

# 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. 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/](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 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 6 tipos de generadores: Eólica, solar, hidráulica, cogenerador y térmica.

Esta variable numérica cuenta con siguiente análisis descriptivo de cada uno de los tipos de generadores. Como se observa en las siguientes gráficas, este análisis realizado para cada hora del día en el marco de fechas establecidas anteriormente.

-*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, al presentarse el fenómeno del niño, los agentes generadores deben producir energía térmica, 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. 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 I 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, 23.
- **Modelo 2**: Predicción del precio de bolsa para las horas: 1, 11, 13, 14, 16, 19, 20 y 21.

### 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

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 202022 y Septiembre 102022 en las Figuras 3 y 4, respectivamente.



Figura 3. Ejemplo para Agosto 20/2022



Figura 4. Ejemplo para Septiembre 10/2022

## V. CONCLUSIONES Y RECOMENDACIONES

- Se encontró como limitación, que debido a que el precio de bolsa energético es muy volátil y depende de 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. 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 diferentes más variables en los modelos por hora.

- Fue un bastante interesante obtener los datos geográficos para utilizarlos como predictores o para imputar datos de variables existentes, fue en la parte del PS3 que más se demoró para obtener y que representó un gran reto y esfuerzo, sin embargo, consideramos la información obtenida muy relevante y que puede servir para futuras aplicaciones.
- Los resultados más aceptables para el ejercicio fueron los que tuvieron en cuenta mayor cantidad de variables para obtener la predicción.
- Para el presente trabajo, el modelo con menor RMSE y que también presentó una menor relación precio vs vivienda, fue el escogido para realizar la predicción y fue el que se entrenó posteriormente para obtener la predicción.
- Como recomendación, con mayor cantidad de tiempo, se hubiese podido realizar otros modelos variando los hiperparámetros del modelo *XGBoost*.

## V-A. 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.

## REFERENCIAS

- 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>
- XM (2022). Sinergox de XM. Recuperado de: <https://sinergox.xm.com.co/trpr/Paginas/Historicos/Historicos.aspx>.
- Climate Prediction Center."Cold & Warm episodes by season" Recuperado de: [https://origin.cpc.ncep.noaa.gov/products/analysis\\_monitoring/ensostuff/ONI\\_v5.php](https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php)
- 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.>
- 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 I

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

Hora	Mín.	Máx.	Media	Moda
h0	1417548.91	7628404.46	4528715.49314105	1417548.91
h1	1330047.13	7647431.1	4398223.23793477	1330047.13
h2	1324454.7	7557410.8	4298109.39673482	1324454.7
h3	1283247.27	7416477.69	4268356.91911647	1283247.27
h4	1183501.72	7536784.76	4424014.27556042	1183501.72
h5	1179017.14	7833471.26	4822082.19976634	5570182.61
h6	970812.04	8091892.65	4993436.19984057	3995987.35
h7	1306117.5	8618320.45	5303482.8637897	1306117.5
h8	1664275.94	9178610.74	5705667.2770415	1664275.94
h9	1806851.95	9442061.34	5952079.71550201	1806851.95
h10	92118.27	9678024.4	6185775.40317756	92118.27
h11	765668.02	9986122.45	6352912.1410819	765668.02
h12	2129648.15	9975165.45	6200989.86083364	2129648.15
h13	1335052.8	9983786.18	6103800.30594012	1335052.8
h14	1970576.06	10130255.68	6110836.72523792	1970576.06
h15	1795877.68	10060447.37	6074021.69582573	1795877.68
h16	1666592.38	9927489.34	6007630.16346112	1666592.38
h17	1268806.53	9897813.48	6063786.19773153	5633259.81
h18	2612639.5	10322512.92	6978153.39616162	2612639.5
h19	3529632.31	10401510.03	7246731.12784228	3529632.31
h20	3350340.67	10158243.39	6938421.98133991	3350340.67
h21	2478786.26	9737813.23	6401386.06922965	2478786.26
h22	1888084.34	9148863.62	5656073.32120239	1888084.34
h23	1106884.57	8589358.18	4989189.83028356	1106884.57

Cuadro II

Hora	Mín.	Máx.	Media	Moda
h0	0	410146	40840.8838773275	0
h1	0	150458.79	40902.0432213703	0
h2	0	149261.96	41014.7000924912	0
h3	0	150293.93	41142.5474285019	0
h4	0	147386.04	41050.6506863819	0
h5	0	151134.87	41022.7245673603	0
h6	0	152285.89	40554.742469271	0
h7	0	148395.45	40456.5829962273	0
h8	0	145334.36	40113.7776645978	0
h9	0	143457.07	39222.8965863454	0
h10	0	141100.67	38617.4831520019	0
h11	0	143005.01	38375.5244128027	0
h12	0	148322.67	38565.8891420226	0
h13	0	144766.57	38777.0870025557	0
h14	0	143230.34	38465.0441681879	0
h15	0	142148	38513.4134611172	0
h16	0	145504.87	38936.497754655	0
h17	0	144053.28	39410.824948278	0
h18	0	138961.56	39291.8108117318	0
h19	0	145042.94	39659.4548156261	0
h20	0	142540.31	40354.9041365462	0
h21	0	144540.66	40660.369437751	0
h22	0	146768.19	40348.7295338932	0
h23	0	143285.56	40450.3159936717	0

Cuadro IV

Hora	Mín.	Máx.	Media	Moda
h0	0	4346958.35	952843.162151637	0
h1	0	4376783.85	866869.172104174	0
h2	0	4389886.65	842063.152084703	0
h3	0	4402622.4	836680.849206523	0
h4	0	4400095.75	850738.130781307	0
h5	0	4400159.45	874619.086799318	0
h6	0	4396572.75	888079.99072776	0
h7	0	4392592.85	908524.226031398	0
h8	0	4397880.86	930736.170962638	0
h9	0	4381757	942592.61255446	0
h10	0	4284917.55	954549.727831325	150000
h11	0	4248899.11	966628.178724595	150000
h12	0	4263505.75	966729.399020324	150000
h13	0	4272879.6	965323.129386637	150000
h14	0	4375807.45	967056.145185591	150000
h15	0	4391872.75	968643.59182305	150000
h16	0	4394330.1	972772.87397347	150000
h17	0	4397210.45	989960.175544603	150000
h18	0	4399168.6	1023728.48206523	150000
h19	0	4405031.2	1040641.06900085	150000
h20	0	4409350.3	1022014.04499696	150000
h21	0	4415668.25	996256.08237678	150000
h22	0	4387023.65	958256.776575392	150000
h23	0	4379308.16	918575.431517586	0

Cuadro V

Hora	Mín.	Máx.	Media	Moda
h0	0	479.26	0.159219909942801	0
h1	0	471.58	0.154964098819521	0
h2	0	459.8	0.152128514056225	0
h3	0	450.15	0.151263234757211	0
h4	0	461.67	0.156680053547523	0
h5	0	940.49	11.7147462577583	0
h6	0	48263.35	1270.43153218936	0
h7	0	122211.63	5003.54400754533	0
h8	0	139131.62	8256.38221005233	0
h9	0	164277.33	10391.7433199465	0
h10	0	167135.36	11581.4211159791	0
h11	0	175221.75	12069.044162103	0
h12	0	164512.27	11996.0893489108	0
h13	0	164651.47	11402.8934136546	0
h14	0	154896.74	10246.818991116	0
h15	0	140096.57	8208.54441036875	0
h16	0	99322.75	4685.86007058537	0
h17	0	25263.59	834.02225508093	0
h18	0	5485.35	3.11297310453937	0
h19	0	520.83	0.219476694657418	0
h20	0	468.67	0.166227333576731	0
h21	0	466.98	0.178140440550079	0
h22	0	468.25	0.162139466958744	0
h23	0	469.52	0.160421078252404	0

Cuadro VI

Hora	Mín.	Máx.	Media	Moda
h0	0	16557.13	3366.2878203724	0
h1	0	16074.14	3146.87241328952	0
h2	0	16136.39	2961.92061701351	0
h3	0	15162.08	2810.42970062066	0
h4	0	15586.59	2656.93201411707	0
h5	0	15242.7	2550.63747961543	0
h6	0	16500.46	2574.47079834489	0
h7	0	16616.68	3010.65102470488	0
h8	0	16917.75	3463.02059754168	0
h9	0	17026.27	3819.57742850189	0
h10	0	17680.5	4338.9201350858	0
h11	0	17564.76	5181.30148107582	0
h12	0	17660.44	6227.22634903249	0
h13	0	17989.05	7006.29132043325	0
h14	0	18423.85	7549.24762078618	0
h15	0	18345.76	7841.00276013144	0
h16	0	17863.57	7736.86381282707	0
h17	0	17865.31	7151.0047864184	0
h18	0	17896.9	6101.10481562614	0
h19	0	17346.61	5195.8023487891	0
h20	0	17366.73	4618.03351344773	0
h21	0	17777.47	4226.12508458075	0
h22	0	17365.74	3903.93365096751	0
h23	0	17181.65	3624.14230254351	0