



UNIVERSIDAD NACIONAL DE COLOMBIA

Herramienta de datos para el análisis de precios de polipropileno en ambientes coberturistas

Carlos Daniel Fernández Múnera

Trabajo final presentado como requisito parcial para obtener el título de:
Magíster en Ingeniería - Analítica

Director:

Carlos Jaime Franco Cardona, Ph.D

Codirector:

Juan David Velásquez Henao, Ph.D

Universidad Nacional de Colombia – Sede Medellín
Facultad de Minas
Área Curricular de Sistemas e Informática
Medellín, Colombia
2020

Resumen

El presente trabajo final de maestría propone una herramienta de datos enfocada en el pronóstico para la gestión del riesgo de precios de polipropileno que permite mejorar los procesos decisionales del perfil coberturista. El trabajo se desarrolla desde el entendimiento del problema en términos de negocio y en términos analíticos, hace una revisión de la literatura y aborda una discusión sobre modelos posibles, para finalmente escoger la familia de modelos con mejor ajuste contrastando su desempeño en ambientes productivos.

Palabras Clave

Riesgo, mercado, polipropileno, industria, derivados financieros, coberturas, políticas

Abstract

Data tool for the analysis of polypropylene prices in hedging environments

The present final master's project proposes a data tool focused on forecasting for the risk management of polypropylene prices that allows improving the decision-making processes of the hedging profile. The work is developed from the understanding of the problem in terms of business and in analytical terms, makes a review of the literature and addresses a discussion on possible models, to finally choose the family of models with the best fit by contrasting their performance in productive environments.

Keywords

Risk, market, polypropylene, industry, financial derivatives, hedges, policies

Tabla de contenido

Resumen.....	2
Palabras Clave	2
Abstract	2
Keywords.....	2
Tabla de contenido.....	3
Introducción	4
1. Definición del problema en términos de negocio.....	6
1.2 Mecanismo de transmisión de precios	8
1.3 Aspectos técnicos.....	9
1.4 Aspectos económicos.....	9
1.5 Gestión del riesgo de precio del polipropileno en la industria colombiana	11
1.6 Las decisiones, la necesidad del pronóstico y las series de tiempo.....	13
2 Definición del problema en términos analíticos	15
2.1 Descripción del problema en términos analíticos.....	15
2.2 Información disponible	15
2.3 Limitaciones en términos analíticos.....	16
3 Revisión de la literatura	17
4 Objetivos	22
4.1 Objetivo general.....	22
4.2 Objetivos específicos.....	22
5 Metodología	23
5.1 Discusión y modelado	24
5.1.1 Análisis de filtro.....	24
5.1.2 Análisis de estacionalidad y tendencia de la variable dependiente.....	26
5.1.3 Análisis de causalidad.....	28
5.2 Modelación y evaluación.....	29
5.2.1 ARIMA dinámico.....	29
5.2.2 ARIMAX dinámico.....	32
5.2.3 Holt Winters	33
5.2.4 Regresión lineal, Lasso y XGBoost.....	35
5.2.5 Modelos a priori como complementos y soporte de la decisión.....	36
5.3 Conclusiones de modelado y análisis de resultados de la fase de modelado.....	37
6 Herramienta de soporte a la decisión	40
7 Conclusiones generales, recomendaciones y trabajo futuro.....	45
Referencias bibliográficas	47

Introducción

La gestión del riesgo de precios del polipropileno es un factor determinante para la rentabilidad de un negocio que se dedique a la transformación de esta resina en producto de uso final. La administración de materiales es la variable con mayor participación en la estructura de costos de una industria manufacturera, en este sentido se vuelve importante generar herramientas que permitan a las organizaciones alcanzar niveles de competitividad desde la gestión estratégica de compras.

Este problema del riesgo de precio puede resumirse en términos analíticos, como la necesidad de desarrollar herramientas de pronóstico a partir de datos propios y datos libres que expliquen los procesos de transmisión de precios en un horizonte futuro, permitiendo al usuario decisor tener una visual más amplia del mercado para que este logre ajustar sus políticas estratégicas con las dinámicas del mercado alcanzando una ventaja competitiva al tomar decisiones mejor informadas. En términos prácticos, el riesgo de precio en se resume como el cambio abrupto de una tendencia.

Teniendo en cuenta que no se han encontrado trabajos públicos que traten directamente el tema de precios de polipropileno, se encuentra útil abordar temáticas relacionadas con el pronóstico de bienes como petróleo, fertilizantes, arroz, energía eléctrica, entre otros; teniendo en cuenta que las metodologías utilizadas para el pronóstico de precios, podrían ser replicadas al polipropileno. El trabajo aborda literatura relacionada con el pronóstico de series de tiempo con el fin de evaluar si existen herramientas previamente desarrolladas que cumplan en alguna medida los objetivos de este trabajo, así como aportes teóricos y metodológicos de diversos autores que permitan hacer una mejor construcción de los modelos que se plantean.

En la sección 5.1 Discusión y modelado, se desarrolla una familia de modelos de pronóstico de precios de polipropileno que pretende ampliar el panorama del usuario decisor, puesto que, de forma recurrente, este se ve enfrentado a volatilidades importantes que ponen en riesgo la rentabilidad y liquidez del negocio. El modelo permite ampliar el panorama de mercado 3 meses delante mediante la combinación de cuatro modelos y un pronóstico combinado que busca encausar al usuario hacia la optimización de sus procesos decisionales a través del manejo de los datos. En el mismo capítulo 5.1, alrededor de los modelos desarrollados se plantea una discusión en la que se escogen los modelos de mejor ajuste.

Para su construcción se utilizaron datos libres que tienen relación con los mercados del petróleo y el dinero. Esta información fue complementada con un conjunto de datos construido manualmente desde el año 2014 hasta marzo del año 2020 y que representa la formación del precio del polipropileno; este conjunto fue construido a partir de las publicaciones semanales de un proveedor de información. El detalle de esta captura de información se encuentra desarrollado en la sección 2.2 Información disponible.

Con la información recaudada, en la sección 5.2 Modelación y evaluación, se desarrollaron modelos de tipo ARIMA, ARIMAX, Holw Winters, regresión Lasso, XG Boost, y modelos a priori que vienen siendo probados en ambientes reales desde septiembre del año 2019, observando resultados con buen ajuste (RMSE bajo) con respecto al comportamiento real de mercado, es decir demostrando poder predictivo, en síntesis el objetivo del presente trabajo es desarrollar un producto de datos que permita a la industria manufacturera de empaques flexibles (coberturista) la realización de un análisis descriptivo y predictivo de precios en el mercado del polipropileno .

El presente trabajo cumple con los objetivos que se plantea, en tanto que logra definir las variables de mayor impacto en términos de precios del polipropileno, las captura y las procesa en un modelo de pronóstico que sirve como herramienta de ayuda y soporte a la decisión a los usuarios finales. Finalmente se plantean las conclusiones generales y se proponen aspectos relevantes para trabajo futuro.

1. Definición del problema en términos de negocio

Pese a lo desarrollado de los mercados, aún persisten sectores con rezagos importantes en lo que a competitividad se refiere, situación que se observa en los sectores orientados a los mercados de energía. Estos rezagos tienen impactos importantes en la rentabilidad del sector industrial; puesto que las fuerzas del mercado no interactúan de forma libre. Como menciona (Benavides & Cardona, 2018) todavía es necesario aumentar competencia, diversificar el portafolio, entre otras.

El consumo de resinas, concretamente de polipropileno es un rubro de gran preponderancia dentro de la producción nacional, el cual, además, se convierte en un *commodity* electrointensivo a la hora de transformarlo en objetos de consumo final. Si bien en el caso puntual del polipropileno podrían estar dadas todas las condiciones para mercados competitivos, la concentración de pocos productores (Jiang, Marsh, & Tozer, 2015) mantiene en desventaja a los compradores. En fábricas con procesos de extrusión de resina, el consumo derivado del proceso puede representar el 60% del gasto energético (Vargas, Posada, Jaramillo, & García, 2015), situación que además impacta el agregado nacional de consumo eléctrico.

De acuerdo con (Trademap, 2019), Colombia exportó en 2018, alrededor de 288.070 toneladas de polipropileno; siendo el único fabricante local Propilco, filial de Ecopetrol, y que comercializa sus productos bajo la marca Esenttia. Por su parte, las importaciones ascendieron a 77.822 toneladas, cuyo origen principal es Brasil, país en donde opera la socia de Odebrecht, Braskem.

Según datos de la Encuesta Anual Manufacturera (DANE, 2019), para el año 2017, el consumo de polipropileno en Colombia fue de 201.200 toneladas. Haciendo cálculos simples con datos de la misma fuente, se puede inferir que, en promedio, de forma mensual para el año 2017, se transaron 64 mil millones de pesos de este material en Colombia.

En términos de precios, al menos de forma teórica se puede hablar de un mercado de libre competencia en tanto que hay opciones de compra pese a que la oferta se encuentra concentrada a nivel mundial en muy pocos productores (Gahleitner & Paulik, 2016). Si bien cualquiera podría establecer su refinería, las barreras de entrada al mercado se dan en tanto que se requiere un gran músculo financiero para establecer una refinería. En el caso de los fabricantes nacionales no es diferente, solamente la ampliación de la Refinería de Cartagena, por ejemplo, le costó al país más de 5.000 millones de dólares (Portafolio, 2013). La producción nacional está concentrada en Propilco y en algunos actores del mercado que fungen como distribuidores locales de material.

De acuerdo con los datos de importaciones (Trademap, 2019), los participantes del mercado cuentan con alternativas en el exterior que les permitirían hacer importaciones cuando los precios del mercado local les resten competitividad o su capacidad de compra les permita mejores opciones de negociación. Podría decirse que, en términos prácticos, este mercado funciona como un oligopolio, puesto que conviven en el ecosistema de negocios con múltiples demandantes y pocos oferentes.

Usualmente en el mercado se toman algunos referentes que determinan la formación de los precios, estos están dados según la geografía de acuerdo con índices de mercado. Algunos proveedores de índices de mercado reconocidos por la industria son Platts e Icis (Independent Chemical Information Service); el primero es una filial de S&P Global (Binary Tribune, 2013).

En conclusión, de cara al consumidor coberturista; entendiendo este como un ente o persona preocupado por el riesgo de los precios y el impacto de los cambios abruptos en la rentabilidad de su negocio, se encuentran una serie de problemas o fallas de mercado que lo posicionan como la parte débil dentro de la negociación, tanto así, que no solo está expuesto a las variaciones de la tasa de cambio, ya que las transacciones se dan en dólares, sino también a la volatilidad de los precios en un mercado desigual y expuesto a la violación del supuesto fundamental de los mercados competitivos: que compradores y vendedores estén debidamente informados para la toma de decisiones.

La figura 1-1 muestra la volatilidad de los precios en los últimos años en las plazas referentes más importantes del mundo, este conjunto de datos fue construido de forma manual, dato a dato con base a los reportes semanales de Platts (2019).

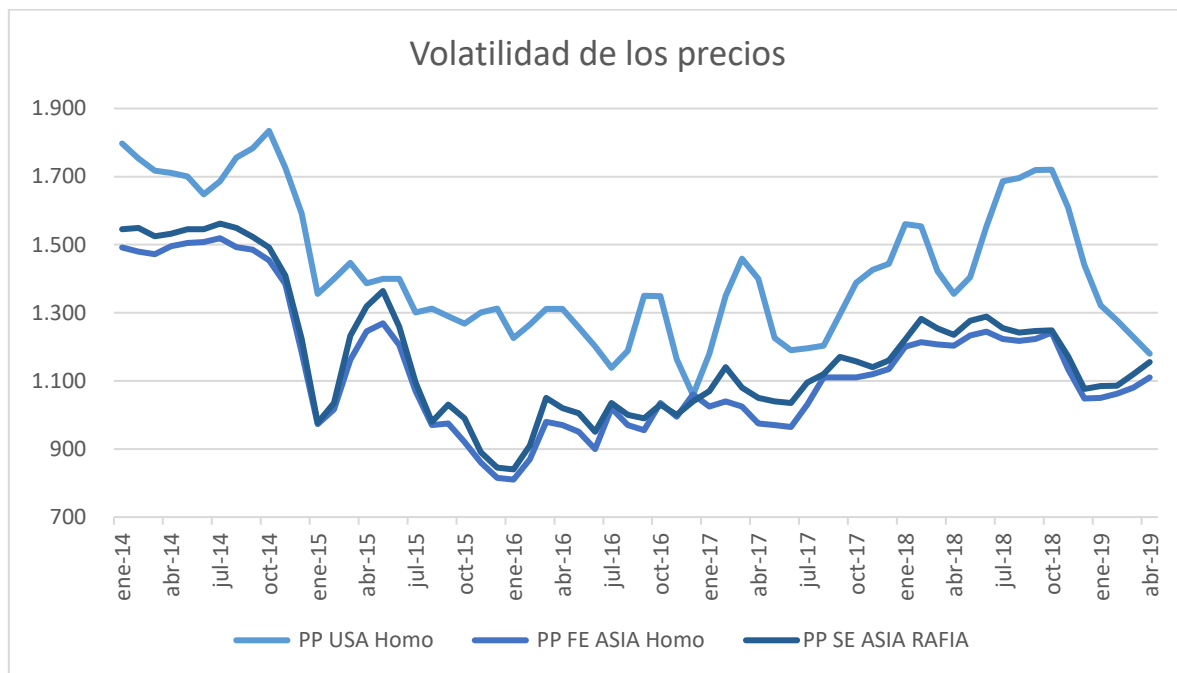


Figura 1-1. Volatilidad de los precios del Polipropileno

Fuente: elaboración propia, con dataset construido dato a dato del reporte semanal de Platts (2019)

Los riesgos mencionados anteriormente, podrían clasificarse en dos: los riesgos hoy gestionables y no gestionables. Sobre los riesgos hoy gestionables, el mercado ofrece alternativas de cobertura, como lo son el riesgo cambiario y el riesgo de tasa de interés sobre los cuales, existe información pública más o menos simétrica para que compradores y vendedores estén informados. En el

segundo grupo están los riesgos asociados al precio del material, que pese a que existan algunos productos financieros estandarizados como el PGP (CME, 2019) o pueda llegarse a la negociación de algún OTC, el mercado no cuenta hoy con la liquidez para ofrecer ese tipo de productos, aún si existiera, sería de alto valor la generación de herramientas que permitan gestionar la incertidumbre de los precios del *commodity* mencionado, dándole a los coberturistas un panorama con posibilidad de pronóstico del mercado del polipropileno.

Después del contexto establecido de forma previa, es válido resumir el problema en términos de negocio como la inexistencia de herramientas de uso libre para la gestión del riesgo de precios de polipropileno para la industria colombiana transformadora de plástico, que como se demostró en la Figura 1-1 es un material que presenta altas volatilidades y podría generar riesgos importantes en la rentabilidad y liquidez de los negocios.

Para generar mayor claridad al respecto, a continuación, se aborda la dinámica de transmisión de precios a lo largo de la cadena refinación, se explican de forma breve aspectos técnicos y económicos, posteriormente, se hacen algunas claridades respecto de la gestión del riesgo de precio. Finalmente se abordan las necesidades del pronóstico como herramienta necesaria para mejorar los procesos decisionales.

1.2 Mecanismo de transmisión de precios

Para abordar el tema de precios de polipropileno, es importante partir de la cadena de transmisión de precios a lo largo de la cadena de refinación, para comprender el origen del polipropileno tanto desde variables que impactan su precio como desde su proceso productivo, evidentemente limitado, puesto que no se trata de un trabajo técnico sobre el *commodity* sino de gestión de riesgo de precio del polipropileno. La Figura 1-2 muestra la forma como se impacta la cadena de refinación del PP para la semana previa al 5 de junio del año 2019.

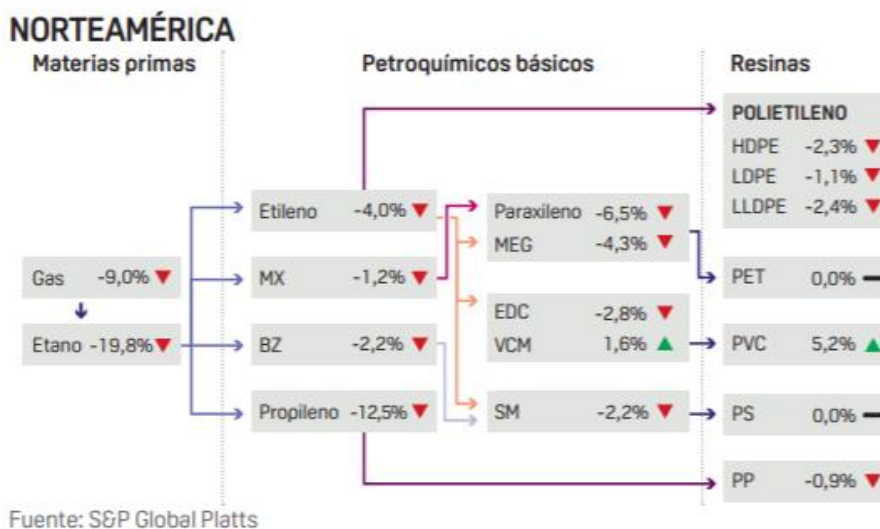


Figura 1-2. Mecánica de transmisión de precios

Fuente: Platts (2019)

1.3 Aspectos técnicos

El Polipropileno se obtiene a partir de la cristalización de propileno; este último es un residuo gaseoso que se consigue de la refinación de petróleo en nafta o del etano, que se surge de partir de la deshidrogenación de gas. Si bien se puede obtener por dos caminos distintos, los costos de producción pueden variar considerablemente entre uno y otro (Gahleitner & Paulik, 2016). Mencionar las alternativas de producción podría inicialmente parecer intrascendente, pero es bien relevante si se toman en cuenta los avances regulatorios en cuanto a la producción de combustibles limpios, que obliga a que los crudos que hoy son destinados a la producción del material, sean desviados a la producción de combustibles más limpios (Platts S&P, 2018); dejando en ventaja a los productores que utilizan el gas como materia prima y obligando a una modernización tecnológica de gran tamaño a los productores a partir del nafta; factores que evidentemente transformarán la actual lógica de precios.

La historia del polipropileno está altamente ligada a la del polietileno y fueron los avances de Ziegler y Natta, condecorados con el premio Nobel de química de 1964, quienes con sus investigaciones permitieron generar los catalizadores tipo Ziegler-Natta, compuestos organometálicos que permiten la polimerización estereoespecífica de olefinas a temperaturas moderadas (Gahleitner & Paulik, 2016), avance que permitió el crecimiento exponencial de las olefinas, como de desarrollar en el capítulo destinado a los aspectos económicos de este hito científico.

1.4 Aspectos económicos

La disponibilidad y lo económico del polímero hicieron que desde 1964 se dieran incrementos exponenciales en su producción y demanda [como se muestra en la figura 1-3] (Gahleitner & Paulik, 2016); de ahí, que los productos derivados de este material fueron tomando grandes participaciones de mercado e incluso se empezaron a crear sustitutos a otros producidos con insumos más costosos, por ejemplo, las botellas de vidrio.

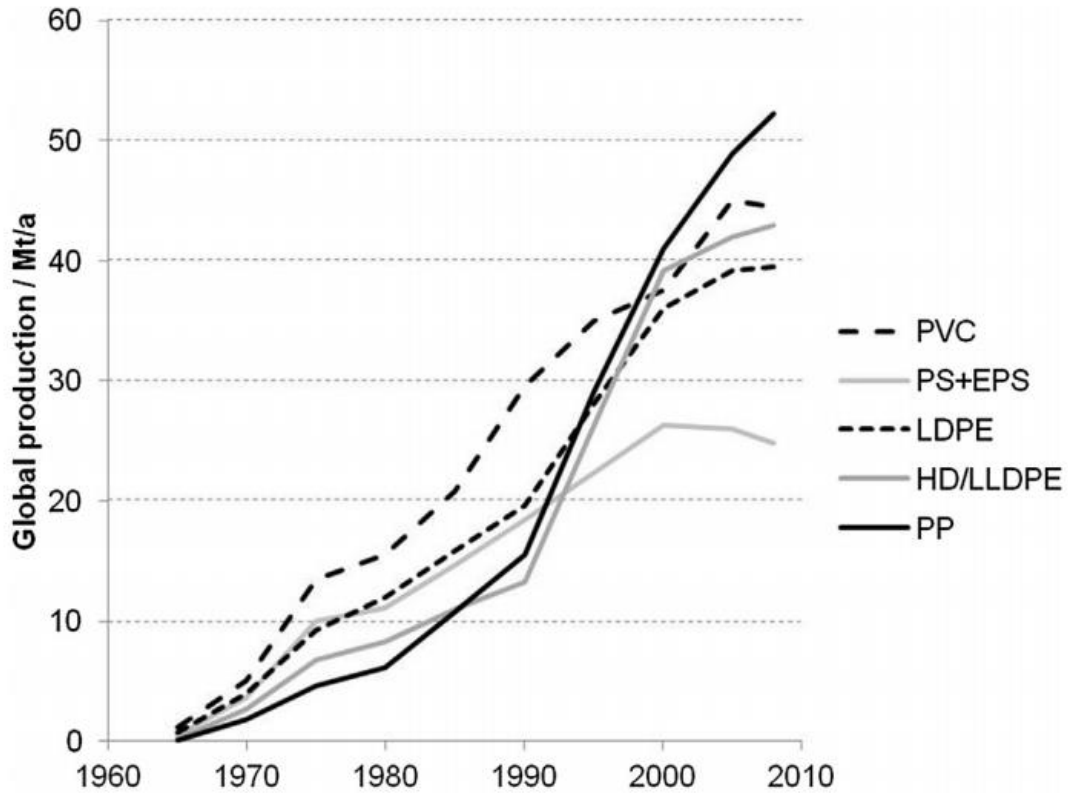


Figura 1-3. Crecimiento exponencial de la demanda de polímeros
Fuente: (Gahleitner & Paulik, 2016)

El uso masivo de estas resinas en la producción de bienes fue posible gracias a los avances importantes que se dieron en los procesos de catalización y polimerización (Gahleitner & Paulik, 2016) que fueron mencionados en el aparte de los aspectos técnicos. De acuerdo con datos de IHS, para el año 2013, China fue el mayor productor de PP del mundo [como se muestra en la figura 1-4]; y dicha producción estuvo concentrada en menos de 10 compañías (Gahleitner & Paulik, 2016).

De lo anterior, es válido afirmar que es un mercado altamente concentrado y pese a la volatilidad de los precios presentada al principio del escrito [figura 1-1], la concentración del mercado en pocos productores genera riesgos que impiden que el activo se transe en un mercado competitivo y deja en desventaja a los consumidores al no tener las mismas fuerzas que tienen los oferentes concentrados.

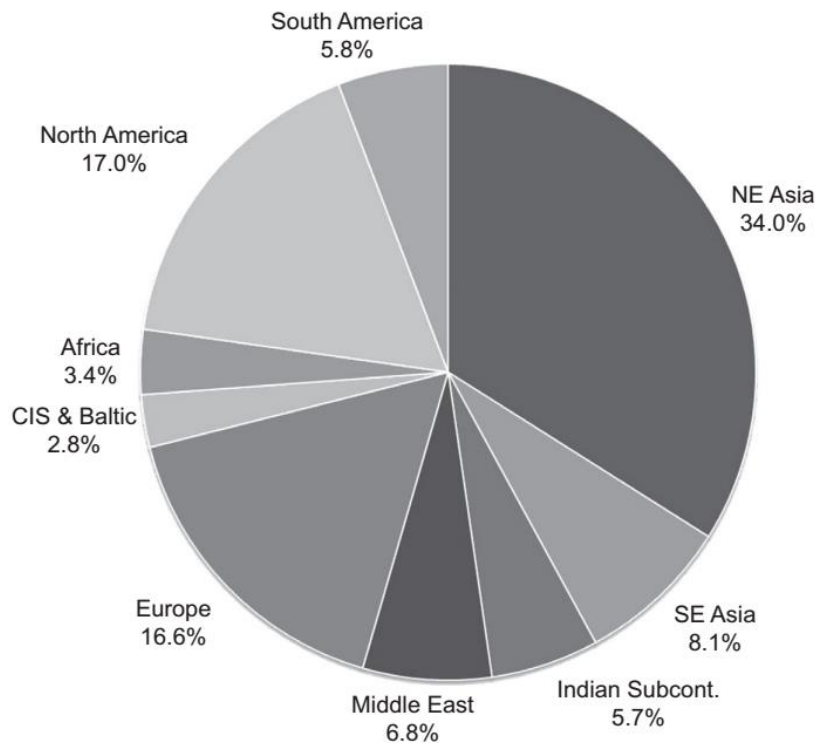


FIGURE 11.2 Regional distribution of global polypropylene production (2012, data from IHS Chemical, 2013).

Figura 1-4. Distribución regional de la producción de polipropileno

Fuente: (Gahleitner & Paulik, 2016)

1.5 Gestión del riesgo de precio del polipropileno en la industria colombiana

Ahora planteados algunos números de la oferta y la demanda, resulta prudente tratar la gestión del riesgo al que se ven expuestos los perfiles coberturistas. Suponga un industrial colombiano con una balanza comercial deficitaria. Importa el 100% del material para su producción; pero sus ventas al exterior son apenas el 30% de su facturación. Adicionalmente mantiene una exposición por su endeudamiento en dólares. En este sentido, el perfil antes descrito se encuentra expuesto fuertemente a los riesgos derivados de la moneda local, es decir, su fortaleza o debilidad frente al dólar estadounidense, así como a la volatilidad de los precios internacionales de su materia prima.

Estos factores son de suma trascendencia puesto que el tipo de perfil industrial en mención se ve impactado en sus costos de producción por los dos factores y debe encontrar herramientas para gestionar la incertidumbre, aun cuando el mercado financiero en Colombia es incipiente. Hoy en el mercado financiero colombiano no se encuentran derivados financieros para cubrir el polipropileno.

En términos de PP es difícil encontrar estrategias de cobertura debido a la fuerte concentración de productores y resulta complicado encontrar mercados de futuros lo suficientemente líquidos para

gestionar sus riesgos como tradicionalmente se expone en la literatura y como normalmente, al menos desde los escenarios académicos se considera la ruta más adecuada (Jiang, Marsh, & Tozer, 2015).

Pese a lo robusto e importante del mercado, no se cuentan con futuros estandarizados con la liquidez suficiente para transar de forma recurrente, además de no contar con herramientas decisionales que permitan a los coberturistas gestionar sus riesgos de precio. Al respecto (Jiang, Marsh, & Tozer, 2015) confirman que existe un mercado muy limitado de futuros para productores de plástico. Esto, debido a la concentración de fuerza que tienen los productores y que, en el hemisferio, podrían contarse con los dedos de las manos.

De acuerdo con la revista (Mundoplast, 2019), en 2020 el consumo mundial de polipropileno alcanzará 62,4 millones de toneladas, de las cuales 20,1 corresponderá al sector del *packaging*; lo que, sin duda, es una oportunidad para generar conocimiento en cuanto a la formación de precio del *commodity*.

El administrador financiero que se desenvuelve en esta industria se ve enfrentado a una serie de complejidades en términos de la gestión. Es una persona que debe incursionar en los mercados del dinero y de capitales; en tanto que su obligación es garantizar la financiación de la operación, el crecimiento (inversión) y los retornos a los accionistas (García, 2009), así como la adecuada gestión de costos y gastos que garanticen la competitividad de la compañía.

Hasta este punto se ha tratado la necesidad de tener información para efectuar compras de materias primas a precios competitivos, pero vale la pena preguntarse, ¿qué impactos tiene esto en la rentabilidad de la empresa si se mira desde la perspectiva de un industrial, tal como se ha enfocado todo el trabajo? Y según la experiencia profesional, puede afirmarse que los precios finales del material puesto en la fábrica serán aproximadamente el 60% de los costos de producción y venta, por tanto, errores de decisión en la negociación podrían ser de alto impacto para la liquidez y rentabilidad del negocio.

El administrador financiero, en este caso se ve enfrentado a una serie de riesgos de tiempo, modo y lugar que, si bien la teoría le ofrece herramientas para manejarlo y gestionarlo, otra es la realidad del mercado colombiano. El financiero del hoy debe valerse de su intuición y astucia para lograr costos competitivos haciendo un delicado juego con el Capital de Trabajo, teniendo en cuenta que este es un concepto ampliamente asociado con la liquidez.

Las operaciones cerradas en el mercado spot, normalmente son de unos volúmenes importantes que van más allá de un mes de funcionamiento; estas exigen la disponibilidad de un importante nivel de capital de trabajo para el adecuado cierre de las operaciones.

En síntesis, de los aciertos o desaciertos en la política de compras y coberturas de materias primas, tendrá fuertes efectos en la rentabilidad de un negocio como el descrito previamente.

1.6 Las decisiones, la necesidad del pronóstico y las series de tiempo

Como se mencionó previamente, el impacto de las decisiones que se tomen en términos de las materias primas en el sector manufacturero configurará un inductor o destructor de valor en lo que a competitividad y gestión del Capital de Trabajo se refiere. En este sentido, es evidente la necesidad del pronóstico y de ayudas que soporten las decisiones estratégicas de la organización. Aunque haciendo referencia al mercado eléctrico, (Velásquez, 2008) hace mención de la necesidad de pronóstico por parte de los distribuidores y los grandes consumidores, haciendo hincapié en el impacto que tiene la predicción en la planeación de operaciones; en tanto que los usuarios bien informados podrán desarrollar negociaciones bilaterales que sean beneficiosas para ambas partes y que mejoren la rentabilidad de sus negocios.

Si bien los modelos de pronóstico son una importante herramienta decisional, es necesario entender que el modelo per se no es una bola de cristal, estos no son más que una interpretación de realidad extendida, sujetos a unas condiciones particulares de tiempo modo y lugar. En este sentido, debido a las complejidades de la realidad y a las nuevas dinámicas que día a día se incorporan en las situaciones de riesgo que enfrentan las organizaciones, se hace necesario, además de los pronósticos, contar con herramientas complementarias, así como lo son los criterios de expertos y las combinaciones de pronóstico, de modo que a la hora de tomar decisiones que puedan afectar de forma importante la rentabilidad del negocio, se tengan suficientes elementos de juicio.

De acuerdo con (Petropoulos, Fildes, & Goodwin, 2005) aun cuando es de suma importancia el criterio experto en las decisiones, es usual que en las empresas se dé una mayor participación al juicio profesional, desviando de forma considerable las decisiones de los resultados que plantean los modelos. (Petropoulos, Fildes, & Goodwin, 2005) demuestran en su estudio que los pronósticos que se ven ajustados únicamente por criterio experto fueron en promedio un 14% peores que los métodos estadísticos por sí solos.

(Petropoulos, Fildes, & Goodwin, 2005) sugieren que se deben aplicar estrategias de corrección un poco más simples: como lo son la combinación no ponderada de pronósticos estadísticos, dicha mezcla permite por un lado mejorar el pronóstico y por el otro, ser un soporte de nivel superior para los expertos que quieran ajustar los pronósticos con base al criterio, en la medida que le suministra al experto visiones distintas así como complementos, por ejemplo una gráfica de acuerdo entre modelos que permite ver una tendencia entre los modelos o detectar una particularidad o anomalía que esté presentado el mercado en un momento puntual.

Haciendo combinación simple de modelos (Petropoulos, Fildes, & Goodwin, 2005) demuestran mejoras de hasta el 16% en sus ejercicios, dada la simplicidad de lo que proponen es de alta utilidad práctica y será aplicado en la fase de experimentación.

De acuerdo con las clasificaciones que da la literatura, estos modelos pueden ser explicativos o predictivos (Velásquez, 2008). Los primeros se basan en el entendimiento y comprensión de los datos, es decir un modelo de orden descriptivo, en síntesis, se basa en entender propiedades o condiciones puntuales de la serie de forma retrospectiva. Los segundos están orientados

exclusivamente a visualizar valores futuros obviado muchos aspectos metodológicos (Velásquez, 2008).

En síntesis, se puede concluir del capítulo anterior dos ideas básicas. Lo primero es la importancia de la planeación y el modelado en términos de operación de negocios, es importante definir una ruta con base en información que conduzca al cumplimiento de los objetivos corporativos, lo segundo, es que los modelos per se no son una bola de cristal pero si un poderoso instrumento de planeación; si se hacen combinaciones simples con otros pronósticos y criterios expertos. El capítulo siguiente se dedica a definir el problema en términos de negocio y en términos analíticos

2 Definición del problema en términos analíticos

2.1 Descripción del problema en términos analíticos

En términos analíticos, el problema se resume como la necesidad de desarrollar una herramienta que permita exponer varios modelos de pronóstico que ayuden al experto a construir un panorama más amplio sobre los precios del polipropileno, permitiendo al usuario mejoras en sus procesos decisionales de compra, apoyando la gestión eficiente de los riesgos de precios y permitiendo competitividad a la industria transformadora de plástico en Colombia. Una herramienta que le permita mediante la combinación simple de pronósticos tomar varias visiones con sustento metodológico que le permitan soportar y complementar sus decisiones.

2.2 Información disponible

La recolección de datos se enmarca en un proceso de búsqueda en la red de información libre, variables que bajo un criterio experto inicial tuviesen algún grado de asociación con el Polipropileno. En esta fase del trabajo, el único fin era construir un conjunto de variables externas que permitiera mezclarse con la información del mercado del polipropileno capturada manualmente desde 2014.

Las variables exógenas debieron extraerse de entidades gubernamentales o servicios de información, no obstante, Quandl (www.quandl.com) es una compañía que se dedica a la consolidación y tratamiento de datos de diversas fuentes y latitudes del mundo.

“[...]es una plataforma para los datos financieros, económicos, y alternativa que sirve a los profesionales de inversión. Quandl es fuente de datos de más de 500 editores. Todos los datos de Quandl son accesibles a través de un API . El acceso a la API es posible a través de paquetes para múltiples lenguajes de programación, incluyendo R , Python , Matlab , arce y Stata . [...] Desde su lanzamiento, Quandl ha sido discutido como un disruptor de datos en el sector financiero tradicional. Con más de 200.000 usuarios Quandl ha sido considerado como una alternativa a la terminal de Bloomberg y Thomson Uno.” (Qwerty Wiki, 2019)

El inventario de datos que resulta después del proceso de captura es el siguiente:

- Dataset de precios de Polipropileno mensual (Construcción propia, los demás son extraídos de Quandl)
- Precios OPEC del petróleo, diarios y recalculado al promedio mensual desde 2003
- Producción de crudo desde 2003: Arabia Saudita, Canadá, China, Rusia, USA
- Producción de gas desde 2012: Rusia, Noruega, Canadá, USA
- Refinación de Nafta desde 2009: Alemania, Arabia Saudita, China, Canadá, USA
- Refinación de Motor y Aviación desde 2002: Alemania, China, USA, Canadá, Arabia Saudita
- Refinación de Kerosene desde 2009: Alemania, USA, Canadá

- Refinación de Diesel desde 2002: Alemania, USA, China, Canadá
- Refinación de Fuel Oil desde 2002: Alemania, Canadá, USA, China
- EUR/USD desde 1999
- Yuan/USD desde 1999
- Tasa de ocupación de las refinerías en estados Unidos

Los datos de este inventario de datos se utilizaron con varios fines: determinar las variables que tienen impacto en la formación de los precios del Polipropileno y adicionalmente, como insumos del análisis descriptivo que acompaña los ejercicios de pronóstico, constituyendo una herramienta de soporte a la decisión.

2.3 Limitaciones en términos analíticos

Teniendo en cuenta la información disponible mencionada en el aparte anterior, es pertinente mencionar limitaciones en términos analíticos. Lo primero, es que no se trata de un problema de *big data*, desde la variable a explicar se cuenta con un conjunto de datos escaso, costoso y apetecido por quienes se mueven en la industria plástica, en la medida que ha sido recaudado de forma mensual desde el año 2014 a través de cálculos propios con base en una suscripción a un proveedor de información que cuesta alrededor de 8 mil dólares al año, acceder a un conjunto más grande podría representar grandes gastos.

Lo segundo, es que de acuerdo con el volumen de datos; se observó en la etapa de construcción de modelos que las herramientas más robustas y sofisticadas como lo son, por ejemplo, las redes neuronales artificiales, no tienen un buen desempeño en productivo, por lo que se optó por aplicar herramientas de pronóstico más clásicas que demostraron un buen desempeño en ambientes productivos.

En síntesis, concluyendo sobre el presente aparte, es válido reafirmar que se trata justamente de un problema que tiene información limitada, por las mismas limitaciones en cantidad de información; es un tipo de problema al que le van mejor herramientas clásicas y menos sofisticadas de pronóstico. En el capítulo siguiente se aborda la revisión de literatura desde varias perspectivas de pronóstico; arroz, petróleo, fertilizantes, gas, energía eléctrica y polipropileno recuperado.

3 Revisión de la literatura

Es importante marcar el inicio de este capítulo clarificando que no se encontraron estudios públicos directamente relacionados con la transmisión de precios de polipropileno y tal vez todo esto puede verse influenciado, como se ha mencionado de forma reiterada, a la concentración del mercado (Jiang, Marsh, & Tozer, 2015).

Al respecto, (Jiang, Marsh, & Tozer, 2015) desarrollaron un trabajo que aborda los cambios en las políticas y la evolución de las tecnologías verdes. Estos autores abordan el comportamiento del mercado de petróleo, maíz, y plásticos en Estados Unidos. Usan un modelo de soporte vectorial para explicar la volatilidad de precios entre los mercados mencionados en Estados Unidos. Además, concluyen que, en el largo plazo, los precios de los plásticos y los futuros de maíz se mueven juntos, esta relación tiene causalidad en la producción de etanol y gasolina.

En este sentido, la relación planteada por los autores sirve en la medida que podría utilizarse el maíz como variable exógena de pronóstico para los precios del polipropileno, o desde la administración financiera, como opción de cobertura teniendo en cuenta las dificultades para encontrar los productos financieros adecuados para la gestión del riesgo, los agentes se ven obligados a explorar opciones de coberturas cruzadas (Jiang, Marsh, & Tozer, 2015), siendo mercado de futuros de maíz y el mercado del crudo mucho más líquidos que los mercados del PP (CME, 2019).

Sobre las coberturas cruzadas, al menos en las configuraciones actuales del mercado local, únicamente se cuenta con opciones de cobertura acudiendo al mercado spot, donde el decisor basado en las expectativas de precios cubre la operación mediante derivados de tasa de interés y moneda.

Sin embargo, surgen cierto tipo de preguntas como cuál sería la forma óptima de garantizar el abastecimiento del material desde las lógicas del Objetivo Básico Financiero (García, 2009), teniendo en cuenta que los usuarios decisores no cuentan hoy con herramientas para el análisis, modelación de precios futuros y creación de estrategias de cobertura que le permitan maximizar la rentabilidad.

De acuerdo con los planteamientos de (Jiang, Marsh, & Tozer, 2015), podría pensarse en acudir a los mercados de futuros de Estados Unidos y trazar una estrategia de cobertura cruzada en mercados con mayor liquidez, coberturas que serían realizadas únicamente mediante formas *non delivery*, ya que al industrial colombiano en cuestión no le interesaría hacerse físicamente al crudo o al maíz, únicamente le interesa cubrir sus riesgos asociados a la materia prima.

Si bien todo lo planteado en este trabajo está direccionado a entender la formación de los precios del polipropileno en el mercado primario, es decir nuevo o virgen, (Stromberg, 2004) aborda una problemática similar, pero en el mercado secundario, es decir, en el mercado de plástico reciclado. El ofrece un estudio de la volatilidad de los precios en Seattle USA. Éste estudio busca explicar la formación de precios mediante herramientas econométricas, toma datos mensuales, datos locales,

y el uso de variables explicativas para abordar los cambios de los precios; como variable explicativa toma básicamente los niveles de existencia o desechos que se generan en el sector primario. Finalmente concluye las elasticidades de los precios entre uno y otro mercado son insignificantes y que más bien depende de las políticas gubernamentales en materia del uso de plásticos.

Es necesario abordar la imperfección del mercado del PP. Al indagar por las imperfecciones particulares que se presentan en este, se confirman las sospechas planteadas párrafos atrás en este documento, donde se aborda el impacto que tiene la concentración del mercado en pocos productores. Los principales fallos de este mercado están asociados con los excesivos costos de transacción, la asimetría de la información y la concentración de la propiedad (Stromberg, 2004), esta última entendida como concentración de la cuota de mercado.

Un direccionador de dinámica de precios en cualquier mercado es la incertidumbre (Stromberg, 2004), que propicia ambientes de abundancia y escasez que tienen impactos directos en la confianza y sentimientos de los actores del mercado. Esto redundando en transmisión de rápidos efectos en los precios, generando enormes volatilidades que forjan escenarios complejos para la industria convertidora.

Otro factor con impactos grandes en la formación de precios es la inelasticidad de la demanda (Stromberg, 2004), es decir, la configuración del mercado tal y como se conoce hoy impide a los convertidores reaccionar al uso de bienes sustitutos que les permitan maniobrar cuando se presentan escenarios desfavorables en los precios, pues se requerirían cambios enormes en tecnología, mientras que el consumidor final de empaque flexible, puede migrar sin mayor esfuerzo del polipropileno al polietileno; empeorando aún más las finanzas del fabricante de empaques flexibles en PP.

Es pertinente mencionar, además, los nuevos riesgos a los que se enfrentan los convertidores en la medida que aparecen regulaciones de combustibles más limpios (Platts S&P, 2018) que impactarían los precios y abastecimiento del material y lo más importante es que la sociedad atraviesa una etapa de transformación en sus valores colectivos; en la medida que día a día aparecen en diferentes lugares del mundo regulaciones antiplástico que dejan el mercado del PP enfrentado a dos caminos: reinventarse o ser reemplazado; situación que tendrá sin duda impactos en la formación de precios.

Volviendo a las coberturas cruzadas como una opción de gestión del riesgo (Gutierrez, 2016), presenta un trabajo en el que amplía los modelos de correlación dinámica para diseñar estrategias de cobertura cruzada dinámica para el petróleo mexicano. Utiliza un modelo de tipo MGARCH-CCD aplicado al mercado de futuros del WTI como cobertura. Sus hallazgos demuestran efectividad en las coberturas cruzadas y logra disminuir los riesgos asociados con el precio.

Está demostrado que en largo plazo el efecto de las coberturas es neutro (Gutierrez, 2016); sin embargo, el objetivo fundamental es generar ambientes de menor volatilidad para mejorar la administración y la competitividad de las empresas. Básicamente el objetivo de las coberturas es fijar el precio, es decir, atar una de las variables para tener certeza sobre ella pero que no necesariamente implica que genere beneficios o pérdidas al momento de su liquidación, por eso, ya

sea que se trate de transar en el mercado spot o en el mercado de futuros, resulta de suma importancia para los entes económicos contar con herramientas de decisión que le permitan con un buen grado de confianza al determinar la ventaja competitiva que tiene su alternativa de abastecimiento y cobertura con base a la información disponible, y tratando de mejorar en algún grado las asimetrías del mercado en términos de información y de desigualdad entre oferentes y demandantes.

Existe una relación causal entre el mercado del gas y el mercado del polipropileno, puesto que como se mencionó en los antecedentes, una de las vías de obtención del PP es mediante la deshidrogenación de propano (Gahleitner & Paulik, 2016), por lo tanto, la relación obedece a que ambos tienen relación con los determinantes de oferta y demanda del petróleo (González, Ávila, Gitter, & Ángel, 2011), en temas referentes a la transmisión de precios se observan similitudes en cuanto al rezago, por tanto, se vuelve importante tratar el mercado del gas en lo que a formación de precios se refiere.

En cuanto al mercado del gas, (González, Ávila, Gitter, & Ángel, 2011) correlacionan la formación de precios del petróleo, el gas y los fertilizantes. Mencionan (González, Ávila, Gitter, & Ángel, 2011) que “empleando una regresión con variable dicótoma se concluye que, el precio nominal del amoníaco al que compra el agricultor tiene una dependencia asimétrica con respecto al precio nominal del gas natural con dos meses de retraso al que vende Pemex Petroquímica” (González, Ávila, Gitter, & Ángel, 2011); y lo anterior no significa que se quiera relacionar las variables estudiadas por los autores con el problema puntual que motiva este escrito, sin embargo suena útil y llamativo la metodología aplicada por estos y que en la experimentación del presente trabajo se encuentran relaciones similares en términos del rezago temporal mediante un test de causalidad de Granger.

El estudio de González, Ávila, Gitter y Ángel (2011) se orienta a las crónicas correlacionales; “en el análisis de correlación el objetivo principal es medir la fuerza o el grado de asociación lineal entre dos variables. El coeficiente de correlación mide esta fuerza de asociación. La serie de datos de los precios nominales del gas natural y del amoníaco anhidro están fuertemente asociados.” (González, Ávila, Gitter, & Ángel, 2011). No obstante, cabe aclarar que correlación no es causalidad, pero sí puede medir comportamientos similares de los precios de diferentes activos, y esto en términos prácticos puede tener poder explicatorio a pesar de que empíricamente no pueda demostrarse la dependencia.

En este sentido, sería válido encontrar variables con asociaciones o comportamientos similares pese a que desde el empirismo sea difícil demostrar su relación directa pero que desde la práctica permitan resolver el problema en cuestión, lo anterior complementado con herramientas estadísticas serán insumos importantes en la fase de selección de datos y características que se aborda más adelante.

Llama también la atención, un estudio de transmisión de precios realizado en el mercado colombiano de arroz (Troncoso, 2019), donde se analizan las variables que tienen poder explicatorio en los precios; pero fundamentalmente en lo relacionado con el presente escrito resultan valiosos

los aportes en términos de los modelos utilizados. En la revisión de literatura del artículo en cuestión, se hacen alusiones a una serie de modelos matemáticos y econométricos, como lo son: *Markov switching vector error correction model* (MS-VECM), bajo este modelo el autor encuentra que hay una transmisión especial de precios en ese mercado e incluso clasifica en regímenes las transiciones de precios; en el mismo trabajo se citan trabajos como los de Farrel (1952), que resultan ser potencialmente útiles en la medida que son modelos transmisión de precios que incorporan umbrales o rezagos en su transmisión (Troncoso, 2019). El tema de los rezagos se convierte en un tema que determina la precisión de los modelos utilizados, como se observa en la fase de modelado del trabajo con el test de causalidad de Granger.

Se encuentra en el artículo de Troncoso, que “Los desarrollos van desde modelos autorregresivos de rezagos distributivos (Farrel, 1952; Wolfram, 1971; Houck, 1977) hasta aproximaciones de corrección de equilibrio, como los modelos de ajuste parcial (Bacon, 1991), y las primeras aplicaciones a modelos de corrección de errores (Engle & Granger, 1987).” De lo anterior, se evidencia la necesidad de incluir modelos autorregresivos en el desarrollo del trabajo.

Si bien hasta este punto se han referenciado una cantidad importante de modelos, no quiere decir eso que el alcance del presente trabajo sea abarcarlos per se, es una revisión del estado del arte en lo que a transmisión de precios se refiere, de activos que por algunas condiciones pudieran asimilarse en sus comportamientos en transmisión de precios al polipropileno.

Una similitud del mercado del arroz con el del PP, es el alto grado de incertidumbre que se presenta en materia de precios, demuestra Troncoso que por lo menos en el caso colombiano, el mercado de arroz presenta altas volatilidades pese a ser un sector protegido por las políticas gubernamentales. El logro final de la investigación de Troncoso fue presentar los resultados para los coeficientes del modelo que se planteó; donde logró explicar cómo se transmiten los precios del arroz en el mercado colombiano.

Si bien hasta el momento no se han encontrado contribuciones literarias directamente relacionadas con el problema de investigación, se han abordado temáticas que persiguen objetivos parecidos aunque desde la relación de los precios de bienes distintos, como los fertilizantes con el crudo (González, Ávila, Gitter, & Ángel, 2011), sobre la transmisión de los precios del arroz en Colombia (Troncoso, 2019).

Ahora se aborda la temática desde los pronósticos de precios de energía eléctrica en Colombia. De acuerdo con (Barrientos, Rodas, Velilla, Lopera, & Villada, 2012) normalmente los precios de los bienes están determinados por la interacción entre las fuerzas de oferta y demanda, sin embargo, el mercado de energía eléctrica presenta una serie de particularidades. Es un mercado que se caracteriza por la existencia de pocos generadores y muchos consumidores con poco poder de negociación; situación similar al mercado del polipropileno.

En la misma línea planteada por (Barrientos, Rodas, Velilla, Lopera, & Villada, 2012), el comportamiento no lineal de los precios del mercado de energía, mercado con volatilidades tan grandes como las planteadas en los antecedentes del mercado del PP, desde la década de los 90 se

han planteado nuevos métodos de pronóstico basados en redes neuronales artificiales, cuyo fin máximo es incorporar en los modelos relaciones no lineales entre variables. Este tipo de modelos han sido ampliamente usados para pronosticar relaciones entre monedas como USD/CAD. Vale la pena aclarar que no se hayan relaciones de causalidad entre el polipropileno y el USD/CAD, sin embargo, resulta interesante probar modelos aplicados a precios altamente volátiles.

Un tema importante para el cumplimiento de los objetivos de este trabajo, es la selección de variables, (Barrientos, Rodas, Velilla, Lopera, & Villada, 2012) también tratan este aspecto en su aporte a la predicción de precios en el mercado energía eléctrica en Colombia; sin embargo en el caso de energía eléctrica se toman como variables dependientes tópicos relativamente evidentes como lo son el PIB, el nivel de los embalses y el agregado de la demanda de energía como determinantes del precio. En el trabajo se prueban diversas metodologías de pronóstico basados en redes neuronales y las herramientas clásicas para el análisis de series de tiempo, como lo son los modelos ARIMA. Sin embargo, estos autores en este artículo concluyen que para el mercado eléctrico colombiano tiene mejor desempeño los modelos con un enfoque econométrico clásico.

En el mismo mercado eléctrico, pero (Muñoz, Urquijo, Castro, & Lombana, 2017) sostienen que la volatilidad de los mercados energéticos hace que los pronósticos de sus precios se conviertan en un reto, su objetivo es demostrar que modelo autorregresivo es el que mejor pronostica los precios de la electricidad en Colombia, reconociendo la importancia de la planeación de los ingresos y los costos para los actores económicos.

Mencionan entre otras como determinantes de los precios los hábitos de la población, fenómenos climáticos y restricciones de transporte, variables que evidentemente tienen impactos en la formación de los precios del PP, en tanto que los precios se ven afectados por la temporada de huracanes en USA, las limitaciones de los *Railcars* en USA entre otras. Debido a lo complejo de incorporar las variables, los autores consideran a los modelos autorregresivos como la mejor opción de pronóstico, esto es explicar la variable con ella misma (Muñoz, Urquijo, Castro, & Lombana, 2017), lo anterior dada la complejidad de identificar e incorporar variables explicatorias.

Un modelo ARIMA es dinámico y de corto plazo, toda vez que para pronosticar el valor de febrero de 2017 debe alimentarse la serie con el valor hallado por el modelo de enero de 2017. Este valor pronosticado de enero se convierte en un valor histórico, con el que se genera medias, varianzas, correlogramas y coeficientes diferentes antes de su ingreso a la serie original mensual de enero de 2004 a diciembre de 2016, siendo necesario iterar para identificar, estimar, verificar y así encontrar un nuevo modelo autorregresivo integrado de media móvil con *igarch* y predecir el periodo de febrero de 2017" (Muñoz, Urquijo, Castro, & Lombana, 2017). Lo anterior es una breve explicación de un tipo de modelo autorregresivo. De acuerdo con los aportes del artículo, en la fase de modelado se toman las recomendaciones de modelos autorregresivos dinámicos y de corto plazo.

4 Objetivos

4.1 Objetivo general

Desarrollar un producto de datos que permita a la industria manufacturera de empaques flexibles (coberturista) la realización de un análisis descriptivo y predictivo de precios en el mercado del polipropileno.

4.2 Objetivos específicos

- Definir las variables de mayor impacto en el mercado mundial de polipropileno
- Construir un conjunto de datos que permita analizar y modelar los precios del polipropileno
- Construir una herramienta para análisis descriptivo y predictivo de los precios de mercado del polipropileno

•

5 Metodología

La metodología utilizada para cumplir los objetivos planteados se desarrolla en fases de acuerdo con el alcance de los objetivos específicos.

Se aplicó de forma parcial la metodología ASUM-DM, ya que es una metodología adecuada y ampliamente reconocida para la solución de este tipo de problemas. Esta consiste básicamente en cinco etapas (Alianza Caoba, 2018), las fases correspondientes a comprensión del negocio, enfoque analítico y captura de los datos se abordaron en los capítulos 1 y 2. El entendimiento, modelamiento y evaluación se aborda en los apartes siguientes. Previo al abordaje de las demás fases, se plantea una breve explicación de cada fase de la metodología:

“Comprensión del Negocio: Esta primera etapa establece las bases para abordar de forma exitosa el proyecto ya que se orienta al entendimiento de un negocio, con el fin de definir proyectos y soluciones alienados con las estrategias de las organizaciones.

Enfoque Analítico: Implica traducir el problema de negocio a un problema técnico. Una vez que el problema de negocio de la empresa ha sido claramente establecido, el científico de datos puede definir el enfoque analítico para resolver el problema. Esta etapa implica expresar el problema en el contexto de las técnicas estadísticas y de aprendizaje automático, para que la organización pueda identificar y seleccionar las más adecuadas para el proyecto.

Requisitos de los datos: La elección del enfoque analítico determina los requisitos de los datos, ya que los métodos analíticos que se utilizan requieren un contenido de datos, formatos y representaciones particulares, guiados por el conocimiento del dominio.

Recolección de los datos: El científico de datos identifica y reúne los recursos de datos necesarios, relevantes para el dominio del problema. Al encontrar brechas en la recopilación de datos, el científico de datos podría necesitar revisar los requisitos de datos y recopilar más información.

Entendimiento de los datos: Técnicas de visualización o estadística pueden ayudar a un científico de datos a comprender el contenido de los datos, evaluar su calidad y tener hallazgos iniciales de interés para el proyecto.

Preparación de los datos: La etapa de preparación de datos comprende aquellas actividades para construir el conjunto de datos que se utilizará en la etapa de modelado. Estos incluyen la limpieza de datos y otras técnicas de análisis de datos para satisfacer la necesidad de tener un conjunto robusto para la construcción de modelos apropiados para abordar el problema.

Modelado: A partir de la primera versión del conjunto de datos preparado, la etapa de modelado se centra en el desarrollo de modelos predictivos de acuerdo con el enfoque

analítico previamente definido. Con modelos predictivos, los científicos utilizan un conjunto de formación (datos históricos en los que se conoce el resultado del interés) para construir el modelo. Esta etapa suele ser iterativa, lo que lleva a refinamientos en la preparación de los datos y la especificación del modelo.

Evaluación: El científico de datos evalúa el modelo para entender su calidad y asegurarse de que aborda adecuada y completamente el problema del negocio. La evaluación del modelo implica el cálculo de diversas medidas de diagnóstico. Despliegue: Después de que se ha desarrollado un modelo con resultados satisfactorios en su evaluación, se despliega en el entorno de producción o en un entorno de prueba comparable.

Realimentación: Al recolectar los resultados del modelo implementado, la organización obtiene realimentación sobre el rendimiento del modelo y observa cómo afecta su entorno de despliegue” (Alianza Caoba, 2018).

5.1 Discusión y modelado

En los siguientes apartes se plantea una discusión desde la selección de variables e inventario de modelos. Donde de acuerdo con los resultados y alcance del trabajo se seleccionan los de mejor desempeño.

Se aborda en primera instancia el análisis de filtro de datos que permite hacerse alguna idea sobre las variables explicatorias de la transmisión de precios. En segunda instancia, se hacen varios análisis que permiten hacer conclusiones sobre la estacionalidad y tendencia de la serie de tiempo a explicar; lo anterior permite enrutar el camino de modelado. Acto seguido, se plantea un análisis de causalidad con el fin de cotejar las variables seleccionadas mediante el análisis de filtro y adicionalmente, en el análisis de causalidad, se determina el rezago de transmisión de los precios entre la variable a explicar y las variables explicatorias.

Frente al modelado, con base en las conclusiones resultantes de los procesos mencionados en el párrafo anterior, se construyen una serie de posibles modelos y se discute sobre su pertinencia, para posteriormente hacer una selección de los mejores modelos y rechazar otros. Lo anterior con el fin de generar una herramienta que contiene una combinación de pronósticos para soportar la decisión de compras de polipropileno.

5.1.1 Análisis de filtro

El primer acercamiento con las variables exógenas fue aplicar métodos de filtrado. En primera instancia se calcula la matriz de correlación del Market Basket versus el resto del inventario de datos.

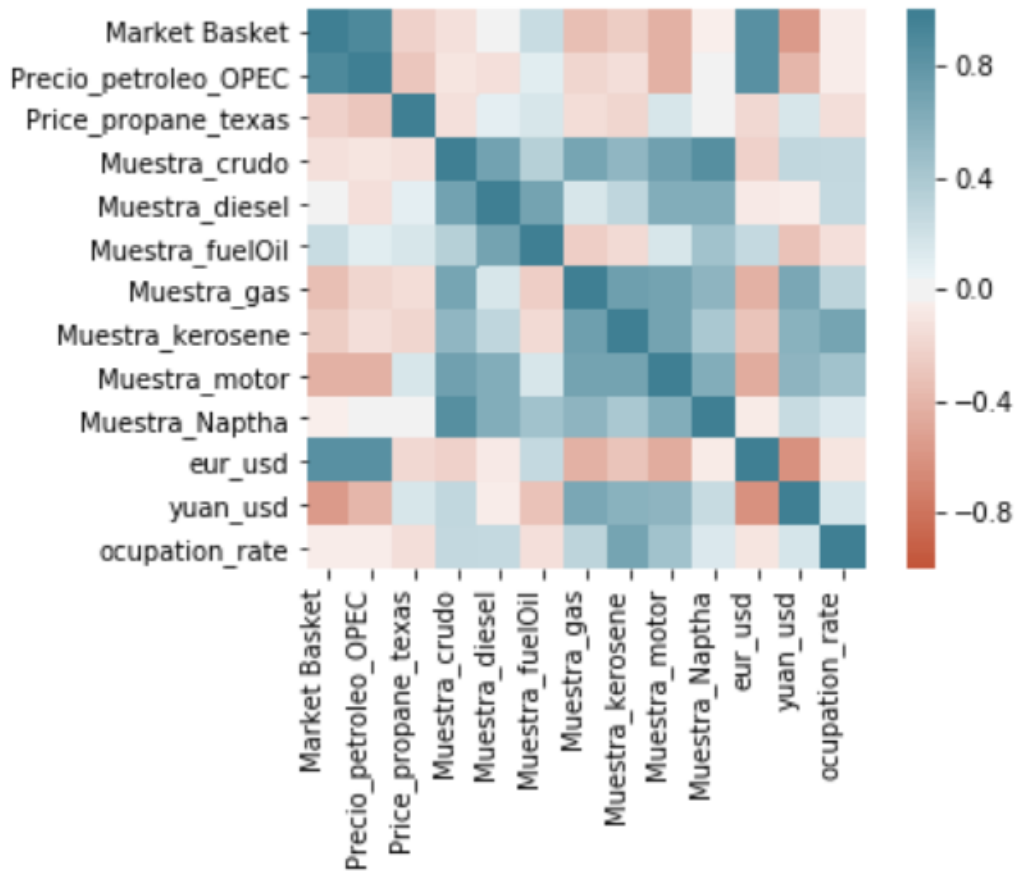


Figura 5-1. Matriz de correlación Market Basket Vs inventario de datos

Fuente: Elaboración propia

De la gráfica anterior, se extraen el siguiente ranking de correlaciones:

Ranking de Correlaciones	
Serie de tiempo	Coefficiente de correlación
Precio de Petróleo OPEC	90,93%
Refinación Fuel Oil Canadá	63,27%
Refinación Motor Arabia	54,29%
Eur/ USD	85,27%
Yuan / USD	55,74%

Figura 5-3. Top 5 de mejores correlaciones

Fuente: Elaboración propia

Sobre el resultado del análisis de correlaciones, vale la pena mencionar que las variables seleccionadas tienen relación con el polipropileno, entendiendo que este es un derivado del petróleo y que tiene en términos transaccionales relación con los mercados del dinero.

5.1.2 Análisis de estacionalidad y tendencia de la variable dependiente

Es común escuchar en el medio que el polipropileno contiene un componente importante de estacionalidad. Sin embargo, no se ha encontrado un análisis público al respecto, por lo tanto, en la fase de comprensión de los datos se pasó la serie por una cantidad importante de test estadísticos que permitieran algún grado de entendimiento sobre la misma.

Una serie temporal es estacionaria si la media y la varianza se fijan entre dos puntos equidistantes. Es decir, no importa dónde tome sus observaciones, los resultados deben ser los mismos. Una serie de tiempo que muestra la estacionalidad no es estacionaria.

“Una prueba de estacionariedad generalmente implica una prueba de hipótesis de raíz unitaria, donde la hipótesis nula H_0 es que la serie no es estacionaria y contiene una raíz unitaria. La hipótesis alternativa H_1 apoya la estacionariedad. Las pruebas aumentadas por Dickey-Fuller and Kwiatkowski-Phillips-Schmidt-Shin son pruebas de estacionariedad”. Traducción libre de (Portilla, 2019).

De acuerdo con el resultado que se presenta en la figura 5-4, esta serie presenta un componente mínimo de estacionalidad. No obstante, se realiza de acuerdo con la sugerencia de (Portilla, 2019), la prueba aumentada de Dickey-Fuller, según la cual se concluye de nuevo que la serie no es estacionaria, tal como se muestra en la figura 5-5.

Finalmente, para tomar una conclusión frente a la estacionalidad y tendencia de la serie, se realizó una descomposición manual basada en meses y trimestres agrupados y separados por año como se muestra en la figura 5-6. Dicha prueba, además, sirve como insumo para soportar decisiones de compra, toda vez que permite al usuario determinar con base en la historia y su criterio experto, más allá de pruebas estadísticas, si la serie tiene alguna incidencia el factor estacional.

Como resultado del análisis planteado en la figura 5-6, surge una inquietud adicional en cuanto al comportamiento de la serie, esta parece ser de regresión a la media, hipótesis que se comprueba en la figura 5-7 donde se observa como la media en cada mes del año es poco distante entre sí durante los 6 años de observación. En conclusión, puede afirmarse que la serie no es estacionaria y presenta un comportamiento de regresión a la media.

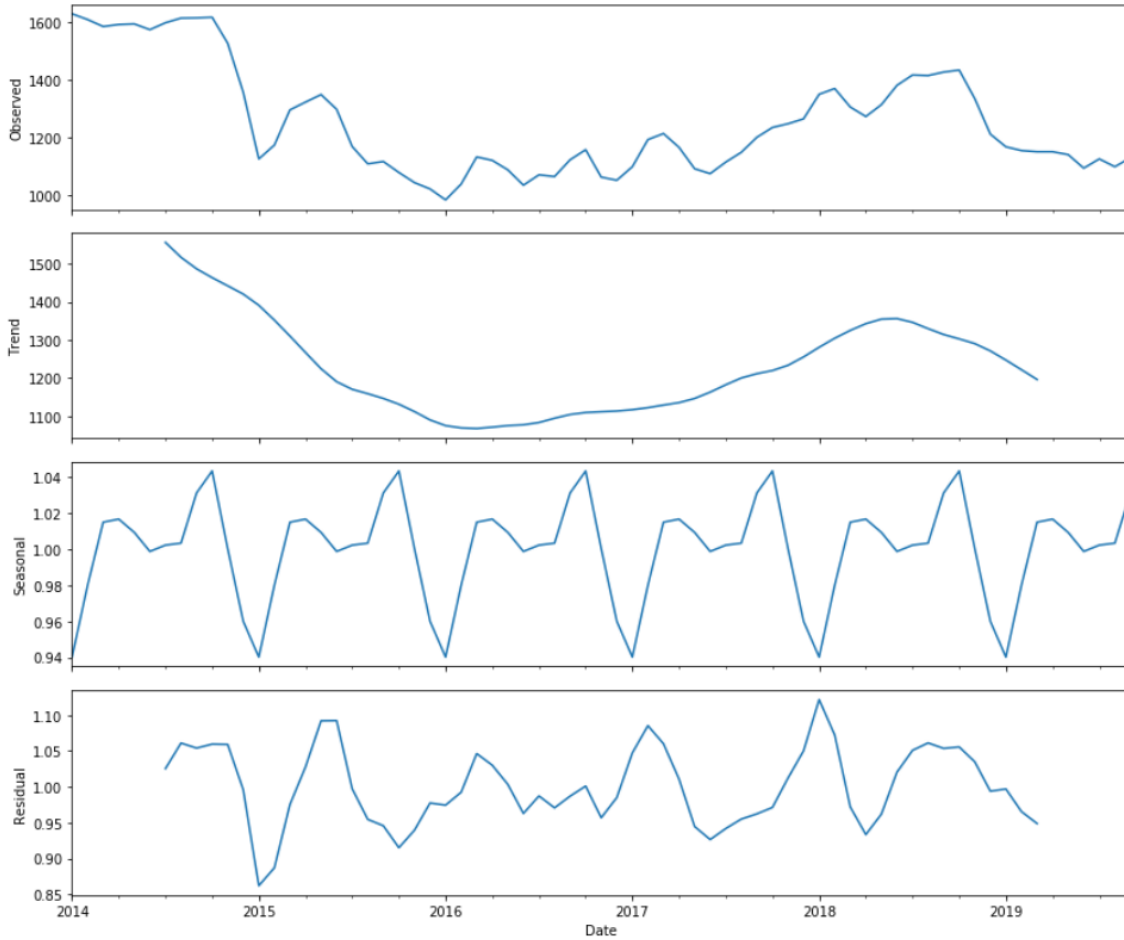


Figura 5-4. Descomposición de la variable dependiente

Fuente: Elaboración propia

```
Augmented Dickey-Fuller Test:  
ADF test statistic      -2.123245  
p-value                0.235271  
# lags used            2.000000  
# observations         66.000000  
critical value (1%)   -3.533560  
critical value (5%)   -2.906444  
critical value (10%)  -2.590724  
Weak evidence against the null hypothesis  
Fail to reject the null hypothesis  
Data has a unit root and is non-stationary
```

Figura 5-5. Resultado de Prueba aumentada de Dickey Fuller

Fuente: Elaboración propia

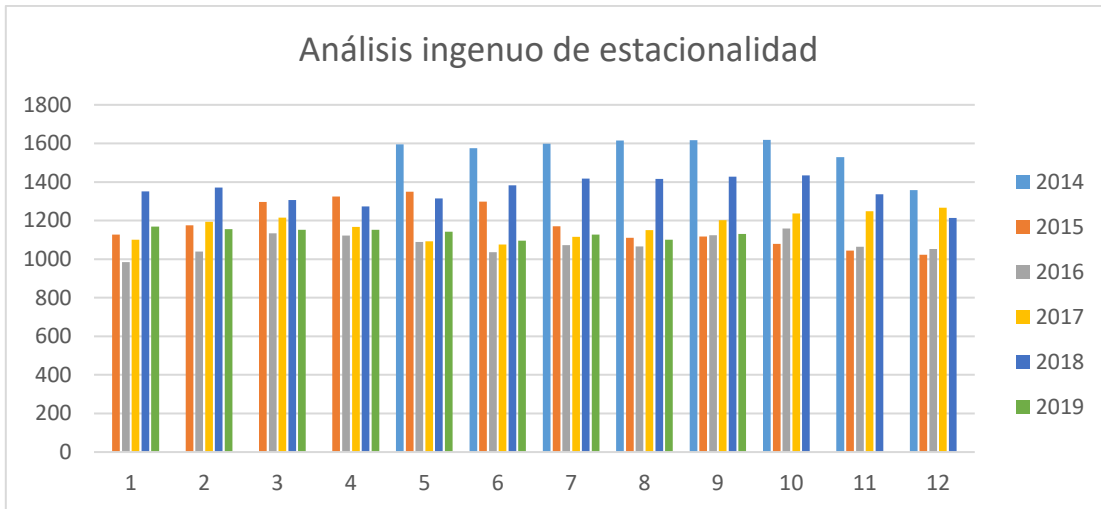


Figura 5-6. Análisis ingenuo de estacionalidad
Fuente: Elaboración propia

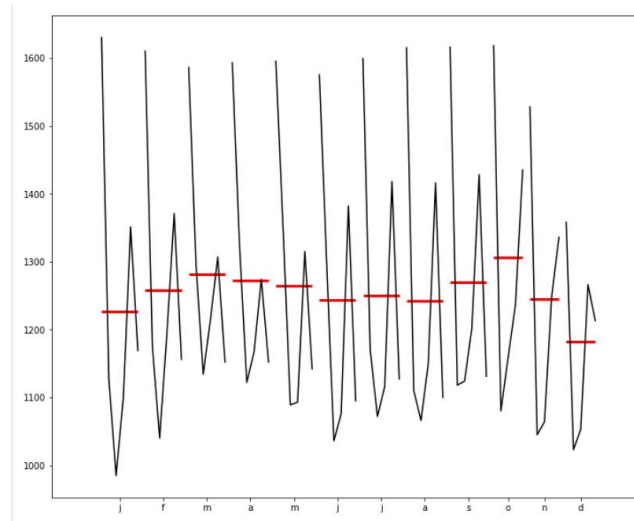


Figura 5-7. Análisis de media mensual
Fuente: Elaboración propia

5.1.3 Análisis de causalidad

Como se mencionó en la revisión de la literatura, de acuerdo con (González, Ávila, Gitter, & Ángel, 2011), es necesario determinar el rezago que tiene la variable dependiente en términos de la transmisión que se da desde las variables exploratorias, en este sentido :

“[l]a prueba de causalidad Granger es una prueba de hipótesis para determinar si una serie de tiempo es útil para pronosticar otra. Si bien es bastante fácil medir las correlaciones entre series (cuando una sube, la otra sube, y viceversa), otra cosa es observar los cambios en una serie correlacionados con los cambios en otra después de un período de tiempo constante. Esto puede indicar la presencia de causalidad, que los cambios en la primera serie influyeron en el comportamiento de la segunda. Sin embargo, también puede ser que ambas series se

vean afectadas por algún tercer factor, solo a tasas diferentes. Aun así, puede ser útil si los cambios en una serie pueden predecir los próximos cambios en otra, ya sea que haya causalidad o no. En este caso decimos que una serie Granger causa a otra” Traducción libre de (Portilla, 2019).

En el caso de dos series, y y x , la hipótesis nula es que los valores rezagados de x no explican las variaciones en y . En otras palabras, supone que xt no causa Granger yt .

En términos generales, después de aplicar el test a las variables explicatorias determinadas por el método de filtro, se puede decir que el retraso de la serie es de 3 meses. Los mejores resultados de la prueba lo ofrecen EUR-USD, si bien esta variable había sido rechazada en el análisis de filtro por su alta correlación con el precio del petróleo, se reintegra posterior a este análisis en la medida que además de correlación, presenta una causalidad de acuerdo con la prueba de Granger. También aportan a la causalidad los niveles de producción de crudo y los precios de la OPEC.

Toma por sorpresa, el poco poder explicatorio de la tasa de ocupación de las refinerías, no obstante, en la fase de experimentación se toma como variable exógena y se le realizan algunas transformaciones intentando evidenciar algún poder explicatorio de esa variable, que por análisis técnico y su relación con la industria debería explicar en gran medida variaciones de precios.

El UR-USD tiene mayor explicatorio causal que incluso los precios del petróleo. Teniendo en cuenta el rezago de tres meses que muestra el test, y teniendo en cuenta la ventana de pronóstico que se plantea para resolver este problema, no sería necesario desarrollar un modelo de pronóstico de moneda, ya que, según el test, los datos actuales de la moneda tienen poder explicatorio de 3 meses posteriores sobre el PP, satisfaciendo el alcance en la necesidad del pronóstico.

5.2 Modelación y evaluación

5.2.1 ARIMA dinámico

Los modelos ARIMA son modelos Autorregresivos Integrados de promedio Móvil, originalmente en inglés *Autoregressive Integrated Moving Average*, es un modelo que se basa en las variaciones de la serie en si misma con el fin de encontrar patrones que permitan realizar predicciones basadas en el pasado de la serie (Portilla, 2019), es decir, un modelo de pronóstico univariable.

En un modelo ARIMA(p,d,q), sus parámetros son enteros no negativos que indican las componentes del modelo así: AR(p) Autorregresiva, I(d) Integrada, MA(q) de medias móviles. Sobre la selección de parámetros (Duke University, 2019) ofrece un documento en línea que permite realizar una selección de parámetros adecuado, en este se explica de acuerdo con el comportamiento de los datos y la fundamentación estadística y matemática la selección más apropiada de parámetros según las condiciones de la serie objeto de estudio, aspectos realmente importantes desde la formalidad científica pero que desde el enfoque analítico pueden abordarse con técnicas basadas en la fuerza de cómputo, como se explica en el párrafo siguiente.

La experimentación en campo de este modelo fue más extensa de lo que en este escrito se plantea, esta contó con una fase de optimización, en la que se optó por programar un algoritmo que iterara

el modelo en todas las combinaciones posibles y se hiciera una selección del mejor modelo de acuerdo a las métricas de error más adecuadas, en este caso el menor AIC de todas las ejecuciones realizadas, lo anterior, con base en los aportes de (Sergeev, 2019).

Esta forma de selección agrega valor, además, en la medida que para corridas posteriores y en ambientes productivos del modelo se garantizará que la selección de hiperparámetros será dinámica en la medida que la serie presente cambios en el tiempo. En el caso puntual de la serie planteada, el modelo se itera 54 veces antes de arrojar como resultado la combinación de hiperparámetros que minimiza el error, tal como se muestra en las dos figuras siguientes (5-8 y 5-9).

```
In [*]: 1 result_table = optimizeSARIMA(parameters_list, d, D, s)
19% ██████████ 10/54 [00:03<00:16, 2.59it/s]
```

Figura 5-8. Muestra de ejecución iterativa del modelo, iteración 10 de 54
Fuente: Captura de pantalla a código del estudio

```

Statespace Model Results
=====
Dep. Variable:          Market Basket   No. Observations:          68
Model:                 SARIMAX(2, 1, 2)x(2, 1, 0, 3)   Log Likelihood             -345.813
Date:                  Sat, 21 Dec 2019   AIC                        705.625
Time:                  14:14:30   BIC                        720.738
Sample:                05-31-2014   HQIC                       711.579
                    - 12-31-2019
Covariance Type:      opg
=====
              coef    std err          z      P>|z|      [0.025    0.975]
-----
ar.L1          0.8861    0.076     11.622    0.000     0.737     1.036
ar.L2         -0.8926    0.085    -10.526    0.000    -1.059    -0.726
ma.L1         -0.3474    0.128     -2.704    0.007    -0.599    -0.096
ma.L2          0.3066    0.180     1.706    0.088    -0.046     0.659
ar.S.L3       -0.4450    0.158     -2.815    0.005    -0.755    -0.135
ar.S.L6       -0.6173    0.099     -6.259    0.000    -0.811    -0.424
sigma2        2675.7007   494.806     5.408    0.000   1705.898   3645.503
=====
Ljung-Box (Q):                29.76   Jarque-Bera (JB):                1.30
Prob(Q):                      0.88   Prob(JB):                        0.52
Heteroskedasticity (H):       0.66   Skew:                            -0.24
Prob(H) (two-sided):          0.35   Kurtosis:                        3.50
=====

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
    
```

Figura 5-9. Extracto del mejor modelo, resultado del proceso iterativo
Fuente: Captura de pantalla a código del estudio

Teniendo en cuenta que el horizonte de pronóstico del trabajo es de 3 meses, se plantea un modelo donde el modelo recibe información hasta septiembre de 2019 y se espera que arroje valores de pronóstico para los meses faltantes del 2019. En la Figura 5-10 se analiza el comportamiento del modelo en el conjunto de datos de prueba y arroja un RMSE de 3.49%, adicionalmente se calculó un pronóstico extendido que posteriormente fue cruzado con datos reales desconocidos por el modelo como se demuestra en la figura 5-11, donde se evidencia un error promedio del modelo en productivo del 5%.

```

2019-10-31    1174.524609
2019-11-30    1106.203877
2019-12-31    1042.752246
2020-01-31    1023.103589
dtype: float64
    
```

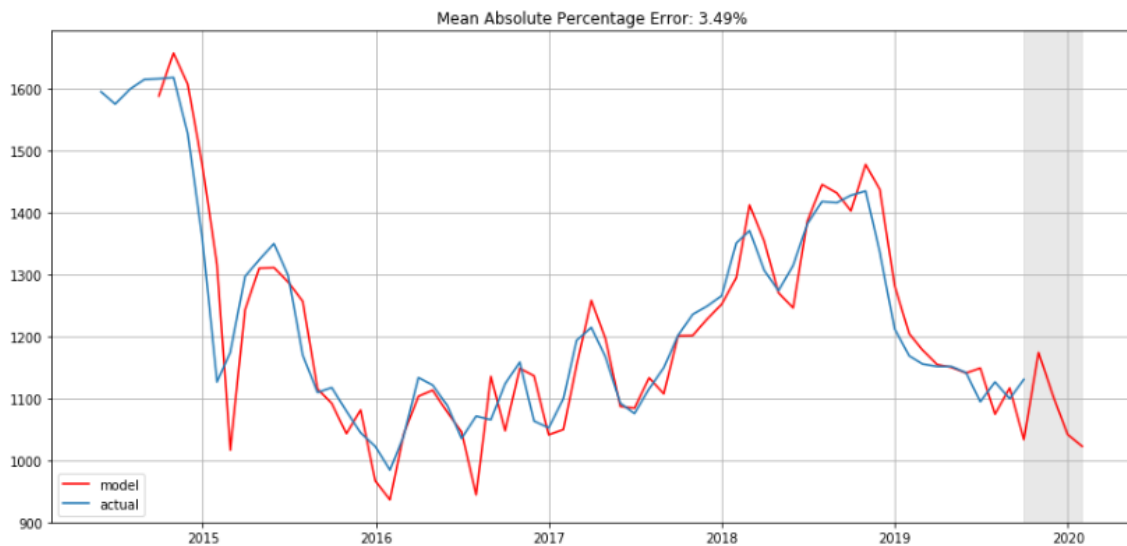


Figura 5-10. Resultado y contraste del conjunto de prueba y pronóstico extendido

Fuente: Captura de pantalla a código del estudio

ARIMA TEST				
	Real	Forecast	Error absoluto	Error %
Octubre 2019	1.102	1174	72	7%
Noviembre 2019	1.028	1106	78	8%
Diciembre 2019	1.040	1042	2	0%
			Error medio	5%

Figura 5-11. Resultado y contraste pronóstico extendido cruzado con datos reales desconocidos por el modelo

Fuente: elaboración propia

5.2.2 ARIMAX dinámico

El modelo ARIMAX es en principio una variedad del ARIMA, pero que acepta variables exógenas, es decir, es un modelo que además de los componentes autorregresivos (AR), Integrados (I), de medias móviles (MA), captura otros estímulos de variables exploratorias exógenas a la serie (X), de ahí su denominación ARIMAX (Portilla, 2019).

En esta versión se aplicaron diversas combinaciones de las variables de las capturadas, sin embargo, se hizo énfasis en las seleccionadas mediante los métodos de selección de características previamente explicados y considerando facilidad y practicidad para la captura de datos en escenarios productivos y las mejoras en cuanto a los resultados que se presentaban en cada una de las versiones. Un apunte interesante es que, a lo largo de la construcción, se encontró que las únicas transformaciones que podían agregar poder predictivo al modelo eran las de tipo retrasado, elemento consecuente con el análisis de Granger planteado previamente, en el que se demuestra que la serie presenta un rezago de entre 2 y 3 meses a los estímulos externos determinados en la etapa de filtrado.

Para explicar esta experimentación, se muestra el error cuadrático medio (RMSE) de las combinaciones de variables exógenas que presentaron mejor ajuste.

	Variables exógenas incorporadas al modelo¹	RMSE Train	RMSE Test
Modelo 1	Precio petróleo países OPEC (lag3)², EUR/USD (lag3), Refinación de Fuel Oil en Canadá (lag3), Producción de Gas Natural en Rusia (lag3), Refinación de gasolina tipo Motor en Arabia Saudita (lag3), Muestra de Diesel (lag3)³, Muestra de crudo (lag3)	3,5%	2,4%
Modelo 2	Precio petróleo países OPEC, EUR/USD, Tasa de ocupación de las refinerías en USA, Precio petróleo países OPEC (Lag2), EUR/USD (lag2), Tasa de ocupación de las refinerías en USA (lag2), Precio petróleo países OPEC (Lag3), EUR/USD (lag3), Tasa de ocupación de las refinerías en USA (Lag3)	3,0%	9,3%
Modelo 3	Precio petróleo países OPEC, EUR/USD, Refinación de Fuel Oil en Canadá, Producción de Gas Natural en Rusia, Refinación de gasolina tipo Motor en Arabia Saudita, Muestra de Diesel, Muestra de Crudo	3,0%	4,3%
Modelo 4	Precio petróleo países OPEC, EUR/USD	3,1%	4,0%

Figura 5-12. Contraste de modelos en modelo ARIMAX

¹ En el aparte dedicado a la captura de datos se encuentra con mayor detalle la descripción de las variables

² 'Lag(n)' quiere expresar que es la variable con un retraso *n* meses

³ Las variables que contienen la palabra '*muestra*' corresponden a variables calculadas a partir de la sumatoria del dato de los países disponibles, mayor ampliación en el capítulo de captura de datos

Fuente: Elaboración propia

Es necesario aclarar que la metodología para la selección de hiperparámetros es la misma basada en fuerza de cómputo que se trató en la experimentación ARIMA bajo las recomendaciones de (Sergeev, 2019).

De la experimentación sobre ARIMAX planteada en la figura 5-12 vale la pena resaltar los siguientes aspectos:

- El modelo 1 tienen un ajuste bastante bueno, deja en evidencia la utilidad de la prueba de causalidad de Granger, puesto que de acuerdo con sus resultados queda demostrado que el mejor rezago es de 3 meses.
- Genera tranquilidad en la medida que la ventana de pronóstico es de tres o cuatro meses, al trabajar con datos previos de 3 meses, implica que para el modelo no se deben desarrollar submodelos de pronóstico adicionales de las variables exploratorias y el ruido que esto implica, bastaría con tomar los datos de mercado disponibles, generando un mayor grado de administración del modelo en un ambiente productivo empresarial no académico.
- Del modelo 2 se observa que al incorporar las tasas de ocupación de las refinerías se dispara el error, corroborando lo planteado en la fase de selección de características y contradiciendo fuertemente cualquier tipo de análisis de negocio a priori, pues supondría cualquier experto en la materia que la ocupación debería tener poder explicatorio. Un problema de investigación futuro podría ser el entendimiento de por qué la tasa de ocupación no tiene poder predictor en este caso, yendo en contravía de las conversaciones usuales con los expertos de mercado
- Del modelo 4 sorprende que es un modelo sumamente básico con solamente dos variables exógenas de fácil captura que presenta un resultado muy bueno, pero no tanto como el modelo 1.

Apreciación final de los modelos ARIMAX: la modelación con el modelo 1 es la que mejor calidad presenta y la que se selecciona para realizar posteriormente un acuerdo entre modelos como herramienta que sirva de apoyo y soporte a la decisión.

5.2.3 Holt Winters

Los métodos Holt Winters son métodos basados en el suavizamiento exponencial, el método Holt considera el doble suavizamiento exponencial y el Holt Winters el triple, de este modo se considera la estacionalidad y tendencia de la serie. Este método considera dos formas de cálculo dependiendo de la estructura de la serie, pueden ser aditivos o multiplicativos (Maguiña, 2016).

Este tipo de modelos son una derivación de los modelos EWMA (Exponentially Weighted Moving Average) (Portilla, 2019). Básicamente aparece un elemento suavizamiento adicional que en la siguiente ecuación puede verse como β

$$\begin{aligned}
 l_t &= (1 - \alpha)l_{t-1} + \alpha x_t, && \text{level} \\
 b_t &= (1 - \beta)b_{t-1} + \beta(l_t - l_{t-1}) && \text{trend} \\
 y_t &= l_t + b_t && \text{fitted model} \\
 \hat{y}_{t+h} &= l_t + hb_t && \text{forecasting model } (h = \# \text{ periods into the future})
 \end{aligned}$$

Figura 5-13. Representación matemática del doble suavizamiento exponencial
Fuente: (Portilla, 2019)

En cuanto al triple suavizamiento exponencial aparece un tercer factor denominado γ que representa la estacionalidad de la serie

$$\begin{aligned}
 l_t &= (1 - \alpha)l_{t-1} + \alpha x_t, && \text{level} \\
 b_t &= (1 - \beta)b_{t-1} + \beta(l_t - l_{t-1}) && \text{trend} \\
 c_t &= (1 - \gamma)c_{t-L} + \gamma(x_t - l_{t-1} - b_{t-1}) && \text{seasonal} \\
 y_t &= (l_t + b_t)c_t && \text{fitted model} \\
 \hat{y}_{t+m} &= (l_t + mb_t)c_{t-L+1+(m-1)\text{mod}L} && \text{forecasting model } (m = \# \text{ periods into the future})
 \end{aligned}$$

Figura 5-14. Representación matemática del triple suavizamiento exponencial
Fuente: (Portilla, 2019)

En cuanto la experimentación planteada se realizaron algunas pruebas para entender el efecto de los hiperparámetros en la serie $(\alpha \beta \gamma)$, sin embargo, su estimación se hizo mediante fuerza de cómputo, se utilizó una función de tipo ‘Grid Search’ de la librería Scikit-learn para Python que mediante procesos iterativos permite encontrar los parámetros de mejor ajuste del modelo, esto garantiza además que el modelo en productivo siempre cuente con la mejor selección de parámetros.

Adicionalmente se implementó una validación cruzada que permitiera dividir el conjunto de datos en 5 particiones, logrando darle mayor precisión al modelo, entendiendo además que las observaciones que se tienen del fenómeno estudiado son pocas.

```

0.1772277469767815 0.03580922821770621 0.7356294040032191
Wall time: 9.18 s
    
```

Figura 5-15. Estimación de los mejores α, β, γ respectivamente
Fuente: Captura de pantalla del código del estudio

68 1043.283894
 69 1050.184270
 70 1043.576665
 71 1050.477040

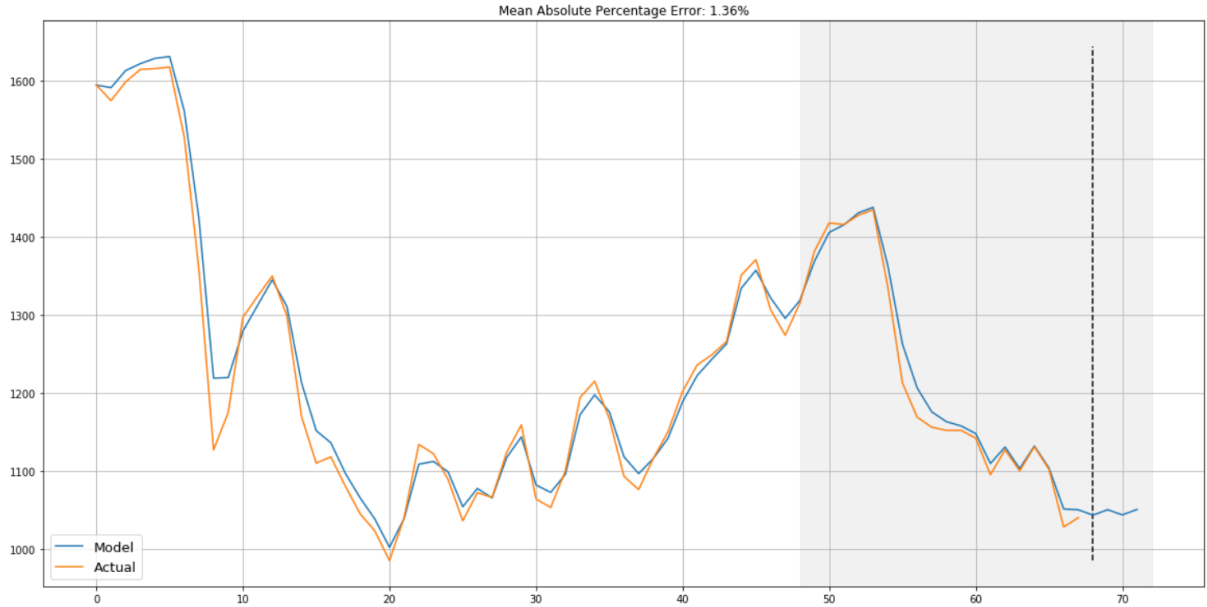


Figura 5-16 Conjunto de entrenamiento, prueba y pronóstico extendido del modelo
 Fuente: Captura de pantalla del código del estudio

Test de pronóstico extendido Holt Winters				
Periodo	Real	Forecast	Error abs	% Error
Octubre 2019	1.102	1043	59	5,4%
Noviembre 2019	1.028	1050	22	2,1%
Diciembre 2019	1.040	1042	2	0,2%
			Error medio	2,6%

Figura 5-17 Conjunto de entrenamiento, prueba y pronóstico extendido del modelo
 Fuente: elaboración propia

De este modelo es válido mencionar algunas conclusiones: se trata de un modelo de fácil implementación y muestra su utilidad en el ambiente productivo después de los ejercicios de pronóstico extendido y posterior contraste con los datos reales de mercado. Es además un modelo univariable que facilita la captura de los datos, la desventaja podría ser la demora en capturar estímulos externos procedentes de variables relacionadas del polipropileno.

5.2.4 Regresión lineal, Lasso y XGBoost

La regresión lineal utilizada, es un algoritmo de aprendizaje no supervisado que permite hacer pronósticos basados en la relación entre una variable dependiente y variables independientes (Portilla, 2019). Para la ejecución de este modelo se tomaron algunas variables de tipo Dummy, basadas en las implementaciones sugeridas por (Sergeev, 2019):

- Se crearon variables correspondientes a 11 retrasos

- Se crearon variables correspondientes al mes, y trimestre, si era comienzo o final de trimestre, adicionalmente otra que considerara si era final del año

	lag_2	lag_3	lag_4	lag_5	lag_6	lag_7	lag_8	lag_9	lag_10	lag_11	Month	Quarter	is_quarter_end	is_year_end
Date														
2015-04-30	1175.0	1127.0	1358.0	1528.0	1618.0	1616.0	1615.0	1599.0	1575.0	1595.0	4	2	False	False
2015-05-31	1297.0	1175.0	1127.0	1358.0	1528.0	1618.0	1616.0	1615.0	1599.0	1575.0	5	2	False	False
2015-06-30	1324.0	1297.0	1175.0	1127.0	1358.0	1528.0	1618.0	1616.0	1615.0	1599.0	6	2	True	False

Figura 5-17 Conjunto de datos con variables tipo Dummy

Fuente: captura de pantalla del código del estudio

Después del ejercicio previamente planteado se corrió el modelo con el conjunto de datos completos, generando un RMSE de 5,28%. Buscando mejorar el resultado, se analizó del conjunto de datos que variables podían tener mayor poder expiatorio, en este análisis se encontró que final de año, final de trimestre, trimestre y mes son las variables de mayor poder predictivo, sin embargo, el RMSE solamente cambia a 5,16%. Vale la pena mencionar que estas experimentaciones también contaron con validación cruzada.

Se experimentó además con una regresión de tipo Lasso donde se observó que el error se disparó, alcanzando un RMSE del 18,74%; por lo que se descarta.

Adicionalmente se implementó una experimentación de tipo XGBoost que arroja un RMSE de 8,99% por lo que se concluye que los modelos planteados en el presente aparte no generan mejoras importantes a la solución del problema, como si se nota en los modelos planteados anteriormente, en tanto que son fácilmente superados por los modelos ingenuos que se plantean a continuación como modelos complementarios.

5.2.5 Modelos a priori como complementos y soporte de la decisión

No se pueden dejar de lado los modelos sencillos y de bajo costo, muchas veces estos pueden generar buenos pronósticos sin ningún desgaste, los modelos que se planteen desde la analítica deben superar este tipo de modelos, por lo que se plantean algunas combinaciones de modelos a priori que serán complementarios a la familia de modelos que se abordarán en el acuerdo entre modelos. A continuación, se plantean dos modelos a priori denominados MA3 y WMA (6,3,1)

MA3: es el pronóstico calculado como el promedio móvil de los últimos 3 meses. Este presenta un RMSE del 7.3%.

WMA (6,3,1): es el pronóstico calculado como el promedio móvil de los últimos 3 meses pero dando un mayor peso al último dato: 60%, 30% al segundo y 10% al tercero, este genera un RMSE de 4,21%

5.3 Conclusiones de modelado y análisis de resultados de la fase de modelado

Modelo	RMSE en productivo
AIRMA	5,0%
ARIMAX 1	2,4%
ARIMAX 3	4,0%
Holt Winters	2,6%
MA3	7,3%
WMA(6,3,1)	4,2%
Media de errores	4,3%

Figura 6-1 Errores de los modelos experimentados

Fuente: elaboración propia

De acuerdo con la Tabla 6-1, el modelo ARIMAX del modelo 1 es el que mejor ajuste presenta, seguido del Holt Winters, recuérdese que se descartaron los que dieron errores considerables.

Se incluyen además los modelos a priori ya que no se puede dejar de lado lo práctico y simple, estos podrían llegar a ser en algún momento un indicador de alerta sobre los desajustes que podrían tener los modelos en el tiempo. Con el fin de mejorar los procesos decisionales de los usuarios, se plantea que, para los análisis predictivos en campo, deban usarse los modelos planteados en la tabla anterior, permitiendo observar posibles acuerdos entre los modelos e impedir que los usuarios se sesguen con uno solo, esto teniendo en cuenta los riesgos que representa tomar una mala decisión en términos del capital de trabajo neto operativo

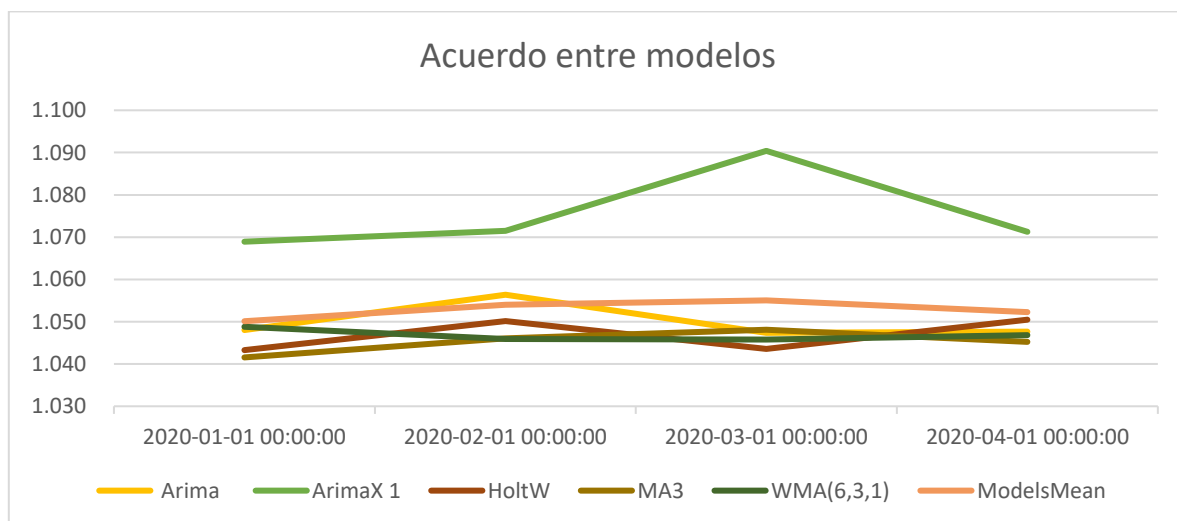


Figura 6-2 Acuerdo entre modelos

Fuente: elaboración propia

Como se observa en la gráfica anterior, los pronósticos coinciden en que el polipropileno calculado En el Market Basket se moverá en el primer trimestre de 2020 entre 1040 y 1090, se prevé cierta estabilidad en los mercados situación que representa lo que perciben otros analistas de mercado. Con corte a enero los mercados reportan en promedio un valor de 1035 para la WCSA (West Coast South America (S&P, 2020), por lo tanto se concluye que pese a que la experimentación se realizó en el último trimestre del 2019, los resultados arrojados son consistentes con las dinámicas actuales de mercado. El modelo es usado, evaluado y actualizado periódicamente en ambientes productivos.

Vale la pena mencionar que el modelo que más de aleja del consenso, es el ARIMAX1, recuérdese que éste se alimenta de variables exógenas con un rezago de 3 meses por lo que podría estar captando estímulos externos que lo llevan a distanciarse un poco de los demás, sin embargo la distancia es corta (entre 10 y 30 dólares), en este sentido los análisis en ambientes reales deberían considerar un margen de posibilidades y adicionalmente ejercicios de simulación y creación que escenarios que son soportados por los modelos ARIMAX planteados anteriormente y que estarán al alcance de usuarios decisores, este permitiría responder preguntas de tipo ¿Qué pasaría si?

Para contrastar el resultado del mercado, se cruzan los valores obtenidos con los mercados del petróleo y del dinero, para el caso de Colombia con el USD COP. Los analistas de Bloomberg citados por (Banco de Bogotá, 2019) coincidían a finales del año 2019, que el año 2020 sería uno de calma y estabilidad en la medida en que se resolverían cierto tipo de conflictos como la guerra comercial US – China y el Brexit, los que han mantenido fuertes volatilidades en el año 2020 debido al riesgo que representan, los resultados de los modelos son consistentes con los pronósticos que tenía el modelo para ese momento.

Consenso Bloomberg – Petróleo Brent (USD x barril)

Consensus	Spot	As Of	Q4 19	Q1 20	Q2 20	Q3 20	Q4 20
Median		11/21/19	62.00	60.00	60.00	60.00	60.00
Mean		11/21/19	62.34	61.77	61.69	61.82	62.94
High		11/21/19	70.00	73.50	73.00	80.00	85.00
Low		11/21/19	58.00	57.50	54.00	53.80	55.00
Forward	62.57	11/29/19	62.36	62.13	60.95	60.11	59.43
Diff (Median - Curr)			-0.36	-2.13	-0.95	-0.11	+0.57

Consenso Analistas Bloomberg USDCOP

Region	Latin America	Currency	Colombian Peso	Q4 19	Q1 20	Q2 20	Q3 20
USDCOP							
Spot	3515	Median		3350	3385	3353	3325
Q3 19	Actual	Mean		3355	3352	3343	3355
	3478	High		3500	3475	3517	3567
Q3 19	Forecast	Low		3200	3175	3150	3150
	3400	Forward		3521	3539	3559	3582

Figura 6-3 Consenso de Analistas Bloomberg
Fuente: (Banco de Bogotá, 2019)

Teniendo en cuenta que las fases de testeo se presentaron en momentos diferentes, las conclusiones planteadas y que fueron contrastadas con el consenso de analistas de Bloomberg (Figura 6-3) puede diferir un poco de la Figura 6-4, en la medida que durante la realización del trabajo fueron apareciendo nuevos sucesos de impacto mundial que modifican la lógica de precios del mercado, es decir, el modelo reaccionó favorablemente a los cambios de información del mercado.

Durante el primer mes del año 2020 aparece el COVID-19, una epidemia de alcance mundial que ha tenido impactos importantes en lo que a precios de petróleo y dinero se refiere. Como se mencionó al principio del escrito, China es un jugador muy importante, en tanto que es el primer productor y consumidor de polipropileno del mundo. Adicionalmente la demanda mundial de los *commodities* se ha visto afectada por esta pandemia, los analistas anticipan una contracción en el consumo mundial, situación que se ven reflejada en la corrida posterior que se hizo del modelo, se observa como posiblemente derivado de esta crisis de orden mundial los precios tendrían incentivos para moverse a la baja, cambiando los fundamentos de mercado planteados por la mayoría de los analistas reputados del mundo a finales del año 2019. En la medida que cambien los fundamentales y el petróleo continúe con su tendencia a la baja, el polipropileno debería mostrar una tendencia a la reducción de precios, no obstante el coberturista puro colombiano tendrá problemas importantes en términos de la devaluación que traerá consigo la caída de los precios del petróleo.

Al cierre del presente trabajo, no han cambiado las condiciones de precios y abastecimiento en los fundamentales de formación de precio en Latinoamérica, en China por su parte se encuentra medianamente superada la emergencia y se reactiva el consumo.

Combinación de pronósticos con la información disponible

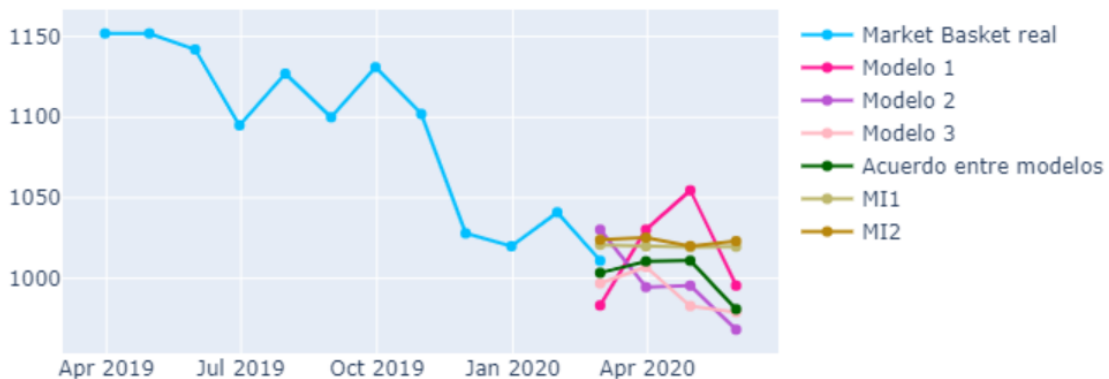


Figura 6-4 Acuerdo entre modelos
Fuente: elaboración propia

En síntesis, es válido mencionar que el modelo presenta resultados acordes con las nuevas realidades y dinámicas del mercado. El capítulo siguiente, explica el entregable del trabajo de cara a los usuarios finales.

6 Herramienta de soporte a la decisión

Es válido mencionar que en el presente trabajo se cumple con los objetivos que se plantea, en tanto que se logran definir las variables de mayor impacto en términos de precios del polipropileno, las captura y las procesa en un modelo de pronóstico que sirve como herramienta de ayuda y soporte a la decisión de los usuarios finales.

Los resultados del trabajo están disponibles en un repositorio Git Hub y reciben actualizaciones recurrentes, se recomienda descargarlo y correrlo en de forma local. Está publicada mediante libros bajo la herramienta Jupyter que permiten entender de una forma dinámica el código, así como la modificación por parte del usuario final si este lo considera pertinente. Esta herramienta viene siendo testeada de forma mensual desde septiembre de 2019 arrojando resultados positivos en ambientes productivos.

El modelo per se, no es una bola de cristal, es una herramienta de análisis que combinado con criterios expertos puede mejorar notablemente los procesos decisionales de una organización que sea intensiva en compas de polipropileno, en tanto que le permite mezclar el conocimiento del mercado con herramientas analíticas, logrando generar en el usuario un ambiente de información más completo. Es válido recordar de nuevo, que el supuesto fundamental de los mercados competitivos es que compradores y vendedores estén debidamente informados.

La herramienta de usuario final se aborda desde tres perspectivas: análisis histórico del polipropileno en los principales mercados del mundo, pronóstico y análisis de las variables explicatorias, este último sirve además como herramienta de calibración y ajuste futuro del modelo.

La primera gráfica le permite al usuario ver de forma actualizada el dato mensual de los precios en las plazas referentes más importantes, así como el Market Basket escogido, el cual podría ser revaluado en el momento que los usuarios decisores lo decidan.

Precios Internacionales del PP



Figura 6-5 Precios de mercado y Market Basket
Fuente: elaboración propia

Seguidamente, el usuario podrá ver el pronóstico individual arrojado por cada uno de los modelos aplicados, que fueron explicados en la fase de modelación de este texto, finalmente verá el pronóstico extendido arrojado por cada modelo, un acuerdo entre modelos que se calcula a partir de la combinación simple de pronósticos sugerida en el capítulo “1.6 la necesidad del pronóstico y las series de tiempo”. Se incluye, además, el pronóstico de dos modelos a priori que le permitirán hacerse un mejor panorama de lo que podrá ocurrir en términos de mercado.

Combinación de pronósticos con la información disponible

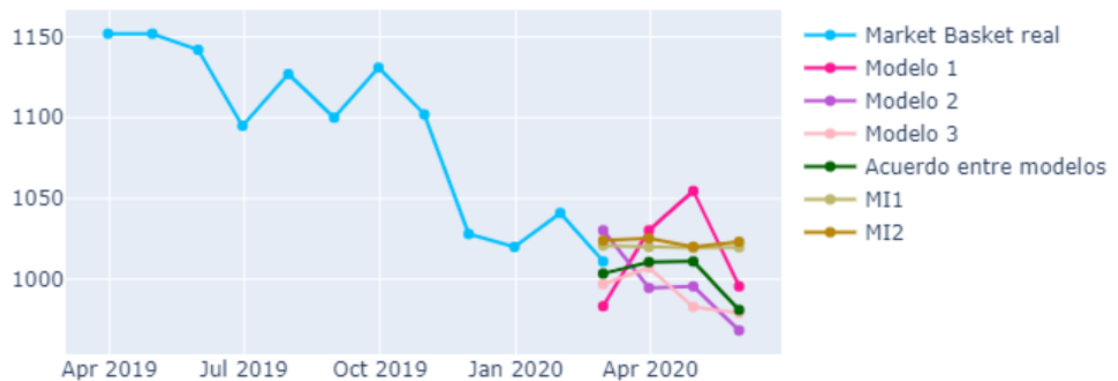


Figura 6-4 Acuerdo entre modelos
Fuente: elaboración propia

Como se muestra en la Figura 6-6, los usuarios pueden analizar en una sola imagen, el comportamiento de las variables explicatorias del modelo orientado en dos vertientes: mercado del dinero y determinantes de la oferta del polipropileno; lo anterior con dos fines, por un lado, complementar la información para las decisiones, de tal modo que el usuario pueda ubicar en el tiempo momentos similares al actual y concluir sobre los rumbos que ha tomado el mercado en esos momentos. Y por el otro, como herramienta de calibración del modelo que deberá someterse por lo menos a dos revisiones profundas al año.

En la Figura 6-7, se muestra de forma actualizada e interactiva el comportamiento del Market Basket versus el mercado del dinero y los precios del petróleo de los países OPEC, esto permite al usuario tomador de decisiones examinar el impacto de cada variable con el PP de una forma sencilla, permite por lo menos evaluar la dirección que siguen las variables y observar hacia el pasado momentos que presenten alguna similitud con el presente con el fin de tomar unas mejores decisiones. Es válido mencionar que, especialmente en el caso de las monedas, la fuerza de transmisión tiene magnitudes diferentes, sin embargo, permite identificar algunos patrones comunes en términos de la dirección que toman las variaciones. Es necesario resaltar que estos análisis serán también un insumo permanente de calibración de los modelos

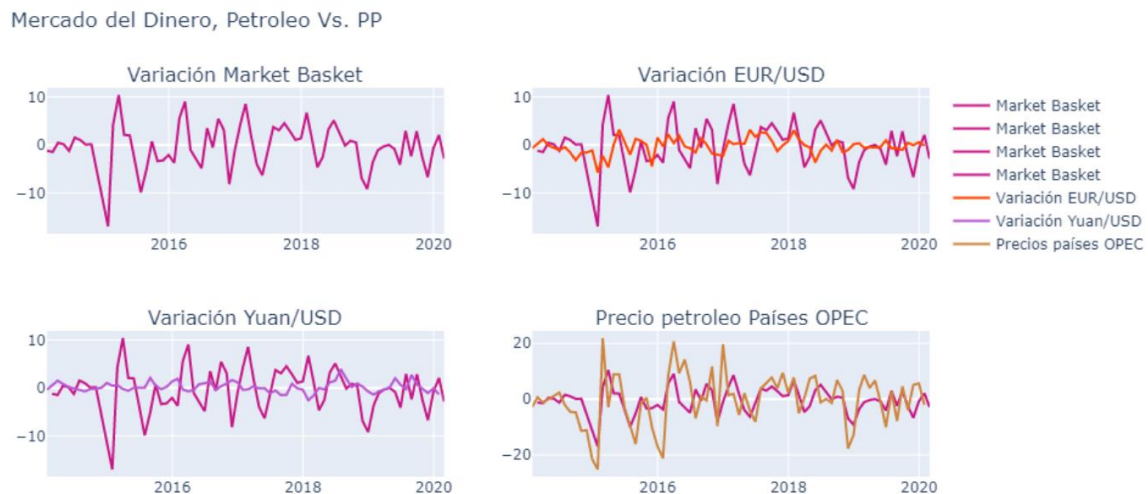


Figura 6-6 Mercado del dinero, el petróleo Vs Market Basket
Fuente: elaboración propia

Determinantes de la oferta y su relación con los precios

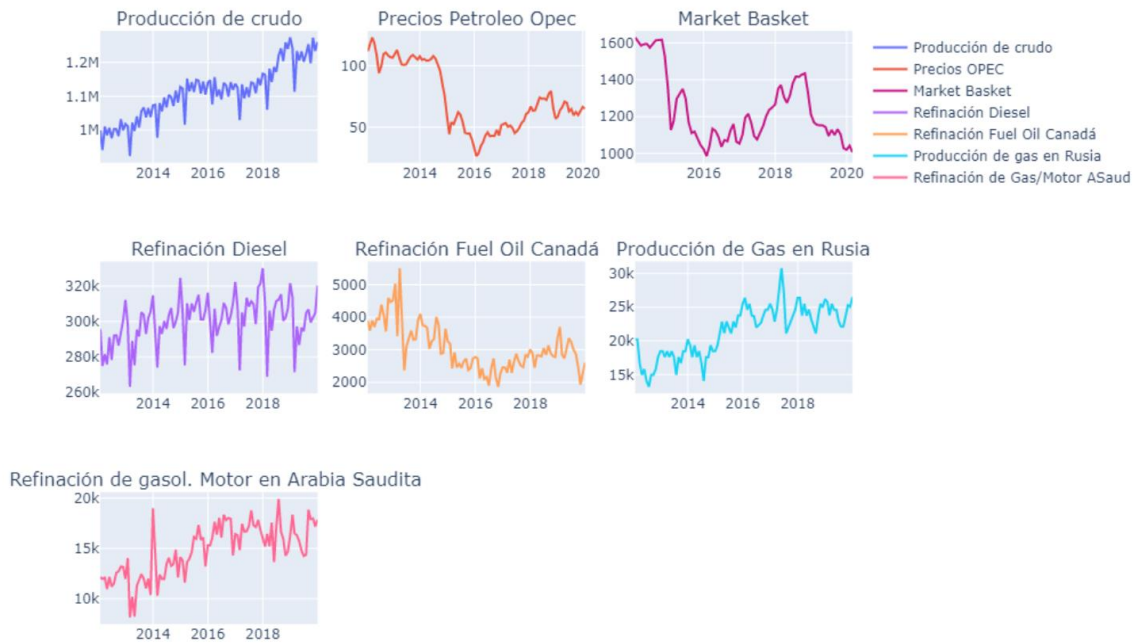


Figura 6-7 Determinantes de los precios para el modelo

Fuente: elaboración propia

Después de la actualización rápida sobre los fundamentos de mercado, el usuario verá la relación del Market Basket con cada uno de los determinantes mencionado, logrando identificar de forma sencilla patrones que sigan las series asociadas, como se muestra en la Figura 6-8:

Variación de los determinantes de la Oferta Vs Market Basket

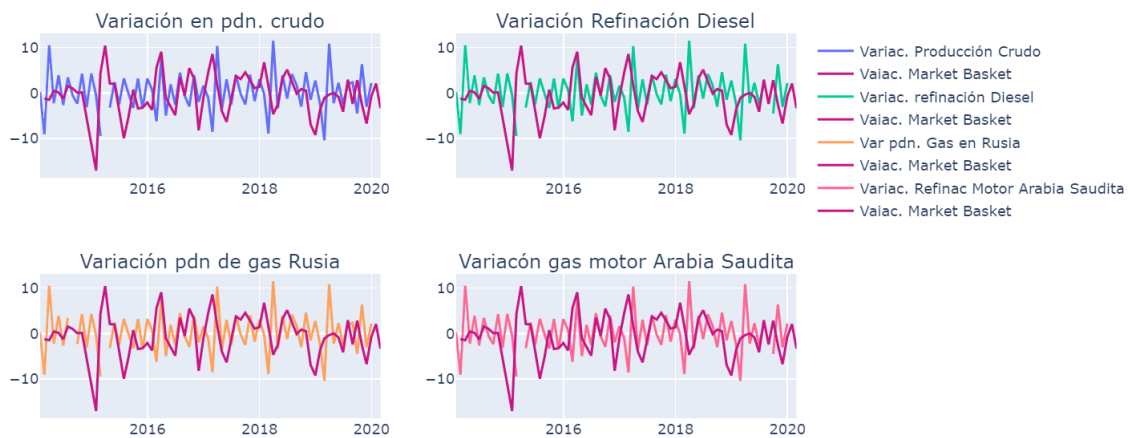


Figura 6-8 Variación de los determinantes de los precios para el modelo Vs Market Basket

Fuente: elaboración propia

Finalmente, como complemento a las decisiones, en la Figura 6-9 se presenta un análisis histórico donde con gran facilidad el usuario puede contrastar los pronósticos propuestos por la familia de modelos con el comportamiento histórico de la variable explicada desde el año 2014, como se muestra en la figura 6-9:

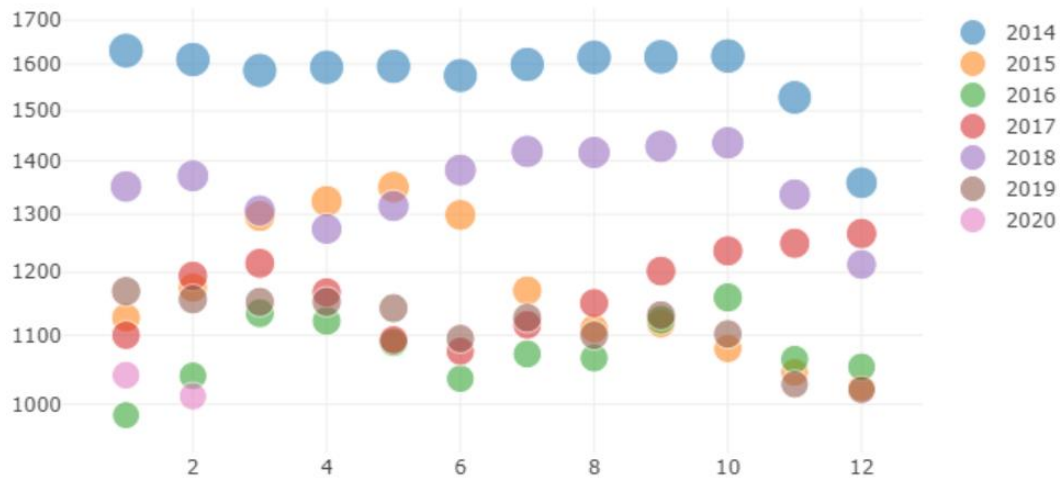


Figura 6-9 Contraste de meses y años Market Basket

Fuente: elaboración propia

Las visualizaciones anteriores están disponibles en un libro interactivo que por un lado permite su fácil mantenimiento en términos de código y permite a los usuarios de las decisiones hacer un análisis completo y rápido de lo que ocurre en materia de polipropileno, logrando mejorar los procesos decisionales.

7 Conclusiones generales, recomendaciones y trabajo futuro

Los modelos de pronóstico seleccionados constituyen una herramienta práctica de soporte a la decisión que debe ser complementada con los criterios expertos, así como con un estricto seguimiento de ajuste. Estos modelos son un insumo importante para el desarrollo de una política de compras, dicha política debería considerar las variables explicativas con el fin de encontrar en mercados líquidos opciones de cobertura cruzada que permitan mejorar la gestión del riesgo de precio en ambientes coberturistas.

Las conclusiones asociadas al cumplimiento de los objetivos planteados se abordarán inicialmente con respecto al cumplimiento de los objetivos específicos y posteriormente al objetivo general.

- Definir las variables de mayor impacto en el mercado mundial de polipropileno
Este objetivo se cumple en los siguientes capítulos: 1.2 Mecanismos de transmisión de precios, 1.3 Aspectos técnicos, 1.4 Aspectos económicos. La síntesis de lo desarrollado en estos capítulos puede verse de forma gráfica en las figuras 6-7 y 6-8, dedicadas a analizar los determinantes de la formación de precios y los determinantes de la oferta respectivamente
- Construir un conjunto de datos que permita analizar y modelar los precios del polipropileno
Este objetivo se cumple en los numerales 2.2 Información disponible y 2.3 Limitaciones en términos analíticos.
- Construir una herramienta para análisis descriptivo y predictivo de los precios de mercado del polipropileno
Este objetivo es el corazón del trabajo, se cumple principalmente en los capítulos 5.1 Discusión y modelado, 5.3 Conclusiones de modelado, 6. Herramienta de soporte a la decisión.

Frente al objetivo general:

- Desarrollar un producto de datos que permita a la industria manufacturera de empaques flexibles (coberturista) la realización de un análisis descriptivo y predictivo de precios en el mercado del polipropileno.
Este se cumple como la agregación de los objetivos específicos, pero de cara al usuario final se cumple mediante la herramienta de uso público que se menciona en el capítulo 6, que es, en resumen, el compendio o memoria de cálculo de los objetivos específicos

Frente al trabajo futuro:

Es necesario reconocer que el mercado atraviesa un importante momento de transformación en los valores colectivos de las personas, las regulaciones verdes, los rechazos a los plásticos de un solo uso y la aparición cada vez más fuerte de los discursos de sostenibilidad que retan a las compañías a encontrar alternativas distintas. Esta transformación social sin duda cambiará la lógica de la

formación de precios y generará unos incentivos distintos para la preferencia por el uso de los plásticos recuperados, en la revisión de la literatura se presentó un estudio en ese campo donde se demostró, que por lo menos en el mercado de plástico recuperado de Seattle no hay correlación entre el mercado de plástico virgen y el de recuperado, de acuerdo con el autor, los precios de la resina recuperada depende en mayor grado de las políticas verdes de los gobiernos (Stromberg, 2004). En este sentido, el modelo debe ser revisado periódicamente en tanto que podrían cambiar los determinantes de los precios en la medida que el mercado posiblemente tome alguna preferencia hacia resinas recuperadas frente a las vírgenes, ocasionando una transformación total en la lógica de formación de precios actual que anticipe la obsolescencia del modelo tal como se plantea en este escrito.

Referencias bibliográficas

- Alianza Caoba. (2018). *Peril Alianza Caoba*. Bogotá: Universidad de los Andes.
- Amghizar, I., Vandewalle, L. a., & Van Geem, K. M. (2017). New Trends in Olefin Production. *Engineering*, 171-178.
- Banco de Bogotá. (2019). *¿Puede seguir subiendo la tasa de cambio?*
- Barrientos, J., Rodas, E., Velilla, E., Lopera, M., & Villada, F. (2012). Modelo para el pronóstico del precio de la energía eléctrica en Colombia. *Lecturas de Economía*, 91-127.
- Benavides, J., & Cardona, Á. (2018). *Mercado eléctrico en Colombia: transición hacia una arquitectura descentralizada*. Bogotá: Fedesarrollo.
- Binary Tribune*. (2013). Obtenido de <http://www.binarytribune.com/2013/05/23/platts-argus-and-icis-separate-reviews/>
- CME. (2019). *CME Group*. Obtenido de <https://www.cmegroup.com/trading/energy/petrochemicals/polymer-grade-propylene-pcw-calendar-swap.html>
- DANE. (2019). *Departamento Administrativo Nacional de Estadística DANE*. Obtenido de Departamento Administrativo Nacional de Estadística DANE: <https://www.dane.gov.co/index.php/estadisticas-por-tema/industria/encuesta-anual-manufacturera-enam>
- Duke University. (21 de Diciembre de 2019). Obtenido de <https://people.duke.edu/~rnau/411arim.htm>
- Gahleitner, M., & Paulik, C. (2016). Polypropylene and Other Polyolefins. En *Brydson's Plastics Materials: Eighth Edition*. (págs. 279-309).
- García, Ó. L. (2009). *Administración financiera fundamentos y aplicaciones*. Cali: Libre Expresión creativos.
- González, H., Ávila, J., Gitter, R., & Ángel, M. M. (2011). Price Transmission in the Production Chain of Natural GastoAnhydrousAmmonia in Mexico. *Terra Latinoamericana*, 1-9.
- Gutierrez, R. (2016). Estrategias dinámicas de cobertura cruzada eficiente para el mercado del petróleo mexicano: Evidencia de dos modelos garch multivariados con término de corrección de error. *Economía teoría y práctica*, 115-146.
- Jiang, J., Marsh, T., & Tozer, P. (2015). Policy induced price volatility transmission: Linking the U.S. crude oil, corn and plastics markets. *Energy economics*, 217-227.

- Kizulj, R. (2012). *Análisis de formación de precios y tarifas de gas natural en América del Sur*. Santiago: CEPAL.
- Maguiña, O. (2016). *Administración 21*. Obtenido de <https://administration21.files.wordpress.com/2017/01/pronc3b3sticos-holt-winters-omr-nov2016.pdf>
- Martínez, G., Cortés, M., & Pérez, A. (2016). Metodología para el análisis de correlación y concordancia en equipos de mediciones similares. *Universidad y Sociedad*, 65-70.
- (2009). *Mecanismos actuales en la formación del precio del gas*.
- Mundoplast. (2019). *Mundoplast*. Obtenido de <https://mundoplast.com/el-consumo-mundial-de-polipropileno-alcanzara-los-624-millones-de-toneladas-en-2020/>
- Muñoz, A., Urquijo, J., Castro, A., & Lombana, J. (2017). Pronóstico del precio de la energía en Colombia utilizando modelos arima con igarch. *Revista de Economía del Rosario*, 125-159.
- Petropoulos, F., Fildes, R., & Goodwin, P. (2005). Do 'big losses' in judgmental adjustments to statistical forecasts affect experts' behaviour? *European Journal of Operational Research*, 842-852.
- Platts S&P. (24 de 10 de 2018). Memorias Foro Anual Platts. Bogotá.
- Portafolio. (2013). *Portafolio.com*. Obtenido de <https://www.portafolio.co/negocios/empresas/refineria-cartagena-costaria-us-000-millones-84248>
- Portilla, J. (2019). *Python for time series analysis*. Obtenido de <https://drive.google.com/file/d/1jHGUMno2qO9x4ZSZ6Us2KC2cJav1Vxld/view>
- Qwerty Wiki. (9 de Diciembre de 2019). *Qwerty Wiki*. Obtenido de <https://es.qwertyu.wiki/wiki/Quandl>
- Sergeev, D. (21 de 12 de 2019). *Open Machine Learning Course. Topic 9. Part 1. Time series analysis in Python*. Obtenido de Medium: <https://medium.com/open-machine-learning-course/open-machine-learning-course-topic-9-time-series-analysis-in-python-a270cb05e0b3>
- Stromberg, P. (2004). Market imperfections in recycling markets: conceptual issues and empirical study of price volatility in plastics. *Resources Conservation & Recycling*, 339.364.
- Trademap. (2019). *See: Importaciones y exportaciones de polipropileno*. Recuperado el 10 de 8 de 2019, de International Trade Centre: <http://www.trademap.org/Index.aspx>
- Troncoso, R. (2019). Transmisión de los precios del arroz en Colombia y el mundo. *Lecturas de Economía*, 151-179.

Vargas, C., Posada, J., Jaramillo, L., & García, L. (2015). Consumo de energía en la industria del plástico: revisión de estudios realizados. *Revista CEA*, 93-107.

Velásquez, J. D. (2008). *Construcción de Escenarios de Pronóstico del Precio de Electricidad en Mercados de Corto Plazo*. Medellín: Universidad Nacional de Colombia.