

**DEMANDA DIARIA DE ENERGÍA EN COLOMBIA 2000-2013:
UNA APLICACIÓN DE P-SPLINES**

Johana Andrea Sanabria Domínguez

Trabajo presentado como requisito para optar el título de Economista

**Universidad del Valle
Facultad de Ciencias Sociales y Económicas
Santiago de Cali
2013**

**DEMANDA DIARIA DE ENERGÍA EN COLOMBIA 2000-2013:
UNA APLICACIÓN DE P-SPLINES**

Johana Andrea Sanabria Domínguez

Trabajo presentado como requisito para optar el título de Economista

Director

Julio Escobar Potes

Jefe de Estudio Económicos del Banco de la República Cali

Universidad del Valle

Facultad de Ciencias Sociales y Económicas

Santiago de Cali

2013

CONTENIDO

0. INTRODUCCIÓN.....	5
1. ANTECEDENTES.....	8
2. REVISIÓN DE LITERATURA SOBRE SPLINES.....	13
3. METODOLOGÍA.....	16
3.1 Datos.....	16
3.2 Metodología.....	17
3.2.1 Extracción de la fluctuación estacional anual: Aplicación de las funciones splines.....	19
3.2.2 Extracción de la fluctuación estacional semanal.....	21
4. RESULTADOS.....	21
COMENTARIOS FINALES.....	31
ANEXOS.....	33
REFERENCIAS.....	34

LISTA DE CUADROS Y GRÁFICOS

Cuadro 1. Revisión de literatura nacional e internacional.....	9
Gráfico 1. Demanda diaria de energía eléctrica no regulada 2000-2013, $\{y_{NR,t}\}$	22
Gráfico 2. Demanda diaria de energía eléctrica regulada 2000-2013, $\{y_{R,t}\}$	22
Gráfico 3. Demanda diaria de energía eléctrica no regulada entre enero 1 y diciembre de 2001.....	23
Gráfico 4. Demanda diaria de energía eléctrica regulada entre enero 1 y diciembre de 2001.....	24
Gráfico 5. Demanda diaria de energía eléctrica no regulada en enero de 2000.....	25
Gráfico 6. Demanda diaria de energía eléctrica regulada en enero de 2000.....	25
Gráfico 7. Componente de Tendencia de $\{y_{NR,t}\}$	26
Gráfico 8. Componente de Tendencia de $\{y_{R,t}\}$	26
Gráfico 9. Componente de estacionalidad semanal de $\{y_{NR,t}\}$	27
Gráfico 10. Componente de estacionalidad semanal de $\{y_{R,t}\}$	27
Gráfico 11. Componente de estacionalidad anual de $\{y_{NR,t}\}$	28
Gráfico 12. Componente de estacionalidad anual de $\{y_{R,t}\}$	29
Gráfico 13. Componente Irregular de $\{y_{NR,t}\}$	30
Gráfico 14. Componente Irregular de $\{y_{R,t}\}$	30

INTRODUCCIÓN

El Mercado de Energía Mayorista (MEM) en Colombia es el conjunto de sistemas de intercambio de información a través del cual se transa la energía producida por medio de dos mecanismos: mediante los contratos en Bolsa y a través de contratos bilaterales de corto y largo plazo. Los contratos bilaterales de compra son contratos financieros¹ pactados libremente entre comercializadores y generadores, con el fin de reducir su exposición a la volatilidad del mercado. Los contratos en bolsa son contratos de corto plazo pactados a través del Administrador del Sistema de Intercambios Comerciales (ASIC)².

El mercado de energía es regulado por el Estado pero su infraestructura y operación está a cargo de agentes privados. La regulación del mercado es llevada a cabo por la Comisión de Regulación de Energía y Gas (CREG), mientras que la provisión de la energía es tarea de cuatro agentes privados: generadores, transmisores, distribuidores y comercializadores. Los generadores producen la energía eléctrica para que se transe en el MEM³. Los transmisores transportan la energía desde los puntos de generación hasta Centros Regionales de Despacho (CRD). Los distribuidores transportan la energía desde los CRD hasta el consumidor final. Por último, los comercializadores⁴ compran la energía en el

¹ Se entiende por contratos financieros aquellos en los cuales un generador, o un comercializador, se obliga a entregar electricidad a un comprador, a un precio acordado previamente. En contraste, en los contratos físicos un generador se obliga a entregar físicamente su producción de electricidad a un comprador, a un precio acordado previamente (Infomemo-XM, 2007).

² El ASIC es el administrador de las transacciones realizadas en el MEM.

³ El total de energía eléctrica generada es la suma de las producciones de plantas hidroeléctricas, plantas térmicas, plantas menores y las cogeneradoras.

⁴ Los comercializadores son los intermediarios entre los generadores y el consumidor; se encargan de comprar electricidad en el MEM, su facturación y cobro a los usuarios finales.

MEM para su venta a los usuarios finales, regulados y no regulados⁵ (Lozano y Rincón, 2010).

El equilibrio permanente del mercado de energía requiere que la demanda de electricidad -de comportamiento inelástico- sea igual al total de la energía ofertada neta de las pérdidas ocasionadas en su transmisión. En este sentido, la capacidad de las plantas generadoras debe superar en cada momento la cantidad estrictamente demandada, por cuanto la electricidad no es un bien acumulable y además, la oferta es fija en el corto plazo, no es posible ajustar los niveles de producción con la suficiente velocidad para mantener el equilibrio del mercado. (Martín, 2002: 169).

Estas características especiales del mercado de la electricidad, su importancia por ser un insumo en la producción de bienes y servicios de una economía, y además el requerimiento de abastecer al mercado con la mayor continuidad, sin interrupciones, y manteniendo la calidad del bien, obligan a realizar estudios de la dinámica del consumo de energía y plantear modelos para efectuar pronósticos en horizontes de corto plazo.

Por lo arriba expuesto, se motiva encontrar en esta investigación, una metodología que facilite identificar las características principales del mercado, específicamente de la demanda diaria de energía eléctrica en Colombia, mediante la descomposición de la serie en sus componentes básicos de tendencia, ciclo, estacionalidad y ruido, para la demanda regulada (principalmente hogares) y no regulada (mayoritariamente industria, comercio y servicios). Adicionalmente, se espera que el modelo planteado facilite relacionar la actividad económica con la demanda de energía y facilite la estimación del PIB.

⁵ Los usuarios regulados están constituidos principalmente por los hogares y los usuarios no regulados están constituidos por la industria, el alumbrado público y los demandantes extranjeros.

De otro lado, la demanda diaria de energía eléctrica de tipo residencial y otros usos (industrial, comercial, etc.) en Colombia, genera extensas series de tiempo de alta frecuencia caracterizadas por un patrón estacional complejo en el que se superponen fluctuaciones semanales, mensuales y anuales (festivos, fines de semana, bisiestos, Semana Santa, temporadas especiales, entre otras). Las técnicas tradicionales alejan el uso de modelos parsimoniosos, por lo que se plantea la técnica de splines⁶ cúbicos para remover el componente estacional de las series, de manera que la contribución de esta investigación es en términos metodológicos, en concreto, el uso de los splines como técnica para suavizar series de tiempo económicas de alta frecuencia (diaria).

El documento contiene cuatro secciones adicionales a esta introducción. En la primera sección se presenta la revisión de literatura sobre la demanda de energía en Colombia y en el exterior. En la segunda sección se revisan investigaciones que abordan la técnica Spline. En la sección tres se describe la metodología empleada. En la sección cuatro se presentan los resultados. Por último, se realizan los comentarios finales.

⁶ Un spline es una curva construida a partir de funciones polinomiales de grado m que se unen en puntos llamados nodos. Son útiles porque facilitan el ajuste de muchos datos con periodicidad sin recurrir al uso elevado de parámetros o variables ficticias para su modelación.

1. ANTECEDENTES

La electricidad es un bien no almacenable, con oferta fija en el corto plazo, y un insumo importante en la producción de bienes y servicios de una economía. Estas características especiales de la electricidad han hecho que el estudio del comportamiento de la demanda de energía eléctrica adquiriera recientemente mayor relevancia en la investigación económica⁷. En ese sentido, ya sea con fines de modelación o pronóstico y, dependiendo de la frecuencia de los datos, se han realizado investigaciones que recurren a distintas metodologías. El cuadro 1 resume los estudios revisados durante esta investigación.

A nivel internacional, es más frecuente encontrar en la literatura reciente el uso de series temporales económicas de alta frecuencia (semanal, diaria, horaria) para la modelación de la demanda de energía eléctrica.

Un primer estudio revisado es el de Poirier, Hendricks y Koenker (1979) en el que estiman la demanda residencial de energía en Estados Unidos. Con información registrada en intervalos de 15 minutos, realizan su estudio para el Electric Power Research Institute mediante la aplicación de modelos con splines.

Por su parte, Engle, Granger, Rice y Weiss (1986) estiman un modelo semiparamétrico para la relación no lineal entre las ventas diarias de electricidad y la temperatura en Georgia. Usan la técnica de suavización Spline y el parámetro de suavización (λ) es estimado minimizando el criterio de validación cruzada generalizado (GCV, por sus siglas en inglés).

⁷ En algunos países menos desarrollados, especialmente de África, instituciones internacionales han utilizado el consumo como una aproximación al PIB.

Cuadro 1
Revisión de literatura nacional e internacional

Autor	País/Región	Técnica/Modelo	Variables	Información
Zapata (2011)	Colombia	ARIMA	PIB, consumo de gas, precios del gas y de la energía.	Mensual, 2003:1- 2010:8
Rueda (2011)	Colombia	Híbrido entre SARIMA y GSMN		Mensual, 1995:8-2010:4
Franco, Velásquez y Olaya (2008)	Colombia	Componentes no observables		Mensual, 1995:8 -2006:1
Barrientos, Olaya y González (2007)	Región Sur-Occidente colombiano: Valle del Cauca	Funciones Spline ARIMA Un modelo para cada hora del día	Hora, día, mes	Horario, 2001:1- 2004:11
Murillo, Trejos y Carvajal (2003)	Demanda de una empresa de energía de Pereira	ARIMA Un modelo para cada hora del día		Horario, 2001:1 - 2001:12
Ramírez (1991)	Colombia	Mínimos cuadrados ordinarios, Mínimos cuadrados ponderados.		Anual
Pérez y Acosta (1987)	Colombia	Revisión de literatura		
Taylor (2003)	Inglaterra y Gales	Doble suavización de Holt-Winters		Intervalos de 30 minutos
Blaconá y Abril (2000)	Argentina	Modelo de Espacio de Estado - Funciones Spline	Temperatura	Diaria, 1-8-1995 a 12-09-1999
Engle, Granger, Rice y Weiss (1986)	Georgia	Modelo semiparamétrico Suavización Spline	Temperatura	Diaria, 1-8-1995 a 12-09-1999
Poirier, Hendricks y Koenker (1979)	Estados Unidos	Funciones splines		Intervalos de 15 minutos
Harvey y Koopman (1993)	Estados Unidos	Modelo Estructural Spline cúbico	Temperatura	Horaria, 1-1-1985-9-9-1991

Hendricks y Koenker (1992) emplean modelos splines jerárquicos para modelar la demanda de electricidad horaria residencial en el área metropolitana de Chicago para 168 observaciones correspondientes a cuatro meses del año 1985.

De otro lado, Harvey y Koopman (1993) utilizan un modelo estructural de series de tiempo para modelar y pronosticar la demanda horaria de electricidad de la compañía Puget Sound Power. Los autores incluyen la temperatura como variable explicativa y emplean un spline cúbico, una técnica parsimoniosa, para recoger los movimientos periódicos intradiarios e intramensuales.

También Blaconá y Abril (2000) emplean un modelo estructural de series de tiempo para estudiar la demanda diaria de energía eléctrica en Argentina. Según los autores, la serie de tiempo de la energía tiene dos efectos estacionales superpuestos, uno semanal y otro anual, que son difíciles de modelar con la teoría convencional de series de tiempo. La estacionalidad semanal es fácilmente modelada con variables ficticias mientras que la estacionalidad anual es modelada aplicando la técnica spline.

Taylor (2003) estudia la demanda de energía eléctrica en Inglaterra y Gales. La información del lunes 5 de junio de 2000 al domingo 18 de junio de 2000 es registrada en intervalos de 30 minutos. Debido a la alta frecuencia de los datos, la serie de tiempo contiene más de un patrón estacional que es modelado mediante la formulación del método de doble suavización exponencial de Holt-Winters, el cual supera al tradicional Holt-Winters y al modelo ARIMA estacional multiplicativo.

A nivel nacional, son pocas las investigaciones publicadas que combinan el estudio de la demanda de energía eléctrica y el uso de series temporales económicas de alta frecuencia. La literatura que aborda ambos aspectos se ha desarrollado principalmente durante la última década.

Una de las primeras aproximaciones al tema de demanda de energía en Colombia es el trabajo de Pérez y Acosta (1987). Los autores presentan un repaso de las metodologías utilizadas para explicar la demanda residencial y comercial de energía eléctrica en Colombia y en otros países, y señalan que las tarifas crecientes por bloques constituyen un problema de la estimación de los modelos de demanda de electricidad pues causan simultaneidad y, en consecuencia, proponen incluir el precio marginal como variable explicativa o utilizar mínimos cuadrados en dos etapas.

Por otra parte, Ramírez (1991) estudia la relación entre la demanda de energía eléctrica en la industria colombiana y los insumos no energéticos, específicamente el capital y el trabajo. Para una muestra de 448 observaciones correspondiente a 28 sectores industriales y 16 años, encuentra que la energía eléctrica en la industria es sustituta del trabajo y complemento del capital.

Posteriormente, Murillo, Trejos y Carvajal (2003) realizan un pronóstico de la demanda de energía eléctrica en Pereira utilizando modelos ARIMA con información provista por la Empresa de Energía de Pereira entre el 1 de enero de 2001 y el 31 de diciembre.

Barrientos, Olaya y González (2007) son pioneros en el uso de modelos con spline cúbicos en Colombia. Los autores estudian la demanda horaria de energía eléctrica en el Suroccidente colombiano con fines de pronóstico, teniendo en cuenta que los datos presentan fluctuaciones periódicas horarias, diarias, semanales, mensuales y anuales, que son complejas y difíciles de modelar con la teoría básica de métodos paramétricos y por ello recurren a un modelo de regresión no paramétrico con splines o formas funcionales parsimoniosas.

Franco, Velásquez y Olaya (2008) utilizan un modelo de componentes no observables para caracterizar la demanda mensual de electricidad en Colombia

entre agosto de 1995 y enero de 2006. Mediante el filtro de Kalman separan la tendencia y el componente estacional de la serie. Los autores se enfrentan a una estacionalidad anual que modelan con variables dummy y logran aislar el efecto de la recesión económica de finales de la década de los noventa en la demanda de electricidad.

Rueda (2011) modela la demanda mensual de energía en Colombia entre agosto de 1995 y abril de 2010. Empleando modelos no lineales (SARIMA, MARS; DAN2, SRT y GSMN) y combinando un modelo SARIMA y un modelo GSMN, logra mostrar que el modelo híbrido presenta mejor desempeño pues es el que mejor captura la dinámica de la serie y presenta los mejores estadísticos de ajuste.

Zapata (2011) estudia el impacto del PIB, del consumo de gas natural y de los precios de la electricidad, en el consumo de energía eléctrica en Colombia entre enero de 2003 y agosto de 2010. Mediante un modelo ARIMA logra demostrar que el PIB es el mayor determinante del consumo de energía eléctrica en Colombia superando los precios de la energía eléctrica y los del gas natural, así como el consumo del gas natural; no obstante el consumo de gas y los precios del mismo en la explicación del consumo total de energía en Colombia no son significativos en el modelo.

En resumen, han sido varias las metodologías empleadas para modelar y pronosticar la demanda de energía eléctrica. En la revisión de literatura se evidencia que en los últimos años el sector eléctrico se ha enriquecido con investigaciones para series temporales de alta frecuencia, caracterizadas por la presencia de múltiples componentes estacionales superpuestos que son complejos y difíciles de modelar con la teoría tradicional de series de tiempo. Para el caso colombiano, no se ha encontrado una caracterización de la demanda de energía con información diaria que combine un modelo estructural y una función spline cúbica.

En concreto, el aporte de la presente investigación radica en los siguientes aspectos:

1. Para el caso colombiano no ha sido publicada alguna investigación que aborde el problema de la caracterización de la demanda nacional diaria de energía eléctrica, y además, que utilice las funciones splines o formas funcionales parsimoniosas para extraer el componente estacional de la serie de consumo.
2. Si bien la investigación de Franco, et al. (2008) plantea un modelo de componentes no observables, igual que en esta investigación, la estacionalidad anual es tratada con variables dummy.
3. En las investigaciones de Franco, et al. (2008) y Barrientos *et al.* (2007) se trabaja con la serie de demanda mensual y horaria, respectivamente, no diaria, tal como se realiza en la presente investigación, que facilitará la estimación diaria y semanal, según el mes del año en curso.
4. Se establece un precedente de la caracterización de la estructura de la demanda de energía para analizar en un trabajo posterior, su grado de asociación con el PIB nacional, lo cual facilitará la construcción de un indicador coincidente de la actividad económica vigente por el lado de la oferta, líder en su estimación, dada la gran oportunidad en tiempo real de las cifras de consumo de electricidad.

2. REVISIÓN DE LITERATURA SOBRE SPLINES

2.1 Estacionalidad en series de tiempo de alta frecuencia

En los últimos años, las características de las series de tiempo a las que pueden acceder los investigadores económicos se han modificado considerablemente como consecuencia del mayor desarrollo tecnológico y el interés en la dinámica de corto plazo, al punto que para cortos periodos se alcanza una gran cantidad de observaciones (Martín, 2008).

El análisis de series de tiempo de alta frecuencia tiene implícito un mayor grado de complejidad, cuando se presentan variaciones estacionales para diversos periodos: horaria, diaria, semanal, mensual y anual, dependiendo de los intervalos en que se registre la información. En el caso de la demanda diaria de energía en Colombia, se presenta un patrón estacional complejo en el que se superponen una fluctuación semanal y una fluctuación anual (días festivos, fines de semana, años bisiestos, Semana Santa, Navidad, Vacaciones, entre otras)

La extracción de los componentes estacionales de una serie es necesaria por la distorsión que generan sobre la dinámica de largo plazo. Sin embargo, con series de tiempo extensas cuya información es registrada en cortos intervalos de tiempo, los métodos existentes pierden la parsimonia al requerir en la estimación un parámetro para cada elemento o estación del componente estacional.

Este problema motiva la búsqueda de modelos más parsimoniosos, es decir, de menor número de parámetros, y es así como las funciones splines aparecen en la literatura especializada de series de tiempo como la metodología más conveniente para capturar gran cantidad de estaciones del componente estacional y la heterogeneidad de la serie como fue planteada inicialmente por Poirier (1973) y más reciente por Martín (2008).

En su estudio pionero, Poirier expone la teoría básica de los splines de regresión cúbicos y los plantea como una herramienta natural para analizar el cambio estructural, en donde los nodos pueden ser puntos en el tiempo cercanos a los puntos de cambio estructural, lo cual ilustra con un ejemplo basado en 500 datos de Indianapolis. Así mismo, plantea que un Spline cúbico es una herramienta útil para modelar las variaciones estacionales complejas y la heterogeneidad en la longitud de las observaciones (años bisiestos y no bisiestos).

Posteriormente, Wegman y Wright (1983) revisaron una gran variedad de trabajos sobre la aplicación de la técnica de Splines en Estadística y la resumen como una técnica de estimación no paramétrica con un rol fundamental en el análisis de series de tiempo con alta estacionalidad. Al comparar la metodología de regresión Spline con la metodología de regresión isotónica⁸ formulan un estimador híbrido que mejora las propiedades de continuidad de los estimadores isotónicos.

De otro lado, Silverman (1985) aporta los principales aspectos metodológicos de suavizado con spline, en especial, los principios, la elección del parámetro de suavizado y el análisis de los residuales, técnica que es ilustrada con dos ejemplos. A su vez, Eilers y Marx (1996) combinan los splines de regresión y los splines de suavizado para obtener un spline penalizado o P-spline. Entre sus ventajas señalan que con la penalización no se requiere la especificación del número de nodos y su localización; además, se penalizan los coeficientes del polinomio y no la curva.

Un modelo estructural de series de tiempo fue planteado por Harvey, Koopman y Riani (1997) para modelar observaciones semanales, con número de semanas variable entre distintos años, bisiestos y no bisiestos, y el comportamiento de la serie en los días festivos. La estacionalidad de la serie es capturada mediante splines cúbicos y su metodología se ilustra con información de la oferta de dinero del Banco Inglaterra.

Finalmente Durbán (2009), realiza una presentación de los Splines Penalizados o P-Splines como técnica de suavizado de mayor flexibilidad, que permite el tratamiento de datos de alta frecuencia temporal caracterizados por una modelación de mayor complejidad.

En resumen, la teoría y aplicación de los splines en estadística ha evolucionado desde un método de regresión hasta el suavizado de series temporales para

⁸ La teoría de la regresión isotónica también es conocida como inferencia estadística bajo restricciones de orden

remover estacionalidad; es decir que, en años recientes, la flexibilidad de la técnica de splines, específicamente de los P-splines, para ajustar los datos ha facilitado ampliamente el tratamiento de variaciones estacionales complejas en series de tiempo económicas de la alta frecuencia que, aún en periodos cortos de estudio, llegan a ser muy extensas.

3 METODOLOGÍA

3. 1 Datos

Se dispone de información diaria sobre la demanda de energía eléctrica regulada y no regulada medida en Kilowatt hora (Kwh) entre enero 1 de 2000 y julio 19 de 2013, disponible en el sistema Neón de X.M. Expertos en Mercados. Filial de ISA S.A. E.S.P. La información es procesada en el software R.

La demanda comercial regulada se define como la demanda de los comercializadores para atender el consumo de los hogares, más la participación en las pérdidas⁹ del Sistema de Trasmisión Nacional (STN). La demanda comercial no regulada es la demanda de los comercializadores para atender el consumo de las industrias, el alumbrado público y las exportaciones, más la participación en las pérdidas del STN (Neón, XM).

⁹ Las pérdidas del Sistema de Transmisión Nacional (STN) están constituidas por la cantidad de energía que se pierde en la transmisión y la distribución y, que por tanto, no llega al consumidor final.

3.2 El Método

El objetivo de esta investigación es plantear un método que permita encontrar las características principales de la demanda diaria de energía eléctrica de usuarios regulados y no regulados en Colombia entre enero 1 de 2000 y julio 19 de 2013, basado en la descomposición de las series en sus componentes básicos de tendencia, ciclo, estacionalidad y ruido.

Con el propósito de obtener el consumo neto de energía de los agentes del sistema se modifican las dos variables anteriores. Por un lado, la demanda comercial regulada menos su participación en las pérdidas del STN constituye la *Demanda Comercial Regulada Neta*, $\{y_{R,t}\}$. Por otro lado, al descontar de la Demanda Comercial No Regulada su participación en las pérdidas del STN y las exportaciones internacionales se obtiene la *Demanda Comercial No Regulada Neta*, $\{y_{NR,t}\}$.

Se utiliza el enfoque de los modelos estructurales como herramienta estadística para recoger la variabilidad en los componentes de la serie. Se especifica entonces el siguiente modelo¹⁰:

$$y_t = T_T + S_t^{semanal} + S_t^{anual} + C_T + \varepsilon_t, \quad \varepsilon_t \sim NID(0, \sigma_\varepsilon^2) \quad (1)$$

donde T_T representa el componente de tendencia, $S_t^{semanal}$ es la variación estacional semanal modelada mediante variables dummy, S_t^{anual} representa la estacionalidad anual que se modelará mediante una función spline. En forma respectiva, C_T y ε_t representan el componente cíclico y el componente irregular.

¹⁰ Una ecuación similar se presenta en Martín (2008).

En primer lugar, se descomponen las series $\{y_{NR,t}\}$ y $\{y_{R,t}\}$ mediante el filtro de Hodrick-Prescott para tener una noción más clara sobre el comportamiento de la tendencia y el ciclo.¹¹ Luego, se remueve el componente estacional semanal a través de siete variables dummy, una para cada día de la semana. Posteriormente, se remueve la fluctuación estacional anual, la que contiene efectos de festivos, Semana Santa y otros por detectar, mediante un *spline cúbico*, concretamente, se usa la metodología P-spline (se explica más abajo con mayor detalle). Al final del proceso, se llega al componente irregular o ruido.

Es importante anotar que durante el periodo de estudio (01-2000-07-2013) existen cuatro años bisiestos, una de las fuentes de heterogeneidad en las series. Para el tratamiento de los años bisiestos se distribuyó el consumo de los días 29 de febrero de 2000, 2004, 2008 y 2012 entre el consumo de los 1460 días correspondientes a los cuatrienios o lapsos respectivos (2000-2003, 2004-2007, 2008-2011, 2012-2013), consiguiendo así homogeneizar la longitud de observaciones anuales¹², como se ilustra en (2)

$$Consumo_{29} * w_i + Consumo_i = Consumo_i^*, \quad i \neq 29 \quad (2)$$

donde $w_i = Consumo_i / Consumo$ y $Consumo = \sum_{i=1}^{1460} Consumo_i$, $i \neq 29$

¹¹ El filtro de Hodrick-Prescott (1980) es una técnica estadística empleada para extraer el componente cíclico y el componente de tendencia de una serie de tiempo.

¹² A pesar de que la longitud de las observaciones anuales es ahora de igual magnitud, no es conveniente la aplicación del operador diferencia para la transformación de las serie en estacionarias porque los días de la semana son diferentes entre sí. Por ejemplo, en el caso de Semana Santa, no es lo mismo comparar un viernes festivo con un viernes hábil.

3.2.1 Extracción de la fluctuación estacional anual: Aplicación de las funciones Splines

Como se resaltó en la revisión de literatura sobre Splines, éstos son útiles para modelar la estacionalidad en series de tiempo de alta frecuencia. Las funciones splines son usadas para ajustar los valores de una serie y en el tiempo. Así, en una nube de puntos secuencial (datos en el tiempo) la técnica de splines cúbicos identifica cambios de pendiente en la serie, crea un tramo polinomial entre uno y otro cambio de pendiente -cada tramo corresponde a una pieza de la curva completa del modelo- y dichos tramos se unen unos con otros en sus extremos en un único punto de tiempo elegido mediante la existencia e igualdad de la segunda derivada de cada tramo en ese punto, con lo cual garantiza la continuidad de la curva completa definida como la suma o unión de todos los tramos.

Formalmente, un spline es una curva construida a partir de un conjunto de piezas polinomiales cada una de grado m que se unen en k puntos llamados nodos, asociados a ciertas condiciones de continuidad. La cantidad de nodos necesarios para construir la curva que ajusta la serie de tiempo y_t constituye la base de la spline, que se denota con la letra B . Entonces, un B-spline es un spline de base B .

En general, una función spline es de grado m cuando existen las m primeras derivadas y las $(m-1)$ primeras son continuas. Así, una función spline de grado m presenta las siguientes propiedades:

- Se compone de $m+1$ trozos de polinomio, cada uno de grado m .
- Los trozos de polinomio se conectan en m nodos internos.
- En los puntos de unión, las derivadas hasta el orden $m-1$ son continuas.
- El B-spline es positivo sobre el dominio expandido por $m+2$ nodos y 0 en el resto.

- Excepto en los extremos, un B-spline se intercepta con $2m$ trozos de polinomios de sus vecinos.
- Para cada x , $m+1$ B-splines son no nulos.

Para pares de datos (x_i, y_i) , un modelo suavizado viene dado por

$$y_i = f(x_i) + \varepsilon_i \quad (3)$$

Donde $f(x_i)$ es una función suave de los datos, a la que no se le impone ninguna forma.

El P-spline es un método de suavizado que encuentra una función con dos derivadas continuas que minimiza la suma de cuadrados penalizada:

$$SCP = \sum_{i=1}^n (y_i - f(x_i))^2 + \lambda \int_{x_{min}}^{x_{max}} (f''(x))^2 dx \quad (4)$$

Donde el ultimo termino es una penalización en la segunda derivada de la curva y λ es el parámetro de suavizado.

En los P-splines la elección del parámetro de suavizado, λ , es de vital importancia. Aquí, el parámetro λ elegido corresponde al valor de λ que minimiza el criterio de validación generalizado (CVG). Dicho criterio ha sido utilizado por Engle et al. (1986), Blaconá y Abril (2000), Barrientos et al. (2007), entre otros, y se define como el promedio ponderado del error al cuadrado:

$$CVG(\lambda) = n^{-1} \sum_{i=1}^n (y_i - f(x_i) / 1 - s_i)^2 \left\{ 1 - s_i / n^{-1} tr[I - S_\lambda] \right\}^2 \quad (5)$$

donde S_λ es una matriz equivalente a la matriz HAT¹³ en regresión paramétrica y s_i es el i -ésimo elemento de la diagonal de S_λ .

¹³ La matriz HAT (sombrero), también llamada matriz de proyección, describe la influencia de cada valor observado sobre los valores ajustados.

3.2.2 Extracción de la fluctuación estacional semanal

La estacionalidad semanal se refiere al comportamiento recurrente y diferenciado de cada uno de los días de la semana. El componente estacional semanal ($s=7$) puede ser removido bien sea a partir del uso de funciones trigonométricas en las frecuencias estacionales $s/2$ (s periodo estacional), o empleando variables dummy. En este trabajo dicho comportamiento periódico semanal es tratado mediante siete variables dummy cuya suma en un periodo de siete días es 0, si se considera una estacionalidad determinística, es decir

$$\sum_{j=1}^7 S_j^{semanal} = 0 \quad (6)$$

Pero si la estacionalidad semanal varía en el tiempo, es decir que se considera un estacionalidad estocástica, debe especificarse

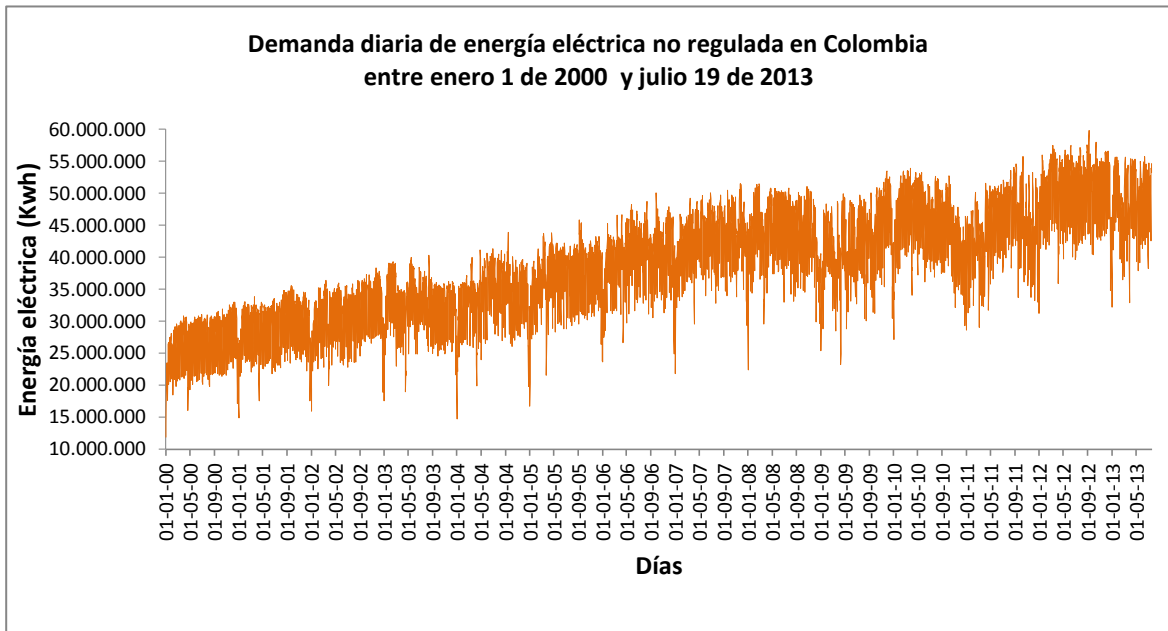
$$S_t^{semanal} = \sum_{j=1}^6 S_{t-j}^{semanal} + \omega_t, \quad t = 1, \dots, T, \quad \omega_t \sim NID(0, \sigma_\omega^2) \quad (7)$$

4. RESULTADOS

El análisis exploratorio de los datos mostró la gran inestabilidad del consumo diario de energía regulada y no regulada y la presencia de una tendencia de largo plazo (gráficos 1 y 2). Además, se detecta un comportamiento regular a nivel de días de la semana (gráficos 5 y 6) y de meses del año (gráficos 3 y 4).

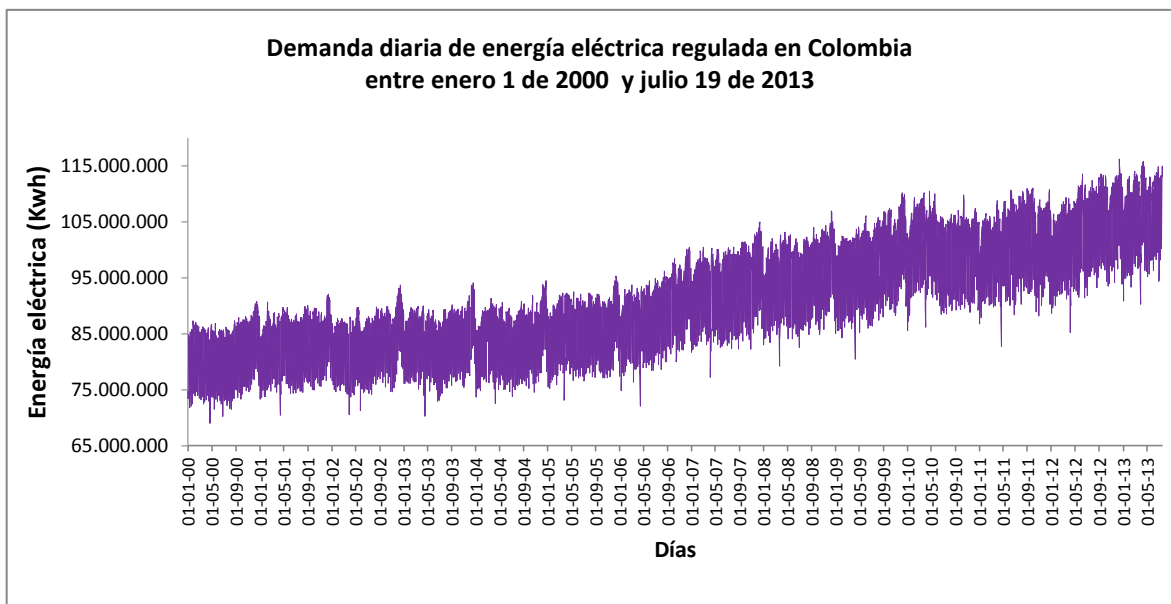
Con respecto a la variación periódica anual, los gráficos 3 y 4 muestran que la demanda de energía disminuye considerablemente el primer día del año, la noche de Navidad y en Semana Santa, festividad que no cae siempre en el mismo periodo del año.

Gráfico 1. Demanda diaria de energía eléctrica no regulada, $\{y_{NR,t}\}$



Fuente: Neón-X.M. Expertos en Mercados. Filial de ISA S.A. E.S.P. Cálculos propios.

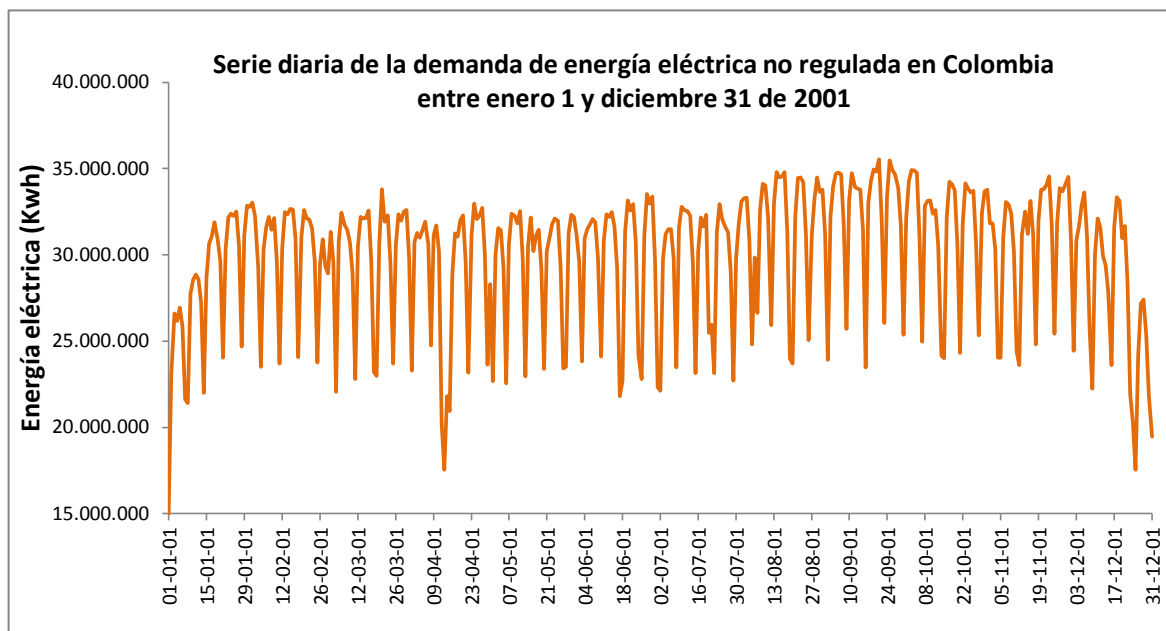
Gráfico 2. Demanda diaria de energía eléctrica regulada, $\{y_{R,t}\}$



Fuente: Neón-X.M. Expertos en Mercados. Filial de ISA S.A. E.S.P. Cálculos propios.

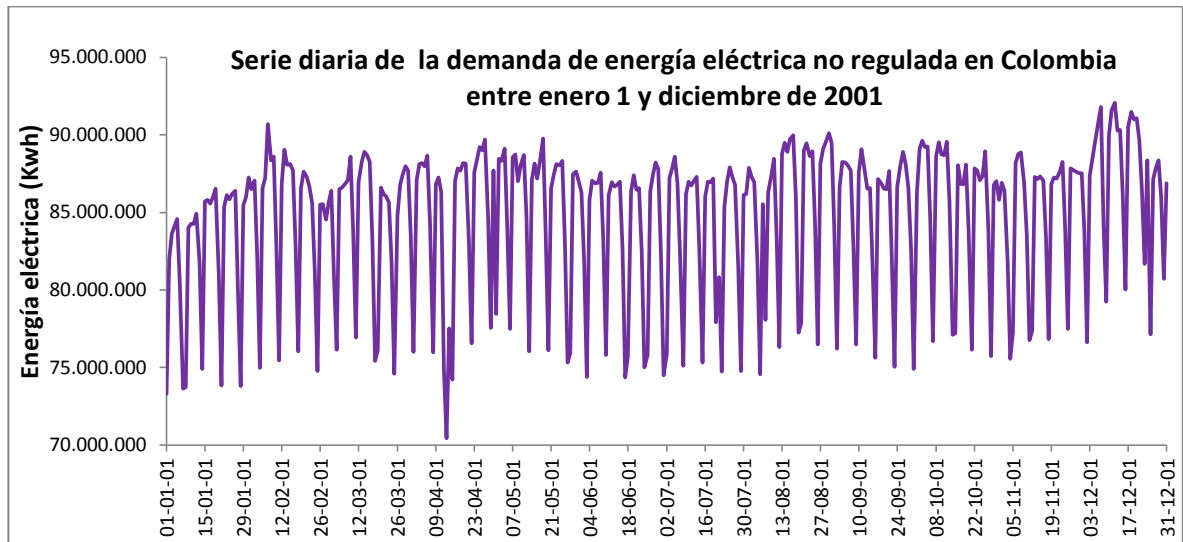
Con respecto a la variación de periodo semanal, los gráficos 5 y 6 muestran que la demanda de energía es mayor en los días hábiles (lunes a viernes) que los días sábado -donde la demanda disminuye- y domingo, donde la demanda de energía presenta el menor valor de la semana. Además, muestran que en los días festivos la demanda presenta una disminución variable de acuerdo al día de la semana. Si se trata de lunes festivo (festividad móvil), la demanda se incrementa relativamente respecto del domingo, sin superar el nivel de un lunes hábil, por ejemplo, véase el 10 de enero de 2000 en los gráfico 5 y 6; y si es un festivo fijo (cualquier día de la semana excepto lunes), la demanda de energía disminuye, como es el caso de Semana Santa (abril 14 de 2001) en los gráficos 3 y 4.

Gráfico 3. Demanda diaria de energía eléctrica no regulada entre enero 1 y diciembre de 2001



Fuente: Neón-X.M. Expertos en Mercados. Filial de ISA S.A. E.S.P. Cálculos propios.

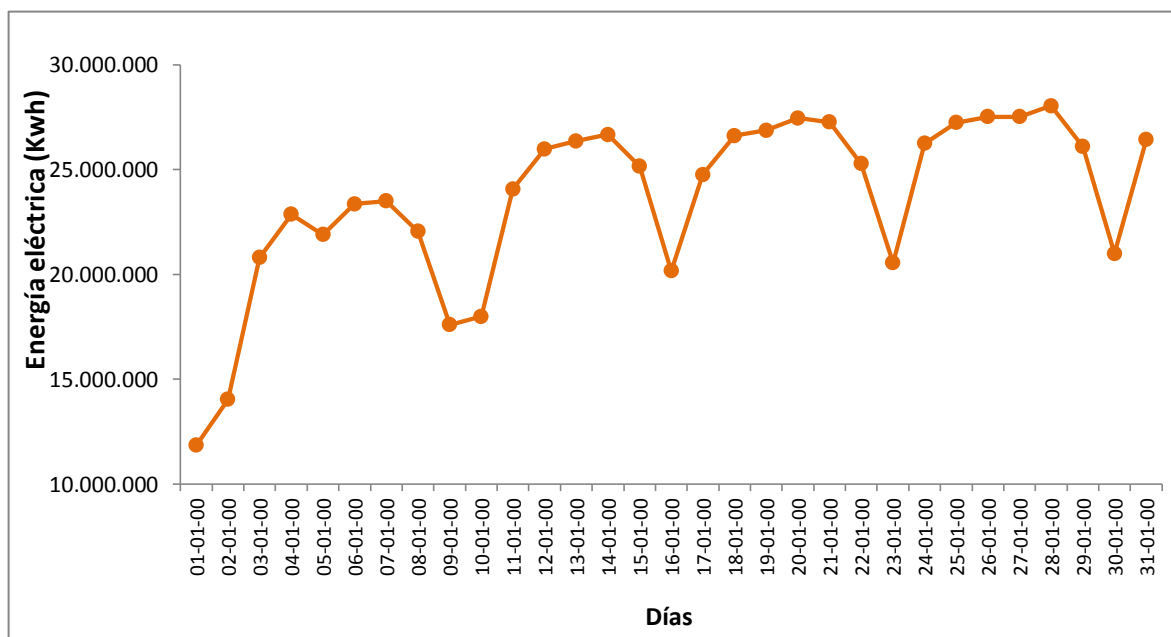
Gráfico 4. Demanda diaria de energía eléctrica regulada entre enero 1 y diciembre de 2001



Fuente: Neón-X.M. Expertos en Mercados. Filial de ISA S.A. E.S.P. Cálculos propios.

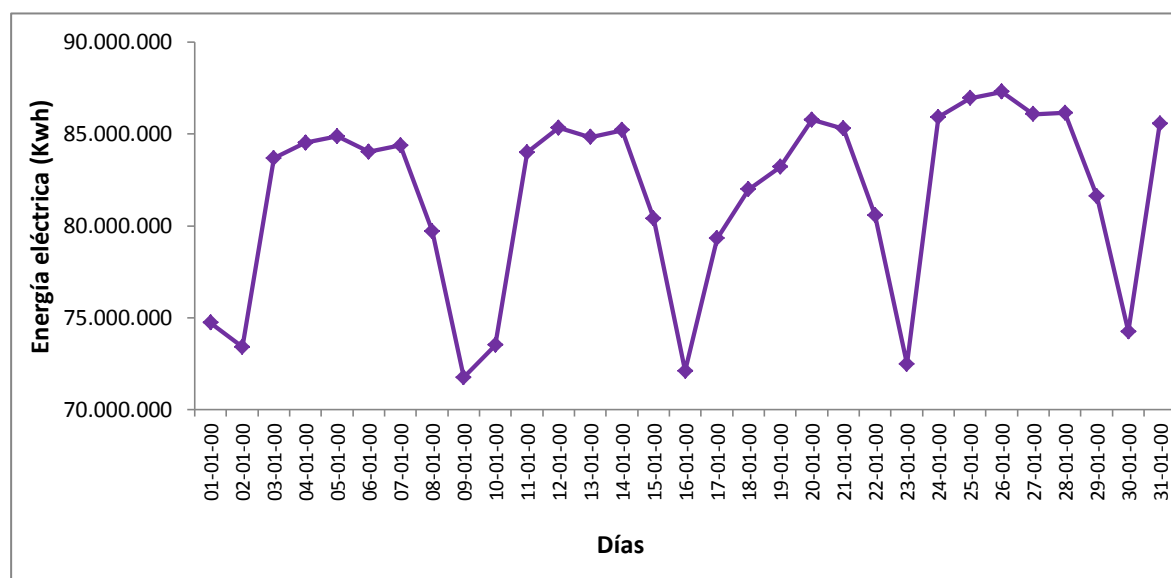
La fluctuación estacional semanal es estable a través de los años, es decir no varía mucho, por tanto se considera un componente estacional determinista o fijo, modelado a partir de siete variables ficticias (ver ecuación 5). La estacionalidad anual no es siempre de la misma longitud debido a la composición de las semanas (y años bisiestos). Además, en una serie diaria, aunque se asuma la estabilidad de la variación de periodo anual a lo largo de los años, difícilmente podrá ser estimada a través de un modelo de variables cualitativas, debido a la gran cantidad de parámetros a estimar, uno para cada elemento del componente estacional. Por lo tanto, se aplica un método más parsimonioso, las funciones splines.

Gráfico 5. Demanda diaria de energía eléctrica no regulada en enero de 2000



Fuente: Neón-X.M. Expertos en Mercados. Filial de ISA S.A. E.S.P. Cálculos propios.

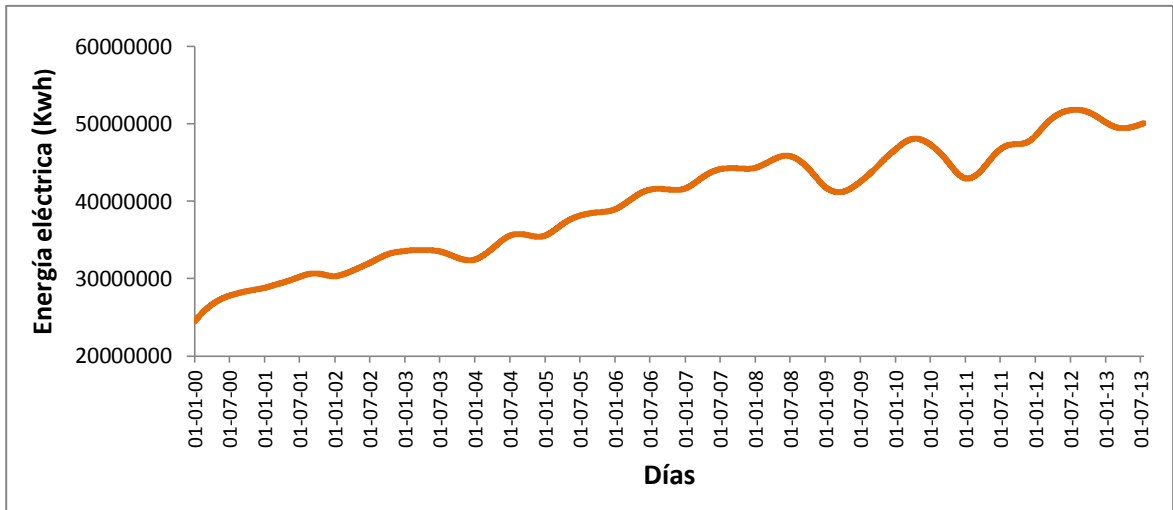
Gráfico 6. Demanda diaria de energía eléctrica regulada en enero de 2000



Fuente: Neón-X.M. Expertos en Mercados. Filial de ISA S.A. E.S.P. Cálculos propios.

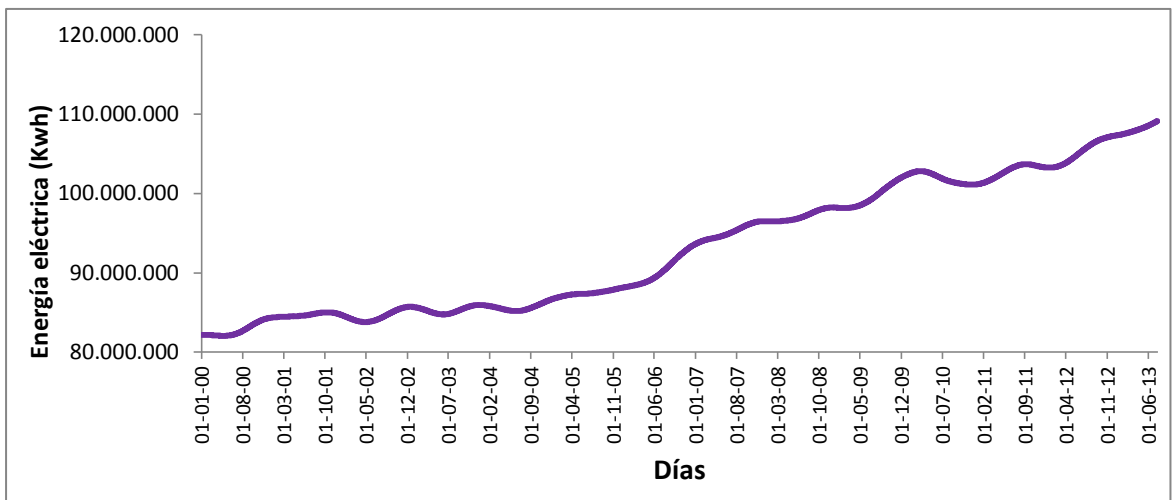
Dado el crecimiento observado en las series a través de los años, conviene remover dicho crecimiento para que no distorsione la extracción de las fluctuaciones estacionales semanal y anual. Con este fin, se aplica el filtro de Hodrick-Prescott que permite extraer el componente tendencial. Las gráficas 7 y 8 muestran la tendencia de las series de demanda regulada y no regulada, respectivamente.

Gráfico 7. Componente de Tendencia de $\{y_{NR,t}\}$



Fuente: Neón-X.M. Expertos en Mercados. Filial de ISA S.A. E.S.P. Cálculos propios.

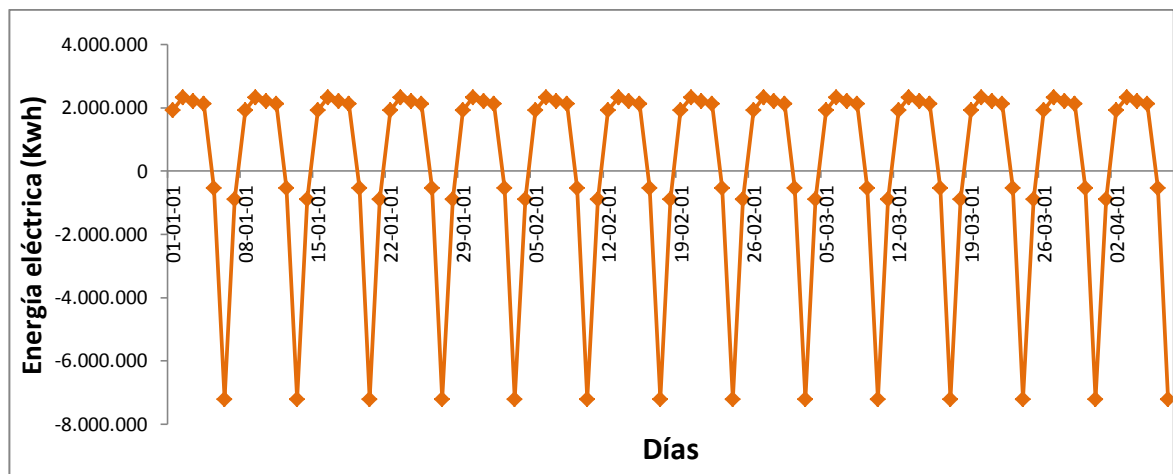
Gráfico 8. Componente de Tendencia de $\{y_{R,t}\}$



Fuente: Neón-X.M. Expertos en Mercados. Filial de ISA S.A. E.S.P. Cálculos propios.

Sobre el residuo estimado que quedó después de aplicar el filtro de Hodrick-Prescott, se aíslan las variaciones estacionales de periodo semanal aplicando el método de variables ficticias, tal que la suma sobre un periodo de siete días es cero, o lo que es lo mismo, la suma de los siete elementos del componente estacional semanal es cero. Las gráficas 9 y 10 muestran el componente estacional semanal estimado.

Gráfico 9. Componente de estacionalidad semanal de $\{y_{NR,t}\}$



Fuente: Neón-X.M. Expertos en Mercados. Filial de ISA S.A. E.S.P. Cálculos propios.

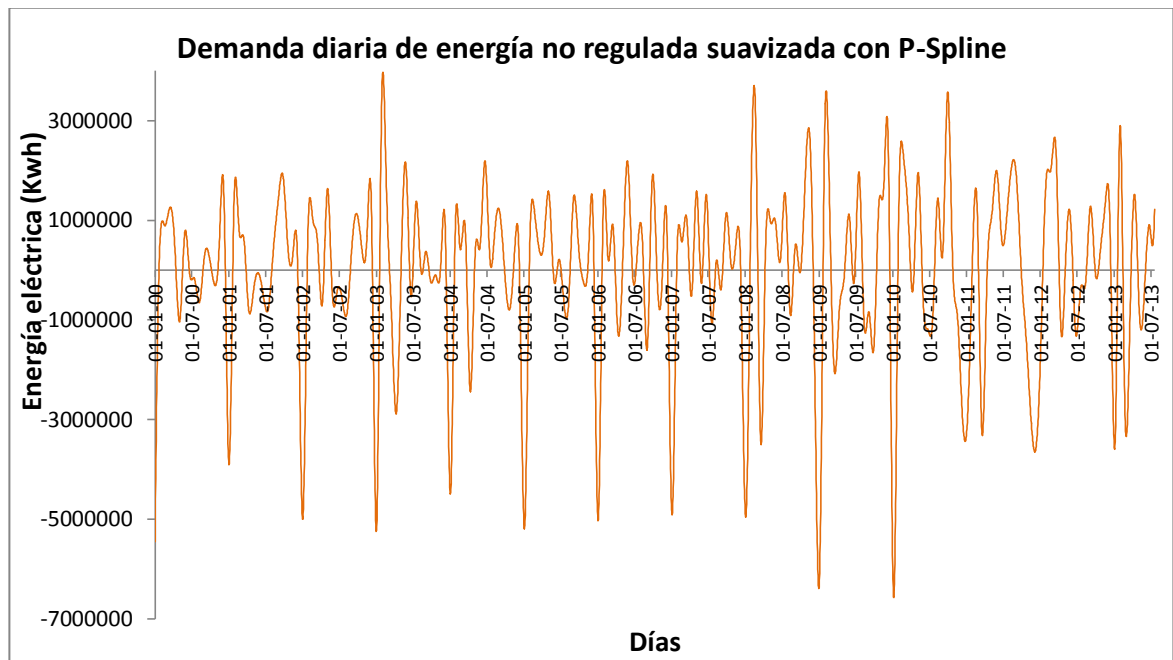
Gráfico 10. Componente de estacionalidad semanal de $\{y_{R,t}\}$



Fuente: Neón-X.M. Expertos en Mercados. Filial de ISA S.A. E.S.P. Cálculos propios.

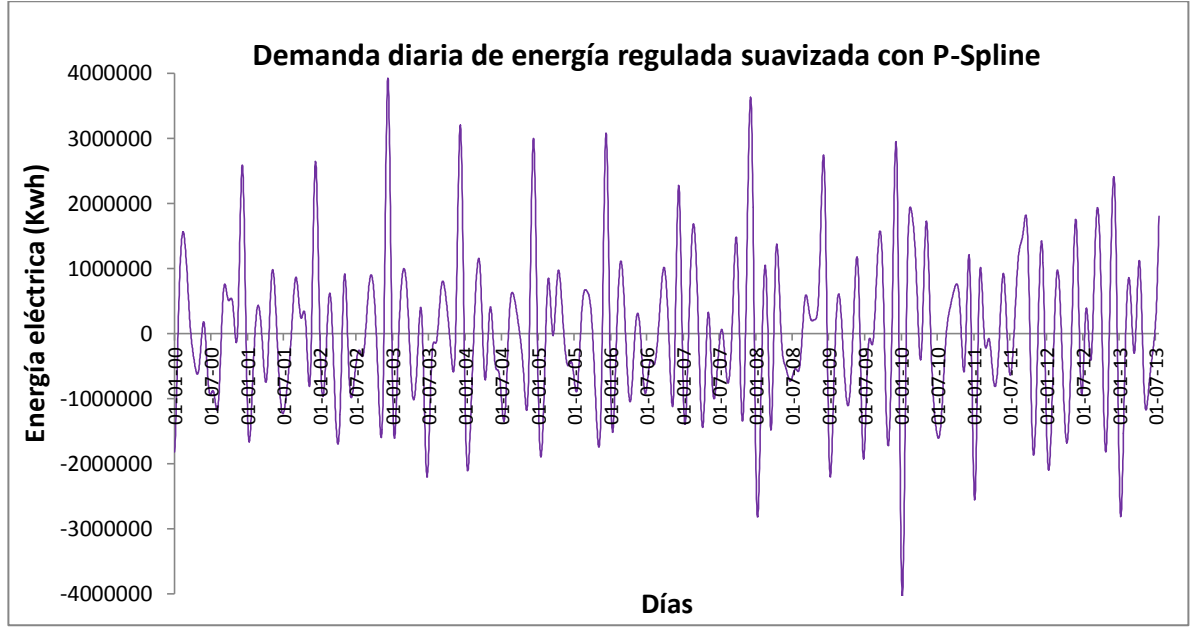
Sobre el error de tendencia y estacionalidad semanal, se aplica el método spline penalizado calculado partir de una base de funciones B-splines, donde el parámetro de suavización optimo se obtuvo mediante el CVG. Se usa la función `smooth.spline()` existente en el software estadístico R y se obtienen curvas suavizadas de las series de demanda regulada y no regulada, libres de tendencia y estacionalidad semanal. Dichas curvas equivalen al componente estacional de periodo anual de las series. En forma respectiva, los gráficos 11 y 12 muestran el componente estacional anual, S_t^{anual} , de la demanda no regulada y regulada.

Gráfico 11. Componente de estacionalidad anual de $\{y_{NR,t}\}$



Fuente: Neón-X.M. Expertos en Mercados. Filial de ISA S.A. E.S.P. Cálculos propios.

Gráfico 12. Componente de estacionalidad anual de $\{y_{R,t}\}$



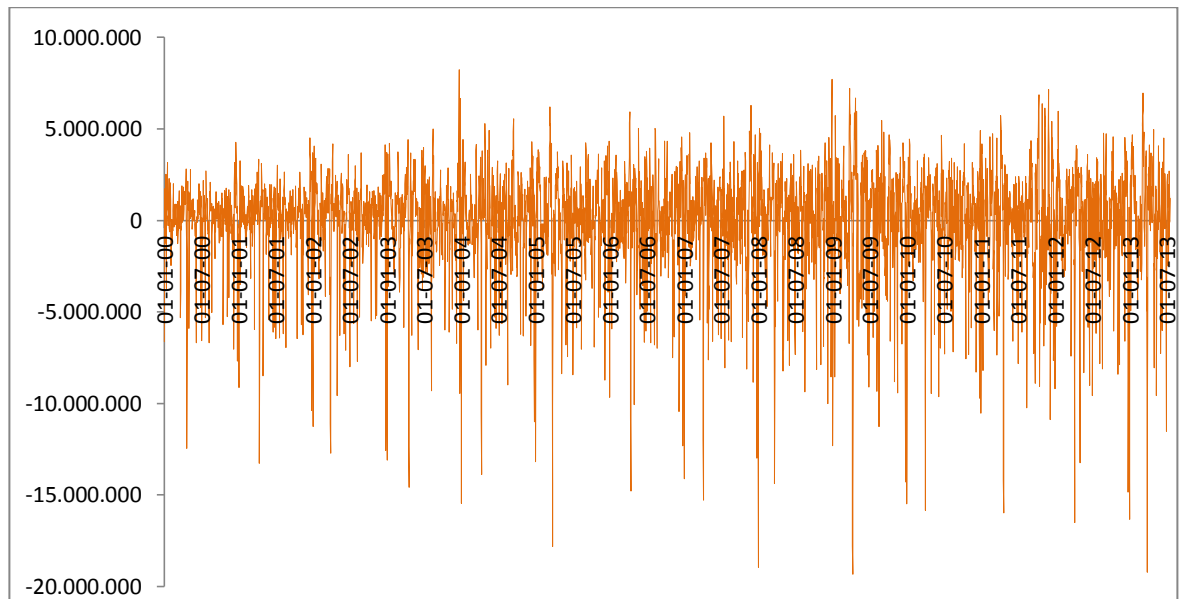
Fuente: Neón-X.M. Expertos en Mercados. Filial de ISA S.A. E.S.P. Cálculos propios.

Al intentar extraer el componente cíclico de las series se encuentra que éste ya no está presente en el error libre de tendencia y estacionalidad tanto semanal como anual. Se infiere que este resultado se debe a que el spline cúbico está capturando las variaciones recurrentes periódicas y no periódicas. En consecuencia, el componente cíclico no aparece explícitamente como un elemento del modelo y la ecuación (1) se reduce a

$$y_t = T_T + S_t^{semanal} + S_t^{anual} + \varepsilon_t, \quad \varepsilon_t \sim NID(0, \sigma_\varepsilon^2) \quad (8)$$

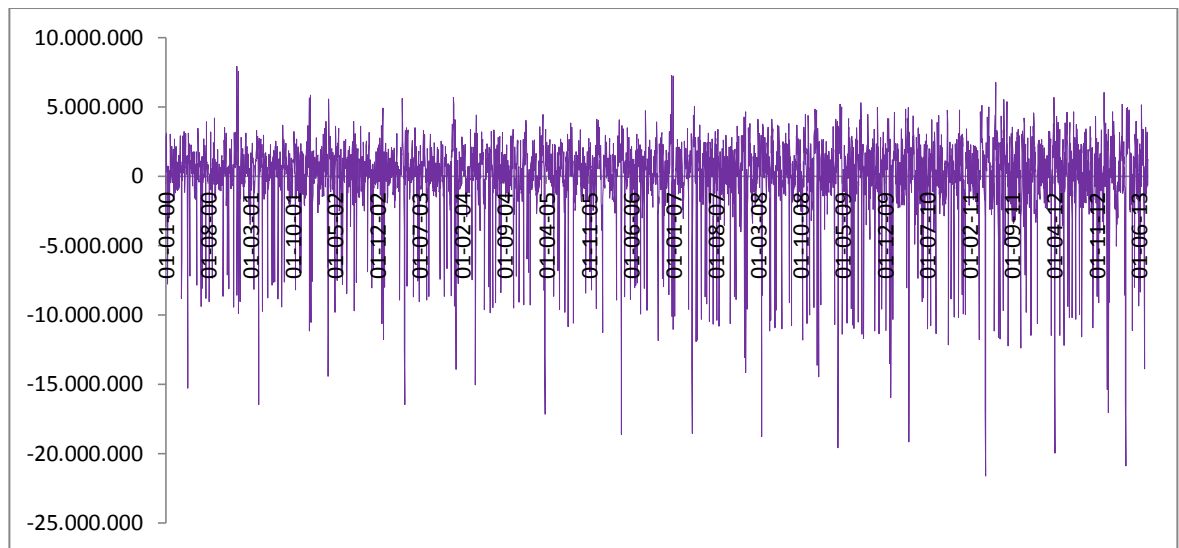
De esta manera, se tiene que el error libre de tendencia y estacionalidad semanal y anual de las series de demanda regulada y no regulada, corresponde al componente irregular o ruido de las series. Los gráficos 13 y 14 muestran el componente irregular de la demanda no regulada y regulada, respectivamente.

Gráfico 13. Componente Irregular de $\{y_{NR,t}\}$



Fuente: Neón-X.M. Expertos en Mercados. Filial de ISA S.A. E.S.P. Cálculos propios.

Gráfico 14. Componente Irregular de $\{y_{R,t}\}$



Fuente: Neón-X.M. Expertos en Mercados. Filial de ISA S.A. E.S.P. Cálculos propios.

COMENTARIOS FINALES

Los resultados de la estimación del modelo estructural muestran características relevantes de la demanda diaria de energía eléctrica regulada y no regulada en Colombia. En primer lugar, se destaca el carácter complejo y fluctuante de las series de demanda, donde se tiene por un lado, el comportamiento diferenciado entre días de la semana, el cual constituye la estacionalidad semanal, y por otro, el comportamiento de la demanda en los días festivos, considerados como estacionalidad anual porque las festividades se presentan una vez en el año. De esta forma, la consideración por separado de las fluctuaciones estacionales de periodo semanal y anual permitió identificar los puntos extremos del consumo de energía según el día de la semana o el mes del año. De hecho, se aprecia que la demanda en los días no hábiles –festivos y fines de semana- es menor que en los días laborables o hábiles. En general, a lo largo de una semana, la demanda es más baja en los días no laborables, especialmente el día domingo, incrementándose el día lunes, excepto si éste es festivo, cuando la demanda se incrementa respecto del domingo pero en menor proporción que si se tratase de un lunes hábil. En resumen, la demanda de energía depende fuertemente del día de la semana, condicionado al tipo de día (hábil, festivo, fin de semana).

En segundo lugar, se ilustra la utilidad de los splines como herramienta para modelar patrones estacionales complejos. La formulación del componente estacional en términos de una función spline conlleva a una ganancia en términos computacionales que de otra forma, con variables ficticias, no podrían ser llevadas a cabo por los softwares estadísticos disponibles, pues el elevado número de variables explicativas o regresoras, causaría problemas computacionales.

Vale la pena resaltar que el modelo desarrollado en esta investigación tenía por objetivo encontrar una metodología para estudiar la dinámica histórica del consumo de energía, más que desarrollar una metodología para fines de pronóstico. En este sentido, dado que no se incorporan variables explicativas que pueden afectar la demanda, como lo es la temperatura, el modelo planteado no es adecuado para desarrollar proyecciones de corto y largo plazo.

Una posible extensión de este trabajo consiste en la modelación de la demanda de energía eléctrica por sectores económicos y regiones, considerando el comportamiento estacional, lo cual permitiría una planificación precisa de la generación de energía. Además, se puede intentar modelar la relación entre la demanda de energía y variables macroeconómicas, lo cual sería un paso para la construcción de un indicador coincidente de la actividad económica por el lado de la oferta, líder en su estimación, dada la disponibilidad en tiempo real de las cifras de consumo de electricidad.

Agradecimientos

Agradecimientos a X.M. Expertos en Mercados. Filial de ISA S.A. E.S.P. por la información suministrada. Igualmente, agradecimientos a Julio Escobar, director de la presente investigación, por sus valiosos comentarios.

ANEXOS

I. Prueba de Raíz Unitaria sobre el componente irregular

Tes de Dickey-Fuller Aumentado				
Prueba	Hipótesis nula	Valor-p	Decisión	Conclusión
Componente Irregular de la demanda de energía no regulada	No estacionariedad	0.000	Rechazo Ho	Estacionaria
Componente Irregular de la demanda de energía regulada	No estacionariedad	0.000	Rechazo Ho	Estacionaria

* Al 5% de significancia estadística

Tras aplicar el test de Dickey-Fuller, se concluye que el componente irregular de las series de demanda de energía regulada y no regulada es estacionario en media.

II. Prueba de Normalidad sobre el componente irregular

Test de Shapiro-Wilk					
Variable	Hipótesis nula	z-estadístico	p-valor	Decisión*	Conclusión
Componente Irregular de la demanda de energía no regulada	Normalidad	15.607	0.0000	Rechazo Ho	No Normalidad
Componente Irregular de la demanda de energía regulada	Normalidad	17.036	0.0000	Rechazo Ho	No Normalidad

El componente irregular de la demanda de energía regulada y no regulada no cumple con el supuesto de Normalidad. Los gráficos 13 y 14 muestran el comportamiento del componente irregular de las dos series. Se observan clusters de volatilidad¹⁴, igual que en las series financieras que también son series temporales de alta frecuencia. Estas características del componente irregular de la demanda de energía motivan el uso de modelos GARCH, los cuales no desarrollados en la presente investigación.

REFERENCIAS

Barrientos, A.F, Olaya, J. y González, V.M. (2007). "Un modelo spline para el pronóstico de la demanda de energía eléctrica", *Revista Colombiana de Estadística*, 30 (2), 187-202.

Blaconá, M. T. y Abril, J. C. (2000). "Modelo estructural de espacio de estado para la demanda diaria promedio de energía eléctrica en la república Argentina", en 'Trabajo Presentado en la Reunión de la Asociación Argentina de Economía Política (AAEP)', Asociación Argentina de Economía Política.

Durbán, M. (2009). "An introduction to smoothing with penalties: P-splines". *Boletín de Estadística e Investigación Operativa*, 3, 195-205.

Eilers, P. y Marx, B. (1996). "Flexible smoothing with B-splines and Penalties" *Statistical Science*, 11, 89-121.

Engle, R. F., Granger, C. W. J., Rice, J. E., and Weiss, A. (1986), "Semi- parametric Estimates of the Relation Between Weather and Between Weather and Electricity Sales," *Journal of the American Statistical Association*, 81, 310-320.

¹⁴ Los Clusters de volatilidad son periodos de alta (baja) volatilidad tienden a venir seguidos por otros periodos de alta (baja).

- Franco, C. J., Velásquez, J. y Olaya, Y. (2008). "Caracterización de la demanda mensual de electricidad en Colombia usando un modelo de componentes no observables", *Cuadernos de Administración de Bogotá*, 21, 221-235.
- Harvey, A. C., and Koopman, S. J. (1993), "Forecasting Hourly Electricity Demand Using Time-Varying Splines", *Journal of the American Statistical Association*, 88, 1228-1236.
- Harvey, A. C., Koopman, S. J. y Riani, M. (1997), "The Modeling and Seasonal Adjustment of Weekly Observations", *Journal of Business and Economic Statistics*, 15, 354-368.
- Hillmer, S. C., and Tiao, G. C. (1982). "An ARIMA-Model-Based Approach to Seasonal Adjustment", *Journal of the American Statistical Association*, 77, 63-70.
- Lozano, I. y Rincón, H. (2010). "Formación de las Tarifas Eléctricas e Inflación en Colombia" *Borradores de Economía*, 634, 1-47.
- Martín, G. (2002). *Modelos estructurales y estacionalidad en series temporales económicas de alta frecuencia*. Tesis doctoral inédita, Universidad de la Laguna.
- Murillo, J., Trejos, A. y Carvajal, P. (2003). "Estudio del pronóstico de la demanda de energía eléctrica utilizando modelos de series de tiempo", *Scientia et Technica*, 23, 37-42.
- Perez, A. J., y Acosta, C. E. (1987). "Modelos de demanda residencial e industrial de electricidad para Colombia y estimaciones de elasticidades precio", *Revista Desarrollo y Sociedad*, 19, 69-93.
- Poirier, D. J. (1973). "Piecewise Regression Using Cubic Splines", *Journal of the American Statistical Association*, 68, 515-524.

- Poirier, D. J., Hendricks, W. y Koenker, R. (1979). "Residential Demand for Electricity: An Econometric Approach", *Journal of Econometrics*, 9, 33-57.
- Ramirez, G. A. (1991). "La Demanda de Energía Eléctrica en la Industria Colombiana", *Revista Desarrollo y Sociedad*, 27, 121-139.
- Rodríguez, G. M. (2008). "Splines como función de la proporción del período estacional" Documento de trabajo
- Rueda, M. V. (2011). *Predicción del consumo de energía en Colombia con modelos no lineales*. Tesis de Maestría, Universidad Nacional de Colombia, Sede Medellín.
- Wegman, E. J. y Wright, I. W. (1983). "Splines in Statistics", *Journal of the American Statistical Association*, 78, 351-365.
- Zapata, J. (2011). *Impacto del PIB, del Gas Natural y de los Precios de la Electricidad, en el Consumo De Energía Eléctrica en Colombia*. . Tesis de Maestría, Universidad EAFIT, Sede Medellín.