

Pronóstico robusto de demanda de energía eléctrica: un enfoque basado en optimización.

Aarón José Casadiegos Osorio, Diego Adolfo Mejía Giraldo.

PALABRAS CLAVE:

Pronostico a corto plazo; demanda eléctrica; optimización robusta; integridad de datos.

RESUMEN: Los datos utilizados para el pronóstico de demanda eléctrica son el insumo principal de los modelos de predicción. Aunque normalmente estos datos son tratados y corregidos cuando presentan errores, algunas veces su integridad puede verse afectada por factores externos, como ataques cibernéticos. A pesar de que ya se han propuesto modelos de pronóstico de demanda eléctrica a corto plazo que buscan mitigar estos problemas, son pocas las soluciones planteadas. Estos nuevos modelos, han tenido algunas mejoras en los resultados probando con nuevas funciones objetivos. En este artículo se propone trabajar este reto, por tres frentes que pueden ser combinados: el primero, con un modelo horario y un modelo único; el segundo, usando 2 funciones de pérdida robustas; el tercero, agregando restricciones a los modelos de optimización. Las funciones de pérdida propuestas son la “valor absoluto” y la “Huber”. Las restricciones están enfocadas a no permitir valores atípicos en el pronóstico. Tomando como ejemplo la información de demanda del sistema de EPM, se pudo observar que el modelo horario propuesto supera los modelos de referencia e incluso el modelo único, cuando los datos dañados son los datos de entrenamiento. Cuando los datos dañados son los de pronóstico y el porcentaje de daño es menor al 20%, el modelo horario también presenta resultados superiores.

1. Introducción

El estudio del pronóstico de demanda de energía eléctrica ha incrementado y evolucionado con el transcurso de los años. Abundante literatura se puede consultar al respecto, donde cada autor presenta enfoques diferentes. En [1]-[3], se pueden leer estudios de pronóstico de demanda eléctrica desarrollados en los años noventa. Los autores daban énfasis en desarrollar modelos con operaciones computacionales menos complejas, que se pudieran resolver con la tecnología de la época, basados en modelos de redes neuronales (RNN) y modelos auto-regresivos. En [4]-[6], estudios de principios del 2000, se puede apreciar que los autores lograron hacer modelos más complejos, gracias a los nuevos desarrollos computacionales. Incluso para el 2010, se había avanzado tanto en los modelos de pronóstico a corto plazo, que Tao Hong escribió el libro [7] de su tesis doctoral. Donde recopila gran información sobre los modelos de pronóstico a corto plazo. Destacando la importancia de las variables explicativas, su forma de seleccionarlas y su aplicación a modelos de regresión. Luego del 2010, se ha seguido investigando el pronóstico a corto plazo, con la intención de hacer las predicciones más robustas, ante nuevas eventualidades en la red y fuera de esta. En [8]-[11], los autores buscan minimizar los efectos negativos en el pronóstico, al integrar la generación solar distribuida a la red. En [12], se estudian los efectos en los modelos de predicción, cuando hay cambios en la demanda, causados por eventos sociales difíciles de predecir, como los del covid-19. En [13], los autores buscan que sus modelos de pronóstico no sea afectado por ataques cibernéticos, que dañen la integridad de los datos.

Siguiendo el enfoque de hacer más robusto los modelos de pronóstico, ante ataques cibernéticos o daños en los datos, en este artículo se propone utilizar la función de pérdida norma- l_1 , y la función “Huber” [14]. Funciones que ya se han usado en [15] y [13], para el pronóstico de demanda. Siendo la contribución clave de este documento, combinar dichas funciones de pérdidas con restricciones en el modelo de optimización. Adicionalmente, se formulan dos tipos de arquitectura para los modelos: uno “Horario”, el cual se compone de 24 modelos entrenados para pronosticar por separado cada hora del día; y otro “Único”, el cual es un solo modelo que tiene la capacidad de predecir cualquier hora del día. Para probar el rendimiento de los modelos, se utiliza la métrica MAPE (Error Porcentual Absoluto Medio, por sus

siglas en ingles), para un horizonte de pronóstico de 14 días adelante. El MAPE permite medir el error del pronóstico cuando se llevan a cero, de forma deliberada, parte de los datos de pronóstico o de entrenamiento. Se presentan los resultados comparando los modelos propuestos, con modelos típicos encontrados en la literatura, como lo son regresión con norma- l_2 y Lasso (Operador de Selección y Contracción Mínima Absoluta, por sus siglas en inglés).

2. Metodología

En esta sección se presenta la fundamentación teórica de los modelos de pronóstico propuestos. Los cuales, en general, buscan mantener un comportamiento robusto del pronóstico, ante daños en los datos de entrenamiento o de prueba. Inicialmente, se presentan las variables explicativas del modelo. Luego, se proponen tres estrategias para hacer más robusto el modelo, las cuales pueden ser combinadas. La primera, desde la función objetivo. La segunda, agregando restricciones al modelo (para el conocimiento de los autores, esto no se ha reportado en la literatura sobre modelos de predicción de demanda eléctrica). La tercera, desde la arquitectura del modelo, pudiendo ser “Horario” o “Único”.

2.1. Variables explicativas.

Todos los modelos propuestos en el documento están compuestos por las mismas variables explicativas. Se dividen en datos históricos de la demanda y variables exógenas. Es decir, la demanda y_t del periodo t es modelada como función del comportamiento de la demanda pasada, considerando diferentes datos históricos:

- Historia inmediata: $y_{t-1}, y_{t-2}, \dots, y_{t-k}$; donde k son la cantidad de rezagos horarios.
- Historia semanal: $y_{t-168}, y_{t-169}, \dots, y_{t-336}, y_{t-337}, \dots, y_{t-n(168)-j+1}$; donde j es el # de rezagos horarios, tomados a partir del rezago $n(168)$. Siendo n el # de semanas anteriores.

Se utilizará $y_{[t]}$ para denotar los datos históricos de demanda. La demanda y_t también es parametrizada en términos de las variables exógenas $x_{1,t}, x_{2,t}, \dots, x_{k,t}$. Considerando una variable meteorológica (temperatura). Además, de los efectos calendarios (variables dummies que indican el día de la semana, los días festivos y los días cercanos a festivos). En general, los modelos de pronóstico presentados en este documento se pueden representar en función de las variables explicativas, así: $y_t = f(y_{[t]}, x_{1,t}, x_{2,t}, \dots, x_{k,t} | \beta) + \varepsilon_t$. Donde β son los coeficientes estimados.

2.2. Función objetivo:

Con el fin de reducir la influencia de las anomalías causadas por daños en los datos, al momento de entrenar, se proponen las funciones de penalización del error: norma- l_1 (1) y la función “Huber” (2). Las cuales, en el caso de “Huber”, da menos importancia a la minimización de los errores al cuadrado. Mientras la norma- l_1 , solo tiene en cuenta la minimización de los errores absolutos.

$$\min \sum_{t \in T} |\tilde{y}_t - y_t| \quad (1) \quad \min \sum_{t \in T} \Phi_{hub_t} \quad (2)$$

$$\Phi_{hub_t} = \begin{cases} (\tilde{y}_t - y_t)^2, & |\tilde{y}_t - y_t| \leq M \\ M(2|\tilde{y}_t - y_t| - M), & |\tilde{y}_t - y_t| > M \end{cases}$$

Siendo: \tilde{y}_t , la demanda histórica utilizada para entrenar; y_t , el modelo de pronóstico que se busca entrenar; M , un parámetro de la función “Huber” que delimita cuando se usa la función valor absoluto y la cuadrática.

2.2. Restricciones:

Las restricciones se plantean con el objetivo de mantener los pronósticos dentro de valores típicos dados por la historia.

- Energía típica diaria (ED): $\left| \sum_{h \in H} (y_{h,d}) - \underline{y}_d \right| \leq n_1(\sigma_d), \forall d \in D$; siendo d el tipo de día ($D = \{\text{lunes}, \dots, \text{domingo}\}$); h la hora del día ($H = \{0, 2, \dots, 24\}$); \underline{y}_d y σ_d , la demanda media y la desviación estándar de los datos históricos en el día d , respectivamente; n_1 , parámetro que indica el número de desviaciones estándar.
- Energía típica horaria (EH): $\left| y_{h,d} - \underline{y}_{h,d} \right| \leq n_2(\sigma_{h,d}), \forall h \in H, \forall d \in D$; siendo $\underline{y}_{h,d}$ y $\sigma_{h,d}$, la demanda media y la desviación estándar de los datos históricos en la hora h , del día d . Manteniendo los mismos conjuntos D y H , de la restricción anterior. n_2 , parámetro que indica el número de desviaciones estándar.
- Pendientes típicas de la demanda (PT): $\left| y_{h,d} - \underline{y}_{h-1,d} \right| \leq \max(\tilde{y}_h - \tilde{y}_{h-1}), \forall h \in H, \forall d \in D$; donde $\max(\cdot)$ indica la máxima pendiente entre los datos históricos, que corresponden a la hora h del día d .

2.3. Arquitectura:

La arquitectura del modelo es una idea fundamental en este documento. Se pretende analizar el desempeño de los modelos cuando se especifican de forma general a cualquier hora del día, o cuando se enfocan en cada hora por separado.

2.3.1. Única.

Consiste en una arquitectura básica, que solo tiene un modelo general, capaz de pronosticar cualquier hora del día. Donde, el pronóstico de demanda producido por un modelo con esta arquitectura, es denotado por: $\hat{y}_{t+m|t} = \hat{f}(\hat{\beta})$; donde, se tiene las variables exógenas en el conjunto J ; con un horizonte horario denotado por m ; y $\hat{\beta}$, corresponde a los parámetros estimados.

2.3.2. Horaria.

Esta arquitectura es más compleja, debido a que permite pronosticar cada hora del día, con un modelo enfocado a esa hora en específico. Donde, el pronóstico de demanda producido por un modelo con esta arquitectura, es denotado por: $\hat{y}_{t+m|t} = \hat{f}(\hat{\beta}_h)$; donde, J y m tienen el mismo significado que en el modelo único; pero, se cuenta con parámetros estimados para cada hora h del día, denotados por $\hat{\beta}_h$.

3. Resultados.

Este documento presenta resultados utilizando datos de la demanda horaria, del operador de red EPM de Colombia, registrada entre el 29/06/2017 y el 16/06/2019. Los datos fueron recopilados de la página oficial del Operador de Sistema Independiente de Colombia XM. Además, se cuenta con información de temperatura de la base de datos “Datos Hidrometeorológicos Crudos – Red de Estaciones IDEAM: Temperatura” extraída de la página datos.gov.co. Los modelos de optimización fueron implementados en el software Pyomo 6.01, usando Python 3.8.5. Se usaron 2 solvers: Mosek 9.2.36 y Ipopt 3.12.13. Según la estructura explicada en 2.1, el proceso de entrenamiento empleó $k = 10$ rezagos inmediatos, $j = 4$ rezagos horarios de las $n = 3$ semanas anteriores. Para probar el rendimiento de los modelos se hicieron dos pruebas, que simulan los ataques o daños de los datos. La primera, llevar a cero un porcentaje de los datos de pronóstico. La segunda, llevar a cero un porcentaje de los datos de entrenamiento.

3.1 Evaluación comparativa del rendimiento de los modelos propuestos.

Al realizar la combinación de las tres estrategias propuestas para hacer más robustos los modelos, se formaron 24 modelos diferentes. Para mayor claridad, en la Figura 1, se muestran como se organizaron los modelos propuestos.

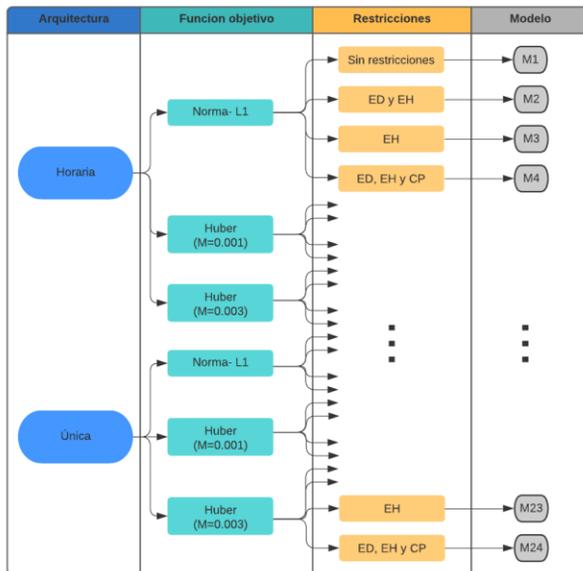


Figura SEQ Figura * ARABIC 1. Modelos

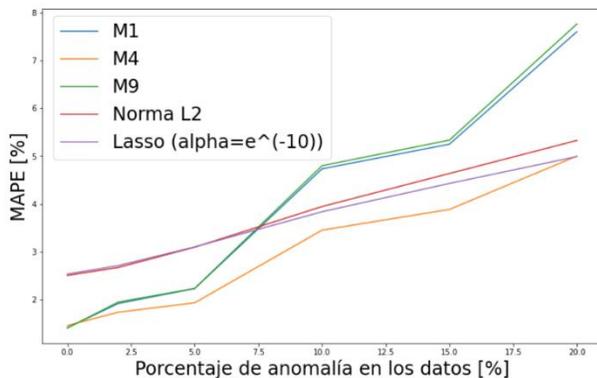


Figura SEQ Figura * ARABIC 2. Progreso del MAPE cuando aumenta el porcentaje de daño en los datos de

Al realizar la predicción sin daños en los datos de pronóstico, se obtuvieron los mejores resultados con los modelos de arquitectura “Horaria”. MAPEs menores a 1.60%, que superan los modelos de arquitectura “Única” y los de referencia (norma- l_2 y Lasso) por más de 1% en el MAPE. Incluso los modelos de arquitectura “Horaria” superan algunos modelos de arquitectura “Única” en más de 5%. Al aumentar los daños en un 10% de los datos de pronóstico, se destaca el comportamiento de los modelos horarios que se combina con las restricciones ED, EH y CP (Modelos M4, M8 y M12).

La Figura 2 muestra el comportamiento del MAPE, de los modelos con mejor desempeño y los de referencia, cuando aumenta el porcentaje de daños en los datos de pronóstico. Como se puede observar, el modelo M4 presenta el mejor comportamiento, hasta cuando se alcanza un 20% de daños. En ese momento, la regresión Lasso se acerca a los resultados del modelo M4. Estos resultados indican que, al combinar las tres restricciones propuestas, con la arquitectura horaria y la función de pérdida norma- l_1 , se logra hacer al modelo más robusto, para altos porcentajes de daños en los datos.

Al hacer pruebas a los modelos con daños en los datos de entrenamiento, se logra observar resultados de los

modelos “Horarios”, que superan con contundencia los resultados de los demás modelos. Para un 5% de daños en los datos de entrenamiento, los modelos del M1 al M12, superan por más de 5% del MAPE a los modelos del M13 al M24 y los de referencia. Siendo mejor calificados los “Horarios” que se usan con la función Huber (M=0.003). En estos casos las restricciones no fueron relevantes para mejorar o empeorar los resultados. Los modelos de referencia Norma- l_2 y Lasso, obtuvieron MAPEs de 9.4672% y 8.1246%, respectivamente. Mientras que los modelos usando la arquitectura “Horaria” no sobrepasaron el 1.7500%.

4. Conclusiones

Este artículo se centró en la consolidación de tres estrategias que permitieran hacer más robustos los modelos de pronóstico de demanda eléctrica ante ataques o daños en los datos. Los modelos de pronóstico tradicionales deben replantearse cuando se trata de implementarlos con datos que pueden ser vulnerables a ataques cibernéticos, ya que no tienen un buen rendimiento, sobre todo cuando los ataques van dirigidos a los datos de entrenamiento. Los modelos con arquitectura “Horaria” se destacan, sobre los demás modelos, al momento de tener daños en los datos de entrenamiento. Si bien cuando se tiene daños en los datos de pronóstico, los resultados del modelo “Horario” con las tres restricciones son muy buenos, los resultados de los demás modelos no difieren tanto. En general, las métricas de desempeño logradas con la arquitectura “Horaria” son prometedoras.

Data Availability Statement

Referencias

- [1] M. A. El-Sharkawi, S. Oh, R. J. Marks, M. J. Damborg, y C. M. Brace, «Short term electric load forecasting using an adaptively trained layered perceptron», en *Proceedings of the First International Forum on Applications of Neural Networks to Power Systems*, Seattle, WA, USA, 1991, pp. 3-6. doi: 10.1109/ANN.1991.213487.
- [2] Y. Al-Rashid y L.D. Paarmann, «Short-term electric load forecasting using neural network models», presentado en *Proceedings of the 39th Midwest Symposium on Circuits and Systems*, Ames, IA, USA, ago. 1996. doi: 10.1109/MWSCAS.1996.593237.
- [3] P. K. Dash, A. C. Liew, y H. P. Satpathy, «A functional-link-neural network for short-term electric load forecasting», *J. Intell. Fuzzy Syst.*, vol. 7, n.º 3, pp. 209-221, ene. 1999.
- [4] M. Tamimi y R. Egbert, «Short term electric load forecasting via fuzzy neural collaboration», *Electr. Power Syst. Res.*, vol. 56, n.º 3, pp. 243-248, dic. 2000, doi: 10.1016/S0378-7796(00)00123-1.
- [5] M. Caciotta, R. Lamedica, V. O. Cencelli, A. Prudenzi, y M. Sforna, «Application of artificial neural networks to historical data analysis for short-term electric load forecasting», *Eur. Trans. Electr. Power*, vol. 7, n.º 1, pp. 49-56, sep. 2007, doi: 10.1002/etep.4450070108.
- [6] K. Metaxiotis, A. Kagiannas, D. Askounis, y J. Psarras, «Artificial intelligence in short term electric load forecasting: a state-of-the-art survey for the researcher», *Energy Convers. Manag.*, vol. 44, n.º 9, pp. 1525-1534, jun. 2003, doi: 10.1016/S0196-8904(02)00148-6.
- [7] T. Hong, «Short Term Electric Load Forecasting», North Carolina State University, Raleigh, North Carolina, 2010.
- [8] F. A. Monforte, C. Fordham, J. Blanco, S. Barsun, A. Kankiewicz, y B. Norris, «Improving Short-Term Load Forecasts by Incorporating Solar PV Generation», Itron, San Diego, California, 2016. Accedido: may 07, 2021. [En línea]. Disponible en: <https://www.itron.com/-/media/feature/products/documents/white-paper/101525wp01-forecast-simulations.pdf>
- [9] F. A. Monforte, «Developing Net Load Uncertainty Forecasts to Support System Operations», mar. 2019. [En línea]. Disponible en: <https://blogs.itron.com/developing-net-load-uncertainty-forecasts-to-support-system-operations/>
- [10] Y. Wang, N. Zhang, Q. Chen, D. S. Kirschen, P. Li, y Q. Xia, «Data-Driven Probabilistic Net Load Forecasting With High Penetration of Behind-the-Meter PV», *IEEE Trans. Power Syst.*, vol. 33, n.º 3, pp. 3255-3264, may 2018, doi: 10.1109/TPWRS.2017.2762599.
- [11] «Italian protocol for massive solar integration: Imbalance mitigation strategies», *Renew. Energy*, vol. 153, pp. 725-739, jun. 2020, doi: 10.1016/j.renene.2020.01.145.
- [12] «A hybrid multi-objective optimizer-based model for daily electricity demand prediction considering COVID-19», *Energy*, vol. 219, p. 119568, mar. 2021, doi: 10.1016/j.energy.2020.119568.
- [13] J. Luo, T. Hong, y S.-C. Fang, «Robust Regression Models for Load Forecasting», *IEEE Trans. Smart Grid*, vol. 10, n.º 5, pp. 5397-5404, sep. 2019, doi: 10.1109/TSG.2018.2881562.
- [14] S. Boyd y L. Vandenberghe, *Convex optimization*. Cambridge, UK ; New York: Cambridge University Press, 2004.
- [15] Z. Zhang, H. Chen, Y. Huang, y W.-J. Lee, «Quantile Huber Function Guided TCN for Short-Term Consumer-Side Probabilistic Load Forecasting», en *2020 IEEE/IAS Industrial and Commercial Power System Asia (I&CPS Asia)*, Weihai, China, jul. 2020, pp. 322-329. doi: 10.1109/ICPSAsia48933.2020.9208545.