

Modelo de previsión de demanda de energía eléctrica para la optimización del despacho energético

FSDA_1_2017_1_143604

Contenido

1. Resumen ejecutivo	2
2. Introducción	3
3. Consideraciones previas.....	4
3.1. Información base y recursos	4
3.1.1. Recursos informáticos	4
3.1.2. Datos de demanda	4
3.1.3. Datos de temperatura	5
3.1.4. Otras series de datos.....	5
3.1.5. Selección de las muestras	5
4. Metodología	8
4.1. Tramos horarios y energía diaria por tramo	9
4.2. Crecimiento vegetativo o tendencia	10
4.3. Construcción de serie de demanda de “día hábil”	12
4.3.1. Tratamiento de días feriados	13
4.3.2. Sustracción del ciclo semanal.....	13
4.4. Construcción de la CEGH.....	14
4.4.1. Series de temperaturas	14
4.4.2. Configuración del Análisis Serial	17
4.5. Metodología para la identificación del comportamiento horario	21
5. Configuración de la sala y el actor Demanda	22
6. Conclusiones.....	23
7. Referencias.....	24

1. Resumen ejecutivo

En el marco del proyecto ANII **FSDA_1_2017_1_143604** titulado “*Modelo de previsión de demanda de energía eléctrica para la optimización del despacho energético*”, se combinaron esfuerzos desde distintas instituciones - UTE (Administración Nacional de Usinas y Trasmisiones Eléctricas), ADME (Administración Nacional del Mercado Eléctrico) e IIE (Instituto de Ingeniería Eléctrica, Facultad de Ingeniería, Universidad de la República) – para buscar soluciones alternativas a las ya existentes a la dificultad de prever la demanda de energía eléctrica en el sistema eléctrico uruguayo.

Conformando un equipo de investigación interdisciplinario e interinstitucional, desde dentro del sistema productivo, se trabajó en el análisis pormenorizado de datos para el posterior desarrollo de una herramienta que satisficiera necesidades de los equipos técnicos que trabajan diariamente para optimizar el despacho energético y garantizar el suministro al mínimo costo para el país, así como, conjuntamente con la previsión de generación de energías eólica y solar y de los aportes hidráulicos, determinar con más certeza y menor riesgo la negociación anticipada de exportaciones de energía a los países vecinos, permitiendo maximizar los beneficios.

La principal variable a analizar fue la demanda de energía eléctrica con paso horario y diario. Los análisis aplicados a la misma fueron desde la evaluación a nivel nacional de la efectividad de la medida de cambio de horario aplicada hasta el año 2015 en términos energéticos, hasta la comparación del crecimiento vegetativo en las distintas horas del día y las estaciones del año. También se analizaron series de temperaturas horarias de distintas fuentes o estaciones meteorológicas [1], a partir de las cuales se elaboraron series de temperaturas máximas y mínimas diarias. Se estudió la dependencia que con ellas tienen las demandas horaria y diaria de energía eléctrica, arribando al uso de las primeras como variables explicativas de las segundas.

Trabajando en la plataforma SimSEE [2], cuya flexibilidad permite la incorporación relativamente sencilla de nuevas fuentes generadores de series sintéticas, se elaboró el modelo objetivo del presente proyecto: una serie de herramientas cuyo uso encadenado permite la generación de previsiones probabilísticas de demanda de energía eléctrica con paso diario, partiendo de previsiones de temperaturas diarias máximas y mínimas, así como información de energía a lo largo del día de los días anteriores a la previsión.

Los resultados obtenidos evidencian un desempeño adecuado de la herramienta, cumpliendo con las expectativas iniciales del proyecto.

2. Introducción

En un contexto de cambio de la matriz energética de nuestro país [3], se evidenció la necesidad de mejorar la precisión al estimar las variables claves para la optimización del despacho energético, en los equipos técnicos que trabajan en este tema, tanto en UTE (Administración Nacional de Usinas y Trasmisiones Eléctricas) como en ADME (Administración Nacional del Mercado Eléctrico). La fuerte incorporación de fuentes renovables no gestionables que se dio en el último decenio, que se sumaron a la ya importante participación de energía hidroeléctrica, hace cada vez más necesaria la toma de decisiones en base a probabilidades, considerando para el cálculo una aproximación estadística a las infinitas combinaciones posibles de disponibilidad de los diferentes recursos y requerimientos de la demanda.

Es por esta razón que hoy en día es necesario contar con un pronóstico probabilístico de demanda, así como también, de cara al futuro, con una herramienta de esta índole para afrontar las exigencias de una tarifa variable horaria, a lo que ya se está tendiendo en otras partes del mundo. Cuanto más precisa sea la estimación de la demanda esperada, junto con el pronóstico del resto de las variables estocásticas involucradas, mayor certeza se tendrá, por ejemplo, sobre la disponibilidad de energía exportable. Esta certeza permite reducir riesgos y maximizar beneficios, como los asociados con la negociación anticipada de exportaciones a los países vecinos.

El objetivo del proyecto es obtener un modelo para el pronóstico probabilístico de demanda en tiempo real, listo para su puesta en operación y definiendo a su vez los parámetros necesarios para el mantenimiento de su calibración con el transcurso del tiempo. Se propone abordar el desarrollo de un modelo asociado a la plataforma SimSEE, utilizada para la programación y seguimiento de la operación del sistema eléctrico, así como para realizar estudios de planificación. Cabe destacar que esta plataforma es de acceso libre.

El proyecto se propuso en un marco de colaboración interinstitucional entre UTE, ADME y el IIE-UdelaR, con una mirada desde el sistema productivo y con el apoyo académico de la universidad.

El modelo desarrollado es una primera aproximación al problema de representar y pronosticar la demanda para la programación semanal en un contexto de alta probabilidad de excedentes exportables de energía. A partir de un análisis profundo de los errores del modelo se podrán realizar mejoras en futuras versiones.

3. Consideraciones previas

Dentro de las consideraciones previas se incluye el análisis y tratamiento de los datos base para la posterior construcción de la herramienta. Algunas de las consideraciones previas listadas podrían leerse como limitaciones del modelado debidas a la información de base disponible y el tratamiento escogido para ésta.

3.1. Información base y recursos

3.1.1. Recursos informáticos

En cuanto a los recursos, el principal es la plataforma SimSEE toda, y particularmente el módulo de Análisis Serial [4]. También se utilizaron herramientas como planillas de cálculo, software libre especializado para análisis de datos y estadística (R) y Matlab para cálculos y programación de rutinas simples para su uso en el producto final.

3.1.2. Datos de demanda

Es preciso, en primer lugar, definir a qué medida o cálculo se le llama demanda. En lo que al despacho energético refiere, la demanda del Sistema Interconectado Nacional (SIN) usualmente se define como la potencia (o energía) que instante a instante es preciso suministrar para satisfacer las necesidades de consumo nacional. Dada la complejidad para su medición, se determina a través de un cálculo que consta de la suma de la energía generada por los grandes generadores de energía eléctrica de UTE conectados al sistema de transmisión (represas, parques eólicos, parques térmicos, parques solares, etc.), la energía generada por los generadores conectados al sistema de red de distribución, las importaciones menos las exportaciones de energía eléctrica¹², y la diferencia entre la energía entrante al sistema (o generación) y saliente (o consumo) de los cogeneradores³. Es preciso destacar que a lo largo de este texto se utilizará el término “demanda” para referirse tanto a la demanda instantánea (expresada en MW) así como a la energía en intervalos de tiempo (horas, días, años, etc).

Existen varias fuentes que registran o calculan la demanda del SIN en distintos intervalos de tiempo. La fuente primaria de datos de demanda que se utilizó en este estudio fue una versión revisada del SCADA (serie de datos que es diariamente revisada y corregida en el Despacho de Cargas de UTE, de aquí en más DCU). Para comparar y complementar la anterior se tomaron datos del Sistema de Medición Comercial (SMEC de DCU) y de ingresos manuales. Cuando la serie obtenida de la combinación de todas estas fuentes presenta datos que no reflejan el comportamiento de los consumidores, por ejemplo, frente a cortes de energía, se realizan modificaciones necesarias para evitar incorporar estas distorsiones a la serie final.

¹ Se destaca que interesa considerar como componente de la demanda solamente la importación para satisfacer el consumo del país. No se considera, por ejemplo, la energía que se importa para ser luego exportada (si ello devenga en una ganancia) pues no va a satisfacer necesidades de consumo nacional.

² La forma de cálculo propuesta tiene un problema cuando hay cantidades importantes de energía en Tránsito en el SIN, pues en ese caso las pérdidas por transmisión quedan registradas como demanda del SIN.

³ El consumo de los cogeneradores para sus procesos internos no se considera en la demanda ya que no es posible predecir las decisiones del agente sobre sus procesos internos.

3.1.3. Datos de temperatura

A nivel nacional, los organismos encargados del despacho de energía eléctrica⁴ han constatado con el pasar del tiempo que las temperaturas máximas y mínimas diarias, son las variables atmosféricas que más influyen en el comportamiento humano en lo que al consumo de energía eléctrica refiere. Por lo antedicho, ante la confección de una herramienta probabilística de previsión de demanda de energía eléctrica que requiere datos históricos para su elaboración, es necesario contar con series históricas de medidas de temperaturas a nivel del suelo en territorio nacional cuya calidad esté comprobada.

En la propuesta original del presente proyecto se detallaba que los datos a utilizarse provendrían principalmente de la base de datos de ADME, que tiene acceso a las series históricas de temperatura de la estación Carrasco de INUMET. Como complemento a los datos de temperatura antes mencionados, se utilizó la serie de temperatura de la estación meteorológica propiedad de UTE instalada en las inmediaciones del DCU en Melilla, cuya calidad es verificada por especialistas de esta institución. Utilizando ambas fuentes de datos, se elaboró una serie de datos horarios de temperatura sin datos faltantes entre las fechas requeridas.

3.1.4. Otras series de datos

Se realizaron múltiples análisis de covarianzas entre las series de demanda y otras series de variables aleatorias además de las temperaturas ya mencionadas: series simuladas de generación eólica y solar, series de medidas de velocidad de viento y de radiación solar. Se variaron los pasos de tiempo de las series (promedios horarios, diarios, semanales) y también se buscaron relaciones no lineales entre las variables utilizando distintas potencias de la temperatura.

Se concluyó que el aporte de mayor peso para la explicación de la demanda, eran las series de temperaturas máximas y mínimas y las series construidas a partir de éstas, y se descartó el aporte de las series de otras variables aleatorias como sol o viento, cuyo aporte no resultó significativo, al menos con el nivel de precisión finalmente alcanzado. No se descarta que, en base a estudios más pormenorizados de casos particulares, pueda detectarse una eventual influencia de estas variables que permita mejorar la predicción en casos particulares y concretos.

3.1.5. Selección de las muestras

Previo a la selección de las muestras que serían el insumo fundamental para la elaboración del modelo, se hicieron estudios varios de modo de identificar los datos cuyo comportamiento se deseaba replicar. Algunos se mencionan a continuación, así como extractos de los resultados obtenidos.

⁴ Históricamente el Despacho Nacional de Carga de UTE (DCU) y desde la creación de la Administración del Mercado eléctrico (ADME), ésta con la colaboración de DCU a través de un contrato de arrendamiento de servicios.

Horario de verano

Se estudió el impacto del cambio de horario en épocas de verano que se aplicó hasta el año 2015. El comportamiento humano y productivo se veía modificado y, por ello resultaba preciso evaluar el comportamiento diario de la demanda de energía eléctrica (también llamado *curva de carga*).

En cuanto a la energía diaria, se concluyó que la variación en la correlación entre los datos diarios de energía eléctrica y la temperatura máxima y mínima del día, entre días con cambio de hora y días sin cambio de hora es imperceptible. En lo que respecta a la distribución horaria de la energía, se observó que la forma de la curva de carga diaria es perceptiblemente diferente entre días con y sin cambio de hora.

Como se dijo anteriormente, no se detectó una variación de importancia en la energía diaria entre días con y sin cambio de horario, por lo tanto, para la modelación de la energía diaria por tramo horario, se decidió que la serie de datos a utilizar sería la demanda horaria medida/calculada entre el 01/07/2012 00:00 y el 08/03/2019 23:00, es decir, se utilizarían datos de años con cambio de horario.

Sin embargo, para el modelado de las formas de las curvas, se decidió tomar la serie de datos de demanda comprendida entre el 09/03/2015 00:00 (día siguiente al domingo en que finalizó el último periodo con horario de verano) y el 08/03/2019 23:00 consiguiendo así tener 4 años móviles completos, evitando el efecto del cambio de forma de la curva de carga por el horario de verano.

Dentro del conjunto de datos utilizado se encuentran tres períodos donde se aplicó el cambio de horario de verano, que dejó de aplicarse a partir de marzo de 2015. Es difícil determinar con exactitud los efectos de este cambio sobre el monto de energía diaria y sobre la forma de la curva horaria. Para poder proceder a la identificación de un modelo, en esta instancia se despreció el efecto del cambio de horario sobre el monto de energía diaria, es decir, se tomaron en cuenta los valores reales de energía diaria sin modificación alguna.

En cuanto a la forma de la curva horaria se observó que los efectos eran significativos debido al desplazamiento horario que se registra en la mayoría de las actividades humanas asociadas al consumo eléctrico. En este sentido lo que se hizo fue eliminar los datos correspondientes al período de cambio de horario a los efectos de la identificación de las curvas base horarias.

En el marco de este proyecto, en el proceso de selección de muestras de datos para ser utilizados en la confección del modelo CEGH, se decidió que los datos claves para la elección del periodo histórico que ambos abarcarían, serían los de demanda y no los de temperatura. Esto se justifica por varias razones. La primera y principal, es que la variable que quiere predecirse es la demanda, utilizando la temperatura como variable explicativa, y no viceversa. Por otro lado, al querer captarse con la selección de datos comportamientos de índole social y/o antropomórfica, se seleccionó un periodo de datos de demanda que caracterizara el comportamiento de la demanda sin incorporar patrones de años anteriores que ya no reflejaran la demanda actual.

De lo antedicho que el periodo de datos de temperatura a ser considerado para la elaboración de esta herramienta y, por ende, para la realización de un análisis de calidad basándose en herramientas estadísticas estándar, viene dado por la selección de datos

históricos de demanda de energía eléctrica. El periodo escogido va desde el 01/07/2012 00:00 hasta el 08/03/2019 23:00.

Se analizaron los datos horarios de temperatura de Carrasco entre las fechas 10-05-2013 a la 01:00 y 29-11-2018 a las 00:00. Por su parte, Los datos de DCU analizados corresponden a las medidas horarias entre el 01-01-2015 00:00 y el 05-05-2019 23:00.

4. Metodología

La metodología de trabajo se compone de varias etapas y se describe a través del siguiente diagrama. Cada etapa se detalla posteriormente.

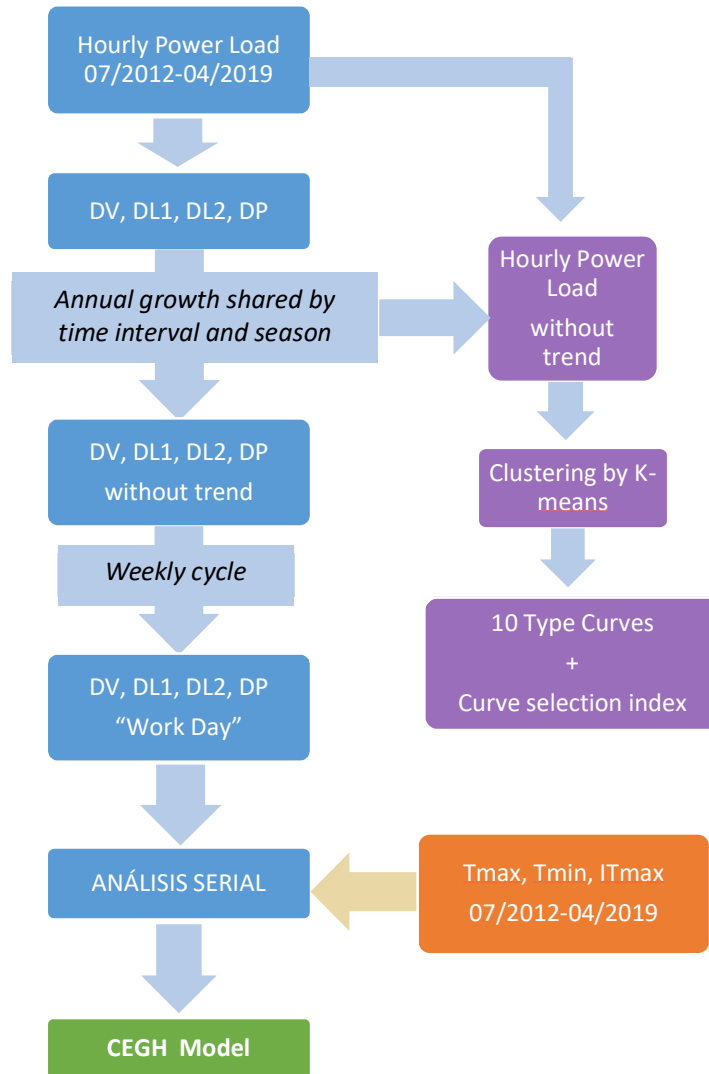


Ilustración 1. Diagrama de pasos seguidos para elaboración del modelo.

El diagrama anterior puede leerse de la siguiente manera:

- i. Se parte de la serie histórica de datos horarios de demanda entre las fechas 01-07-2012 al 08-03-2019. La calidad de dichos datos ya fue revisada como se detalló en 3.1.2 Datos de demanda.
- ii. Se definen cuatro tramos horarios que serán Valle, Llano 1, Llano 2 y Pico. Se elaboran series de energías diarias para cada tramo horario.
- iii. Se calcula el crecimiento anual para cada tramo horario y estación del año. Con dicha información se elaboran series de energía diaria por tramo sin tendencia o crecimiento vegetativo.
 - a. Extrayéndole a dichas series sin tendencia el ciclo semanal (entiéndase por ciclo semanal la variación promedio de potencia registrada entre los días hábiles y los fines de semana), se obtienen series de energía diaria por tramo horario sin tendencia de “día hábil”.
 - i. Combinando esta información con las series diarias de temperaturas máximas, mínimas y lo que se denomina iTmax, a través del Análisis Serial se obtiene un CEGH de energías diarias de día hábil por tramo horario.
 - b. Combinando la información de la composición del crecimiento anual por tramo horario y época del año, con las series originales de demanda horaria, se generan series de demanda horaria sin tendencia. Para evitar el efecto sobre la curva de carga diaria del horario de verano en los años 2012 a 2014, sólo se utilizan datos desde el año 2015, cuando por primera vez en los años de los datos, no se aplicó el cambio de horario.
 - i. Utilizando el método de clusterización por k-means, se agrupan las curvas de carga diarias en 10 clústers para cada tramo horario.

4.1. Tramos horarios y energía diaria por tramo

Si se observa la curva de carga diaria típica, se pueden reconocer tramos horarios donde el comportamiento del consumo es bien diferente; corresponden a momentos del día donde las actividades que gobiernan el uso de la energía eléctrica son diferentes y, por lo tanto, se estima que también es diferente su dependencia con la temperatura en el corto plazo y a lo largo del año. Similar a los tramos horarios (también llamados “postes”) que se utilizan regularmente por los operadores del SIN, se definen cuatro tramos horarios de 6 horas de duración cada uno, de la siguiente manera:

- i. Tramo 1, Valle, entre las 0:00 y las 5:59 hs de cada día
- ii. Tramo 2, Llano 1, entre las 6:00 y las 11:59 hs de cada día
- iii. Tramo 3, Llano 2, entre las 12:00 y las 17:59 hs de cada día
- iv. Tramo 4, Pico, entre las 18:00 y las 23:59 hs de cada día.

Las series de energías diarias por tramo horario se calculan simplemente acumulando las energías en las horas que corresponden a cada tramo.

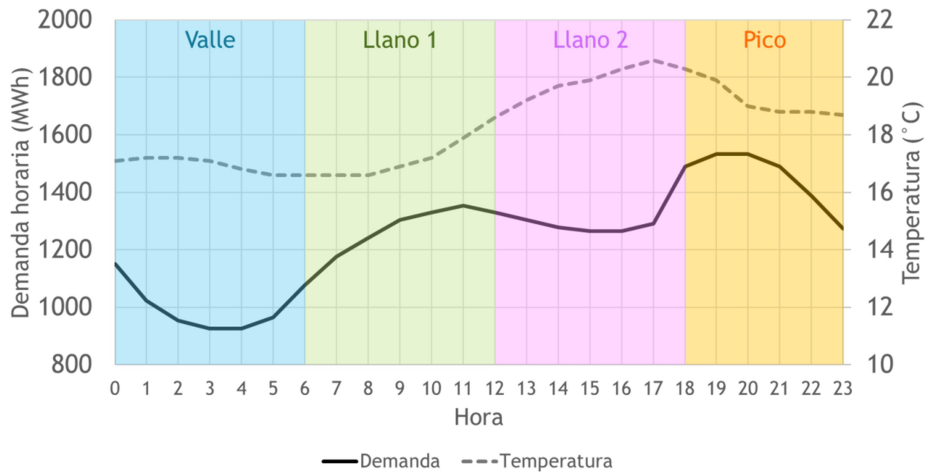


Ilustración 2. Superpuesto en el gráfico de demanda horaria y temperatura del día 08-05-2018, se colorean los cuatro tramos horarios definidos.

4.2. Crecimiento vegetativo o tendencia

El módulo Análisis Serial de SimSEE requiere que la serie histórica a representar (en este caso, los datos históricos de demanda de energía eléctrica en cada tramo horario), no tenga crecimiento inter-anual. Se sabe que la demanda de energía eléctrica varía a lo largo de los años por variadas razones: coyuntura económica nacional, desarrollos tecnológicos, cambios a nivel de comportamiento poblacional, etc. Por ello, fue necesario estudiar el crecimiento vegetativo de las series de demanda de energía eléctrica diaria por tramo horario.

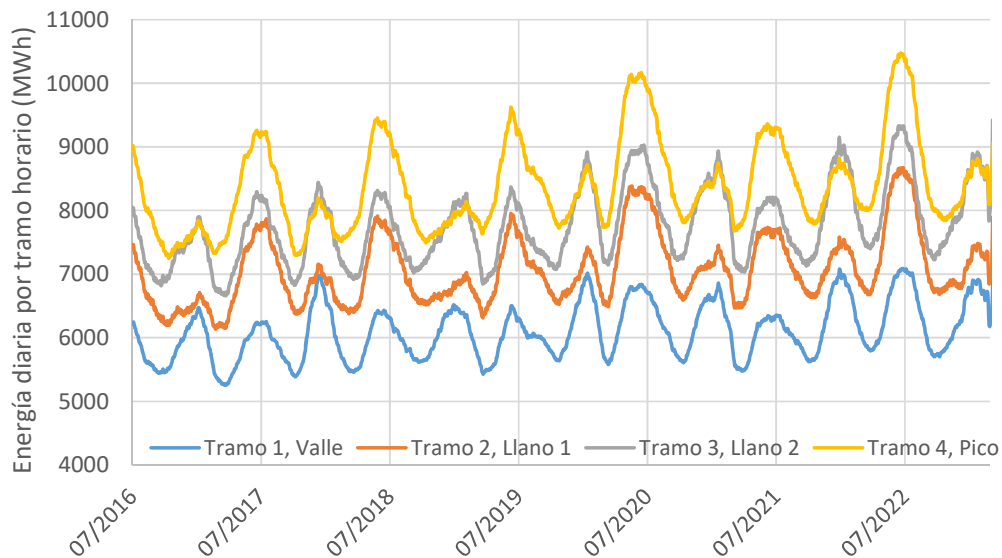


Ilustración 3. Promedio móvil de 45 días de la demanda diaria por tramo horario, para el conjunto completo de datos.

Luego de haber estudiado las series temporales, se concluyó que, habiéndose identificado que algunos tramos horarios habían crecido más que otros para un mismo año (por ejemplo, creció más el pico que el valle) y que el crecimiento era diferente en las distintas [10]

épocas del año (por ejemplo, creció proporcionalmente más la demanda en verano que en primavera), se decidió calcular los crecimientos por tramo horario y por época del año. Lo anterior puede observarse en la Ilustración 3.

Se definieron las épocas del año como se detalla a continuación:

1. Verano, durante los meses de diciembre, enero y febrero,
2. Otoño, durante los meses de marzo, abril y mayo,
3. Invierno, durante los meses de junio, julio y agosto,
4. Primavera, durante los meses de setiembre, octubre y noviembre de cada año.

Se aproximó el crecimiento interanual por tramo horario y época del año a través de ajustes lineales con el método de mínimos cuadrados. Se muestra a continuación el ajuste para el tramo 4 (pico) en la primavera.

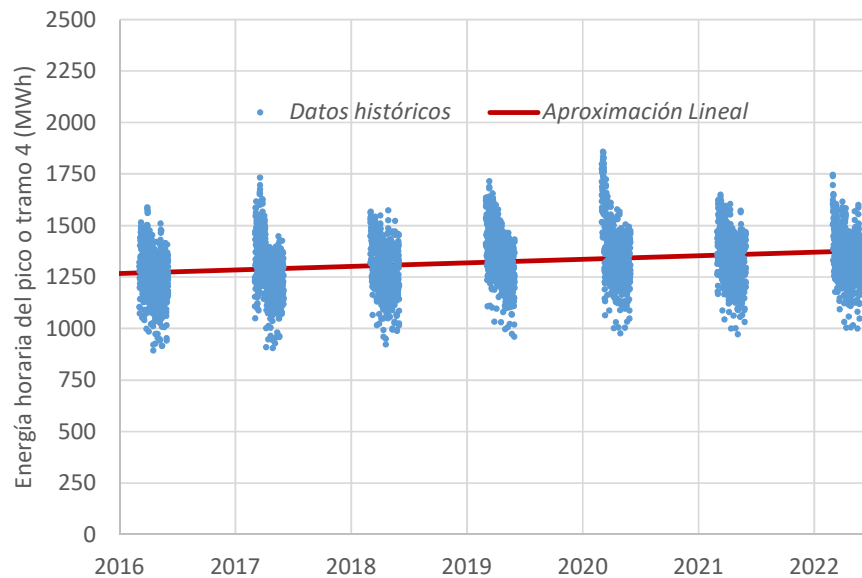


Ilustración 4. Ejemplo de aproximación lineal calculada para los datos horarios en horario de pico o tramo 4, durante la primavera.

Siendo el polinomio que mejor ajusta a los datos $y(x) = mx + b$, se procedió a elaborar series de energía horaria sin tendencia para todo el periodo de los datos. Cada medida de demanda horaria “se llevó a términos de energía del año 2012”, que era el origen de los datos.

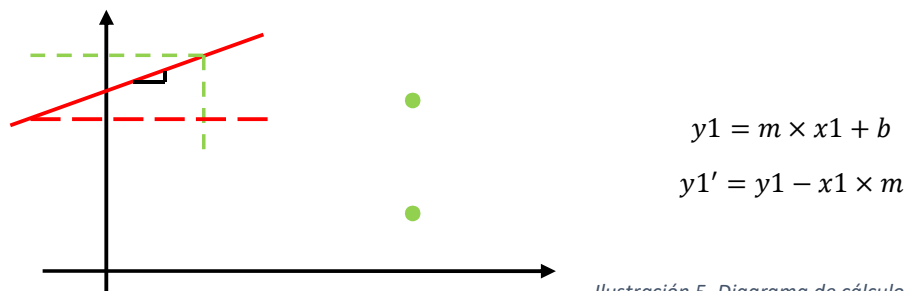


Ilustración 5. Diagrama de cálculo de datos

Con las series sin tendencia antes mencionadas, se elaboraron nuevamente series de energía diaria por tramo horario, ahora sin tendencia. Los valores de los términos m se muestran en la siguiente tabla por tramo horario y época del año. Al colorearlos según su valor, se puede concluir que el tramo horario que tuvo mayor crecimiento en una época del año específica fue el pico, y esto ocurrió en el invierno. El valle, para cualquier estación del año, fue el rango horario cuya demanda de energía eléctrica creció menos en el periodo de estudio.

Tabla 1. Coeficientes para quitar la tendencia a las medidas de demanda horaria histórica.

		SEASONS			
		SUMER	AUTUMN	WINTER	SPRING
Tramo horario	Valle	0,029	0,046	0,055	0,025
	Llano 1	0,051	0,048	0,068	0,035
	Llano 2	0,071	0,065	0,072	0,035
	Pico	0,070	0,050	0,087	0,046

4.3. Construcción de serie de demanda de “día hábil”

Es preciso destacar que para la utilización de la herramienta Análisis Serial [4], que captará las correlaciones de histogramas en el espacio Gaussiano entre series de datos (que serán la demanda de energía eléctrica por tramo horario y series diarias de temperaturas máximas y mínimas, e $iT_{máx}$, cuya construcción se detalla más adelante) es preciso despojar las series de toda aquella información que se sepa no es explicada por el resto de las variables involucradas, de modo que capte únicamente la dependencia de unas con otras y no con variables ajenas. Por ejemplo, el ciclo semanal de la demanda de energía eléctrica está fuertemente afectado por la construcción cultural y social de los días de la semana: de lunes a viernes son días “hábiles” y sábados y domingos, “fin de semana”, los dos últimos además diferentes entre sí.

A través de una simple inspección visual de los datos históricos de demanda de energía eléctrica (tanto horarios como diarios), se puede observar el ciclo semanal asociado más que nada a los usos de la energía y el ciclo semanal humano o cultural. Además, se detecta que los días feriados, tanto laborables como no laborables, tienen un comportamiento diferente al resto de los días del año, ya sean hábiles como fines de semana.

Por mencionado al comienzo de esta sección, es indispensable despojar a las series de estudio de toda información ajena a la dependencia entre ellas, como lo es el ciclo semanal y los días feriados. Cabe destacarse que no se debe extraer el ciclo anual (la demanda también depende fuertemente del momento del año, o sea, la estación) pues las variaciones que la demanda pueda tener con el momento del año, se espera sean explicadas por las series de temperaturas.

4.3.1. Tratamiento de días feriados

No se especificó un procesamiento especial para los días feriados laborables y no laborables. Actualmente no son tenidos en cuenta los datos correspondientes a días feriados laborables y no laborables, a menos que coincidan con un día domingo.

Como oportunidad de mejora a futuro, se deja planteado la realización de un análisis pormenorizado del comportamiento de cada tipo de feriado (laborable o no laborable), en relación también con el día de la semana en que se registra.

4.3.2. Sustracción del ciclo semanal

Se analizó el ciclo semanal, es decir, la afectación sobre la demanda de la sucesión de semanas con sus días hábiles y fines de semana. Se observó que el ciclo semanal en relación a la energía diaria de cada tramo horario, se mantiene independientemente de la estacionalidad anual. Es decir, el ciclo semanal no se ve afectado por los cambios de temperatura, o lo hace en grado despreciable e indetectable con las técnicas y herramientas utilizadas. Para lo anterior, se calculó, para cada día de la semana y estación del año, el cociente promedio de la demanda diaria por tramo y la demanda media del tramo de los días hábiles del total de los datos.

Esto permite establecer coeficientes multiplicadores de cada tramo horario de demanda para “trasladar” los valores de energía de cada día de la semana a una referencia única de energía de día hábil medio, definiéndose a partir de allí las relaciones demanda-temperatura para la definición del modelo. Los coeficientes utilizados fueron los promedios de todo el año (sin diferenciar estaciones) que se muestran en la Tabla 2.

	dom	lun	mar	mie	jue	vie	sab
valle	0.9458	0.9645	1.0071	1.0064	1.0082	1.0098	1.0021
llano1	0.7909	0.9926	1.0006	0.9996	0.9948	0.9927	0.8967
llano2	0.8121	1.0049	1.0062	1.0044	0.9996	0.9972	0.9092
pico	0.8985	1.0006	1.0020	1.0048	0.9997	0.9844	0.9326

Tabla 2. Multiplicadores de la demanda según día de la semana, referido al día hábil promedio del tramo horario correspondiente.

Los valores hallados son casi coincidentes con los presentados en [5] donde se trabajó con datos hasta el año 2010 expresando “En cuanto a la estacionalidad semanal, no se aprecia un cambio en este patrón a lo largo del período analizado. Más precisamente, para el total de la muestra el consumo promedio de sábado y domingo representa el 93% y 84% del promedio del consumo en de los días hábiles respectivamente, cifras que continúan representando el patrón de consumo intra-semanal en 2010.” Los valores presentados en la Tabla 2 prorrateados según cantidad de horas de participación en la demanda, equivalen a 93% para el sábado y 86% para el domingo, lo que muestra que esencialmente se mantiene el mismo comportamiento con un muy leve aumento para el consumo del domingo.

4.4. Construcción de la CEGH

Como ya se mencionó y como se detalla en [4] y [5], a través del módulo de SimSEE llamado Análisis Serial, es posible detectar la interdependencia o covarianzas entre series de datos en el espacio Gaussiano, y a partir de ellas, elaborar un generador de series sintéticas de una o más variables involucradas, a partir de sus variables explicativas. Particularmente en este proyecto, se busca encontrar variables explicativas de la variable demanda de energía eléctrica consumida en cuatro tramos horarios preestablecidos.

4.4.1. Series de temperaturas

Como ya se mencionó en 3.1.3 Datos de temperatura, estas series son de gran importancia para explicar el comportamiento de la demanda diaria de energía eléctrica. En particular, se detectó que la mejor forma de captar la covarianza entre ambas era a través de las temperaturas máximas y mínimas diarias, y la demanda diaria en cada tramo horario de los definidos previamente.

Como forma de aproximarse a la dependencia entre las variables antes de la confección de la CEGH, se elaboran las siguientes ilustraciones. Por ejemplo, en la Ilustración 6 se observa que existe una clara dependencia de la demanda diaria con la temperatura media del día. La forma de “U” se debe a que el aumento de la demanda se da tanto por la necesidad de calefacción en épocas de temperaturas bajas, como por la necesidad de refrigeración en épocas de temperaturas elevadas, en adición a los niveles básicos de consumo no influenciados (o no tan directamente influenciados) por la temperatura. Al observarse el mismo conjunto de datos separando por época del año e identificando a través de colores el tipo de día (según sea hábil, sábado o domingo) en la Ilustración 8, se observa que en las estaciones de temperaturas máximas y mínimas alcanzadas (verano e invierno respectivamente), los comportamientos son bien diferenciados y “opuestos” en sus pendientes, mientras que para las estaciones intermedias el comportamiento es variable, sobretodo en la primavera donde se mantiene la forma de “U” de la curva.

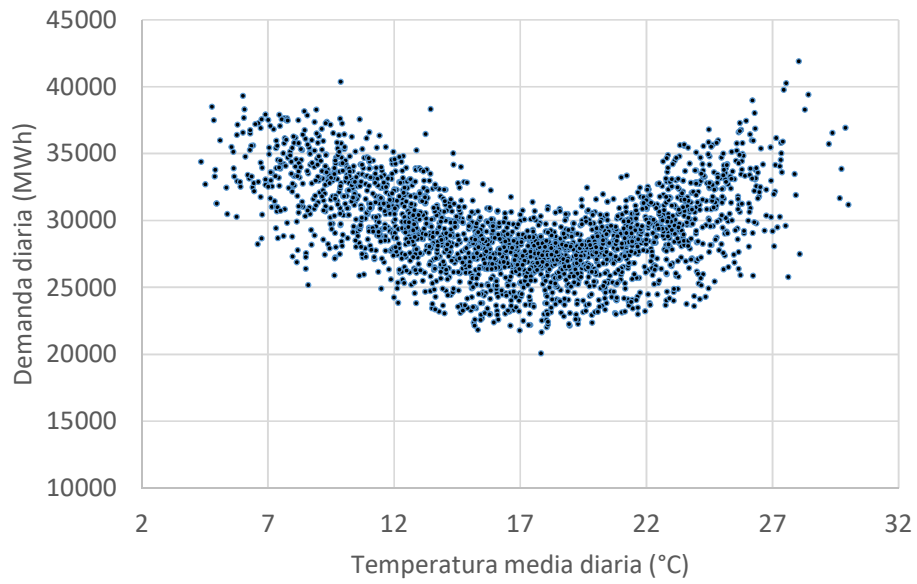


Ilustración 6. Demanda diaria (excluyendo feriados) en función de la temperatura media diaria.

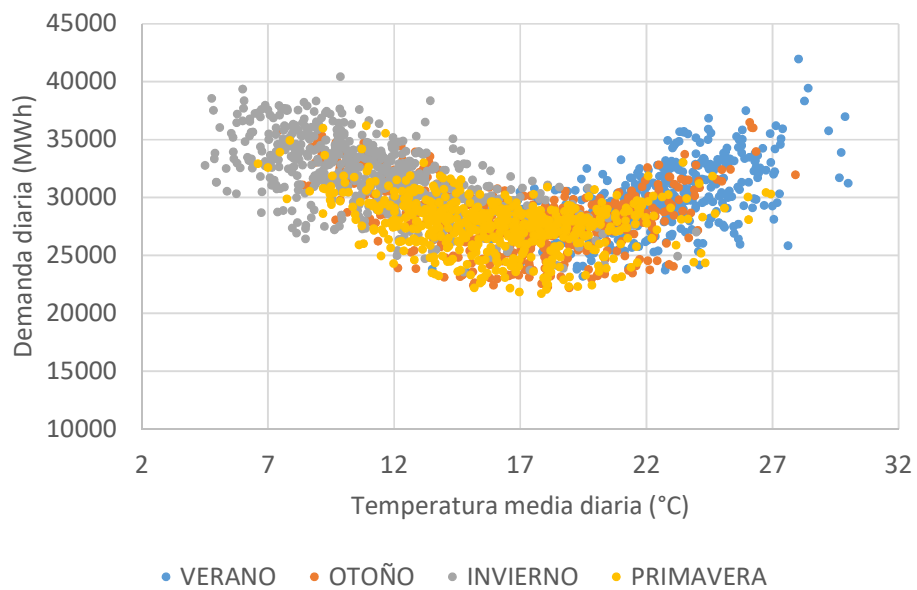


Ilustración 7. Demanda diaria (excluyendo feriados) en función de la temperatura media diaria, diferenciando por época del año.

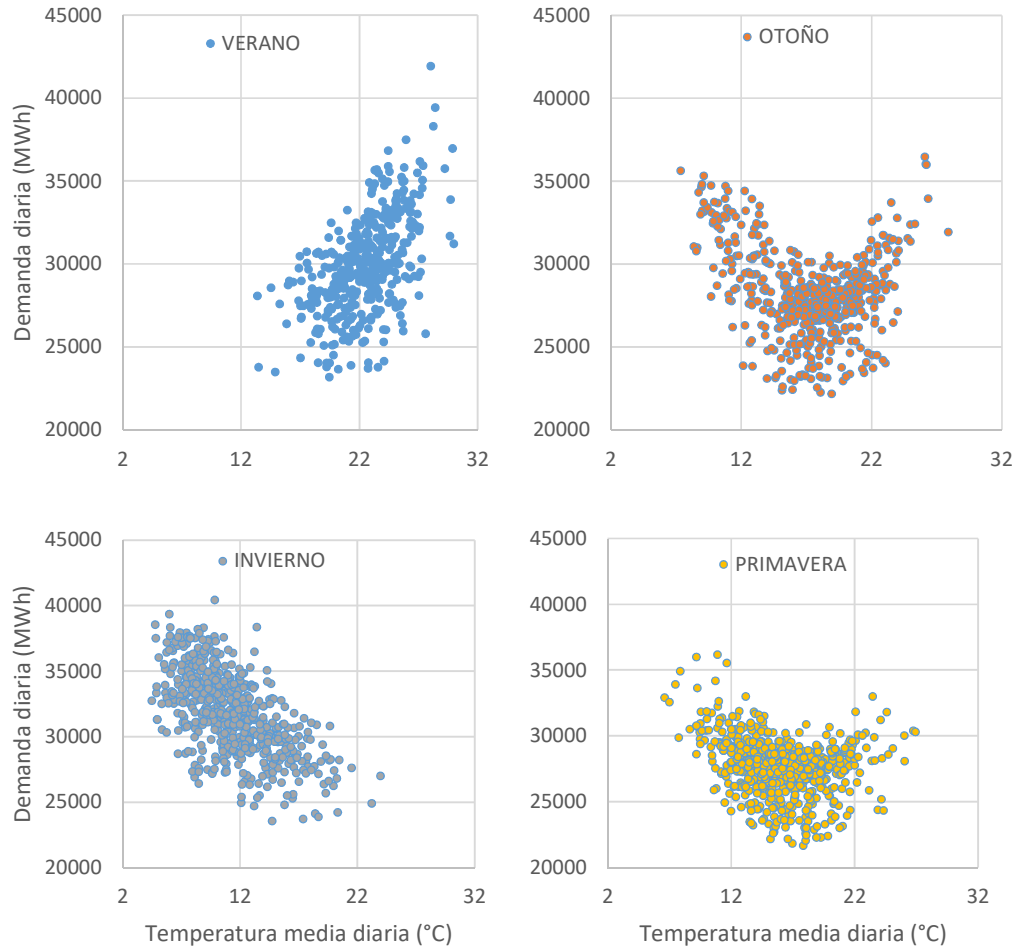


Ilustración 8. Demanda diaria en función de la temperatura para las cuatro estaciones del año.

De la Ilustración 7 (también Ilustración 8) se desprende que las relaciones buscadas entre la demanda y la temperatura para predecir su comportamiento, son diferentes a lo largo del año. También se observa que las variaciones de la demanda no dependen exclusivamente del valor de la temperatura, sino que también, de la estación del año. Por ejemplo, la pendiente de la demanda entre los 20° y los 25° es de 0,4 GWh/día/°C en primavera aproximadamente (va de 25 a 27 GWh/día) y de 0,8 GWh/día/°C en verano aproximadamente (va de 24,5 a 28,5 GWh/día). Algo similar sucede al comparar una diferencia de 5°C de temperatura en otoño e invierno, por ejemplo, entre 10° y 15°.

En conclusión, el modelo a identificar debe tener en cuenta la estacionalidad anual y no solo la temperatura aisladamente como variable explicativa (ya sea temperatura media u otra serie).

También se detectó que una serie construida a partir de la serie de temperatura máxima diaria como

$$iT_{Máx}^t(T_{Máx}^t) = \begin{cases} (T_{Máx}^t - 20^\circ C)^2, \wedge KDIA_INI_ITMAX < t < KDIA_FIN_ITMAX \\ (T_{Máx}^t - 22.5^\circ C)^2, \wedge KDIA_INI_ITMAX \geq t \cup t \geq KDIA_FIN_ITMAX \end{cases}$$

denominada $iT_{M\acute{a}x}$ explicaba significativamente las variables demanda diaria por tramo horario, siendo

$KDIA_INI_ITMAX = 102$; y $KDIA_FIN_ITMAX = 287$;

La variable t diferencia las épocas cálida y fría del año definiendo la época fría entre los días 102 y 287 del año y la época cálida en el resto de los días. Esto es, entre los días 11 de abril y 13 de octubre el valor que se resta a las temperaturas máximas es 20° y en el resto del año el valor que se resta es $22,5^\circ$.

Por otra parte, como se menciona en 3.1.4 Otras series de datos, se hicieron pruebas similares para otras variables aleatorias como velocidad del viento o radiación solar, pero no se detectó que existieran covarianzas que favorecieran a la precisión de la herramienta.

4.4.2. Configuración del Análisis Serial

En base a lo mencionado respecto a que el modelo debe tener en cuenta la estacionalidad anual además de la temperatura, inicialmente se analizó la identificación de modelos diferentes para cada estación del año. Los resultados no fueron acordes a lo esperado. Se continuó con las pruebas de identificación de modelos diferentes para las cuatro estaciones del año agregando cuatro modelos más para las “transiciones entre estaciones” del año (por ejemplo, los meses en que el verano “pasa” a otoño, etc.), totalizando ocho “estaciones”. En todos los casos anteriores, se usaron filtros constantes al momento de generar la CEGH.

Se utiliza la palabra filtro para nombrar el conjunto de funciones lineales hallada para para estimación del siguiente valor de cada variable en función de los datos de todas las variables en el paso anterior. Un filtro constante refiere a la utilización del mismo conjunto de funciones lineales para todo el período mientras que un filtro variable correspondería a la utilización de un conjunto de funciones lineales diferentes para cada paso de tiempo.

Finalmente, se concluyó que lo más apropiado era identificar un modelo diferente para cada día del año. Es decir, generar una CEGH con filtro variable estableciendo la cantidad de pasos en 365. Con estos parámetros, el módulo Análisis Serial identifica un filtro para cada día del año, de modo que, de manera continua, los coeficientes de la matriz de correlación entre las variables predictivas y explicativas (demanda diaria “de día hábil” sin tendencia por tramo horario y series de temperaturas, respectivamente) que se aplicarán, cambian a lo largo del año, captando la relación entre las demandas diarias “de día hábil” sin tendencia por tramo horario y las series de temperaturas máxima, mínima e $iT_{M\acute{a}x}$. Es decir, se identifica una matriz de correlación para cada punto del ciclo (que es anual), en este caso, una matriz por día. Esto se hace activando la casilla “Filtro Lineal Variable” de la Ilustración 9.

Cabe destacar en este punto que como se está pretendiendo hacer una herramienta que prediga la energía de día hábil por tramo horario para un tiempo t , conociendo el pasado de todas las variables a predecir hasta el tiempo $t-1$ (las demandas diarias de día hábil sin tendencia por tramo horario), y asumiendo que se conoce el valor de las variables explicativas (las tres series de temperaturas diarias) desde el pasado hasta el tiempo t inclusive (sea t mañana, es conocido pronóstico de temperaturas para el día de mañana), entonces las series a

ingresar en el módulo de Análisis Serial estarán “desfasadas”. Es decir, con la identificación de los modelos para cada día del año (filtro variable ya mencionado), se intenta captar las covarianzas que existen entre cada serie a predecir para un paso o tiempo t , con las series explicativas para un paso o tiempo $t+1$. En palabras, conociendo las temperaturas y demandas históricas, conociendo la demanda de hoy y la temperatura del día de mañana, se quiere predecir la demanda de mañana.

Como se mencionó en 4.3.1. Tratamiento de días feriados, los días feriados, tanto laborables como no laborables, excepto los que coincidieran con día domingo, fueron considerados datos inválidos en la generación de la CEGH con el fin de no interferir con la identificación de las relaciones de dependencia de las variables.

La configuración de parámetros utilizada en la interfaz del módulo Análisis Serial son los que se muestran en la Ilustración 9.

Ilustración 9. Parámetros del módulo Análisis Serial para identificación del modelo CEGH.

El modelo utilizado es $X[k+1]=AX[k] + BR[k]$, es decir, se predice el valor más probable del paso $k+1$ en base a todos los valores del paso anterior y utilizando orden del filtro igual a 1 ($NR=1$). Sin embargo, como ya se mencionó, en la construcción del conjunto de series se desplazó la temperatura un paso de tiempo hacia adelante, de manera que la temperatura de la fila k , en base a la cual se estimará el valor de la demanda del paso $k+1$ corresponde en realidad a la temperatura registrada en el paso $k+1$. De esta manera el modelo predice el valor más probable de demanda para el día k en función de la demanda registrada en el día $k-1$ y las temperaturas pronosticadas para el día k .

Se observa también que se seteó un overlapping de 15 días, sensiblemente menor a los 45 días que se utilizaban en la versión anterior. Dicho valor es suficiente para poblar los histogramas aunque genera factores resultantes con importantes oscilaciones de alta frecuencia en la evolución de los coeficientes del filtro. Esto se mejora realizando un

overlapping posterior, es decir, un suavizado de los factores resultantes para el cual se utilizó un buffer de 15 días.

El Traslapping se setó en cero pues al haber eliminado el ciclo semanal con la aplicación de los coeficientes antes mostrados para uniformizar las energías a las de día hábil promedio, no se detectó otro miniciclo relevante en la serie de datos.

En la Ilustración 10 se observan los factores de la combinación lineal de cada una de las demandas con las 3 variables vinculadas a las temperaturas (Tmax, Tmin e ITmax), para cada uno de los días del año. No se representaron en dicho gráfico los factores de la combinación lineal de las 4 demandas, que también son variables de entrada.

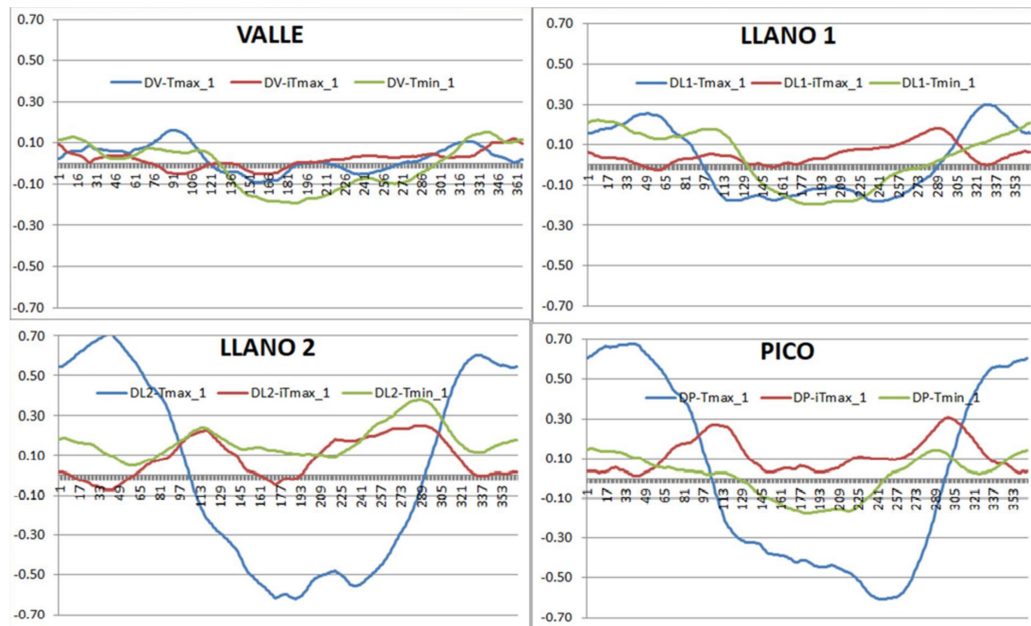


Ilustración 10 coeficientes del Filtro A para las demandas Pico, Valle, Llano1 y Llano2 respecto de las variables función de las temperaturas, graficados en función del día del año

Luego de identificado el modelo CEGH, se editaron las matrices de los deformadores para llevarlos a “por unidad de energía” ya que las series de entrada utilizadas corresponden a series de energía diaria acumulada en cada uno de los horarios, referenciada al año 2012 (al haberle quitado la tendencia creciente según se explicó anteriormente).

Para cada una de las matrices del deformador de las series (DP, DL1, DL2 y DV) se divide cada término de la matriz entre el promedio total de ésta, obteniéndose los deformadores por unidad de energía o potencia media de cada tramo con lo cual puede luego escalarse la demanda salida del modelo a los valores correspondientes al año de interés, multiplicando los valores por la energía o potencia media anual correspondiente de cada tramo horario.

En la Ilustración 11 se muestran los promedios de los deformadores ya pasados a p.u. para las variables demanda pico (DP), llano1 (DL1), llano2 (DL2) y valle (DV).

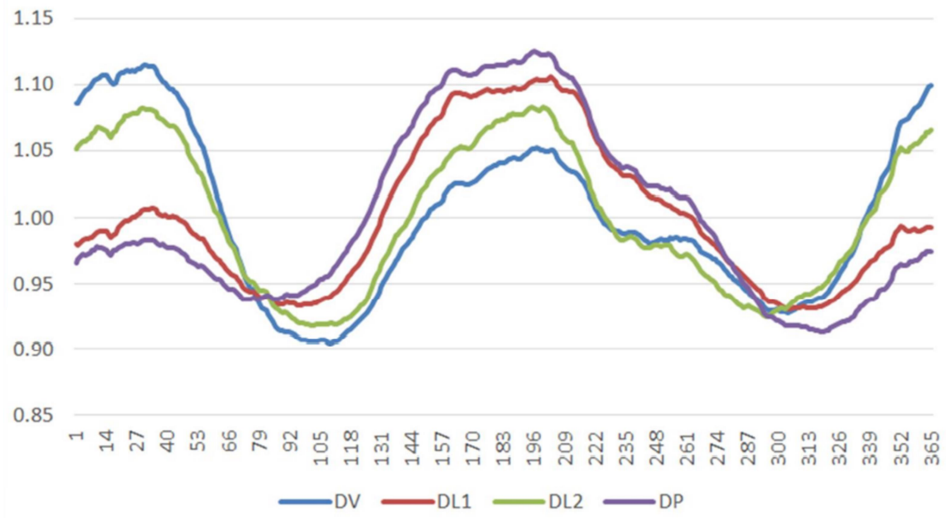


Ilustración 11 promedio de los deformadores del CEGH para las variables DP, DL1, DL2 y DV

4.5. Metodología para la identificación del comportamiento horario

El modelo CEGH elaborado genera valores diarios de demanda de cada uno de los tres horarios elegidos. Para representar el comportamiento horario de la demanda se distribuye la energía de acuerdo a la forma esperada de la curva horaria con el método que se describe aquí.

La forma de la curva horaria de la demanda es muy estable es decir, no presenta oscilaciones de baja frecuencia y se puede describir básicamente en tres tramos. Para el caso de los días hábiles, el primer tramo, desde las cero horas hasta las 5 o 6 de la mañana, presenta concavidad positiva alcanzando el mínimo entre las 3 y 4 de la mañana. El segundo tramo presenta concavidad negativa en su mayor parte y en algunas estaciones presenta un cambio de concavidad sobre las últimas horas (13 o 14 horas). El último tramo presenta también concavidad negativa mayor que la del tramo anterior, y alcanza el máximo entre las 20 y las 21 horas, dependiendo de la época del año.

Esta diferenciación de la curva en tres tramos es lo que sugirió realizar el modelado de la energía diaria en tres tramos ya que evidentemente, los tres tramos tienen un forzante principal diferente en cada caso, que responde a diferentes actividades humanas principales en cada horario y por lo tanto se esperaba que la relación con la temperatura pudiera ser diferenciada.

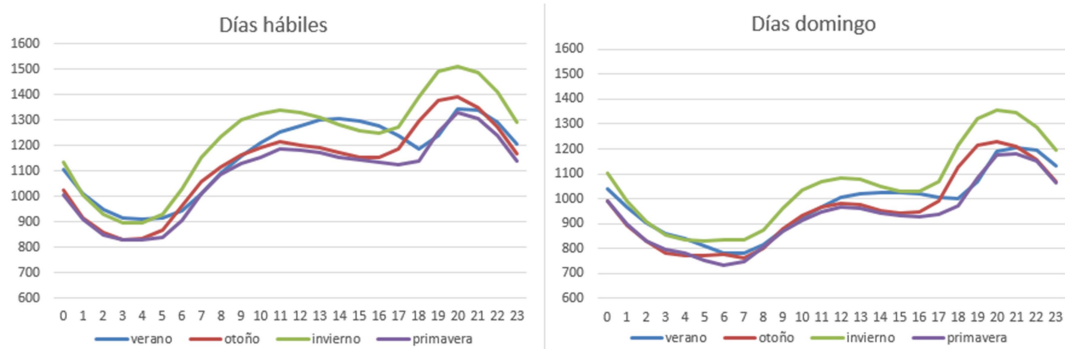


Ilustración 12 curvas horarias medias por estación para días hábiles y días domingo

Dada la baja variabilidad de la forma de la curva horaria a lo largo de las distintas estaciones del año, la representación elegida se basó en un procedimiento de agrupamiento de curvas históricas y selección e interpolación a partir de este índice de curvas históricas y las variables conocidas vinculadas que en este caso son, la temperatura media, el tipo de día (hábil, feriado, semiferiado) y el día de la semana.

El procedimiento se realizó mediante la aplicación Cluster_dem.lpi.

Previo al agrupamiento de curvas se realizó la misma división por tramos horarios que para el modelo CEGH, pico, valle, llano1 y llano2. Sin embargo, debido a la concavidad y continuidad presentada normalmente por la curva formada por los tramos llano1 y llano2, se optó por trabajar con ambos tramos de manera conjunta. En la Ilustración 13 se muestran las 10 curvas representativas de los agrupamientos de curvas para el tramo horario que abarca los tramos llano1 y llano2 del reparto de energías por tramo.

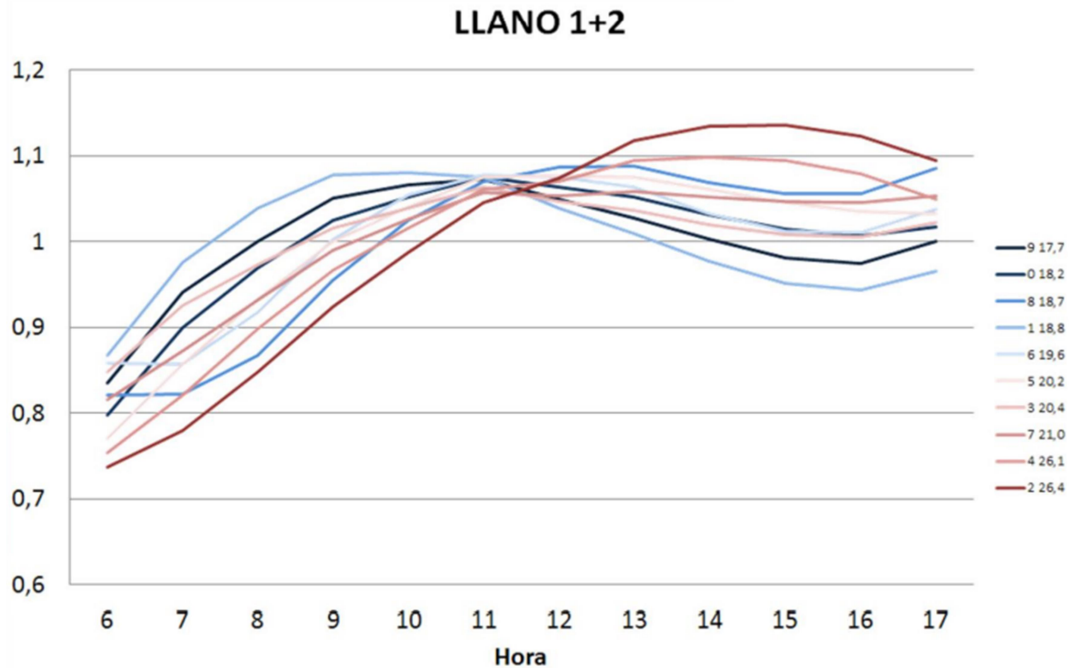


Ilustración 13 curvas representantes de los agrupamientos de curvas para el tramo horario que abarca los tramos llano1 y llano2 del reparto de energías por tramo.

Como se mencionó anteriormente, el procedimiento de selección de curvas para cada tramo horario se realiza en base al día del año y la temperatura media del día correspondiente, y también al tipo de día. El procedimiento se realiza tomando 20 días en torno al día del año en cuestión, filtrando de allí los días que no correspondan con el tipo de día del día en cuestión e interpolando linealmente entre las curvas resultantes según su temperatura media correspondiente y la temperatura media del día en cuestión.

Para la generación de pronósticos de demanda de los siguientes 14 días esta selección se realiza a partir de pronósticos de temperatura utilizando el programa "selectcurva_vates_pvl.m"

5. Configuración de la sala y el actor Demanda

Para la configuración del actor Demanda dentro de la Sala SimSEE se deben seguir los siguientes pasos:

1. Ingreso de la fuente CEGH y fuentes "producto" para multiplicar cada borne de la fuente por el valor medio anual del período de interés
2. Ingreso de valores de partida del día inicial de la corrida:
 - a. Demanda de cada tramo horario, en por unidad y llevado al equivalente de día hábil según los coeficientes de la Tabla 2
 - b. Temperaturas máxima y mínima del día siguiente al primer día de la corrida (debido al que el CEGH está construido con las temperaturas del día k+1)

3. Configuración de tres actores de demanda detallada, PICO, VALLE, LLANO1 y LLANO2, ingresando en cada actor el archivo correspondiente con los parámetros de forma de la curva en por unidad el horario correspondiente y ceros en las horas no correspondientes. La demanda total del sistema será la suma de las tres demandas.
4. Conexión de cada una de las tres demandas a la fuente de ruido correspondiente (sin sumar ruido), es decir a la fuente resultante de multiplicar el borne del CEGH del horario correspondiente por el valor de potencia media anual correspondiente.

En caso de que la corrida sea de largo plazo se utiliza un actor tipo “Demanda generada a partir de un año base y vector de energías anuales” teniendo en cuenta que el escalado del crecimiento anual debe hacerse en base a la energía ficticia que aparece en el formulario calculada como la suma de las energías del año base (que en este caso no será una energía ya que los valores del año base serán por unidad). Ver Ilustración 14.

Actor	Tipo de actor	Información adicional
D_Llano	Demanda generada a partir de un año base y vector de en llano.bin	[Edit] [Delete] [Export]
D_Pico	Demanda generada a partir de un año base y vector de en pico.bin	[Edit] [Delete] [Export]
D_Valle	Demanda generada a partir de un año base y vector de en valle.bin	[Edit] [Delete] [Export]

Escalón	Profundidad [p.u.]	Costo [USD/MWh]
1	0.02	175
2	0.05	600
3	0.075	2400
4	0.855	4000

Año	Demanda del año [GW]
2017	4.2
2018	4.3071466666
2019	4.4016

Ilustración 14 Ejemplo de ingreso de actor Demanda en corrida de largo plazo

6. Conclusiones

El nuevo modelo desarrollado y ya incorporado a las herramientas de optimización y despacho presenta un probado mejor comportamiento que su anterior versión, atacando los principales puntos débiles detectados durante la evaluación del primer modelo.

De todas maneras son muchas las oportunidades de mejora que se intentarán abordar en el futuro, algunas de las cuales ya estaban planteadas y no se alcanzaron a abordar durante

el proyecto y otras que han surgido a partir de la ampliación y profundización del estudio de datos realizado.

7. Referencias

[1] T. Hong y S. Fan, «Probabilistic electric load forecasting: a tutorial review,» *International Journal of Forecasting*, nº 32, pp. 914-938, 2016.

[2] R. Chaer y G. Casaravilla, «Sistemas de Simulación de Energía Eléctrica,» Proyecto PDT 47/12, <http://iie.fing.edu.uy/publicaciones/2008/CCA08a/CCA08a.pdf>, Montevideo, 2008.

[3] Dirección de Planificación, Oficina de Planeamiento y Presupuesto, Presente y futuro de las energías renovables en Uruguay, Montevideo: https://www.opp.gub.uy/sites/default/files/inline-files/12_%20Presente%20y%20futuro%20de%20las%20Energ%C3%ADas%20Renovables%20en%20Uruguay.pdf, 2019.

[4] R. Chaer, «Manual de usuario de Análisis Serial, IIE-Fing, UdelaR,» 2014. [En línea]. Available: https://simsee.org/simsee/manuales/SimSEE_mu5_Analisisserial.pdf.

[5] Lanzilotta, B. and Rodríguez, S. (2014), Modelos de predicción de energía eléctrica con datos horarios para Uruguay, Informe de investigación, Fondo Sectorial de Energía. ANII.

[6] R. Chaer, «Fundamentos de modelo CEGH de procesos estocásticos multivariados - SimSEE, IIE-Fing, UdelaR,» 2013. [En línea]. Available: <https://iie.fing.edu.uy/publicaciones/2011/Cha11>.

[7] Hong, T. (2010). Short term electric load forecasting. North Carolina State University

[8] Hong, T., & Shahidehpour, M. (2015). Load forecasting case study. National association of regulatory utility commissioners. Retrieved from http://naruc.org/Grants/Documents/Load%20Forecasting%20Case%20Study%20_%20UNC%20C.pdf.

[9] Xie, J., Hong, T., & Stroud, J. (2015). Long term retail energy forecasting with consideration of residential customer attrition. *IEEE Transactions on Smart Grid*, 6(5), 2245–2252.

[10] Tao Hong, Shu Fan (2015). Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting* 32 (2016) 914–938

[11] Electric Power Systems-Advanced Forecasting Techniques and Optimal Generation Scheduling - J. P. S. Catalão - CRC Press