



PRONÓSTICO DE DEMANDA DE ENERGÍA ELÉCTRICA USANDO PROCESOS GAUSSIANOS: UN ANÁLISIS COMPARATIVO

Short-term load demand forecasting using gaussian processes: a comparative analysis

Abstract

Load demand forecasting is an essential component for planning power systems, and it is an invaluable tool to grid operators or customers. Many methods have been proposed to provide reliable estimates of electric load demand, but few methods can address the problem of predicting energy demand from a probabilistic point of view. One of them is the Gaussian processes (GP) that considering an adequate covariance function are suitable tools to carry out this load forecasting task. In this article, we show how to use Gaussian processes to predict electrical energy demand. Additionally, we thoroughly test various covariance functions and provide a new one. The performance of the proposed methodology was tested on two real data sets, showing that GPs are

• Jorge Eliécer Loaiza Muñoz - Jorge.loaiza@utp.edu.co - Universidad Tecnológica de Pereira

AUTOR

competitive alternatives for short-term load demand forecasting compared to other state-of-the-art methods.

Keywords: Covariance functions, Gaussian process, short-term load forecasting.

Resumen

El pronóstico de consumo de energía eléctrica es un componente esencial para los sistemas de distribución y es

una herramienta invaluable para los comercializadores y clientes de este recurso. Muchos métodos han sido propuestos para proveer estimaciones confiables de la demanda de energía eléctrica, pero pocos abordan el problema desde un punto de vista probabilístico. Un método probabilístico que emplea una función de covarianza para abordar este tipo de tareas, son los procesos Gaussianos (GPs). En este artículo, mostramos cómo usar los GP para predecir la demanda de energía eléctrica. Adicionalmente, probamos varias funciones de covarianza y proponemos una. El rendimiento de la metodología propuesta fue probado en dos conjuntos de datos diferentes, mostrando que los GPs son alternativas competitivas para tareas de estimaciones en cortos horizontes al compararlo con los métodos propuestos en el estado del arte.

Palabras clave: Función de covarianza, proceso Gaussiano, pronóstico a corto plazo, demanda, energía eléctrica.

1. INTRODUCCIÓN

La predicción de la demanda de energía eléctrica es una componente clave para el planeamiento y operación de los sistemas eléctricos de potencia. Es también empleada para el análisis de confiabilidad y seguridad de los sistemas de potencia en mercados de energía competitivos (Fan G., Peng L., & Hong W., 2018). De lo anterior, es necesario que esta componente deba contemplar la variabilidad intrínseca en la operación de los sistemas de potencia, como por ejemplo el comportamiento del clima, la anticipación para comprar o vender

energía eléctrica, o la alta penetración de generación distribuida. Todos estos factores introducen incertidumbre en la predicción de la demanda de energía y hacen que esta tarea sea un reto actual de investigación (Fan G., Peng L., & Hong W., 2018).

Con el fin de abordar lo anterior, muchos métodos han sido propuestos en la literatura (Fan G., Peng L., & Hong W., 2018), (Van Der Meer D. et al., 2018), (Zhang J. et al., 2018), (Tarsitano A. & Amerise I., 2017). Estos métodos pueden ser clasificados en dos grupos: los métodos estadísticos y métodos basados en inteligencia artificial (Fan G., Peng L., & Hong W., 2018). Por un lado, los enfoques estadísticos usan datos de la demanda de energía para encontrar relaciones lineales entre variables exógenas o periodos de tiempo y la demanda de energía (Fan G., Peng L., & Hong W., 2018), (Van Der Meer D. et al., 2018). Entre las técnicas más empleadas dentro de este primer grupo se pueden encontrar los modelos autorregresivos integrados de media móvil (ARIMA) (Zhang J. et al., 2018), (Tarsitano A. & Amerise I., 2017), modelos lineales de regresión (Dudek G., 2016), (Zamo M. et al., 2014), modelos de suavizado exponencial (Oliveira E. & Oliveira F., 2018), filtros de Kalman (Takeda H., Tamura Y. & Sato S., 2016), entre otros. Estos métodos presentan facilidad en su implementación, sin embargo, el principal inconveniente es el uso de simplificaciones o relaciones lineales para el modelado de los datos de la demanda de energía (Fan G., Peng L., & Hong W., 2018). Por otro lado, los métodos basados en inteligencia artificial usan grandes cantidades de datos para realizar la predicción de la demanda

de energía. Los más usados son: sistemas expertos (I. da Silva & L. de Andrade, 2016), (Duan Q., Liu J. & Zhao D., 2017), máquinas de vectores de soporte (Chen Y. et al., 2017), (Fan G. et al., 2016), y redes neuronales artificiales (Jimenez J., Donado K. & Quintero C., 2017), (Singh P. Dwivedi P., 2018). Una de las dificultades de estas metodologías es depender del uso del conocimiento de un experto para la sintonización o la especificación de parámetros para realizar una correcta predicción (Van Der Meer D. et al., 2018). Es el caso de las redes neuronales donde es necesario que un experto defina la cantidad de capas y neuronas por capa para el buen funcionamiento de este modelo, es decir, el número óptimo de capas y neuronas en una red neuronal se determinan de manera empírica o basada en experiencias previas (X. Cao et al., 2015).

En la actualidad muchos trabajos han empleado enfoques profundos de aprendizaje, dentro de los cuales se pueden destacar las redes neuronales profundas (G. Sideratos, A. Ikononopoulos & N. Hatziaargyriou, 2020). (Deep Neural Network-DNN) las cuales son redes que se caracterizan por la utilización de redes neuronales compuestas de múltiples capas de extracción de características y transformación de las entradas. Otro enfoque de aprendizaje profundo son las redes neuronales convolucionales (Q. Huang, J. Li & M. Zhu, 2020). Son redes neuronales más potentes que generan mapas de características permitiendo una mejor toma de decisiones a la hora de realizar tareas de clasificación o regresión. Así mismo, se han usado redes neuronales recurrentes y redes neuronales de larga memoria a corto plazo

en la predicción de la demanda de energía eléctrica (W. Kong et al., 2019). La falencia más representativa de estas redes profundas es la cantidad de datos requeridos para ser entrenadas, los cuales son del orden de millones. Esto sumado a su compleja arquitectura, las hace poco atractivas por su alto costo computacional dado que requieren de hardware especial como puede ser una Unidad de Procesamiento Gráfico o GPU en inglés, para el entrenamiento (Q. Huang, J. Li & M. Zhu, 2020), (Z.-Y. Dou et al., 2020).

Como un esfuerzo para solucionar los inconvenientes anteriormente mencionados, los procesos Gaussianos (GPs) para regresión han sido empleados como alternativa para modelar la demanda de energía eléctrica. Por ejemplo, en (Van Der Meer D. et al., 2018) han realizado una predicción de la radiación solar y demanda de energía, mostrando mejor rendimiento ante la regresión basada en máquinas de vectores de soporte. Por otro lado, en (Sheng H et al., 2017) utilizaron un GP incorporando detección de datos espurios en la predicción de la demanda de energía. En (Lauret P., David M. & Calogine D., 2012), han comparado redes neuronales, redes neuronales Bayesianas y los GPs para la predicción de la demanda de energía. A partir de este estudio, los autores observan que los GPs superan en rendimiento a las redes neuronales. Sin embargo, estos estudios no hacen énfasis en el análisis de la función de covarianza, la cual es un componente clave para la utilización de los GPs en tareas de clasificación o regresión (Shepero, M. et al., 2018).

Este artículo muestra como usar los GPs para la predicción de la demanda de energía

eléctrica usando diferentes funciones de covarianza base, tales como: exponencial cuadrática (SE), periódica (PER), cuadrática racional (RQ), y se propone una combinación de dos de ellas. Con el fin de realizar un análisis comparativo se aplicarán otros métodos de predicción del estado del arte usando dos bases de datos reales de demanda de energía.

2. MATERIALES Y MÉTODOS

Las simulaciones realizadas en este trabajo de investigación fueron realizadas en MATLAB empleando el Toolbox GPML versión 4.2.

2.1 Bases de datos

Para este análisis comparativo se han empleado datos de consumo de demanda de energía eléctrica en Colombia y en Bélgica. Los datos de la demanda de energía en Colombia fueron extraídos del operador del sistema interconectado nacional a través del portal de XM (XM, 2020), y como datos alternos para este análisis, se usaron los datos de un operador del sistema de transmisión en Europa del grupo Elia (E. Group, 2020).

2.2 Procesos Gaussianos para el modelado probabilístico de la demanda de energía

Dentro de los métodos propuestos para modelar la demanda de energía, los GPs pueden ser vistos como una herramienta poderosa que proporciona un modelado con diferentes virtudes (Shepero, M. et al., 2018). Con el uso de los GPs se puede obtener un modelo probabilístico de la demanda

de energía y así mismo se puede evitar la dependencia de una definición específica de la estructura del modelo, lo que si pasa con las redes neuronales artificiales (Van Der Meer D. et al., 2018). Una de las maneras de lograr este modelado es considerando que la demanda de energía está dada por la ecuación (1).

$$y_t = f(x_t) + \varepsilon_t \quad (1)$$

donde y_t es el valor de la demanda de energía en el instante t ; x_t es el vector de entrada, el cual está formado por eventos pasados de la demanda de energía. ε_t es la incertidumbre presente en los datos del consumo energético, los cuales se asumirán que se distribuyen normalmente (Kevin P., 2012), es decir, $\varepsilon_t \sim N(0, \sigma^2)$ donde σ^2 es la varianza presente en las observaciones, f es la función donde está consignada la representación no lineal de la demanda de energía. Con lo anterior definido, un proceso Gaussiano puede ser especificado por una función media $m(x)$ como se detalla en la expresión (2), y una función de covarianza o kernel $k(x, x')$ (C. Rasmussen & C. Williams, 2006) como se referencia en la expresión (3), es decir,

$$m(x) = E[f(x)] \quad (2)$$

$$k(x, x') = E[(f(x) - m(x))(f(x') - m(x')))] \quad (3)$$

donde $E[f(x)]$ corresponde al valor esperado de $f(x)$; x y x' representan los datos de entrenamiento y de prueba respectivamente. A partir de lo anterior, un proceso gaussiano se puede representar como en (4):

$$f(x) \sim \mathcal{GP}(m(x), k(x, x')) \quad (4)$$

Una componente crucial en un proceso Gaussiano es la función de covarianza (Bagas Abisena & Swastanto, 2016). Esta función es la encargada de codificar hipótesis iniciales (por ejemplo, variaciones suaves o periodicidad) del problema de modelado. Además, la selección de una función de covarianza está sujeta a que sea una matriz de covarianza semi-definida positiva (C. Rasmussen & C. Williams, 2006). Las funciones de covarianza más usadas son la exponencial cuadrática, cuadrática racional y la periódica (Shepero, M. et al., 2018).

2.3 Funciones de covarianza base

2.3.1 Función de covarianza exponencial cuadrática

Es una función de covarianza entre dos variables aleatorias y está especificada por la ecuación (5).

$$k(x, x') = \sigma^2 \exp\left(-\frac{\|x - x'\|^2}{2\lambda^2}\right) \quad (5)$$

donde σ^2 y λ^2 son parámetros de varianza y escala de la función de covarianza. Esta función es ampliamente usada en los procesos gaussianos por su capacidad de lizado o smoothing (Wilson, A., & Adams, R. (2013, February). En este estudio, esta función de covarianza se denotará como **SE**.

2.3.2 Función de covarianza cuadrática racional

Es una función de covarianza que permite modelar datos variantes en múltiples escalas y puede ser vista como una función SE con diferentes escalas de longitud (C. Rasmussen & C. Williams, 2006). Esta función está dada

por la expresión (6).

$$k(x, x') = \exp\left(1 + \frac{(x - x')^2}{2\alpha\lambda^2}\right)^{-\alpha} \quad (6)$$

donde α y λ son parámetros de escala. En este estudio, esta función se nombrará como **RQ**.

2.3.3 Función de covarianza periódica

Otra función de covarianza muy común es la periódica (C. Rasmussen & C. Williams, 2006), la cual está definida en (7).

$$k(x, x') = \sigma^2 \exp\left(-\frac{\text{sen}^2(\pi(x - x')/\rho)}{2\lambda^2}\right) \quad (7)$$

Este kernel periódico ofrece una estructura repetitiva sobre la función, la cual es controlada por el parámetro del periodo ρ . Un valor de ρ grande ocasiona oscilaciones bajas, mientras que un valor de ρ pequeño conlleva a oscilaciones altas (Bagas Abisena & Swastanto, 2016). La longitud de escala λ^2 y la varianza σ^2 tienen el mismo efecto que se mencionó para el kernel **SE**. De igual forma, esta función base se denotará como **PER**.

2.4 Composición de funciones de covarianza

Según Rasmussen y Williams en (C. Rasmussen & C. Williams, 2006), es posible proponer o desarrollar nuevas funciones de covarianza a través de la adición o multiplicación de funciones de covarianza base como se muestra en (8) y (9). Es decir, se puede tener:

$$k(x, x') = k_1(x, x') + k_2(x, x') \quad (8)$$

O también,

$$k(x, x') = k_1(x, x') \cdot k_2(x, x') \quad (9)$$

Es primordial que la nueva función de covarianza resultante sea semi-definida positiva. La idea al realizar una composición de una función de covarianza es que cada función base sea responsable de modelar una cierta estructura o patrón en las series de tiempo de acuerdo con sus características (Bagas Abisena & Swastanto, 2016). A partir de lo anterior y para llevar a cabo el pronóstico de la demanda de energía eléctrica, se propone una nueva función de covarianza la cual está dada por (10),

$$k(x, x') = \sigma^2 \exp\left(\frac{x - x'}{2\lambda^2}\right) +$$

$$k(x, x') = \sigma^2 \exp\left(-\frac{\text{Sen}^2(\pi(x-x')/\rho)}{2\lambda^2}\right) \quad (10)$$

Esta iniciativa surge del análisis de los patrones de periodicidad que se presentan en las curvas de demanda de energía. Se decide emplear el operador de adición pues según (Pasan Karunaratne, s.f.), cuando se emplea este operador cada función involucrada trabaja de manera independiente con sus propias características. Por otro lado, en la literatura se han analizado diversas composiciones de funciones de covarianza (Van Der Meer D. et al., 2018), pero no se ha tenido en cuenta la composición anteriormente expuesta. Esta composición de funciones de covarianza se denotará como **SE+PER** y se expresa en (10).

2.5 Hiperparámetros

Además de ser una componente clave dentro de los GPs, la función de covarianza depende de varios parámetros (varianzas, periodos y escalas) que codifican los patrones de las observaciones que se están analizando (C. Rasmussen & C. Williams, 2006), (K. Chalupka, 2011). Estos parámetros

deben ser estimados de manera óptima para garantizar un correcto modelado de la demanda de energía. Cuando se usan los procesos Gaussianos para regresión o para clasificación, los hiperparámetros se obtienen mediante aprendizaje de los datos de entrenamiento. Existen diferentes técnicas para la estimación de estos valores. Una de ellas es el estimador de probabilidad máxima o máxima verosimilitud conocida como MLE (Maximum Likelihood Estimation), el cual fue empleado en este estudio para el cálculo de los hiperparámetros.

2.6 Implementación

Con el fin de realizar un pronóstico en un horizonte corto, es decir, de 24 horas para los días lunes, martes, miércoles, jueves y viernes, se toman como datos de entrenamiento los cuatro días similares inmediatamente anteriores al día que se desea predecir. Por ejemplo, para predecir el primer miércoles del mes de febrero, se toma como datos de entrenamiento los cuatro miércoles del mes anterior. Se avanza un día hacia adelante y se repite el proceso. En total se toman 48 días de los cuales 34 son para entrenamiento y 14 para prueba. La matriz de los datos de entrenamiento $D = (X, y)$ para la predicción horaria queda definida así:

$$X = \begin{bmatrix} D_1 & D_2 & D_3 & D_4 \\ D_2 & D_3 & D_4 & D_5 \\ D_3 & D_4 & D_5 & D_6 \\ \vdots & \vdots & \vdots & \vdots \\ D_{n-3} & D_{n-2} & D_{n-1} & D_n \end{bmatrix} \quad y = \begin{bmatrix} D_5 \\ D_6 \\ D_7 \\ \vdots \\ D_{n+1} \end{bmatrix}$$

donde n es la dimensión de los datos de entrenamiento, lo cual para un horizonte horario $n = 816$ que equivale a observaciones de 24 horas por 34 días. Para el conjunto de datos de prueba $D_* = (X_*, y_*)$, se tiene:

$$X_* = \begin{bmatrix} D_{1*} & D_{2*} & D_{3*} & D_{4*} \\ D_{2*} & D_{3*} & D_{4*} & D_{5*} \\ D_{3*} & D_{4*} & D_{5*} & D_{6*} \\ \vdots & \vdots & \vdots & \vdots \\ D_{*m-3} & D_{*m-2} & D_{*m-1} & D_{*m} \end{bmatrix}$$

$$y_* = \begin{bmatrix} D_{5*} \\ D_{6*} \\ D_{7*} \\ \vdots \\ D_{*m+1} \end{bmatrix}$$

donde $m = 336$ equivale a 24 horas por 14 días. Para realizar el pronóstico diario (por semana) se sigue un proceso similar al realizado en la predicción horaria. Es decir, se inicia con cuatro semanas de entrada y la siguiente como objetivo a predecir. Luego, se avanza una semana y se repite el proceso. Para este horizonte se toman 47 semanas en total, de las cuales 38 son para entrenamiento y 9 para prueba.

2.7 Métodos alternativos para la predicción

Con el fin de llevar a cabo el análisis comparativo, se emplearon otros métodos de predicción. Específicamente se usaron máquinas de vectores de soporte para regresión (SVR) (Bing Donga et al., 2015), modelos ARIMA (Zhang J. et al., 2018), redes neuronales artificiales (NN) (Bing Donga et al., 2015) y redes neuronales profundas (DNN) (G. Din & A. Marnerides, 2017). Las SVR son ampliamente usadas en el pronóstico de la demanda de energía (Bing Donga et al., 2015) y se basan en la construcción de un hiperplano óptimo en forma de superficie de decisión. Los vectores de soporte hacen referencia a un pequeño subconjunto de las observaciones de entrenamiento que se utilizan como soporte para la ubicación óptima de la superficie de decisión (Kevin P., 2012). Por otro lado, los

modelos ARIMA son métodos que se basan en la comción lineabinal de datos y errores pasados (Zhang J. et al., 2018), (Tarsitano A. & Amerise I., 2017). Por último, las NN y DNN son arreglos de unidades o nodos (neuronas) que se interconectan entre sí para predecir un nuevo dato de demanda de energía (G. Din & A. Marnerides, 2017).

2.8 Métricas de rendimiento

El desempeño de los modelos implementados en este trabajo es medido mediante el cálculo de dos métricas de error asociados a los resultados obtenidos. Estas dos medidas son: el error cuadrático medio estandarizado (SMSE) (K. Chalupka, 2011), el cual está dado por (11).

$$SMSE = \frac{1}{\sigma_{test}^2 + \mu_{test}^2} \left(y_{*i} - \bar{f}(x_{*i}) \right)^2 \quad (11)$$

donde x_{*i} y y_{*i} son los valores de un dato de entrada y salida del conjunto de datos de prueba, σ_{test}^2 y μ_{test}^2 son la varianza y la media del conjunto de datos de prueba respectivamente.

La otra métrica de rendimiento es la pérdida logarítmica estandarizada de la media (MSLL) (K. Chalupka, 2011), la cual está expresada en (12).

$$MSLL = \frac{1}{2} \log(\sigma_{*i}^2) + \frac{(y_{*i} - \bar{f}(x_{*i}))^2}{\sigma_{*i}^2} - \frac{(\sigma_{test}^2 + \mu_{test}^2)}{2} \quad (12)$$

donde $\bar{f}(x_{*i})$ y σ_{*i}^2 son la media y varianza predictiva en el punto de prueba x_{*i} , respectivamente.

3. RESULTADOS Y DISCUSIÓN

Esta sección tiene como objeto mostrar un análisis comparativo entre las técnicas que se han usado para la predicción en un horizonte

corto de la demanda de energía frente a los procesos Gaussianos. Inicialmente se probaron las técnicas anteriormente descritas usando datos de demanda de energía en Colombia. Posteriormente se emplea el GP para realizar la predicción de demanda de energía usando datos proporcionados por un operador del sistema de transmisión en Bélgica.

3.1 Predicción horaria en Colombia

Teniendo en cuenta lo anterior se presentan los resultados de las estimaciones por horas de lunes a viernes del año 2015 empleando la función de covarianza (kernel) SE+PER usando datos extraídos de la plataforma de XM. Con el fin de resaltar el desempeño del modelo, se decide mostrar los días hábiles de la semana como se observa en la tabla 1.

Tabla 1. Resultados de la predicción horaria. Fuente: Autores

Día	SMSE	MSLL
Lunes	0.0008 ± 0.00040	-3.2506 ± 0.1631
Martes	0.0003 ± 0.00011	-3.3259 ± 0.0928
Miércoles	0.0002 ± 0.00018	-3.3132 ± 0.0947
Jueves	0.0002 ± 0.00017	-3.3517 ± 0.0764
Viernes	0.0004 ± 0.00008	-3.3204 ± 0.0663

La anterior tabla muestra que el método propuesto presenta un menor SMSE y un alto MSLL los días lunes. Esto se debe a la presencia de días festivos los cuales se manifiestan más en los días en cuestión. Estos días feriados presentan un comportamiento irregular respecto a aquellos días normales hábiles, y son una razón para obtener unas pobres estimaciones. En contraste, los días miércoles y jueves son los que

presentan mejores índices de error. Este buen comportamiento puede ser explicado por la nula o poca presencia de días festivos en estos días. El experimento se repite con algunas funciones de covarianza para constatar el desempeño de la función compuesta. Dichas funciones base son: SE, PER y RQ. La tabla II muestra los SMSE y MSLL usando estas funciones de covarianza base. Los resultados con la función base SE son un poco mejores o iguales a los obtenidos con la función compuesta. Esto nos indica que la función PER en la composición no está realizando un trabajo significativo. Es posible, entonces, que el comportamiento periódico de los datos de entrenamiento presente un patrón cíclico que la función PER no puede manejar adecuadamente, razón por la cual dicha función arroja pobres estimaciones como se evidencia en la tabla 2.

Tabla 2. Resultados de predicciones horarias por días usando diferentes funciones base. Fuente: Autores

Día	SE	PER	RQ
Lunes	0.00082 ± 0.00027	0.02218 ± 0.0014	0.00070 ± 0.00018
Martes	0.00028 ± 8.3e-05	0.02106 ± 0.0023	0.00028 ± 8.3e-05
Miércoles	0.00020 ± 0.00018	0.01648 ± 0.0016	0.00022 ± 4.4E-05
Jueves	0.00024 ± 0.00016	0.01740 ± 0.0010	0.00036 ± 0.00015
Viernes	0.00028 ± 0.00010	0.01930 ± 0.0023	0.00042 ± 8.3e-05

A partir de los datos reportados en la tabla 1 se decidió presentar gráficamente la predicción de la demanda para el día jueves 29 de octubre de 2015. Como se mencionó anteriormente, las buenas estimaciones para este día hábil, pueden ser explicadas por el comportamiento regular del consumo de energía que se presenta normalmente en los días de este tipo.

Cabe resaltar que para el entrenamiento de los días jueves, se excluyó las observaciones del día primero de enero, pues este es un día que presenta demasiada irregularidad en el consumo de energía por tratarse del primer día del año. La Fig. 1 nos ofrece una perspectiva gráfica de la predicción para el día mencionado con la variabilidad de las estimaciones representada en gris.

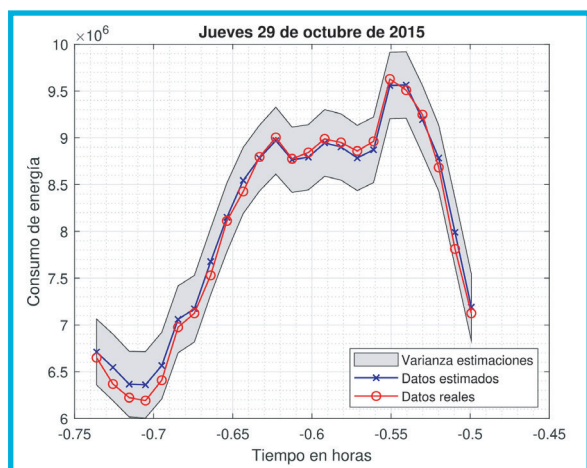


Fig. 1. Predicción para el jueves 29 de octubre de 2015. Fuente: Autores

3.2 Predicción diaria en Colombia

Una vez presentado el comportamiento del método propuesto para un horizonte horario, se procede a observar el rendimiento de este ante diferentes estimaciones diarias (por semana) para el año 2015 usando la función de covarianza propuesta. La Fig. 2 permite apreciar gráficamente la estimación para la semana 44 del año 2015, la cual va desde el 26 de octubre hasta el 1 de noviembre, y cuyo SMSE fue de 0,0001.

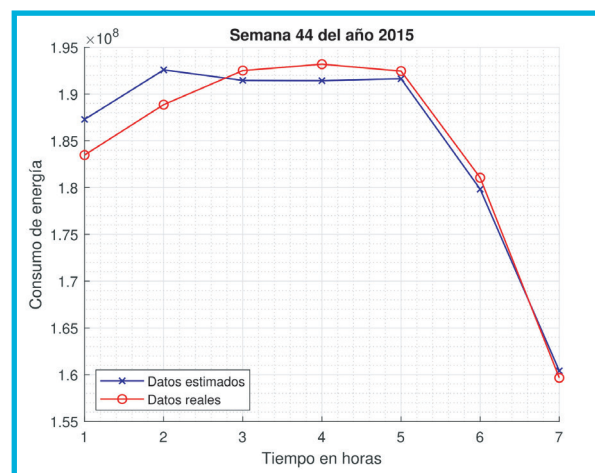


Fig. 2. Estimación de la semana 44 del año 2015. Fuente: Autores

Con el fin de contrastar la función de covarianza compuesta, el experimento se repite usando las funciones base SE y RQ. Los resultados SMSE para la semana 44 del año 2015 con estas funciones base son 0,0002 y 0,0002 respectivamente. La Fig. 3 y 4 permiten visualizar el resultado para SE y RQ, respectivamente. No se muestran gráficamente los resultados usando la función de covarianza PER dado que presente estimaciones pobres en este horizonte de predicción.

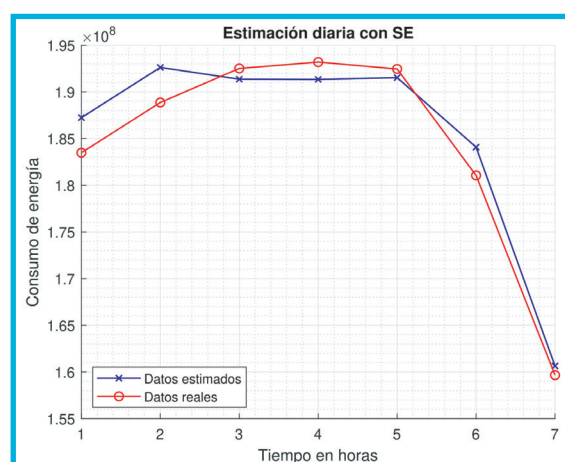


Fig. 3. Estimación diaria (por semana) con la función de covarianza base SE. Fuente: Autores

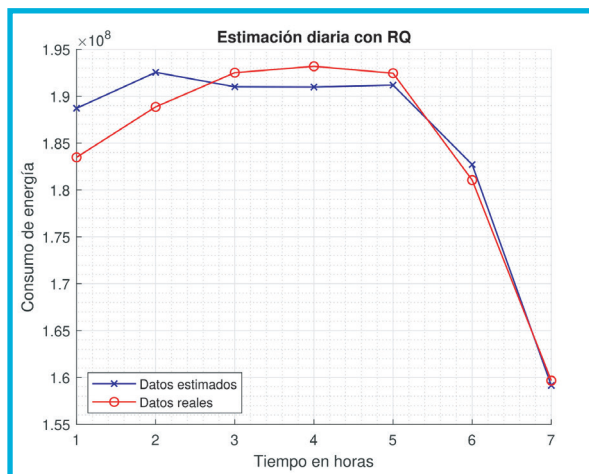


Fig. 4. Estimación diaria (por semana) con la función de covarianza base RQ. Fuente: Autores

De lo anterior, se decide repetir el experimento para los años 2013 y 2014 usando la función de covarianza propuesta. En la tabla 3 se presenta las métricas de rendimiento para estos años. Dicha tabla informa que la tendencia del rendimiento se mantiene estable, lo que indica que el modelo se comporta de manera similar ante diferentes periodos de tiempo.

Tabla 3: Resultados de la predicción diaria usando SE+PER entre los años 2013 y 2015

Año	SMSE	MSLL
2015	0.002372±0.002634	-1.6883 ± 8.8903
2014	0.002946±0.003662	-1.6206 ± 1.0773
2013	0.002761±0.002992	-1.4923 ± 8.2823

3.3 Comparación usando métodos alternativos de la literatura

Esta sección presenta los resultados obtenidos al comparar el GP con otros métodos de predicción como se mencionaron anteriormente. Específicamente, se usó las máquinas de vectores soporte para

regresión (SVR) con una cantidad de vectores de soporte en el entrenamiento de 86 y un kernel Gaussiano. También se usó un modelo ARIMA con $p = 26$, $d = 1$, $q = 1$ y sus parámetros fueron estimados por máxima verosimilitud. Para la NN se usó una red con 10 neuronas en la capa oculta. Finalmente, se implementó una DNN con tres capas ocultas con 48, 20 y 50 neuronas por capa, respectivamente. Todas estas metodologías fueron implementadas en Matlab. La Fig. 5 muestra el comportamiento de cada uno de los métodos para predecir la demanda para el caso horario. Se puede notar que todas las metodologías siguen el patrón de la demanda de energía.

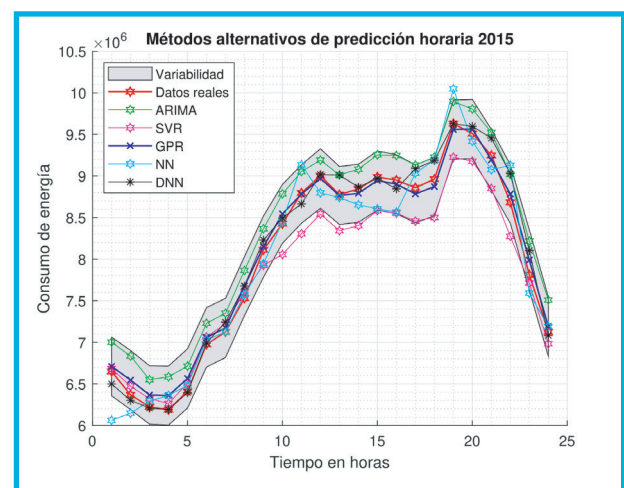


Fig. 5. Métodos alternativos predicción horaria vs GP. Fuente: Autores

Tabla 4: Métricas de rendimiento usando métodos alternativos para un horizonte horario.

	SMSE	MSLL
ARIMA	0.0065 ± 0.0025	-2.7340 ± 0.2177
SVR	0.0124 ± 0.0032	-2.1247 ± 0.5760
GP	0.0003 ± 0.0002	-2.9908 ± 0.2730
NN	0.0042 ± 0.0027	-2.2816 ± 0.3526
DNN	0.0017 ± 0.0011	-3.2816 ± 0.4098

De la Fig. 5 se observa que las SVR no logra obtener predicciones acertadas en horas de mayor demanda. Similarmente pasa con los modelos ARIMA. Por un lado, se nota que las NN entrega predicciones cercanas a los datos reales de la demanda de energía. Por otro lado, se puede notar que las DNN y el enfoque propuesto obtiene predicciones similares a los datos de la demanda. En la tabla 4 están consignados los valores de SMSE y MSLL para cada una de las técnicas. De esta tabla, se puede apreciar que las SVR y los modelos ARIMA obtienen los menores índices de rendimiento. Los enfoques de las NN, DNN y los GP presentan índices de error comparables, donde los dos últimos métodos son los que presentan los mejores registros para el SMSE y MSLL. De lo anterior cabe mencionar que sin duda las redes neuronales son elementos valiosos para realizar tareas de predicción en series temporales. Sin embargo, también se debe señalar que los GPs al basarse en modelos probabilísticos (distribuciones de probabilidad) entregan una región o intervalo de confianza en sus predicciones. Esta característica sería un valor agregado para la técnica propuesta en aplicaciones como la operación de los mercados de energía en sistemas de potencia. Otra ventaja de los GPs frente a los otros métodos de predicción es que al usar un enfoque probabilístico se puede incorporar información extra o particular en el modelo probabilístico (R. Adams, G. Dahl, & I. Murray, 2010).

3.4 Predicción de demanda usando datos alternos

Luego de mostrar las predicciones horarias y diarias con datos de demanda de energía de Colombia, se procede a verificar el comportamiento del GP usando datos de

demanda de energía proporcionada por un operador del sistema de transmisión en Bélgica (ver [?]). Sólo se presentarán las predicciones con la función de covarianza compuesta SE+PER. Estas predicciones pueden observarse en las Fig. 6 y 7, en horizontes horarios y semanales, respectivamente.

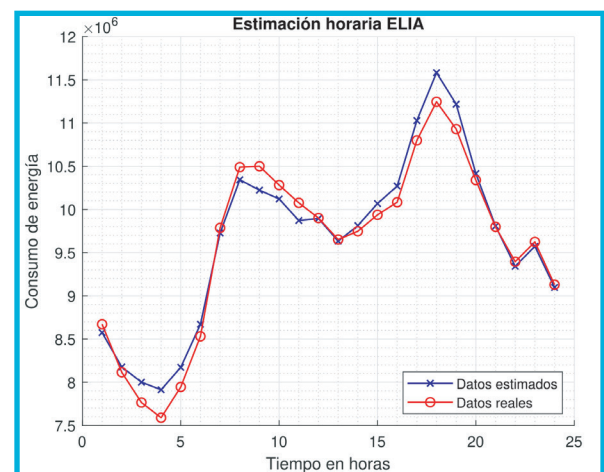


Fig. 6. Estimaciones para el día jueves de 2015 con la función SE+PER. Fuente: Autores

En la Fig. 6 se muestra el comportamiento del método propuesto en un horizonte horario y cuyo SMSE es de 0,0003. De manera general los SMSE y MSLL son $0,008375 \pm 0,01114$ y $-4,4304 \pm 6,3446$, respectivamente. Por otro lado, la Fig.7 permite apreciar el resultado de las estimaciones siguiendo la misma metodología para el horizonte semanal. La mejor estimación se presenta para la semana número 2 de los datos de prueba y presenta un error SMSE de 0,0002. Los resultados de manera general son $0,005496 \pm 0,007369$ para SMSE y $-2,4186 \pm 0,2611$ para MSLL.

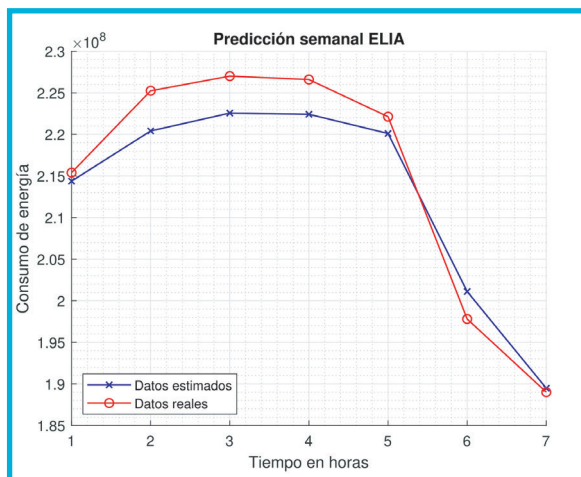


Fig. 7. Estimación semanal de los datos alternos. Fuente: Autores.

4. CONCLUSIONES

Los resultados obtenidos en esta propuesta de investigación permiten determinar que los GP con la función de covarianza propuesta, son otra posible herramienta para la predicción de la demanda de energía eléctrica en horizontes cortos. El buen desempeño se debe a que son flexibles y tienen en cuenta la variabilidad del comportamiento de la demanda de energía. Debido a que sus parámetros (hyperparámetros) son optimizados usando estimación por máxima verosimilitud, posee una ventaja importante frente al aprendizaje basado en redes neuronales. Por otro lado, el éxito de un proceso Gaussiano radica en la adecuada selección de la función de covarianza, pues esta debe ir acorde al comportamiento que presenten los datos considerando las características propias de cada una de ellas.

Como trabajos futuros, sería interesante analizar más composiciones de funciones de covarianza, proponer modelos es-

calables con procesos Gaussianos debido a su régimen específico de cantidad de datos y analizar el comportamiento de las predicciones usando GP multitareas.

Referencias

- Bagas Abisena, Swastanto. (2016). "Gaussian Process Regression for Long-Term Time Series Forecasting". Delft University of Technology.
- Bing Donga, Zhao Xuan Lia, S.M. Mahbobur Rahmana, Rolando Vega. (2015). A hybrid model approach for forecasting future residential electricity consumption. 17 June 2015. Energy and Buildings 117 (2016) 341–351.
- C. E. Rasmussen and C. Williams (2006). Gaussian Processes for Machine Learning. The MIT Press, 2006.
- Chen Y., Xu P., Chu Y., Li W., Ni L. (2017). Short-term electrical load forecasting using the support vector regression. ScienceDirect, 195, Pages 659-670. <https://doi.org/10.1016/j.apenergy.2017.03.034>
- De Oliveira E., Oliveira F. (2018). Forecasting mid-long-term electricity energy consumption through bagging ARIMA and exponential smoothing methods. ScienceDirect, 144, Pages 776-788. <https://doi.org/10.1016/j.energy.2017.12.049>
- Duan Q., Liu J., Zhao D. (2017). Shortterm electric load forecasting using an automated system of model choice. ScienceDirect, 91, Pages 92-100. <https://doi.org/10.1016/j.ijepes.2017.03.006>
- Dudek G. (2016). Pattern based local linear regression models for short term load forecasting. ScienceDirect, 130, Pages 139147. <https://doi.org/10.1016/j.epr.2015.09.001>
- E. Group (Junio de 2020). Transparency on grid data, <https://www.elia.be/en/grid-data>
- Fan G., Peng L., Hong W., (2018). Short term load forecasting based on phase space reconstruction algorithm and bisquare kernel regression model. ScienceDirect, 224,13-33. <https://doi.org/10.1016/j.apenergy.2018.04.075>
- Fan G., Peng L., Hong W., Sun F. (2016). Electric load forecasting by the SVR model with differential empirical mode decomposition and auto regression. ScienceDirect, 173, Pages 958-970. <https://doi.org/10.1016/j.neucom.2015.08.051>
- G. M. U. Din and A. K. Marnerides. (2017). "Short term power load forecasting using Deep Neural Networks," 2017 International Conference on Computing, Networking and Communications (ICNC), Santa Clara, CA, 2017, pp. 594-598.
- G. Sideratos, A. Ikonomopoulos, and N. D. Hatziargyriou. (2020). "A novel fuzzy-based ensemble model for load forecasting using hybrid deep neural networks," Electric Power Systems Research, vol. 178, p. 106025, 2020. <http://www.sciencedirect.com/science/article/pii/S037877961930344X>
- Geng J. Huang M., Li M., Hong W. (2015). Hybridization of seasonal chaotic cloud simulated annealing algorithm in a SVR based load forecasting model. ScienceDirect, 151, Pages 1362-1373. <https://doi.org/10.1016/j.neucom.2014.10.055>
- I. N. da Silva and L. C. M. de Andrade (2016). "Efficient neurofuzzy model to very short-term load forecasting," IEEE Latin America Transactions, vol. 14, no.

- 2, pp. 721–728, 2016.
- Jimenez J., Donado K., and Quintero C. (2017). “A methodology for shortterm load forecasting,” *IEEE Latin America Transactions*, vol. 15, no. 3, pp.400–407, 2017.
- K. Chalupka. (2011). “Empirical evaluation of gaussian process approximation algorithms,” Ph.D.
- Kevin P. Murphy. (2012). *Machine Learning: A Probabilistic Perspective*. Adaptive computation and machine learning series. MIT Press, Cambridge, MA, 2012. ISBN 978-0-262-01802-9.
- Lauret P., David M., Calogine D. (2012). Nonlinear models for short-time load forecasting. *ScienceDirect*, 14, Pages 1404-1409. <https://doi.org/10.1016/j.egypro.2011.12.1109>
- Pasan Karunaratne, Masud Moshtaghi, Shanika Karunasekera, Aaron Harwood, Trevor. (s.f.) Q. Huang, J. Li, and M. Zhu. (2020). “An improved convolutional neural network with load range discretization for probabilistic load forecasting,” *Energy*, vol. 203, p. 117902, 2020. <http://www.sciencedirect.com/science/article/pii/S0360544220310094>
- R. P. Adams, G. E. Dahl, and I. Murray. (2010) “Incorporating side information in probabilistic matrix factorization with gaussian processes,” in *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*, ser. UAI10. Arlington, Virginia, USA: AUAI Press, 2010, p. 1–9.
- Residential probabilistic load forecasting: A method using Gaussian process designed for electric load data. *Applied Energy*, 218, 159172.
- Sheng H., Xiao J. Cheng Y., Ni Q., Wang S. (2017). Short-term solar power forecasting based on weighted Gaussian process regression. *IEEE Xplore*, 65, pages 300-308. 10.1109/TIE.2017.2714127
- Shepero, M., van der Meer, D., Munkhammar, J., & Widén, J. (2018).
- Singh P. Dwivedi P. (2018). Integration of new evolutionary approach with artificial neural network for solving short term load forecast problem. *ScienceDirect*, 217, Pages 537-549. <https://doi.org/10.1016/j.apenergy.2018.02.131>
- Takeda H., Tamura Y., Sato S. (2016). Using the ensemble Kalman filter for electricity load forecasting and analysis. *ScienceDirect*, 104, Pages 184-198. <https://doi.org/10.1016/j.energy.2016.03.070>
- Tarsitano A., Amerise I. (2017). Short term load forecasting using a twostage sarimax model. *ScienceDirect*, 133, Pages 108–114. <https://doi.org/10.1016/j.energy.2017.05.126>
- Van Der Meer D., Shepero M., Svensson A., Widén J., Munkhammar J. (2018). Probabilistic forecasting of electricity consumption, photovoltaic power generation and net demand of an individual building using Gaussian Process. *ScienceDirect*, 213, 195-207. <https://doi.org/10.1016/j.apenergy.2017.12.104>
- W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang. (2019). “Short-term residential load forecasting based on lstm recurrent neural network,” *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 841–851, 2019.

- Wilson, A., & Adams, R. (2013, February). Gaussian process kernels for pattern discovery and extrapolation. In International Conference on Machine Learning (pp. 10671075).
- X. Cao, S. Dong, Z. Wu, and Y. Jing. (2015). "A data-driven hybrid optimization model for short-term residential load forecasting," in 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing, 2015, pp. 283–287
- XM. (junio de 2020). Curva de demanda promedio en Colombia, <https://www.xm.com.co/Paginas/Consumo/historico-de-demanda.aspx>
- Z.-Y. Dou, X. Wang, S. Shi, and Z. Tu. (2020). "Exploiting deep representations for natural language processing," Neurocomputing, vol. 386, pp. 1 – 7, 2020. <http://www.sciencedirect.com/science/article/pii/S0925231219317692>
- Zamo M. Mestre O., Arbogast P., Pannekouke O. (2014). A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production. ScienceDirect, 105, Pages 792-803. <https://doi.org/10.1016/j.solener.2013.12.006>
- Zhang J., Wei Y., Li D., Tan Z., Zhou J. (2018). Short term electricity load forecasting using a hybrid model. ScienceDirect, 158, Pages 774781. <https://doi.org/10.1016/j.energy.2018.06.012>