

Predicción del precio de bolsa de energía eléctrica Del Mercado Energía Mayorista Colombiano

Andrea Margarita Beleño
200620739
E-mail: a.beleno@uniandes.edu.co

María Valeria Gaona
202214418
E-mail: mv.gaona@uniandes.edu.co

Resumen—El precio de bolsa de energía eléctrica del Mercado de Energía Mayorista (MEM) Colombiano está dado por diversos factores para generar un ambiente propicio de competencia y formación eficiente de precios, que le permitan a la demanda tener precios óptimos para el Sistema Interconectado Nacional-SIN. Es por ello que es fundamental el análisis de este mercado para que a medida que pase el tiempo, los agentes que participan en este mercado puedan identificar factores de riesgo más rápido y así, tomar las mejores decisiones para la economía y distribución de energía del país. Por lo tanto, de acuerdo con el análisis anterior, contar con una herramienta como una aplicación web que pueda predecir este valor genera una dinámica más eficiente para los agentes que están dentro del sector y a su vez, a ciudadanos que estén interesados en el campo energético. Para acceder a la aplicación web, se presenta el siguiente enlace: https://aplicaciones-web-abeleno.shinyapps.io/pred_precio_bolsa_horario/ A su vez, el link al repositorio del presente documento se encuentra en el siguiente enlace: <https://github.com/mvgaona/Proyecto-final-MEcA-4107>

I. INTRODUCCIÓN

El mercado eléctrico colombiano es un mercado competitivo en donde participan generadores, transmisores, distribuidores, consumidores y comercializadores de energía. Este mercado se divide en dos segmentos: Corto y largo plazo, sin embargo, en el siguiente documento se presentará el análisis del mercado en corto plazo por medio de la bolsa de energía de Colombia, la cual es administrada por XM, en donde se presenta la participación de generadores y comercializadores de energía para la compra y venta a precio de bolsa de energía eléctrica, con el objetivo de suplir la demanda adecuada de energía en el país.

De acuerdo con Trespalacios, Pantoja & Fernandez (2017) [5], el mercado spot o la bolsa de energía hace referencia al mercado en donde se obtiene la energía eléctrica de forma instantánea, con el objetivo de lograr un balance entre oferta y demanda. A su vez, los autores afirman que dicho precio de bolsa se define mediante un conjunto de normas que buscan precisar el nivel de referencia en caso de escasez.

Poveda (2012) [1] afirma que el despacho ideal es el programa de generación que está dado por el uso de los recursos más económicos hasta cubrir la demanda doméstica real, más las Transacciones Internacionales de Electricidad de Corto Plazo - TIE (Exportaciones como demanda e importaciones como generación), más las pérdidas del STN (Sistema de Transmisión Nacional). Teniendo en cuenta lo anterior, el precio de bolsa está dado por el precio de oferta obtenido por

medio del despacho ideal, el cual es utilizado para valorar los intercambios en bolsa.

El correcto funcionamiento del mercado eléctrico es fundamental para el análisis de la demanda energética en el país: si es necesario realizar estrategias en el manejo de los recursos naturales con los que se genera energía, si la dinámica de compra y venta de energía está siendo óptima para la economía y sociedad colombiana. Generar una proyección de estos precios permite poder hacer inferencia acerca de cómo el mercado puede estar funcionando y aunque este sea un sistema fluctuante, se puede generar predicciones acerca de su comportamiento.

En el siguiente documento se encuentra el análisis acerca de los datos recaudados para la predicción del precio de bolsa de energía eléctrica del mercado de energía mayorista colombiano, en donde se implementará un modelo de Machine Learning automático en una aplicación web, que permitirá modelar futuros precios de bolsa y con ello, tomar decisiones comerciales basadas en los datos adquiridos

II. DATOS

El precio de bolsa nacional de energía puede estar dado por diferentes factores que representan las condiciones del mercado energético colombiano, por ejemplo. Este precio se genera cada hora y, por lo tanto, es necesario considerar 24 modelos correspondientes a cada momento del día en donde se presentan cambios en este valor. Por lo tanto, para considerar realizar un modelo de predicción es necesario contar con factores determinantes que expliquen cómo se produce este precio. Por otra parte, dentro del análisis se recopiló información desde el 01/01/2000 hasta 30/06/2022, en donde no se presenta información faltante que deba ser imputada. En total la base cuenta con 8.217 observaciones.

Los datos a utilizar para el desarrollo de la predicción, se encuentran en la página del operador del mercado eléctrico colombiano XM. Esta empresa concentra todos los parámetros del sistema eléctrico colombiano que son importantes a la hora de realizar la predicción del precio de bolsa. Los predictores están dados por:

-*Generación por tipo de recursos (kWh)*: Los recursos naturales permiten abastecer de energía eléctrica a las comunidades. Además, el precio de bolsa está representando por la generación de energía por medio de estos recursos, ya que esto

permite que los agentes puedan continuar con su producción de energía y así, establecer dicho valor. Por lo tanto, existen 5 tipos de generación: Eólica (generación a partir del viento), solar (generación a partir de la radiación solar), hidráulica (generación a partir del agua), cogenerador (generación de energía eléctrica a partir de procesos industriales) y térmica (a partir de combustibles fósiles).

Esta variable numérica cuenta con el análisis descriptivo de cada uno de los tipos de generadores por cada hora, que se muestran en los Cuadros III, IV, V, VI y VII. Como se observa en los Cuadros, este análisis realizado para cada hora del día en el marco de fechas establecidas anteriormente. Se observa que el tipo de generación que abastece la demanda de energía eléctrica en el país es la generación hidráulica y en segundo lugar la generación térmica. Para la generación solar, eólica y de cogeneración, como no existen muchas plantas, estas presentan valores de generación iguales a cero; para la generación solar, se observa que para las horas de la noche, los valores son bajos a nulos.

-Índice Interoceánico de El Niño (ONI): Este indicador refleja la variación de la temperatura en el océano pacífico. Por lo tanto, al contar con un índice alto o mayor a 0.5 se presentan sequías, es decir, hay fenómeno del niño y si, por el contrario, el índice presenta valores pequeños o inferiores a -0.5 existe un fenómeno de la niña, en donde se presentan muchas lluvias. Por lo tanto, por ejemplo, al presentarse el fenómeno del niño, los niveles de los ríos bajan, se debe priorizar la generación a través de plantas térmicas, aumentando costos y llevando a que el precio de bolsa energético también aumente. Este indicador se da de manera mensual, se realizó la interpolación a datos diarios y se asume que es la misma para cada hora del día. Se obtuvo la información a través de web scrapping de la página web en [3].

-Aportes totales (Aportes Hídricos en kWh): La variable representa cuanta energía se puede obtener de los ríos en Colombia. El precio de bolsa energético está altamente representado por esta variable ya que, la fuente principal energética del país es por medio del agua, ya que la matriz energética es 69 % hidráulica citeb2. Por lo tanto, Colombia depende de los ríos abastecer de energía a la sociedad. Es decir, a medida que aumenta el caudal en los ríos, el precio de bolsa disminuye. Este indicador se da de manera diaria, se asume que es la misma para cada hora del día.

-Tasa Representativa del mercado-TRM: Este indicador hace referencia a la cantidad de pesos colombianos por un dólar estadounidense. Esta variable es relevante dentro de los modelos, debido a que el combustible térmico es pagado con la moneda dólar y por lo tanto, la variación de este valor genera un impacto, ya que estos recursos son distribuidos por medio de una moneda internacional, es decir, al contar con la moneda local devaluada, el precio de estos insumos aumenta. Este indicador se da de manera diaria, se asume que es la misma para cada hora del día.

Para las variables numéricas mencionadas anteriormente, se presenta la Figura 1 donde se encuentra el análisis descriptivo.

Análisis Descriptivo/ Variable	Mínimo	Media	Máximo	Moda
TRM	1624,41	2536,017	4153,91	1907,06
Aportes Totales	25253400	142961979	591245300	82310700
ONI	-1,7	-0,09062918	2,6	-0,3

Figura 1. Cuadro descriptivo para TRM, ONI y Aportes totales

III. MODELO

Con base en las variables presentadas en la sección II, se procedió a realizar, considerando series de tiempo, los modelos con diferentes hiperparámetros teniendo como base el *XGBoost*. Se debe tener en cuenta, que el precio de bolsa del despacho ideal se da de manera horaria, por lo cual, se realizaron las predicciones para cada hora del día, considerando los predictores: *TRM*, *Aportes hídricos*, *ONI*, la fecha (se dividió en *año*, *mes* y *día* y el tipo de Generación, 5 tipos de generación por 24 horas del día), dependiendo del recurso. Al ser series de tiempo los datos utilizados para desarrollar el presente trabajo, se dividió la base de datos en *train* con el 70 % de los datos desde el 2000-01-01 y el restante 30 % fueron las fechas subsiguientes. Cabe anotar que para series de tiempo, las muestras de las bases a analizar deben ser secuenciales, por lo cual, dichas bases no se obtuvieron de manera aleatoria.

Ahora bien, se consideró que el modelo *XGBoost* era el más apto para realizar la predicción del precio bolsa, teniendo en cuenta que este modelo se considera el mejor para aplicarlo a casos donde la no linealidad y la volatilidad son importantes. Se entrenaron varios modelos, de los cuales se presentan resultados de 2, en donde los hiperparámetros variaron para estos modelos para cada predicción del precio de bolsa para cada hora del día. Se explicará en breve los aspectos relevantes de cada modelo. Se debe tener en cuenta que para todos los modelos se utilizaron todas las variables contempladas en la sección II, es decir, 151 variables en total.

- **Modelo 1:** Hiperparámetros importantes: $\lambda = 0,3$ (tasa de aprendizaje), $d = 3$ (número de bifurcaciones), $n = 100$ (número de iteraciones). Esto se aplicó para la predicción del precio de bolsa horario para las 24 horas.
- **Modelo 2:** Hiperparámetros importantes: $\lambda = 0,3$ (tasa de aprendizaje), $d = 100$ (número de bifurcaciones), $n = 1000$ (número de iteraciones). Esto se aplicó para la predicción del precio de bolsa horario para las 24 horas.

El parámetro de comparación entre modelos fue el RMSE. Para cada precio de bolsa horario se observó que los modelos presentados anteriormente, tenían diferente RMSE, tomándose el más bajo para realizar la predicción en la aplicación web. En el Cuadro II se presentan los resultados del RMSE para cada predicción de bolsa. En resumen, los precios de bolsa horarios a los cuales se les aplicó el modelo 1 o 2, se presentan a continuación:

- **Modelo 1:** Predicción del precio de bolsa para las horas: 0, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 17, 18, 22 y 23
- **Modelo 2:** Predicción del precio de bolsa para las horas: 1, 11, 13, 14, 16, 19, 20 y 21.

Para el modelo 2, las horas predichas coinciden con las de mayor demanda en el sistema eléctrico colombiano.

III-A. Implementación en App web

Ahora bien, para la implementación del modelo en la aplicación web, no fue necesario correr todo el código desarrollado en el modelo, sin embargo, para los predictores correspondientes a generación, se realizó una restricción de los datos a ingresar. Al tener los datos de las capacidades instaladas en generación del sistema (potencia), se obtuvo la suma de las capacidades de todos los generadores por tipo de generación, y así se obtuvo el rango en dichos predictores. Los resultados se muestran en el Cuadro I. No se puede asignar más capacidad de la que se encuentra instalada en el sistema de manera horaria.

Cuadro I
CAPACIDADES MÁXIMAS EN EL SIN COLOMBIANO POR TIPO DE GENERACIÓN

Tipo de Generación	Capacidad (kW)
Cogenerador	178.400
Eólica	20.000
Hidráulica	11.790.431
Térmica	8.823.450
Solar	157.726

IV. RESULTADOS

Para implementar el modelo en una aplicación web, utilizamos el paquete de *Shiny* de R con los códigos especiales para que sean ejecutados mediante una página web. Para la implementación, se guardaron los modelos que menor RMSE tuvieron para cada modelo de precio horario y estos son los que se ejecutan en la página para la visualización e ingreso de variables por parte de un usuario. En la Figura 2 se muestra el entorno de la aplicación web de acceso público pero con limitación en el tiempo, con posibilidad de variar los campos a criterio y conocimiento del usuario, y después de unos segundos, se muestra el resultado de la aplicación.



Figura 2. Resultado de la aplicación web



Figura 3. Ejemplo para Agosto 20/2022



Figura 4. Ejemplo para Septiembre 10/2022

De acuerdo con la aplicación web descrita anteriormente, cuyo enlace se encuentra en el abstract del presente documento, se presentan dos ejemplos de predicciones con diferentes valores reportados en las variables para Agosto 20/2022 y Septiembre 10/2022 en las Figuras 3 y 4, respectivamente.

V. CONCLUSIONES Y RECOMENDACIONES

- El precio de bolsa energético está caracterizado por los factores anteriormente descritos y analizados. Sin embargo, existen factores exógenos a los modelos, como la volatilidad que puede presentarse dentro de un mercado eléctrico y sus operaciones cada hora. Por otra parte, aunque se esté analizando en términos económicos este valor, es fundamental considerar que este precio está influenciado por el impacto ambiental y el calentamiento global, haciendo que cuando se generen sequías, este precio aumente considerablemente, ya que otro tipo de generadores son más costosos y por ende, el precio de energía también.
- La implementación de modelos de series de tiempo requiere mantener la secuencia de los datos, lo cual fue un reto al ejecutar ambos modelos.
- Realizando un análisis de las variables utilizadas, se observa que el precio está influenciado principalmente por la generación hidráulica, predictor para el cual no se tiene valor de generación 0 y que al variarse genera un cambio en la predicción.
- Por otra parte, la predicción de series de tiempo por medio de XGBoost aporta sustancialmente al proceso académico, ya que hace que se afronten retos en su análisis e interpretación de resultados. A su vez, el proceso de creación de aplicación web generó un desafío

adicional dentro del análisis, además, permite contar con herramientas futuras para diferentes proyectos en los que una aplicación web aporte un valor agregado y sea útil.

Finalmente, mediante la predicción del precio de bolsa energética por hora, se pudo evidenciar limitaciones y recomendaciones.

V-A. Limitaciones

- Se encontró como limitación, que debido a que el precio de bolsa energético es muy volátil y depende las ofertas en generación que realicen los agentes, se puede presentar algunas imprecisiones en la predicción de este valor, debido a que no fue posible disminuir el RMSE.

V-B. Recomendaciones

- Debido al hallazgo de la volatilidad de este precio, es necesario crear más estrategias que mitiguen este impacto volátil inevitable por medio de inclusión de más variables en los modelos por hora, por ejemplo, los parámetros técnicos de los generadores (información no disponible de manera pública por parte de XM).
- En cuanto a la Aplicación Web, se puede extender para que el usuario de la interfaz pueda escoger con qué otros modelos realice la predicción.
- Se observa que al variar las fechas, los precios no se afectan significativamente, por lo cual, se debe profundizar en la implementación de modelos de Boosting para series de tiempo.

REFERENCIAS

- [1] Poveda Núñez, M. A. (2012) "Modelamiento del precio de bolsa. Universidad Nacional de Colombia." Bogotá D.C, Recuperado de: <https://repositorio.unal.edu.co/bitstream/handle/unal/21159/300038.2012.pdf?sequence=1&isAllowed=y>
- [2] XM (2022). Sinergox de XM. Recuperado de: <https://sinergox.xm.com.co/trpr/Paginas/Historicos/Historicos.aspx>.
- [3] Climate Prediction Center. "Cold & Warm episodes by season" Recuperado de: https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php
- [4] XGBoost time series forecast in R. "xgboost time series forecast in R" Recuperado de: <http://datasideoflife.com/?p=1009#:~:text=xgboost%2C%20or%20Extreme%20Gradient%20Boosting,while%20doing%20time%20series%20predictions.>
- [5] Trespalacios Carrasquilla, A., Pantoja Robayo, J. O., & Fernández Taborda, Ó. A. (2017). Análisis de mercados de electricidad. EAFIT.

APÉNDICE

CUADROS DE VARIABLES DESCRIPTIVAS Y RMSE

Se presentan los Cuadros donde se consignan las estadísticas descriptivas de las variables explicadas en la sección II y III.

Cuadro II
COMPARACIÓN DE RMSE PARA MODELO CON BIFURCACIÓN 3 Y 100

Modelo	RMSE_3	RMSE_100
Predicción_h0	112,29	141,76
Predicción_h1	120,89	119,11
Predicción_h2	109,63	112,36
Predicción_h3	112,78	115,21
Predicción_h4	109,59	115,15
Predicción_h5	114,35	127,95
Predicción_h6	139,17	146,15
Predicción_h7	126,32	133,02
Predicción_h8	131,53	138,11
Predicción_h9	131,62	136,65
Predicción_h10	172,78	217,50
Predicción_h11	213,59	201,64
Predicción_h12	126,65	137,01
Predicción_h13	126,91	124,82
Predicción_h14	134,25	129,88
Predicción_h15	129,50	135,12
Predicción_h16	133,62	130,76
Predicción_h17	111,18	125,89
Predicción_h18	121,82	141,82
Predicción_h19	158,05	136,94
Predicción_h20	150,23	136,19
Predicción_h21	125,02	120,78
Predicción_h22	126,08	151,35
Predicción_h23	120,46	132,46

Cuadro III
CUADRO DESCRIPTIVO PARA TIPO DE GENERACIÓN MEDIANTE COGENERACIÓN

Hora	Mín.	Máx.	Media	Moda
h0	0	410.146	40.840,883	0
h1	0	150.458,79	40.902,0432	0
h2	0	14.9261,96	41.014,700	0
h3	0	150.293,93	41.142,547	0
h4	0	147.386,04	41.050,650	0
h5	0	151.134,87	41.022,724	0
h6	0	152.285,89	40.554,742	0
h7	0	148.395,45	40.456,582	0
h8	0	145.334,36	40.113,777	0
h9	0	143.457,07	39.222,896	0
h10	0	141.100,67	38.617,483	0
h11	0	143.005,01	38.375,524	0
h12	0	148.322,67	38.565,889	0
h13	0	144.766,57	38.777,087	0
h14	0	143.230,34	38.465,044	0
h15	0	142.148	38.513,413	0
h16	0	145.504,87	38.936,497	0
h17	0	144.053,28	39.410,824	0
h18	0	138.961,56	39.291,810	0
h19	0	145.042,94	39.659,454	0
h20	0	142.540,31	40.354,904	0
h21	0	144.540,66	40.660,369	0
h22	0	146.768,19	40.348,729	0
h23	0	143.285,56	40.450,315	0

Cuadro IV
CUADRO DESCRIPTIVO PARA TIPO DE GENERACIÓN HIDRÁULICA

Hora	Mín.	Máx.	Media	Moda
h0	1.417.548,91	7.628.404,46	4.528.715,493	1.417.548,91
h1	1.330.047,13	7.647.431,1	4.398.223,237	1.330.047,13
h2	1.324.454,7	7.557.410,8	4.298.109,397	1.324.454,7
h3	1.283.247,27	7.416.477,69	4.268.356,919	1.283.247,27
h4	1.183.501,72	7.536.784,76	4.424.014,275	1.183.501,72
h5	1.179.017,14	7.833.471,26	4.822.082,199	5.570.182,61
h6	970.812,04	8.091.892,65	4.993.436,200	3.995.987,35
h7	1.306.117,5	8.618.320,45	5.303.482,864	1.306.117,5
h8	1.664.275,94	9.178.610,74	5.705.667,277	1.664.275,94
h9	1.806.851,95	9.442.061,34	5.952.079,715	1.806.851,95
h10	92.118,27	9.678.024,4	6.185.775,403	92.118,27
h11	765.668,02	9.986.122,45	6.352.912,141	765.668,02
h12	2.129.648,15	9.975.165,45	6.200.989,860	2.129.648,15
h13	1.335.052,8	9.983.786,18	6.103.800,306	1.335.052,8
h14	1.970.576,06	10.130.255,68	6.110.836,725	1.970.576,06
h15	1.795.877,68	10.060.447,37	6.074.021,695	1.795.877,68
h16	1.666.592,38	9.927.489,34	6.007.630,163	1.666.592,38
h17	1.268.806,53	9.897.813,48	6.063.786,198	5.633.259,81
h18	2.612.639,5	10.322.512,92	6.978.153,396	2.612.639,5
h19	3.529.632,31	10.401.510,03	7.246.731,128	3.529.632,31
h20	3.350.340,67	10.158.243,39	6.938.421,981	3.350.340,67
h21	2.478.786,26	9.737.813,23	6.401.386,069	2.478.786,26
h22	1.888.084,34	9.148.863,62	5.656.073,321	1.888.084,34
h23	1.106.884,57	8.589.358,18	4.989.189,830	1.106.884,57

Cuadro V
CUADRO DESCRIPTIVO PARA TIPO DE GENERACIÓN TÉRMICA

Hora	Mín.	Máx.	Media	Moda
h0	0	4.346.958,35	952.843,162	0
h1	0	4.376.783,85	866.869,172	0
h2	0	4.389.886,65	842.063,152	0
h3	0	4.402.622,4	836.680,849	0
h4	0	4.400.095,75	850.738,131	0
h5	0	4.400.159,45	874.619,085	0
h6	0	4.396.572,75	888.079,991	0
h7	0	4.392.592,85	908.524,226	0
h8	0	4.397.880,86	9.30.736,171	0
h9	0	4.381.757	942.592,613	0
h10	0	4.284.917,55	954.549,728	150.000
h11	0	4.248.899,11	966.628,179	150.000
h12	0	4.263.505,75	966.729,394	150.000
h13	0	4.272.879,6	965.323,129	150.000
h14	0	4.375.807,45	967.056,145	150.000
h15	0	4.391.872,75	968.643,591	150000
h16	0	4.394.330,1	972.772,873	150000
h17	0	4.397.210,45	989.960,17	150000
h18	0	4.399.168,6	1.023.728,482	150000
h19	0	4.405.031,2	1.040.641,069	150000
h20	0	4.409.350,3	1.022.014,044	150000
h21	0	4.415.668,25	996.256,082	150000
h22	0	4.387.023,65	958.256,776	150000
h23	0	4.379.308,16	918.575,431	0

Cuadro VI
CUADRO DESCRIPTIVO PARA TIPO DE GENERACIÓN SOLAR

Hora	Mín.	Máx.	Media	Moda
h0	0	479,26	0,159	0
h1	0	471,58	0,155	0
h2	0	459,8	0,152	0
h3	0	450,15	0,151	0
h4	0	461,67	0,157	0
h5	0	940,49	11,715	0
h6	0	48.263,35	1.270,431	0
h7	0	122.211,63	5.003,544	0
h8	0	139.131,62	8.256,382	0
h9	0	164.277,33	10.391,743	0
h10	0	167.135,36	11.581,421	0
h11	0	175.221,75	12.069,044	0
h12	0	164.512,27	11.996,089	0
h13	0	164.651,47	11.402,893	0
h14	0	154.896,74	10.246,819	0
h15	0	140.096,57	8.208,544	0
h16	0	993.22,75	4.685,860	0
h17	0	25.263,59	834,022	0
h18	0	5.485,35	3,113	0
h19	0	520,83	0,219	0
h20	0	468,67	0,166	0
h21	0	466,98	0,178	0
h22	0	468,25	0,162	0
h23	0	469,52	0,160	0

Cuadro VII
CUADRO DESCRIPTIVO PARA TIPO DE GENERACIÓN EÓLICA

Hora	Mín.	Máx.	Media	Moda
h0	0	16.557,13	3.366,288	0
h1	0	16.074,14	3.146,872	0
h2	0	16.136,39	2.961,920	0
h3	0	15.162,08	2.810,429	0
h4	0	15.586,59	2.656,932	0
h5	0	15.242,7	2.550,637	0
h6	0	16.500,46	2.574,470	0
h7	0	16.616,68	3.010,651	0
h8	0	16.917,75	3.463,020	0
h9	0	17.026,27	3.819,577	0
h10	0	17.680,5	4.338,920	0
h11	0	17.564,76	5.181,301	0
h12	0	17.660,44	6.227,226	0
h13	0	17.989,05	7.006,291	0
h14	0	18.423,85	7.549,247	0
h15	0	18.345,76	7.841,002	0
h16	0	17.863,57	7.736,863	0
h17	0	17.865,31	7.151,004	0
h18	0	17.896,9	6.101,1048	0
h19	0	17.346,61	5.195,802	0
h20	0	17.366,73	4.618,033	0
h21	0	17.777,47	4.226,125	0
h22	0	17.365,74	3.903,933	0
h23	0	17.181,65	3.624,142	0