
ANÁLISIS DE LA AFECTACIÓN DEL PRECIO *SPOT* DE ENERGÍA ELÉCTRICA EN COLOMBIA DEBIDO AL RETRASO DEL PROYECTO HIDROELÉCTRICO ITUANGO*

Alfredo Trespalacios¹
Luis Guillermo Vélez²
Luis Eduardo Franco³
Felipe Carrasquilla⁴
Yorley Palacios⁵
Edison Giraldo⁶

* DOI: <https://doi.org/10.18601/01245996.v27n52.10>. Esta investigación no recibió ninguna subvención específica de agencias de financiación del sector público, comercial o sin fines de lucro. Recepción: 01-06-2023, modificación final: 23-10-2024, aceptación: 29-10-2024. Sugerencia de citación: Trespalacios, A., Vélez, L. G., Franco, L. E., Carrasquilla, F., Palacios, Y., & Giraldo, E. (2024). Análisis de la afectación del precio spot de energía eléctrica en Colombia debido al retraso del proyecto hidroeléctrico Ituango. *Revista de Economía Institucional*, 27(52), 285-317.

¹ Doctor en Economía, Magíster en Finanzas, Ingeniero Electricista. Director ELAE

² Doctor en Economía, Economista. Investigador Fundación ECSIM

³ Magíster en Finanzas, Ingeniero Financiero. Instituto Tecnológico Metropolitano.

⁴ Ingeniero Administrador. Investigador, ELAE

⁵ Especialista en Analítica, Ingeniero Industrial. Investigador Fundación ECSIM

⁶ Ingeniero Industrial. Investigador Fundación ECSIM

Análisis de la afectación del precio spot de energía eléctrica en Colombia debido al retraso del proyecto hidroeléctrico Ituango

Resumen. Este artículo presenta un análisis contrafactual del impacto que el retraso en la ejecución del proyecto hidroeléctrico Ituango ha tenido en el precio *spot* de la energía eléctrica en bolsa en Colombia. Como es el proyecto de generación más grande del país, con una participación esperada del 17% en el mercado, su retraso ha supuesto un desafío para la estabilidad de la oferta energética. Se analiza cómo la falta de este aporte de energía ha afectado la dinámica de precios, considerando los *Objetivos de Desarrollo Sostenible* (ODS7) y los factores que inciden en la oferta y demanda de electricidad.

Palabras clave: Mercado Eléctrico, ODS7, Precio *Spot*, Hidroeléctrica Ituango; JEL: Q41, Q42, L94, O13

Analysis of the Impact on the Spot Price of Electricity in Colombia Due to the Delay of the Ituango Hydroelectric Project

Abstract. This article presents a counterfactual analysis of the impact that the delay in the execution of the Ituango hydroelectric project has had on the spot price of wholesale electricity in Colombia. As the country's largest generation project, with an expected market share of 17%, its delay has posed a challenge to the stability of the energy supply. The analysis examines how the absence of this energy contribution has affected price dynamics, considering the Sustainable Development Goals (SDG7) and the factors influencing electricity supply and demand.

Keywords: Electricity Market, SDG7, Spot Price, Ituango Hydroelectric Project; JEL: Q41, Q42, L94, O13

Análise do Impacto no Preço Spot da Eletricidade na Colômbia Devido ao Atraso do Projeto Hidrelétrico de Ituango

Resumo. Este artigo apresenta uma análise contrafactual do impacto que o atraso na execução do projeto hidrelétrico Ituango teve sobre o preço *spot* da energia elétrica no mercado colombiano. Sendo o maior projeto de geração do país, com uma participação de mercado esperada de 17%, seu atraso representou um desafio para a estabilidade da oferta energética. O estudo examina como a ausência desse aporte de energia afetou a dinâmica dos preços, considerando os *Objetivos de Desenvolvimento Sustentável* (ODS7) e os fatores que influenciam a oferta e demanda de eletricidade.

Palavras-chave: Mercado Elétrico, ODS7, Preço *Spot*, Hidrelétrica Ituango; JEL: Q41, Q42, L94, O13

INTRODUCCIÓN

A nivel mundial, se reconoce la importancia de la disponibilidad y funcionalidad de la infraestructura para el cumplimiento de los Objetivos de Desarrollo Sostenible (ODS) establecidos en 2015 por la Asamblea de las Naciones Unidas (ONU) y previstos para ser alcanzados en 2030. Concretamente, el ODS 9 hace referencia al desarrollo de una infraestructura de calidad, fiable, sostenible y resiliente. Además, los ODS 6, 7 y 11 se refieren a la gestión sostenible del agua y el saneamiento, y a la creación de ciudades inclusivas y resilientes (Jaimurzina y Sánchez, 2017).

El sector eléctrico colombiano fue transformado profundamente con la reforma de las Leyes 142 y 143 de 1994. En estas leyes aparecen elementos clave, como la liberalización del mercado, la introducción de formación de precios competitivos en la actividad de generación, la regulación por incentivos en transmisión, distribución y comercialización, y los avances institucionales en regulación y supervisión de las firmas reguladas. Con el paso del tiempo, se han identificado necesidades de incrementar la competencia, diversificar el portafolio de generación, monetizar los recursos locales embebidos en las redes de distribución y modernizar la arquitectura de mercado y la regulación (Benavides et al., 2018, p. 4).

En Colombia predomina la generación hidráulica, que representa entre el 65 y el 80 por ciento de la generación del país (XM, 2019). Lo anterior implica una matriz energética menos contaminante, pero al mismo tiempo vulnerable al cambio climático (Vega Plata, 2019; Fedesarrollo, 2013). En ese contexto, como plantean Melo et al. (2017), la disminución del recurso hídrico ocasionada por eventos de variabilidad climática, como *El Niño*, puede generar incrementos importantes en el precio de la energía, afectar la competitividad y causar impactos negativos en el Producto Interno Bruto (PIB).

Actualmente, la expansión de la generación en el país depende de la entrada en operación completa de Ituango. La confiabilidad en la operación de este recurso es clave para el abastecimiento energético a mediano y largo plazo en Colombia.

Desde 1962, el ingeniero Lucio Chiquito Caicedo identificó el potencial del Proyecto Hidroeléctrico Pescadero Ituango, con estudios iniciados en 1974, cuya evolución detallada se encuentra en Vélez (2021). Posteriormente, el 30 de enero de 2009, el Ministerio de Ambiente concedió la Licencia Ambiental del proyecto, y se definió su fecha de entrada en operación para octubre de 2018, con ocho

unidades de 300 MW de potencia nominal cada una y una capacidad instalada total de 2.400 MW.

De acuerdo con la ANLA, luego de la evolución en la ejecución técnica y financiera del proyecto, desde el 28 de abril de 2018, cuando ocurrió la contingencia, el proyecto se encuentra en emergencia. La contingencia surgió debido a una obstrucción en el Sistema Auxiliar de Desviación (SAD), lo que llevó a un llenado anticipado del embalse.

En un sistema de energía eléctrica con una potencia instalada cercana a 17,000 MW, es razonable que el retraso en la entrada en operación de un proyecto como Hidroituango, que representa un incremento en la oferta de energía eléctrica nacional equivalente al 17% del consumo actual, afecte el precio del mercado. Hidroituango contará finalmente con 2,400 MW de capacidad instalada, repartida en ocho máquinas de 300 MW.

Los retrasos en los proyectos hidroeléctricos han sido un tema relevante a nivel mundial. Por ejemplo, Bacon y Besant (1998), del Banco Mundial, mostraron que los proyectos hidroeléctricos en todo el mundo incurrieron en un promedio de 27% de sobre costo y un 28% de retraso en el tiempo. De acuerdo con el Encuentro Mundial de Gestión de Proyectos (PIPC, 2005), el 31% de todos los proyectos no se construyen en el tiempo definido.

Según la International Hydropower Association (IHA), en el sector hidroeléctrico, los retrasos suelen ser causados por múltiples factores que se pueden dividir en dos grupos: los riesgos técnicos, que incluyen problemas comerciales y de ingeniería; y los riesgos no técnicos, que abarcan factores ambientales y sociales, problemas comunitarios y desafíos de salud y seguridad. Los retrasos en los proyectos hidroeléctricos generan costos económicos para todas las partes interesadas. Para el inversor, después de haber completado la financiación de la deuda, el retraso genera riesgos financieros, pues podría reducirse el valor presente neto del proyecto debido a los ingresos diferidos y al posible aumento de los costos. Adicionalmente, para los desarrolladores del proyecto, los retrasos pueden causar impactos negativos en la reputación de la empresa, los ingresos, la retención del personal, la contratación y la posibilidad de obtener contratos futuros. Además, el retraso de un proyecto hidroeléctrico limita los beneficios económicos, ambientales y sociales del proyecto, incluidas las medidas de alivio de la pobreza, la mitigación del carbono y la generación y transferencia de rendimientos financieros para el sector público.

La literatura sobre las causas y efectos del retraso en la construcción de hidroeléctricas evalúa las causas y efectos de estos retrasos,

pero no considera los impactos que puedan quedar fuera del marco contractual de la relación entre el propietario, el desarrollador y el contratista (Sun y Meng, 2009).

Bohórquez, López y Marín (2013, p.1) proponen un modelo exponencial para la función de oferta en el pool eléctrico en Colombia, similar al que se utiliza en este documento para el precio *spot*, donde las variables exógenas son los costos, el clima, la demanda, los momentos de intervención por parte del regulador, entre otras. Asimismo, otros trabajos, como los de Santa María et al. (2009), Maya y Gil (2008) o Franco, Velásquez y Olaya (2008), explican las relaciones entre los precios de mercado y sus fundamentos, así como propuestas para su modelación.

Desde otra perspectiva, Henao y Dyner (2020), mediante un modelo de optimización robusto, evalúan los efectos de la incorporación de energía solar y eólica en un sistema eléctrico dominado por la energía hidroeléctrica, que ha estado en riesgo por problemas en proyectos como el megaproyecto Hidroituango. Los autores evidencian que invertir en unidades de energía hidroeléctrica o térmica podría ser más riesgoso y hasta el doble de costoso que invertir en energías renovables. Además, muestran que las tecnologías solar y eólica, en una proporción de 2,5 a 1, forman parte de un plan óptimo y seguro para la expansión del sistema eléctrico colombiano. Adicionalmente, Henao et al. (2020) destacan que los recursos solares y eólicos sirven de complemento al sistema, reduciendo su exposición al riesgo durante las estaciones secas del ciclo climatológico anual y las fases cálida y fría del fenómeno *El Niño*-Oscilación del Sur (ENSO).

Rodríguez, Trespalacios y Galeano (2021) emplean un modelo de Rezago Distribuido Autorregresivo para analizar la transferencia de incertidumbre de las variables del mercado eléctrico colombiano a la Prima de Riesgo Forward (FRP) en contratos de largo plazo, utilizando información sobre demanda eléctrica, precio *spot*, generación de energía a través de diversas tecnologías y el Índice Oceánico del Niño. Los resultados evidencian que el FRP responde a corto plazo a variables como la generación hidroeléctrica, la generación de carbón, la demanda eléctrica y el Índice del Niño Oceánico, aunque la transacción se materialice uno o dos años después de la ocurrencia del evento.

Por otro lado, García y Stacchetti (2011) presentan un modelo dinámico con crecimiento aleatorio de demanda y equilibrios markovianos no colusorios. Los autores logran identificar un precio tope mínimo, no relacionado con el valor de la energía o la carga perdida

(VoLL), que permite a las empresas recuperar sus costos de inversión y producción en equilibrio.

En cuanto a la transición energética y su impacto en la oferta de energía, Pérez y García-Rendon (2021) evalúan el impacto de los recursos renovables no convencionales en el mercado eléctrico mayorista colombiano. Los autores emplean un escenario contrafactual basado en un modelo estructural del comportamiento de una empresa de energía que ofrece 1000 MW en 2018, encontrando que el precio *spot* disminuye en promedio 12,75 COP\$/kWh (4,31 US\$/MWh) por día. Concluyen que 6,75 COP\$/kWh (2,28 US\$/MWh) se debe a las estrategias de las empresas, y 6 COP\$/kWh (2,03 US\$/MWh) es un efecto de orden de mérito. Para quienes deseen profundizar en el sector eléctrico colombiano, Maya, Hernández y Gallego (2012) explican sus particularidades.

No se encontró literatura local relacionada con el impacto en los precios de energía debido al retraso de un proyecto hidroeléctrico, lo que motiva a los autores a realizar esta investigación. En este artículo se presenta un análisis contrafactual que tiene por objetivo conocer el impacto del retraso del proyecto hidroeléctrico Ituango en el precio de bolsa de la energía eléctrica en Colombia. Se utilizan tres modelos para contrastar los datos reales y estimados de la curva de oferta del mercado eléctrico. Para tal fin, se realizan tres aproximaciones a la curva de oferta del mercado eléctrico colombiano: una de tipo lineal con función logarítmica natural, otra con árboles de decisión, y otra usando el modelo XGBoost.

En la siguiente sección de este documento, se hace una descripción de la curva de oferta del mercado eléctrico colombiano. Luego, en la tercera sección, se presenta el análisis econométrico. En la cuarta sección, se expone el análisis realizado a partir de herramientas propias de la inteligencia artificial. En la quinta y sexta sección se presentan, respectivamente, los resultados de la investigación y sus conclusiones.

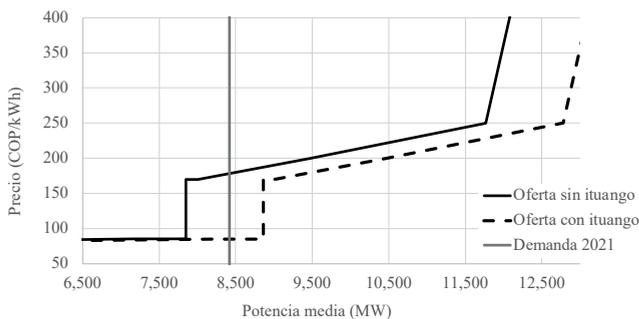
CURVA DE OFERTA DEL MERCADO ELÉCTRICO COLOMBIANO

Una curva de oferta muestra la cantidad que un productor está dispuesto a entregar de un bien en función del precio de mercado, manteniendo constantes los demás factores que afectan la decisión de venta. Una curva de suministro de energía es una serie de funciones escalonadas en las que cada paso representa un grupo o categoría particular de recurso energético (Denholm & Margolis, 2008).

Como plantean Sierra y Castaño (2010) y Villaplana (2003, p. 26), en los mercados eléctricos más liberalizados, los generadores ofertan la cantidad de energía que están dispuestos a entregar y el precio mínimo al que lo harían. En el caso colombiano, la demanda de energía puede asumirse como inelástica en el corto plazo. Esto significa que, ante una demanda fija, las variaciones del precio *spot* de la electricidad dependen exclusivamente de los cambios en la curva de oferta del mercado. La disponibilidad de recursos de bajo costo es, por lo tanto, esencial para la reducción esperada de los precios de la energía eléctrica.

La entrada de Ituango a la bolsa de energía tiene un efecto al aumentar la cantidad de recursos disponibles a bajo costo, lo que se representa como un desplazamiento hacia la derecha de la curva de oferta del sistema, como se ilustra en la Figura 1. En consecuencia, ante un mismo nivel de demanda de energía eléctrica, el precio marginal del sistema se reduce con la entrada en operación del proyecto.

Figura 1.
Curva de oferta del sistema con y sin Ituango



Fuente: Elaboración propia.

Cuadro 1.
Cronograma inicial del Megaproyecto Hidroituango

	Inicial compromisos	Antes de emergencia*
Unidad 1	julio 2018	noviembre 2018
Unidad 2	octubre 2018	febrero 2019
Unidad 3	enero 2019	mayo 2019
Unidad 4	marzo 2019	agosto 2019

Fuente: Elaboración propia.

En 2010, año en que se inició la construcción de Hidroituango, se esperaba que la primera unidad de generación comenzara su operación en julio de 2018 y que el resto de las unidades se integraran una a

una aproximadamente cada tres meses. Sin embargo, hubo varios inconvenientes que retrasaron inicialmente la obra, tales como demoras en la adquisición de tierras y en la reubicación de las comunidades locales, problemas en la construcción de túneles y otros trabajos de infraestructura, litigios y disputas legales relacionados con la construcción del proyecto y sus impactos en las comunidades locales y el medio ambiente, así como retrasos en la aprobación de permisos y licencias ambientales necesarios.

Estos problemas generaron un retraso de aproximadamente 20 meses en el cronograma del proyecto. Sin embargo, a finales de 2015 se implementó un plan de aceleración para recuperar parte del retraso, y se esperaba que la primera unidad entrara en operación en noviembre de 2018. La situación más crítica se presentó en mayo de 2018, cuando una falla en uno de los túneles de desvío causó la inundación de la zona cercana a la presa y la evacuación de miles de personas. Esta emergencia obligó a tomar la decisión de inundar el cuarto de máquinas y retrasó aún más la entrada en operación de las unidades.

Cuadro 2.

Cronograma últimos acontecimientos del Megaproyecto Hidroituango

Fecha	Evento
7 de noviembre de 2022	El Ministerio de Minas y Energía solicitó formalmente a EPM que entregara todos los análisis hechos para verificar las obras.
11 de noviembre de 2022	El alcalde de Medellín declaró que la operación iniciaría el 15 de noviembre
26 de noviembre de 2022	La turbina 1 pasó la prueba de entrega de energía a XM
29 de noviembre de 2022	La turbina 2 pasó la prueba de entrega de energía a XM
30 de noviembre de 2022	Las turbinas 1 y 2 entran en operación comercial en modalidad de prueba
14 de diciembre de 2022	Última prueba exitosa de las turbinas 1 y 2 y entrada en operación comercial de la turbina 1
23 de enero de 2023	Entrada en operación de la turbina 2
30 de noviembre de 2023	Se espera que las turbinas 3 y 4 entren en operación para esta fecha

Fuente: Elaboración propia.

El 7 de noviembre de 2022, el Ministerio de Minas y Energía informó que se había solicitado oficialmente a EPM la entrega de todos los análisis realizados hasta el momento para verificar las obras. Se esperaba que el 15 de noviembre se iniciara la operación comercial, siempre y cuando se superaran todas las pruebas exitosamente. Sin embargo, las turbinas 1 y 2 pasaron las pruebas de envío de energía

los días 26 y 29 de noviembre, respectivamente, para comenzar la operación comercial finalmente el 30 de noviembre en modalidad de pruebas. El 14 de diciembre se realizaron otras pruebas en las unidades 1 y 2, las cuales fueron exitosas, lo que permitió iniciar la operación comercial plena. Respecto a las unidades 3 y 4 del proyecto, se espera que entren en funcionamiento antes del 30 de noviembre de 2023.

ANÁLISIS ECONOMETRICO

APROXIMACIÓN DE LA CURVA DE OFERTA

Si se considera que el costo de oportunidad del sistema lo marcan tecnologías no hidráulicas, se puede hacer una relación exponencial para el precio de bolsa en función de la generación no hidráulica GNh y algunos otros factores determinantes. Si P es el precio *spot* o precio de bolsa de energía. Considerando factores adicionales, se puede extender el modelo así:

$$P = A \cdot e^{\alpha \cdot GNh + \sum \beta_i \cdot S_i + \sum \gamma_i \cdot f_i} \quad (1)$$

Donde A , α , β_i y γ_i son estimadas a partir de los datos originales. S_i corresponden con exógenas que sí se modifican con la generación de Ituango y f_i son variables de mercado que si bien modifican el precio, se mantienen constantes con la generación de Ituango. Otra forma de escribir esta representación es tomando el logaritmo natural del precio como en (2). Los parámetros de la ecuación (2) pueden ser estimados a partir de una regresión lineal múltiple, donde el logaritmo del precio *spot* es la variable endógena y las variables exógenas son la generación no hidráulica y los factores S_i y f_i necesarios.

$$\ln(P) = \beta_0 + \alpha \cdot GNh + \sum \beta_i \cdot S_i + \sum \gamma_i \cdot f_i \quad (2)$$

RELACIÓN PRECIO

Consideremos dos casos hipotéticos para un mismo periodo de tiempo. Uno de los casos considera que Ituango entra en operación según el cronograma inicial, y el segundo caso en donde Ituango no entra en operación. Si se define la relación precio -RP como la relación que existe entre el precio *spot* con Ituango P^I y el precio *spot* sin Ituango P^N para el mismo instante de tiempo, se obtienen las expresiones (3).

$$RP = \frac{P^I}{P^N} \quad P^I = RP \cdot P^N \quad (3)$$

Considerando la aproximación (1) para el precio *spot*, en el anexo de este documento se demuestra que RP tiene una relación exponencial decreciente con la generación estimada de Ituango, como en (4). Esto, siempre que la única diferencia entre los casos analizados sea la generación hidráulica aportada por el proyecto. Cualquier diferencia adicional que pudiese producir entre los escenarios debería tratarse de otra forma, como también se muestra en el anexo 1.

$$RP = e^{-\alpha \cdot Gh_{ituango}} \quad (4)$$

Donde $Gh_{ituango}$ es la generación de Ituango. Lo interesante de esta aproximación radica en que basta conocer el valor de α para la curva de oferta del mercado eléctrico colombiano a fin de estimar el efecto que tiene la generación del proyecto sobre el precio *spot* de la energía eléctrica.

MODELO ECONOMÉTRICO

El modelo econométrico propuesto es el siguiente:

$$\ln(P) = \beta_0 + \alpha \cdot GNh + \sum \beta_i \cdot S_i + \sum \gamma_i \cdot f_i + \epsilon$$

Donde GNh es la generación de energía eléctrica mensual que proviene de tecnologías no hidráulicas, medida en unidades equivalentes promedio día (GWh-día).

ϵ representa los residuales del modelo econométrico, calculados como la diferencia entre el valor observado y el valor estimado del logaritmo del precio *spot*.

Se utiliza un modelo de regresión lineal, donde la variable dependiente PPP (precio) se transforma aplicando la función logarítmica natural para obtener una distribución más simétrica de los datos y reducir la influencia de valores atípicos. La variable independiente GNh representa la cantidad de energía generada a través de fuentes no hidráulicas y se emplea para predecir el precio de la energía eléctrica.

El objetivo de este modelo es proporcionar una aproximación de cómo la generación de energía no hidráulica y otras variables influyen en el precio *spot* de la energía. La transformación logarítmica ayuda a mitigar la influencia de valores extremos o atípicos que puedan distorsionar los resultados.

El modelo exponencial utilizado en este estudio se justifica porque refleja la relación no lineal entre la generación no hidráulica y el precio *spot*. Pequeños cambios en la oferta de energía pueden provocar aumentos exponenciales en los precios, especialmente en

mercados con oferta limitada. Además, este modelo captura el efecto de la sustitución de fuentes de energía más costosas por energía hidroeléctrica, mostrando cómo la entrada de Hidroituango podría reducir significativamente el precio de la electricidad. Este enfoque es coherente con la teoría económica, que sugiere que, en mercados con demanda inelástica, los cambios en la oferta generan variaciones significativas en los precios.

Se espera que este modelo permita identificar relaciones lineales entre variables como la generación de energía térmica, solar y eólica, y el precio *spot* de la energía, destacando la relación directa que el retraso en la generación hidroeléctrica tiene sobre el mercado energético. Se prevé que el precio *spot* aumente de forma lineal con la disminución en la participación de la generación hidroeléctrica.

RESULTADOS MODELO ECONOMETRICO

Para el entrenamiento del modelo se tomaron los precios mensuales de la energía en Colombia y la generación de energía no eléctrica (suma de generación eólica, solar, térmica y por cogenerador) entre enero de 2000 y julio de 2018.

Cuadro 3.

Resultados del modelo de regresión lineal con función logarítmica natural

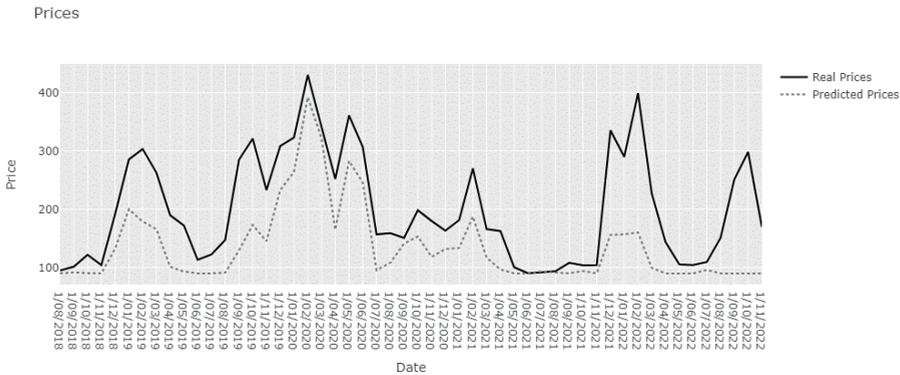
	Std. Err.	t	P > t	[0,025	0,975]
$\alpha = 10,99 \times 10^{-4}$	$4,767 \times 10^{-5}$	23,06	$9,75 \times 10^{-61}$	$10,05 \times 10^{-4}$	$11,93 \times 10^{-4}$
$\beta_0 = 3,42$	0,06	62,08	$7,47 \times 10^{-142}$	3,311	3,5281

Fuente: Elaboración propia.

Para la elaboración del modelo, se asume que existe una relación lineal entre $\ln(P)$ y GNb , lo que significa que el cambio en $\ln(P)$ es proporcional al cambio en GNb . El coeficiente α representa la pendiente de la línea de regresión y muestra cuánto cambia $\ln(P)$ por cada unidad de cambio en GHN . El coeficiente β_0 representa la intersección de la línea de regresión con el eje y, es decir, el valor de $\ln(P)$ cuando x es igual a cero.

Teniendo en cuenta el Costo Equivalente Real en Energía del Cargo por Capacidad Real (CERE) establecido por la CREG, se restringe la variable dependiente en el modeloparea que tenga un precio mínimo de 90.

Figura 2.
Comparación de precios reales de la energía comparada con la predicción hecha co



Fuente: Elaboración propia.

Cuadro 4.
Indicadores de errores obtenidos con regresión lineal con función logarítmica natural

MAE	MSE	RMSE	r ²	MAPE
64,66	7193,57	84,81	0,14	28,70

Fuente: Elaboración propia.

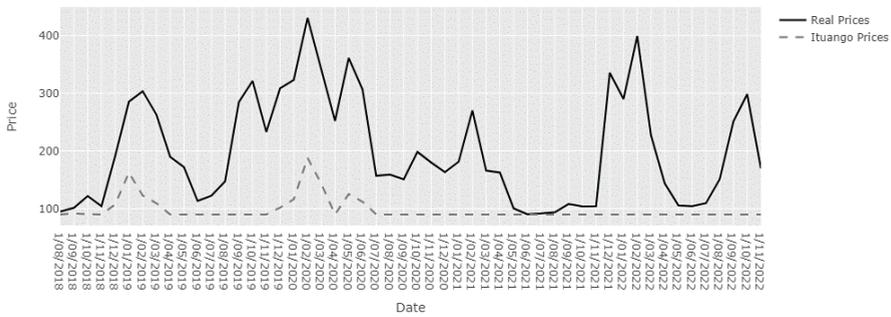
Se observa que el modelo se ajusta relativamente bien, aunque presenta diferencias más notables en los últimos dos picos del precio; sin embargo, los valores de error son aceptables. Al aplicar el modelo a los datos que incluirían la energía generada por Hidroituango, se obtienen los siguientes resultados.

Se observa una diferencia muy notable entre los precios reales y los que el modelo predice cuando se considera la energía que se habría generado con Hidroituango en operación desde noviembre de 2018, como estaba previsto. Esta diferencia se explica porque el modelo asume que toda la energía que Hidroituango hubiera generado reemplazaría la energía que realmente se produjo mediante fuentes no hídricas, y que, además, el precio de la energía se determina en función de la generación no hidráulica.

De acuerdo con los hallazgos, se puede graficar la relación de precios entre las situaciones “con Ituango” y “sin Ituango”. La siguiente gráfica se puede interpretar de la siguiente manera: en caso de que la generación de Ituango hubiese sido, en promedio, 6 GWh-día, el precio observado, habría sido, en promedio, un 87% del precio real.

Figura 3.

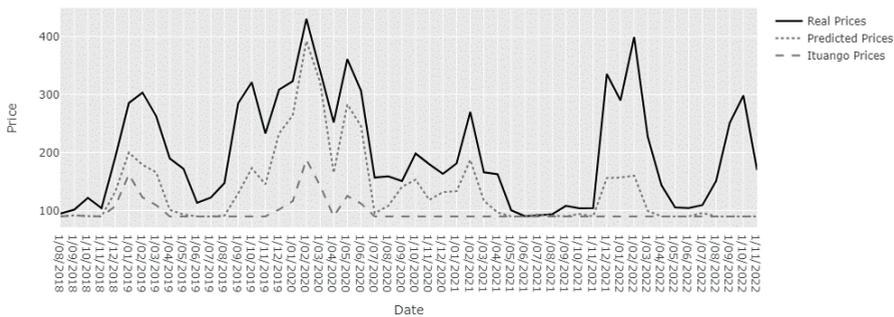
Comparación de precios reales de la energía comparada con la predicción obtenida mediante el modelo de regresión lineal con función logarítmica natural, teniendo en cuenta la generación de energía de Hidroituango



Fuente: Elaboración propia

Figura 4.

Comparación de precios reales de la energía con la predicción obtenida con el modelo de regresión lineal con función logarítmica natural, considerando la generación de energía de Hidroituango y sin ella

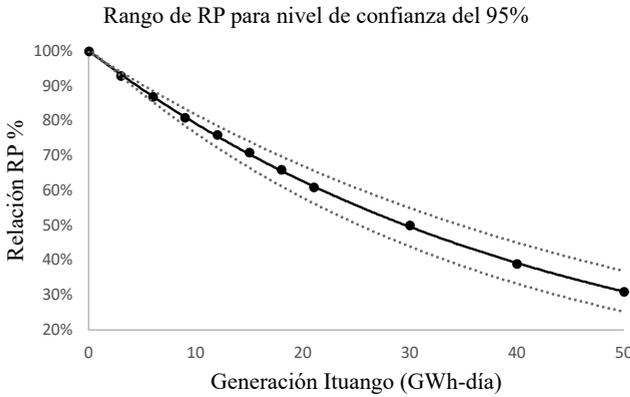


Fuente: Elaboración propia

Así, si para un periodo específico del periodo analizado el precio *spot* de electricidad en Colombia hubiese sido de 200 COP/kWh, el precio en caso de que Ituango hubiese entrado sería el 87% de 200 COP/kWh, es decir 174 COP/kWh. Asimismo, se muestra que si la generación hubiese sido de 15GWh-día, el precio habría estado en 71% del precio real.

Figura 5.
Curva relación precio - RP en función de la generación promedio de Ituango

$$Pb^{CI} = RP \cdot PbPB^{SI}$$



Curva relación precio - RP en función de la generación promedio de Ituango											
Generación Ituango promedio mes (GWh-día)	0	3	6	9	12	15	18	21	30	40	50
RP promedio	100%	93%	87%	81%	76%	71%	66%	61%	50%	39%	31%
RP mínimo (2.5%)	100%	94%	89%	83%	79%	74%	70%	66%	55%	45%	37%
RP máximo (97.5%)	100%	92%	85%	79%	73%	68%	62%	56%	45%	33%	25%

Fuente: Elaboración propia.

ÁRBOLES DE DECISIÓN

Los árboles de decisión (DT, por sus siglas en inglés) son un método de aprendizaje supervisado no paramétrico, utilizado para resolver problemas de clasificación y regresión. Este método crea un modelo que predice el valor de la variable exógena a partir de la inferencia de reglas de decisión extraídas de los datos suministrados. Puede entenderse como una aproximación constante por partes (Pedregosa et al., 2011).

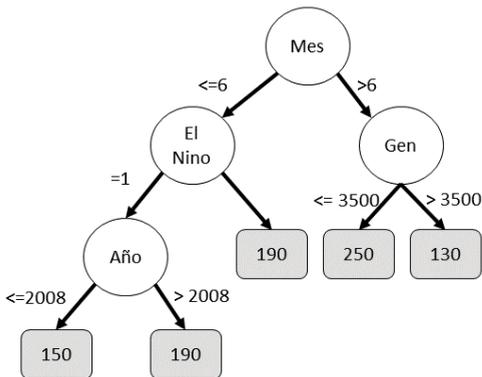
Están compuestos por nodos que en su conjunto forman un árbol. El primero de estos se llama nodo raíz, mientras que todos los demás nodos tienen exactamente un vértice entrante.

Un nodo con vértices salientes se denomina nodo interno, y los nodos finales se llaman hojas (Rokach, 2005).

En un árbol de decisión, cada nodo interno divide el espacio de instancias en dos subespacios según una determinada función discreta de los valores de los datos de entrada. Habitualmente, para esta división se considera solo uno de los atributos de los datos, definiendo límites de acuerdo con su rango de valores (Rokach, 2005). En la siguiente ilustración se presenta un ejemplo del resultado de

un modelo de árboles de decisión para predecir el precio de bolsa, el cual se lee desde el nodo raíz en la parte superior hacia las hojas (rectángulos) en la parte inferior, siguiendo los caminos establecidos por las condiciones en cada nodo.

Ilustración 1.
Ejemplo de árbol de decisiones



Fuente: Elaboración propia

Este modelo permite analizar relaciones no lineales y complejas entre variables. Los árboles de decisión pueden dividir el conjunto de datos en diferentes escenarios o “ramas,” lo que facilita el análisis de diversas combinaciones de factores que afectan el precio *spot*. A través de múltiples divisiones, se busca identificar patrones clave en los datos, considerando la interacción entre múltiples variables.

Se espera que este modelo identifique escenarios específicos en los que el precio *spot* varía significativamente. Por ejemplo, podría encontrar que bajo condiciones de bajos niveles de embalses, junto con alta demanda y bajas reservas de generación térmica, el precio *spot* tiende a dispararse. Este modelo debe proporcionar información más específica para gestionar escenarios críticos.

SOFTWARE Y DATOS

Para la realización de este trabajo, se utilizaron los datos de precio de bolsa mensual de energía eléctrica (en pesos por kilovatio hora - kWh) del mercado colombiano, de la generación de energía total del sistema (en GWh) desagregada por tipo de fuente de generación (hidráulica, térmica, eólica, solar, entre otras), y del total de la demanda

energética. Todos estos datos fueron publicados en el Portal BI del Operador y Administrador del Sistema Eléctrico Colombiano (XM) entre el año 2000 y 2022. Por último, también se incluyeron los datos meteorológicos de los fenómenos de *El Niño* y *La Niña*.

Para la implementación de los modelos econométricos y de *machine learning*, se utilizó Python como lenguaje de programación, empleando las bibliotecas *statsmodels* para el modelo de regresión lineal, *scikit-learn* para los árboles de decisión, y *XGBoost* para el modelo de *boosting*. Adicionalmente, se utilizó *GridSearchCV* para la optimización de los hiperparámetros y mejorar el rendimiento de los modelos.

PARAMETRIZACIÓN DEL MODELO

Para la parametrización del modelo, se utilizó *GridSearchCV*, un proceso de ajuste de parámetros que permite determinar los valores que brindan un mejor desempeño. Dado que inicialmente no se tiene una idea clara de cuáles parámetros son los óptimos para el problema planteado, *GridSearchCV* prueba una combinación de distintos valores para cada parámetro hasta obtener los mejores resultados posibles, utilizando por defecto el coeficiente de determinación R^2 como métrica de desempeño.

Los parámetros analizados con la función *GridSearchCV* del paquete *scikit-learn* de Python fueron los siguientes:

Max depth: profundidad máxima del árbol de decisión

Max features: el número de características a considerar al buscar la mejor división en los nodos internos

Max leaf nodes: crea un árbol con el número máximo de nodos hoja (*max_leaf_nodes*) en modo *best-first*. Los mejores nodos se definen como aquellos que generan la mayor reducción relativa de la impureza.

Min samples leaf: establece el número mínimo de muestras necesarias en un nodo hoja.

Min impurity decrease: un nodo se dividirá si esta división produce una disminución de la impureza mayor o igual a este valor.

Min samples split: el número mínimo de muestras necesarias para dividir un nodo interno.

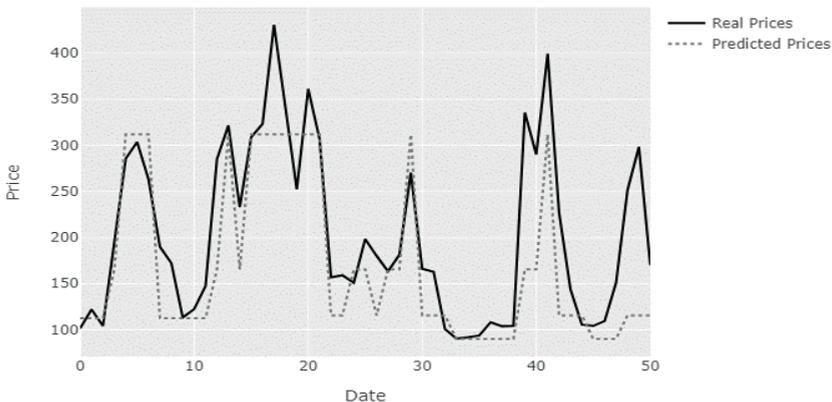
Para el entrenamiento del modelo se usaron los datos recogidos entre enero de 2000 y julio de 2018.

RESULTADOS DEL ÁRBOL DE DECISIÓN

En la siguiente gráfica se muestra la proyección del modelo frente a los datos reales desde noviembre de 2018 hasta noviembre de 2022. En

este caso, el objetivo es evaluar la capacidad del modelo para predecir los datos reales, los cuales se utilizan como referencia. Posteriormente, se realiza una nueva predicción considerando la variación en el precio de bolsa debido a la energía adicional que se habría generado si el proyecto Hidroituango hubiera entrado en funcionamiento en su momento.

Figura 6.
Comparación entre el precio real y el resultado del modelo de árbol de decisión



Fuente: Elaboración propia.

Al igual que en el modelo de regresión lineal, se considera el Costo Equivalente Real en Energía del Cargo por Capacidad Real (CERE) establecido por la CREG y se restringe la variable dependiente para que tenga un precio mínimo de 90.

Se observa que el modelo se ajusta adecuadamente a los valores reales, aunque presenta un mayor desfase en los periodos de precios altos. Usando métricas de error, como el error absoluto medio (MAE, por sus siglas en inglés) o el coeficiente de determinación (r^2), se confirma el buen desempeño del modelo.

Cuadro 5.
Indicadores de error obtenidos con árboles de decisión

<i>MAE</i>	<i>MSE</i>	<i>RMSE</i>	r^2	<i>MAPE</i>
43,35	3843,22	61,99	0,54	19,79

Fuente: Elaboración propia.

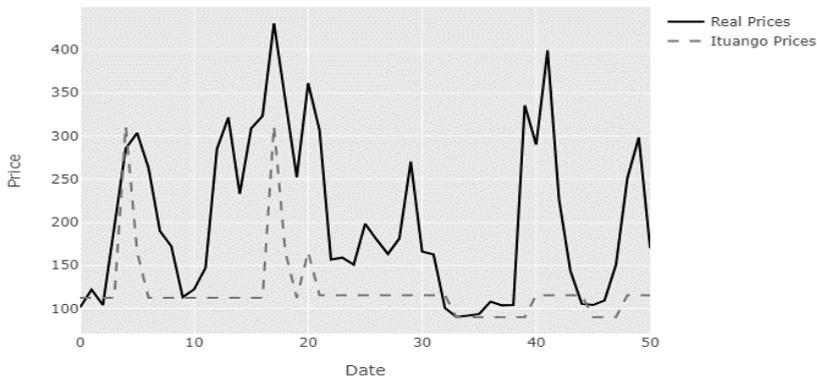
El modelo de árbol de decisión presenta resultados significativamente superiores a los obtenidos por el modelo de regresión lineal, en parte

porque considera una mayor cantidad de variables independientes que explican el precio de la energía.

Los resultados de la estimación del precio de bolsa de energía, considerando la entrada de Hidroituango en comparación con el precio real observado, se muestran en la figura 7. Se observa que durante los primeros meses el precio se mantiene sin una variación muy alta; sin embargo, posteriormente, el precio estimado llega a ser mucho más bajo que los precios reales, logrando ser en promedio un 42% más bajo durante el período analizado. Además, al comparar la predicción sin tener en cuenta Hidroituango con la predicción que incluye la generación que este proyecto hubiera aportado, se encuentra que el precio habría sido en promedio un 30% más bajo.

Figura 7.

Comparación del precio real con el resultado del modelo de árbol de decisión considerando la generación de Hidroituango



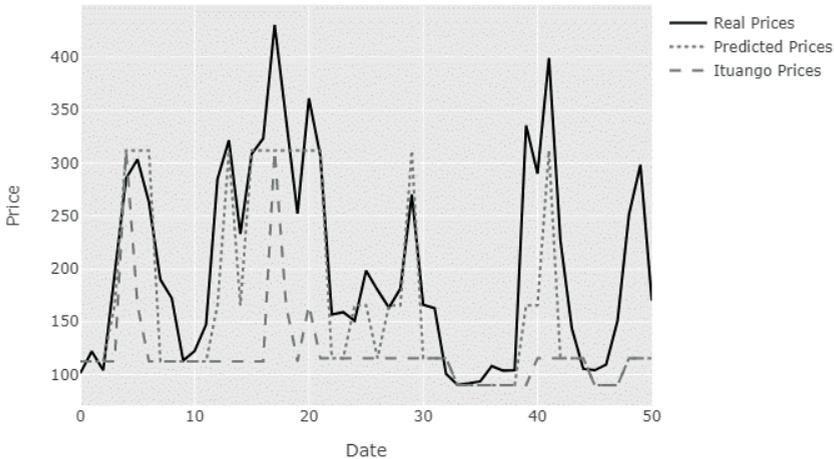
Fuente: Elaboración propia.

En la siguiente gráfica se puede apreciar mejor la diferencia entre la predicción del modelo que incluye los datos de generación de energía de Hidroituango, la predicción sin los datos de Hidroituango y los precios reales observados.

Usando el método de árbol de decisión, también se observa una diferencia considerable entre los precios reales de la energía y los precios que el modelo habría predicho si Hidroituango hubiera comenzado sus operaciones en la fecha inicialmente prevista.

Figura 8.

Comparación de los precios reales de la energía con la predicción del modelo de árbol de decisión considerando la generación de energía de Hidroituango y sin ella



Fuente: Elaboración propia.

EXTREME GRADIENT BOOSTING - XGBOOST

El XGBoost es una técnica avanzada de aprendizaje automático para clasificación y regresión que combina una gran cantidad de árboles de decisión individuales para formar un modelo más robusto (Chen y Guestrin, 2016).

XGBoost utiliza un algoritmo de optimización para mejorar la predicción de cada árbol de decisión que lo compone y aplica técnicas de regularización para reducir el sobreajuste en el modelo. Además, es una herramienta altamente eficiente en términos de tiempo de ejecución, lo que la hace ideal para trabajar con grandes conjuntos de datos (Chen et al., 2016).

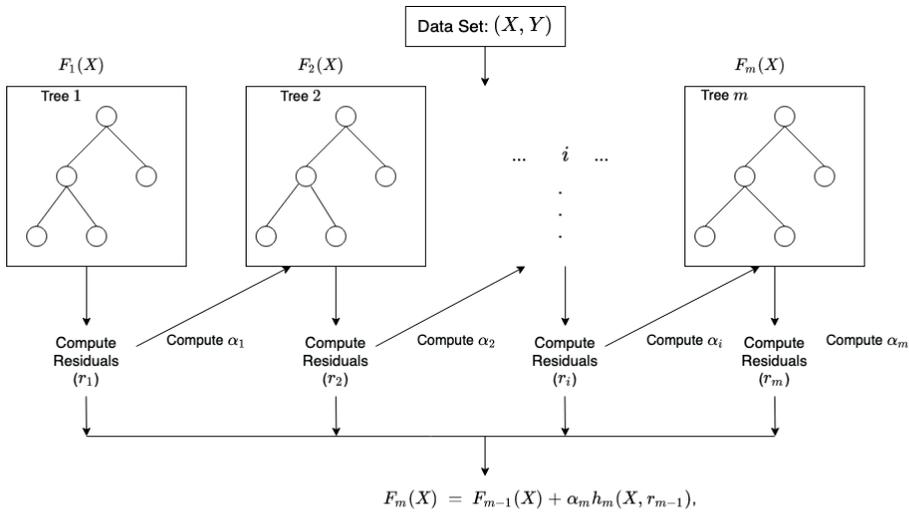
En términos de su uso y aplicaciones, XGBoost ha demostrado ser una herramienta efectiva en una amplia gama de industrias, incluyendo la banca, el marketing digital y, en general, cualquier sector que pueda beneficiarse de la ciencia de datos (Chen et al., 2020).

XGBoost consiste en una secuencia de árboles de decisión, donde cada árbol corrige los errores del anterior. El proceso de construcción de cada árbol en la secuencia es iterativo, ajustando en cada iteración la distribución de pesos de los datos de entrenamiento para dar más

peso a los puntos de datos mal clasificados en la iteración anterior. Este proceso se conoce como “boosting”.

La principal diferencia con el modelo de árbol de decisión es que XGBoost comienza con un árbol de decisión entrenado con una submuestra de los datos de entrenamiento; posteriormente, calcula el error residual producido por este primer árbol. Ese error residual se usa para asignarle un peso a ese primer árbol y para construir un segundo árbol de decisión orientado a reducir el error residual. Este proceso se repite de forma iterativa hasta que se obtiene un conjunto de árboles de decisión, cada uno con un peso asignado inversamente proporcional a su error residual. Finalmente, se combinan los resultados de todos los árboles para obtener la predicción final.

Ilustración 2.
Ejemplo de modelo XGBoost



Fuente: Amazon Web Services. (s.f.)

Se utiliza la librería XGBoost de gradient boosting y se seleccionan los siguientes parámetros para realizar la optimización:

n_estimators: e

representa el número de árboles que se van a utilizar en el modelo.

max_depth: Representa la profundidad máxima que puede tener cada árbol individual.

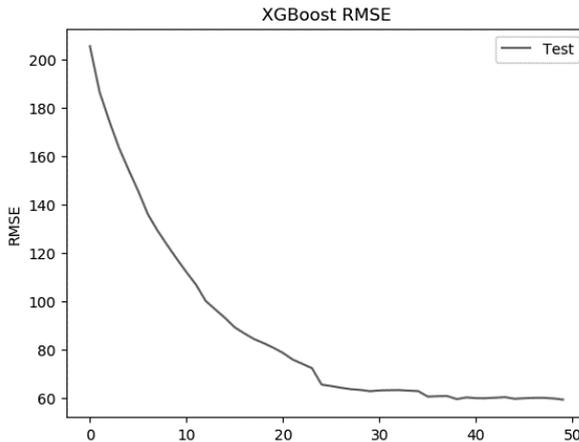
eta: representa la tasa de aprendizaje, que controla cuán rápido se actualizan los pesos en el modelo.

`subsample`: representa la fracción de datos de entrenamiento que se utilizarán para entrenar cada árbol individual.

`colsample_bytree`: representa la fracción de características que se utilizarán para entrenar cada árbol individual. Un valor más alto significa un modelo más robusto, pero también más lento en el entrenamiento.

Ilustración 3.

Secuencia de mejora de cada árbol de decisión que compone el modelo XGBoost



Fuente: Elaboración propia.

Este es un modelo más avanzado de aprendizaje automático que combina múltiples árboles de decisión para reducir los errores de predicción. Su principal ventaja radica en la mejora continua del ajuste del modelo, gracias a su capacidad para corregir iterativamente los errores de los árboles anteriores.

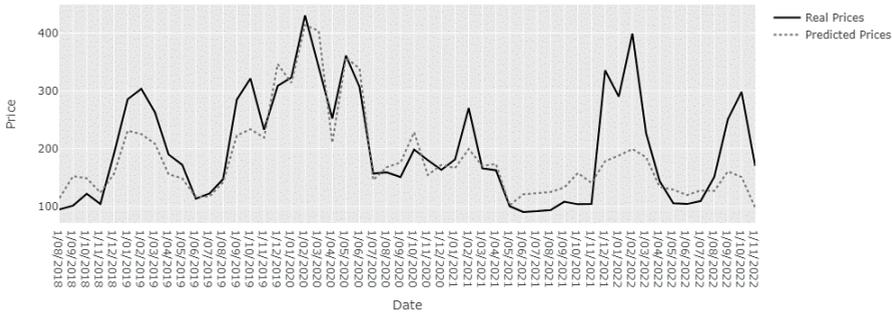
Se espera que el modelo XGBoost ofrezca la predicción más precisa de los tres modelos, dado su enfoque en mejorar continuamente los resultados mediante un proceso iterativo. Al ajustar los hiperparámetros, se busca minimizar el error de predicción y capturar mejor las fluctuaciones del precio *spot* bajo escenarios complejos, como cambios en la oferta y demanda energética a corto plazo.

RESULTADOS XGBOOST

Se presenta la proyección del modelo contra los datos reales desde agosto de 2018 hasta noviembre de 2022. Al igual que en los casos

anteriores, se utiliza como referencia la capacidad del modelo para predecir los datos históricos disponibles y, posteriormente, realizar la predicción considerando la variación del precio de bolsa debido a la energía hidráulica adicional que se habría generado si el proyecto Hidroituango hubiera entrado en funcionamiento en la fecha originalmente prevista.

Figura 9. Comparación del precio real con el resultado del modelo de XGBoost



Fuente: Elaboración propia.

Se observa que el modelo se ajusta adecuadamente a los valores reales. Utilizando métricas de error como el error absoluto medio (MAE, por sus siglas en inglés) y el coeficiente de determinación (R^2), se confirma el buen desempeño del modelo.

Cuadro 6. Indicadores de errores obtenidos con árboles de decisión

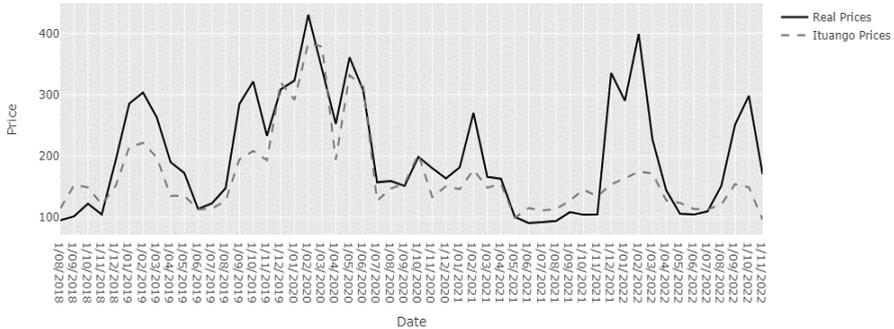
<i>MAE</i>	<i>MSE</i>	<i>RMSE</i>	r^2	<i>MAPE</i>
40,40	3248,18	56,99	0,61	19,87

Fuente: Elaboración propia.

Cuando se aplica el mismo modelo de regresión a los datos, incluyendo la generación energética que Hidroituango habría ofertado si hubiera comenzado su operación comercial en noviembre de 2018, el modelo se comporta como se muestra en la figura 10.

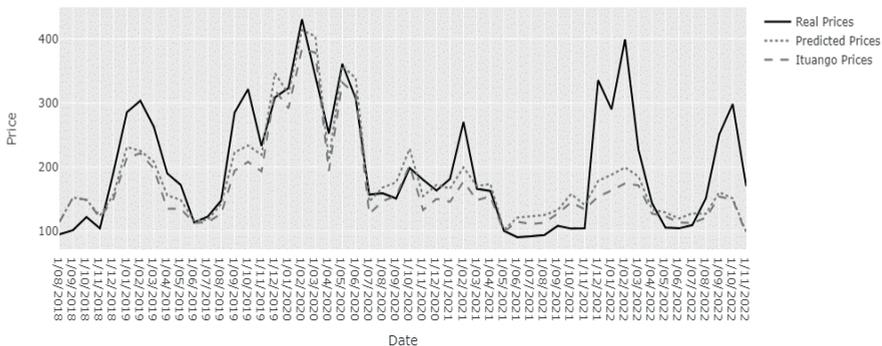
Al comparar el precio real con las predicciones del modelo, tanto con Ituango como sin Ituango, la gráfica muestra que el modelo XG-Boost también indica que el precio de la energía habría sido más bajo si Hidroituango hubiera iniciado sus operaciones en las fechas originalmente previstas en el cronograma, y no como finalmente ocurrió.

Figura 10.
Comparación del precio real con el resultado del modelo XGBoost teniendo en cuenta la generación de Hidroituango



Fuente: Elaboración propia.

Figura 11.
Comparación de precios reales de la energía con la predicción hecha con el modelo XGBoost, teniendo en cuenta la generación de energía de Hidroituango y sin ella.



Fuente: Elaboración propia.

RESULTADOS GENERALES

Para la evaluación del desempeño de los modelos se utilizaron los siguientes indicadores: MAE (Error absoluto medio), MSE (Error cuadrático medio), RMSE (Raíz del error cuadrático medio), (coeficiente de determinación) y MAPE (Error porcentual absoluto medio). Estos indicadores de evaluación de modelos son una parte fundamental del análisis de datos y del aprendizaje automático, ya que permiten medir la calidad de las predicciones realizadas por un

modelo, evaluando su capacidad para hacer predicciones precisas y confiables.

MAE (error absoluto medio): el MAE mide el promedio de las diferencias absolutas entre las predicciones del modelo y los valores reales, es decir, el promedio de la magnitud de los errores absolutos. Este indicador refleja el error promedio en la escala original de la variable de respuesta y se expresa en las mismas unidades que dicha variable. Un MAE más bajo indica un mejor desempeño del modelo.

MSE (error cuadrático medio): el MSE calcula el promedio de las diferencias cuadradas entre las predicciones del modelo y los valores reales, es decir, el promedio de los errores al cuadrado. El MSE penaliza más los errores grandes que los pequeños y se expresa en las unidades cuadradas de la variable de respuesta. Un MSE más bajo indica un mejor desempeño del modelo.

RMSE (raíz del error cuadrático medio): el RMSE es la raíz cuadrada del MSE. Este indicador mide el error promedio en la misma escala que la variable de respuesta y se expresa en sus mismas unidades. Un RMSE más bajo indica un mejor desempeño del modelo.

r^2 (Coeficiente de determinación): el r^2 mide la proporción de la varianza total en la variable de respuesta que es explicada por el modelo. El r^2 es un valor entre 0 y 1, donde 0 indica que el modelo no explica la variabilidad en los datos y 1 indica que el modelo explica toda la variabilidad. Un r^2 más alto indica un mejor desempeño del modelo.

MAPE (error porcentual absoluto medio): el MAPE mide el error porcentual medio entre las predicciones del modelo y los valores reales, es decir, el promedio de la magnitud de los errores porcentuales. Este indicador expresa el error promedio como un porcentaje de la variable de respuesta. Un MAPE más bajo indica un mejor desempeño del modelo. Es importante destacar que el MAPE puede dar lugar a problemas matemáticos cuando los valores reales son cercanos a cero, por lo que debe ser utilizado con precaución.

ERRORES EN ETAPA DE ENTRENAMIENTO DEL MODELO

Cuadro 7.

Comparación de los indicadores de error en la etapa de entrenamiento de cada modelo

	Modelo lineal	XGBoost	Árbol de decisiones
<i>MAE</i>	0,27	7,27	39,71
<i>MSE</i>	0,11	233,92	8371,13
<i>RMSE</i>	0,34	15,29	91,49
r^2	0,71	0,99	0,49
<i>MAPE</i>	6,10	6,23	24,85

Fuente: Elaboración propia.

En la etapa de entrenamiento, el modelo de regresión lineal con transformación logarítmica es el que mejor se ajusta y presenta los mejores indicadores de error, seguido del modelo XGBoost. Por último, el modelo con peores indicadores es el de árbol de decisiones.

Cuadro 8.

Comparación de los indicadores de error en la etapa de testeo de cada modelo

	Modelo lineal	XGBoost	Árbol de decisiones
<i>MAE</i>	64,66	40,40	43,35
<i>MSE</i>	7193,57	3248,17	3843,22
<i>RMSE</i>	84,81	56,99	61,99
r^2	0,14	0,61	0,54
<i>MAPE</i>	28,70	19,87	19,79

Fuente: Elaboración propia.

En la etapa de prueba, el modelo de regresión lineal, a diferencia de la etapa de entrenamiento, presenta los peores indicadores de error. Los modelos de árboles de decisión y el XGBoost muestran un desempeño similar en términos de indicadores de error, con un rendimiento ligeramente superior en el caso del XGBoost.

IMPACTO SOBRE EL PRECIO SPOT

En el siguiente cuadro se muestra la comparación entre el precio de bolsa ocurrido realmente y el precio de bolsa con la entrada a tiempo de Ituangó para los diferentes modelos (COP/kWh). El precio promedio ocurrido entre noviembre de 2018 y noviembre de 2022 fue de 202,75 COP/kWh, mientras que, para ese mismo período, el modelo de regresión lineal, sin considerar la entrada de Hidroituangó, predijo un precio promedio de 136,28 COP/kWh; el modelo XGBoost, 186,43 COP/kWh; y el modelo de árboles de decisión, 171,27 COP/kWh. El modelo que más se ajustó a los precios reales fue el XGBoost.

Cuadro 9.

Comparación de los precios reales y los estimados con cada uno de los modelos empleados

Año*	Ocurrido	Precio de bolsa (COP/kWh)		
		Modelo lineal	XGBOOST	Árbol de decisión
2018	129,69	103,93	140,15	138,98
2019	228,67	137,97	197,13	204,26
2020	251,76	201,63	256,83	226,01
2021	150,63	108,65	149,07	126,42
2022	204,46	99,14	147,13	130,36
Promedio	202,75	136,28	186,43	171,27

*2018 desde noviembre y 2022 hasta noviembre

Fuente: Elaboración propia.

Suponiendo que Hidroitungo hubiese iniciado su operación comercial en noviembre de 2018, el modelo de regresión lineal predijo un precio promedio para el periodo evaluado de 97,92 COP/kWh, el modelo XGBoost de 170,90 COP/kWh y el modelo de árbol de decisión de 120,49 COP/kWh.

Cuadro 10.

Comparación de los precios estimados con cada uno de los modelos empleados, suponiendo la entrada a tiempo de Hidroitungo

Año*	Ocurrido	Desviación (COP/kWh)		
		Modelo lineal	XGBoost	Árbol de decisión
2018	NA	98,67	135,55	112,51
2019	NA	101,36	180,89	133,51
2020	NA	109,55	233,86	139,65
2021	NA	90	135,16	100,61
2022	NA	90	136,74	108,53
Promedio	NA	97,92	170,90	120,49

*2018 desde noviembre y 2022 hasta noviembre.

Fuente: Elaboración propia.

En el siguiente cuadro se muestra la diferencia entre el precio real de la energía y el precio estimado por cada modelo en el caso de que Hidroitungo hubiera entrado en operación en noviembre de 2018, como estaba previsto inicialmente. La diferencia promedio entre el precio real y el modelo lineal, considerando la entrada de Ituango, es de 108,55 COP/kWh, lo que representa un 53% menos que el precio real. La diferencia promedio entre el precio real y el resultado del modelo XGBoost, incluyendo la generación de Hidroitungo, fue de 35,57 COP/kWh, equivalente al 17% del precio real registrado. En el caso del modelo de árbol de decisión, la diferencia promedio entre el precio real y la predicción considerando la entrada de Ituango fue de 85,98 COP/kWh, es decir, un 42% menos que el precio real. Al tener en cuenta la energía que Hidroitungo habría generado, los tres modelos predicen precios inferiores al precio real observado.

Cuadro 11.

Diferencia de precio real vs. Modelos con Ituango

Año*	Real - Modelo lineal con Ituango	Real - XGBoost con Ituango	Real - Árbol de decisiones con Ituango
2018	49,06	12,18	35,22
2019	127,31	47,77	95,16
2020	142,21	17,90	112,11
2021	60,63	15,47	50,01
2022	114,46	67,72	95,93
Promedio	108,55	35,57	85,98

Fuente: Elaboración propia.

En el siguiente cuadro se muestra la diferencia entre el precio de la energía predicho por cada modelo sin tener en cuenta la energía que se hubiera generado en Hidroituango y el precio estimado por cada modelo en el caso de que Hidroituango hubiera entrado en operación el 11 de noviembre de 2018. La diferencia promedio entre las predicciones hechas con el modelo de regresión lineal, considerando y sin considerar Hidroituango, fue de 40,94 COP/kWh. Con el modelo XGBoost, esta diferencia promedio fue de 15,53 COP/kWh (8%), y con el modelo de árbol de decisiones fue de 51,14 COP/kWh. Los tres modelos indican que el retraso en la entrada en operación de Hidroituango implicó un aumento en el precio spot de la energía.

Cuadro 12.

Diferencia de cada modelo sin Ituango vs. cada modelo con Ituango

Año*	Modelo lineal sin Ituango - Modelo lineal con Ituango	XGBoost sin Ituango - XGBoost con Ituango	Árbol de decisiones sin Ituango - Árbol de decisiones con Ituango
2018	12,44	4,60	26,47
2019	39,55	16,24	70,75
2020	92,08	22,97	86,36
2021	20,79	13,91	26,79
2022	13,84	10,39	22,37
Promedio	40,94	15,53	51,14

Fuente: Elaboración propia.

IMPACTO SOBRE LAS COMPRAS DE ENERGÍA

La tarifa de energía eléctrica que paga el usuario en Colombia se compone de los costos de generación, transmisión, distribución, comercialización y otros aspectos. La componente de generación de energía permite remunerar las compras de energía que realiza el comercializador para abastecer la demanda de sus usuarios, ya sea a través de contratos de largo plazo o en la bolsa de energía. Los niveles de exposición a la bolsa de energía para cada comercializador en el país son diferentes, y sus valores dependen de diversas circunstancias.

Con el objetivo de estimar el ahorro potencial que los usuarios del país habrían obtenido con la entrada a tiempo del proyecto, se puede asumir un porcentaje fijo de compras en bolsa del 20% durante el periodo, lo cual implicaría que el 80% restante de la energía se transa en contratos de largo plazo.

Aunque los movimientos en el precio de la bolsa de energía deberían generar alguna variación en el mercado de contratos, en esta

sección se asumirá que el precio de los contratos permanece invariante ante los dos escenarios. Si se considera una demanda promedio nacional de 70 TWh de energía en el periodo y una exposición promedio a bolsa del 20%, el potencial de ahorro para los usuarios con la entrada en operación del proyecto se estima en 6,2 billones de COP según la aproximación lineal, 4,9 billones de COP con la aproximación mediante árboles de decisión y 2 billones de COP según el modelo XGBoost.

Cuadro 13.

Diferencia de cada modelo sin Ituango vs. cada modelo con Ituango

Referente	Precio de bolsa (COP/kWh)	Compras en bolsa (Mill COP)	Diferencia con ocurrido (Mill COP)
Real	206	2,890,603	
Regresión Lineal	98	1,370,940	-6,205,290
Árbol de decisión	120	1,686,889	-4,915,165
XGBoost	171	2,392,637	-2,033,361
Promedio modelos		1,816,822	-4,384,605

Fuente: Elaboración propia.

CONCLUSIONES

En este trabajo se realiza un análisis contrafactual para estimar el impacto sobre el precio *spot* (de bolsa) de la energía eléctrica en Colombia debido al retraso del proyecto de generación de 2400 MW Pescadero Ituango. Para ello, se efectúan tres aproximaciones a la curva de oferta del mercado eléctrico colombiano: regresión lineal con función logarítmica natural, árboles de decisión y XGBoost, para el periodo comprendido entre noviembre de 2018 y noviembre de 2022. Los resultados indican que el precio *spot* de la energía eléctrica en Colombia habría tenido una reducción significativa si el proyecto hubiera entrado en operación según el cronograma inicialmente pactado; en promedio, el precio *spot* de la energía podría haber sido un 36% inferior al precio observado, lo cual habría representado un ahorro potencial para el país de 4,4 billones de pesos en dicho periodo. El retraso ocasionado por la contingencia de 2018 en la Central de Ituango ya ha afectado al usuario final, al impedirle acceder a un precio más bajo de la energía.

Los tres modelos empleados coinciden en que el precio *spot* de la energía habría sido inferior si el proyecto Hidroituango hubiera iniciado operaciones en noviembre de 2018. El modelo XGBoost, que presenta los mejores indicadores de error en la etapa de prueba,

sugiere que el precio habría sido un 17% inferior al registrado si la hidroeléctrica hubiese cumplido con el cronograma inicial. El modelo de árbol de decisiones arrojó indicadores de error muy similares a los del modelo XGBoost y muestra una diferencia de un 42% en el escenario que incluye a Hidroituango. Por su parte, el modelo de regresión lineal con función logarítmica presentó indicadores de error significativamente más altos que los otros dos modelos, posiblemente porque solo considera la generación no hidráulica, a diferencia de los otros modelos, que incluyen muchas más variables. Este modelo mostró una diferencia promedio entre escenarios del 53%.

Es importante destacar que el precio *spot*, o precio de bolsa, no es el mismo que los consumidores observan en sus facturas. Los agentes especializados del mercado, como generadores y comercializadores, utilizan el precio de bolsa para realizar sus transacciones de compra y venta de electricidad. A pesar de esto, el precio de bolsa tiene un gran impacto en los precios de los contratos de energía a largo plazo y en las licitaciones de los mercados regulados. Suponiendo una demanda promedio nacional de 70 TWh de energía en el periodo y una exposición promedio a bolsa del 20%, el potencial de ahorro para los usuarios, en caso de haber entrado en operación el proyecto, se estima en 6,2 billones de pesos según la aproximación lineal, 4,9 billones con la aproximación mediante árboles de decisión y 2 billones según el XGBoost. Se sugiere que futuros estudios analicen el impacto económico que ha tenido el aumento en los precios del mercado *spot* en Colombia sobre la tarifa de los usuarios finales, los costos de producción en el país y el incremento en los gastos gubernamentales en operaciones y subsidios debido al retraso de esta central hidroeléctrica. Asimismo, será necesario identificar los impactos negativos en el producto interno bruto.

El aumento del riesgo de desabastecimiento por el retraso de Ituango también ha estado acompañado por un alza en el precio de referencia de la energía eléctrica para todo el país. Esta situación amerita una revisión del mecanismo de expansión en generación del sistema. El actual mecanismo de subastas de cargo por confiabilidad asume como determinística la entrada en operación comprometida por cada empresa participante en la subasta, aunque la experiencia ha demostrado que, incluso con los incentivos establecidos, los proyectos suelen enfrentar retrasos importantes que afectan económica y energéticamente a la sociedad. Los proyectos de transmisión, por ejemplo, también están experimentando este fenómeno, ya que su construcción toma cada vez más años. Se invita al Ministerio de Minas

y Energía y a los organismos de planeación y regulación a explorar nuevas formas de ejecutar la expansión del sistema interconectado nacional, de manera que se garantice la seguridad energética y se promueva la energía asequible.

REFERENCIAS

- Amazon Web Services. (s.f.). *XGBoost: How It Works*. Recuperado de <https://docs.aws.amazon.com/sagemaker/latest/dg/xgboost-HowIt-Works.html>
- Autoridad Nacional de Licencias Ambientales (ANLA). (2022). *Proyecto Hidroeléctrico Pescadero Ituango*. Recuperado de <https://www.anla.gov.co/proyectos/proyecto-de-interes-en-seguimiento-hidroelectrico-pescadero-ituango#:~:text=El%20proyecto%20Hidroel%C3%A9ctrico%20Pescadero%20Ituango,la%20vertical%20de%20800%20m.>
- Bacon, R. W., & Besant-Jones, J. E. (1998). Estimating construction costs and schedules: experience with power generation projects in developing countries. *Energy Policy*, 26(4), 317-333.
- Benavides, J., Cadena, Á., González, J. J., Hidalgo, C., y Piñeros, A. (2018). *Mercado eléctrico en Colombia: Transición hacia una arquitectura descentralizada*. FEDESARROLLO.
- Bohórquez, J., López, A., y Marín, J. (2013). Un modelo exponencial para la función de oferta en el pool eléctrico colombiano. *Revista de Métodos Cuantitativos para la Economía y la Empresa*, 1(1), 1-15. <https://doi.org/10.46661/revmetodoscuanteconempresa.1234>
- Chen, R. C., Caraka, R. E., Arnita, N. E. G., Pomalingo, S., Rachman, A., Toharudin, T., & Pardamean, B. (2020). An end-to-end scalable tree boosting system. *Sylwan*, 165(1), 1-11.
- Chen, T., & Guestrin, C. (2016, August). XGBoost: A scalable tree boosting system. En *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 785-794).
- Denholm, P., & Margolis, R. (2008). *Supply Curves for Solar PV-Generated Electricity for the United States* (No. NREL/TP-6A0-44073). National Renewable Energy Lab. (NREL), Golden, CO (United States).
- Fedesarrollo. (2013). *Análisis costo beneficio de energías renovables no convencionales en Colombia*.
- Franco, J., Velásquez, J., y Olaya, J. (2008). Análisis de la demanda mensual de electricidad en Colombia: un enfoque de componentes no observables. *Revista de Análisis Económico*, 23(1), 25-40. <https://doi.org/10.1016/j.rae.2008.01.003>
- Henao, A., y Dyner, I. (2020). Evaluación de la inserción de energías renovables en un sistema eléctrico dominado por la hidroeléctrica. *Revista de Energía Renovable*, 15(2), 45-60.
- International Hydropower Association. (2016). *How to address the causes of delay in hydropower projects*. Recuperado de <https://www.hydropower.org/blog/how-to-address-the-causes-of-delay-in-hydropower-projects#:~:text=Causes%20of%20project%20delay,and%20health%20and%20safety%20challenges.>

- Jaimurzina, A., y Sánchez, R. (2017). *Gobernanza de la infraestructura para el desarrollo sostenible en América Latina y el Caribe: una apuesta inicial*. CEPAL.
- Maya, J., y Gil, J. (2008). Dependencia de los precios spot de la energía eléctrica en Colombia. *Revista de Economía Energética*, 12(1), 15-30. <https://doi.org/10.1016/j.reven.2008.01.001>
- Maya, J., Hernández, M., y Gallego, J. (2012). Potencial de recursos energéticos en Colombia: un análisis de fuentes tradicionales y no tradicionales. *Revista de Energía y Recursos Naturales*, 10(1), 5-20. <https://doi.org/10.1016/j.rern.2012.01.001>
- Melo, S., Riveros, L., Romero, G., Álvarez, A., Díaz, C., y Calderón, S. (2017). Efectos económicos de futuras sequías en Colombia: Estimación a partir del Fenómeno El Niño 2015. *Archivos de Economía*, 466, 1-34.
- Pedregosa, F., et al. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
- Pedregosa, F., Varoquaux, G., Gramfort, A., & Thirion, B. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
- Pérez, A., García-Rendon, M. (2021). Impacto de recursos renovables no convencionales en el mercado eléctrico mayorista colombiano. *Revista de Energía y Desarrollo Sostenible*, 18(3), 100-115. <https://doi.org/10.1016/j.reds.2021.04.002>
- PIPC. (2005). *Global Project Management Survey*. PIPC Global Holding, London, UK.
- Plummer Braeckman, J., & Guthrie, P. (2015, October). Loss of value: effects of delay on hydropower stakeholders. En *Proceedings of the Institution of Civil Engineers-Engineering Sustainability* (Vol. 169, No. 6, pp. 253-264). Thomas Telford Ltd.
- Rodríguez, J., Trespalcacios, J., y Galeano, M. (2021). Transferencia de incertidumbre en el mercado eléctrico colombiano: un enfoque de Rezago Distribuido Autorregresivo. *Revista de Economía y Finanzas*, 17(2), 75-90. <https://doi.org/10.1016/j.ref.2021.02.004>
- Rokach, L., & Maimon, O. (2005). Decision Trees. En *The Data Mining and Knowledge Discovery Handbook* (pp. 165-192). https://doi.org/10.1007/0-387-25465-X_9
- Santa María, J., López, A., y Rodríguez, C. (2009). Análisis de los precios de generación en el mercado mayorista de electricidad en Colombia. *Revista de Energía*, 24(1), 1-30. <https://doi.org/10.1016/j.reven.2009.01.002>
- Sierra, J., y Castaño, E. (2010). Pronóstico del precio spot del mercado eléctrico colombiano con modelos de parámetros variantes en el tiempo y variables fundamentales. *Estadística Aplicada: Didáctica de la Estadística y Métodos Estadísticos en Problemas Socioeconómicos*. Universidad Nacional de Colombia.
- Sun, M., & Meng, X. (2009). Taxonomy for change causes and effects in construction projects. *International Journal of Project Management*, 27(6), 560-572.

- Vega Plata, M. A. (2019). Aproximación portafolio eficiente de la matriz eléctrica de Colombia: un modelo de optimización lineal bajo diferentes escenarios de precios por emisiones de CO₂ y CH₄.
- Vélez, L. G. (2021). *Hidroituango para Dummies: Toda la Verdad*. Recuperado de <http://luisguillermovelezalvarez.blogspot.com/2021/10/hidroituango-para-dummies-toda-la-verdad.html>
- Villaplana Conde, P. (2003). *Valoración y gestión de riesgos en mercados eléctricos liberalizados*. Tesis Doctoral, Universidad Carlos III de Madrid.
- XM. (2019). *Informe Oferta y Generación*. Recuperado de <https://www.xm.com.co/nuestra-empresa/informes/informes-de-la-operacion-y-el-mercado/informes-mensuales-de-analisis-del-mercado>

ANEXO 1.

DEMOSTRACIÓN RELACIÓN PRECIO -RP

Si para el precio *spot* P se considera como aproximación adecuada la siguiente relación exponencial: $P = A \cdot e^{\alpha \cdot GNh + \sum \beta_i S_i + \sum \gamma_i f_i}$, donde S_i son variables fundamentales que sí se modifican con la generación de Ituango, f_i son variables fundamentales que no se modifican con la generación de Ituango y se define RP como la relación entre el precio *spot* en el caso con Ituango P^I y el precio *spot* en el caso sin Ituango P^N , entonces:

$$RP = \frac{P^I}{P^N} = \frac{e^{\beta_0} \cdot e^{\alpha \cdot GNh^I + \sum \beta_i S_i^I + \sum \gamma_i f_i}}{e^{\beta_0} \cdot e^{\alpha \cdot GNh^N + \sum \beta_i S_i^N + \sum \gamma_i f_i}}$$

Donde se cancelan algunos factores que coinciden en numerador y denominador:

$$\Rightarrow RP = \exp(\alpha \cdot GNh^I - \alpha \cdot GNh^N) \cdot \exp\left(\sum \beta_i \cdot S_i^I - \sum \beta_i \cdot S_i^N\right)$$

$$\Rightarrow RP = \exp(\alpha \cdot GNh^I - \alpha \cdot GNh^N) \cdot \exp\left(\sum \beta_i \cdot (S_i^I - S_i^N)\right)$$

Como α es factor común en el primer exponente:

$$\Rightarrow RP = \exp\{\alpha(GNh^I - GNh^N)\} \cdot \exp\left(\sum \beta_i \cdot (S_i^I - S_i^N)\right)$$

En este contexto, la diferencia entre la generación no hidráulica de los dos casos solo obedece a la aparición de la generación de Ituango. La generación no hidráulica para el caso con Ituango (GNh^I) se calcula restando la generación de Ituango ($GH_{ituango}$) de la generación no hidráulica en el caso sin Ituango (GNh^N): $GNh^I = GNh^N - GH_{ituango}$.

$$\Rightarrow RP = \exp\{\alpha(GNh^N - Gh_{ituango} - GNh^N)\} \cdot \exp\left(\sum \beta_i \cdot (S_i^I - S_i^N)\right)$$

$$\Rightarrow RP = \exp\{\alpha(-Gh_{ituango})\} \cdot \exp\left(\sum \beta_i \cdot (S_i^I - S_i^N)\right)$$

$$\Rightarrow RP = \exp\{-\alpha \cdot Gh_{ituango}\} \cdot \exp\left(\sum \beta_i \cdot (S_i^N - S_i^I)\right)$$

Si se asume que todos los factores fundamentales considerados en la modelación del precio se mantienen constantes ante la entrada de Ituango, se demuestra que RP tiene una relación exponencial decreciente con la generación estimada del proyecto:

$$\Rightarrow RP = e^{-\alpha \cdot Gh_{ituango}}$$