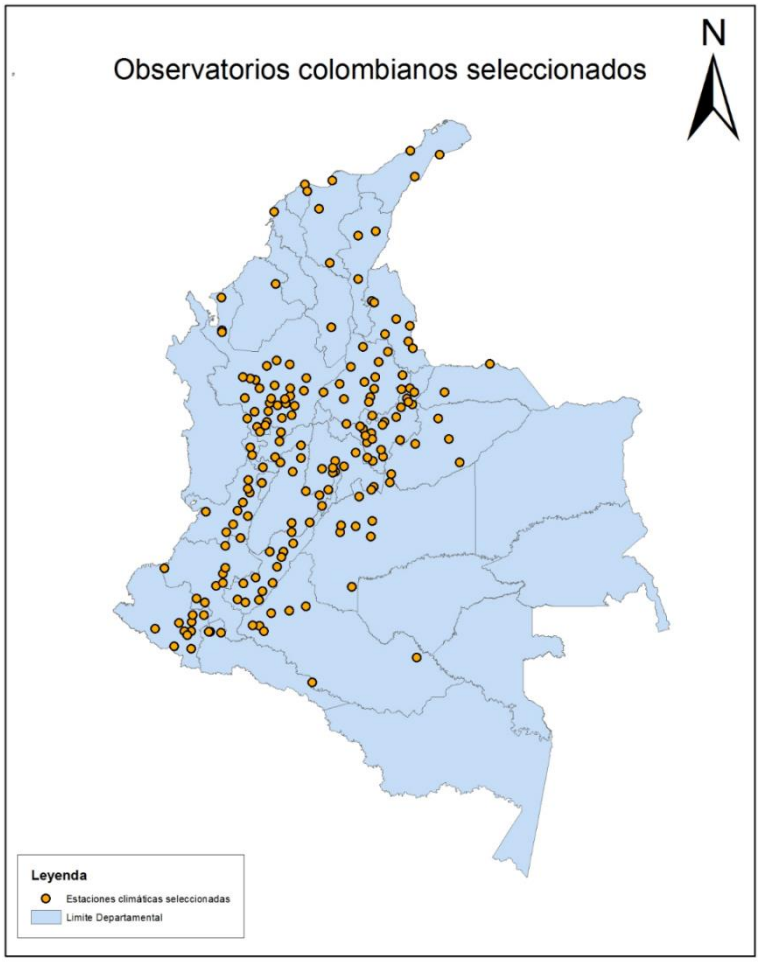
Fuente: Elaboración propia.

3.1.6 Indicador de temperatura por área operativa

Como se indicó en el numeral g de la sección 2.2.1.1, en la presente investigación se pretende evaluar el efecto de las variaciones de la temperatura media del aire sobre la demanda de energía eléctrica en cada una de las áreas operativas que integran el Sistema Interconectado Nacional. Para ello, fue considerada la metodología propuesta por Moral et al. (2005). No obstante, cabe notar que algunas de las series seleccionadas tienen un valor del coeficiente de variación estadística superior al 10 %, pero fueron consideradas en la estimación del indicador, ya que su porcentaje de completamiento inicial era cercano al 100%.

La lista de las estaciones climáticas seleccionadas para el cálculo del indicador de temperatura en cada una de las áreas operativas, y algunos de los parámetros básicos de estadística descriptiva para las series de temperatura de dichos observatorios meteorológicos se presentan en el **Anexo 8.7**. Las estaciones climáticas seleccionadas se muestran en la siguiente figura.

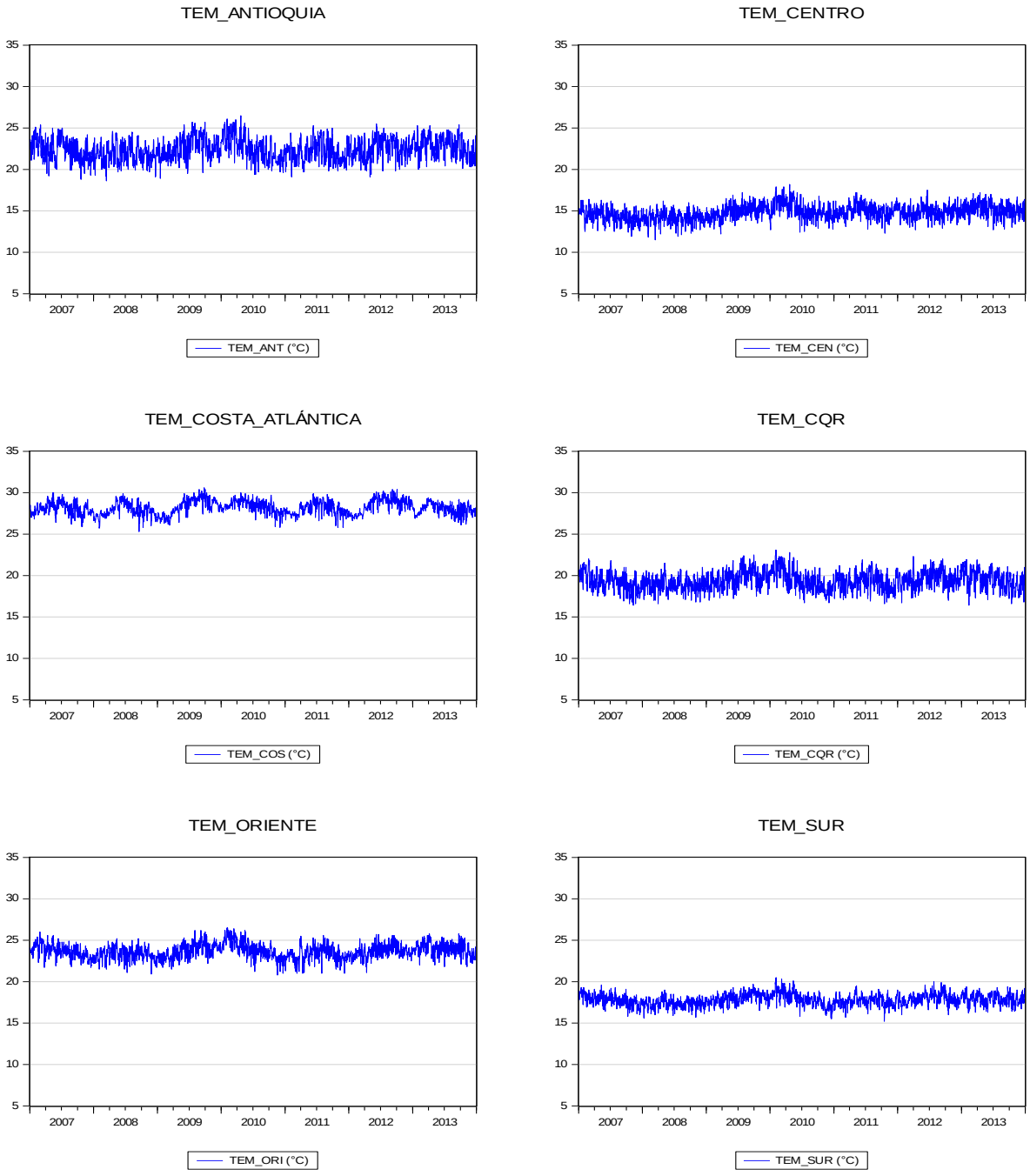
Figura 3.4 Observatorios climatológicos colombianos seleccionados.



Fuente: Elaboración propia.

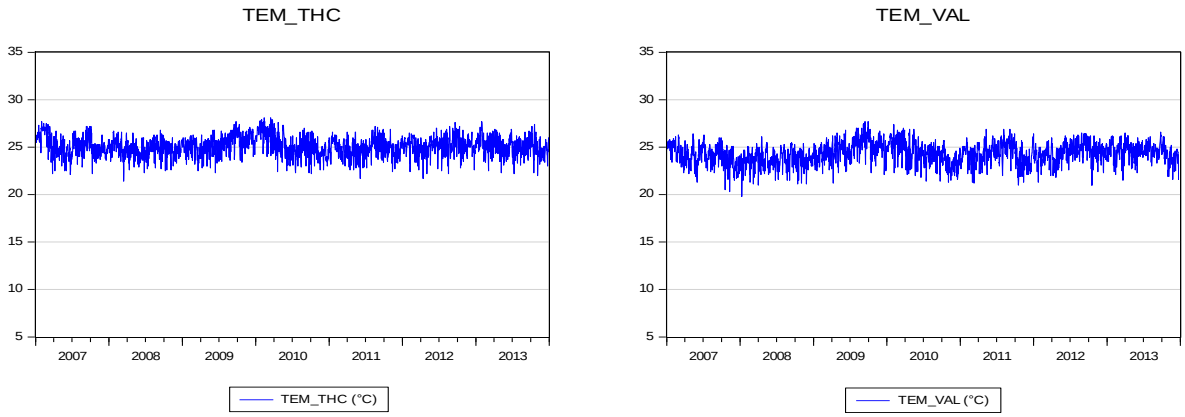
A continuación, se presentan las gráficas de las series de indicadores de temperatura para cada una de las áreas operativas que integran el Sistema Interconectado Nacional. Dichas series fueron estimadas como una media ponderada de las temperaturas de cada serie de acuerdo con la población de los municipios donde éstas están localizadas.

Figura 3.5 Serie de Indicador de Temperatura para las ocho áreas operativas (a).



Fuente: Elaboración propia a partir de Eviews.

Figura 3.6 Serie de Indicador de Temperatura para las ocho áreas operativas (b).



Fuente: Elaboración propia a partir de Eviews.

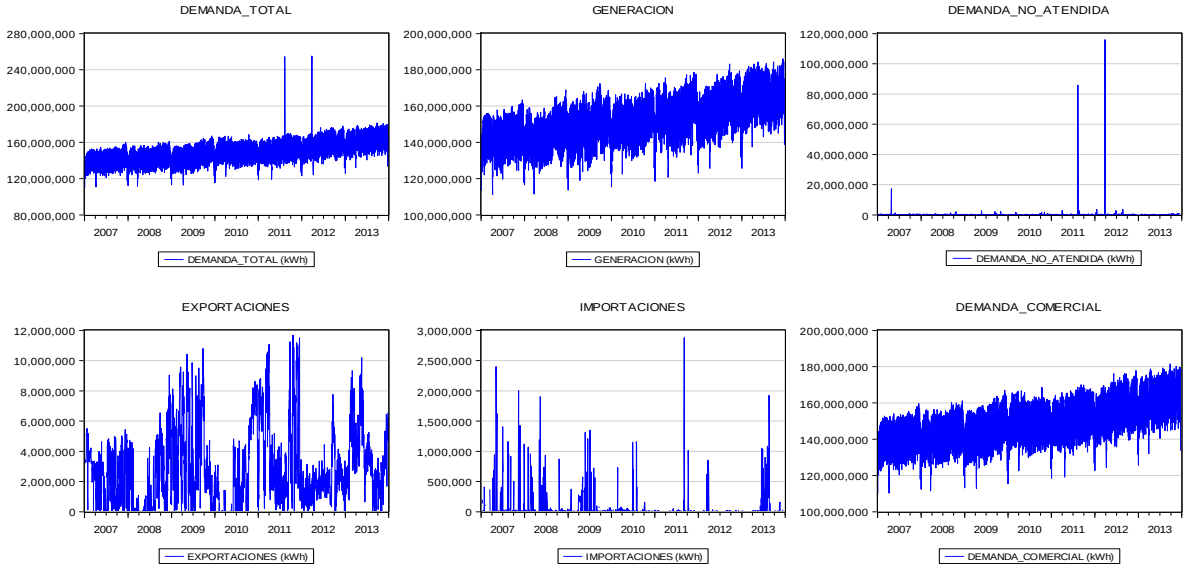
La metodología propuesta por Moral et al. (2005) permite estimar series de con un indicador de temperatura para un área geográfica dada tal y como se observa en la **Figura 3.5** y en la **Figura 3.6**. Vale la pena resaltar que en todas las series estimadas se observa un patrón similar, no hay una tendencia significativa en ningún sentido ya que las series oscilan alrededor de sus medias. Por tal motivo, a priori podría considerarse que dichas series son estacionarias, lo cual es favorable ya que el objetivo de este trabajo es analizar el patrón de consumo de energía eléctrico como resultado de las variaciones en las condiciones climáticas.

4. PRE PROCESAMIENTO: DEMANDA DE ELECTRICIDAD

4.1 La demanda de energía eléctrica

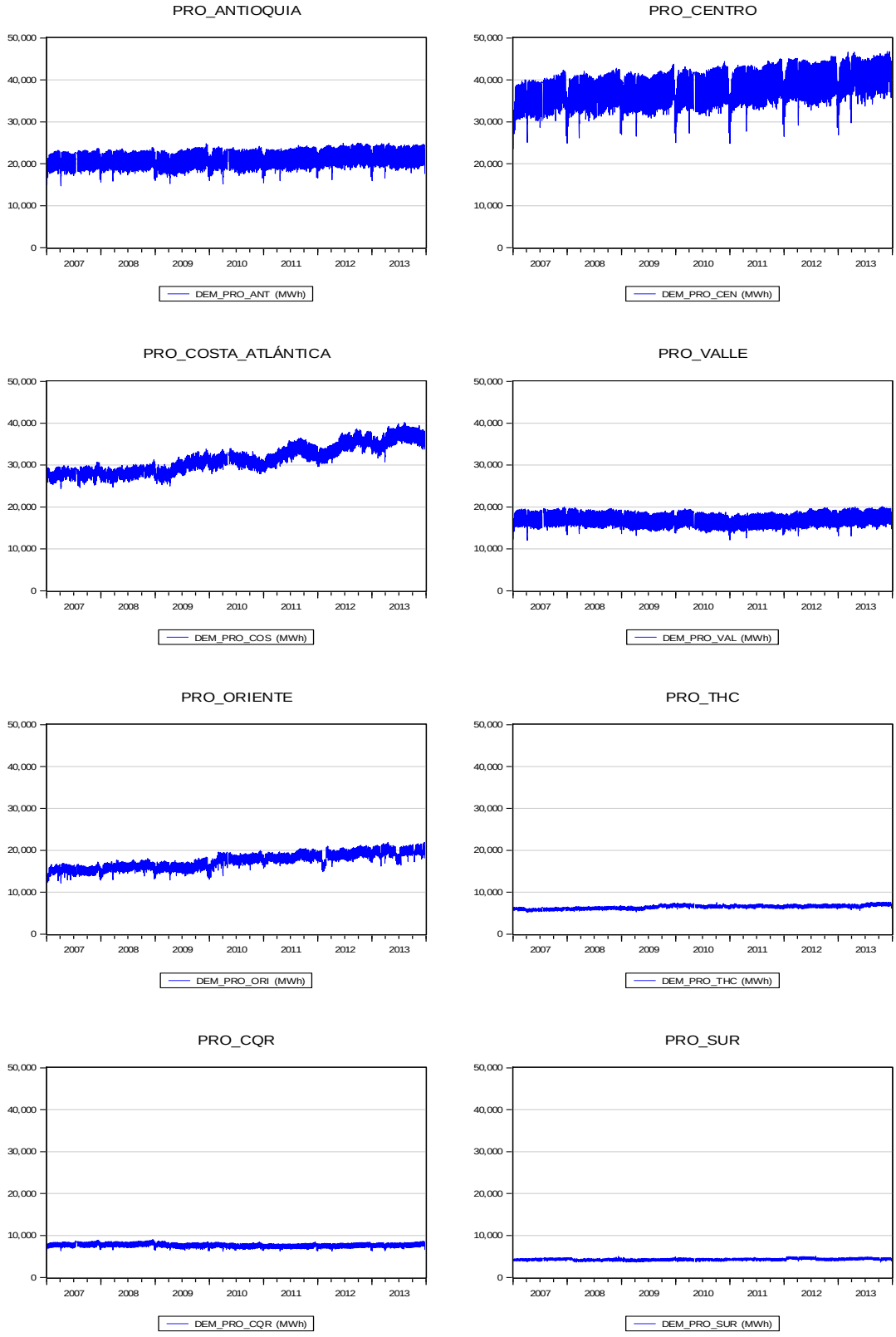
Como se indicó en la sección 2.1.2, se emplearon las series de demanda total del SIN, generación, demanda no atendida, exportaciones, importaciones, demanda comercial no regulada y las series de pronóstico oficial de demanda de electricidad para cada una de las áreas operativas, en resolución diaria y para el periodo comprendido entre el 01:01:2007 y el 31:12:2013; a fin de estimar series que reflejen el comportamiento o la dinámica de la demanda de electricidad en dichas áreas operativas. Las series no procesadas de demanda se muestran en las **Figura 4.1** y **Figura 4.2**.

Figura 4.1 Series de demanda de energía eléctrica del SIN.



Fuente: Elaboración propia a partir de Eviews.

Figura 4.2 Series de pronóstico de demanda de energía eléctrica del SIN en cada área operativa.

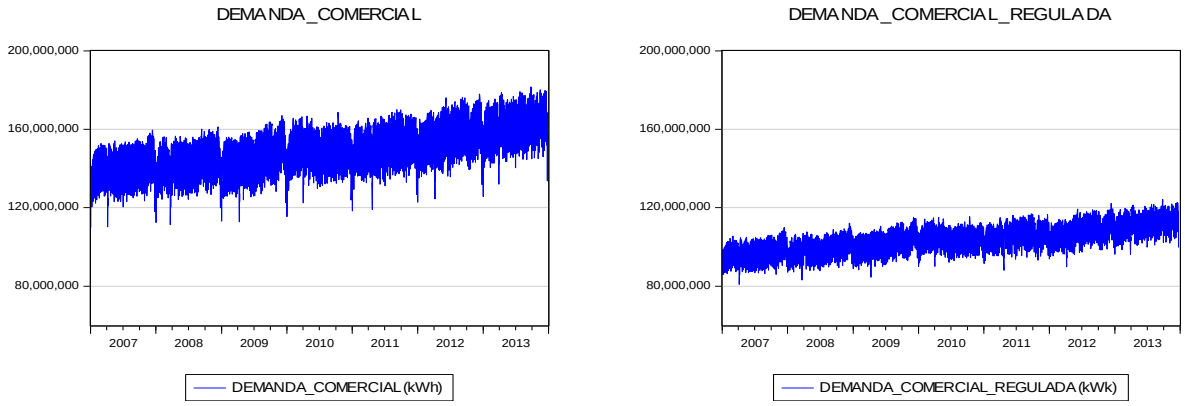


Fuente: Elaboración propia a partir de Eviews.

4.1.1 Demanda comercial y demanda regulada

De acuerdo con las definiciones presentadas en el numeral a de la sección 2.2.1.2 y las expresiones matemáticas utilizadas para estimar de demanda comercial y de demanda regulada de energía eléctrica del SIN, a continuación se presentan las series de dichas demandas.

Figura 4.3 Series de demanda comercial y regulada de energía eléctrica en el SIN.



Fuente: Elaboración propia a partir de Eviews.

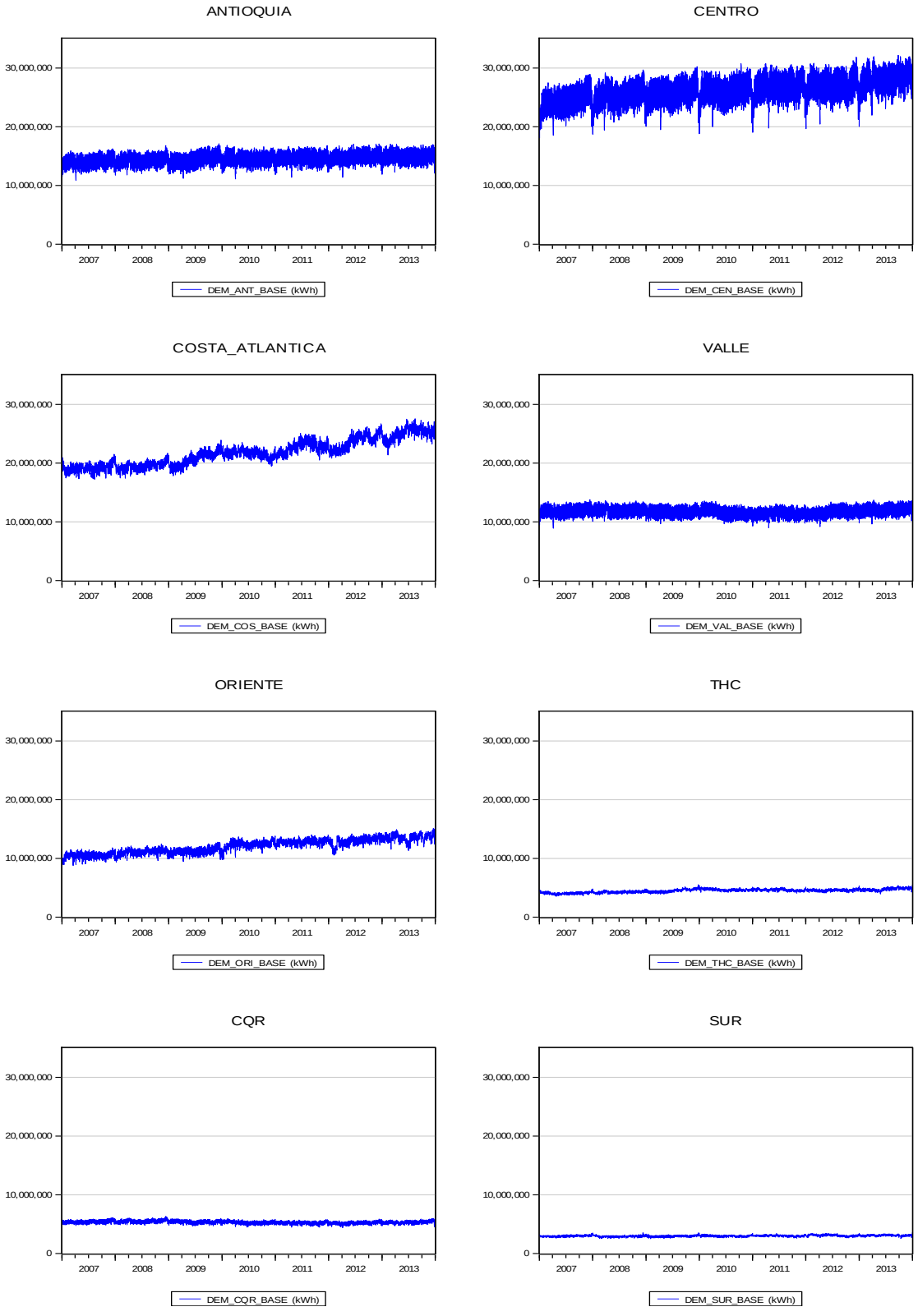
4.1.2 Demanda comercial, regulada y no regulada por área operativa

Teniendo en cuenta que la demanda comercial, la demanda regulada y la demanda no regulada por área operativa, no se pueden calcular de forma directa con la información disponible en el portal del operador del mercado, como se indicó en el numeral b de la sección 2.2.1.2, se supuso que la demanda comercial, la demanda la regulada y demanda no regulada son proporcionales al pronóstico de la demanda de electricidad de cada área operativa, es decir, que la demanda comercial diaria de un área operativa será igual al producto entre la demanda comercial del SIN y la relación entre el pronóstico de la demanda de energía en dicha área y el pronóstico para la demanda total del SIN. Lo anterior se aplica de manera análoga a las demandas regulada y no regulada del Sistema Interconectado Nacional.

Debido a que en el presente trabajo se utilizan únicamente las series de demanda comercial regulada de cada una de las áreas operativas, ya que se pretende establecer si los cambios en la temperatura media del aire generan variaciones en el consumo de electricidad en el sector residencial.

En la **Figura 4.4** se presentan las gráficas de la demanda regulada de energía eléctrica de cada una de las áreas operativas que conforman el Sistema Interconectado Nacional.

Figura 4.4 Series de demanda regulada de cada una de las áreas operativas.



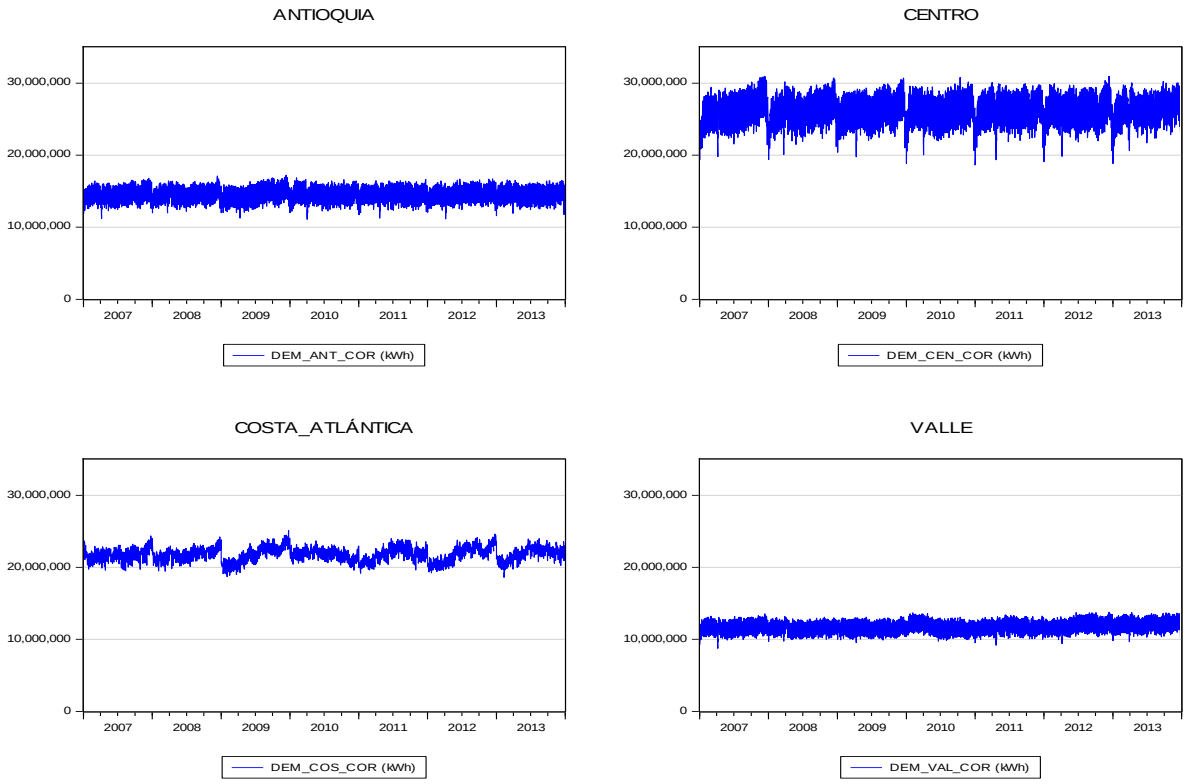
Fuente: Elaboración propia a partir de Eviews.

4.2 Eliminación de efectos económicos en las series de demanda

Como se presentó en la sección 2.2.1.2, para determinar si las variaciones en la temperatura generan cambios en el comportamiento de la demanda de electricidad, las series de demanda deben ser transformadas previamente. Para ello, es importante tener en cuenta que en el presente trabajo final de maestría sólo se pretende analizar el efecto de los cambios en la temperatura diaria del aire. Por lo tanto, variables como la evolución demográfica y el número de usuarios dejan de ser representativas en el presente estudio ya que éstas ganan mayor representatividad en el caso de análisis para horizontes de tiempo de mediano y largo plazo (Rueda, Velásquez, & Franco, 2011). No obstante, como se expresó en la sección 1.3, para evitar un desequilibrio en la estimación, se incorpora el efecto día o laboralidad a través de variables dummy que permiten acotar el comportamiento de la demanda de electricidad en las distintas áreas operativas.

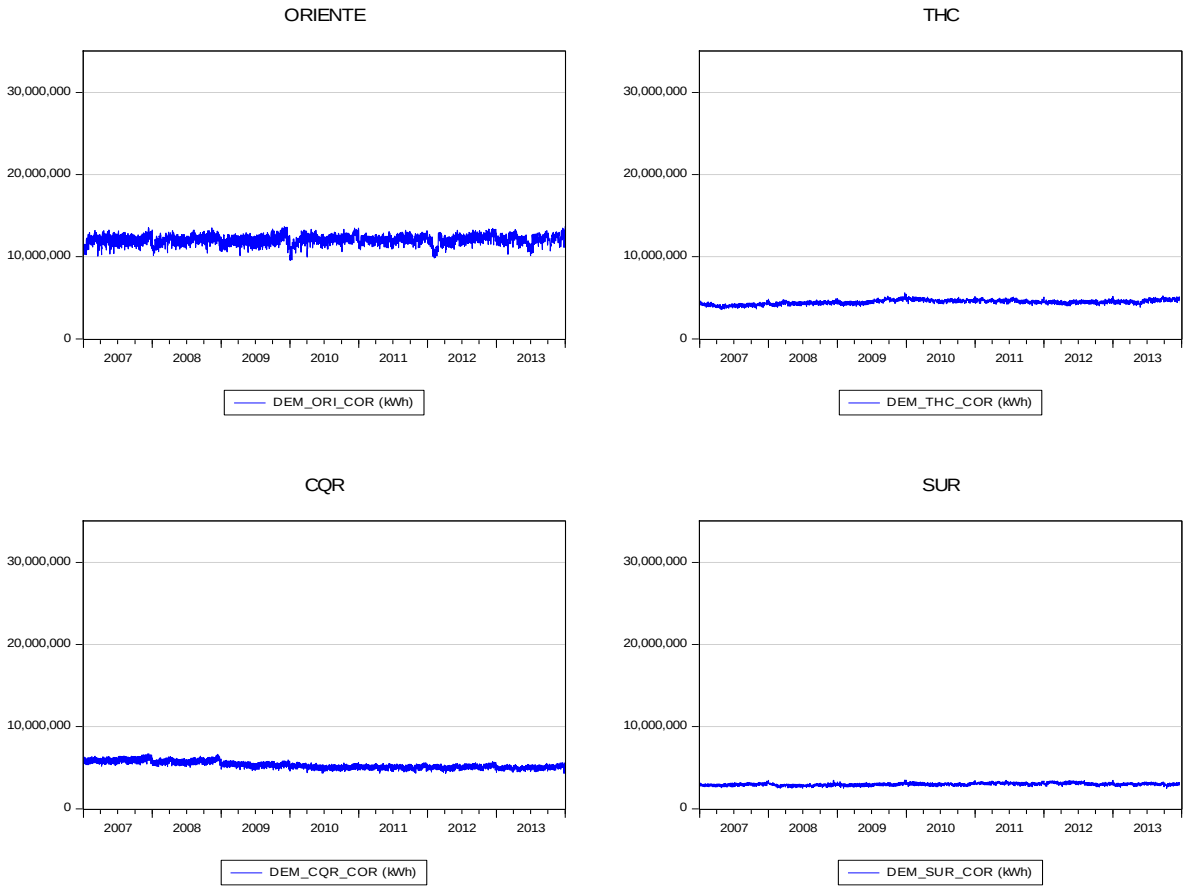
Dado que Franco et al (2008), determinaron que la demanda presenta una tendencia de largo plazo que está relacionada con el crecimiento del Producto Interno Bruto (PIB), se deben eliminar los efectos asociados con el crecimiento económico del país, y para ello se utilizó el factor de corrección propuesto por Sailor et al. (1997) adaptado para series con resolución diaria. Los factores de corrección anuales estimados se muestran en el **Anexo 8.8**. Teniendo en cuenta la metodología descrita en la sección 2.2.1.2, a continuación se presentan las series de demanda de energía eléctrica corregidas para cada una de las áreas operativas.

Figura 4.5 Demanda comercial regulada corregida de cada una de las áreas operativas (a).



Fuente: Elaboración propia a partir de Eviews.

Figura 4.6 Demanda comercial regulada corregida de cada una de las áreas operativas (b).



Fuente: Elaboración propia a partir de Eviews.

Como puede observarse en la **Figura 4.4**, las series de demanda de energía eléctrica de cada una de las áreas operativas tienen una tendencia al alza como resultado del crecimiento de la economía. Sin embargo los efectos del crecimiento económico no son igual de pronunciados en todas ellas, por ejemplo, en las áreas operativas Centro, Costa Atlántica y Oriente la demanda de energía eléctrica aumentó de forma considerable entre el año 2007 y año 2013, en comparación con la evolución observada en las demás áreas. Un patrón similar fue evidenciado por Climent Diranzo et al. (2003) para el caso español.

Una vez eliminados los efectos del crecimiento económico en la demanda de energía eléctrica, en el siguiente capítulo se analiza si los cambios en el comportamiento de ésta pueden ser explicados por las variaciones en las condiciones climáticas, particularmente en la temperatura media del aire.

5. RESULTADOS

5.1 Estadística descriptiva de las series

Como se indicó en la sección 2.1, las series de temperatura media del aire y de demanda de energía eléctrica contienen información diaria del periodo comprendido desde el 01:01:2007 hasta el 31:12:2013, lo cual equivale a 2.557 observaciones por serie, es decir, 40.912 registros en las 16 series analizadas.

Para realizar el análisis estadístico de las series se estimaron cada uno de los estadísticos descriptivos indicados en la sección 2.2.2 (véase **Tabla 5.1** y **Tabla 5.2**) y se graficaron los histogramas (véase **Figura 5.1** y **Fuente:** Elaboración propia a partir de Eviews.
Figura 5.2) tanto para la serie de indicador de temperatura media del aire como para la demanda de energía eléctrica de cada una de las áreas operativas que integran el Sistema Interconectado Nacional.

Tabla 5.1 Estadísticos descriptivos de las series de indicador de temperatura media del aire.

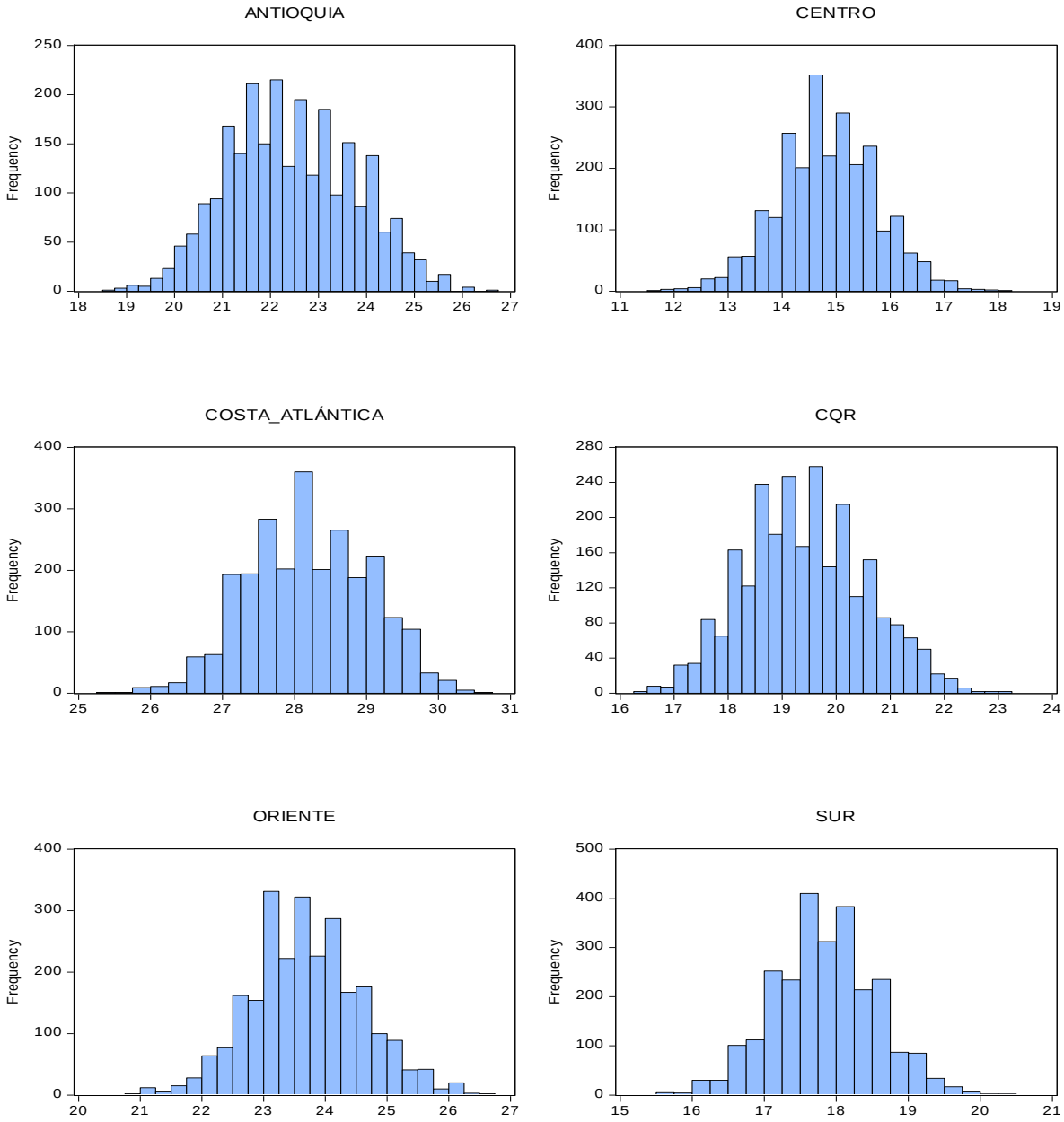
Estadístico Descriptivo	Indicador de temperatura media del aire (ºC) de cada área operativa							
	ANT	CEN	COS	CQR	ORI	SUR	THC	VAL
Media	22.4212	14.8111	28.1745	19.4113	23.6723	17.8164	25.0967	24.3130
Mediana	22.3000	14.8000	28.1000	19.4000	23.6000	17.8000	25.1000	24.4000
Máximo	26.5000	18.2000	30.6000	23.1000	26.5000	20.5000	28.1000	27.7000
Mínimo	18.6000	11.5000	25.3000	16.4000	20.8000	15.2000	21.4000	19.8000
Desv. Tip.	1.31701	0.91576	0.83769	1.12314	0.89852	0.71291	1.04616	1.14329
Asimetría	0.13290	0.01574	-0.01736	0.17244	0.11308	0.03352	-0.09167	-0.20989
Curtosis	2.48564	3.11717	2.55820	2.68064	3.09872	3.13692	2.96730	2.90945
JB	35.7150	1.56846	20.9238	23.5389	6.48795	2.47660	3.69541	19.6482
Prob.	0.00000	0.45647	0.00002	0.00000	0.03900	0.28987	0.15759	0.00005

Fuente: Elaboración propia.

De la **Tabla 5.1** se puede inferir que las series de indicador de temperatura media del aire de las ocho áreas operativas no tienen una distribución normal, ya que la prueba de hipótesis de normalidad Jarque Bera se rechaza con un nivel de confianza del 95% para un valor crítico de 5,99 (Jarque & Bera, 1987). Así mismo, dado que a excepción de las series de las áreas operativas Centro, Oriente y Sur, la curtosis es superior a tres, los valores de cada una de estas series de tiempo se aproximan mejor a una distribución de tipo leptocúrtica. Para los demás casos, los datos se ajustan mejor a una distribución de tipo platicúrtica.

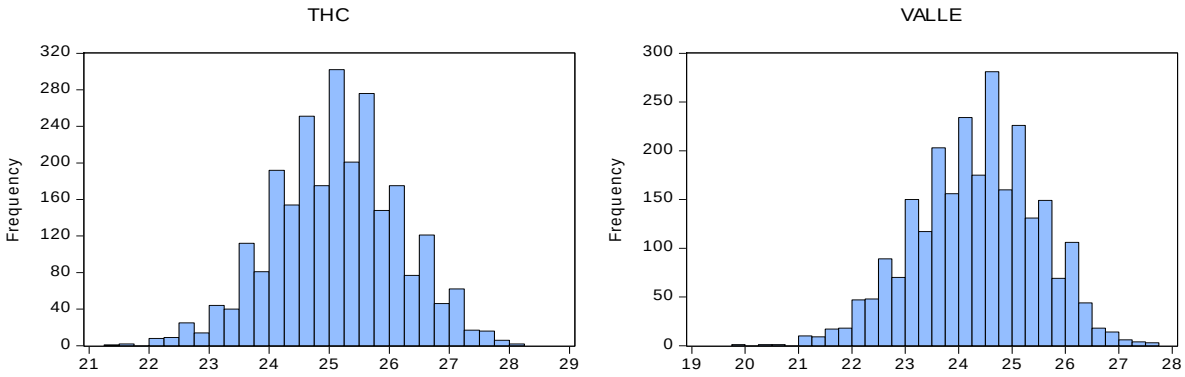
Adicionalmente, se puede ver en los histogramas de las series y en la desviación estándar calculada para cada una, que la dispersión de cada serie de tiempo podría estar relacionada con variaciones en la altitud y en la densidad urbana en cada una de las estaciones de medición de temperatura consideradas. De hecho, los resultados indican que las áreas operativas Antioquia, Valle y CQR (Caldas, Quindío & Risaralda) son aquellas en las cuales se presenta mayor desviación estándar, casualmente estos departamentos figuran entre los que se encuentran poblaciones en los diferentes pisos términos: cálido, medio, frio e incluso páramo. Lo contrario se observa en las áreas operativas Costa Atlántica, Oriente y Sur donde las estaciones están en altitudes menos variables.

Figura 5.1 Histogramas de indicador de temperatura media del aire (a).



Fuente: Elaboración propia a partir de Eviews.

Figura 5.2 Histogramas de indicador de temperatura media del aire (b).



Fuente: Elaboración propia a partir de Eviews.

En la **Tabla 5.2** se presentan los estadísticos descriptivos de las series de demanda de energía eléctrica de cada una de las áreas operativas que conforman el Sistema Interconectado Nacional. De esta tabla se puede inferir que la dispersión de las series aumenta con la media de la demanda histórica. Esto puede deberse a que la demanda de energía eléctrica se concentra en los centros urbanos más poblados del país, en los cuales la laboralidad, las temporadas vacacionales o el tipo de día de la semana podrían influir más en el comportamiento de la demanda (Moral & Vicéns, 2003).

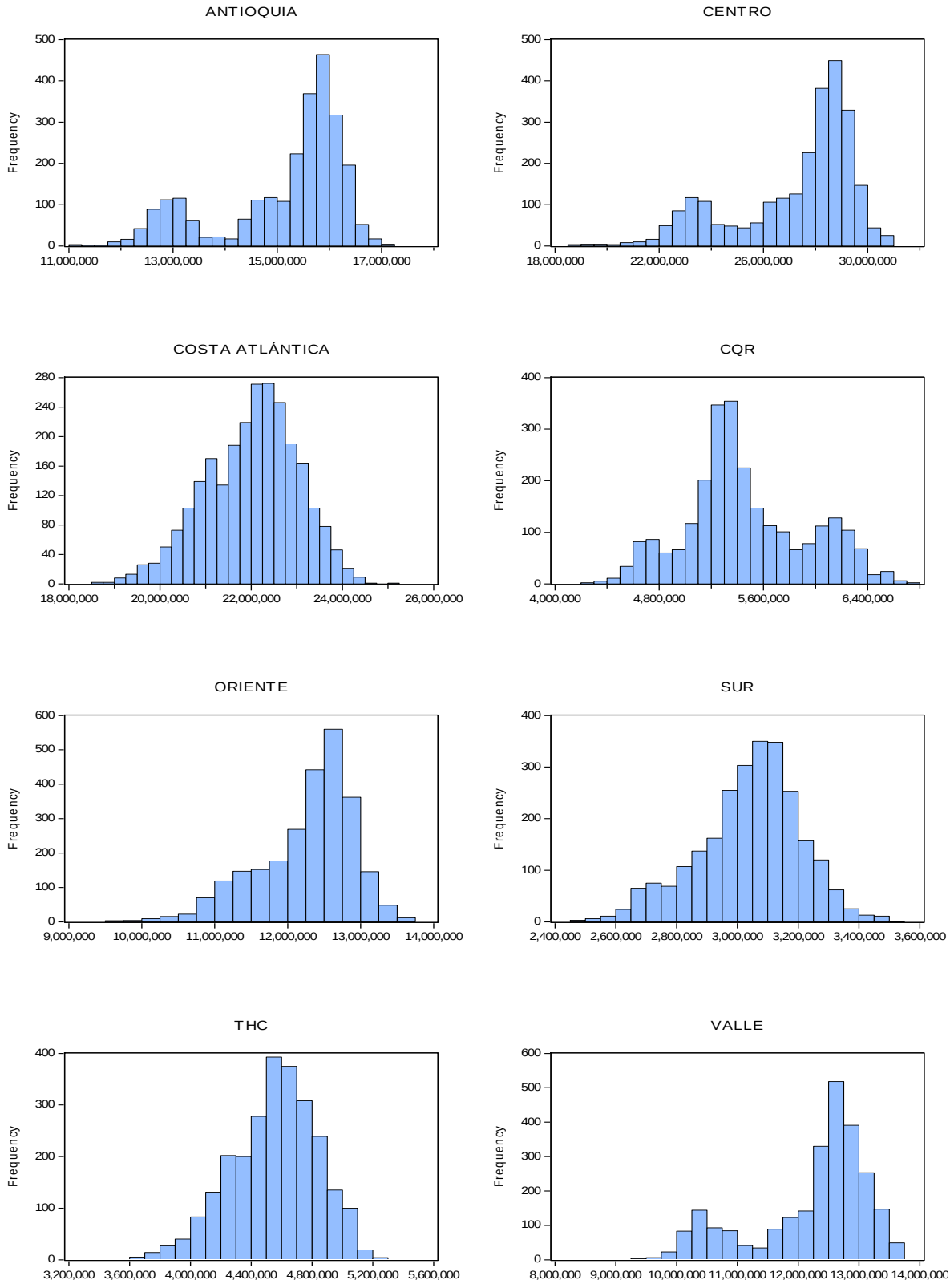
Por otra parte, tal y como también se observa en los histogramas de la **Figura 5.3**, a excepción del área operativa CQR (Caldas, Quindío y Risaralda), las series de demanda de energía eléctrica son de cola izquierda. Excepto en el área operativa Costa Atlántica, la asimetría aumenta a medida que disminuye la media histórica de la demanda de energía eléctrica. Así mismo, en dichos histogramas parece haber dos medias, particularmente en las áreas operativas Antioquia, Centro y Valle, una de los días ordinarios (lunes a viernes) y otra para los fines de semana y días festivos, esto último se conoce como efecto calendario, cuyo análisis no se realiza en el presente trabajo final de maestría, pero el cual podría ser relevante en futuras investigaciones en el proceso de modelamiento de los movimientos de la demanda de energía eléctrica en el corto plazo (Moral & Vicéns, 2003).

Tabla 5.2 Estadísticos descriptivos de las series de demanda de energía eléctrica.

	Demanda de Energía Eléctrica (kWh)							
Estadístico Descriptivo	**ANT**	**CEN**	**COS**	**CQR**	**ORI**	**SUR**	**THC**	**VAL**
Media	15133200	27256894	22015099	5462228.	12263562	3039157.	4552890.	12227024
Mediana	15614194	28149505	22112797	5374755.	12422509	3059152.	4578166.	12548615
Máximo	17228522	30947384	25095079	6706663.	13724915	3506071.	5597187.	13781808
Mínimo	11093279	18624522	18604266	4284709.	9577958.	2456591.	3543855.	8750776.
Desv. Tip.	1214917.	2315540.	1014144.	468445.3	651991.3	168692.1	284148.0	956770.4
Asimetría	-1.084042	-1.056127	-0.308979	0.292782	-0.865346	-0.43191	-0.33590	-1.005074
Curtosis	3.008066	3.167056	2.814401	2.602158	3.528760	3.140861	2.984943	2.984327
JB	500.8153	478.3216	44.35523	53.39483	348.9115	81.61618	48.10932	430.5283
Prob.	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

Fuente: Elaboración propia.

Figura 5.3 Histogramas de demanda de energía eléctrica.



Fuente: Elaboración propia a partir de Eviews.

5.2 Análisis de estacionariedad

5.2.1 Correlogramas

a. Indicador de temperatura media del aire

En la sección 8.9.1 del **Anexo 8.9** se presentan los correlogramas de las series de indicadores de temperatura media del aire de cada una de las áreas operativas en estudio para 36 retardos, con el objetivo de contrastar si las observaciones se encuentran autocorrelacionadas y si hay patrones estacionales en las series. En la **Figura 5.4** se presenta el correlograma del área operativa Antioquia con siete rezagos.

Figura 5.4 Correlograma del indicador de temperatura media del aire de Antioquia.

	AC	PAC	Q-Stat	Prob
1	0.682	0.682	1190.5	0.000
2	0.532	0.125	1914.8	0.000
3	0.438	0.067	2405.5	0.000
4	0.380	0.060	2775.1	0.000
5	0.330	0.029	3053.6	0.000
6	0.307	0.057	3295.2	0.000
7	0.287	0.039	3507.2	0.000

Fuente: Elaboración propia a partir de Eviews.

En la **Figura 5.4** se observa que los coeficientes de autocorrelación (AC) disminuyen lentamente mientras que los coeficientes de autocorrelación parcial (PAC) lo hacen de manera exponencial. Note que se espera que la autocorrelación sea mayor en los retardos 1 y 2, dado que la variable de análisis es la temperatura, y que disminuya para días posteriores. También es posible que esta serie presente estacionalidad anual. Vale la pena notar que las características identificadas en el correlograma de la **Figura 5.4** se observan en los correlogramas del indicador de temperatura media del aire de las demás áreas operativas. (**Véase Anexo 8.9**).

b. Demanda de energía eléctrica

Se efectuó el mismo procedimiento de análisis realizado a las series de indicador de temperatura media del aire para contrastar si las series de demanda de energía eléctrica de cada una de las áreas operativas se encuentran autocorrelacionadas y si hay patrones estacionales en éstas. En la sección 8.9.1 del **Anexo 8.9** se presentan los correlogramas de dichas series. Es importante notar que se identificaron dos tipos de correlogramas, en los cuales se observa que las series tienen estacionalidad semanal o efecto calendario. En las **Figura 5.5** y **Fuente:** Elaboración propia a partir de Eviews.

Figura 5.6 se presentan los correlogramas de las áreas operativas Antioquia y Costa Atlántica con siete rezagos.

Figura 5.5 Correlograma de la demanda de energía eléctrica de Antioquia.

	AC	PAC	Q-Stat	Prob
1	0.361	0.361	334.10	0.000
2	-0.104	-0.269	361.57	0.000
3	-0.237	-0.114	505.18	0.000
4	-0.252	-0.166	667.98	0.000
5	-0.157	-0.075	730.90	0.000
6	0.241	0.308	880.08	0.000
7	0.691	0.575	2105.6	0.000

Fuente: Elaboración propia a partir de Eviews.

Figura 5.6 Correlograma de la demanda de energía eléctrica de la Costa Atlántica.

	AC	PAC	Q-Stat	Prob
1	0.698	0.698	1246.7	0.000
2	0.472	-0.030	1816.8	0.000
3	0.384	0.128	2194.7	0.000
4	0.364	0.108	2534.7	0.000
5	0.412	0.194	2970.1	0.000
6	0.587	0.413	3854.1	0.000
7	0.765	0.462	5356.1	0.000

Fuente: Elaboración propia a partir de Eviews.

En el correlograma de la **Figura 5.5** se observa un patrón estacional, en el cual tanto la función (AC) como la (PAC) disminuyen de manera sinusoidal y alternando coeficientes positivos y negativos con agrupación de rezagos. Este comportamiento también se presenta en las áreas operativas Centro, Oriente y Valle. Es probable que dicho patrón sea resultado de la influencia de otras variables que explican el comportamiento de la demanda como la laboralidad, las temporadas vacacionales o el tipo de día de la semana (Moral & Vicéns, 2005). Véase sección 5.1.

Por otra parte, en el correlograma de la **Fuente:** Elaboración propia a partir de Eviews.

Figura 5.6 se observa un patrón distinto al identificado en las áreas operativas Antioquia, Centro, Oriente y Valle, ya que los coeficientes de la función (AC) no cambian de positivo a negativo y viceversa. Es posible que este comportamiento se deba a que en las áreas operativas CQR, Sur y THC se presentan las menores demandas de energía eléctrica del SIN (véase **Figura 5.3** y sección 5.1).

5.2.2 Test Dickey-Fuller Aumentado

a. Indicador de temperatura media del aire

Con el objetivo de determinar si las series son estacionarias o no y si las observaciones están correlacionadas entre sí, se realizó la prueba Dickey-Fuller Aumentado a las series de indicador de temperatura media del aire, a partir de los cuales fue posible establecer que dichas series son estacionarias hasta con siete rezagos y sus observaciones están correlacionadas entre sí. Vale la pena notar, que en el análisis de estacionariedad fueron considerados sólo siete rezagos, ya que en los

correlogramas de las series de demanda de energía eléctrica de todas las áreas operativas se observa un comportamiento estacional semanal. Los resultados de las pruebas se presentan en el **Anexo 8.10**.

b. Demanda de energía eléctrica

Se realizó el procedimiento descrito en el numeral a de la presente sección para determinar si las series de demanda de energía eléctrica son estacionarias o no y si sus observaciones están correlacionadas entre sí. Los resultados obtenidos indican que con excepción de la serie del área operativa CQR para seis rezagos, las demás series son estacionarias hasta con siete rezagos. Las series de demanda de energía eléctrica presentan autocorrelación negativa y positiva con cinco y seis rezagos respectivamente. Además, las series de las áreas operativas Costa Atlántica, CQR, Sur y THC evidencian autocorrelación negativa con cuatro rezagos y la serie del área operativa Oriente para el escenario con siete rezagos. Por otra parte, las áreas operativas Centro y Costa Atlántica están en zona de indecisión con siete rezagos, de manera que no es posible concluir sobre la autocorrelación de las observaciones en este caso.

En cuanto a la autocorrelación con cuatro y siete rezagos, el patrón no es claro, surgen dudas respecto a si estos dos componentes debieran incluirse en un modelo de regresión que explique el comportamiento de todas las áreas operativas. No obstante, en los casos con cinco y seis rezagos, el comportamiento es uniforme por lo que podría llegarse a concluir que se deberían incluir dichos componentes en dicho modelo de regresión. Los resultados de los test ADF de las series de demanda de energía eléctrica de cada una de las áreas operativas y para todos los rezagos considerados se presentan en el **Anexo 8.11**.

5.2.3 Test Phillips-Perron

a. Indicador de temperatura media del aire

A fin de contrastar los resultados obtenidos con la prueba DFA, se realizó la prueba Phillips-Perron a las series de indicador de temperatura media del aire. Los resultados obtenidos en cuanto a estacionariedad son exactamente iguales, pero respecto a la autocorrelación no lo son. Dichos resultados indican que las series analizadas tienen autocorrelación negativa para todos los rezagos considerados. En este sentido, es importante tener en cuenta que aunque ambas pruebas tienen el mismo objetivo, están basadas en análisis matemáticos diferentes, lo cual puede llevar a conclusiones contradictorias. Los resultados de los test PP de las series de indicador de temperatura media del aire de cada una de las áreas operativas y para todos los rezagos considerados se presentan en el **Anexo 8.12**.

b. Demanda de energía eléctrica

Al igual que en el caso de las series de indicador de temperatura media del aire, se realizó la prueba Phillips-Perron para contrastar los resultados obtenidos en la prueba DFA. En este caso, respecto a la estacionariedad de las series, la conclusión es la misma. La única diferencia se encuentra en la serie del área operativa CQR, la cual según la prueba PP es estacionaria con seis rezagos.

En cuanto a la autocorrelación, la prueba indica, excepto para la serie del área operativa Costa Atlántica, que las observaciones están autocorrelacionadas. Dicha relación es positiva para las áreas operativas Antioquia, Centro, Oriente, Sur y Valle; y negativa para las áreas operativas CQR y THC. Estos resultados indicarían que se deben incluir dichos componentes en el modelo de estudio. Los resultados de los test PP de las series de demanda de energía eléctrica de cada una de las áreas operativas y para todos los rezagos considerados se presentan en el **Anexo 8.13**.

5.2.4 Discusión sobre las pruebas de raíz unitaria

Dado que los resultados de los test ADF y PP indican que las series de indicador de temperatura media del aire y de demanda de energía eléctrica de cada una de las ocho áreas operativas que integran el Sistema Interconectado Nacional de Colombia son estacionarias integradas de orden cero, ya que no es necesario diferenciarlas, es relativamente fácil construir un modelo predictivo para la demanda de energía eléctrica.

Sin embargo, los resultados de las pruebas ADF y PP no coinciden siempre, lo cual puede deberse, como se dijo en el numeral a de la sección 5.2.3 a las características y restricciones de cada prueba y a que las series fueron pre procesadas.

En el caso del indicador de temperatura media del aire, la prueba ADF la prueba sugiere que la serie diaria de temperatura no presenta autocorrelación. Por otra parte, para la demanda de energía eléctrica hay correlación entre rezagos de diferente orden, dependiendo del área. Esto puede deberse a que la demanda tiene un comportamiento cíclico o estacional como se evidencia en los correlogramas para cada área operativa. Por lo anterior, en el estudio de las variables objeto de investigación, para efectos del presente trabajo final de maestría se consideran únicamente los resultados obtenidos en la prueba ADF.

5.3 Test de causalidad de Granger

Si bien la prueba de causalidad de Granger permite determinar si existe una relación de precedencia entre la variable explicativa y la variable respuesta (Gutiérrez Moya & Arévalo Quijada, 2003), antes de realizar dicha prueba es importante analizar la correlación lineal entre las series de demanda de energía eléctrica y del indicador de temperatura media del aire. Por lo anterior, en la **Tabla 5.3** se presentan los coeficientes de correlación lineal de Pearson para las ocho áreas operativas.

Tabla 5.3 Matriz de coeficientes de correlación entre las series en estudio.

	Indicador de temperatura media del aire							
	ANT	**CEN**	**COS**	**CQR**	**ORI**	**SUR**	**THC**	**VAL**
Demanda de energía eléctrica	0,1031	0,0720	0,4378	-0,0328	0,1329	0,1757	0,1810	0,1320

Fuente: Elaboración propia.

En la **Tabla 5.3** se observa que el coeficiente de correlación lineal entre la serie de indicador de temperatura media del aire y la serie de demanda de energía eléctrica más alto es el de la Costa Atlántica, lo cual sugiere una posible relación causal de la primera variable sobre la segunda. Por otra

parte, en las áreas operativas Centro y CQR, el coeficiente de correlación lineal de Pearson es cercano a cero, lo cual indica que podría no existir una relación de precedencia entre dichas variables. En las demás áreas operativas se espera que la prueba sugiera que la demanda de electricidad responde a los cambios en la temperatura.

La prueba de causalidad de Granger es una manera más cercana de establecer si una variable X, en este caso el indicador de temperatura media del aire, causa una variable Y, demanda de energía eléctrica. La aplicación de la prueba está sujeta a la estacionariedad de las series. En este sentido, en la sección 0 se determinó que las series de las variables objeto de investigación en las ocho áreas operativas son estacionarias integradas de orden cero. Note que la prueba de causalidad de Granger si bien es una condición necesaria no es suficiente para garantizar la existencia de una verdadera causalidad entre las variables (Montero Granados, 2013).

En el procedimiento de desarrollo la prueba de causalidad de Granger, se realiza una regresión de la variable endógena, en este caso de la demanda de energía eléctrica, sobre su propio pasado, sobre la variable exógena: el indicador de temperatura media del aire, y sobre el pasado de ésta última. Ya que se pretende determinar si los cambios en la temperatura modifican el comportamiento de la demanda de energía eléctrica, se realizó la prueba de causalidad de Granger a cada una de las áreas operativas con un rezago, ya que éste sólo incluye los efectos de la demanda de energía del día anterior y del valor actual y del día anterior del indicador de temperatura media del aire. Los resultados de la prueba de causalidad de Granger del indicador de temperatura media del aire sobre la demanda de energía eléctrica se presentan en la **Tabla 5.4** y en el **Anexo 8.14**.

Tabla 5.4 Resultados de la prueba de causalidad de Granger.

Resultados de la prueba de Causalidad de Granger								
	ANT	**CEN**	**COS**	**CQR**	**ORI**	**SUR**	**THC**	**VAL**
F-Statistic	16,4149	1,84755	90,0633	0,23022	12,457	34,5762	14,6247	36,8949
Probabilidad	0,00005	0,17420	0,00000	0,63140	0,00040	0,00000	0,00010	0,00000
¿Tem→Dem?	**Sí**	No	**Sí**	No	**Sí**	**Sí**	**Sí**	**Sí**

Fuente: Elaboración propia.

Los resultados obtenidos validan la conclusión del análisis de la matriz de coeficientes de correlación lineal de Pearson, ya que, excepto en las áreas operativas Centro y CQR, indican que las variaciones en la temperatura media del aire generan cambios en el comportamiento de la demanda de energía eléctrica. Esto último coincide con lo observado en países como España donde la demanda de energía eléctrica aumenta cuando la temperatura media del aire incrementa o disminuye respecto a un valor de referencia estimado (Moral & Vicéns, 2005). En Colombia XM S.A. E.S.P., empresa que opera el mercado eléctrico del país, ha publicado en su portal institucional que la demanda regulada de energía eléctrica ha aumentado por el incremento de la temperatura en varias regiones del país (XM Compañía de Expertos en Mercados S.A. E.S.P., 2015f), lo cual ha contribuido a un mayor consumo de energía para refrigeración y acondicionamiento del ambiente (XM Compañía de Expertos en Mercados S.A. E.S.P., 2014).

5.4 Modelos de regresión

En esta sección se exponen algunos análisis complementarios al alcance definido inicialmente en el presente trabajo final de maestría. Se propone un modelo de regresión para cada área operativa a fin de determinar cuán influyentes son las variaciones de la temperatura media del aire sobre la demanda de energía eléctrica de cada una de estas. Vale la pena resaltar que dichos modelos de regresión no están concebidos como una herramienta de predicción sino como un instrumento de diagnóstico. No obstante, a fin de establecer conclusiones veraces y útiles para el desarrollo de futuros proyectos de investigación, se evalúa si tanto los residuos en niveles como al cuadrado presentan autocorrelación, ya que esto es indicativo de la ausencia de ruido blanco en los mismos y de la presencia de heterocedasticidad condicional autorregresiva (ARCH) en los distintos modelos.

Para la formulación de los modelos de regresión se tuvo en cuenta que los histogramas de las series de demanda de energía eléctrica indican que en algunas áreas operativas parecen haber dos medias, una relacionada con los días ordinarios y otra con los festivos, lo cual sugiere que la laboralidad y el efecto calendario podrían explicar el comportamiento de la demanda de energía eléctrica (Climent Diranzo et al., 2003). Por otra parte, a partir de los correlogramas de las series de demanda de electricidad se estableció cuáles son los rezagos susceptibles de ser considerados en cada modelo econométrico. De igual forma, los resultados del test de causalidad de Granger también fueron considerados con el fin de no incluir la temperatura como variable explicativa en las áreas operativas donde no hay evidencia de relación causal de la temperatura media del aire sobre la demanda de energía eléctrica.

En la **Tabla 5.5** se presentan las expresiones funcionales de los modelos base de cada área operativa, cuyo horizonte temporal de estudio contiene información diaria desde el 01:01:2007 hasta el 31:12:2013. Las variables demanda con rezagos están expresadas en unidades de kWh. En dicho modelos no fue necesario realizar la prueba de variables redundantes ya que todos los coeficientes de regresión resultaron ser significativos. Por otra parte, los resultados del contraste Breusch-Godfrey-LM y los correlogramas de las series de demanda de energía eléctrica indicaron que debían incluirse el componente AR(1) y MA(1) en las áreas operativas Costa Atlántica y CQR (Véase **Anexo 8.15**).

En la **Tabla 5.6** se presentan las variables de los modelos de regresión definidos en la **Tabla 5.5**.

Tabla 5.5 Modelos de demanda de energía eléctrica de cada una de las áreas operativas.

Área Operativa	Expresión funcional
Antioquia	$D_t = C + \alpha_1 D_{t-1} + \alpha_2 D_{t-6} + \alpha_3 D_{t-7} + \alpha_4 T_t + \alpha_5 Lun + \alpha_6 Mar + \alpha_7 Mie + \alpha_8 Jue + \alpha_9 Vie + \alpha_{10} Sab$
Centro	$D_t = C + \alpha_1 D_{t-1} + \alpha_2 D_{t-3} + \alpha_3 D_{t-6} + \alpha_4 D_{t-7} + \alpha_5 Lun + \alpha_6 Mar + \alpha_7 Mie + \alpha_8 Jue + \alpha_9 Vie + \alpha_{10} Sab$
Costa Atlántica	$D_t = \alpha_1 D_{t-1} + \alpha_2 D_{t-2} + \alpha_3 D_{t-3} + \alpha_4 D_{t-7} + \alpha_5 T_t + \alpha_6 Lun + \alpha_7 Mar + \alpha_8 Mie + \alpha_9 Jue + \alpha_{10} Vie + \alpha_{11} Sab + ar(1) + ma(1)$
CQR	$D_t = \alpha_1 D_{t-1} + \alpha_2 D_{t-3} + \alpha_3 D_{t-6} + \alpha_4 D_{t-7} + \alpha_5 Lun + \alpha_6 Mar + \alpha_7 Mie + \alpha_8 Jue + \alpha_9 Vie + \alpha_{10} Sab + ar(1) + ma(1)$
Oriente	$D_t = \alpha_1 D_{t-1} + \alpha_2 D_{t-2} + \alpha_3 D_{t-5} + \alpha_4 D_{t-7} + \alpha_5 T_t + \alpha_6 Lun + \alpha_7 Mar + \alpha_8 Mie + \alpha_9 Jue + \alpha_{10} Vie + \alpha_{11} Sab$
Sur	$D_t = \alpha_1 D_{t-1} + \alpha_2 D_{t-2} + \alpha_3 D_{t-3} + \alpha_4 D_{t-5} + \alpha_5 D_{t-6} + \alpha_6 D_{t-7} + \alpha_7 T_t + \alpha_8 Lun + \alpha_9 Mar + \alpha_{10} Mie + \alpha_{11} Jue + \alpha_{12} Vie + \alpha_{13} Sab$
THC	$D_t = \alpha_1 D_{t-1} + \alpha_2 D_{t-2} + \alpha_3 D_{t-3} + \alpha_4 D_{t-5} + \alpha_5 D_{t-7} + \alpha_6 T_t + \alpha_7 Lun + \alpha_8 Mar + \alpha_9 Mie + \alpha_{10} Jue + \alpha_{11} Vie + \alpha_{12} Sab$
Valle	$D_t = C + \alpha_1 D_{t-1} + \alpha_2 D_{t-6} + \alpha_3 D_{t-7} + \alpha_4 T_t + \alpha_5 Lun + \alpha_6 Mar + \alpha_7 Mie + \alpha_8 Jue + \alpha_9 Vie + \alpha_{10} Sab$

Fuente: Elaboración propia.

Tabla 5.6 Variables utilizadas en los modelos de regresión.

Variable	Definición
D_t	Demanda de energía eléctrica en cada área operativa en el día t.
C	Constante que se interpreta como la media de la demanda de energía en los modelos que la incluyen.
D_{t-i}	Demanda de energía eléctrica en cada área operativa en el día $t-i; i = 1 \ldots 7$
T_t	Registro del indicador de temperatura media del aire en el día t.
Lun	Dummy que toma un valor de uno cuando el día t es un lunes.
Mar	Dummy que toma un valor de uno cuando el día t es un martes.
Mie	Dummy que toma un valor de uno cuando el día t es un miércoles.
Jue	Dummy que toma un valor de uno cuando el día t es un jueves.
Vie	Dummy que toma un valor de uno cuando el día t es un viernes.
Sab	Dummy que toma un valor de uno cuando el día t es un sábado.
ar	Componente autorregresivo
ma	Componente de media móvil

Fuente: Elaboración propia.

Los correlogramas de los errores en niveles y al cuadrado de los modelos planteados fueron analizados con el fin de determinar si eran o no correctos para modelar la demanda de cada una de las áreas operativas. Los resultados indican que si bien los errores se comportan como ruido blanco, los modelos presentan heterocedasticidad condicional autorregresiva o efecto ARCH, por lo que deben ser modificados mediante la estimación de la volatilidad, en este caso, la asociada con la demanda de energía eléctrica. Para ello, en primer lugar se realizó un análisis gráfico de los modelos planteados a fin de determinar si era necesario o no incluir variables adicionales que permitieran agrupar comportamientos o eliminar observaciones atípicas, ya que los outliers transmiten su efecto sobre la volatilidad, lo cual ocasiona que los errores subsiguientes tiendan exponencialmente a su media anterior (Carnero, Peña, & Ruiz, 2012). La evaluación gráfica permitió establecer que se debían incluir variables adicionales, dummies, para capturar comportamientos particulares de la demanda en cada una de las áreas operativas como por ejemplo la eliminación de outliers y la representación de días con alta y baja demanda. En segundo lugar, se evaluaron el correlograma de los residuos en niveles y al cuadrado para determinar si debían incluirse componentes AR(p) y MA(p) y eliminarse variables antes incluidas, así como también para validar que los errores se comportaran como ruido blanco y que los modelos no presentaran efecto ARCH, para ello fue necesario modelar la volatilidad en cada uno de los modelos. En la **Tabla 5.7** se presentan las expresiones funcionales modificadas de los modelos de cada área operativa.

Tabla 5.7 Modelos de demanda de energía eléctrica modificados.

Área Operativa	Expresión funcional
Antioquia	$D_t = C + \alpha_1 D_{t-1} + \alpha_2 D_{t-7} + \alpha_3 T_t + \alpha_4 Lun + \alpha_5 Mar + \alpha_6 Mie + \alpha_7 Jue + \alpha_8 Vie + \alpha_9 Sab + \alpha_{10} Alta + \alpha_{11} Outliers + \alpha_{12} Ar(6)$
Centro	$D_t = C + \alpha_1 D_{t-1} + \alpha_2 D_{t-6} + \alpha_3 Lun + \alpha_4 Mar + \alpha_5 Mie + \alpha_6 Jue + \alpha_7 Vie + \alpha_8 Sab + \alpha_9 Alta + \alpha_{10} Baja + \alpha_{11} Ar(1) + \alpha_{12} Ar(7)$
Costa Atlántica	$D_t = C + \alpha_1 D_{t-1} + \alpha_2 D_{t-2} + \alpha_3 D_{t-3} + \alpha_4 D_{t-4} + \alpha_5 D_{t-7} + \alpha_6 T_t + \alpha_7 Lun + \alpha_8 Mar + \alpha_9 Mie + \alpha_{10} Jue + \alpha_{11} Vie + \alpha_{12} Sab + \alpha_{13} Ar(7)$
CQR	$D_t = \alpha_1 D_{t-1} + \alpha_2 D_{t-3} + \alpha_3 D_{t-6} + \alpha_4 D_{t-7} + \alpha_5 Lun + \alpha_6 Mar + \alpha_7 Mie + \alpha_8 Jue + \alpha_9 Vie + \alpha_{10} Sab$
Oriente	$D_t = \alpha_1 D_{t-1} + \alpha_2 D_{t-2} + \alpha_3 D_{t-5} + \alpha_4 D_{t-7} + \alpha_5 T_t + \alpha_6 Lun + \alpha_7 Mar + \alpha_8 Mie + \alpha_9 Jue + \alpha_{10} Vie + \alpha_{11} Sab + \alpha_{12} Alta + \alpha_{13} Ar(1)$
Sur	$D_t = \alpha_1 D_{t-1} + \alpha_2 D_{t-2} + \alpha_3 D_{t-3} + \alpha_4 D_{t-5} + \alpha_5 D_{t-6} + \alpha_6 D_{t-7} + \alpha_7 T_t + \alpha_8 Lun + \alpha_9 Mar + \alpha_{10} Mie + \alpha_{11} Jue + \alpha_{12} Vie + \alpha_{13} Sab + \alpha_{14} Alta + \alpha_{15} Baja + \alpha_{16} Outliers + \alpha_{17} Ar(6)$
THC	$D_t = \alpha_1 D_{t-1} + \alpha_2 D_{t-2} + \alpha_3 D_{t-3} + \alpha_4 D_{t-5} + \alpha_5 D_{t-7} + \alpha_6 T_t + \alpha_7 Lun + \alpha_8 Mar + \alpha_9 Mie + \alpha_{10} Jue + \alpha_{11} Vie + \alpha_{12} Sab + \alpha_{13} Alta + \alpha_{14} Outliers + \alpha_{15} Ar(2)$
Valle	$D_t = C + \alpha_1 D_{t-1} + \alpha_2 T_t + \alpha_3 Lun + \alpha_4 Mar + \alpha_5 Mie + \alpha_6 Jue + \alpha_7 Vie + \alpha_8 Sab + \alpha_9 Alta + \alpha_{10} Outliers + \alpha_{11} Ar(2)$

Fuente: Elaboración propia.

En la siguiente sección se presenta el análisis estructural de los modelos. Para ello primero se evalúa la coherencia de los signos y la significancia estadística de los coeficientes de regresión de todas las variables consideradas en cada uno de los modelos de las ocho áreas operativas.

5.4.1 Análisis estructural de los modelos

A fin de validar cada modelo de regresión, a continuación se presenta un análisis estructural de cada uno, en el cual se evalúa la consistencia de los coeficientes de regresión, así como también su significancia conjunta.

a. Coeficientes de regresión

En primer lugar se analiza si los signos de los coeficientes de regresión son coherentes, es decir si existe una relación directa o inversa entre las variables explicativas y la demanda de energía eléctrica En el modelo de regresión seleccionado se espera que la asociación de la constante, la temperatura, la demanda de electricidad y las variables dummy utilizadas para analizar el efecto día en relación con la variable respuesta, sea como se describe a continuación.

- **Constante:** Este término incorpora de cierta manera la demanda media de energía eléctrica. Se espera entonces que la relación sea positiva. Sin embargo, en caso de ser necesario modelar la volatilidad, dicha relación podría ser negativa ya que la ecuación de volatilidad también incluye un término constante que representa la media o el valor en torno al cual se presentan las variaciones, lo que podría contribuir a que la constante del modelo de regresión sea inferior a cero.
- **Temperatura:** Teniendo en cuenta que con el aumento en la temperatura media del aire se espera un aumento en la demanda de energía eléctrica, la relación de este componente y la variable respuesta debe ser positiva.
- **Demanda de electricidad con i rezagos:** La relación podría ser positiva o negativa ya que la demanda podría aumentar o disminuir respecto a días anteriores sin patrón alguno en cada área operativa.
- **Variables Dummy (Lunes a Sábado):** Variables ficticias que incorporan el efecto de la demanda de energía de cada uno de los días de la semana. Por lo anterior, la relación esperada debe ser positiva.
- **Variables Dummy (Outliers):** Tienen como objetivo excluir valores atípicos de la serie de tiempo. Por tanto, la relación puede ser positiva o negativa según sea el caso.
- **Variables Dummy (Alta):** Dado que tienen como objetivo capturar el comportamiento de los días con mayor demanda, la relación debe ser positiva respecto a la media histórica.
- **Variables Dummy (Baja):** Dado que tienen como objetivo capturar el comportamiento de los días con menor demanda, la relación debe ser negativa respecto a la media histórica.

En la **Tabla 5.8** se muestran los coeficientes de regresión de los modelos para las ocho áreas operativas y el $T - Statistic$. Nótese que en algunos modelos se incluyeron o excluyeron variables respecto a los modelos inicialmente considerados.

Tabla 5.8 Coeficientes de regresión de los modelos modificados.

Área Operativa	Stat	C	Tem	D(-1)	D(-2)	D(-3)	D(-4)	D(-5)	D(-6)	D(-7)	Lun	Mar	Mie	Jue	Vie	Sab	Alta	Outliers	Baja	Ar(1)	Ar(2)	Ar(3)	Ar(6)	Ar(7)
Antioquia	**Coe.**	8820314	18537	0,25	---	---	---	---	---	0,04	2315391	1958930	1958013	1918957	1875602	1107525	495625	-1613720	---	---	---	---	0,03	---
	Prob.	0,00	0,00	0,00	---	---	---	---	---	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	---	---	---	---	0,03	---
Centro	**Coe.**	4682274	---	0,69	---	---	---	---	0,03	---	6173609	3777670	3628119	3605640	3532702	1623021	678327	---	-	0,00	---	---	---	0,04
	Prob.	0,00	---	0,00	---	---	---	---	0,00	---	0,00	0,00	0,00	0,00	0,00	0,00	0,00	---	0,00	0,74	---	---	---	0,00
Costa Atlántica	**Coe.**	-2324472	215908	0,52	0,09	-0,24	0,24	---	-0,02	0,19	1356669	918571	748585	1152929	882074	286210	---	---	---	---	---	0,32	---	---
	Prob.	0,00	0,00	0,00	0,00	0,00	0,00	---	0,51	0,00	0,00	0,00	0,00	0,00	0,00	0,00	---	---	---	---	---	0,00	---	---
CQR	**Coe.**	---	---	0,63	---	0,05	---	---	0,06	0,19	647826	420219	423112	372673	395080	262794	---	---	---	---	---	---	---	---
	Prob.	---	---	0,00	---	0,00	---	---	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	---	---	---	---	---	---	---	---
Oriente	**Coe.**	---	71603	0,50	0,12	---	---	0,09	---	0,09	1292201	1044974	888829	853726	914818	459290	279941	---	---	0,08	---	---	---	---
	Prob.	---	0,00	0,00	0,00	---	---	0,00	---	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	---	---	0,14	---	---	---	---
Sur	**Coe.**	---	24153	0,43	0,08	0,07	---	0,08	0,08	0,08	253002	216916	200461	180249	194723	107850	86194	---	-131324	0,08	---	---	---	0,06
	Prob.	---	0,00	0,00	0,00	0,00	---	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	---	0,00	0,03	---	---	---	0,02
THC	**Coe.**	---	9191035	0,50	---	0,13	0,05	0,06	0,01	0,14	416204	325067	350949	332343	340504	207402	104265	-85305	---	---	0,09	---	---	---
	Prob.	---	0,00	0,00	---	0,00	0,00	0,00	0,65	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	---	---	0,00	---	---	---
Valle	**Coe.**	5363768	45257	0,35	---	---	---	---	---	---	1871062	1763802	1758030	1679917	1706216	955996	443729	-563006	---	---	---	---	---	0,17
	Prob.	0,00	0,00	0,00	---	---	---	---	---	---	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	---	---	---	---	---	0,00

Fuente: Elaboración propia a partir de Eviews.

Tabla 5.9 Modelos de volatilidad de las áreas operativas. Excepto Centro.

Área Operativa	Stat	C	Componentes ARCH							Componentes GARCH						
			RED(-1)^2	RES(-2)^2	RES(-3)^2	RES(-4)^2	RES(-5)^2	RES(-6)^2	RES(-7)^2	GARCH(-1)	GARCH(-2)	GARCH(-3)	GARCH(-4)	GARCH(-5)	GARCH(-6)	GARCH(-7)
Antioquia	**Coe.**	1,35E+11	0,54	---	---	---	---	---	---	0,14	-0,05	0,01	-0,02	0,00	0,08	-0,06
	Prob.	0,00	0,00	---	---	---	---	---	---	0,00	0,00	0,00	0,00	0,68	0,00	0,00
Costa Atlántica	**Coe.**	1.16E+11	0,23	-0,03	0,00	-0,00	-0,02	-0,00	0,21	---	---	---	---	---	---	---
	Prob.	0.0000	0,00	0,0000	0,98	0,28	0,00	0,24	0,00	---	---	---	---	---	---	---
CQR	**Coe.**	9,11E+09	0,32	---	---	---	---	---	---	-0,06	-0,02	-0,04	0,01	-0,05	0,38	0,11
	Prob.	0,00	0,00	---	---	---	---	---	---	0,00	0,03	0,00	0,23	0,00	0,00	0,00
Oriente	**Coe.**	5,42E+10	0,27	---	---	---	---	---	---	0,13	-0,07	0,05	-0,02	0,13	0,15	-0,11
	Prob.	0,00	0,00	---	---	---	---	---	---	0,00	0,00	0,07	0,29	0,00	0,00	0,00
Sur	**Coe.**	1,82E+09	0,30	---	---	---	---	---	---	0,04	-0,03	0,02	-0,02	0,09	0,24	-0,02
	Prob.	0,00	0,00	---	---	---	---	---	---	0,36	0,32	0,50	0,33	0,02	0,00	0,62
THC	**Coe.**	4,65E+09	0,39	---	---	---	---	---	---	0,05	-0,02	-0,03	-0,01	0,06	0,24	-0,07
	Prob.	0,00	0,00	---	---	---	---	---	---	0,06	0,14	0,00	0,42	0,00	0,00	0,00
Valle	**Coe.**	9,88E+10	0,16	0,01	0,01	-0,01	-0,01	0,00	0,30	-0,16	-0,01	-0,05	-0,01	-0,05	-0,03	0,34
	Prob.	0,00	0,00	0,23	0,00	0,00	0,00	0,30	0,00	0,00	0,47	0,00	0,05	0,00	0,00	0,00

Fuente: Elaboración propia a partir de Eviews.

Nótese que en la **Tabla 5.9** no se presenta el modelo de volatilidad del área operativa centro, ya que ésta fue estimada mediante el método EGARCH. Los resultados para dicha área operativa se presentan en la **Tabla 5.10**.

Tabla 5.10 Modelos de volatilidad del área operativa Centro.

Coeficientes de Regresión		
Variable	**C**	**Prob**
C(14)	7,61	0,00
C(15)	0,75	0,00
C(16)	-0,90	0,00
C(17)	0,19	0,00
C(18)	-0,01	0,43
C(19)	-0,36	0,00
C(20)	1,00	0,00
C(21)	-0,67	0,00
C(22)	-0,22	0,00
C(23)	1,25	0,00
C(24)	-0,67	0,00
C(25)	0,23	0,00
C(26)	0,41	0,00
C(27)	-1,13	0,00
C(28)	1,16	0,00
C(29)	-0,53	0,00

Fuente: Elaboración propia a partir de Eviews.

En la **Tabla 5.8** se observa los signos de las variables consideradas en los distintos modelos de regresión son consistentes con las relaciones esperadas. Por otra parte, si bien los términos AR(1) de las áreas Centro y Oriente y Dem(-6) de las áreas Costa Atlántica y THC parecer no ser significativos, las pruebas de variables redundantes indican lo contrario (**Véase Anexo 8.16**), de manera que todas las variables consideradas son útiles para representar el comportamiento de la demanda.

En la **Tabla 5.11** se presenta el resumen de las ecuaciones de volatilidad de cada área operativa. En ella se detalla el modelo de estimación y su correspondiente orden, se presenta el coeficiente que denota la media en torno a la cual se presentan las variaciones y la sumatoria de los coeficientes de los componentes ARCH y GARCH.

Tabla 5.11 Resumen de los modelos de volatilidad.

Área Operativa	Modelo	Coeficiente C	$\sum C_{i_{ARCH}} + C_{i_{GARCH}}$	$\sum C_{i_{EGARCH}}$
Antioquia	GARCH (1,7)	1,35E+11	0,64	---
Centro	EGARCH (7,7)	--	--	8,11
Costa Atlántica	GARCH (7,0)	1,16E+11	0,39	---
CQR	GARCH (1,7)	9,11E+09	0,65	---
Oriente	GARCH (1,7)	5,42E+10	0,53	---
Sur	GARCH (1,7)	1,82E+09	0,63	---
THC	GARCH (1,7)	4,65E+09	0,61	---
Valle	GARCH (7,7)	9,88E+10	0,49	---

Fuente: Elaboración propia.

Nótese que el término $\sum C_{i_{EGARCH}}$ sólo aplica al área operativa Centro, dado que la volatilidad fue estimada mediante un modelo EGARCH, ya que mediante los modelos ARCH y GARCH no fue posible garantizar ausencia de autocorrelación en los errores en niveles y al cuadrado. En las demás áreas, la volatilidad se estimó a través de modelos GARCH (p, q). No obstante, en algunas de ellas uno o más de los coeficientes son negativos (**Véase Tabla 5.11**), por lo que no se puede asegurar que las varianzas condicionales sean positivas en todos los periodos (Sosa et al., 2014). Sin embargo, según Sosa et al. (2014) algunos coeficientes de un modelo GARCH (p. q) pueden ser negativos y aún así cumplirse que las varianzas condicionales sean positivas (Nelson & Cao, 1992). Es importante resaltar que en la **Tabla 5.11** se observa que la sumatoria de los componentes ARCH y GARCH es inferior a la unidad, lo que indica que la volatilidad estimada no es explosiva y por tanto que el proceso es estacionario. Respecto al nivel de significancia de los coeficientes, algunos de ellos resultan no ser significativos. No obstante, cuando se estiman modelos de orden superior al GARCH (1,1) algunos coeficientes de la ecuación de varianza pueden ser no significativos (Tecano Molano, 2010). Por lo anterior, dado que los modelos planteados no tienen fines de pronóstico, a que los análisis planteados en la presente sección son complementarios y a que de acuerdo con los resultados presentados en la **Tabla 5.11** se puede concluir que el proceso es estacionario en varianza (Sosa et al., 2014), se asume que los modelos de regresión planteados son correctos.

b. Significancia conjunta del modelo

En la presente sección, se analiza la significancia conjunta de cada modelo. En primer lugar, se evalúa el $\boldsymbol{T-Statistic}$, a fin de determinar si se puede asegurar que los coeficientes de todas las variables consideradas en la regresión no son nulas. En la **Tabla 5.12** se presentan los parámetros estadísticos de los modelos de todas las áreas operativas que conforman el Sistema Interconectado Nacional.

Tabla 5.12 Parámetros estadísticos de los modelos de cada área operativa.

	Área Operativa							
Estadístico	**ANT**	**CEN**	**COS**	**CQR**	**ORI**	**SUR**	**THC**	**VAL**
R-squared	0.81	0.77	0.82	0.88	0.73	0.84	0.87	0.80
Adj R-Squared	0.81	0.76	0.82	0.88	0.73	0.83	0.87	0.80
S.E. Regression	535450	1120581	428283	160413	339929	68744	103622	425084
Sum Squared	7.27E+14	3.19E+15	4.65E+14	6.54E+13	2.93E+14	1.20E+13	2.72E+13	4.60E+14
Log likehood	-36,976	-38210	-36529	-33900	-35905	-31885	-32872	-36344
Mean Dep. Var	15136236	27266440	22011785	5460794	12266834	3039253	4553110	12228180
S.D. Dependent Var	1213346	2307996	1013122	467964	649517	168808	28441	955169
Akaike Criterion	29.02	29.98	28.67	26.60	28.18	25.03	25.80	28.46
Schwarz Criterion	29.07	30.05	28.72	26.65	28.23	25.09	25.86	28.52
Hannan-Quinn Cri.	29.04	30.00	28.69	26.62	28.20	25.05	25.82	28.48
DW Statistic	2.13	2.44	2.12	2.15	2.14	2.09	2.03	2.10

Fuente: Elaboración propia a partir de Eviews.

Dado que cada área operativa está representada por un modelo diferente, con distinto número de variables explicativas, y a que todos tratan distintas muestras no se deben comparar los modelos entre sí. Sin embargo, sí es posible presentar una caracterización general de los mismos.

De acuerdo con los valores del R cuadrado y del R cuadrado ajustado, se puede afirmar que los modelos incorporan entre el 73 y 88 % de la variación de la demanda de energía eléctrica. Nótese que

en las áreas operativas Oriente y Centro los modelos explican el 73 y 77 % de las variaciones respectivamente; mientras que en THC y CQR, los modelos recogen el 87 y 88 % de los cambios en la demanda. Es probable, que estas diferencias sean resultado de otros elementos que no se pueden explicar con los modelos como por ejemplo las principales actividades económicas desarrolladas en cada región, la población de los municipios localizados en cada área operativa, entre otros. Es importante tener en cuenta el modelo no presenta problemas de multicolinealidad, ya que si bien en todas las áreas el R cuadrado y el R cuadrado ajustado son moderadamente elevados (mayores a 0,7), todas las variables explicativas son significativas (véase numeral *a* de la presente sección).

En relación con el error estándar de la regresión (S.E. of Regression), el cual es una medida de la varianza de los residuales, se concluye que la varianza es proporcional a la demanda de cada área operativa. Este comportamiento se observa en los demás estadísticos presentados en la **Tabla 5.12**.

c. Contrastes de diagnóstico para evaluación de los modelos

De acuerdo con lo planteado en los numerales a y b de la presente sección, los modelos definidos resultaron ser significativos al 95% de confianza con las variables explicativas consideradas en cada uno de ellos, a pesar de que fue necesario realizar el contraste de variables irrelevantes para validar la significancia de los términos AR(1) de las áreas Centro y Oriente, y Dem(-6) en las áreas Costa Atlántica y THC.

Por otra parte, a fin de analizar el ajuste del ruido blanco, a continuación se presentan los resultados del análisis de los residuos en niveles y al cuadrado para cada uno de los modelos de las ocho áreas operativas que conforman el Sistema Interconectado Nacional (SIN). Es importante aclarar, que los correlogramas se analizan hasta con siete rezagos, ya que la demanda de energía eléctrica presenta estacionalidad semanal (Véase **numeral b de la Sección 5.2.1**).

- **Análisis de los residuos en niveles**

El análisis de los correlogramas de los residuos en niveles permite determinar, para los siete retardos, si éstos se encuentran autocorrelacionados (Gutiérrez Moya & Arévalo Quijada, 2003). A continuación se presentan los correlogramas de los residuos en niveles para las ocho áreas operativas.

Tabla 5.13 Correlograma de los residuos. Antioquia.

	AC	PAC	Q-Stat	Prob*
1	-0.009	-0.009	0.2123	
2	0.013	0.013	0.6466	0.421
3	-0.003	-0.003	0.6671	0.716
4	-0.010	-0.010	0.9348	0.817
5	0.005	0.005	1.0008	0.910
6	-0.034	-0.033	3.8939	0.565
7	0.032	0.031	6.4668	0.373

Fuente: Elaboración propia a partir de Eviews.

Tabla 5.14 Correlograma de los residuos. Centro.

	AC	PAC	Q-Stat	Prob*
1	-0.003	-0.003	0.0208	
2	-0.010	-0.010	0.2602	
3	-0.003	-0.003	0.2813	0.596
4	0.027	0.027	2.1178	0.347
5	-0.006	-0.006	2.2198	0.528
6	-0.007	-0.006	2.3290	0.675
7	0.057	0.057	10.665	0.058

Fuente: Elaboración propia a partir de Eviews.

Tabla 5.15 Correlograma de los residuos. Costa Atlántica.

	AC	PAC	Q-Stat	Prob*
1	-0.011	-0.011	0.2983	
2	0.022	0.022	1.5592	0.212
3	0.013	0.014	2.0113	0.366
4	-0.011	-0.011	2.3327	0.506
5	0.052	0.051	9.2907	0.054
6	-0.015	-0.014	9.8684	0.079
7	0.027	0.025	11.701	0.069

Fuente: Elaboración propia a partir de Eviews.

Tabla 5.16 Correlograma de los residuos. CQR.

	AC	PAC	Q-Stat	Prob*
1	0.024	0.024	1.4780	0.224
2	-0.007	-0.008	1.6163	0.446
3	-0.018	-0.017	2.4267	0.489
4	0.023	0.024	3.7483	0.441
5	0.044	0.042	8.6183	0.125
6	-0.007	-0.009	8.7326	0.189
7	-0.024	-0.022	10.175	0.179

Fuente: Elaboración propia a partir de Eviews.

Tabla 5.17 Correlograma de los residuos. Oriente.

	AC	PAC	Q-Stat	Prob*
1	-0.013	-0.013	0.4297	
2	0.002	0.002	0.4410	0.507
3	0.026	0.026	2.2262	0.329
4	0.024	0.025	3.7273	0.292
5	-0.025	-0.024	5.3221	0.256
6	0.027	0.026	7.2189	0.205
7	0.031	0.031	9.6663	0.139

Fuente: Elaboración propia a partir de Eviews.

Tabla 5.18 Correlograma de los residuos. Sur.

	AC	PAC	Q-Stat	Prob*
1	-0.006	-0.006	0.0906	
2	0.018	0.018	0.9410	
3	-0.009	-0.008	1.1274	0.288
4	0.023	0.022	2.4533	0.293
5	-0.019	-0.018	3.3699	0.338
6	-0.027	-0.028	5.2592	0.262
7	0.025	0.026	6.8250	0.234

Fuente: Elaboración propia a partir de Eviews.

Tabla 5.19 Correlograma de los residuos. THC.

	AC	PAC	Q-Stat	Prob*
1	0.036	0.036	3.3090	
2	0.007	0.005	3.4228	0.064
3	0.013	0.012	3.8441	0.146
4	0.007	0.006	3.9531	0.267
5	-0.015	-0.016	4.5507	0.337
6	-0.001	-0.000	4.5541	0.473
7	-0.003	-0.003	4.5756	0.599

Fuente: Elaboración propia a partir de Eviews.

Tabla 5.20 Correlograma de los residuos. Valle.

	AC	PAC	Q-Stat	Prob*
1	0.000	0.000	4.E-05	
2	0.002	0.002	0.0092	0.924
3	-0.031	-0.031	2.4166	0.299
4	0.001	0.001	2.4183	0.490
5	0.011	0.011	2.7339	0.603
6	0.017	0.016	3.4406	0.632
7	-0.001	-0.001	3.4449	0.751

Fuente: Elaboración propia a partir de Eviews.

De acuerdo con los valores del estadístico Q de Ljung-Box y de sus correspondientes probabilidades presentados de la **Tabla 5.13** a la **Tabla 5.20**, se puede concluir con un 95% de confianza que los residuos hasta con siete rezagos no están autocorrelacionados entre sí, lo que indica que los modelos planteados son correctos y que por tanto los residuos se comportan con arreglo a un ruido blanco.

- **Análisis de los residuos al cuadrado**

El análisis de los correlogramas de los residuos al cuadrado permite determinar si existe heterocedasticidad condicional autorregresiva (ARCH) en los residuos (Gutiérrez Moya & Arévalo Quijada, 2003). A continuación se presentan los correlogramas de los residuos al cuadrado para las ocho áreas operativas.

Tabla 5.21 Correlograma de los residuos al cuadrado. Antioquia.

	AC	PAC	Q-Stat	Prob*
1	-0.006	-0.006	0.0823	0.774
2	-0.043	-0.043	4.8204	0.090
3	-0.006	-0.006	4.9072	0.179
4	0.011	0.009	5.2220	0.265
5	-0.016	-0.016	5.8745	0.319
6	-0.008	-0.007	6.0271	0.420
7	0.022	0.021	7.3055	0.398

Fuente: Elaboración propia a partir de Eviews.

Tabla 5.22 Correlograma de los residuos al cuadrado. Centro.

	AC	PAC	Q-Stat	Prob*
1	0.001	0.001	0.0013	0.971
2	-0.018	-0.018	0.7906	0.673
3	0.011	0.011	1.1128	0.774
4	0.005	0.004	1.1671	0.883
5	-0.000	0.000	1.1672	0.948
6	-0.026	-0.026	2.9573	0.814
7	0.021	0.021	4.0659	0.772

Fuente: Elaboración propia a partir de Eviews.

Tabla 5.23 Correlograma de los residuos al cuadrado. Costa Atlántica.

	AC	PAC	Q-Stat	Prob*
1	0.004	0.004	0.0341	0.853
2	-0.000	-0.000	0.0343	0.983
3	0.001	0.001	0.0390	0.998
4	-0.014	-0.014	0.5605	0.967
5	-0.017	-0.016	1.2567	0.939
6	-0.015	-0.015	1.8410	0.934
7	-0.017	-0.017	2.5893	0.920

Fuente: Elaboración propia a partir de Eviews.

Tabla 5.24 Correlograma de los residuos al cuadrado. CQR.

	AC	PAC	Q-Stat	Prob*
1	0.015	0.015	0.6038	0.437
2	-0.036	-0.036	3.8818	0.144
3	-0.016	-0.015	4.5214	0.210
4	0.005	0.004	4.5783	0.333
5	-0.013	-0.014	5.0117	0.414
6	0.006	0.006	5.1016	0.531
7	0.045	0.044	10.214	0.177

Fuente: Elaboración propia a partir de Eviews.

Tabla 5.25 Correlograma de los residuos al cuadrado. Oriente.

	AC	PAC	Q-Stat	Prob*
1	0.009	0.009	0.2302	0.631
2	-0.025	-0.025	1.8025	0.406
3	-0.015	-0.015	2.3979	0.494
4	-0.004	-0.004	2.4386	0.656
5	0.017	0.016	3.1407	0.678
6	-0.008	-0.009	3.2994	0.770
7	0.042	0.042	7.7093	0.359

Fuente: Elaboración propia a partir de Eviews.

Tabla 5.26 Correlograma de los residuos al cuadrado. Sur.

	AC	PAC	Q-Stat	Prob*
1	-0.000	-0.000	0.0003	0.986
2	-0.024	-0.024	1.5063	0.471
3	-0.001	-0.001	1.5071	0.681
4	0.013	0.012	1.9269	0.749
5	-0.015	-0.015	2.4975	0.777
6	0.000	0.001	2.4977	0.869
7	0.032	0.032	5.1868	0.637

Fuente: Elaboración propia a partir de Eviews.

Tabla 5.27 Correlograma de los residuos al cuadrado. THC.

	AC	PAC	Q-Stat	Prob*
1	-0.012	-0.012	0.3766	0.539
2	-0.012	-0.012	0.7188	0.698
3	-0.035	-0.035	3.8674	0.276
4	-0.009	-0.010	4.0957	0.393
5	-0.006	-0.007	4.1900	0.522
6	-0.000	-0.002	4.1907	0.651
7	0.024	0.023	5.6829	0.577

Fuente: Elaboración propia a partir de Eviews.

Tabla 5.28 Correlograma de los residuos al cuadrado. Valle.

	AC	PAC	Q-Stat	Prob*
1	0.027	0.027	1.8542	0.173
2	-0.015	-0.016	2.4423	0.295
3	-0.018	-0.017	3.2886	0.349
4	0.009	0.010	3.4885	0.480
5	-0.011	-0.012	3.7782	0.582
6	-0.015	-0.015	4.3748	0.626
7	-0.006	-0.005	4.4757	0.724

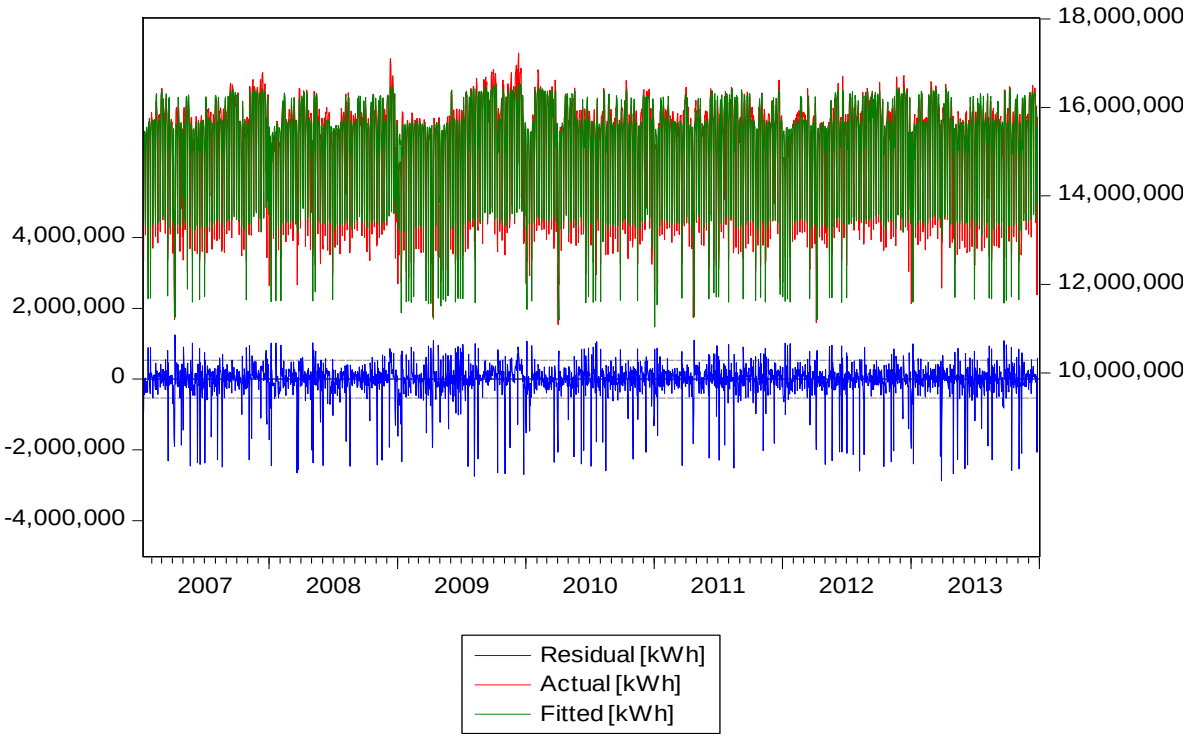
Fuente: Elaboración propia a partir de Eviews.

De acuerdo con los valores del estadístico Q de Ljung-Box y de sus correspondientes probabilidades en los correlogramas de residuos al cuadrado presentados de la **Tabla 5.21** a la **Tabla 5.28**, se puede concluir con un 95% de confianza que los residuos hasta con siete rezagos no presentan heterocedasticidad condicional autorregresiva, ya que en todos los modelos la probabilidad del estadístico Q resultar ser superior a 0,05. En este sentido, los correlogramas de los residuos al cuadrado indican que no existe una relación de dependencia cuadrática, exponencial o de cualquier otro tipo entre los errores de los ocho modelos.

d. Análisis gráfico de los modelos

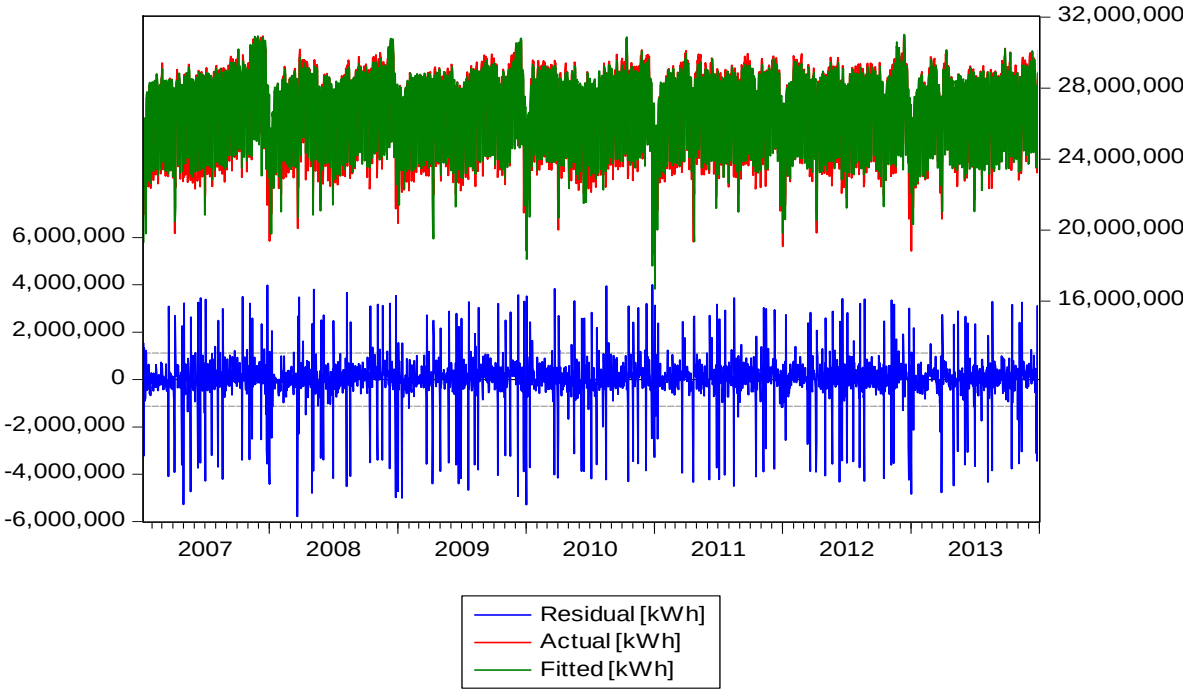
En la presente sección se muestran las gráficas de los distintos modelos (real, estimado y residual) a fin de mostrar el ajuste de los mismos.

Figura 5.7 Modelo de regresión y residuales de Antioquia.



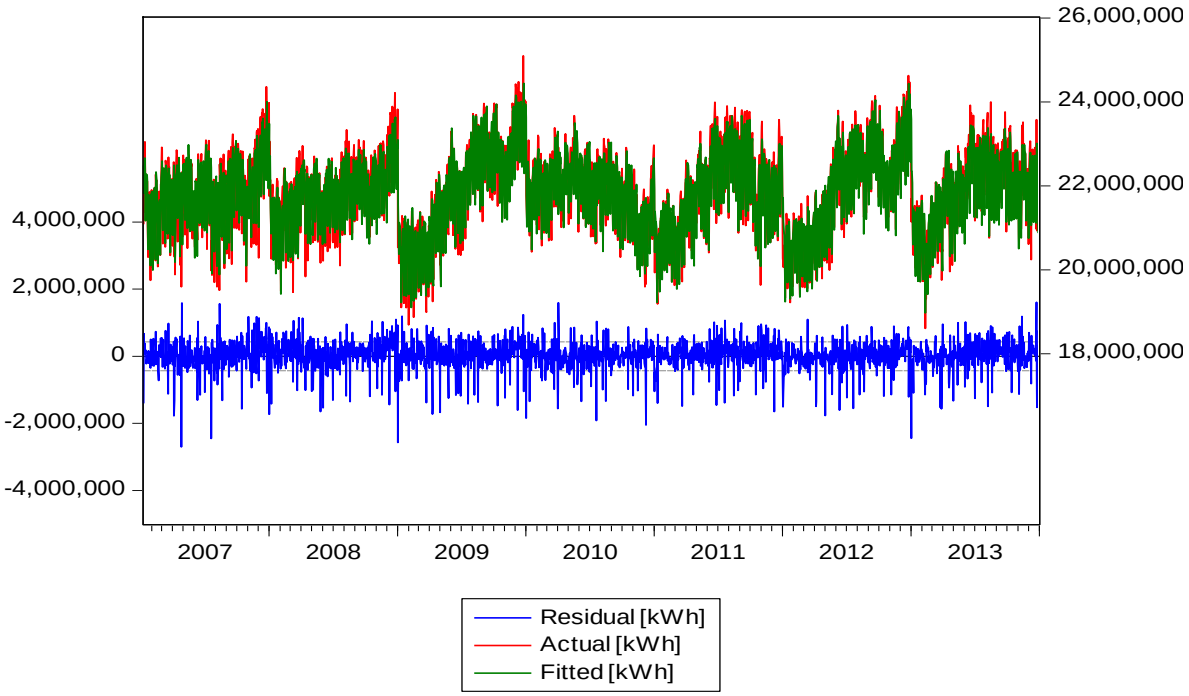
Fuente: Elaboración propia a partir de Eviews.

Figura 5.8 Modelo de regresión y residuales de Centro.



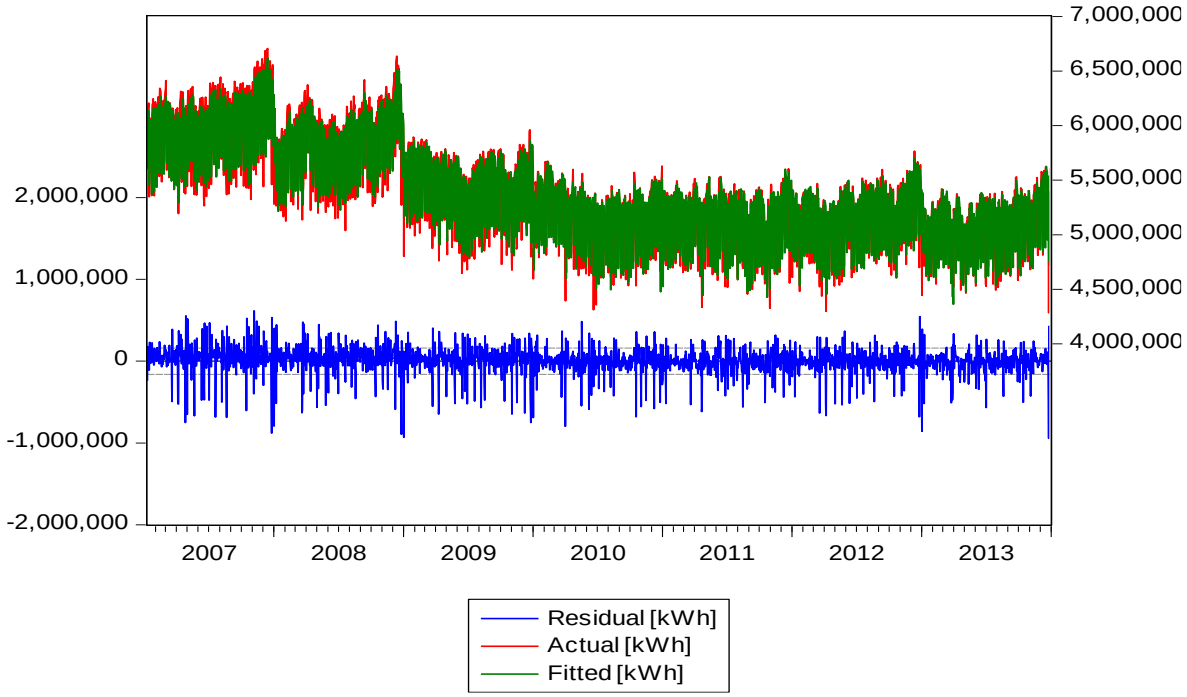
Fuente: Elaboración propia a partir de Eviews.

Figura 5.9 Modelo de regresión y residuales de Costa Atlántica.



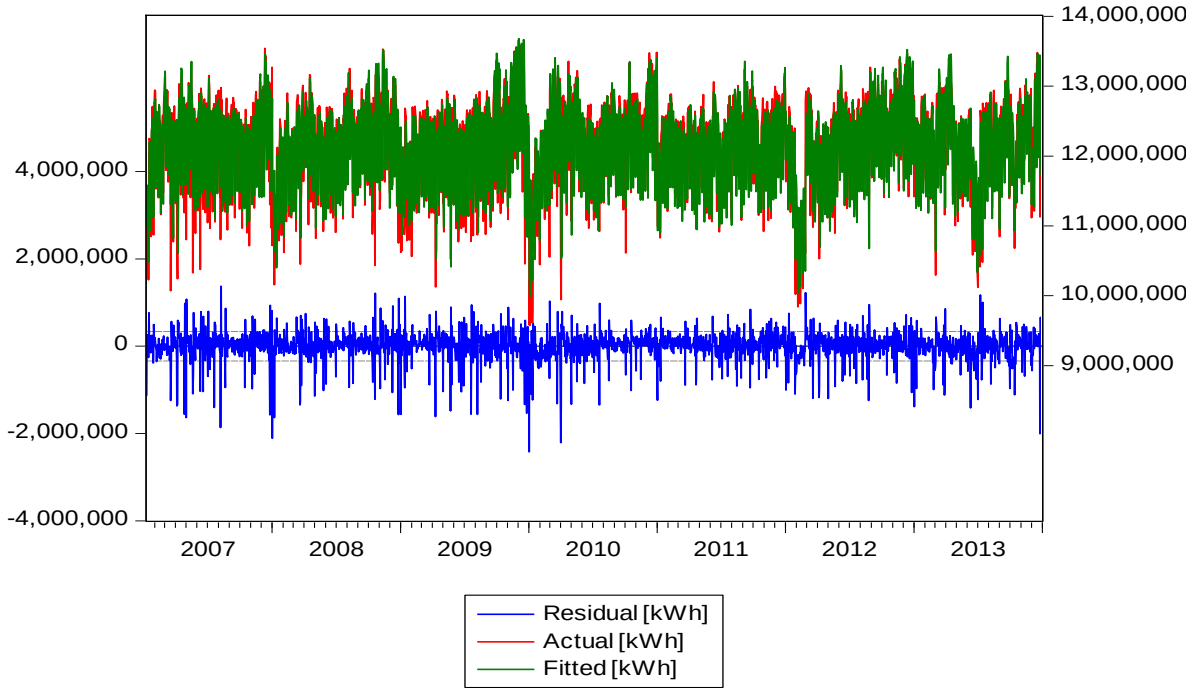
Fuente: Elaboración propia a partir de Eviews.

Figura 5.10 Modelo de regresión y residuales de CQR.



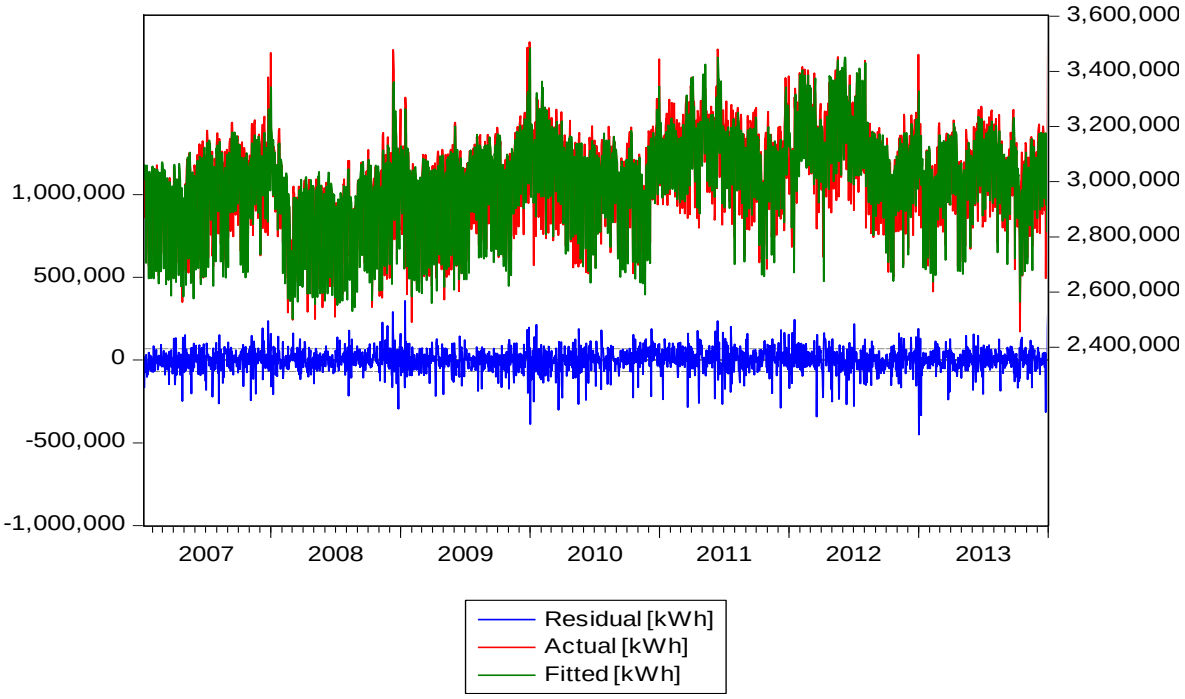
Fuente: Elaboración propia a partir de Eviews.

Figura 5.11 Modelo de regresión y residuales de Oriente.



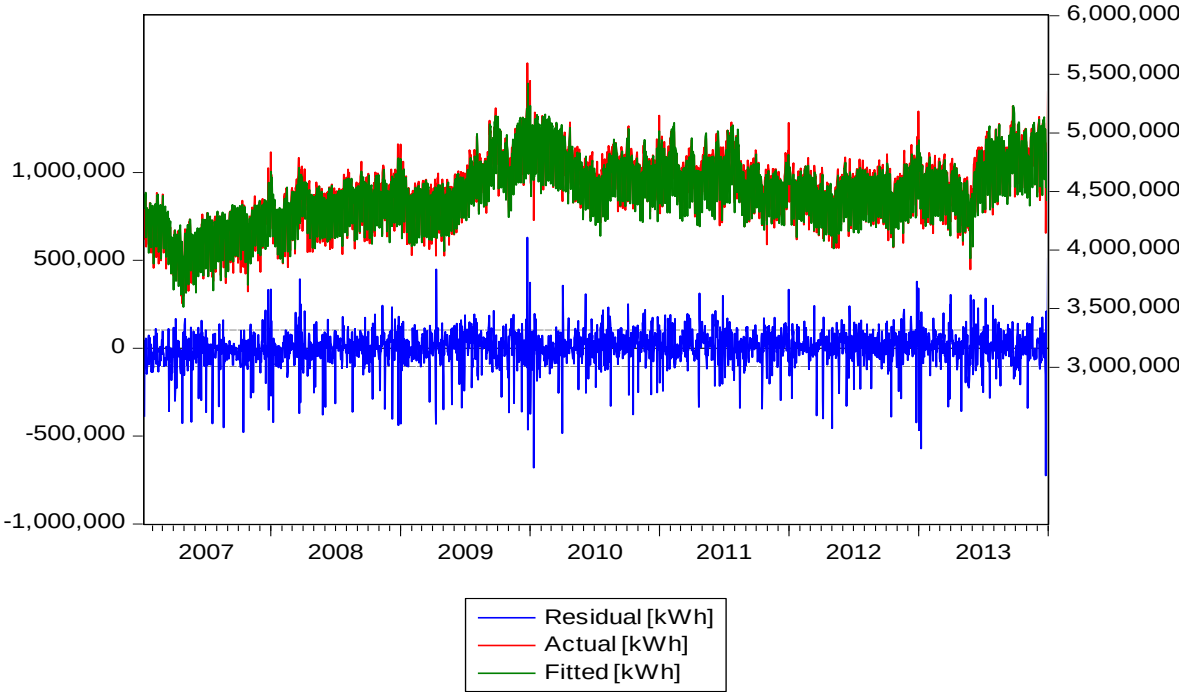
Fuente: Elaboración propia a partir de Eviews.

Figura 5.12 Modelo de regresión y residuales de Sur.



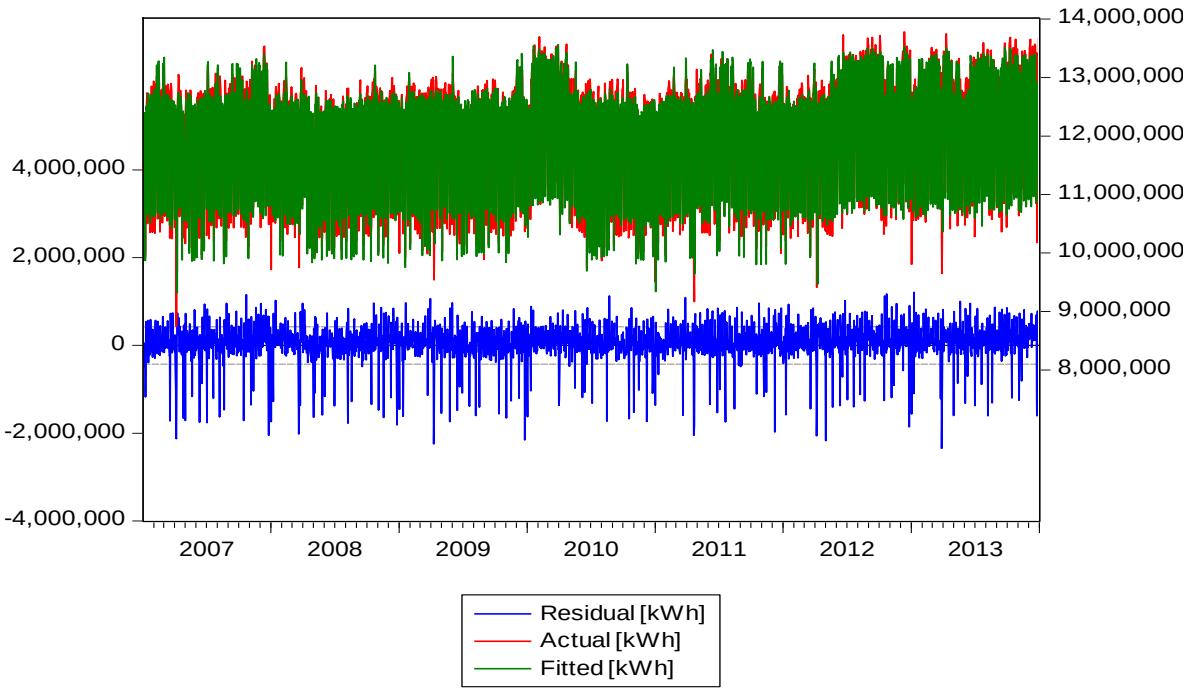
Fuente: Elaboración propia a partir de Eviews.

Figura 5.13 Modelo de regresión y residuales de THC.



Fuente: Elaboración propia a partir de Eviews.

Figura 5.14 Modelo de regresión y residuales de Valle.



Fuente: Elaboración propia a partir de Eviews.

En las series de demanda de energía eléctrica que se presentan de la **Figura 5.7** a la **Figura 5.14**, se observa que los modelos econométricos propuestos describen la demanda de electricidad en cada área operativa. Así mismo, de dichas figuras es posible concluir que los residuales, en todos los casos, oscilan de manera uniforme alrededor de cero, lo cual es, de manera gráfica, una muestra de que los errores no están autocorrelacionados y de que no existe presencia del efecto ARCH.

5.5 Conclusiones

A diferencia de los países con estaciones, donde la temperatura influye de forma clara en la demanda de electricidad (Climent Diranzo et al., 2003), en países ubicados en el trópico, como Colombia, ésta relación no es clara. En primer lugar, un porcentaje importante de la población tiene consumos de subsistencia por sus bajos ingresos y en segundo lugar, en sitios geográficamente cercanos la temperatura cambia con la altura. Por otra parte, la literatura no reporta estudios de la relación entre temperatura y demanda de electricidad en Colombia (véase sección 1.3).

Respecto al problema de investigación, con el desarrollo de este trabajo se pudo establecer que las variaciones en la temperatura media del aire generan cambios en la demanda de energía eléctrica en seis de las ocho áreas operativas que conforman el Sistema Interconectado Nacional. En el Centro y CQR la prueba de causalidad de Granger indica que no hay relación de precedencia. Es probable que la ausencia de dicha relación entre el indicador de temperatura, como variable explicativa, y la demanda de electricidad en estas áreas operativas, sea resultado de la tolerancia al cambio de las

condiciones climáticas (aumento de la temperatura) en sus principales centros poblados, Bogotá y Manizales respectivamente, ya que ambas ciudades se encuentran por encima de los 2.200 msnm. Esto último es coherente con los resultados obtenidos para la Costa Atlántica, ya que en ella, cinco de las siete capitales de los departamentos que la conforman se encuentran localizadas a menos de 20 msnm (Cartagena, Barranquilla, Montería, Santa Marta y Riohacha), y a que además la prueba de causalidad de Granger indica que ésta área operativa tiene la mayor relación de causalidad de la temperatura sobre la demanda de energía eléctrica del país.

A continuación se presentan las principales conclusiones encontradas con el desarrollo del presente trabajo final de maestría respecto a los objetivos específicos formulados en la sección 1.5.2.

5.5.1 Objetivo 1

Objetivo: Identificar las fuentes de información y construir, para cada área operativa, las series de demanda de energía eléctrica desagregada y el indicador diario de temperatura media del aire.

A partir de la información provista por el Instituto de Hidrología, Meteorología y Estudios Ambientales de Colombia (IDEAM), se obtuvieron registros diarios de temperatura media del aire de todas las estaciones climáticas localizadas en territorio colombiano en el periodo comprendido entre el 01:01:2007 y el 31:12:2013. Por otra parte, del portal institucional del Departamento Administrativo Nacional de Estadística (DANE), se descargaron las proyecciones de población de todos los municipios del país para el mismo periodo de tiempo. Finalmente, en el caso de la demanda de energía eléctrica, la información fue tomada del portal del operador del mercado eléctrico nacional XM Compañía de Expertos en Mercados S.A. E.S.P. En el proceso de tratamiento previo de datos se identificaron lagunas de información en las series de temperatura media del aire y limitaciones para el tratamiento de las series de demanda, para lo cual se implementaron las estrategias descritas en las secciones 2.2.1.1 y 182.2.1.2 respectivamente.

5.5.2 Objetivo 2

Objetivo: Determinar el grado de relación, en cada área operativa, entre las series históricas de demanda de energía eléctrica y el indicador diario de temperatura media del aire.

Una vez transformadas las series de demanda de energía eléctrica, se estableció que la demanda responde a cambios en la temperatura en seis de las ocho áreas operativas que conforman el Sistema Interconectado Nacional. Los resultados obtenidos indican que en la Costa Atlántica los usuarios residenciales son menos tolerantes al aumento de la temperatura media del aire en comparación con las demás áreas operativas, y por ello, la relación de causalidad y/o precedencia de la temperatura sobre la demanda de energía eléctrica es más importante en esta región del país. Note en la **Tabla 5.4** que el estadístico F disminuye en el siguiente orden: Costa Atlántica, Valle, Sur, Antioquia, THC, Oriente, Centro y CQR. En las dos últimas áreas operativas no se puede rechazar la hipótesis nula (Ho: Temperatura no causa demanda), ya que el estadístico F es significativo.

Respecto a los resultados obtenidos, es probable que la relación de causalidad se haga menos evidente a medida que la altura media de los principales centros urbanos localizados en cada área operativa es mayor respecto al nivel del mar. Sin embargo, esto no se cumple en las áreas operativas THC y Oriente. En estos casos, es posible que otras variables expliquen mejor el comportamiento de la demanda de electricidad que la temperatura media del aire, como por ejemplo la estacionalidad anual (periodos de vacaciones, semana santa, navidad, entre otros), el precio de la energía y de combustibles sustitutos y el crecimiento de la población, entre otros; ya que estas variables también inciden en la evolución de la demanda de energía eléctrica (Medina & García, 2005).

5.5.3 Objetivo 3

Objetivo: Evaluar el impacto de las variaciones de la temperatura en la demanda de energía eléctrica en cada una de las áreas operativas que integran el Sistema Interconectado Nacional – SIN.

Para la evaluación del impacto de las variaciones de la temperatura media del aire sobre la demanda de energía eléctrica, se propusieron de manera inicial ocho modelos de regresión, uno para cada área operativa. Dichos modelos debieron ser ajustados mediante la estimación de la volatilidad, ya que en todos los casos se encontró evidencia de autocorrelación en los errores al cuadrado, en otras palabras, de la existencia del efecto ARCH (heterocedasticidad condicional autorregresiva).

Respecto a la significancia conjunta de los modelos, se encontró que éstos describen entre el 73 y el 88% de la demanda de energía eléctrica en las diferentes áreas operativas, así como también que ninguno presenta problemas de multicolinealidad a pesar de tener el R cuadrado y el R cuadrado ajustado moderadamente elevados.

En relación con las variables explicativas consideradas, sólo la variable Dem(-1) es común en los ocho modelos planteados, lo cual es evidencia de la influencia del corto plazo en la previsión de la demanda de energía eléctrica en Colombia.

Finalmente, vale la pena resaltar que la prueba de causalidad de Granger indica que en la Costa Atlántica se presenta la relación más fuerte entre el cambio en las condiciones térmicas y la demanda de electricidad, probablemente porque en esta área operativa se utilizan de forma masiva aires acondicionados para satisfacer necesidades de enfriamiento de espacios. En las demás áreas operativas: Antioquia, Oriente, Sur, THC y Valle, la relación de causalidad debe ser generada por el uso (apertura y cierre) colectivo de electrodomésticos de enfriamiento como neveras, y en menor medida, de aires acondicionados para refrigeración. En este sentido, los resultados obtenidos indican que el cambio climático, medido a través de las variaciones en la temperatura media del aire, genera cambios importantes en los perfiles de demanda de energía eléctrica en el corto plazo como también lo concluyeron Oliveira et. al (2012) en su trabajo investigativo.

5.6 Trabajo futuro

De acuerdo con las conclusiones presentadas en la sección anterior, se plantea como posible trabajo futuro incluir registros de temperatura media del aire y de demanda de energía eléctrica posteriores

al 31:12:2013. Esto con el objetivo de identificar si eventos climáticos extremos como el fenómeno del niño 2015-2016 acentúan la respuesta de la demanda de energía eléctrica, en cada área operativa, a las variaciones en la temperatura media del aire. Así mismo, porque dichos eventos podrían dejar en evidencia además una relación de precedencia del cambio en las condiciones térmicas del aire sobre la demanda de electricidad en las áreas operativas Centro y CQR, por lo menos en el último año.

Por otra parte, podría ser útil analizar registros de temperatura y de demanda de energía eléctrica antes del 01:01:2007, incluyendo otras variables como el ENSO, para determinar si la temperatura media del aire aumentó o disminuyó cuando tuvieron lugar fenómenos climáticos extremos como el niño o la niña. A partir de esto último, debería ser posible establecer si es necesario construir un modelo de pronóstico que segmente la respuesta de la demanda a la temperatura dependiendo del evento climático más relevante del momento y que incluya además los periodos de tiempo en los cuales ninguno de los dos (niño o niña) estén sucediendo.

Así mismo, se propone incluir en el modelo econométrico estudiado otras variables que expliquen el comportamiento de la demanda de energía eléctrica como la laboralidad y la estacionalidad asociada con periodos vacacionales como la navidad y semana santa, los precios de la electricidad o de algunos de sus sustitutos. Esto último a fin de tratar de explicar aún más las variaciones de la demanda de energía eléctrica en las distintas áreas que conforman el Sistema Interconectado Nacional.

Finalmente, si bien para efectos del presente trabajo final de maestría se consideró que los modelos obtenidos son correctos con un 95% de confianza, esto último teniendo en cuenta que según Sosa et. al (2014) a pesar de que uno o más coeficientes de la ecuación de varianza sean inferiores a cero puede cumplirse que la varianza condicional no sea negativa; se recomienda explorar otros métodos de estimación de la volatilidad para garantizar que la varianza condicional sea positiva en todo momento, en particular si el objetivo de la construcción del modelo de regresión es realizar pronóstico, en este caso, de demanda de energía eléctrica.

6. REFERENCIAS

Ahmed, T., Muttaqi, K. M., & Agalgaonkar, a. P. (2012). Climate change impacts on electricity demand in the State of New South Wales, Australia. *Applied Energy*, *98*, 376–383. http://doi.org/10.1016/j.apenergy.2012.03.059

Arranz, J. M., & Zamora, M. M. (2010). Análisis de Autocorrelación. Universidad de Zaragoza.

Barrera. (2004). *Técnicas de completado de series mensuales y aplicación al estudio de la influencia de la NAO en la distribución de la precipitación en España*. Universidad de Barcelona.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327. http://doi.org/10.1016/0304-4076(86)90063-1

Carnero, M. A., Peña, D., & Ruiz, E. (2012). Estimating GARCH volatility in the presence of outliers. *Economics Letters*, *114*(1), 86–90. http://doi.org/10.1016/j.econlet.2011.09.023

Catarlis, C., Synodinou, A., Proedrou, M., & Tsanfrassoulis, A. (2001). Modifications in energy demand in urban areas as a result of climate changes: an assessment for the southeast Mediterranean region. *Energy Conversion and Management2*, *42*(14), 1647–1656. http://doi.org/10.1016/S0196-8904(00)00156-4

Chandramowli, S. N., & Felder, F. A. (2014). Impact of climate change on electricity systems and markets – A review of models and forecasts. *Sustainable Energy Technologies and Assessments*, *5*, 62–74. http://doi.org/10.1016/j.seta.2013.11.003

Chica, H., Peña, A., Giraldo, J., Obando, D., & Riaño, N. (2014). SueMulador: Herramienta para la Simulación de Datos Faltantes en Series Climáticas Diarias de Zonas Ecuatoriales. *Revista Facultad Nacional de Agronomia Medellin*, *67*(2), 7365–7373.

Climent Diranzo, F. J., Valor, E., Torró, H., & Caselles Miralles, V. (2003). Incidencia de la climatología en el consumo de gas y electricidad en España. *Infraestructura: Transporte E Industrias de Red*, (808), 55–70.

Compañía Energética del Tolima S.A. E.S.P. (2015). Mercado regulado y no regulado. Ibagué, Tolima. Retrieved from http://www.enertolima.com/portal/index.php/venta-de-energia/mercado-regulado-y-no-regulado.html

Considine, T. (2000). The Impacts of Weather Variations on Energy Demand and Carbon Emissions. *Resource and Energy Economics*, *22*(4), 295–314. http://doi.org/10.1016/S0928-7655(00)00027-0

Departamento Administrativo Nacional de Estadística. (2008). Estimación e interpretación del coeficiente de variación de la encuesta cocensal. Bogotá.

Dickey, D., & Fuller, W. (1979). Distribution of the Estimator for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, *74*(366a), 427–431.

Durbin, J., & Watson, G. S. (1950). Testing for serial correlation in least squares regression. I.

Biometrika, *37*, 409–428.

Durbin, J., & Watson, G. S. (1951). Testing for serial correlation in least squares regression. II. *Biometrika*, *38*(1/2), 159–177. http://doi.org/10.2307/2332325

Eischeid, J., Pasteris, P., Diaz, H., Plantico, M., & Lott, N. (2000). Creating a Serially Complete , National Daily Time Series of Temperature and Precipitation for the Western United States. *Journal of Applied Meteorology*, *39*, 1580–1591.

Engle, R., Granger, C. W. J., Rice, J., & Weiss, A. (1986). Semiparametric Estimates of the Relation Between Weather and Electricity Sales. *Journal of American Statistical Association*, *81*(394), 310–320.

Fernández, F. (1995). *Manual de Climatología Aplicada. Clima, Medio Ambiente y Planificación. Colección Espacios y Sociedades, Serie Mayor No 2*. (Síntesis, Ed.) (Ilustrada). Madrid.

Filippini, M. (1995). Swiss Residential Demand for Electricity by time-of-use. *Resource and Energy Economics*, *17*(3), 281–290. http://doi.org/10.1016/0928-7655(95)00005-I

Granados Castro, J. C. (2015). *Transmisión de volatilidades entre variables macroeconómicas en América Latina*. Universidad Nacional de Colombia.

Guégan, M., Uvo, C. B., & Madani, K. (2012). Developing a module for estimating climate warming effects on hydropower pricing in California. *Energy Policy*, *42*, 261–271. http://doi.org/10.1016/j.enpol.2011.11.083

Gutiérrez Moya, E., & Arévalo Quijada, M. T. (2003). *La demanda residencial de energía eléctrica en la Comunidad Autónoma de Andalucía: Un análisis cuantitativo*. Universidad de Sevila. Retrieved from http://dialnet.unirioja.es/servlet/tesis?codigo=23084&orden=0&info=link\nhttp://dialnet.unirioja.es/servlet/exttes?codigo=23084

Henley, A., & Peirson, J. (1998). Residential Energy Demand and the Interaction of Price and Temperature: British Experimental Evidence. *1Energy Economics*, *20*(2), 157–171. http://doi.org/10.1016/S0140-9883(97)00025-X

Hernández, R., Fernández, C., & Baptista, P. (2003). *Metodología dela investigación*. (McGraw-Hill, Ed.) (Tercera Ed).

IDEAM. (2014). Series de temperatura histórica en estaciones climáticas de Colombia.

Jaglom, W. S., McFarland, J. R., Colley, M. F., Mack, C. B., Venkatesh, B., Miller, R. L., … Kayin, S. (2014). Assessment of projected temperature impacts from climate change on the U.S. electric power sector using the Integrated Planning Model®. *Energy Policy*, *73*, 524–539. http://doi.org/10.1016/j.enpol.2014.04.032

Jarque, C., & Bera, A. (1987). A Test for Normality of Observations and Regression Residuals. *International Statistical Review*, *55*(2), 163–172. http://doi.org/10.2307/1403192

Juan, Á., Kizys, R., & Manzanedo, L. (2006). Autocorrelación. UNiversitat Oberta de Catalunya.

Justicia, A., & Dominguez, R. (1992). Notas metodológicas para la cumplimentación de series climáticas y extrapolación de datos. Su aplicación al mapa de temperaturas de Andalucía. *Baetica. Estudios de Arte, Geografía E Historia.*, *14*, 55–80.

Karl, T. R., & Williams, C. N. (1987). An Approach to Adjusting Climatological time Series for Discontinuous Inhomogeneities. *Journal of Climate and Applied Meteorology*, *26*, 1744–1763.

León Aristizábal, G. E. (2000). Tendencia de la Temperatura del Aire en Colombia. *Meteorología Colombiana*, (2), 57–65.

McCandless, G., Gabrielli, M. F., & Murphy, T. (2001). *Modelos econométricos de predicción*

macroeconómica en la Argentina.

Medina, S., & García, J. (2005). Predicción de demanda de energía en colombia mediante un sistema de inferencia difuso neuronal. *Energética 33*, 15–24.

Mirasgedis, S., Sarafidis, Y., Georgopoulou, E., Kotroni, V., Lagouvardos, K., & Lalas, D. P. (2007). Modeling framework for estimating impacts of climate change on electricity demand at regional level: Case of Greece. *Energy Conversion and Management*, *48*, 1737–1750. http://doi.org/10.1016/j.enconman.2006.10.022

Montero Granados, R. (2013). *Test de causalidad*. *Documentos de Trabajo en Economía Aplicada. Universidad de Granada. España*.

Moral, J., & Vicéns, J. (2003). *Un modelo de previsión de la demanda de energía eléctrica: Thor II*. *Instituto L.R. Klein - Centro Gauss*.

Moral, J., & Vicéns, J. (2005). Modelos STR para el análisis de la respuesta no lineal de la demanda de electricidad a la temperatura. *Instituto L. R. Klein - Centro Gauss*, (9), 1–24.

Murillos, J., Trejos, Á., & Carvajal, P. (2003). Estudio del pronóstico de la demanda de energia eléctrica, utilizando modelos de series de tiempo. *Scientia et Technica*, *23*(23), 37–40.

Nelson, D. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Source: Econometrica Econometrica*, *59*(2), 347–370. http://doi.org/10.2307/2938260

Nelson, D., & Cao, C. (1992). Inequality Constraints in the Univariate GARCH Model. *Journal of Business & Economic Statistics*, *10*(2), 229–235. http://doi.org/10.2307/1391681

Oliveira, P. M. O., Marzec, D. P., Bordin, G., Bretas, A. S., & Bernardon, D. P. (2012). Efecto del cambio climático sobre la previsión de demanda de energía eléctrica, 72–78.

Paulus, J., & Kohler, M. (1952). Interpolation of Missing Precipitation Records. *Monthly Weather Review*, *80*, 129–133.

Peterson, T., & Easterling, D. (1994). Creation of Homogeneous Composite Climatological Reference Series. *International Journal of Climatology*, *14*, 671–679.

Phillips, P., & Perron, P. (1988). Testing for a Unit Root in Time Series Regression. *Biometrika*, *75*(2), 335–346. http://doi.org/10.2307/2336182

Pilli-Sihvola, K., Aatola, P., Ollikainen, M., & Tuomenvirta, H. (2010). Climate change and electricity consumption—Witnessing increasing or decreasing use and costs? *Energy Policy*, *38*(5), 2409–2419. http://doi.org/10.1016/j.enpol.2009.12.033

Press, W., Flannery, B., Teukolsky, S., & Vetterling, W. (1992). *Numerical Recipes: The Art of Scientific Computing*. *Technometrics* (Second Edi, Vol. 29). University of Cambridge. http://doi.org/10.2307/1269484

Rodríguez, R., Llasat, M., & Martín Vide, J. (1999). Análisis de Series Temporales en Climatología: Modelización y Homogeneidad. In U. de Barcelona (Ed.), *Textos Docents 165* (p. 89). Barcelona.

Rueda, V., Velásquez, J., & Franco, C. (2011). Avances recientes en la predicción de la demanda de electricidad usando modelos no lineales. *DYNA*, *78*(167), 36–43.

Sailor, D., & Muñoz, J. (1997). Sensitivity of Electricity and Natural Gas Comsumption to Climate in USA-methodology and Results for Eight Sates. *Energy*, *22*(10), 987–998.

Sánchez González, C. (1999). *Métodos Econométricos*. (Ariel S.A., Ed.) (1st ed.). Barcelona, España.

Sathaye, J. a., Dale, L. L., Larsen, P. H., Fitts, G. a., Koy, K., Lewis, S. M., & de Lucena, A. F. P. (2013). Estimating impacts of warming temperatures on California's electricity system. *Global*

Environmental Change, *23*, 499–511. http://doi.org/10.1016/j.gloenvcha.2012.12.005

Schaeffer, R., Szklo, A. S., Pereira de Lucena, A. F., Moreira Cesar Borba, B. S., Pupo Nogueira, L. P., Fleming, F. P., … Boulahya, M. S. (2012). Energy sector vulnerability to climate change: A review. *Energy*, *38*, 1–12. http://doi.org/10.1016/j.energy.2011.11.056

Searcy, J., & Hardison, C. (1960). *Double-Mass Curves*. Washington.

Sosa, J. A., Batún Cutz, J. L., Guerrero Lara, E., Kantún Chim, D., Medina Peralta, S., & Pantı Trejo, H. (2014). Un paseo por el modelo GARCH y sus variantes. *Abstraction & Application*, *10*(2014), 35–50.

Startz, R. (2015). *EViews Illustrated for Version 9*. (Meredith Startz and IHS Global Inc., Ed.). Santa Barbara.

Tecano Molano, M. C. (2010). *Volatilidad de la tasa de cambio peso dólar: Estimación de un indicador para la gestión del riesgo y la planeación de estrategias de inversión*. Universidad Nacional de Colombia.

Valencia, C. (2012). AV_SIM_Charla_03. Ciencia, Tecnología y Técnica.

Velásquez, J. (2008). *Construcción de Escenarios de Pronóstico del Precio de Eléctricidad en Mercados de Corto Plazo*. *Cuadernos de Administración*. Universidad Nacional de Colombia.

XM Compañía de Expertos en Mercados S.A. E.S.P. (2014). Informe de Operación del SIN y Administración del Mercado. Retrieved May 14, 2016, from http://informesanuales.xm.com.co/2014/SitePages/operacion/3-1-Demanda-de-energia-nacional.aspx

XM Compañía de Expertos en Mercados S.A. E.S.P. (2015a). Glosario XM. Retrieved from http://www.xm.com.co/Pages/GlosarioXM.aspx

XM Compañía de Expertos en Mercados S.A. E.S.P. (2015b). Información Inteligente. Retrieved from http://informacioninteligente10.xm.com.co/demanda/Paginas/HistoricoDemanda.aspx?RootFolder=/demanda/Historico Demanda/Demanda Nacional&FolderCTID=0x012000DD2EF1C31BC7034A8E758B7C0E7FDADE&View={1C098CA5-E12F-43E3-8474-5787D51C0ED2}

XM Compañía de Expertos en Mercados S.A. E.S.P. (2015c). Información Inteligente. Retrieved from http://informacioninteligente10.xm.com.co/demanda/Paginas/HistoricoDemanda.aspx?RootFolder=/demanda/Historico Demanda/Demanda Comercial&FolderCTID=0x012000DD2EF1C31BC7034A8E758B7C0E7FDADE&View={1C098CA5-E12F-43E3-8474-5787D51C0ED2}

XM Compañía de Expertos en Mercados S.A. E.S.P. (2015d). Información Operativa y Comercial. Retrieved from http://www.xm.com.co/Pages/PronosticoOficialdeDemanda.aspx

XM Compañía de Expertos en Mercados S.A. E.S.P. (2015e). Informe de demanda de energía y fronteras - SIN. Retrieved from http://www.xm.com.co/Informes Mensuales de Anlisis del Mercado/01_Informe_Demanda_y_Fronteras_TXR_10_2015.pdf

XM Compañía de Expertos en Mercados S.A. E.S.P. (2015f). Sala de Prensa - Comunicados. Retrieved May 14, 2016, from http://www.xm.com.co/Pages/Comportamiento-demanda-energia-junio-2015.aspx

Zachariadis, T. (2010). Forecast of electricity consumption in Cyprus up to the year 2030: The potential impact of climate change. *Energy Policy*, *38*, 744–750. http://doi.org/10.1016/j.enpol.2009.10.019