



**UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL**

MODELOS DE PROYECCIÓN DEL PRECIO DE LA CELULOSA

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL INDUSTRIAL

CAMILA ANDREA CASTILLO CARRASCO

**PROFESOR GUÍA:
WILLIAM BAEZA LÓPEZ**

**MIEMBROS DE LA COMISIÓN:
JAVIER SUAZO SÁEZ
MANUEL REYES JARA**

**SANTIAGO DE CHILE
2017**

**RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE:** Ingeniera Civil Industrial
POR: Camila Andrea Castillo Carrasco
FECHA: 09/03/2017
PROFESOR GUÍA: Sr. William Baeza López

MODELOS DE PROYECCIÓN DEL PRECIO DE LA CELULOSA

La fibra de celulosa es el elemento constitutivo de la madera y se utiliza principalmente en la producción de distintos tipos de papeles, teniendo una gran demanda debido a su uso diario en distintos productos. El presente estudio tiene como objetivo generar una proyección del precio global que permita explicar las fluctuaciones de los dos tipos más comunes: fibra corta química blanqueada y fibra larga química blanqueada. Chile se posiciona en el décimo lugar en cuanto a producción y en el cuarto lugar con respecto a exportaciones globales de celulosa representando un 4% de las exportaciones nacionales para el año 2016, por esto, la proyección realizada resulta de interés para inversionistas y productores de celulosa, quienes pueden relacionar el precio proyectado con resultados financieros de los productores chilenos.

Se desarrollan proyecciones utilizando tres modelos econométricos sencillos y ampliamente utilizados en la proyección de commodities: modelo autorregresivo integrado de media móvil, el cual usa observaciones pasadas para predecir variaciones y es usado como referencia; modelo autorregresivo con variables explicativas, el cual relaciona las fluctuaciones con cuatro indicadores macroeconómicos globales (precio petróleo WTI, inflación estadounidense, índice accionario global e índice de producción industrial global); y modelo de corrección de errores, el cual captura la relación de largo plazo entre el precio de la celulosa y sus inventarios globales. No se han realizado estudios similares aplicados a ambos tipos de fibra, por lo que la elección de los modelos mencionados se basa en su capacidad de predicción en distintos horizontes y su capacidad para explicar las fluctuaciones a través de factores globales y relaciones económicas de largo plazo.

Las proyecciones obtenidas mejoran los resultados del modelo de referencia a través del uso de variables explicativas, ofreciendo un R^2 ajustado entre 0,53 y 0,75; presentando mejores resultados en el caso de la fibra larga, debido a las diferencias entre los mercados involucrados. Para evaluar el desempeño de los modelos se calcula el error cuadrático medio sobre múltiples proyecciones obtenidas a través de la técnica de origen móvil, los modelos desarrollados ofrecen la posibilidad de proyectar en el corto y mediano plazo, obteniendo una desviación inferior a US\$35 para un horizonte de seis meses, variación que se traduce en un error de $\pm 4\%$ aproximadamente.

Las variables utilizadas muestran un gran poder explicativo, entregando información que puede ser integrada en distintas aplicaciones y podría ser replicable en otros commodities. Para futuras investigaciones se propone analizar el comportamiento estratégico de consumidores chinos, quienes tienen un gran impacto sobre el precio representando un 31% de las importaciones globales; adicionalmente se plantea incorporar al análisis los distintos tipos de cambio involucrados en el mercado asociado a cada fibra.

Agradecimientos

Agradezco a mi familia, sobre todo a mi mamá, papá, abuelas y tata, por su paciencia, cariño y apoyo.

A mis amigos por hacer mi paso por la universidad infinitamente más agradable, principalmente Eial y Claudio por aguantarme, hacer reír y acompañarme siempre.

Finalmente quiero agradecer a mi profesor guía, William Baeza, por la gran disposición y comprensión mostrada durante este periodo, y a todos los que me ayudaron a desarrollar este trabajo, especialmente a Marcelo Catalán, Rodrigo Perera, Leopoldo Silva y Cristian Palacios.

Tabla de contenido

1	Introducción.....	1
1.1	Mercado de celulosa	3
1.2	Determinantes del precio de la celulosa.....	18
2	Marco Conceptual	20
3	Descripción del proyecto	24
3.1	Justificación.....	24
3.2	Objetivos	25
3.3	Alcances.....	25
4	Metodología	26
4.1	Series de tiempo y desafíos en la proyección de commodities	27
4.2	Estimación mediante modelo autorregresivo	30
4.3	Estimación mediante modelo de indicadores macroeconómicos	33
4.4	Estimación mediante modelo de corrección de errores.....	35
4.5	Origen móvil y métricas de ajuste	38
4.6	Descomposición del error cuadrático medio	40
5	Fuentes de datos	41
5.1	Precio de la celulosa	41
5.2	Indicadores macroeconómicos.....	41
5.3	Días de inventario	42

6	Resultados	43
6.1	Estacionariedad series	45
6.2	Resultados modelo autorregresivo	46
6.3	Resultados modelo de indicadores macroeconómicos	49
6.4	Resultados modelo de corrección de errores	54
6.5	Error de proyecciones	57
6.6	Análisis de especificación	59
7	Conclusiones y recomendaciones	63
	Glosario	66
	Bibliografía	67
	Anexo	70
	Anexo A: Estimaciones modelo autorregresivo	70
	Anexo B: Estimaciones modelo de indicadores macroeconómicos	81
	Anexo C: Estimaciones modelo corrección de errores	90

Índice de tablas

Tabla 1: Resumen revisión bibliográfica.....	23
Tabla 2: Propiedades para ACF y PACF.....	31
Tabla 3: Descripción datos de inventario	42
Tabla 4: Resumen principales resultados obtenidos	44
Tabla 5: Resultados test de raíz unitaria sobre precios.....	45
Tabla 6: Resultados modelo autorregresivo	48
Tabla 7: Efectos netos de indicadores macroeconómicos	51
Tabla 8: Resultados modelo de indicadores macroeconómicos.....	53
Tabla 9: Resultados modelo corrección de errores	56
Tabla 10: Análisis estacionalidad variables dependientes	74
Tabla 11: Identificación de parámetros p y q para fibra corta (log).....	75
Tabla 12: Identificación de parámetros p y q para fibra larga (log).....	75
Tabla 13: Identificación de parámetros p y q para ratio.....	76
Tabla 14: Comparación de modelos estimados para fibra corta	76
Tabla 15: Comparación de modelos estimados para fibra larga	77
Tabla 16: Comparación de modelos estimados para ratio	77
Tabla 17: Estimación modelo ARIMA para fibra corta.....	78
Tabla 18: Estimación modelo ARIMA para fibra larga.....	78
Tabla 19: Estimación modelo ARIMA para ratio.....	78
Tabla 20: Resultados test de raíz unitaria sobre indicadores macroeconómicos.....	82
Tabla 21: Estacionalidad variables exógenas	83
Tabla 22: Test de causalidad de Granger sobre indicadores macroeconómicos	85
Tabla 23: Estimación modelo de indicadores macroeconómicos para fibra corta	86

Tabla 24: Estimaciones modelo de indicadores macroeconómicos para fibra larga	87
Tabla 25: Estimaciones modelo de indicadores macroeconómicos para ratio	87
Tabla 26: Test de raíz unitaria sobre series de inventario	90
Tabla 27: Resultados procedimiento de Engle-Granger sobre precios e inventarios	91
Tabla 28: Evaluación de colinealidad de fibras	92
Tabla 29: Test de causalidad de Granger sobre inventarios	92
Tabla 30: Ajuste de proyecciones para modelo de corrección de errores	93
Tabla 31: Estimaciones modelo de corrección de errores para fibra corta.....	97
Tabla 32: Estimaciones modelo de corrección de errores para fibra larga.....	97

Índice de gráficos

Gráfico 1: Producción global de fibra de celulosa 2000-2015	4
Gráfico 2: Demanda agregada por fibra química blanqueada 2004-2015.....	5
Gráfico 3: Producción de fibra química blanqueada por región 2000-2015.....	6
Gráfico 4: Distribución de la producción de fibra química blanqueada 2015.....	6
Gráfico 5: Importaciones de fibra química blanqueada por región 2000-2015	7
Gráfico 6: Distribución importaciones fibra química blanqueada 2015.....	7
Gráfico 7: Evolución exportaciones de fibra química blanqueada por región 2000-2015.	8
Gráfico 8: Distribución exportaciones fibra química blanqueada 2015.....	8
Gráfico 9: Curvas de oferta globales para fibra corta y fibra larga	10
Gráfico 10: Adiciones de capacidad para principales productores 2016-2020	11
Gráfico 11: Capacidad y producción principales productores 2015.....	11
Gráfico 12: Días de inventario de demanda versus precio para fibra corta	12
Gráfico 13: Días de inventario de demanda versus precio para fibra larga	13
Gráfico 14: Días de inventario de oferta versus precio para fibra corta.....	14
Gráfico 15: Días de inventario de oferta versus precio para fibra larga.....	14
Gráfico 16: Estacionalidad días inventario de oferta BCP 07-2004 a 04-2016.....	15
Gráfico 17: Exportación de bienes 2003 a 2015.....	17
Gráfico 18: Principales productos exportados 2015	17
Gráfico 19: Proyecciones modelo autorregresivo fibra corta	47
Gráfico 20: Proyecciones modelo autorregresivo fibra larga	47
Gráfico 21: Precio fibras vs Inflación USA.....	49
Gráfico 22: Precio fibras vs precio WTI	49
Gráfico 23: Precio fibras vs CLI G7	50

Gráfico 24: Precio fibras vs ACWI	50
Gráfico 25: Proyecciones modelo de indicadores macroeconómicos fibra corta	52
Gráfico 26: Proyecciones modelo de indicadores macroeconómicos fibra larga	52
Gráfico 27: Proyecciones modelo corrección de errores fibra corta	55
Gráfico 28: Proyecciones modelo corrección de errores fibra larga	56
Gráfico 29: RSME a distintos alcances para fibra corta	57
Gráfico 30: RMSE a distintos alcances para fibra larga	58
Gráfico 31: Descomposición MSE modelo autorregresivo para fibra corta	59
Gráfico 32: Descomposición MSE modelo autorregresivo para fibra larga	59
Gráfico 33: Descomposición MSE modelo de indicadores macroeconómicos para fibra corta	60
Gráfico 34: Descomposición MSE modelo de indicadores macroeconómicos para fibra larga	60
Gráfico 35: Descomposición MSE modelo de corrección de errores para fibra corta	61
Gráfico 36: Descomposición MSE modelo de corrección de errores para fibra larga	61
Gráfico 37: Correlograma precio de fibra corta (log)	71
Gráfico 38: Correlograma precio de fibra larga (log)	71
Gráfico 39: Correlograma ratio	72
Gráfico 40: Correlograma primera diferencia precio de fibra corta (log).....	73
Gráfico 41: Correlograma primera diferencia precio de fibra larga (log).....	73
Gráfico 42: Correlograma primera diferencia ratio	74
Gráfico 43: Correlograma residuos estimación fibra corta	79
Gráfico 44: Correlograma residuos estimación fibra larga	79
Gráfico 45: Correlograma residuos estimación ratio	80

Gráfico 46: Correlación dinámica entre fibra corta e indicadores macroeconómicos	83
Gráfico 47: Correlación dinámica entre fibra larga e indicadores macroeconómicos	84
Gráfico 48: Correlación dinámica entre ratio e indicadores macroeconómicos	84
Gráfico 49: CUSUM modelo de indicadores macroeconómicos fibra corta	88
Gráfico 50: CUSUM modelo indicadores macroeconómicos fibra larga	88
Gráfico 51: CUSUM modelo indicadores macroeconómicos ratio	89
Gráfico 52: Proyección fibra corta con datos originales de inventario	93
Gráfico 53: Proyección fibra larga con datos originales de inventario	94
Gráfico 54: Proyección sobre inventarios de demanda fibra corta	94
Gráfico 55: Proyección sobre inventarios de demanda BCP	95
Gráfico 56: Proyección fibra corta con datos estimados de inventarios	95
Gráfico 57: Proyección fibra larga con datos estimados de inventarios	96

1 Introducción

La fibra de celulosa es el elemento constitutivo de la madera y se utiliza principalmente para la producción de tissue, embalajes y distintos tipos de papeles. Es un commodity de consumo masivo, siendo la materia prima más abundante del mundo, y sus precios de referencia globales son reportados semanalmente.

Existen variados tipos de fibra dependiendo tanto del tipo de árbol utilizado como de los procesos involucrados en su obtención, en el presente estudio se desarrolla una proyección para las dos variedades más consumidas en el mercado global: fibra larga química blanqueada (proveniente de maderas coníferas, principalmente pino) y fibra corta química blanqueada (proveniente de maderas latifoliadas, principalmente eucaliptus). Cada una de estas fibras tiene distintas características, en particular con respecto a sus usos se tiene que la fibra larga es una fibra más resistente y se utiliza principalmente para la producción de embalajes y en la producción de distintos papeles y filtros, mientras que la corta es una fibra más suave y se utiliza principalmente para la producción de papeles finos y tissue.

Al ser un commodity transado internacionalmente, existe una gran cantidad de variables que influyen en su precio, en particular existen factores propios del mercado y variables externas que resultan útiles en la proyección de commodities. En el presente estudio se desarrolla un modelo sencillo que permita tanto proyectar como describir las variaciones en los precios internacionales de fibra corta y fibra larga de celulosa. Con este fin se consideran tres modelos econométricos distintos seleccionados según su capacidad predictiva en distintos horizontes y su capacidad de explicar las fluctuaciones importantes en el precio. Los modelos utilizados son: modelo autorregresivo; modelo autorregresivo con variables explicativas, el cual relaciona las fluctuaciones de precio de la celulosa con cuatro factores globales; y modelo de corrección de errores, el cual captura la relación de largo plazo entre el precio de la celulosa y los inventarios globales.

La proyección obtenida resulta de interés tanto para inversores como para productores de celulosa, ya que esta puede ser relacionada con indicadores financieros o bursátiles de los productores chilenos, obteniendo una estimación de sus resultados esperados.

Dicha información también resulta de interés para entidades financieras, quienes pueden generar recomendaciones para sus clientes en base a dichas estimaciones y/o replicar el análisis con el fin de proyectar diversos commodities.

A continuación se realiza una descripción del mercado de celulosa considerando los factores más relevantes para el mercado internacional y los principales determinantes que influyen en el precio. En la sección 2 se resume el marco conceptual describiendo la literatura en el modelamiento del mercado forestal y en la proyección de

commodities. En la sección 3 se describe el proyecto explicitando sus objetivos y alcances. En la sección 4 se explica la metodología utilizada y se describen los tres modelos utilizados, con sus respectivos procedimientos y fundamentos teóricos. En la sección 5 se detallan los datos utilizados en la realización del estudio. En la sección 6 se analiza en orden de integración de las variables en estudio, se presentan los resultados obtenidos con los modelos desarrollados analizando sus errores de proyección, y adicionalmente se analiza la estabilidad de los parámetros obtenidos mediante descomposición del error cuadrático medio y aplicación de origen móvil al modelo estimado. Finalmente, en la sección 7, se concluye, brindando recomendaciones para estudios futuros.

1.1 Mercado de celulosa

A principios de la década del 2000 se producen distintos cambios estructurales a nivel de demanda en el mercado de productos forestales debido a la saturación del mercado en Norte América y Europa, además de la entrada de Rusia y China, dos grandes economías con expectativas de rápido crecimiento de demanda. Dichos cambios incrementaron los flujos comerciales de productos forestales y también afectaron las decisiones de locación relativas a las inversiones de la industria. Por otra parte, la presión competitiva generada por nuevos productores con bajos costos en Asia y Sudamérica obliga a los productores estadounidenses y europeos a obtener economías de escala (Toppinen, y otros, 2010).

La fibra de celulosa es la principal materia prima en la producción de papel y cartón. La mayor parte de la pulpa virgen se extrae de chips de madera, el resto se deriva de fuentes alternativas como algodón, bambú, paja y cáñamo. El objetivo del proceso de fabricación de pulpa es romper la estructura de la madera y obtener sus elementos constitutivos mediante una variedad de procesos.

El proceso químico más común es el kraft o al sulfato, en el cual se cuecen los chips de madera bajo presión en una mezcla de hidróxido de sodio y sulfuro de sodio, conocido como licor blanco. La fibra resultante se lava para extraer los químicos y luego se seca. Otros métodos que utilizan distintos químicos para la extracción de pulpa son el proceso al sulfito y el método soda.

El proceso mecánico implica la molienda de chips de madera para obtener fibra de celulosa. Cuando se aplica presión y temperatura para ablandar los chips, el proceso se llama termomecánico. Los métodos que involucran químicos suelen ser más caros que los métodos mecánicos, pero entregan un producto mucho más suave y refinado.

Las variedades de fibra se pueden clasificar según la madera utilizada o el proceso utilizado para su obtención.

Según el tipo de madera utilizada se tienen dos tipos de fibra:

- Fibra larga: producida a partir de maderas coníferas como el pino, alerce, entre otros. Es una fibra resistente utilizada principalmente en la producción de embalajes y papeles. También se usa para producir fluff pulp para pañales.
- Fibra corta: producida a partir de maderas latifoliadas como el eucalipto, acacia, álamo, entre otros. Es una fibra más densa, pero menos resistente que la fibra larga. Se utiliza principalmente para producir papeles y tissue

Según el proceso utilizado se obtienen las siguientes variedades de fibra:

- Fibra química: incluye fibra kraft blanqueada y no blanqueada, y fibra producida por proceso al sulfito y por método soda.
- Fibra mecánica
- Fibra semi-química

Por el lado de la oferta, el precio de la fibra puede verse afectado por la demanda en el sector construcción, así como también por condiciones climáticas, ya que el calor

en conjunto con fuertes vientos son condiciones propicias para incendios forestales, lo que afecta a la producción en el largo plazo.

En el caso de la demanda, el precio de la celulosa se ve influenciado por la demanda de papel, ya que este es su principal uso. Esta puede variar en función de las preferencias de los consumidores, aunque estas variaciones son limitadas, ya que hay productos como tissue o embalajes que no presentan grandes variaciones.

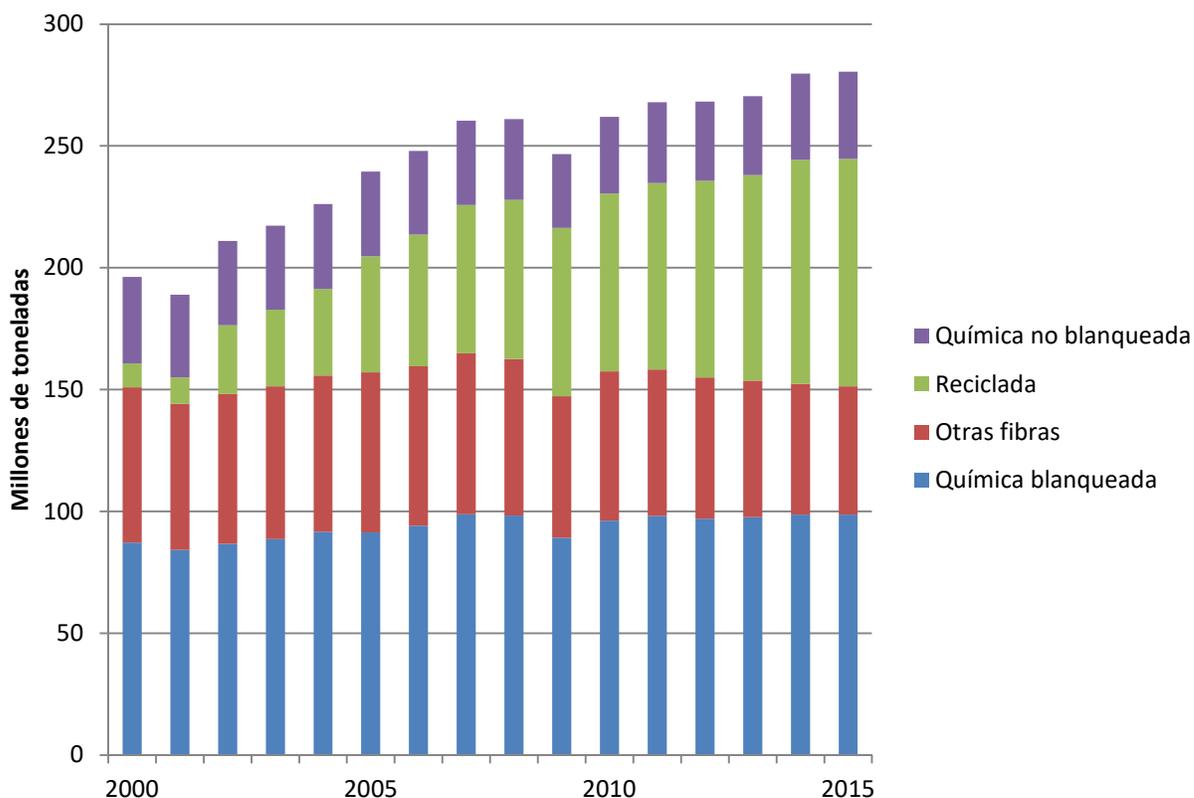
Los costos energéticos pueden variar entre un 15% y un 30% del costo de producción, por esto las plantas de producción se ubican en zonas donde estos costos sean menores, como por ejemplo cerca de planta hidroeléctricas.

Se espera que la demanda por fibra larga crezca con mayor rapidez que la demanda por fibra corta, ya que en los últimos años ha habido una tendencia a sustituir el uso de fibra corta por fibra larga en la producción de papel.

El impacto ambiental de la producción de papel y celulosa es significativo y la preocupación sobre la generación de deforestación y contaminación está en aumento. La regulación en este tema aumenta los costos de producción y también ha guiado una tendencia hacia la sustentabilidad y reciclaje de fibras y papeles.

La fibra más utilizada es la fibra química en sus distintas variedades. El Gráfico 1 muestra la evolución de la producción global de los distintos tipos de fibra de celulosa:

Gráfico 1: Producción global de fibra de celulosa 2000-2015



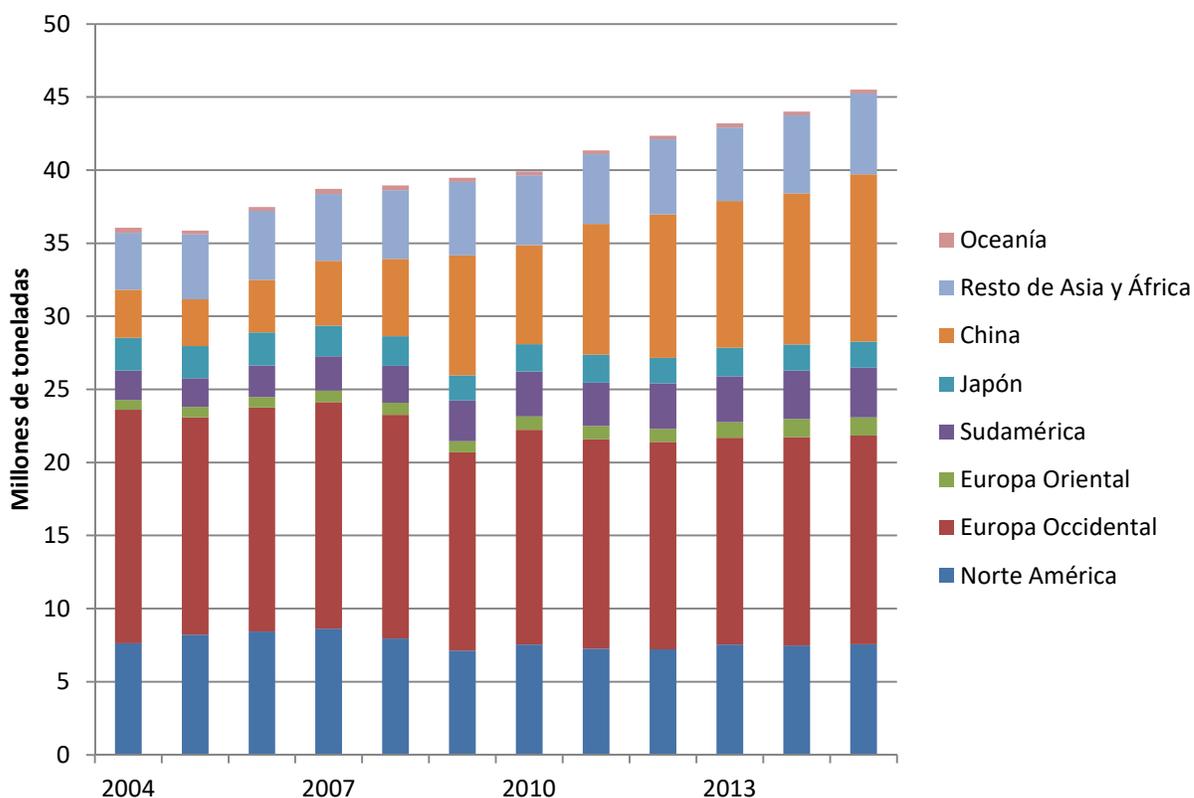
Fuente: Elaboración propia en base a datos de FAOSTAT

La fibra que presenta mayor incidencia dentro del periodo bajo estudio corresponde a la fibra reciclada con una incidencia de 42,5%. Si bien podría pensarse que dicha fibra actúa como sustituto de las fibras vírgenes, según (Szabó, y otros, 2009) estas actúan como productos complementarios y su crecimiento no debería afectar en mayor medida al mercado de la fibra virgen. La fibra química blanqueada presenta una incidencia de 5,8%, la que puede explicarse por un efecto sustitución con las fibras catalogadas en “otras fibras”, las que tienen una incidencia de -5,6%.

En el presente estudio se analizan los precios de la fibra química blanqueada o BCP¹ por su sigla en inglés (Bleached Chemical Pulp), la cual representa la mayor proporción de la fibra virgen, con una tasa de crecimiento anual compuesto o CAGR por su sigla en inglés (Compound Annual Growth Rate) de 0,8%.

El Gráfico 2 muestra la demanda agregada de BCP a nivel global.

Gráfico 2: Demanda agregada por fibra química blanqueada 2004-2015



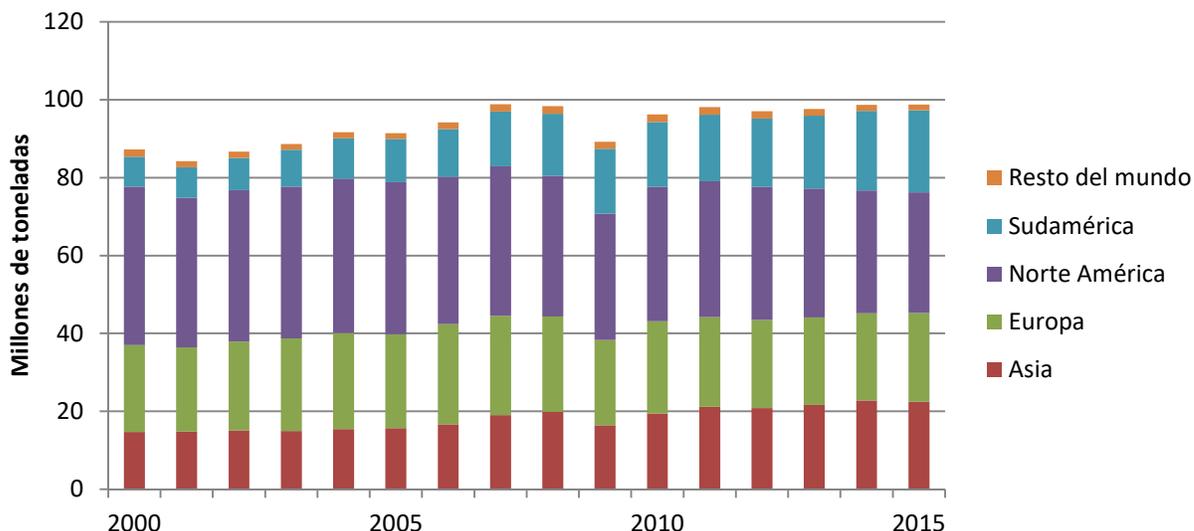
Fuente: Datos obtenidos de Arauco y Hawkins Wright

A pesar que la producción de BCP no ha aumentado considerablemente, la demanda agregada por fibra de mercado (fibra que es exportada desde su país de origen) sí lo ha hecho con un CAGR de 2,1% anual (explicado principalmente por el crecimiento de la demanda china con un CAGR de 12,1%), lo que indica que el mercado se ha vuelto más competitivo. Cabe destacar que no existe producto sustituto para la fibra de celulosa, por lo que la competencia se produce dentro del mercado, sin verse afectado por competencia externa.

¹ En adelante se habla de BCP para hacer referencia a la fibra química blanqueada.

El Gráfico 3 muestra la evolución de la producción por región para el periodo 2000-2015. Norte América y Europa muestran una saturación del mercado, presentando una producción constante en el tiempo, o incluso con tendencia a la baja, como es el caso de Norte América, que presenta un CAGR de -1,3%. En el caso de Asia y Sudamérica, ambos muestran un CAGR de 6,5% y 1,5% respectivamente, lo que podría indicar un potencial de crecimiento. Para el resto del mundo la producción es marginal y se mantiene constante.

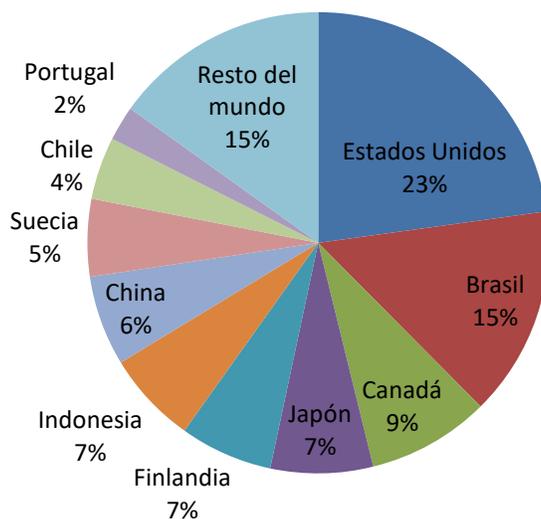
Gráfico 3: Producción de fibra química blanqueada por región 2000-2015



Fuente: Elaboración propia en base a datos de FAOSTAT

El Gráfico 4 presenta los 10 principales productores de BCP para el año 2015, donde se evidencia el importante posicionamiento de Chile en la producción de fibra de celulosa, alcanzando el noveno lugar con más de 4,3 millones de toneladas anuales.

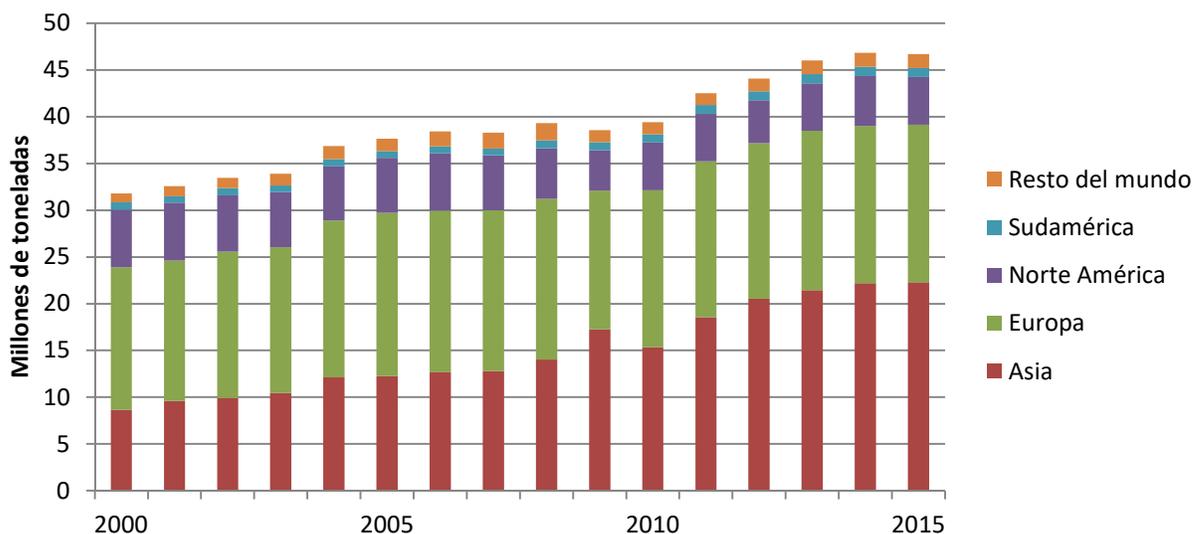
Gráfico 4: Distribución de la producción de fibra química blanqueada 2015



Fuente: Elaboración propia en base a datos de FAOSTAT

El Gráfico 5 muestra la evolución de las importaciones por región para el periodo 2000-2015. Asia presenta una importante tendencia con un CAGR de 6,5% anual, posicionándose como el principal importador a nivel mundial. Europa y Norte América muestran niveles relativamente constantes, pero considerables dentro del total de importaciones, mientras que Sudamérica y el resto del mundo presentan niveles despreciables.

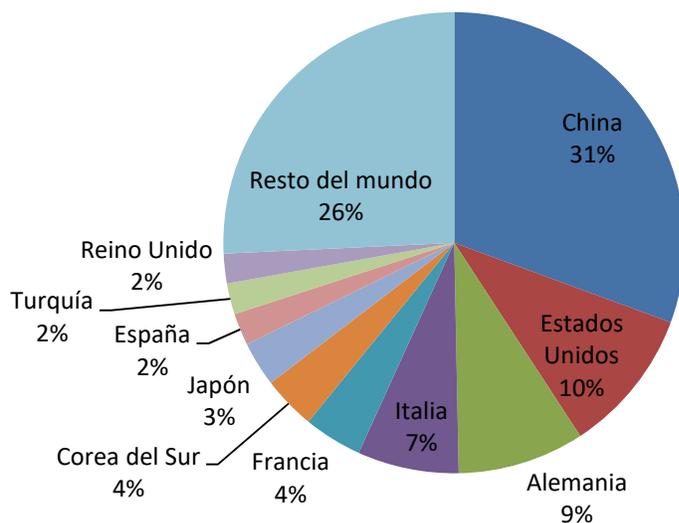
Gráfico 5: Importaciones de fibra química blanqueada por región 2000-2015



Fuente: Elaboración propia en base a datos de FAOSTAT

El aumento de las importaciones de Asia se explica principalmente por la importante tendencia al alza por parte de importadores chinos, quienes logran abarcar un 31% de las importaciones globales para el año 2015. El Gráfico 6 muestra los 10 principales importadores de celulosa para el año 2015:

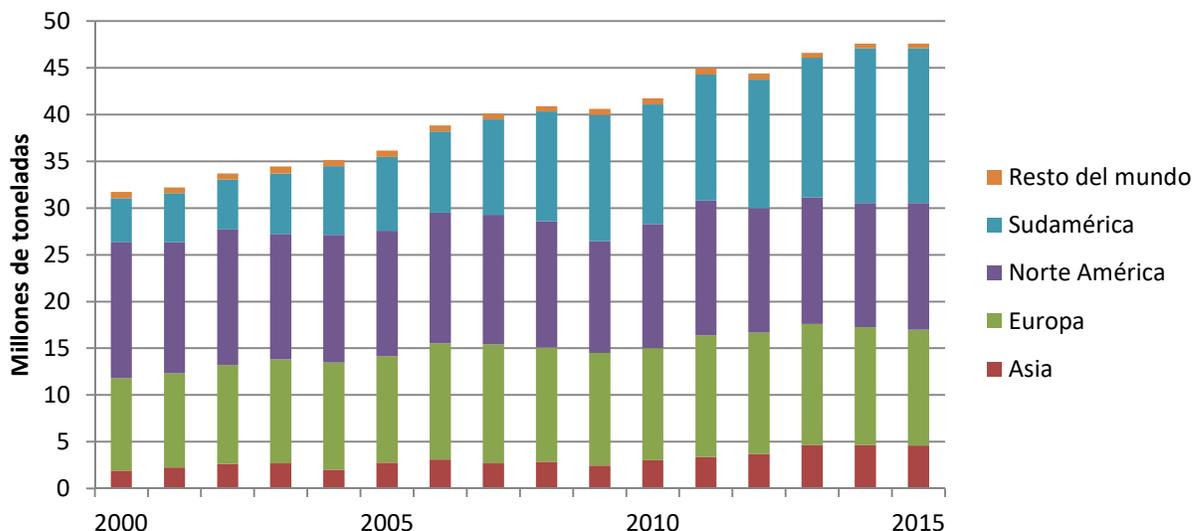
Gráfico 6: Distribución importaciones fibra química blanqueada 2015



Fuente: Elaboración propia en base a datos de FAOSTAT

Finalmente, el Gráfico 7 muestra la evolución de las exportaciones por región para el periodo 2000-2015, donde se aprecia el importante crecimiento de las exportaciones en Sudamérica, las que presentan un CAGR de 1,5%, logrando posicionarse como el principal exportador de celulosa.

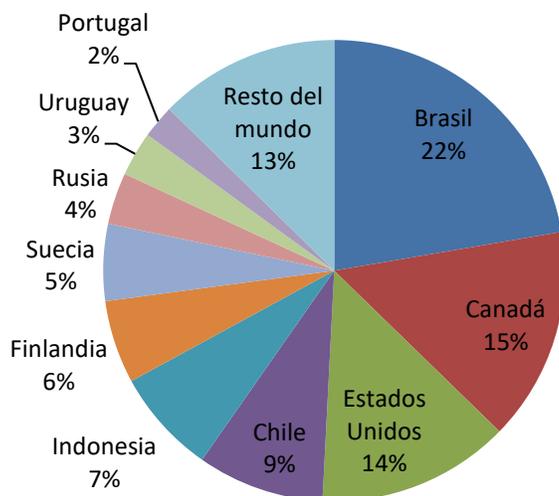
Gráfico 7: Evolución exportaciones de fibra química blanqueada por región 2000-2015



Fuente: Elaboración propia en base a datos de FAOSTAT

El Gráfico 8 muestra los 10 principales exportadores de celulosa para el año 2015. Este evidencia la relevancia de Norte América y Sudamérica como exportadores, regiones en las que se encuentran los cuatro principales exportadores, alcanzando un 60% del total global.

Gráfico 8: Distribución exportaciones fibra química blanqueada 2015



Fuente: Elaboración propia en base a datos de FAOSTAT

Existen distintas consultoras como RISI Inc, Hawkins Wright Ltd, entre otras, que se dedican a recolectar información a nivel global de empresas del sector forestal con el fin de vender análisis especializados para productores e instituciones financieras que realizan seguimiento a las empresas del rubro. Estas publicaciones comprenden análisis de evolución de precios y mercado, noticias del rubro (principalmente adiciones de capacidad) e información sobre capacidades y estructura de costos para los principales productores.

FOEX Indexes Ltd. reporta precios de referencia semanales para fibras, papeles, embalajes, papel reciclado y biocombustibles. Dependiendo del producto se puede obtener un índice global o desagregado por región. Estos índices son calculados a partir de la recolección de datos reportados por productores y compradores de celulosa, los cuales posteriormente son ponderados por el volumen anual de ventas (o compras) según tipo de fibra de la empresa que reporta el precio. Dado que los índices son sólo precios globales de referencia y no se trata de un commodity tradicional, se realizan transacciones que tienen grandes variaciones con respecto a dichos índices. Se destaca que tampoco existen futuros asociados a la celulosa, ni se negocian contratos a largo plazo.

En este estudio se realiza una proyección para los precios globales reportados para BCP desagregada en fibra corta y fibra larga², las cuales son las principales fuentes de ingresos de los productores chilenos.

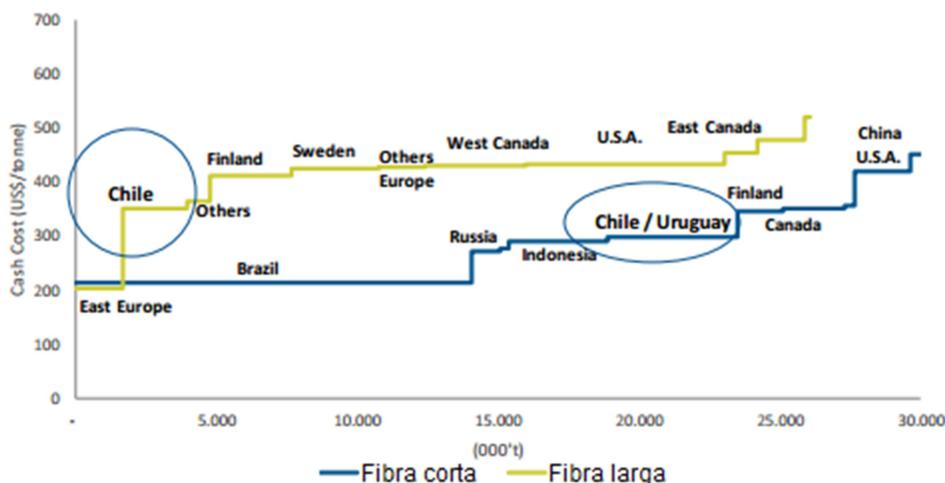
En general, la fibra larga es más cara que la fibra corta, ya que sus usos son más variados: se utiliza para la producción de papeles, cartones, filtros y hasta pañales, mientras que los usos de la fibra corta se reducen prácticamente a papeles y tissue. El spread existente entre las fibras es relevante, ya que existe un efecto sustitución entre las fibras en el sentido de que en la producción de papeles y tissue se puede reemplazar parte de la fibra corta usando fibra larga, por lo que es necesario tomar en cuenta la valoración relativa de las fibras. Es necesario destacar que este trade-off no se produce en el sentido inverso, ya que no es posible usar fibra corta en embalajes.

En particular, en Chile se puede producir eficientemente ambos tipos de fibra, ya que las condiciones climáticas permiten tener ciclos de crecimiento corto para las dos especies de árboles involucrados.

² En adelante se habla de fibra larga para hacer referencia a fibra larga química blanqueada, y fibra corta para hacer referencia a fibra corta química blanqueada.

Por condiciones climáticas el eucalipto (fibra corta) se produce de forma más eficiente en el hemisferio sur, mientras que el pino (fibra larga) se produce más en el hemisferio norte, donde existen máquinas menos eficientes, debido principalmente a su antigüedad. El Gráfico 9 muestra la curva de oferta para fibra larga y fibra corta.

Gráfico 9: Curvas de oferta globales para fibra corta y fibra larga



Fuente: Presentación corporativa Empresas Copec, Octubre 2016

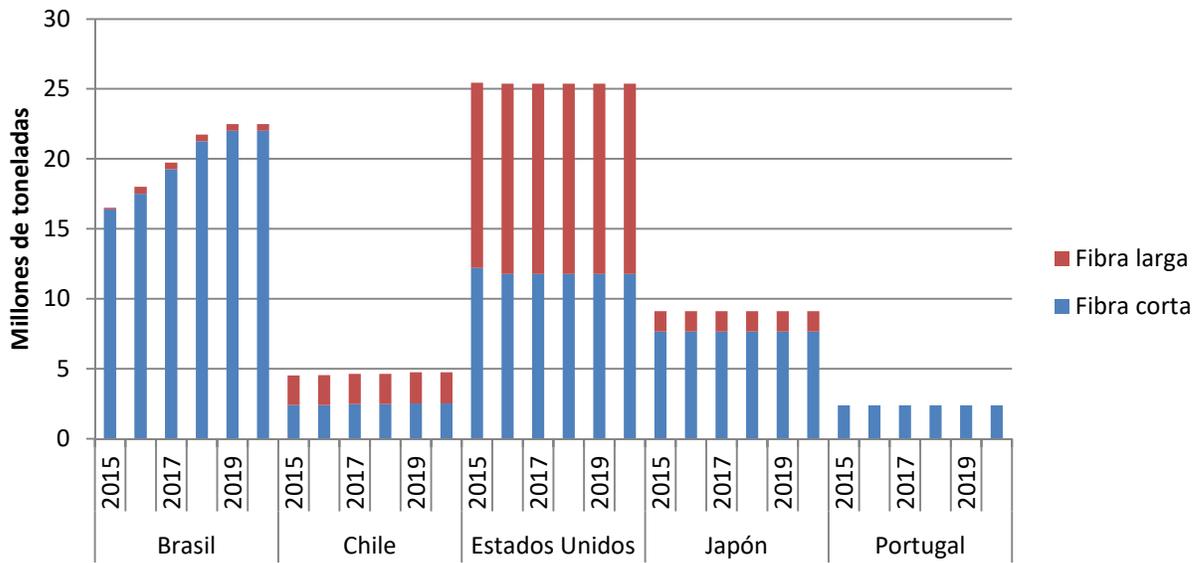
Estas curvas de oferta no son estáticas, dado que en el caso de los productores chilenos aproximadamente un 30% de los costos está indexado en moneda local y el resto está indexado en dólares, mientras que para los productores brasileños un 70% está indexado en moneda local y el resto está indexado en dólares. Esto produce movimientos en la curva de oferta debido a la apreciación o depreciación relativa de sus monedas. Una aplicación útil sería proyectar la curva de costos usando estimaciones del precio de la celulosa y los tipos de cambio, con el fin de optimizar los costos y las decisiones de capacidad, aunque su uso sería bastante restringido, dependiendo de la flexibilidad de la estructura de costos. Por otro lado, la curva de oferta se está aplanando, lo que restringe el plazo de aplicación.

Uno de los principales indicadores del sector corresponde a las capacidades de los distintos productores. Entidades financieras evalúan cómo impactan las adiciones de capacidad y anuncios de la competencia sobre los resultados financieros de los productores.

Dichas adiciones de capacidad no siempre se realizan, dado el alto costo de inversión; y cuando se realizan no siempre cumplen con los plazos programados, por lo que esta información podría resultar poco confiable en el largo plazo.

La Organización de las Naciones Unidas para la Alimentación y la Agricultura, conocida como FAO reporta anualmente las proyecciones de capacidad para un horizonte de cinco años. Estos datos son recopilados mediante encuestas realizadas a los productores de celulosa, aunque estas sólo son respondidas por un grupo pequeño de países (30 para la publicación más reciente). El Gráfico 10 muestra la adición de capacidad, desagregada por tipo de fibra, para los principales productores de los que se tienen datos.

Gráfico 10: Adiciones de capacidad para principales productores 2016-2020

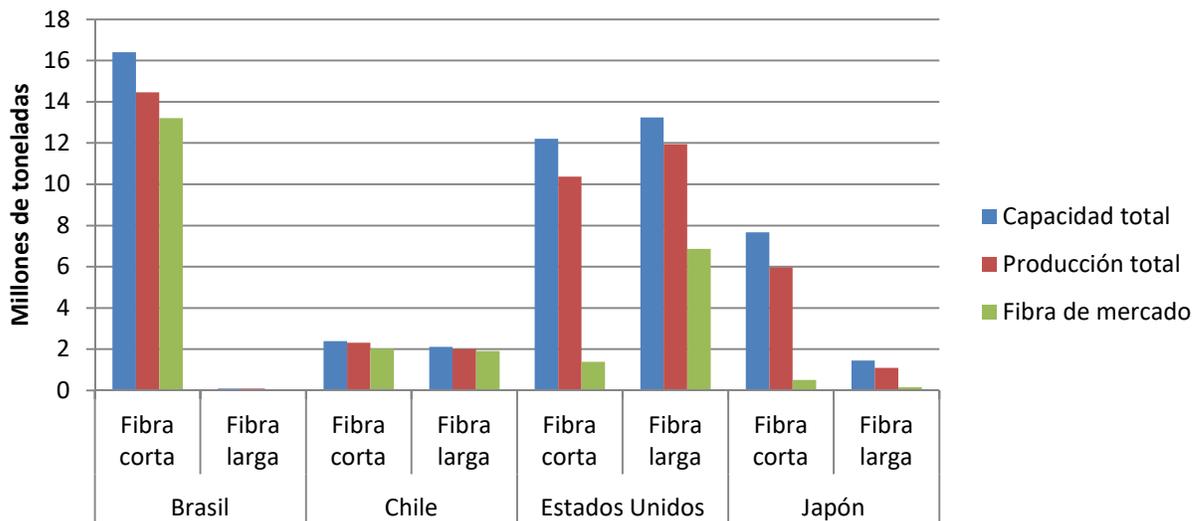


Fuente: Elaboración propia en base a datos de FAOSTAT

Dado que los anuncios de proyecciones de capacidad no siempre son efectivos, no es evidente la identificación del momento óptimo en que una empresa debe modificar su capacidad de producción. Una posible aplicación de un modelo de largo plazo sería la identificación oportunidades de ampliación considerando distintos escenarios para el precio de la celulosa.

Por otro lado, distintos productores tienen distintos niveles de utilización de su capacidad, lo que implica complicaciones o posibles errores sobre el análisis de esta. El Gráfico 11 muestra el nivel de producción versus la capacidad instalada y el nivel de fibra de mercado para los principales productores.

Gráfico 11: Capacidad y producción principales productores 2015



Fuente: Elaboración propia en base a datos de FAOSTAT

Los inventarios globales y/o regionales son fundamentales en los mercados de commodities, ya que representan el principal indicador de desequilibrio entre oferta y demanda. Se han utilizado en variados estudios para proyectar el precio de la celulosa, cobre y otros commodities, por ejemplo (Ye, y otros, 2002) realizan una proyección del precio del cobre usando nivel de inventario OECD obteniendo una proyección dinámica de corto plazo con un R-cuadrado ajustado de 0,935.

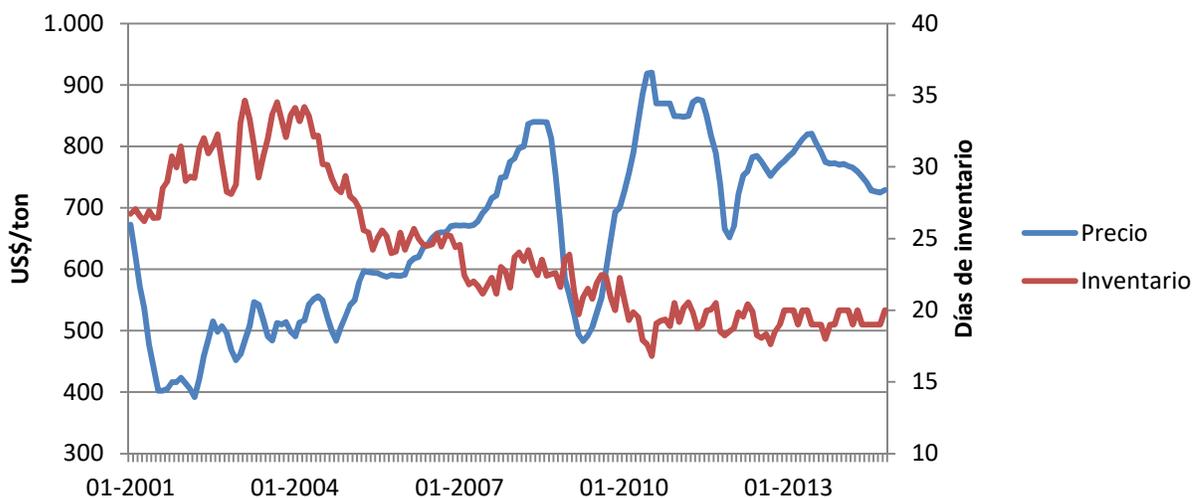
En el caso de la proyección del precio de la celulosa también se ha realizado una cantidad considerable de estudios. En particular, (Toppinen, y otros, 1996) investigan la relación de largo plazo entre el precio de exportación de la fibra finlandesa y el nivel de inventarios NORSCAN (inventarios de productores norte americanos y escandinavos). Se obtiene un equilibrio de largo plazo, cuyo poder explicativo sólo es aceptable bajo condiciones de mercado estables y se genera una proyección de corto plazo con un R-cuadrado de 0,61.

A pesar de que en estos estudios se suele utilizar el volumen de inventario, en el caso de la celulosa el indicador relevante son los días de inventario, definidos como la cantidad de días que el inventario actual permanecerá en stock basado en los niveles de ventas. Este es el indicador al que hacen seguimiento bancos de inversión y productores forestales. En el presente estudio se analiza si los días de inventario globales tienen efecto sobre el precio de la celulosa.

Existen diferentes entidades que reportan mensualmente los inventarios en volumen o en días. A continuación se describen los principales indicadores.

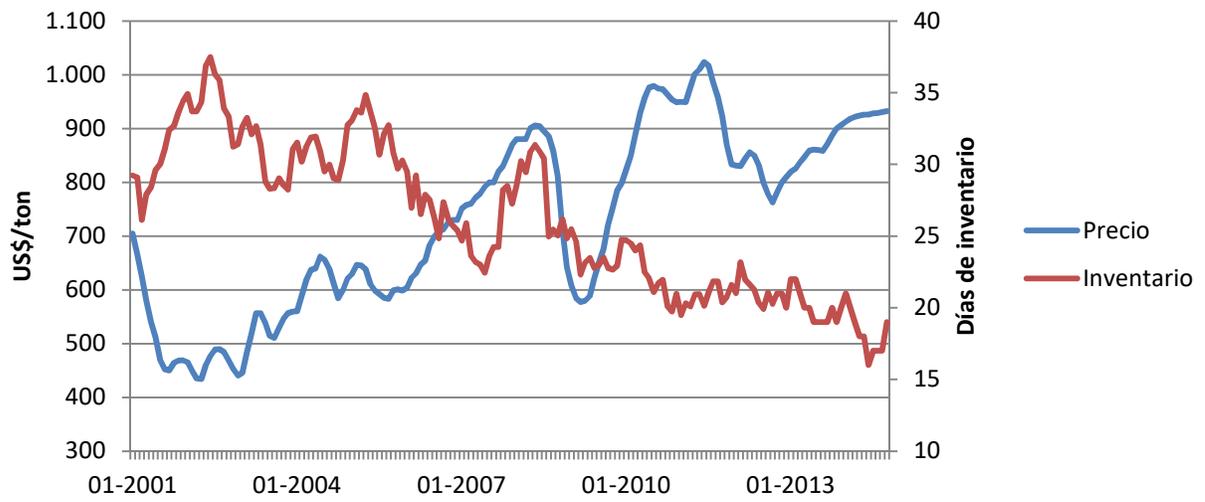
UTIPULP recopila los datos de inventario y consumo de BCP de 11 países europeos que representan una parte importante de la demanda global (alrededor de 11 millones de toneladas de un total de 44 millones demandadas el año 2014). Con esta información reporta los niveles de inventario, consumo y los días de inventario de demanda. Dichos índices son reportados a nivel desagregado por tipo de fibra. Los Gráficos 12 y 13 muestran los días de inventario de demanda versus los precios de las fibras estudiadas para el periodo enero 2001 hasta octubre 2014.

Gráfico 12: Días de inventario de demanda versus precio para fibra corta



Fuente: Elaboración propia

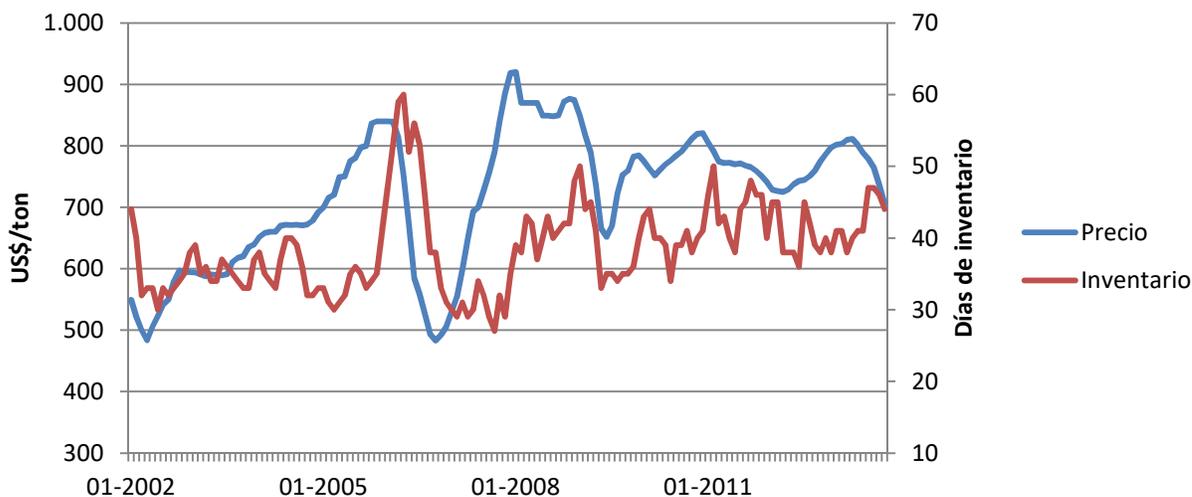
Gráfico 13: Días de inventario de demanda versus precio para fibra larga



Fuente: Elaboración propia

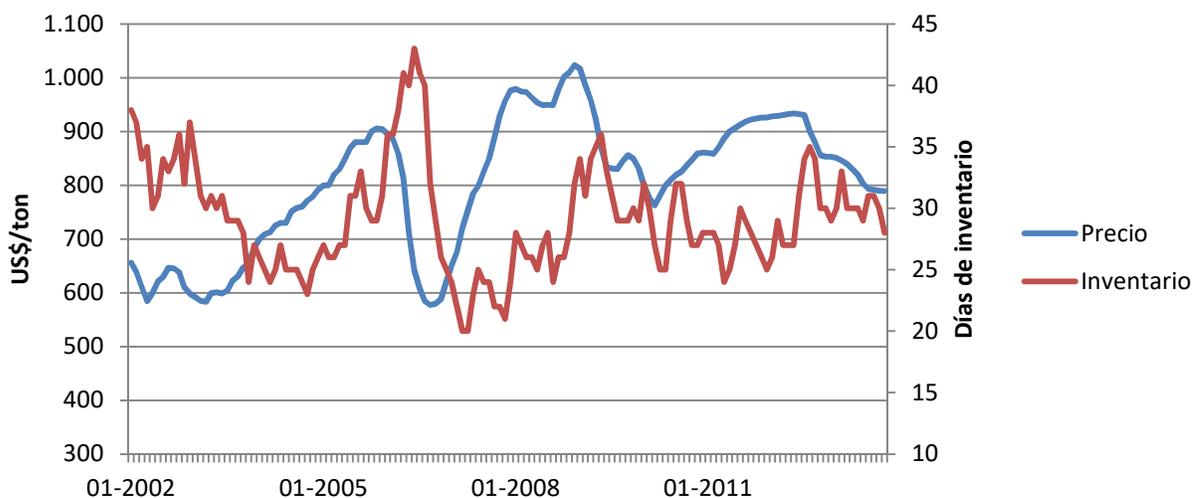
Pulp and Paper Products Council (PPPC) publica los días de inventario de oferta de los 20 principales productores de celulosa a nivel global. Hawkins Wright Ltd. (HW) reporta un indicador similar desagregado por grado de fibra, pero su metodología es menos explícita. Los Gráficos 14 y 15 muestran los días de inventario de oferta versus los precios de las fibras estudiadas para el periodo enero 2002 hasta abril 2016.

Gráfico 14: Días de inventario de oferta versus precio para fibra corta



Fuente: Elaboración propia

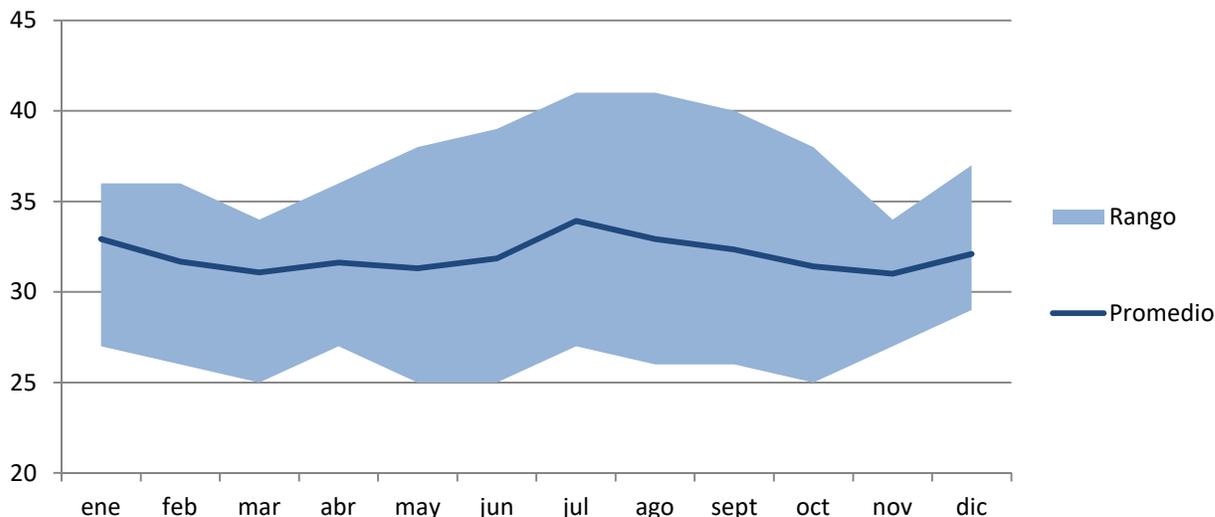
Gráfico 15: Días de inventario de oferta versus precio para fibra larga



Fuente: Elaboración propia

Se analiza el promedio mensual anual de los días de inventario de demanda para verificar si estos presentan estacionalidad. Para esto se suprime el periodo agosto 2008 a marzo 2009 y así corregir el efecto de la crisis. Según el Gráfico 16 los días de inventario de oferta presentan una estacionalidad marcada en enero y julio.

Gráfico 16: Estacionalidad días inventario de oferta BCP 07-2004 a 04-2016



Fuente: Elaboración propia

El mercado forestal es un mercado complejo que involucra diversos productos que son transados internacionalmente, por lo que se pueden identificar distintas perturbaciones que repercuten en los precios, como por ejemplo variaciones en los costos de embarque.

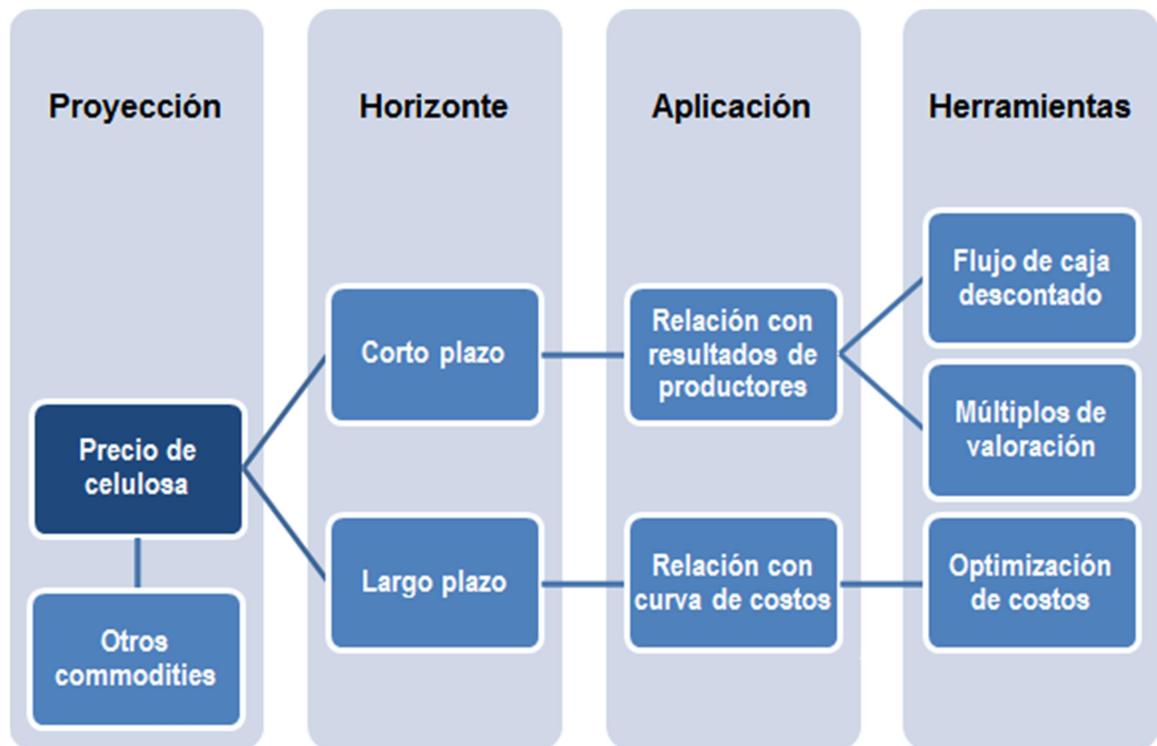
Resulta costoso realizar un análisis permanente sobre todos los indicadores que podrían afectar el precio global, ya que un modelamiento más complejo no necesariamente entregará mejores resultados que uno más simple. Por esto último resultaría interesante investigar el efecto de indicadores globales que puedan incorporar información sobre ciclos económicos y otras perturbaciones que puedan incidir en el precio tanto en el corto como en el mediano plazo.

A pesar de que los datos descritos son los más utilizados en el modelamiento del sector forestal, dado que la fibra es un commodity, existe gran variedad de información complementaria que actualmente no es utilizada con dicho propósito.

Por otro lado, al utilizar indicadores globales, el análisis podría replicarse para la proyección de distintos commodities, resultado de interés para entidades financieras quienes pueden relacionar las proyecciones obtenidas con los futuros ingresos de empresas relacionadas.

Existen diversas posibilidades de aplicación para una proyección del precio de la celulosa. En particular se identifican dos tipos de empresas para los que podría ser de interés un estudio del mercado forestal: productores de celulosa y entidades financieras. El Diagrama 1 muestra algunas de las posibilidades o potenciales aplicaciones de una proyección de precios.

Diagrama 1: Potenciales aplicaciones y estudios



Fuente: Elaboración propia

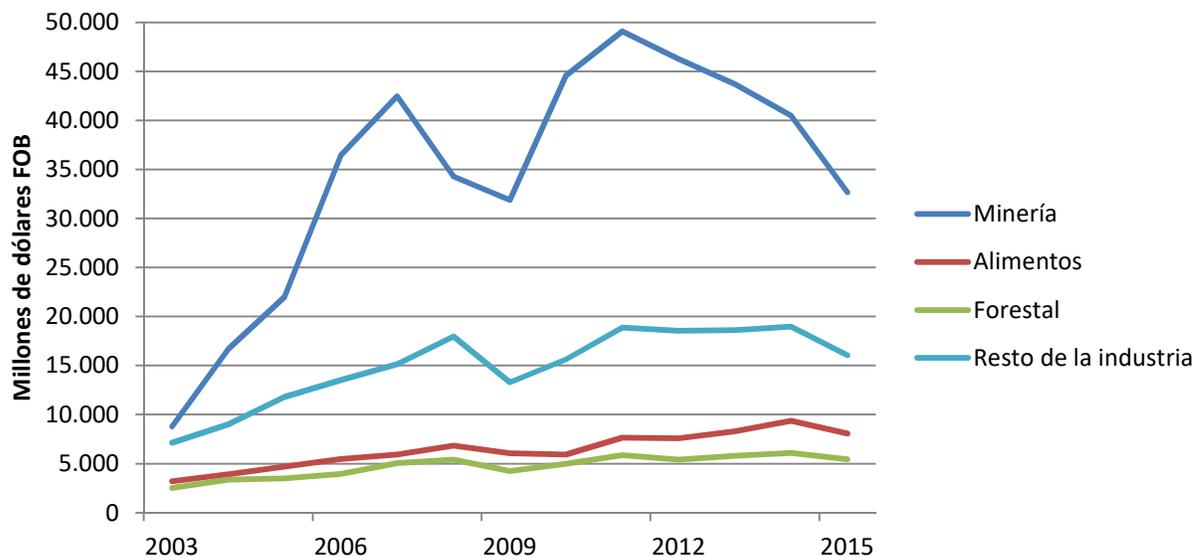
Las fluctuaciones del precio de las fibras inciden sobre los ingresos y utilidades de los principales productores chilenos, dado que la celulosa representa una de las principales fuentes de ingreso.

Al relacionar las proyecciones para los precios con indicadores financieros o bursátiles se puede obtener una estimación de sus resultados esperados. En particular, mediante el uso de múltiplos financieros, se puede obtener una estimación del precio objetivo para las acciones de las empresas involucradas, información relevante para entidades financieras, quienes generan recomendaciones para sus clientes en base a dichas estimaciones.

Por otro lado, al relacionar las proyecciones de precio con la curva de costos observada y posibles escenarios de apreciación o depreciación del precio del dólar con respecto a la moneda local, se podría estimar la indexación óptima al precio del dólar sobre la estructura de costos. Esta información sería de interés para productores de celulosa, quienes podrían usarla dentro de su planificación.

El sector forestal tiene una gran importancia dentro de la industria nacional. Junto con la minería y el sector de alimentos son los que generan más exportaciones de bienes. El Gráfico 17 muestra la evolución de la exportación de bienes entre los años 2003 y 2015.

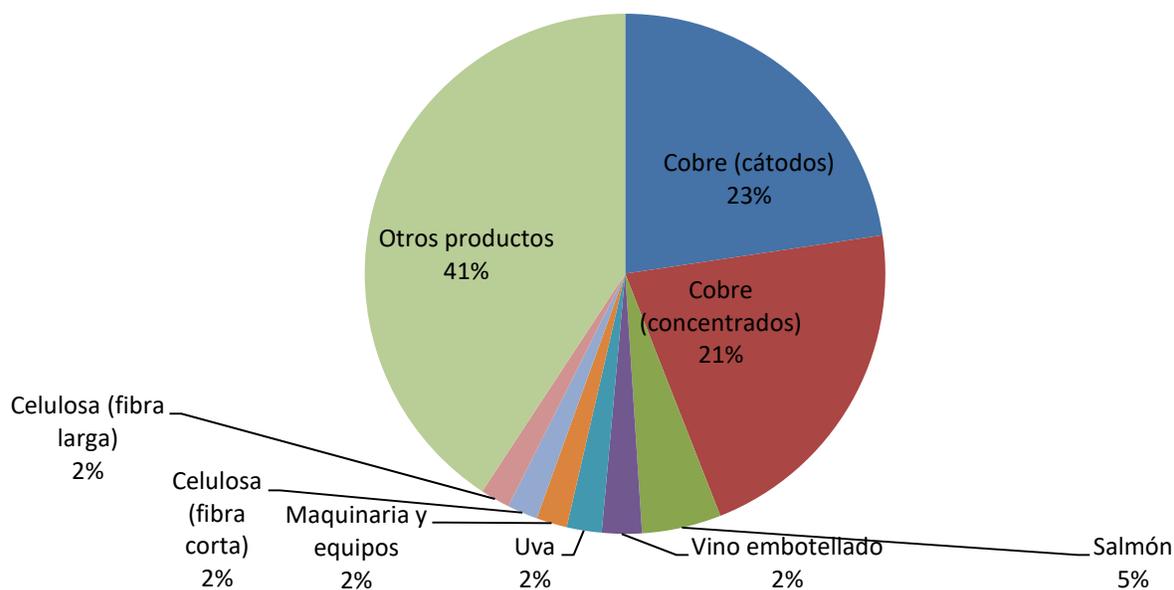
Gráfico 17: Exportación de bienes 2003 a 2015



Fuente: Elaboración propia en base a datos del Banco Central de Chile

Las fibras bajo estudio representan un 3,7% de las exportaciones del año 2015, como se muestra en el Gráfico 18.

Gráfico 18: Principales productos exportados 2015



Fuente: Elaboración propia en base a datos del Banco Central de Chile

1.2 Determinantes del precio de la celulosa

Existen distintos determinantes que influyen en el precio de la fibra:

- **Inventarios:** distintas entidades registran indicadores de inventarios globales o regionales para productores y consumidores de celulosa a nivel global. Dichos inventarios representan desequilibrios entre oferta y demanda, y pueden influir en el precio en el corto y mediano plazo
- **Efecto sustitución entre fibras:** existe un efecto sustitución entre las fibras en el sentido de que en la producción de papeles y tissue se puede reemplazar parte de la fibra corta usando fibra larga, por lo que es necesario tomar en cuenta la valoración relativa, ya que dicha sustitución se gatilla cuando el spread varía erráticamente. Es necesario destacar que este trade-off no se produce en el sentido inverso, ya que no es posible usar fibra corta en embalajes
- **Precio del algodón:** en algunas telas, parte del algodón es sustituible por fibra de celulosa, por lo que un aumento excesivo del precio del algodón puede gatillar un aumento en la demanda de fibra
- **Precio de papel reciclado:** a pesar de que se podría pensar que el papel reciclado actúa como sustituto de la fibra de celulosa, se ha demostrado que ambos actúan como complementos, ya que la fibra del papel reciclado se va acortado en los distintos procesos a los que es sometida, por lo que se puede reutilizar sólo hasta 4 o 6 veces, haciendo que siempre sea necesario el uso de fibra nueva (Szabó, y otros, 2009). Su precio siempre es inferior al precio de la fibra virgen y al aumentar no es conveniente su uso debido a que es costoso blanquear el papel y deshacerse de la tinta
- **Crecimiento de la demanda mundial de papel:** es un mercado con un crecimiento constante que bordea el 2 o 3% anual. Esto se debe principalmente al crecimiento del mercado de tissue y el mercado de embalajes, que se relacionan con el crecimiento de la población
- **Comportamiento de compradores chinos:** China es el principal demandante de fibra a nivel global y produce distintas distorsiones de mercado debido a su alta demanda. Los compradores chinos tienen comportamiento estratégico y aprovechan su poder de mercado generando sobre stock, lo que produce una modificación de la demanda de fibra en el corto plazo y de esta forma se genera un impacto en el precio. Otra distorsión importante se produce durante el año nuevo chino, cuando se reduce la demanda por fibra de celulosa, dado que es feriado nacional y no se realizan transacciones. Al tener un calendario lunar, esta fecha varía todos los años entre enero y febrero.
- **Adiciones de capacidad:** se realizan anuncios de adiciones periódicamente, los cuales no siempre se realizan en la fecha anunciada o simplemente no son ejecutados; además, las plantas nuevas comienzan a utilizar su capacidad gradualmente, existiendo un periodo de ramp up, lo que añade dificultades a la proyección del precio con este método. Dichos anuncios podrían ejercer un efecto psicológico sobre los involucrados, pudiendo tener impactos sobre el precio.

En el presente estudio sólo se analizarán los inventarios y el efecto sustitución entre fibras, dado que se cuenta con datos suficientes y su relación se puede modelar mediante relaciones simples, que en el caso de los inventarios han sido ampliamente estudiadas.

Si bien resultaría interesante estudiar los efectos de los otros determinantes sobre el precio, en el caso del precio del algodón, precio del papel reciclado y crecimiento de demanda global, que son los determinantes para los que podrían obtenerse datos, su efecto resultaría marginal en comparación a otras variables comúnmente usadas en la proyección de commodities a través de modelos más estructurales. En el caso del precio algodón, su aplicación se ve demasiado limitada, ya que sólo un shock importante sobre este podría tener efectos sobre el precio de la celulosa, lo que no es suficiente para generar una proyección del precio. Por otro lado, en el caso del papel reciclado, si bien un aumento en su precio puede gatillar un efecto sustitución en un horizonte inmediato, al ser productos complementarios, ambos deben verse afectados por efectos externos y no tiene sentido utilizar uno para la proyección del otro. La demanda agregada se puede obtener de forma anual, pero con una publicación muy retrasada, lo que limita tanto la proyección como su evaluación; además, como lo muestra el Gráfico 1, si bien la demanda muestra crecimiento, el mercado no, lo que limita su aplicación.

En el caso de los compradores chinos, existen al menos dos efectos identificados que podrían ser estudiados, pero lamentablemente los datos correspondientes son privados y no se tiene acceso a estos. Con respecto a las capacidades, la actualización de las adiciones tiene una frecuencia relativamente baja, y aunque estas pueden tener efectos sobre el precio, este podría no ser suficiente para explicar las fluctuaciones en el corto plazo.

Adicionalmente se desea estudiar el efecto de factores externos al mercado que podrían tener un alto impacto sobre el precio y actualmente no han sido considerados al momento de estudiar sus fluctuaciones. Dada la disponibilidad de datos podría generarse una proyección estable que pueda representar las variaciones en el corto plazo.

2 Marco Conceptual

Se han llevado a cabo variados estudios para modelar el sector forestal, la mayor parte de estos en la zona Norte de Europa. (Toppinen, y otros, 2010) realizan una revisión bibliográfica de los principales estudios realizados en el tema entre los años 1998 y 2007, explicando tanto las técnicas como la finalidad de dichos estudios.

A grandes rasgos las investigaciones se pueden dividir en dos principales tipos de metodologías: modelos econométricos y modelos técnicos. Los primeros no reaccionan bien ante cambios estructurales en los mercados y presentan problemas estadísticos en la estimación de parámetros debido a cantidad y calidad de los datos utilizados, mientras que los segundos son modelos altamente complejos que resultan en falta de transparencia sobre las estimaciones obtenidas, y no manejan los cambios estructurales de mejor manera que los modelos econométricos.

Los modelos técnicos se pueden clasificar en tres categorías: los que estudian el desarrollo futuro del sector forestal bajo supuestos pertinentes (crecimiento del PIB, crecimiento de bosques y cambios tecnológicos), los que estudian los efectos de la conservación de bosques y los que se concentran en la bioenergía forestal. Estos corresponden a modelos multivariados de alta complejidad que utilizan programación lineal y simulaciones, entre otras técnicas, transformando los resultados de estudios econométricos en escenarios para el desarrollo futuro de la industria forestal. Algunos de estos estudios han intentado proyectar la evolución del precio de productos forestales para un país en particular, pero se ha demostrado que, pese a su complejidad, no entregan resultados mejores que los modelos econométricos.

En el caso de los modelos econométricos, estos se han enfocado en cinco temas principales: modelamiento de oferta y demanda de productos, análisis de precios y testeo de integración de mercados, demanda de mercado y proyección de precios, modelamiento de localización de la industria, y tecnología de producción y demanda de factores.

Para la proyección de mercados y precios del sector forestal, los modelos más utilizados corresponden a modelos AR(1), ARIMA, Kalman-filter y VAR, siendo este último el que entrega una leve mejora en la precisión. Una de las metodologías más novedosas consiste en un modelo multivariado basado en factores de máxima autocorrelación y movimientos comunes, pero no se logra obtener mejoras importantes en el corto plazo.

Los resultados sugieren que los modelos econométricos y el uso de información de expertos son necesarios para complementar la evidencia utilizada en métodos más sofisticados, ya que estos logran extraer información cuantitativa de la estructura y comportamiento del mercado.

(Abildtrup, y otros, 1999) recopilan una serie de estudios basados en análisis de series de tiempo para modelar mercados forestales europeos. La mayor parte de los estudios existentes realizan estimaciones de precios para mercados locales usando principalmente datos de cantidad transada, precio de exportación/importación,

inventarios de productores y tasa de cambio involucrada. En estos estudios se estima la elasticidad del mercado y otros indicadores claves para la industria al considerar el logaritmo de las variables mencionadas. Posteriormente se usan los vectores de cointegración para estimar las curvas de oferta y demanda del sistema, generando proyecciones en el largo plazo.

Uno de los commodities que ha concentrado más estudios de proyecciones de precio es el petróleo. (Bashiri Mehmiri, y otros, 2013) realizan una revisión bibliográfica de los principales estudios realizados en el tema. En esta revisión los estudios se dividen en dos categorías: estudios cuantitativos y estudios cualitativos.

Los métodos cuantitativos entregan buenos resultados en el mediano-corto plazo, y a su vez se dividen en métodos econométricos y métodos computacionales. Dentro de los métodos econométricos existen tres grupos: modelos basados en series de tiempo, modelos financieros y modelos estructurales. Los más utilizados son los modelos basados en series de tiempo, en segundo lugar los modelos financieros, y por último los modelos estructurales y computacionales.

Cuando los datos exhiben un patrón sistemático como autocorrelación, el número de posibles variables explicativas es amplio y sus interacciones sugieren un modelo estructural complejo, o la proyección de la variable dependiente requiere la proyección de las variables independientes, los modelos de series de tiempo son los más adecuados. Estos incluyen tres principales categorías: modelos naive, modelos de suavizamiento exponencial y modelos autorregresivos, como por ejemplo ARIMA y GARCH.

Los modelos de series de tiempo presentan una capacidad de proyección limitada en el mediano y largo plazo y, a pesar de que son capaces de ajustar la volatilidad del precio del petróleo, no se debería usar sólo un modelo para realizar una proyección, ya que distintos modelos podrían variar su desempeño en diferentes horizontes. La principal debilidad de estos modelos es que no son adecuados para describir los componentes no lineales que presentan los precios, los que implican que un pequeño shock en la economía puede tener un impacto impredecible en el precio y su volatilidad.

Los modelos financieros estiman la relación entre el precio de un commodity y el precio de su respectivo futuro, investigando si este tipo de contrato es un estimador insesgado y eficiente. Este tipo de modelo no logra predecir con precisión ya que la relación entre el precio futuro y el precio del futuro varía en el tiempo y no logra ser explicada con los modelos actuales.

En los modelos estructurales, las fluctuaciones del precio del petróleo son modeladas en función de variables fundamentales. Las variables comúnmente utilizadas son nivel de inventario, consumo y producción de petróleo, actividad económica, tasas de interés, entre otras. Su capacidad de predicción se reduce drásticamente al existir saltos en el precio.

Los métodos computacionales son aproximaciones no lineales que han ganado notoriedad recientemente; sugieren que las técnicas tradicionales utilizadas para proyección no serían las más adecuadas en la proyección de series de tiempo no lineales, dado que se este tipo de técnica asume que las series son generadas por

procesos lineales. Los principales métodos utilizados son Artificial Neural Networks y Support Vector Machines, dada su alta capacidad de clasificación y reconocimiento de patrones. A pesar de esto, su desempeño en contraste con modelos basados en series de tiempo depende del horizonte de evaluación y no presentan una predicción categóricamente superior.

Finalmente, los métodos cualitativos estudian el impacto de eventos infrecuentes en el precio del petróleo, como por ejemplo guerras o desastres naturales utilizando factores cualitativos como especulación o factores políticos. Son los métodos menos utilizados, pero al ser utilizados en conjunto con otras técnicas, reportan mejoras en la capacidad de predicción.

Diversos estudios se han enfocado en identificar factores comunes que influyan en los precios de distintos commodities, siendo tipos de cambio, indicadores macroeconómicos y precios de commodities relacionados las variables más utilizadas. En particular un estudio realizado por (Kagraoka, 2015) analiza los factores comunes que influyen en las variaciones de precios de 84 commodities. Se identifican cuatro factores que explican en promedio un 68,2% de la varianza de los commodities estudiados; dichos factores son: la tasa de inflación estadounidense, el precio del petróleo, el índice de producción industrial global y el índice accionario global. En particular, la fibra larga es uno de los commodities considerados, y para este se obtiene que los cuatro factores comunes explican más de un 50% de su varianza. A pesar de que no se obtiene un resultado tan potente como para otros commodities, este resulta de interés, ya que los factores obtenidos podrían ayudar a explicar la evolución del precio de la celulosa.

Finalmente, existe un debate sobre si los precios de commodities sirven para predecir tipos de cambio, o si la causalidad es inversa. Se ha publicado una gran variedad de estudios al respecto sin llegar a un consenso con respecto a este tema. En particular, (Chen, y otros, 2008) estudian la relación dinámica entre variaciones de precios agregados de commodities y fluctuaciones de tipos de cambio de cinco pequeñas economías en desarrollo que son exportadores de estos (Australia, Canadá, Chile, Nueva Zelanda y Sudáfrica). Se obtiene como principal resultado una relación robusta interpretable como causalidad desde tipos de cambio hacia precios de commodities, siendo los primeros de gran utilidad en la proyección de precios de los últimos.

En la siguiente tabla se muestra un resumen de la revisión bibliográfica, considerando los principales modelos, sus usos y resultados:

Tabla 1: Resumen revisión bibliográfica

Técnica	Modelos de factores de producción	Machine Learning	Modelos de factores comunes	Modelos Econométricos
Usos	Analizar el impacto de políticas climáticas, estimar cambio tecnológico, consumo de energía y emisiones de CO ₂ .	Proyectar precios y patrones de oferta y demanda en el corto plazo.	Estudiar movimientos comunes entre commodities e identificar indicadores que influyen en sus precios.	Proyectar precios y otras variables del mercado. Identificar relaciones que brinden un entendimiento más profundo del mercado.
VARIABLES	Producción, capacidades, elasticidad precio, características de los bosques, régimen de cosecha, entre otros. Data anual, desagregada por país y región.	Patrones de consumo de oferta y demanda, precios históricos. Data desagregada desde anual hasta diaria en algunos casos.	Precios de commodities e indicadores macroeconómicos. Generalmente data trimestral o mensual en nivel.	Endógenas I(1) propias del mercado (cantidad, precio, inventarios). En log o en nivel, generalmente mensuales o trimestrales.
Observaciones	Modelos complejos que no necesariamente influyen en el precio.	No se utilizan para proyecciones de largo plazo.	Fines investigativos, no apropiados para proyección.	Dan resultados tan potentes como otras técnicas de proyección.
Resultados	Proyecciones de largo plazo realizadas mediante simulaciones y análisis de sensibilidad.	Modelos relativamente complejos, con buen desempeño. En el corto plazo pueden dar mejores resultados que modelos econométricos.	Se encuentran factores que explican en gran proporción las variaciones de distintos commodities.	Co-integración fuerte dentro de los mercados, curvas de oferta y demanda inferidas de las relaciones de cointegración, proyecciones de corto y largo plazo.

Fuente: Elaboración propia

3 Descripción del proyecto

3.1 Justificación

Los precios de commodities son complejos de estimar, dado que se caracterizan por periodos de alzas y bajas en su cotización debido a shocks de demanda. Su valoración es compleja debido a la alta variabilidad, y existen pocas estimaciones realizadas para la fibra de celulosa al ser un commodity poco común.

Los estudios encontrados en la literatura sólo se enfocan en investigar los productos de fibra larga que son principalmente producidos en Europa y Norte América. En el presente estudio se incorpora también el estudio de la fibra corta y el ratio entre ambas para entregar información complementaria para estudios posteriores.

En este estudio se pretende generar un modelo para proyectar el precio de ambas fibras considerando información de mercado. La proyección obtenida resulta de interés para dos tipos de involucrados: productores de celulosa y bancos de inversión. En el caso de los primeros, con los resultados obtenidos se podría realizar tanto una mejor planificación de la capacidad necesaria (fibra larga o fibra corta) como una optimización de la indexación de costos sobre el precio del dólar. En el caso de las entidades financieras, con los resultados obtenidos se puede estimar los resultados esperados de las empresas productoras al relacionar los precios con los ingresos esperados; de esta forma, las entidades financieras podrán estimar el desempeño y nivel de endeudamiento de las empresas y así tomar mejores decisiones de inversión.

Actualmente no se realizan proyecciones para el precio de la celulosa en ninguna de las empresas mencionadas, ya que resulta costoso realizar este tipo de estudios, pero estas podrían ayudar a mejorar los resultados actuales o aportar información para la toma de decisiones pertinentes.

En este estudio se desarrollan tres modelos distintos, comparando su precisión en distintas ventanas temporales. Para esto se eligen modelos que sean capaces de proyectar el precio de la fibra en distintos horizontes además de explicar sus variaciones relacionándolas con indicadores clave de la industria e indicadores macroeconómicos, con el fin de obtener una proyección simple, capaz de integrar información pública de distintas fuentes que permita explicar las fluctuaciones de manera transparente. Por otro, la selección de variables exógenas permite desarrollar un modelo eventualmente replicable en otros mercados.

3.2 Objetivos

3.2.1 Objetivo General

Desarrollar un modelo capaz de predecir y explicar las fluctuaciones del precio de la fibra de celulosa a corto y mediano plazo, usando indicadores macroeconómicos y datos de inventarios globales.

3.2.2 Objetivos específicos

A partir del objetivo general declarado, se desprenden los siguientes objetivos específicos:

- Estudiar el mercado global de la fibra de celulosa, identificando determinantes que influyen en su precio
- Evaluar relaciones entre las variables y sus efectos en el precio de la fibra
- Evaluar desempeño de las proyecciones estimadas en distintos horizontes
- Analizar la estabilidad y robustez de la proyección desarrollada

3.3 Alcances

Una de las principales dificultades de este tipo de estudios se relaciona con la obtención de los datos adecuados, ya que diversas variables propias del mercado son datos privados. Por esto se usarán datos mensuales usando indicadores que se encuentren disponibles públicamente o puedan ser obtenidos a través de productores de celulosa.

En la realización de este tipo de estudios, al incluir variables explicativas externas, es necesario contar también con proyecciones para dichas variables. Se destaca que queda fuera del alcance de este estudio desarrollar proyecciones para las variables que resulten significativas y, en el caso de obtener un modelo capaz de predecir las fluctuaciones en el precio de la celulosa, para la implementación se usarán proyecciones para las variables externas desarrolladas por entidades especializadas.

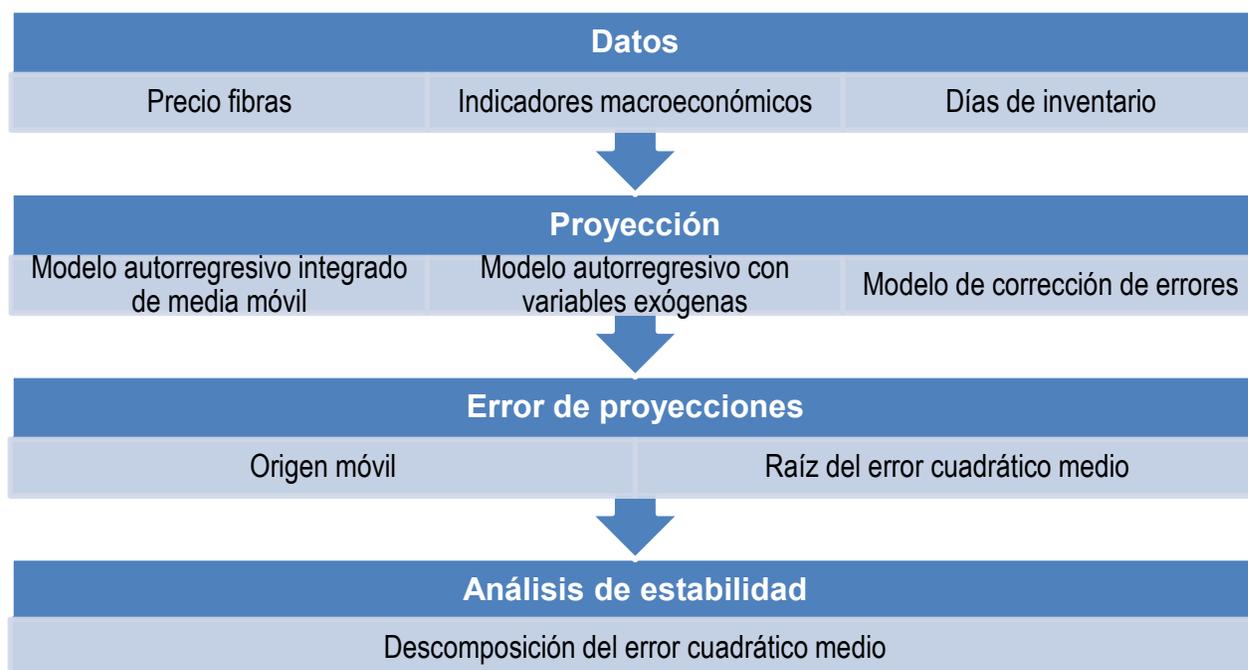
Finalmente, dados los cambios estructurales ocurridos en el sector forestal, el estudio se limita al análisis del periodo enero 2000 a abril 2016.

4 Metodología

Se realizan estimaciones usando tres modelos econométricos distintos, los cuales son seleccionados en base a su capacidad predictiva en distintos horizontes y su capacidad de explicar las fluctuaciones importantes en el precio incorporando distintas fuentes de perturbaciones.

El siguiente diagrama muestra la estructura del estudio desarrollado:

Diagrama 2: Estructura estudio



Fuente: Elaboración propia

A continuación se describen los tres modelos utilizados, con sus respectivos procedimientos y fundamentos teóricos. Se incorpora una breve introducción a modelamiento de series de tiempo y su aplicación a la proyección de precios de commodities, adicionalmente se explican las técnicas usadas para su evaluación y análisis de estabilidad.

4.1 Series de tiempo y desafíos en la proyección de commodities

La proyección de commodities es un tema que ha probado dificultades debido a varios factores: sus precios presentan una alta varianza, sus saltos son erráticos en presencia de shocks en los mercados, no hay un consenso sobre los factores externos que influyen en sus precios, y además se transan en mercados internacionales donde existen variadas fuentes de fluctuación.

Se han realizado variados estudios en torno a precios de commodities, pero no existe un consenso sobre las variables que influyen en sus precios, y, al ser mercados internacionales complejos, existen distintas fuentes de información sobre las que se pueden realizar diversos tipos de análisis, pero resulta imposible identificar todas las fuentes de fluctuación.

Adicionalmente se destaca que algunos mercados de commodities son relativamente estáticos, pero otros tienen un dinamismo mucho mayor, y aunque se logre identificar las fuentes de fluctuación y cuantificar su impacto sobre el precio, esta relación podría no resultar estable en el tiempo, provocando que los estudios realizados queden obsoletos en el corto o mediano plazo.

En particular, con respecto a las proyecciones obtenidas en distintos estudios, se observan métricas de ajuste variadas, las que no siempre muestran la precisión adecuada para el uso de las proyecciones obtenidas en distintas aplicaciones.

Los modelos más utilizados en proyección de commodities son los modelos de series de tiempo, los cuales, a pesar de su sencillez, suelen ofrecer resultados mejores que modelos más complejos en el corto plazo.

Una serie de tiempo es una secuencia de observaciones (muestras) de una variable particular donde cada observación se registra dentro de un tiempo específico. Dichas observaciones se denotan por Y_1, Y_2, \dots, Y_n , donde Y_t es el valor tomado por la serie en el instante t . En el caso de las series de tiempo discretas, estas observaciones son registradas en periodos de tiempo regulares.

Una serie de tiempo resulta pronosticable cuando su comportamiento histórico revela patrones o relaciones que permitan estimar su evolución futura. Existen distintos tipos de series de tiempo, pero los que resultan de interés para este estudio son:

- **Serie estacionaria:** su media y varianza permanecen constantes en el tiempo y su covarianza es sólo función del rezago. Gráficamente la serie oscilará en torno a su media con una amplitud relativamente constante. Estas características permiten modelar el proceso subyacente a través de una ecuación que utilice los datos pasados, ya que una perturbación aleatoria tiene efectos transitorios.
- **Serie no estacionaria:** sus características de media, varianza y covarianza varían en función del tiempo y los efectos de una perturbación sobre la variable son permanentes, lo que dificulta su modelamiento. Sin embargo, al diferenciar series no estacionarias, en muchos casos se obtienen series estacionarias.

Dicho lo anterior es necesario definir la diferenciación de una serie de tiempo. Se define la primera y n -ésima diferencia de una serie de la siguiente manera:

$$\Delta Y_t = Y_t - Y_{t-1}$$

$$\Delta^n Y_t = \Delta^{n-1} Y_t - \Delta^{n-1} Y_{t-1}$$

Las series integradas son un caso particular de series no estacionarias. En este caso, si al diferenciar d veces una serie integrada se obtiene una serie estacionaria, se dice que la serie es integrada de orden d y se denota $I(d)$. Generalmente las series económicas resultan ser series integradas de primer orden.

La econometría tradicional se basa en el supuesto de que las series son estacionarias, pero dado que las series de tiempo económicas suelen ser variables integradas, es necesario realizar otro tipo de análisis, ya que en estas, los estimadores de mínimos cuadrados ordinarios no son consistentes y cualquier shock sobre la serie tendrá un impacto permanente. Esto tiene importantes consecuencias, tanto para la formulación de los modelos como para la distribución de sus estimadores

Para verificar la estacionariedad de una serie existen distintos test de hipótesis que permiten determinar con cierto nivel de confianza si una serie es estacionaria o no. El test más utilizado es el de Dickey-Fuller Aumentado (ADF), el cual prueba la presencia de autocorrelación entre los términos de error.

El test mencionado asume que los datos son generados por un proceso autorregresivo de orden p :

$$y_t = \alpha_0 + \sum_{i=1}^p \alpha_i y_{t-i} + \varepsilon_t$$

Después de una manipulación algebraica de la relación anterior, y la incorporación de rezagos de la variable diferenciada para que desaparezca la autocorrelación entre los residuos y así asegurar que son ruido blanco, se obtiene la siguiente relación:

$$\Delta y_t = \alpha_0 + \alpha_2 t + \gamma y_{t-1} + \sum_{i=2}^p \beta_i \Delta y_{t-i+1} + \varepsilon_t$$

Bajo este test la hipótesis nula es la existencia de raíz unitaria o no estacionariedad de la variable versus la hipótesis alternativa de estacionariedad. El coeficiente de interés en la ecuación anterior es γ , dado que si $\gamma = 0$ se acepta la hipótesis nula. La existencia tanto de intercepto como de tendencia para testear la existencia de raíz unitaria depende de la significancia estadística de los parámetros α_0 y α_1 respectivamente. La cantidad de rezagos a incluir corresponde al que minimiza el valor del criterio de información de Schwartz obtenido en la regresión.

Para rechazar la hipótesis nula, se debe considerar el estadístico que se obtiene por medio del test t estándar. La distribución límite del estadístico es una función de procesos brownianos, por lo que no son válidos los valores críticos usuales para el t-estadístico. Por esto último el estadístico se debe comparar con las tablas de Dickey-

Fuller, las cuales son calculadas mediante simulaciones numéricas de Monte Carlo para diferentes tamaños de la muestra.

Como regla general se tiene que si el estadístico reportado es menor que el valor crítico indicado y la tendencia resulta no significativa, la variable no tendrá raíz unitaria al nivel de confianza señalado.

Al agregar demasiados rezagos el test de Dickey-Fuller Aumentado pierde poder, mientras que al limitar demasiado la cantidad, los errores podrían no ser ruido blanco. Para minimizar el error asociado se contrasta con el test de Phillips-Perron, uno de los más utilizados en la literatura, el cual corrige la correlación serial sin asociarla a rezagos de la variable de interés.

4.2 Estimación mediante modelo autorregresivo

Los modelos autorregresivos se utilizan para realizar proyecciones de corto plazo sobre series de tiempo utilizando las observaciones anteriores para predecir los datos futuros. Estos identifican patrones existentes en la variable estudiada y entregan buenas predicciones cuando no existen alteraciones bruscas.

El modelo autorregresivo integrado de media móvil o ARIMA(p,d,q) permite describir un valor como una función lineal de los datos históricos y errores. Para esto utiliza tres componentes: autorregresivo o AR, integrado o I, y medias móviles o MA, admitiendo la siguiente representación:

$$\Delta^d x_t = \sum_{i=1}^p \beta_i \Delta^d x_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t$$

La primera componente corresponde al modelo AR(p) que representa una suma ponderada de las observaciones pasadas de la serie, donde p representa el orden del modelo. El segundo término corresponde al modelo MA(q) que representa una suma ponderada de los errores previos, donde q determina el orden del modelo. Con una combinación de los anteriores se obtiene un modelo autorregresivo de media móvil o ARMA(p,q), el cual es útil cuando la serie estudiada es una función de su comportamiento previo y shocks externos. Para la aplicación de los modelos anteriores es necesario que las series bajo estudio sean estacionarias, en el caso de que sea una variable integrada de orden d , esta debe ser diferenciada hasta lograr estacionariedad, lo que determina el orden del componente integrado o I(d).

4.2.1 Metodología de Box-Jenkins

La metodología de Box-Jenkins considera cuatro pasos para el desarrollo del modelo.

Identificación: el objetivo de esta etapa es determinar el orden tentativo de los parámetros p , d y q utilizando el principio de parsimonia.

Con respecto a la selección de los parámetros p y q , la principal herramienta utilizada es el correlograma de la variable bajo estudio, el cual entrega las funciones de autocorrelación y autocorrelación parcial, definidas a continuación.

La función de autocorrelación o ACF por su sigla en inglés (Autocorrelation Function) mide la correlación en función de las covarianzas entre valores de la serie con distinto número de retardos. La autocorrelación con la k -ésima muestra de retardo se define como:

$$\rho_k = \frac{\sum_{i=1}^n (x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

Donde \bar{x} es el promedio de las n observaciones de la serie.

La función de autocorrelación parcial o PACF por su sigla en inglés (Parcial Autocorrelation Function) mide la correlación entre un intervalo t y k retardos removiendo los efectos de las observaciones entre los valores originales. La autocorrelación parcial con la k -ésima muestra de retardo (para $k = 3, \dots, r$) se define como:

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{kj} \rho_j}{1 - \sum_{j=1}^{k-1} \phi_{kj} \rho_j}$$

Donde $\phi_{kj} = \rho_{k-j} - \sum_{l=1}^{j-1} \phi_{kl} \rho_l$ para $k = 2, j = 1, \dots, k-1$, con $\phi_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$ y $\phi_{11} = \rho_1$.

Comparando los correlogramas con parámetros conocidos se obtiene información sobre cuál es el orden que mejor se ajusta a la serie en estudio. La Tabla 2 indica el modelo a utilizar según el correlograma obtenido.

Tabla 2: Propiedades para ACF y PACF

Proceso	ACF	PACF
Ruido blanco	Todo $\rho_x = 0$	Todo $\phi_{xx} = 0$
AR(1): $\beta_1 > 0$	Decaimiento exponencial directo: $\rho_s = \beta_1^s$	$\phi_{11} = \rho_1, \phi_{ss} = 0$ para $s \geq 2$
AR(1): $\beta_1 < 0$	Decaimiento oscilante: $\rho_s = \beta_1^s$	$\phi_{11} = \rho_1, \phi_{ss} = 0$ para $s \geq 2$
AR(p)	Decae a cero, coeficientes pueden oscilar	Puntas hasta rezago p . Todo $\phi_{ss} = 0$ para $s > p$
MA(1): $\theta_1 > 0$	Punta positiva en primer rezago, $\rho_s = 0$ para $s \geq 2$	Decaimiento oscilante: $\phi_{11} > 0$
MA(1): $\theta_1 < 0$	Punta negativa en primer rezago, $\rho_s = 0$ para $s \geq 2$	Decaimiento: $\phi_{11} < 0$
ARMA(1,1): $\beta_1 > 0$	Decaimiento exponencial desde primer rezago, $\text{signo}(\rho_1) = \text{signo}(\beta_1 + \theta_1)$	Decaimiento oscilante desde primer rezago, $\phi_{11} = \rho_1$
ARMA(1,1): $\beta_1 < 0$	Decaimiento oscilante desde primer rezago, $\text{signo}(\rho_1) = \text{signo}(\beta_1 + \theta_1)$	Decaimiento exponencial desde primer rezago, $\phi_{11} = \rho_1$ y $\text{signo}(\phi_{ss}) = \text{signo}(\phi_{11})$
ARMA(p,q)	Decaimiento (directo u oscilante) desde rezago p	Decaimiento (directo u oscilante) desde rezago q

Fuente: Enders

Estimación: los parámetros deben ser estimados para los modelos potenciales, para posteriormente seleccionar el mejor modelo según un criterio adecuado.

En este caso se minimiza el criterio de información de Schwarz o SIC por su sigla en inglés (Schwarz Information Criterion), ya que es el criterio que mejor optimiza el orden de los parámetros según (Koebler, y otros, 1988).

$$SIC = -2\widehat{LL} + k \ln(n)$$

Donde \widehat{LL} es el valor de la log-verosimilitud reportada, n es el tamaño de la muestra utilizada y k la cantidad de regresores considerados. Al penalizar según la cantidad de parámetros se obtiene el modelo más parsimonioso, es decir, el que entrega el mayor poder de predicción usando la menor cantidad de variables independientes posible.

Diagnóstico: una vez seleccionado el modelo final se debe comprobar su adecuación.

Con este fin se verifica que los residuos obtenidos presenten distribución de ruido blanco, es decir, tanto la función de autocorrelación como la función de autocorrelación parcial deben encontrarse dentro de las bandas de confianza del correlograma.

Proyección: finalmente se estiman los pronósticos a través del modelo desarrollado.

4.3 Estimación mediante modelo de indicadores macroeconómicos

Los modelos univariantes no necesitan conocer ninguna relación de causalidad, ni ninguna información relativa al comportamiento de la variable dependiente. En cambio, los modelos multivariantes tratan de explicar el comportamiento de una o más variables dependientes en función de la evolución de otras variables que se consideran explicativas dejando poca variación inexplicada en la perturbación. Estos contemplan de forma explícita la información que aportan las variables causales del fenómeno de interés de acuerdo con una determinada teoría económica, ofreciendo resultados más eficientes y con mayor poder explicativo que los modelos univariantes.

Algunos de los problemas en el uso de este tipo de modelos radican en la dificultad asociada a la elección de las variables explicativas a considerar y la dificultad que conlleva el pronóstico de las mismas, problema que podría ser aún más difícil que el pronóstico de la variable de interés; adicionalmente el desconocimiento de los valores de las variables explicativas en el futuro determina la necesidad de utilizar predicciones para éstas, lo cual incrementa el nivel de incertidumbre con que se realiza la proyección. En el caso particular de los commodities, no existe un consenso sobre las variables que afectan sus variaciones de precio, lo que provoca un bajo uso de modelos multivariantes.

4.3.1 Modelo autorregresivo con variables explicativas

Se desea estudiar la influencia de cuatro indicadores macroeconómicos sobre las variables en estudio; para esto se utiliza un modelo ARX, es decir, un modelo autorregresivo con variables exógenas, el cual se presenta a continuación:

$$x_t = \sum_{i=1}^p \varphi_i x_{t-i} + \sum_{i=1}^n \sum_{j=1}^m \beta_{j,i} y_{j,t-i} + \varepsilon_t$$

Donde x_t corresponde a la variable dependiente en el periodo t , $y_{j,t}$ corresponde al valor de la variable exógena j en el periodo t , p es la cantidad de rezagos considerados para la variable dependiente, m es la cantidad de variables exógenas y n es la cantidad de rezagos considerados para dichas variables.

El uso de variables integradas podría dar lugar a correlaciones engañosamente elevadas; por ende, acorde con la metodología de Box-Jenkins, se diferencian las series hasta lograr su estacionariedad para trabajar con las series diferenciadas.

4.3.2 Stepwise regression

Para la obtención de este modelo se utiliza stepwise regression, el cual es un procedimiento automático donde se seleccionan regresores según algún estadístico previamente seleccionado, como por ejemplo t-estadístico o R-cuadrado ajustado.

Existen dos principales enfoques para su implementación:

- Forward selection: el procedimiento comienza sin regresores en el modelo y su adición se testea variable a variable según su cumplimiento con el criterio establecido.
- Backward elimination: el procedimiento comienza utilizando todas las variables a testar en el modelo, y su eliminación se testea variable a variable según su cumplimiento con el criterio establecido.

En el presente estudio se utiliza forward selection y se seleccionan las variables que resulten significativas con un 90% de significancia estadística según su t-estadístico, es decir, las que obtienen un t-estadístico superior a 1,65, valor crítico para el nivel de confianza mencionado. De esta forma se seleccionan todas las variables que presentan una importante influencia sobre el precio de la fibra, mejorando el ajuste obtenido con el modelo anterior.

4.4 Estimación mediante modelo de corrección de errores

El modelo de corrección de errores es un tipo de modelo estructural que pretende representar la variable en estudio en función de su relación económica con alguna variable explicativa a través de una ecuación.

En la proyección del precio de la celulosa, la variable más utilizada y ampliamente documentada es el inventario global o regional del bien bajo estudio. En este estudio se evaluará la relación entre el precio de la fibra de celulosa y los días de inventario de demanda y oferta descritos anteriormente.

Bajo condiciones particulares, las series de tiempo integradas pueden ser modeladas asumiendo una relación de largo plazo entre ellas. A continuación se explica el procedimiento para verificar si se admite la representación mediante modelo de corrección de errores y el modelo propuesto.

4.4.1 Concepto de Cointegración

Dos series de tiempo integradas están cointegradas cuando tienen el mismo orden de integración y a la vez existe una relación estacionaria entre ambas en el largo plazo, es decir, ambas crecen (o decrecen) a una tasa aproximadamente idéntica.

Cuando las variables tienen una relación de largo plazo en común, aunque estas puedan verse desviadas del equilibrio en el corto plazo, fuerzas económicas o acciones gubernamentales las hacen volver a la relación estacionaria. Este equilibrio de largo plazo resulta útil para generar proyecciones de largo plazo basadas en una relación estable y durable en el tiempo.

Las relaciones de cointegración generan estimadores superconsistentes, es decir, presentan el doble de la velocidad de convergencia hacia los parámetros poblacionales que al estimar relaciones entre variables estacionarias

4.4.2 Procedimiento de Engle-Granger

El procedimiento de Engle-Granger determina la existencia de cointegración entre dos variables relacionadas, con el fin de identificar la relación estacionaria de largo plazo.

En primer lugar se debe verificar que las variables en estudio tengan el mismo orden de integración, luego se debe estimar mediante mínimos cuadrados ordinarios la relación de largo plazo usando las variables en nivel. Para que exista cointegración entre las variables en estudio, los residuos obtenidos de la relación estimada deben tener un orden de integración menor al de las variables bajo estudio. En particular, si se está trabajando con variables integradas de primer orden, los residuos de la relación estimada deben resultar estacionarios.

Para verificar si la relación obtenida no es espuria (es decir, ambas variables se encuentran altamente correlacionadas pero no existe una relación estable de largo plazo) se debe realizar un test de raíz unitaria sobre los residuos estimados.

En este caso, como se está testeando estacionariedad sobre los residuos, se asume que estos son ruido blanco, por lo que no se pueden utilizar los mismos valores críticos usados anteriormente. (Engle, y otros, 1987) tabulan los valores críticos, los cuales son ligeramente más altos que los de las tablas de Dickey-Fuller.

La relación de largo plazo estimada se denomina relación de cointegración. Una vez estimada la relación de cointegración entre las variables, se impone el valor estimado de esta a través del término de corrección de errores, el cual brinda la posibilidad de modificar las perturbaciones observadas en el corto plazo, retornando el sistema a su relación estacionaria.

Una manera análoga de testear cointegración es mediante el test de Johansen, quien desarrolla un test basado en una generalización multivariada del procedimiento de Engle-Granger para testear cointegración cuando se consideran dos o más variables. Dicho test no será aplicado en el estudio, ya que sus resultados no presentan grandes diferencias en comparación a los obtenidos con el procedimiento de Engle-Granger.

4.4.3 Modelo de corrección de errores

Cuando existe cointegración entre dos variables se puede aplicar un modelo de corrección de errores.

El modelo de corrección de errores refleja el cambio en una variable relativo al cambio en la variable independiente, el error de equilibrio previo y cambios previos tanto en las variables independientes como en la dependiente.

Se considera la relación estacionaria de largo plazo entre dos variables $I(1)$:

$$z_t = y_t - \beta x_t$$

Dado que z_t es una variable $I(0)$, se puede considerar un modelo ARMA estacionario. Por simplicidad se ejemplifica un modelo AR(2):

$$z_t = \theta_1 z_{t-1} + \theta_2 z_{t-2} + \varepsilon_t$$

Reemplazando por la relación estacionaria y desarrollando la ecuación se obtiene:

$$(y_t - \beta x_t) = \theta_1 (y_{t-1} - \beta x_{t-1}) + \theta_2 (y_{t-2} - \beta x_{t-2}) + \varepsilon_t$$

$$y_t = \beta x_t + \theta_1 y_{t-1} - \theta_1 \beta x_{t-1} + \theta_2 y_{t-2} - \theta_2 \beta x_{t-2} + \varepsilon_t$$

$$\Delta y_t = \beta \Delta x_t + \theta_2 \beta \Delta x_{t-1} - \theta_2 \Delta y_{t-1} - (1 - \theta_1 - \theta_2) \{y_{t-1} - \beta x_{t-1}\} + \varepsilon_t$$

El último término $(y_{t-1} - \beta x_{t-1})$ corresponde al término corrección de errores, el cual mide la relación de equilibrio de largo plazo, mientras que los rezagos incluidos capturan las dinámicas de corto plazo.

El coeficiente asociado al término de corrección de errores se denomina velocidad de ajuste que presenta el modelo, este debe ser significativo y menor a uno, y el signo negativo asociado a este término se debe a que el modelo de corrección de errores tiende a sobreestimar la variable bajo estudio; por esto, el término de corrección de errores corrige el sesgo, haciendo que la relación vuelva a su estado estacionario. Un coeficiente de 0,2 se interpreta como que el modelo corrige en un 20% en cada periodo el desvío de la relación de largo plazo.

Este modelo se basa en el argumento de que la dinámica de un sistema se puede modelar en el largo plazo mediante un sistema de ecuaciones que varía en torno a una tendencia común, la cual describe su estado de equilibrio.

4.5 Origen móvil y métricas de ajuste

En primer lugar es necesario realizar una distinción en cuanto a las proyecciones realizadas. Existen dos tipos de pronósticos: pronósticos estáticos, los que están basados en la última información efectiva disponible, por lo que están limitados a proyectar a un periodo hacia adelante, y pronósticos dinámicos, los que utilizan el último pronóstico disponible como dato para obtener el siguiente pronóstico, permitiendo la realización de proyecciones a dos y más periodos hacia adelante.

En adelante las proyecciones desarrolladas se refieren solamente a proyecciones dinámicas, dado que el objetivo del estudio es desarrollar una proyección a corto y largo plazo, y, por ende, sólo tiene sentido desarrollar y evaluar proyecciones dinámicas.

Para realizar una evaluación de los modelos propuestos, una de las técnicas más usadas en econometría es la división de la muestra de datos en dos conjuntos distintos: el primero es utilizado para estimar el modelo y el segundo es utilizado para evaluar el ajuste del modelo. En general se habla de evaluación dentro de la muestra cuando esta se realiza sobre el conjunto de entrenamiento y evaluación fuera de la muestra cuando se evalúa sobre el conjunto de evaluación.

La división del conjunto de datos se realiza para replicar el funcionamiento real del modelo estimado, de esta forma la evaluación de su desempeño es un referente válido para futuros errores. Asimismo se reduce el sobreajuste, ya que un modelo con buen desempeño dentro de la muestra no necesariamente da buenos resultados fuera de esta porque puede estar memorizando los datos debido a que los modelos de proyección se construyen minimizando los errores dentro de la muestra; de esta forma la evaluación de errores dentro de la muestra sobrestima el potencial predictivo de las herramientas generadas. Para evitar los problemas mencionados, a lo largo del estudio se realizan evaluaciones fuera de muestra.

Para evaluar el ajuste de los modelos estimados se utiliza la técnica de origen móvil con el fin de reducir el sesgo generado por el punto de origen de la evaluación. Esta es la técnica más utilizada para evaluar el error de las proyecciones estimadas, ya que reduce la influencia de la muestra original sobre la que fue estimado el modelo.

Para su aplicación se considera una submuestra de largo fijo sobre la que se generan estimaciones, las cuales son evaluadas para los periodos posteriores. Se define como el origen de los pronósticos al índice T correspondiente al último dato considerado dentro de la muestra, y de esta forma se obtiene la proyección a $T + 1$, $T + 2$, ..., $T + N$ periodos, donde $T + N$ representa el índice del último dato disponible.

Una vez estimada la proyección, se hace avanzar en un periodo la submuestra elegida para la estimación, es decir, el nuevo origen de la muestra será la observación del precio de la fibra para el periodo $T + 1$ en el paso anterior. Se generan las proyecciones hasta el último periodo disponible, y se vuelve a actualizar el origen de la evaluación.

Gracias a esta técnica, en la que se actualiza sucesivamente el origen de los pronósticos, se generan múltiples proyecciones para cada alcance, a diferencia de las

proyecciones con origen fijo, donde sólo se obtiene un pronóstico para cada alcance, lo que es insuficiente para evaluar su desempeño.

Dado que con la técnica de origen móvil se obtienen múltiples proyecciones para cada alcance, el error de estas se estima calculando la raíz del error cuadrático medio sobre los distintos pronósticos.

La raíz del error cuadrático medio o RMSE por su sigla en inglés (Root Mean Squared Error) es una medida para evaluar el desempeño de la predicción que cuantifica qué tan cerca está el pronóstico de su serie de datos correspondiente, penalizando de manera más fuerte los pronósticos más alejados del dato real.

El RMSE se calcula mediante la siguiente ecuación:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{n}}$$

Donde y_t representa el valor real de la variable para el periodo t , \hat{y}_t representa el valor predicho de la variable para el periodo t y n representa la cantidad de pronósticos.

Usando los pronósticos obtenidos mediante la técnica de origen móvil se pueden obtener métricas de ajuste a distintos alcances contrastando los pronósticos obtenidos para un determinado alcance versus los datos reales.

4.6 Descomposición del error cuadrático medio

El error cuadrático medio o MSE por su sigla en inglés (Mean Squared Error) mide el promedio de los errores de proyección al cuadrado. Para analizar la estabilidad de las proyecciones obtenidas (Theil, 1971) propone la descomposición del error cuadrático medio en términos de media, varianza y covarianza:

$$MSE = \sum_{t=1}^T \frac{1}{T} (p_t - \hat{p}_t)^2 = (\bar{p} - \bar{\hat{p}})^2 + (s_m - s_{\hat{m}})^2 + 2(1 - r)s_m s_{\hat{m}}$$

Donde \bar{p} y $\bar{\hat{p}}$ son las medias del precio observado y predicho, respectivamente; s_m y $s_{\hat{m}}$ son las desviaciones estándar del precio observado y precio, respectivamente; r es la correlación entre el dinero observado y predicho; y T es el número de datos de la muestra.

El error cuadrático medio es descompuesto en las proporciones asociadas a cada uno de los términos mencionados:

- **Media:** el primer término mide la desviación entre las medias de las series observadas y predichas (mide la presencia de errores sistemáticos en la proyección)
- **Varianza:** el segundo término mide la desviación entre varianzas de ambas series (si es alta, la variabilidad de la serie observada es mayor que la de la serie estimada)
- **Covarianza:** el último término mide los errores no sistemáticos.

Si la proyección es robusta, se espera que las proporciones asociadas a la media y varianza sean bajas y la mayor parte del error esté concentrada en la proporción asociada a la covarianza.

5 Fuentes de datos

A continuación se describen los datos utilizados en la realización del estudio.

Para todas las series mencionadas se cuenta con información mensual para el periodo enero 2000 hasta abril 2016, a menos que se destaque lo contrario.

5.1 Precio de la celulosa

Se utilizan los precios de referencia globales publicados para fibra corta y fibra larga publicados por FOEX Indexes Ltd. Estos datos no son públicos y se obtienen de entidades financieras que realizan seguimiento a los productores involucrados.

5.2 Indicadores macroeconómicos

Se consideran cuatro variables globales:

- Inflación estadounidense: variación mensual anualizada del IPC de Estados Unidos
- Precio del petróleo: promedio mensual del precio spot del petróleo WTI
- Índice accionario global: se utiliza el All Country World Index o ACWI, el cual es un índice de capitalización de mercado ponderado que integra información de 23 mercados desarrollados y 23 mercados emergentes entregando una medida del desempeño de la capitalización de mercado global.
- Índice de producción industrial global: se considera el Indicador Líder Compuesto G7 o CLI G7 por su sigla en inglés (Composite Leading Indicator), el cual es un indicador reportado por la OECD utilizado para predecir cambios de tendencia en la actividad económica; la serie utilizada corresponde al G7, es decir, los siete países más importantes de la OECD.

5.3 Días de inventario

Se cuenta con datos de días de inventario, los cuales no son públicos y fueron obtenidos de productores de celulosa. La Tabla 3 resume los datos utilizados.

Tabla 3: Descripción datos de inventario

	Días de inventario de oferta	Días de inventario de demanda
Descripción	Pulp and Paper Products Council (PPPC) y Hawkins Wright Ltd. (HW) recopilan información de los principales productores de celulosa a nivel global para publicar datos de inventarios globales (PPPC considera los 20 principales productores de celulosa).	UTIPULP recopila información de los principales consumidores de celulosa en 11 países de Europa (aprox. 25% de la demanda global) para publicar datos de inventarios desagregados por fibra mensualmente.
Cálculo de series	$Días_t = \frac{Nivel\ inventario_t}{Promedio\ embarcaciones\ diarias\ últimos\ tres\ meses}$	$Días_t = \frac{Nivel\ inventario_t \times 360}{\sum_{i=1}^{12} Consumo_{t-i}}$
Series disponibles	<ul style="list-style-type: none"> • Oferta fibra corta (HW) • Oferta fibra larga (HW) • Oferta BCP (HW) • Oferta BCP (PPPC) 	<ul style="list-style-type: none"> • Demanda fibra corta • Demanda fibra larga • Demanda BCP
Periodo disponible	Julio 2004 a Abril 2016 para las series publicadas por Hawkins Wright y enero 2002 a octubre 2014 para la serie publicada por Pulp and Paper Products Council.	Enero 2001 a Octubre 2014 para todas las series.
Características	Series no presentan tendencia, pero sí una marcada estacionalidad.	Series no presentan estacionalidad, pero sí una tendencia variable a lo largo del periodo, coincidente con la tendencia presente en las series de precios.

Fuente: Elaboración propia

6 Resultados

En la presente sección se describe el desarrollo de los modelos detallados en la sección 4, explicando los procedimientos realizados y resultados obtenidos.

Dado que en todos los modelos es fundamental estudiar la estacionariedad de las series, en la sección 6.1 se realizan los test de raíz unitaria descritos en la sección 4.1, a través de los cuales se obtiene que todas las variables en estudio corresponden a series integradas de primer orden. En las secciones 6.2 a 0 se presentan los principales resultados obtenidos con los tres modelos econométricos utilizados. Finalmente se evalúa el ajuste de las proyecciones desarrolladas y la adecuación de su especificación en las secciones 6.5 y 6.6 respectivamente.

El análisis se realiza para fibra corta, fibra larga y el ratio entre estas, dado que en el caso de obtener buenos resultados con su proyección, esta se puede utilizar para proyectar la fibra con la que se obtengan los peores resultados.

Con los modelos propuestos se obtienen resultados comparables con los que presenta la proyección de referencia correspondiente al modelo autorregresivo, pero estos además logran mejorar el ajuste logrando explicar las fluctuaciones del precio en el corto plazo.

El modelo de indicadores macroeconómicos resulta útil para la proyección de fibra larga, pero no en la de fibra corta, lo que puede deberse a que los indicadores seleccionados no representan las fluctuaciones en mercados productores de fibra corta. Por otro lado, en ambas fibras el modelo de corrección de errores es el que reporta mejores resultados, aunque su aplicación depende de un adecuado estudio de los inventarios utilizados.

La Tabla 4 presenta un resumen de los modelos aplicados y los resultados obtenidos.

Tabla 4: Resumen principales resultados obtenidos

	Modelo autorregresivo	Modelo de indicadores macroeconómicos	Modelo de corrección de errores
Cita	(Khan, 2013)	(Kagraoka, 2015)	(Riis, 1999)
Especificación	$\Delta^d y = \sum_{i=1}^p \phi_i \Delta^d y_{t-i}$ $+ \sum_{i=1}^q \theta_i \varepsilon_{t-1}$ $+ \varepsilon_t$	Δx_t $= \sum_{i=1}^{12} \beta_{1,i} \Delta Infl USA_{t-i}$ $+ \sum_{i=1}^{12} \beta_{2,i} \Delta WTI_{t-i}$ $+ \sum_{i=1}^{12} \beta_{3,i} \Delta ACWI_{t-i}$ $+ \sum_{i=1}^{12} \beta_{4,i} \Delta CLI G7_{t-i}$ $+ \sum_{i=1}^{12} \beta_{5,i} \Delta p_{t-i}^{corta}$ $+ \sum_{i=1}^{12} \beta_{6,i} \Delta p_{t-i}^{larga}$	$\Delta \log(p_t^i)$ $= \delta \Delta \log(p_{t-1}^i)$ $+ \varphi \Delta \log(inv_{t-1}^i)$ $- \alpha (\log(p_{t-1}^i))$ $- \beta \log(inv_{t-1}^i) + c + \varepsilon_t$
Alcance	Corto plazo	Corto-mediano plazo	Largo plazo
Signos esperados	(+ -)	Efecto neto sobre fibras: Inflación USA (-) Precio WTI (+) MSCI ACWI (+) CLI G7 (+)	Precio (+) Inventarios (-) Término de corrección de errores [0, 1]
Resultados obtenidos	R-cuadrado ajustado de 0,44 para fibra corta y 0,59 para fibra larga	R-cuadrado ajustado de 0,66 para fibra corta y 0,75 para fibra larga Se obtienen signos esperados en la mayoría de los casos	R-cuadrado ajustado de 0,53 para fibra corta y 0,64 para fibra larga No se obtiene signo esperado para coeficiente de inventarios
Observaciones	No muestra una mejora en la capacidad de proyección en comparación a los otros modelos. Horizonte de proyección muy limitado	Modelo resulta aplicable en el corto-mediano plazo sólo para fibra larga. Es el que genera peores resultados para fibra corta en todos los alcances considerados	Elasticidad con respecto a inventarios de orden similar a las encontradas en la literatura. Signo inverso puede deberse a efecto látigo por los datos utilizados

Fuente: Elaboración propia

6.1 Estacionariedad series

En todos los modelos descritos es fundamental investigar el orden de integración de las variables a proyectar, las variables a analizar son:

- $\text{Log}(\text{Precio fibra larga})$
- $\text{Log}(\text{Precio fibra corta})$
- $\text{Ratio} = \frac{\text{Precio fibra larga}}{\text{Precio fibra corta}}$

Los precios son tratados en logaritmo para considerar la variación porcentual en lugar de la variación absoluta y así mejorar la calidad de la proyección en series de tiempo económicas. Adicionalmente se desea estudiar el ratio entre ambas series, ya que en el caso de obtener una proyección adecuada, esta podría ser de utilidad para proyectar la fibra para la que se obtenga un menor ajuste.

La Tabla 5 muestra los resultados de los test de raíz unitaria descritos en la sección 4.1 para las variables en nivel y en primera diferencia.

Tabla 5: Resultados test de raíz unitaria sobre precios

Muestra: 01-2001 a 12-2012					
Variable	Dickey-Fuller Aumentado				Phillips-Perron
	t-Statistic	K	Tendencia	Intercepto	t-Statistic
Nivel					
Fibra corta (log)	-3,68 *	1	Sí	Sí	-2,84
Fibra larga (log)	-3,78 *	1	Sí	Sí	-2,48
Ratio	-3,22	2	No	Sí	-2,98
Primera diferencia					
Fibra corta (log)	-5,63 *	0	No	No	-5,67 *
Fibra larga (log)	-5,70 *	1	No	No	-4,36 *
Ratio	-9,97 *	0	No	No	-10,24 *

Fuente: Estimaciones obtenidas usando EViews 9 SV

* denota rechazo de la hipótesis nula de raíz unitaria con 95% de confianza

Según los resultados de los test realizados, tanto el precio de la fibra corta como el de la fibra larga resultan no estacionarias con el resultado del test de Phillips-Perron, ya que el t-estadístico reportado es menor al valor crítico. A pesar de que según el test ADF se puede rechazar la hipótesis nula de raíz unitaria, dado que en ambos casos la tendencia resulta estadísticamente significativa, se asume que ambas series resultan no estacionarias en nivel. En el caso del ratio, a pesar de que ambos test reportan un t-estadístico superior al valor crítico rechazando la hipótesis nula de raíz unitaria, dado que el intercepto resulta significativo se considera que la serie es no estacionaria.

Para la primera diferencia de las series consideradas, en todos los casos se rechaza la hipótesis nula a través de los dos test considerados, ya que reportan un t-estadístico mayor al valor crítico.

Se puede concluir que las variables en estudio son series integradas de primer orden, y los análisis restantes se realizan bajo dicha restricción.

6.2 Resultados modelo autorregresivo

Es el modelo más simple de los tres considerados en el estudio y es considerado como proyección de referencia para comparar los otros modelos propuestos, ya que es uno de los más usados en la proyección de commodities y ofrece buenos resultados en el corto plazo. A continuación se detalla el análisis realizado (para más detalle de las estimaciones realizadas, ver Anexo A: Estimaciones modelo autorregresivo).

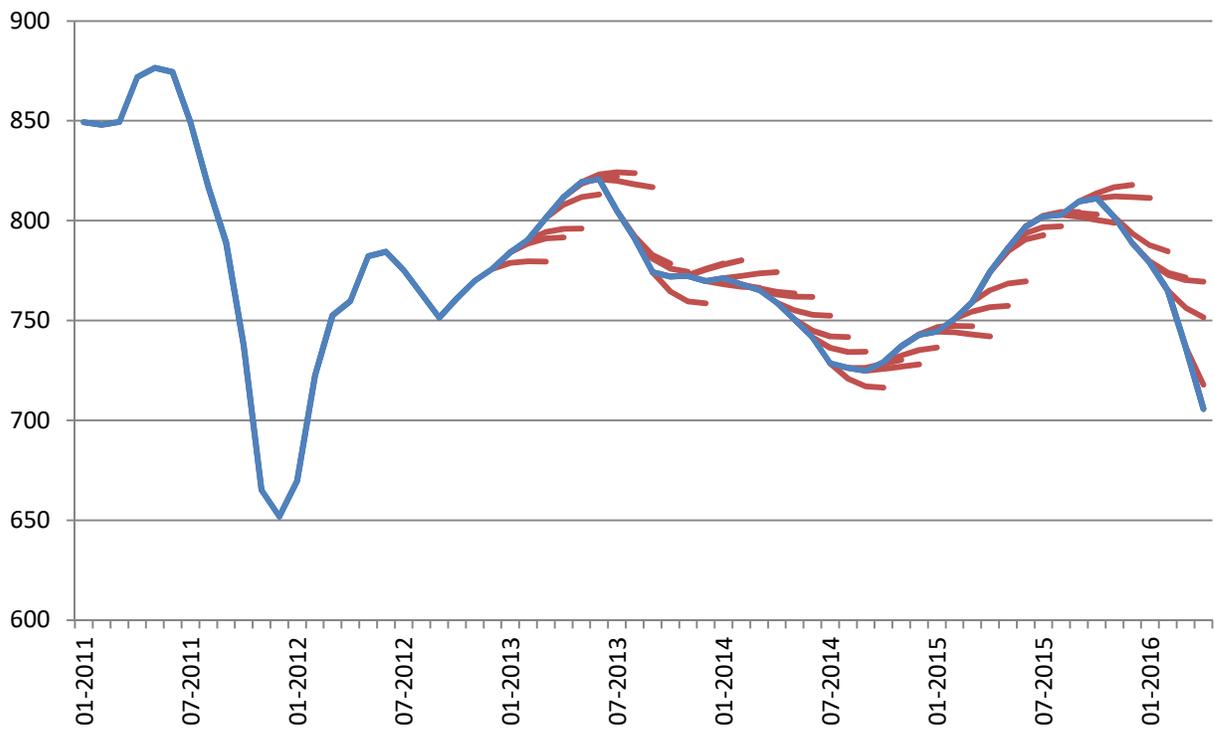
Siguiendo la metodología de Box-Jenkins se procede a diferenciar las series de tiempo en estudio, dado que todas corresponden a series integradas de primer orden según la sección 6.1. Adicionalmente se analiza el correlograma de las variables tanto en nivel como en diferencia, confirmando la necesidad de diferenciación, ya que no se logra obtener niveles no significativos de la función de autocorrelación luego de varios periodos.

Una vez diferenciadas las series es necesario identificar los parámetros p y q , correspondientes a los rezagos de las variables a proyectar y las medias móviles de los errores respectivamente. Con este fin se estiman modelos haciendo variar los parámetros p y q entre cero y cinco. El criterio de selección de modelos se basa en minimizar el SIC reportado, dado que este estadístico permite seleccionar los modelos más parsimoniosos.

De los resultados obtenidos se eligen los cinco candidatos con menor valor de SIC (destacados en las tablas anteriores), se estiman los parámetros y se evalúa el ajuste de predicción en el periodo enero 2012 a marzo 2012 para seleccionar el modelo con mayor poder predictivo. El corto horizonte de proyección considerado se justifica por la limitada capacidad predictiva demostrada por los modelos autorregresivos, los que tienden a generar proyecciones constantes después de cierta cantidad de periodos.

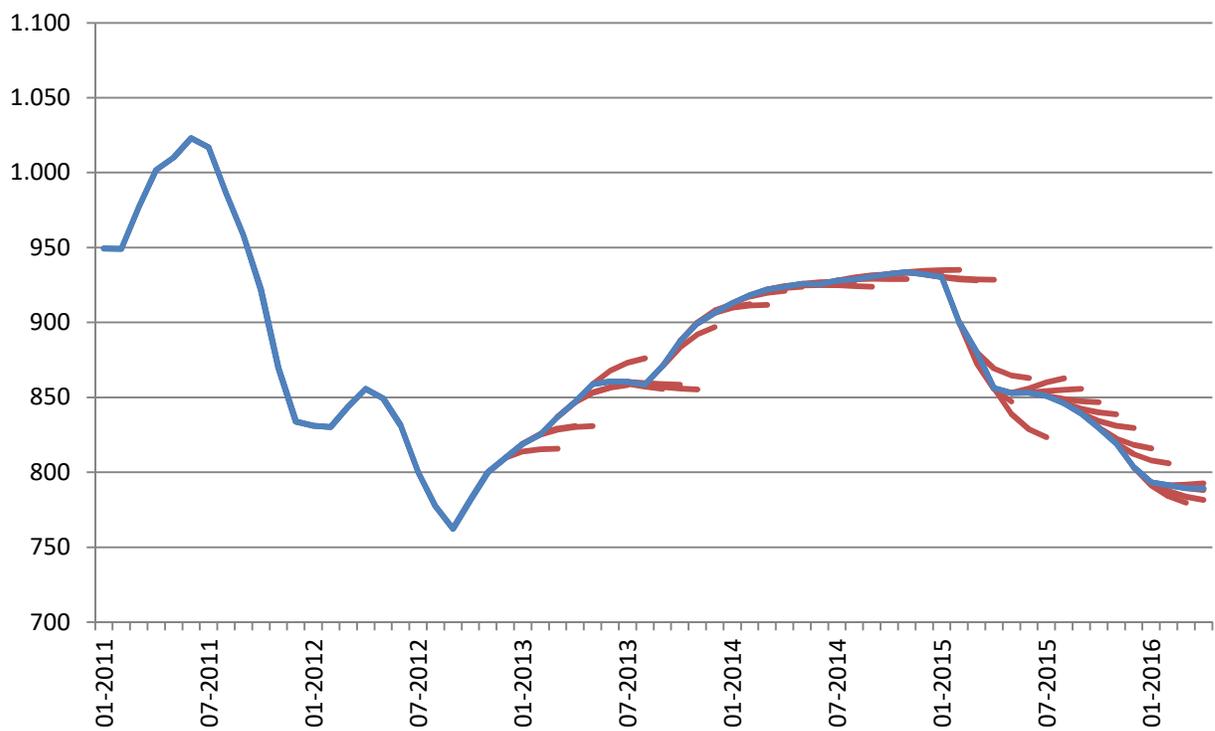
Los Gráficos 19 y 20 muestran las proyecciones realizadas para las variables en estudio. Se grafican las proyecciones dinámicas obtenidas mediante la técnica de origen móvil a un horizonte de tres meses, con un intervalo de un mes entre proyecciones. Estas proyecciones pueden ser de utilidad en el corto plazo, ya que siguen la tendencia que presenta la serie, aunque no son recomendables en el largo plazo al generar proyecciones prácticamente constantes luego de algunos periodos, dado el bajo orden de los modelos estimados.

Gráfico 19: Proyecciones modelo autorregresivo fibra corta



Fuente: Elaboración propia

Gráfico 20: Proyecciones modelo autorregresivo fibra larga



Fuente: Elaboración propia

La Tabla 6 resume los resultados obtenidos para las distintas variables dependientes, para mayor detalle ver Anexo A: Estimaciones modelo autorregresivo.

Tabla 6: Resultados modelo autorregresivo

	Fibra corta	Fibra larga	Ratio
AR(1)	0,67	0,94	0,67
AR(2)	0,08	-0,26	
AR(3)	-0,17		
MA(1)			-0,47

Fuente: Elaboración propia

6.3 Resultados modelo de indicadores macroeconómicos

Se estudia la relación entre la fluctuación de las variables en estudio y los factores comunes para commodities encontrados por (Kagraoka, 2015). Dichos factores explican más del 50% de la varianza de la fibra larga y son:

- tasa de inflación estadounidense
- precio del petróleo
- índice accionario global
- índice de producción industrial global

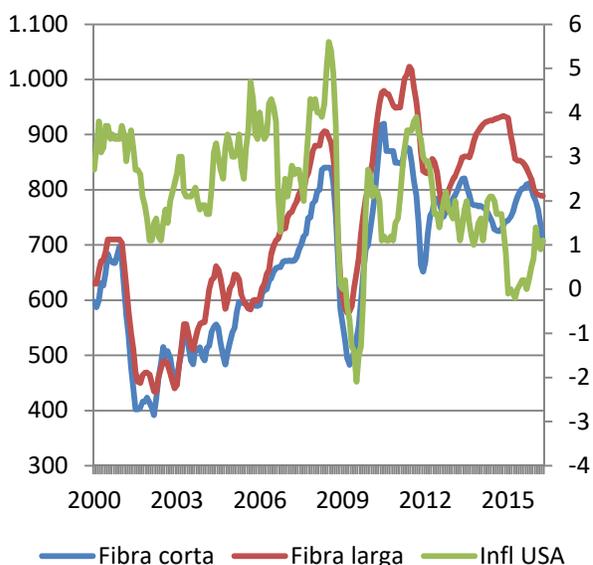
En este estudio se utilizan los datos descritos en la sección 5.2 y el análisis realizado se describe a continuación (para más detalle de las estimaciones realizadas, ver Anexo B: Estimaciones modelo de indicadores macroeconómicos).

En primer lugar, de manera análoga a la sección 6.1, se verifica el orden de integración de las variables para considerar su aplicación en nivel o en diferencia. Se obtiene que todas las variables consideradas corresponden a series integradas de primer orden, por lo que serán diferenciadas para evaluar su impacto sobre el precio.

Adicionalmente se testea la posibilidad de que las series consideradas presenten comportamiento estacional, concluyendo que ninguna presenta estacionalidad; resultado evidente al considerar las posibilidades de arbitraje que esto generaría.

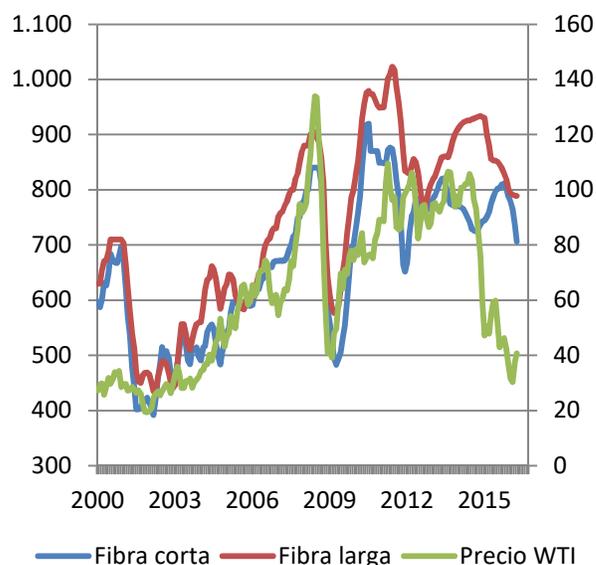
Los Gráficos 21 a 24 muestran la evolución de las variables con respecto a los precios de la celulosa. En todos los casos se tiene una alta correlación entre las variables, por lo que los indicadores en estudio podrían tener influencia sobre las variaciones de precio de la celulosa.

Gráfico 21: Precio fibras vs Inflación USA



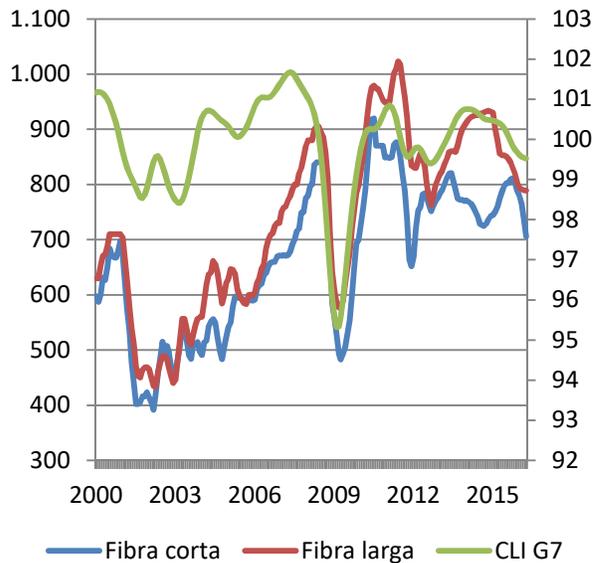
Fuente: Elaboración propia

Gráfico 22: Precio fibras vs precio WTI



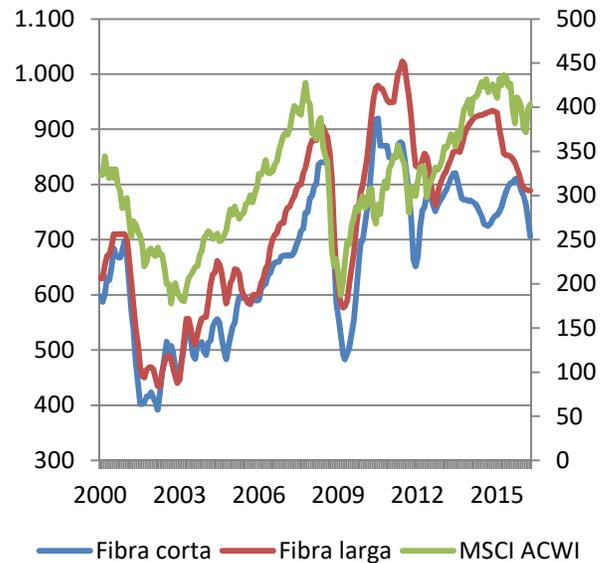
Fuente: Elaboración propia

Gráfico 23: Precio fibras vs CLI G7



Fuente: Elaboración propia

Gráfico 24: Precio fibras vs ACWI



Fuente: Elaboración propia

Para verificar el periodo de influencia de los indicadores sobre las variables en estudio, se grafica la correlación dinámica al incorporar hasta 12 rezagos sobre los indicadores macroeconómicos. Para ambos tipos de fibra se aprecia que el precio del petróleo, el CLI G7 y el ACWI empiezan a perder relevancia desde el periodo cinco, mientras que para la inflación ocurre lo contrario, comenzando a adquirir importancia desde ese punto, aunque es la variable que presenta una correlación menos importante. En el caso del ratio se obtienen las correlaciones más bajas y con patrones menos estables, lo que hace intuir que con los indicadores macroeconómicos utilizados podría no obtenerse un buen modelo para su proyección. Se podría esperar cierta coherencia entre los periodos con mayor correlación y los regresores significativos reportados por el modelo.

A priori se podría esperar que los indicadores utilizados tengan cierto efecto sobre los precios de las fibras, estos se pueden resumir de la siguiente manera:

- al aumentar el precio del petróleo, este produciría un aumento sobre el precio de la celulosa al aumentar los costos
- un aumento en la inflación estadounidense debería tener un efecto negativo al generar una reducción en la demanda
- el efecto del índice de producción industrial global debería ser positivo al aumentar la demanda
- un aumento en el índice accionario global debería producir un aumento en el precio por medio del aumento de la demanda

Dichos efectos se esperan sobre los precios de las fibras, no así sobre el ratio, donde resulta menos evidente el impacto de cada uno de los regresores mencionados.

Se propone el siguiente modelo:

$$\Delta x_t = \sum_{i=1}^{12} \beta_{1,i} \Delta Infl USA_{t-i} + \sum_{i=1}^{12} \beta_{2,i} \Delta WTI_{t-i} + \sum_{i=1}^{12} \beta_{3,i} \Delta ACWI_{t-i} + \sum_{i=1}^{12} \beta_{4,i} \Delta CLI G7_{t-i} + \sum_{i=1}^{12} \beta_{5,i} \Delta p_{t-i}^{corta} + \sum_{i=1}^{12} \beta_{6,i} \Delta p_{t-i}^{larga}$$

Donde Δx_t corresponde a la primera diferencia de la variable dependiente a estudiar (descritas en la sección 6.1) en el periodo t , $\Delta Infl USA_t$ es la primera diferencia de la inflación estadounidense en el periodo t , ΔWTI_t es la primera diferencia del precio del petróleo en el periodo t , $\Delta ACWI_t$ es la primera diferencia del índice accionario global en el periodo t , $\Delta CLI G7_t$ es la primera diferencia del índice de producción industrial en el periodo t , Δp_t^{corta} es la primera diferencia del logaritmo del precio de la fibra corta y Δp_t^{larga} es la primera diferencia del logaritmo del precio de la fibra larga. Los rezagos de ambas fibras se introducen para considerar el efecto sustitución entre fibras, y en el caso del ratio se incorporan rezagos de la misma variable.

Mediante stepwise regression se estima el modelo mencionado considerando sólo las variables significativas con 90% de confianza estadística. La Tabla 7 muestra el efecto neto de los regresores, definido como la suma de los coeficientes asociados a los rezagos que resulten significativos para cada regresor sobre las variables independientes.

Tabla 7: Efectos netos de indicadores macroeconómicos

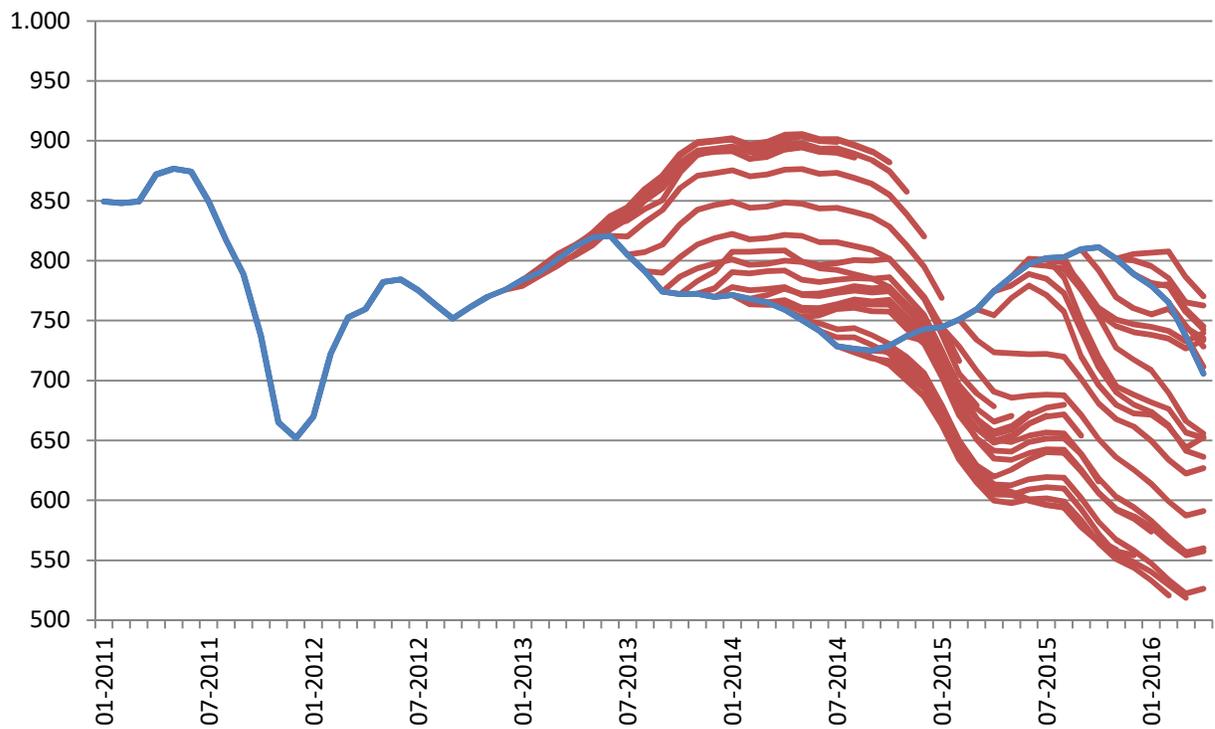
Muestra: 01-2000 a 12-2012				
Regresor	Variable dependiente			Ratio
	Fibra corta	Fibra larga		
Inflación Estados Unidos	-0,01	-0,01		0,01
Precio petróleo WTI	0,16	0,05		-0,17
CLI G7	4,34	5,23		-
ACWI	-	-0,10		-
Fibra corta	0,35	0,15		0,14
Fibra larga	0,11	0,39		0,12
Ratio	-	-		-0,02

Fuente: Estimaciones obtenidas usando EViews 9 SV

En la mayor parte de los casos se obtienen los signos esperados, y adicionalmente se destaca que no se obtienen signos negativos para los rezagos de la fibra usada como variable independiente, lo que demostraría que el efecto sustitución entre fibras resulta muy bajo o nulo.

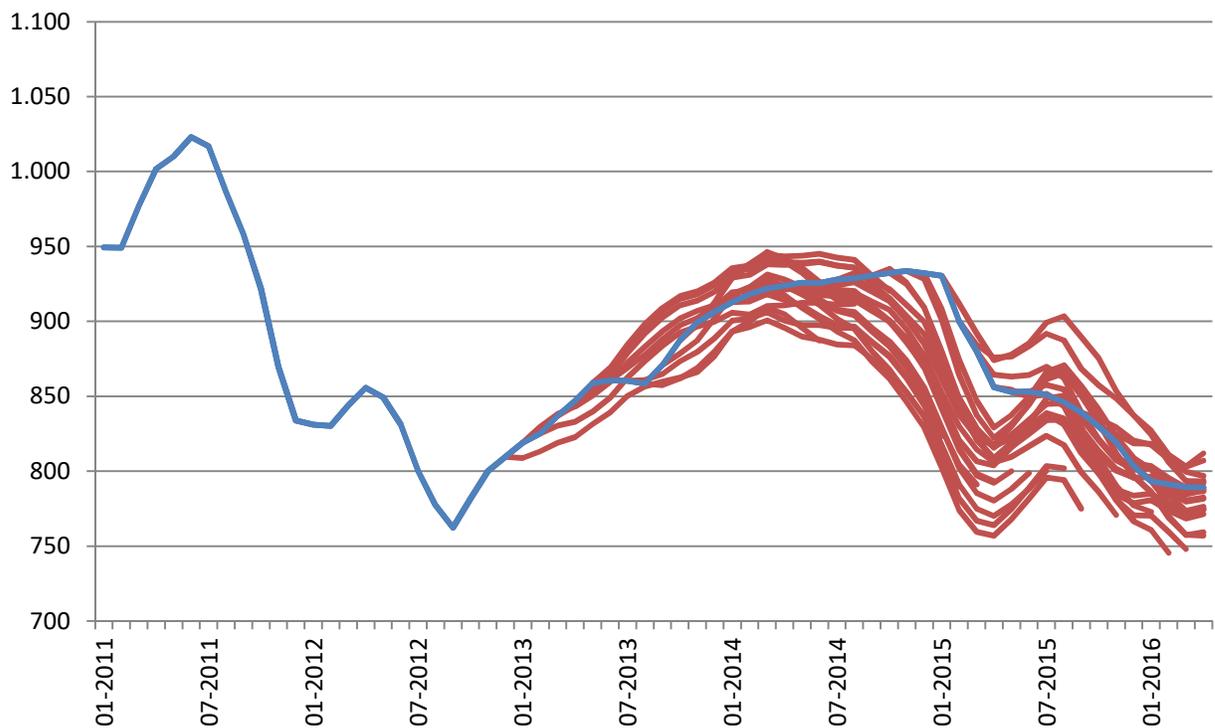
Los Gráficos 25 y 26 muestran las proyecciones obtenidas mediante la técnica de origen móvil a un horizonte de 18 meses, con un intervalo de un mes entre proyecciones. Se realizan seis proyecciones dinámicas a un horizonte de dos años, con intervalo de tres meses entre proyecciones. En el caso de la fibra larga se obtiene una proyección capaz de anticipar ciertas fluctuaciones sobre el precio, aunque en el caso de la fibra corta resulta ser el modelo con peores resultados.

Gráfico 25: Proyecciones modelo de indicadores macroeconómicos fibra corta



Fuente: Elaboración propia

Gráfico 26: Proyecciones modelo de indicadores macroeconómicos fibra larga



Fuente: Elaboración propia

La Tabla 8 muestra los principales regresores obtenidos para las distintas variables dependientes, para mayor detalle ver Anexo B: Estimaciones modelo de indicadores macroeconómicos.

Tabla 8: Resultados modelo de indicadores macroeconómicos

Fibra corta		Fibra larga		Ratio	
Variable	Coefficiente	Variable	Coefficiente	Variable	Coefficiente
$\Delta \log(p_{t-2}^{larga})$	-0,36	$\Delta \log(p_{t-1}^{larga})$	0,50	$\Delta \log(p_{t-10}^{corta})$	0,14
$\Delta \log(p_{t-1}^{corta})$	0,35	$\Delta \log(p_{t-3}^{corta})$	-0,21	$\Delta Infl USA_{t-2}$	0,01
$\Delta \log(CLI G7_{t-2})$	4,34	$\Delta \log(p_{t-1}^{corta})$	0,22	$\Delta \log(WTI_{t-3})$	-0,06

Fuente: Elaboración propia

6.4 Resultados modelo de corrección de errores

Con este modelo se pretende verificar si existe una relación entre el precio de la celulosa y los días de inventario que logre explicar las fluctuaciones de precio en el largo plazo. A continuación se presenta un resumen de los resultados obtenidos (para mayor detalle ver Anexo C: Estimaciones modelo corrección de errores).

En el estudio se consideran los datos de días de inventario de oferta y demanda descritos en la sección 5.3. En primer lugar se debe determinar el grado de integración de las variables a utilizar a fin de verificar si el modelo es aplicable. Para esto se aplican los test de raíz unitaria descritos en la sección 4.1.

Todas las variables en nivel resultan no estacionarias, mientras que para la primera diferencia se obtiene rechazo de la hipótesis nula de raíz unitaria, por ende se asume que todas resultan integradas de primer orden y se evaluará la existencia de cointegración entre estas y el precio de la fibra.

Se procede a verificar la existencia de cointegración mediante el procedimiento de Engle-Granger descrito en la sección 4.4.2. Para esto se considera la siguiente relación de largo plazo:

$$\log(p_t^i) = \alpha + \beta \log(inv_t^i) + \varepsilon_t$$

Donde p_t^i corresponde al precio de la fibra i en el periodo t , e inv_t^i corresponde al logaritmo del inventario considerado para la fibra i en el periodo t . Se espera una relación negativa entre el precio de la fibra y los días de inventario, ya que un mayor nivel de precios produce estrechez en el mercado reduciendo los inventarios, mientras que un menor nivel de precios produce holguras en el mercado, aumentando los inventarios.

Al analizar los resultados de los test de raíz unitaria sobre los residuos de la relación propuesta, se obtiene existencia de cointegración entre precios y días de inventario de demanda considerados, dado que estos presentan una tendencia similar a la del nivel de precios, como se puede ver en los Gráficos 12 y 13. La relación obtenida efectivamente es negativa, pero se destaca que los inventarios considerados no reflejan las perturbaciones producidas por crisis o efectos externos.

Se considera el siguiente modelo utilizado por (Riis, 1999):

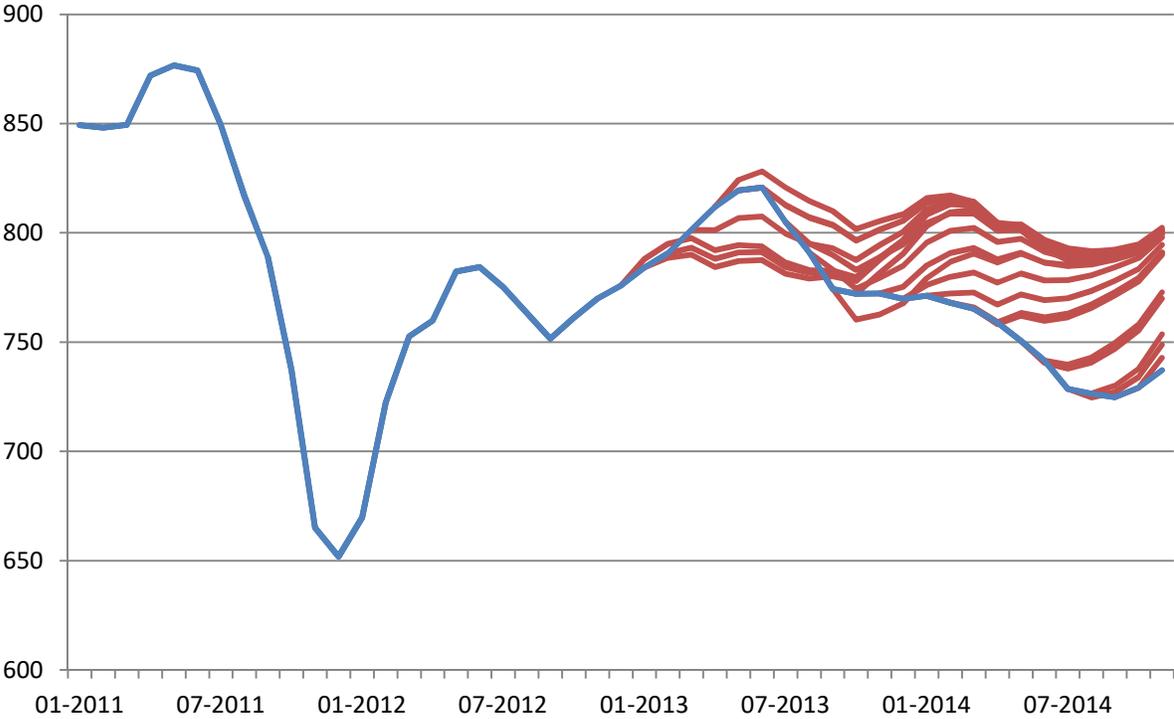
$$\Delta \log(p_t^i) = \delta \Delta \log(p_{t-1}^i) + \varphi \Delta \log(inv_{t-1}^i) - \alpha (\log(p_{t-1}^i) - \beta \log(inv_{t-1}^i)) + c + \varepsilon_t$$

Donde p_t^i corresponde al logaritmo del precio de la fibra i en el periodo t e inv_t^i corresponde a los días de inventario considerados para la fibra i en el periodo t .

Resulta evidente que para proyectar el precio en el largo plazo es necesario también realizar una proyección sobre los días de inventario. Para esto se realiza una proyección de los inventarios a utilizar mediante suavizamiento exponencial, manteniendo la tendencia y estacionalidad que presentan estos en los periodos anteriores.

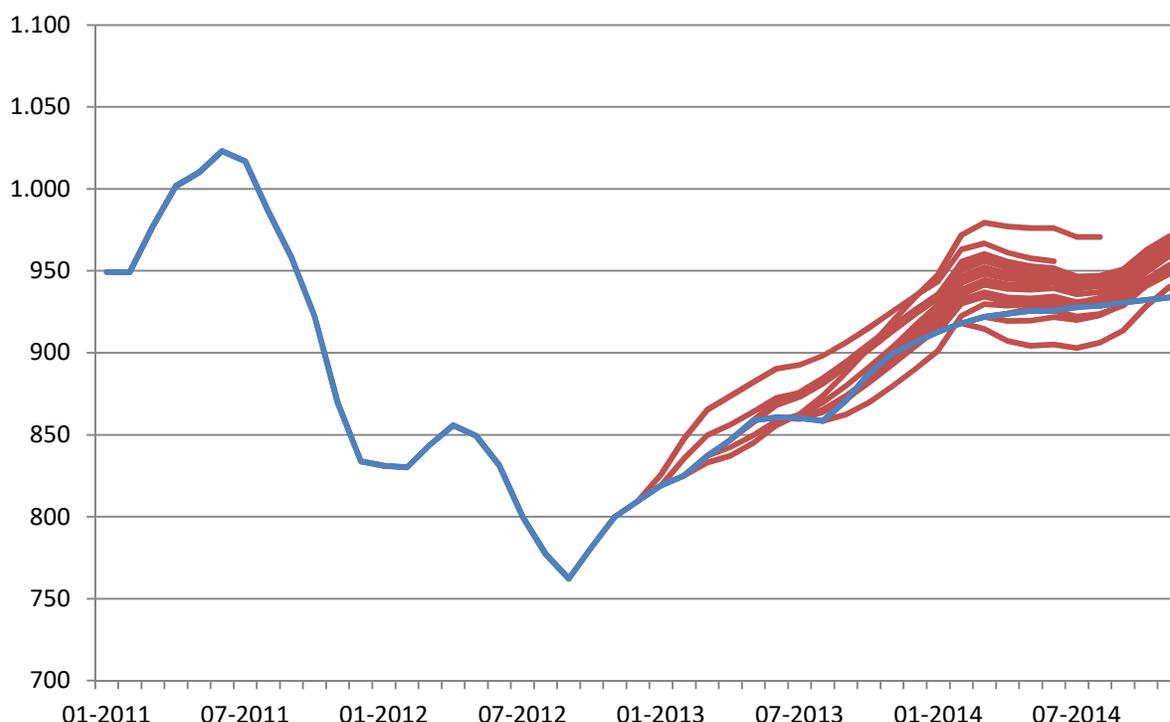
Los Gráficos 27 y 28 muestran las proyecciones dinámicas obtenidas mediante la técnica de origen móvil a un horizonte de 18 meses, con un intervalo de un mes entre proyecciones para el periodo enero 2013 a octubre 2014 usando la proyección auxiliar de datos de inventarios.

Gráfico 27: Proyecciones modelo corrección de errores fibra corta



Fuente: Elaboración propia

Gráfico 28: Proyecciones modelo corrección de errores fibra larga



Fuente: Elaboración propia

En ambos casos el modelo ofrece una proyección estable basada en la relación estacionaria entre precios e inventarios; si bien el precio puede verse afectado por efectos externos en el corto plazo, este debería retornar al equilibrio mencionado.

A pesar de que las elasticidades entre inventarios y precios obtenidas con los modelos estimados no resultan comparables con las encontradas en la literatura (entre -0,14 y -0,21 para fibra larga), ya que estas últimas se obtienen usando nivel de inventarios de productores, mientras que las del presente estudio se obtienen usando días de inventario de consumidores; se destaca que ambas presentan el mismo orden de magnitud, y el cambio de signo se puede justificar por un efecto látigo que gatilla una respuesta rezagada por parte de consumidores.

La Tabla 9 resume los resultados obtenidos para las distintas variables dependientes, para mayor detalle ver Anexo C: Estimaciones modelo corrección de errores.

Tabla 9: Resultados modelo corrección de errores

Parámetro	Fibra corta	Fibra larga
δ	0,68	0,74
φ	0,21	0,17
α	0,11	0,05
β	0,91	1,11
c	1,07	0,55

Fuente: Elaboración propia

6.5 Error de proyecciones

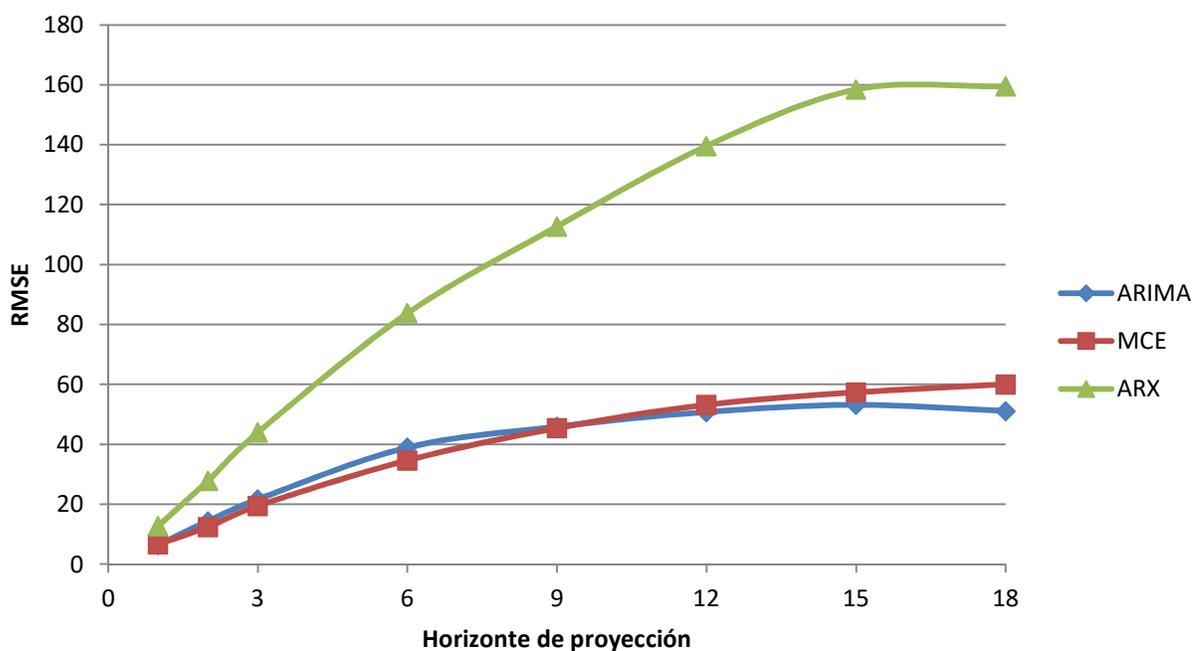
Dado que en ninguno de los modelos considerados se obtienen resultados fiables para la proyección del ratio entre las fibras, el análisis usando origen móvil se realiza sólo para los precios de fibra corta y fibra larga.

Para comparar los diferentes modelos estimados se utiliza la raíz del error cuadrático medio o RMSE, ya que todos son estimados usando muestras del mismo tamaño, de esta forma la métrica de error obtenida resulta comparable y es adecuada para evaluar su desempeño.

Con el fin de evaluar el error de las proyecciones obtenidas con cada modelo se estima el RMSE sobre todos los pronósticos obtenidos a un mes, consiguiendo de esta forma un estimador general del desempeño de los modelos propuestos a través del tiempo. Este procedimiento se repite para distintos alcances hasta 18 meses, en todos los modelos propuestos.

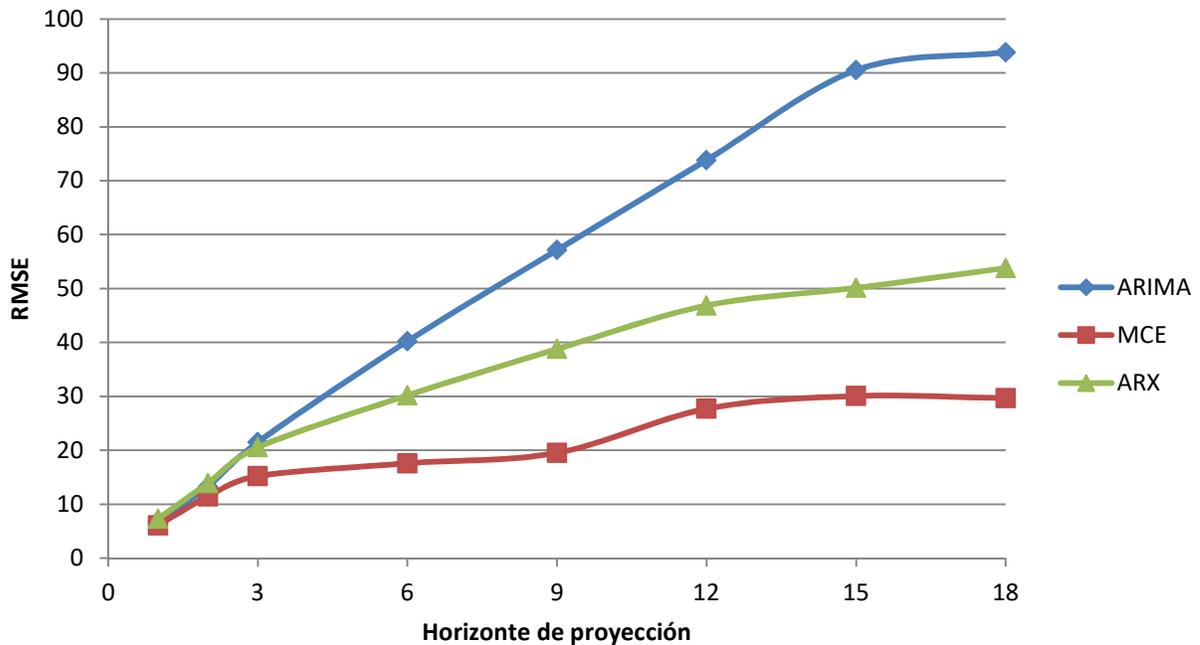
Los Gráficos 29 y 30 muestran el error obtenido con los modelos desarrollados a distintos alcances para fibra corta y fibra larga respectivamente.

Gráfico 29: RSME a distintos alcances para fibra corta



Fuente: Elaboración propia

Gráfico 30: RMSE a distintos alcances para fibra larga



Fuente: Elaboración propia

En todos los casos el RMSE obtenido resulta creciente al aumentar el alcance, lo que es evidente al tratarse de proyecciones dinámicas, donde al utilizar los pronósticos obtenidos como dato para obtener los pronósticos de mayor alcance, el grado de incertidumbre aumenta. Los modelos para fibra corta resultan menos fiables que los de fibra larga, ya que en un horizonte de proyección de 18 meses, convergen a un RSME notoriamente mayor.

Para la fibra corta, el modelo que produce el peor desempeño en todos los alcances considerados es el modelo de indicadores macroeconómicos, lo que podría deberse a que las variables explicativas consideradas no son las adecuadas, dado que el mercado de la fibra corta es más dinámico debido a ciclos de crecimiento cortos, y además es producida en economías emergentes, lo que puede influir en la baja relación con las variables consideradas.

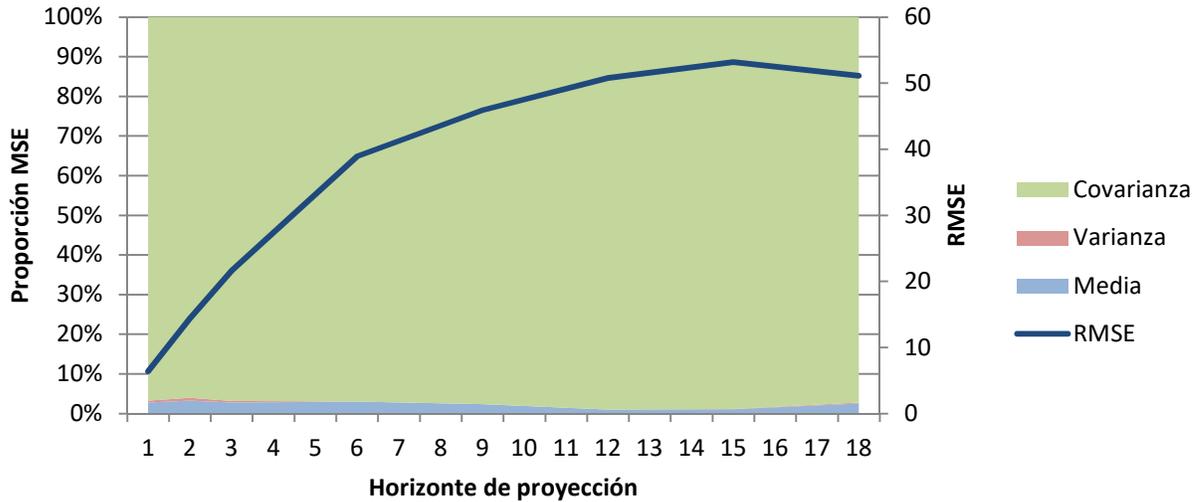
El modelo de corrección de errores resulta ser el que brinda mejores resultados para ambas fibras, aunque es importante destacar que mediante este modelo no se logran capturar efectos externos (por ejemplo periodos de crisis o tendencias de actividad económica) y su desempeño depende de una correcta estimación del inventario utilizado. Por esta razón es necesario complementar la información entregada (al menos para fibra larga) con la obtenida mediante el modelo de indicadores macroeconómicos, el cual sí considera efectos exógenos.

El RMSE puede interpretarse como la desviación estándar generada por las proyecciones con respecto al precio real, es decir, para el modelo de corrección de errores se obtiene una desviación de alrededor de US\$20 ($\pm 3\%$ aprox.) a un alcance de nueve meses. La proyección obtenida puede ser utilizada con distintos propósitos teniendo en cuenta las métricas de error obtenidas sobre los distintos alcances.

6.6 Análisis de especificación

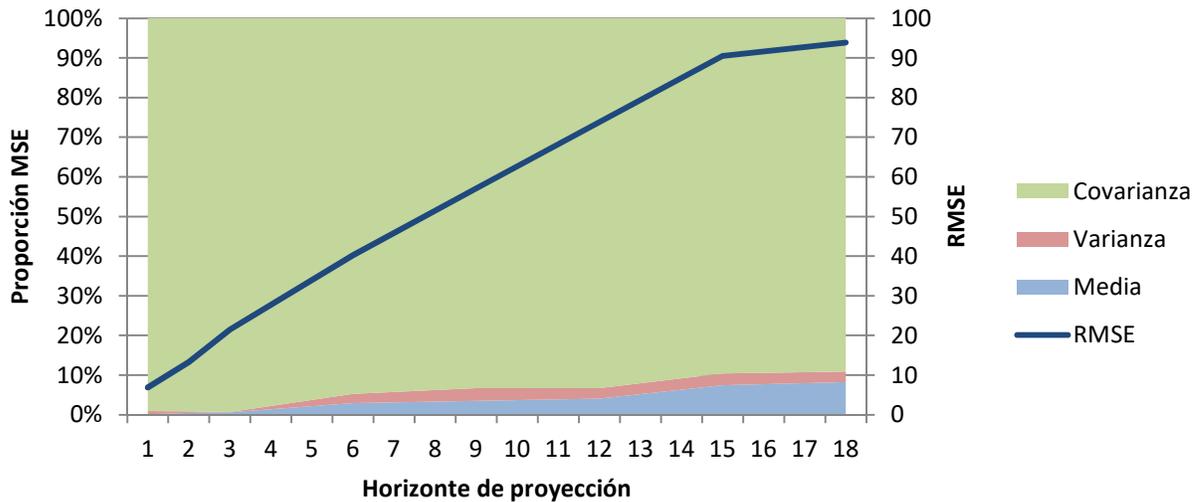
Los Gráficos 31 a 36 muestran la descomposición del MSE en las proporciones de media, varianza y covarianza (descritos en la sección 4.6) de los modelos estimados junto con el RMSE (desviación con respecto al precio real) obtenido.

Gráfico 31: Descomposición MSE modelo autorregresivo para fibra corta



Fuente: Elaboración propia

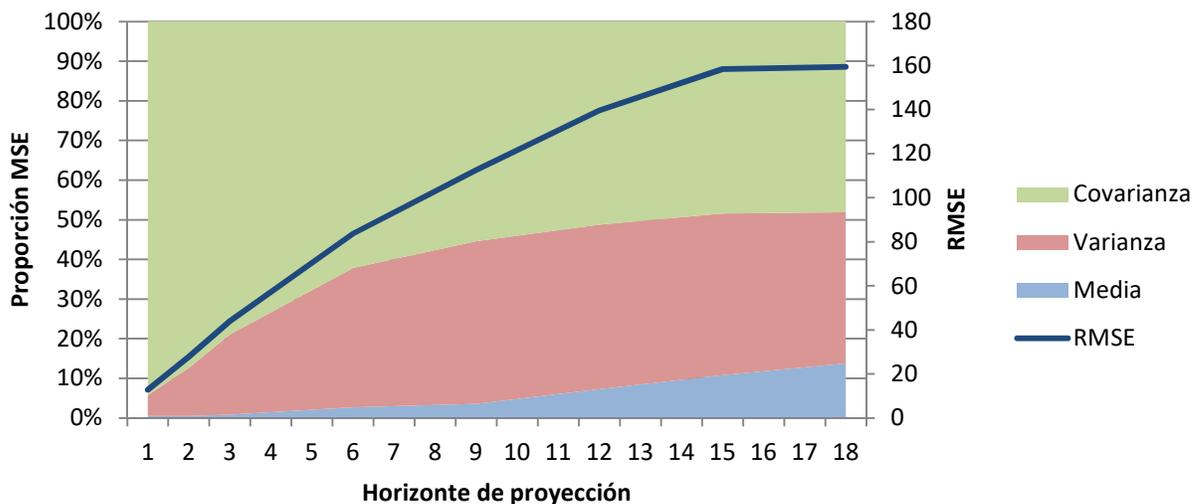
Gráfico 32: Descomposición MSE modelo autorregresivo para fibra larga



Fuente: Elaboración propia

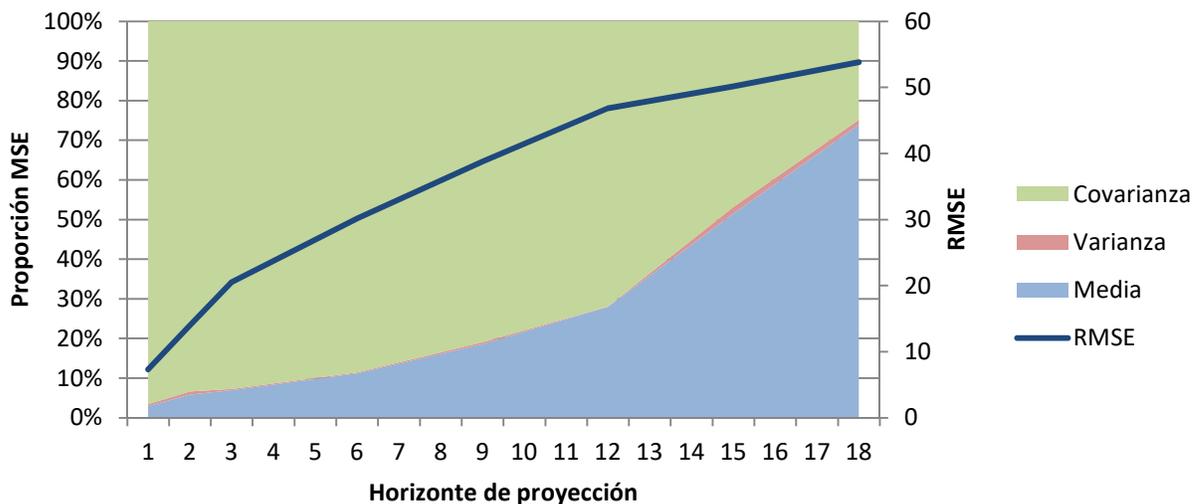
En el caso del modelo autorregresivo el término de covarianza concentra más de un 90% del error hasta los 15 meses en fibra larga y 18 meses en fibra corta, resultado difícil de justificar, dado que en el largo plazo dicho modelo no debería ser considerado al limitarse a presentar una proyección prácticamente constante.

Gráfico 33: Descomposición MSE modelo de indicadores macroeconómicos para fibra corta



Fuente: Elaboración propia

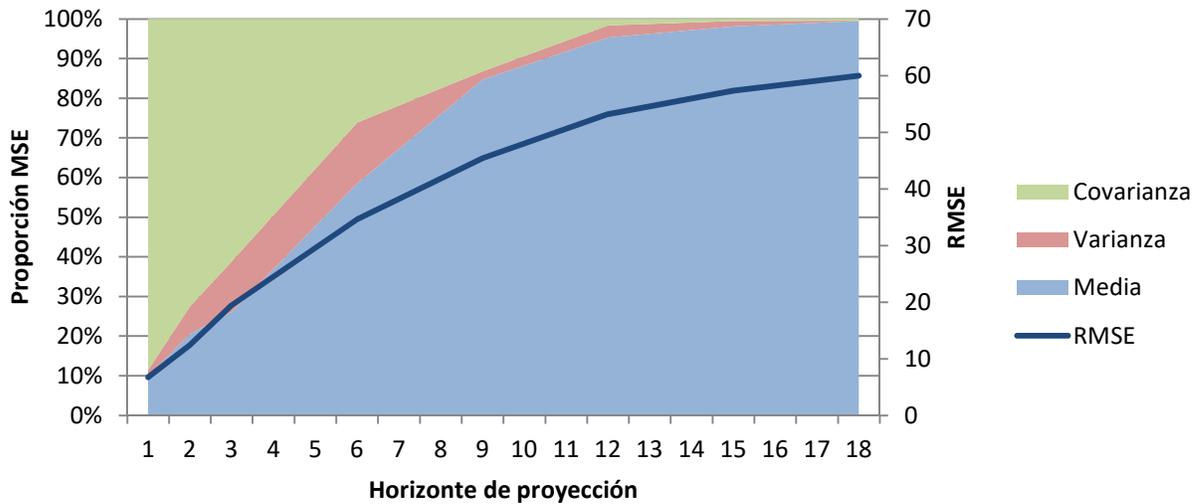
Gráfico 34: Descomposición MSE modelo de indicadores macroeconómicos para fibra larga



Fuente: Elaboración propia

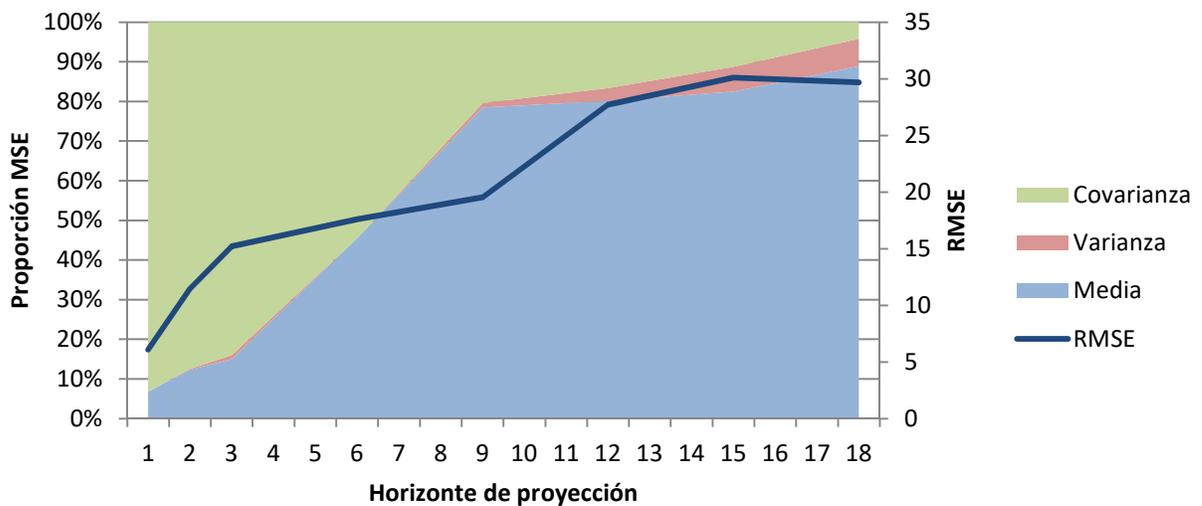
Con respecto al modelo de indicadores macroeconómicos, en el caso de la fibra larga la covarianza concentra más del 90% del error para la fibra larga hasta los tres meses de alcance, mientras que para fibra corta sólo para el primer mes. El importante aumento de la proporción asociada a la varianza en el caso de la fibra corta se asocia con el hecho de que es el que entrega los peores resultados en todos los horizontes de proyección, lo que puede estar relacionado con las variables explicativas seleccionadas.

Gráfico 35: Descomposición MSE modelo de corrección de errores para fibra corta



Fuente: Elaboración propia

Gráfico 36: Descomposición MSE modelo de corrección de errores para fibra larga



Fuente: Elaboración propia

Para el modelo de corrección de errores la proporción asociada al término de covarianza disminuye rápidamente concentrando menos de un 90% desde los dos meses para la fibra larga y desde el primer mes para la fibra corta, resultados poco coherentes con el modelo, el cual pretende identificar relaciones de largo plazo y no debería parecer menos estable que los anteriores.

En general se obtiene que el RMSE crece a medida que va aumentando el horizonte de proyección, resultado lógico, dado que junto con el horizonte de proyección también se incrementa la incertidumbre, aunque en algunos gráficos tiende a decaer al considerar 18 meses de alcance. Esto puede deberse a que el tamaño de la

muestra se reduce al aumentar el horizonte de proyección, ofreciendo resultados menos precisos.

Con respecto a la descomposición del error cuadrático medio, se obtienen resultados que no se condicen con la especificación del modelo, ya que los modelos que ofrecen mejores resultados presentan una peor especificación. (Granger, y otros, 1973) demuestran que una descomposición alternativa propuesta por Theil en los mismos términos entrega resultados cuestionables para la estimación de precios de acciones. Estos resultados podrían ser estudiados posteriormente considerando variaciones a la descomposición propuesta por Theil, dado que esta podría no ofrecer buenos resultados en modelos no lineales o en la comparación entre varios modelos.

Se puede concluir que la descomposición utilizada no sería adecuada para realizar comparaciones entre modelos, o bien sería necesario utilizar una variación de esta, por ejemplo considerando el nivel del error obtenido en cada dimensión para cada modelo, ya que al utilizar sólo la proporción obtenida se puede sobreestimar o subestimar el desempeño en cada término. De todas formas, al evaluar cada modelo por separado, los resultados obtenidos no armonizan con la especificación y resultados esperados para cada uno, por lo que los resultados no serían fiables.

7 Conclusiones y recomendaciones

Los modelos propuestos mejoran los resultados obtenidos con la proyección de referencia, aunque su desempeño depende de la fibra a proyectar. En general el modelo de corrección de errores es que tiene un mejor desempeño en ambos tipos de fibra y el modelo de indicadores macroeconómicos también ofrece un buen desempeño en el caso de la fibra larga, incorporando información externa al mercado.

El ajuste entregado por el **modelo autorregresivo**, usado como proyección de referencia, se ve superado al incorporar variables explicativas, además al aumentar el alcance de la proyección, sus resultados se ven rápidamente superados por los otros modelos, siendo el que entrega los peores resultados desde los tres meses de alcance para la fibra larga. Queda en evidencia el limitado alcance de este tipo de modelos y (al menos en la proyección de este producto) su uso no sería recomendable para obtener pronósticos a más de un mes, ya que, al no presentar un comportamiento estacional, el orden de los parámetros significativos resulta muy bajo, generando proyecciones prácticamente constantes en el mediano plazo.

El **modelo de indicadores macroeconómicos** entrega resultados con alto poder explicativo (R^2 ajustado de 0,75) y con métricas de error relativamente bajas (en particular inferiores al modelo autorregresivo desde los dos meses) en el caso de la fibra larga, situación que no se replica al proyectar la fibra corta. Esta situación puede deberse a la selección de las variables explicativas: existen varias diferencias entre ambos tipos de fibra, como por ejemplo sus usos, estructura de costos y región/países en que son producidas. Con respecto a este último punto, los principales productores de fibra larga corresponden a Norte América y Europa, mientras que la producción de fibra corta se centra principalmente en Sudamérica y Asia, regiones en las que existe menor cantidad de estudios sobre las fuerzas económicas que perturban sus mercados. Es posible que las variables seleccionadas no guarden una fuerte relación con las perturbaciones económicas en los últimos mercados, dado que tienen una mayor influencia en las regiones productoras de fibra larga; además, la fibra corta presenta ciclos productivos más cortos, lo que influye en dichas relaciones.

En cuanto a los indicadores macroeconómicos seleccionados para la realización del estudio, los resultados podrían mejorarse al utilizar variables que tengan un mayor impacto económico en la producción de celulosa. En particular, en la producción de productos forestales se produce biomasa, la que es utilizada para generar electricidad que alimenta los distintos procesos, y es posible que el mercado sea menos sensible que otros commodities frente a variaciones en el precio del petróleo. Por otro lado, dadas las diferencias existentes entre fibra corta y fibra larga, se podría evaluar una mejora del modelo de indicadores macroeconómicos utilizando otro Composite Leading Indicator, ya que el G7 podría tener una mayor influencia sobre la fibra larga, dados los países involucrados. También se podría evaluar si los resultados mejoran al utilizar un índice accionario de mercados emergentes para la fibra corta y uno de mercados desarrollados para fibra larga, dadas las regiones en que son producidas.

En ambas fibras el **modelo de corrección de errores** es el que reporta mejores resultados en la proyección entre los dos y nueve meses, entregando una desviación menor a US\$35 en un alcance de nueve meses. Es importante destacar que una aplicación confiable de dicho modelo en el largo plazo dependerá de la capacidad de proyección de los inventarios utilizados y de un adecuado estudio y anticipación de su tendencia, de forma que esta no genere sesgos sobre la proyección del precio de la celulosa. Por otro lado, dado que los inventarios no reflejan shocks externos, estos presentan tendencias más estables que los precios, lo que puede aportar una mejor referencia sobre la tendencia en el largo plazo.

(Bashiri Mehmiri, y otros, 2013) plantean que al momento de desarrollar proyecciones se debería tener en cuenta el uso de más de un modelo, ya que distintos modelos podrían variar su desempeño en diferentes horizontes y además el origen de la proyección (por ejemplo un periodo de crisis versus un periodo de estabilidad económica) podría tener efectos significativos sobre los diferentes modelos. Estos resultados son totalmente compatibles con los obtenidos en el presente estudio, donde se observa que los inventarios utilizados no reflejan los efectos de las distintas crisis económicas; mientras que los indicadores macroeconómicos (en específico el ACWI y CLI G7), al ser específicamente diseñados para prever tendencias de la economía, sí son capaces de incorporar dichas fluctuaciones, por ende es necesario tener ambos en consideración para tener un panorama más amplio sobre los orígenes de las fluctuaciones.

Con respecto a la descomposición del error cuadrático medio, esta genera resultados que parecieran no ser coherentes con los modelos estimados, en particular se obtiene que el porcentaje de error asociado a covarianza (término que mide los errores no sistemáticos) en el modelo de corrección de errores concentra menos de un 50% desde el mes seis en adelante para fibra corta y desde el mes nueve en adelante para fibra larga, indicando que el modelo estaría mal especificado para dicha proyección. Según la misma descomposición para el modelo autorregresivo estimado, el término de covarianza concentra al menos un 89% del error cuadrático medio para todos los alcances considerados en ambas fibras. Estos resultados son bastante llamativos, ya que un modelo que ofrece una proyección constante en el largo plazo no debería producir resultados más estables que uno que considera variables explicativas propias del mercado. En estudios posteriores se podría verificar si los resultados mejoran al realizar la comparación sobre la magnitud de los términos, ya que al comparar sólo la proporción se pasa por alto el desempeño general y tampoco se considera la no linealidad de la estimación de parámetros comprendida en el modelo de corrección de errores.

Se logra desarrollar una proyección estable en el tiempo basada en información pública de alta calidad y proveniente de fuentes confiables. Al incorporar indicadores macroeconómicos que no son propios del mercado se ofrece la posibilidad de agregar efectos externos con un gran poder explicativo. Este análisis puede resultar de interés para productores de celulosa y entidades financieras, ya que se obtienen resultados acertados sin la necesidad de obtener acceso a información privada del mercado forestal y se incorpora un análisis que no es realizado actualmente en estas instituciones. Por otro lado se destaca que la información utilizada es más estable y genera impactos medibles que pueden ser determinados en distintos horizontes, a

diferencia de la información relacionada con mercados forestales, donde se ven involucradas dinámicas más complejas y sus efectos resultan menos evidentes sobre el precio global. La proyección obtenida se puede relacionar con indicadores bursátiles o financieros de productores de celulosa, generando estimaciones de sus resultados esperados. Por otro lado, al utilizar indicadores globales, el análisis podría replicarse para la proyección de distintos commodities, ampliando las posibilidades de aplicación.

Se concluye que para realizar proyecciones confiables es necesario utilizar toda la información disponible, integrándola en distintos modelos que capturen perturbaciones originadas por distintas fuerzas económicas. De esta forma se obtiene una perspectiva más amplia de los posibles escenarios a futuro, evitando posibles sesgos de cada modelo independiente.

Uno de los determinantes importantes en el precio de la celulosa es el comportamiento estratégico de consumidores chinos, quienes representan un 31% de las importaciones globales. En este estudio no se logra tener acceso a los inventarios de dichos consumidores debido a que estos datos no son públicos, pero en estudios posteriores sería relevante investigar el impacto de dicho comportamiento en el precio de la celulosa, dada su alta participación de mercado, creciente demanda y gran influencia sobre los productores chilenos, dado que comprenden una gran proporción de las exportaciones chilenas.

El presente estudio se enfoca sólo en modelos econométricos, dada su capacidad de proyectar y explicar las fluctuaciones de precios; de esta forma se logra obtener los principales factores que influyen en el precio de la celulosa. A pesar de que los modelos basados en machine learning son menos transparentes en su aplicación, habiendo identificado dichos factores, estos modelos se podrían utilizar para complementar las proyecciones realizadas.

La información existente para modelar mercados forestales es amplia y proviene de distintas entidades de gran reputación. El principal problema en el modelamiento del sector forestal es que existe poca información pública, y la mayor parte de esta es agregada y/o su publicación tiene una baja frecuencia (por lo general anual), lo que dificulta la realización de estudios. En el caso de obtener información de estas fuentes privadas, existe una gran variedad de estudios complementarios que resultaría interesante realizar, por ejemplo incorporar análisis de noticias relativas al sector, las cuales pueden reflejar efectos psicológicos que afecten el precio en el corto plazo. Un modelo de balance de oferta y demanda también podría resultar más adecuado para la proyección en el largo plazo, incluyendo la planificación de capacidad y la probabilidad de realización de esta.

Una última extensión que se deja planteada para futuras investigaciones es el uso de los principales tipos de cambio involucrados en el mercado de la celulosa. Las fibras bajo estudio se producen en distintas regiones debido a condiciones climáticas, y resultaría interesante analizar el impacto de las variaciones de los tipos de cambio asociados a cada fibra sobre el precio de estas. En el caso de obtenerse un buen desempeño con esta modificación, una proyección de largo plazo podría obtenerse mediante el uso de los futuros asociados a los tipos de cambio relevantes.

Glosario

ACF: función de autocorrelación

ADF: test de Dickey-Fuller aumentado

AR: modelo autorregresivo

ARMA: modelo autorregresivo de media móvil

ARX: modelo autorregresivo con variables exógenas

ARIMA: modelo autorregresivo integrado de media móvil

BCP: fibra química blanqueada

CAGR: tasa de crecimiento anual compuesto

CUSUM: gráfico de sumas acumuladas de las desviaciones

DW: estadístico de Durbin-Watson, mide autocorrelación entre residuos

Fibra corta: fibra química blanqueada proveniente de maderas latifoliadas (principalmente eucaliptus)

Fibra larga: fibra química blanqueada proveniente de maderas coníferas (principalmente pino)

MA: modelo de media móvil

MAE: error absoluto medio

MAPE: error porcentual absoluto medio

MCE: modelo de corrección de errores

MSE: error cuadrático medio

PACF: función de autocorrelación parcial

R-cuadrado: estadístico que determina la proporción de la varianza de la variable dependiente que es explicada por el modelo

R-cuadrado ajustado: estadístico que penaliza al R-cuadrado por la cantidad de regresores utilizados en el modelo

RMSE: raíz del error cuadrático medio

SARIMA: modelo autorregresivo integrado de media móvil con estacionalidad

SIC: criterio de información de Schwarz

Stepwise regression: procedimiento automático para selección de regresores

t-estadístico: estadístico utilizado para determinar la significancia estadística de un regresor

Bibliografía

Abildtrup, Jens, y otros. 1999. *Modern Time Series Analysis in Forest Products Markets*. Dordrecht : Springer Science+Business Media, B.V, 1999.

Banco Central de Chile. Base de datos estadísticos. [En línea] [Citado el: 03 de octubre de 2016.]

http://si3.bcentral.cl/Siete/secure/cuadros/arboles.aspx?idCuadro=BP6M_EXPORT.

Bashiri Mehmiri, Niaz y Pires Manso, José R. 2013. Crude Oil Price Forecasting Techniques: a Comprehensive Review of Literature. *Social Science Research Network*. 2013.

Buongiorno, Joseph. 1996. Forest sector modeling: a synthesis of econometrics, mathematical programming, and system dynamics methods. *International Journal of Forecasting*. 1996, Vol. 12, págs. 329-343.

Chen, Yu-Chin, Rogoff, Kenneth y Rossi, Barbara. 2008. Can exchange rates forecast commodity prices? 2008.

Departamento de Estudios Dirección General de Relaciones Económicas

Internacionales, Ministerio de Relaciones Exteriores. 2015. *Informe Anual Comercio Exterior de Chile 2014-2015*. 2015.

Empresas CMPC. 2016. Presentation for Itaú BBA's Breakfast. [En línea] julio de 2016. [Citado el: 01 de agosto de 2016.]

http://s2.q4cdn.com/093357708/files/doc_presentations/english/2016/Presentation-Itau-BBA-Breakfast.pdf.

Enders, Walter. 1995. *Applied Econometric Time Series*. 1995.

Engle, Robert F. y Yoo, Byung Sam. 1987. Forecasting and Testing in Co-Integrated Systems. *Journal of Econometrics*. 1987.

FAOSTAT. 2016. Forestry Production and Trade. [En línea] 2016. [Citado el: 22 de 09 de 2016.] <http://faostat.fao.org/beta/en/#data/FO>.

Food and Agriculture Organization of the United Nations. 2016. *Capacidades de pasta y papel, 2015-2020*. Roma : s.n., 2016.

Granger, C. W J. y Newbold, P. 1973. Some Comments on the Evaluation of Economic Forecasts. *Applied Economics*. 1973, Vol. 5.

Granger, C. W. J. 1969. Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*. 1969, Vol. 37.

Hyndman, Rob J. y Koehler, Anne B. 2006. Another look at measures of forecast accuracy. *International Journal of Forecasting*. 2006, Vol. 22, 4.

Kagraoka, Yusho. 2015. Common dynamic factors in driving commodity prices: Implications of a generalized dynamic factor model. *Economic Modelling*. 2015, Vol. 52, págs. 609-617.

Khan, Massarrat Ali. 2013. Forecasting of Gold Prices (Box Jenkins Approach). *International Journal of Emerging Technology and Advanced Engineering*. 2013, Vol. 3, 3.

Koehler, Anne B. y Murphree, Emily S. 1988. A comparison of the Akaike and Schwarz Criteria for Selecting Model Order. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*. 1988, Vol. 37, 2.

Riis, J. 1999. Forecasting Danish Timber Prices With an Error Correction Model. [aut. libro] Jens Abildtrup, y otros. *Modern Time Series Analysis in Forest Products Markets*. Dordrecht : Springer Science+Business Media, B.V, 1999.

Szabó, L, y otros. 2009. A world model of the pulp and paper industry: Demand, energy consumption and emission scenarios to 2030. *Environmental Science & Policy*. 2009, Vol. 12, 3.

Theil, H. 1971. Principles of Econometrics. 1971.

Toppinen, Anne y Kuuluvainen, Jari. 2010. Forest sector modelling in Europe - the state of the art and future research directions. *Forest Policy and Economics*. 2010, Vol. 12, 1.

Toppinen, Anne, Laaksonen, Susanna y Hänninen, Riitta. 1996. A Dynamic Forecasting Model for the Finnish Pulp Export Price. *Silva Fennica*. 1996, Vol. 30, 4.

UNECE/FAO. 2016. *Forest Products Annual Market Review 2015-2016*. 2016.

Ye, Michael, Zyren, John y Shore, Joanne. 2002. Forecasting Crude Oil Spot Price Using OECD Petroleum Inventory Levels. *International Advances in Economic Research*. 2002, Vol. 8, 4.

Anexo

A continuación se detalla el análisis realizado explicando los resultados obtenidos en cada caso. En todos los casos se corrigen los periodos de crisis y/o outliers identificados por medio de dummies.

Anexo A: Estimaciones modelo autorregresivo

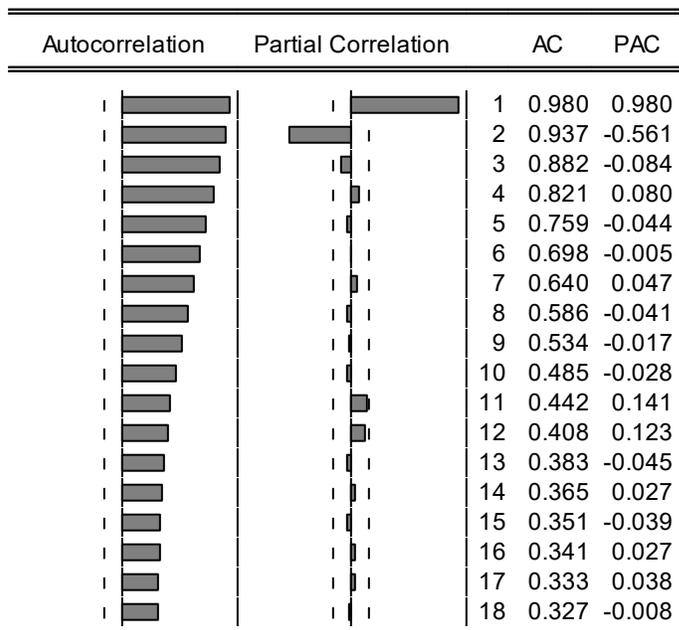
Se desea utilizar un modelo autorregresivo integrado de media móvil (ARIMA) para utilizar como proyección de referencia. Este utiliza información histórica de la variable en estudio para estimar los valores futuros mediante la siguiente representación:

$$\Delta^d x_t = \sum_{i=1}^p \beta_i \Delta^d x_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t$$

Donde p corresponde a la cantidad de rezagos de la variable dependiente a utilizar, d corresponde al orden de integración de dicha variable y q corresponde a la cantidad de errores previos a considerar.

Acorde con la metodología de Box-Jenkins el primer paso es identificar el orden de integración. A pesar que se demuestra en la sección 6.1 que todas son series integradas de primer orden, a continuación se presentan los correlogramas de las variables en nivel y en primera diferencia, considerando hasta 18 rezagos, para obtener un contraste.

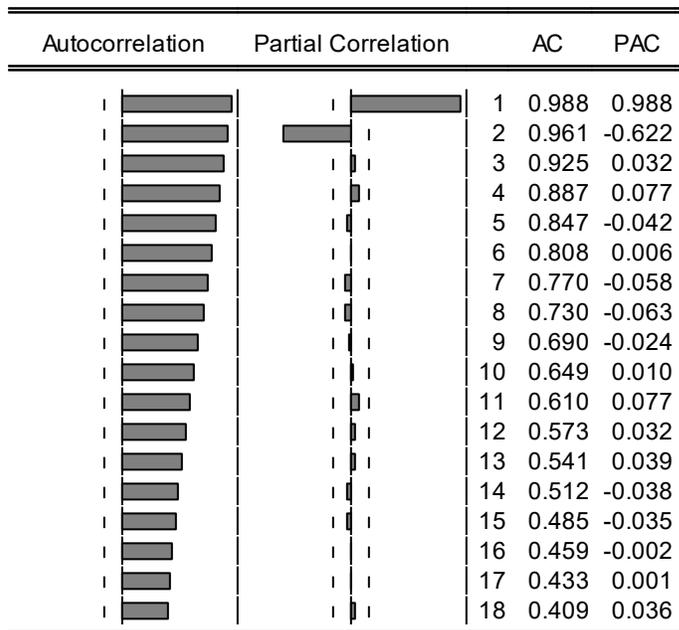
Gráfico 37: Correlograma precio de fibra corta (log)



Muestra: 01-2000 a 12-2012

Fuente: Elaboración propia

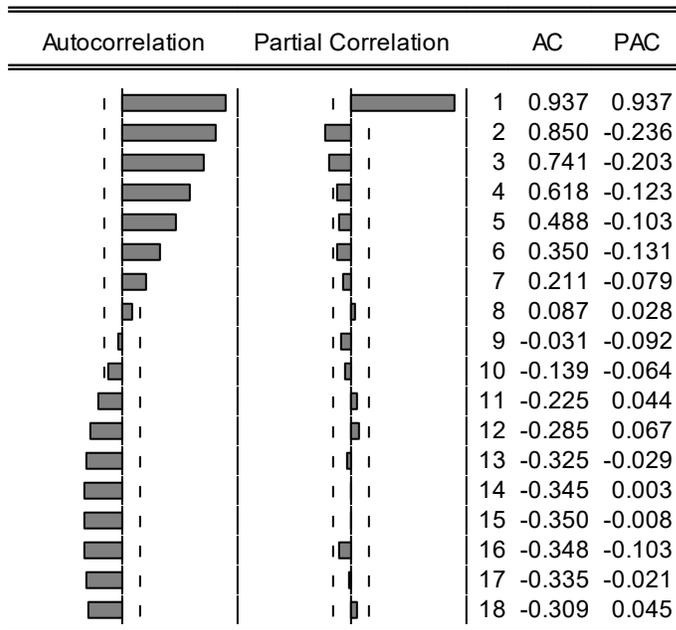
Gráfico 38: Correlograma precio de fibra larga (log)



Muestra: 01-2000 a 12-2012

Fuente: Elaboración propia

Gráfico 39: Correlograma ratio



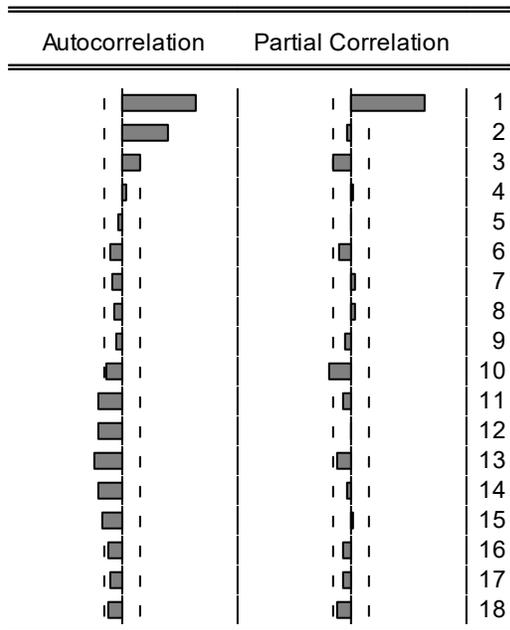
Muestra: 01-2000 a 12-2012

Fuente: Elaboración propia

Los valores “AC” y “PAC” corresponden a las funciones de autocorrelación y autocorrelación parcial definidas en la sección 4.2.1, respectivamente. Para series estacionarias se espera un alisamiento exponencial de ambas funciones, es decir, que decaigan rápidamente a cero. En todos los casos en estudio se obtiene que la autocorrelación no logra llegar a niveles no significativos después de varios periodos, esto es una señal de que se trata de series integradas y es necesario diferenciar para obtener una variable a la que se pueda aplicar el modelo ARMA(p,q).

Los Gráficos 40 a 42 muestran los correlogramas para la primera diferencia de las variables en estudio. En todos se observa que tanto la autocorrelación como la autocorrelación parcial caen rápidamente dentro de las bandas de confianza definidas por $1,96 \pm \frac{1}{\sqrt{n}}$, donde n es el tamaño de la muestra (delimitados por la línea discontinua que rodea el cero); manteniendo la mayoría de los rezagos dentro de las bandas.

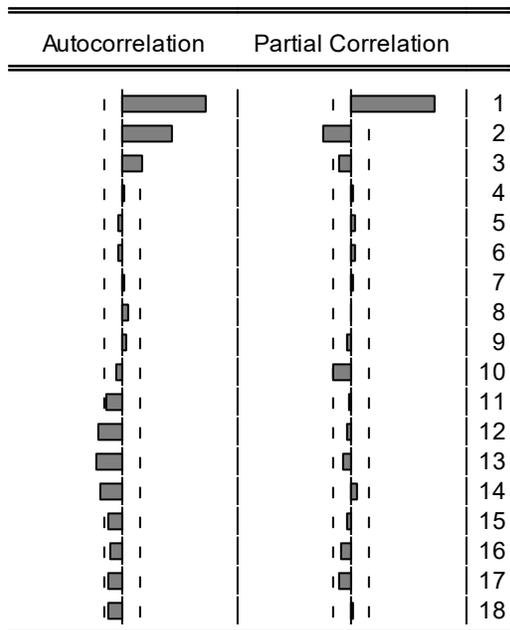
Gráfico 40: Correlograma primera diferencia precio de fibra corta (log)



Muestra: 01-2000 a 12-2012

Fuente: Elaboración propia

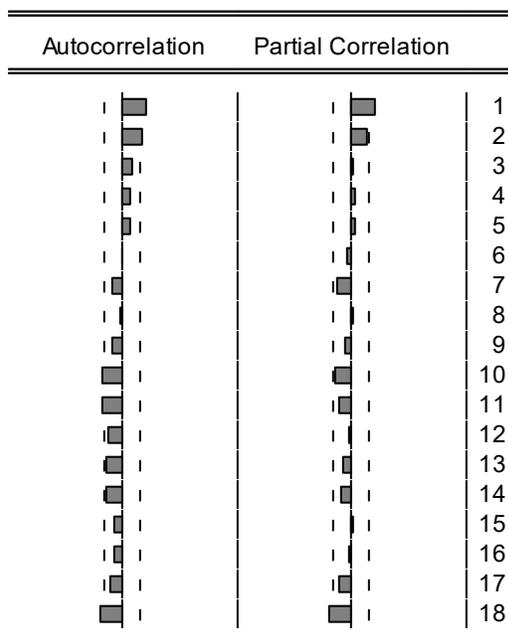
Gráfico 41: Correlograma primera diferencia precio de fibra larga (log)



Muestra: 01-2000 a 12-2012

Fuente: Elaboración propia

Gráfico 42: Correlograma primera diferencia ratio



Muestra: 01-2000 a 12-2012

Fuente: Elaboración propia

Adicionalmente se verifica si las series presentan comportamiento estacional mediante su descomposición en términos de tendencia, estacionalidad y error. La Tabla 10 presenta los resultados obtenidos:

Tabla 10: Análisis estacionalidad variables dependientes

Muestra: 01-2000 a 12-2012					
Variable	Fibra corta		Fibra larga		t-estadístico
	Coefficiente	t-estadístico	Coefficiente	t-estadístico	
Tendencia	0,01	13,34	0,02	14,94	
Intercepto	512,65	18,28	546,76	17,85	
Enero	12,06	0,32	19,22	0,47	
Febrero	15,43	0,41	20,69	0,51	
Marzo	16,85	0,45	27,19	0,67	
Abril	31,25	0,84	35,80	0,88	
Mayo	29,12	0,78	28,97	0,71	
Junio	31,08	0,83	32,53	0,80	
Julio	29,07	0,78	27,03	0,66	
Agosto	17,91	0,48	18,36	0,45	
Septiembre	21,75	0,58	19,66	0,48	
Octubre	16,32	0,44	15,82	0,39	
Noviembre	8,02	0,22	7,25	0,18	

Fuente: Estimaciones obtenidas usando EViews 9 SV

Las variables no presentan comportamiento estacional, dado que el t-estadístico no resulta significativo al presentar valores inferiores a 1,96, valor crítico para una confianza estadística de 95% (además los correlogramas tampoco muestran indicios de estacionalidad). Se descarta el uso de un modelo SARIMA.

Como la estructura de los correlogramas no coincide con las descritas en la sección 4.2.1, se asume que responden a un proceso más complejo. Con el fin de identificar potenciales parámetros p y q asociados a rezagos y promedios móviles de los errores, se estiman modelos haciendo variar dichos parámetros entre cero y cinco. El criterio de selección se basa en minimizar el SIC reportado, obteniendo los modelos más parsimoniosos.

Las Tablas 11 a 13 muestran los valores SIC obtenidos al hacer variar p y q:

Tabla 11: Identificación de parámetros p y q para fibra corta (log)

		Muestra: 01-2000 a 12-2012					
		p					
		0	1	2	3	4	5
q	0	-	-4,05	-4,02	-4,00	-3,97	-3,94
	1	-3,85	-4,02	-3,99	-3,97	-3,96	-3,93
	2	-4,00	-4,01	-3,97	-3,96	-3,93	-3,91
	3	-4,00	-3,97	-3,96	-3,93	-3,93	-3,89
	4	-3,97	-3,95	-3,93	-3,91	-3,90	-3,88
	5	-3,94	-3,91	-3,92	-3,89	-3,86	-3,85

Fuente: Elaboración propia

Tabla 12: Identificación de parámetros p y q para fibra larga (log)

		Muestra: 01-2000 a 12-2012					
		p					
		0	1	2	3	4	5
q	0	-	-4,71	-4,74	-4,71	-4,68	-4,64
	1	-4,47	-4,72	-4,71	-4,68	-4,64	-4,61
	2	-4,66	-4,71	-4,68	-4,65	-4,64	-4,61
	3	-4,70	-4,68	-4,64	-4,64	-4,61	-4,61
	4	-4,67	-4,64	-4,62	-4,61	-4,58	-4,59
	5	-4,64	-4,61	-4,59	-4,58	-4,55	-4,57

Fuente: Elaboración propia

Tabla 13: Identificación de parámetros p y q para ratio

		Muestra: 01-2000 a 12-2012					
		p					
		0	1	2	3	4	5
q	0	–	-4,58	-4,57	-4,53	-4,50	-4,47
	1	-4,56	-4,56	-4,53	-4,50	-4,53	-4,44
	2	-4,56	-4,53	-4,61	-4,47	-4,55	-4,45
	3	-4,54	-4,50	-4,47	-4,48	-4,52	-4,44
	4	-4,45	-4,50	-4,55	-4,46	-4,50	-4,47
	5	-4,48	-4,48	-4,52	-4,44	-4,47	-4,46

Fuente: Elaboración propia

Se destacan los cinco candidatos con menor valor de SIC, para los cuales se estiman los parámetros y se evalúa el ajuste de proyección en el periodo enero 2012 a marzo 2012 para seleccionar el modelo con mayor poder predictivo. El corto horizonte de proyección considerado se justifica por la limitada capacidad predictiva demostrada por los modelos autorregresivos, los que tienden a generar proyecciones constantes después de cierta cantidad de periodos.

Las Tablas 14 a 16 presentan la significancia estadística de cada regresor.

Tabla 14: Comparación de modelos estimados para fibra corta

		Muestra: 01-2000 a 12-2012				
		ARIMA(1,1,0)	ARIMA(2,1,0)	ARIMA(3,1,0)	ARIMA(1,1,1)	ARIMA(1,1,2)
R-cuadrado		0,43	0,43	0,44	0,43	0,44
RMSE		129,55	130,06	103,80	129,89	109,02
MAE		124,77	125,27	100,49	125,11	105,29
MAPE		19,34	16,40	13,17	16,38	13,79
AR(1) p-valor		0,00	0,00	0,00	0,00	0,00
AR(2) p-valor		–	0,31	0,04	–	–
AR(3) p-valor		–	–	0,04	–	–
MA(1) p-valor		–	–	–	0,98	0,14
MA(2) p-valor		–	–	–	–	0,03

Fuente: Estimaciones obtenidas usando EViews 9 SV

Tabla 15: Comparación de modelos estimados para fibra larga

Muestra: 01-2000 a 12-2012					
	ARIMA(1,1,0)	ARIMA(2,1,0)	ARIMA(3,1,0)	ARIMA(1,1,1)	ARIMA(2,1,1)
R-cuadrado	0,57	0,59	0,59	0,59	0,56
RMSE	63,19	33,22	32,62	42,71	33,19
MAE	58,50	27,94	28,62	36,08	29,41
MAPE	7,12	3,38	3,51	4,35	3,61
AR(1) p-valor	0,00	0,00	0,00	0,00	0,00
AR(2) p-valor	–	0,00	0,32	–	0,05
AR(3) p-valor	–	–	0,33	–	–
MA(1) p-valor	–	–	–	0,06	0,41

Fuente: Estimaciones obtenidas usando EViews 9 SV

Tabla 16: Comparación de modelos estimados para ratio

Muestra: 01-2000 a 12-2012					
	ARIMA(1,1,0)	ARIMA(2,1,0)	ARIMA(3,1,0)	ARIMA(1,1,1)	ARIMA(1,1,2)
R-cuadrado	0,38	0,52	0,35	0,56	0,36
RMSE	0,06	0,06	0,06	0,06	0,06
MAE	0,04	0,04	0,04	0,04	0,04
MAPE	3,73	3,60	3,78	3,55	3,56
AR(1) p-valor	0,01	0,02	–	0,00	0,05
AR(2) p-valor	–	0,07	–	–	–
MA(1) p-valor	–	–	0,05	0,03	0,15
MA(2) p-valor	–	–	–	–	0,61

Fuente: Estimaciones obtenidas usando EViews 9 SV

Se observa que existe una correlación entre el ajuste obtenido (R-cuadrado) y la significancia de los coeficientes para cada regresor entregada por su p-valor (interpretable como la probabilidad de que el regresor no resulte significativo, debe ser menor a 0,1 para una confianza estadística de 90%).

Al comparar tanto la significancia estadística de los regresores como las métricas de ajuste reportadas por cada modelo se elige un modelo ARIMA(3,1,0) para la fibra corta, un modelo ARIMA(2,1,0) para la fibra larga y un modelo ARIMA(1,1,1) para el ratio.

A continuación se muestran las estimaciones realizadas mediante modelos autorregresivos para fibra corta, fibra larga y ratio, respectivamente:

Tabla 17: Estimación modelo ARIMA para fibra corta

Muestra: 01-2000 12-2012			
Variable	Coeficiente		p-valor
AR(1)		0,67	0,00
AR(2)		0,08	0,04
AR(3)		-0,17	0,04
Estadísticos			
R-cuadrado ajustado	0,44	AIC	-4,13
DW	1,99	SIC	-4,15

Fuente: Estimaciones obtenidas usando EViews 9 SV

Tabla 18: Estimación modelo ARIMA para fibra larga

Muestra: 01-2000 12-2012			
Variable	Coeficiente		p-valor
AR(1)		0,94	0,00
AR(2)		-0,26	0,00
Estadísticos			
R-cuadrado ajustado	0,59	AIC	-4,84
DW	2,05	SIC	-4,78

Fuente: Estimaciones obtenidas usando EViews 9 SV

Tabla 19: Estimación modelo ARIMA para ratio

Muestra: 01-2000 12-2012			
Variable	Coeficiente		p-valor
AR(1)		0,67	0,00
MA(1)		-0,47	0,03
Estadísticos			
R-cuadrado ajustado	0,56	AIC	-4,62
DW	2,04	SIC	-4,56

Fuente: Estimaciones obtenidas usando EViews 9 SV

En todos los modelos estimados se observa que presentan un buen ajuste según su R-cuadrado ajustado, los parámetros resultan significativos y no se aprecia autocorrelación sobre los errores, ya que el estadístico de Durbin-Watson o DW es suficientemente cercano a dos (un valor menor a dos denota autocorrelación positiva, mientras que uno mayor denota autocorrelación negativa).

Los Gráficos 43 a 45 muestran los correlogramas sobre los residuos de los modelos estimados. Se observa que los residuos son ruido blanco (todos los rezagos de ACF y PACF se encuentran dentro de las bandas de confianza), por lo tanto no habría errores de especificación.

Gráfico 43: Correlograma residuos estimación fibra corta

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.003	0.003	0.0014	
		2	0.001	0.001	0.0017	
		3	-0.023	-0.023	0.0848	
		4	0.013	0.014	0.1141	0.736
		5	0.075	0.075	1.0179	0.601
		6	-0.084	-0.085	2.1648	0.539
		7	-0.067	-0.067	2.9003	0.575
		8	0.001	0.005	2.9004	0.715
		9	0.098	0.095	4.5126	0.608
		10	-0.082	-0.091	5.6283	0.584
		11	-0.129	-0.120	8.4508	0.391
		12	-0.033	-0.023	8.6356	0.472
		13	-0.099	-0.116	10.314	0.413
		14	-0.086	-0.116	11.581	0.396
		15	-0.057	-0.030	12.138	0.435
		16	-0.026	-0.014	12.254	0.507
		17	0.038	0.003	12.509	0.566
		18	-0.138	-0.169	15.883	0.390

Fuente: Elaboración propia

Gráfico 44: Correlograma residuos estimación fibra larga

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.027	-0.027	0.1146	
		2	0.084	0.083	1.2386	
		3	-0.031	-0.027	1.3959	0.237
		4	-0.048	-0.056	1.7600	0.415
		5	-0.047	-0.045	2.1213	0.548
		6	-0.017	-0.012	2.1703	0.704
		7	0.009	0.013	2.1836	0.823
		8	0.017	0.015	2.2307	0.897
		9	0.123	0.118	4.7506	0.690
		10	-0.051	-0.051	5.1788	0.738
		11	-0.069	-0.095	5.9893	0.741
		12	-0.031	-0.019	6.1565	0.802
		13	-0.144	-0.124	9.6993	0.558
		14	-0.058	-0.063	10.287	0.591
		15	0.003	0.014	10.289	0.670
		16	0.045	0.039	10.644	0.714
		17	-0.067	-0.093	11.432	0.721
		18	-0.048	-0.095	11.839	0.755

Fuente: Elaboración propia

Gráfico 45: Correlograma residuos estimación ratio

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.020	-0.020	0.0601	
		2 0.047	0.047	0.4099	
		3 -0.017	-0.016	0.4575	0.499
		4 0.014	0.012	0.4910	0.782
		5 0.044	0.046	0.8097	0.847
		6 0.003	0.004	0.8115	0.937
		7 -0.099	-0.103	2.4133	0.789
		8 0.029	0.026	2.5480	0.863
		9 -0.035	-0.026	2.7570	0.907
		10 -0.122	-0.134	5.2576	0.730
		11 -0.120	-0.122	7.6918	0.565
		12 -0.042	-0.029	7.9892	0.630
		13 -0.071	-0.073	8.8552	0.635
		14 -0.094	-0.113	10.375	0.583
		15 -0.005	0.012	10.379	0.663
		16 -0.007	0.003	10.388	0.733
		17 -0.035	-0.067	10.599	0.780
		18 -0.176	-0.205	16.085	0.447

Fuente: Elaboración propia

Anexo B: Estimaciones modelo de indicadores macroeconómicos

Se desea estudiar el impacto de cuatro indicadores globales sobre el precio de la celulosa, estos son:

- Inflación estadounidense
- Precio petróleo WTI
- All Country World Index o ACWI (indicador accionario global)
- Indicador Líder Compuesto G7

Estos demuestran tener un gran poder explicativo sobre el precio de commodities, y para la proyección se utiliza un modelo autorregresivo con variables exógenas representado a través de la siguiente ecuación:

$$\Delta x_t = \sum_{i=1}^{12} \beta_{1,i} \Delta Infl USA_{t-i} + \sum_{i=1}^{12} \beta_{2,i} \Delta WTI_{t-i} + \sum_{i=1}^{12} \beta_{3,i} \Delta ACWI_{t-i} + \sum_{i=1}^{12} \beta_{4,i} \Delta CLI G7_{t-i} \\ + \sum_{i=1}^{12} \beta_{5,i} \Delta p_{t-i}^{corta} + \sum_{i=1}^{12} \beta_{6,i} \Delta p_{t-i}^{larga}$$

Donde Δx_t corresponde a la primera diferencia de la variable dependiente a estudiar (descritas en la sección 6.1) en el periodo t , $\Delta Infl USA_t$ es la primera diferencia de la inflación estadounidense en el periodo t , ΔWTI_t es la primera diferencia del precio del petróleo en el periodo t , $\Delta ACWI_t$ es la primera diferencia del índice accionario global en el periodo t , $\Delta CLI G7_t$ es la primera diferencia del índice de producción industrial en el periodo t , Δp_t^{corta} es la primera diferencia del logaritmo del precio de la fibra corta y Δp_t^{larga} es la primera diferencia del logaritmo del precio de la fibra larga. Los rezagos de ambas fibras se introducen para considerar el efecto sustitución entre fibras, y en el caso del ratio se incorporan rezagos de la misma variable.

Para la obtención de este modelo se utiliza stepwise regression, un procedimiento automático que iterativamente selecciona los regresores estadísticamente significativos. De esta forma se utilizan todas las variables que presentan una importante influencia sobre el precio de la fibra, en particular se obtienen los rezagos de las variables exógenas que tienen una mayor influencia sobre las variables en estudio que los rezagos de la misma.

Dado que este modelo es aplicable a series estacionarias, en primer lugar es necesario investigar el orden de integración de las variables consideradas. La Tabla 20 muestra los resultados de los test de raíz unitaria sobre las variables exógenas a utilizar.

Tabla 20: Resultados test de raíz unitaria sobre indicadores macroeconómicos

Muestra: 01-2000 a 12-2012					
Variable	Dickey-Fuller Aumentado			Phillips-Perron	
	t-estadístico	K	Tendencia	Intercepto	t-estadístico
Nivel					
Inflación Estados Unidos	-0,96	12	No	No	-1,49
Precio WTI (log)	-3,04	1	Sí	Sí	-2,86
MSCI ACWI (log)	-0,01	1	No	No	-0,00
CLI G7 (log)	-3,08 *	4	No	Sí	-2,60
Primera diferencia					
Inflación Estados Unidos	-7,44 *	11	No	No	-7,97 *
Precio WTI (log)	-9,52 *	0	No	No	-9,50 *
MSCI ACWI (log)	-10,18 *	0	No	No	-10,32 *
CLI G7 (log)	-4,03 *	3	No	No	-3,36 *

Fuente: Estimaciones obtenidas usando EViews 9 SV

* denota rechazo de la hipótesis nula de raíz unitaria con 95% de confianza

En el caso de la inflación estadounidense, el precio del petróleo y el MSCI ACWI se acepta la hipótesis nula de raíz unitaria para las variables en nivel al usar ambos test, ya que el t-estadístico reportado resulta menor al valor crítico, indicando que las series son no estacionarias. En el caso del CLI G7, sólo el test de Phillips-Perron acepta la hipótesis nula de raíz unitaria, pero dado que al usar el test ADF el t-estadístico se encuentra suficientemente cerca de la zona de rechazo, y además el intercepto resulta significativo, se asume que la serie también resulta no estacionaria en nivel.

Para todas las variables en estudio la hipótesis nula de raíz unitaria es rechazada para la primera diferencia de las variables; por lo tanto se asume que todas las variables consideradas resultan integradas de primer orden.

Se estudia la estacionalidad de las series para considerar su variación mensual o anual en el modelo. La Tabla 21 muestra la descomposición en términos de estacionalidad, tendencia y error.

Tabla 21: Estacionalidad variables exógenas

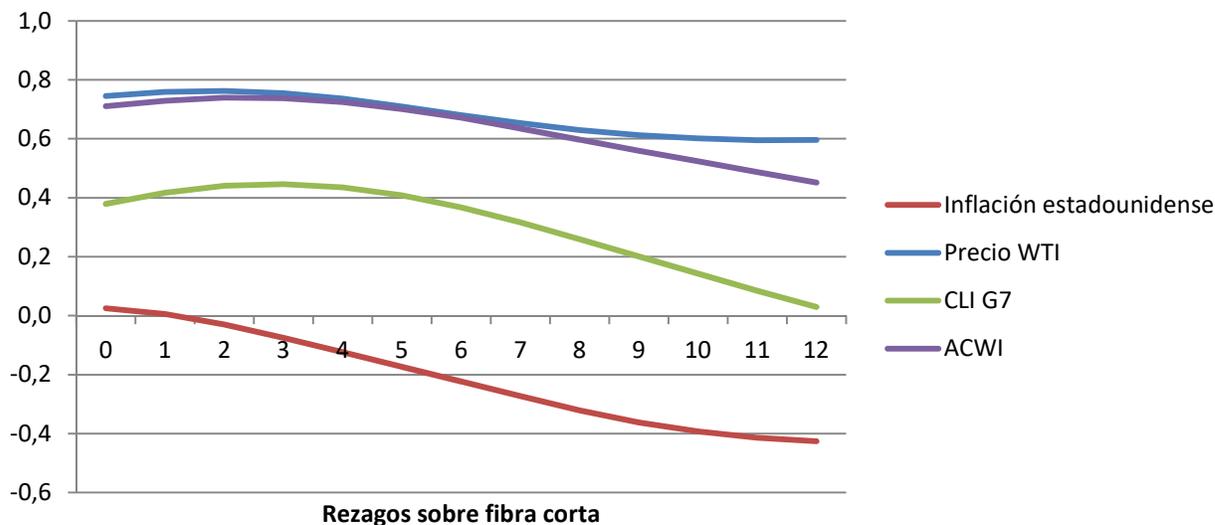
Muestra: 01-2000 a 12-2012									
Variable	Inflación USA		WTI		ACWI		CLI G7		
	Coef	t-estad	Coef	t-estad	Coef	t-estad	Coef	t-estad	
Tendencia	-	-	0,53	0,10	0,50	5,19	-	-	
Intercepto	2,74	6,85	13,25	5,10	253,14	17,10	99,91	233,48	
Enero	0,01	0,02	2,31	0,42	1,69	0,08	0,11	0,20	
Febrero	0,04	0,07	4,11	0,74	0,21	0,01	0,12	0,23	
Marzo	0,10	0,18	7,59	1,37	2,85	0,13	0,13	0,25	
Abril	0,06	0,12	6,68	1,21	8,02	0,38	0,14	0,26	
Mayo	0,08	0,14	6,21	1,12	4,09	0,19	0,13	0,25	
Junio	0,07	0,12	6,86	1,24	0,55	0,03	0,11	0,22	
Julio	0,03	0,06	7,53	1,36	0,02	0,01	0,08	0,16	
Agosto	0,06	0,12	6,82	1,23	-1,88	-0,09	0,05	0,09	
Septiembre	0,06	0,11	5,03	0,91	-6,23	-0,29	0,01	0,03	
Octubre	0,03	0,07	3,18	0,57	-4,83	-0,23	-0,01	-0,01	
Noviembre	0,01	0,02	1,95	0,35	-5,16	-0,24	-0,01	-0,02	

Fuente: Estimaciones obtenidas usando EViews 9 SV

Dado que no se aprecia estacionalidad en ninguno de los casos, sólo se considera la variación mensual de los indicadores.

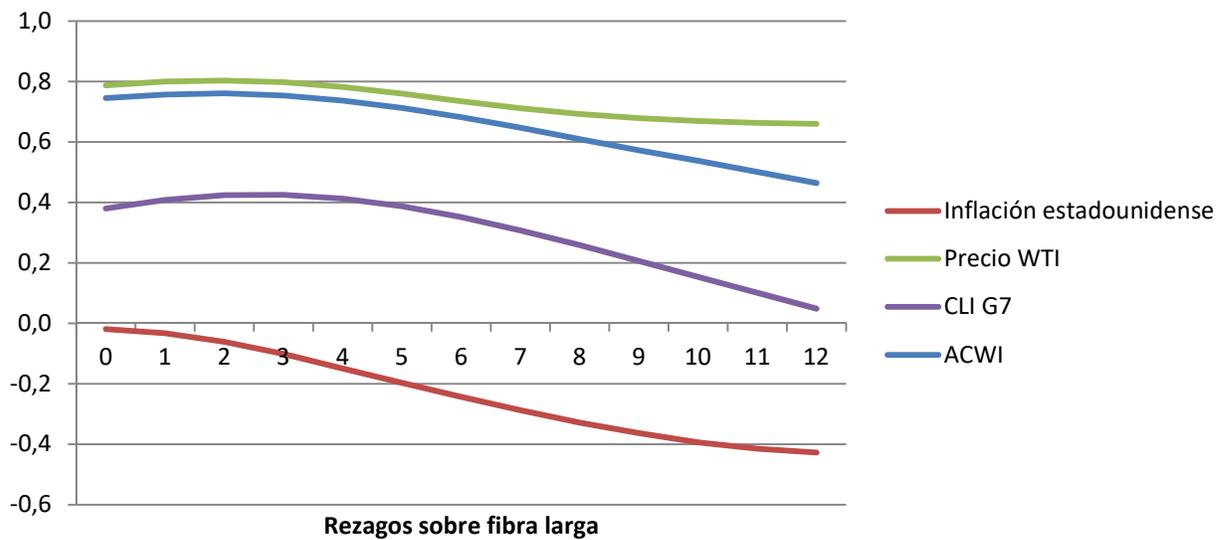
Los Gráficos 46 a 48 muestran la correlación dinámica entre las variables exógenas y las variables estudiadas.

Gráfico 46: Correlación dinámica entre fibra corta e indicadores macroeconómicos



Fuente: Elaboración propia

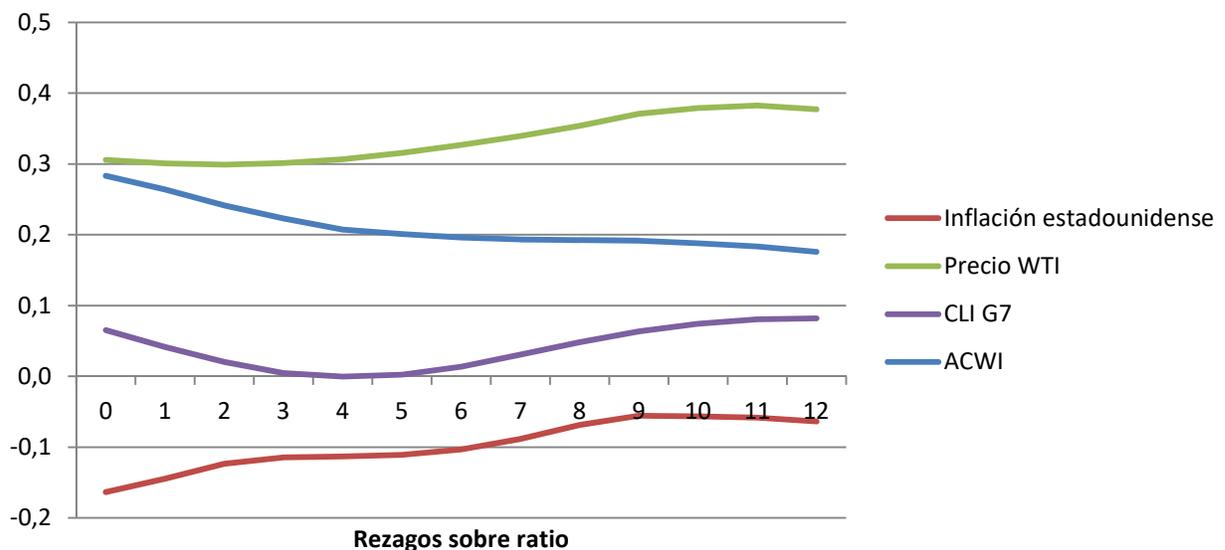
Gráfico 47: Correlación dinámica entre fibra larga e indicadores macroeconómicos



Fuente: Elaboración propia

En ambos casos se aprecia que el precio del petróleo, el CLI G7 y el ACWI empiezan a perder relevancia desde el periodo seis, mientras que para la inflación ocurre lo contrario, comenzando a adquirir importancia desde ese punto, aunque es la variable que presenta una correlación menos importante. Se podría esperar cierta coherencia entre los periodos con mayor correlación y los regresores reportados por el modelo.

Gráfico 48: Correlación dinámica entre ratio e indicadores macroeconómicos



Fuente: Elaboración propia

Finalmente el ratio es la variable que presenta las correlaciones más bajas y con menor estabilidad, lo que hace intuir que con los indicadores macroeconómicos utilizados podría no obtenerse un buen modelo para su proyección.

Adicionalmente se analizan los resultados del test de causalidad de Granger sobre las variables bajo estudio. Este es un test de precedencia estadística desarrollado por (Granger, 1969), usado para analizar si una variable resulta útil en la proyección de otra, y si esta relación es unidireccional o bidireccional.

Tabla 22: Test de causalidad de Granger sobre indicadores macroeconómicos

Hipótesis nula	p-valor
Fibra larga no causa fibra corta	0,00
Fibra corta no causa fibra larga	0,05
Ratio no causa fibra corta	0,00
Fibra corta no causa ratio	0,33
ACWI no causa fibra corta	0,00
Fibra corta no causa ACWI	0,31
CLI G7 no causa fibra corta	0,00
Fibra corta no causa CLI G7	0,08
Precio WTI no causa fibra corta	0,00
Fibra corta no causa precio WTI	0,58
Inflación USA no causa fibra corta	0,01
Fibra corta no causa inflación USA	0,10
Ratio no causa fibra larga	0,03
Fibra larga no causa ratio	0,32
ACWI no causa fibra larga	0,00
Fibra larga no causa ACWI	0,43
CLI G7 no causa fibra larga	0,00
Fibra larga no causa CLI G7	0,03
Precio WTI no causa fibra larga	0,00
Fibra larga no causa precio WTI	0,49
Inflación USA no causa fibra larga	0,00
Fibra larga no causa inflación USA	0,26
ACWI no causa ratio	0,58
Ratio no causa ACWI	0,33
CLI G7 no causa ratio	0,21
Ratio no causa CLI G7	0,53
Precio WTI no causa ratio	0,65
Ratio no causa precio WTI	0,88
Inflación USA no causa ratio	0,40
Ratio no causa inflación USA	0,06

Fuente: Elaboración propia

Se considera la siguiente relación:

$$y_t = \alpha_0 + \sum_{i=1}^n \alpha_i y_{t-i} + \sum_{i=1}^n \beta_i x_{t-i} + \varepsilon_t$$

Si los valores rezagados de x añaden información significativa para la proyección de y , se dice que x causa a y . Para seleccionar el número de rezagos a considerar se considera el criterio de información de Schwarz, seleccionando el número de rezagos que minimiza dicho valor.

La hipótesis nula de no causalidad se puede rechazar al obtener un p-valor igual o inferior a 0,5. En la Tabla 22 se aprecia que los indicadores macroeconómicos seleccionados son útiles en la proyección de las fibras, aunque esta causalidad no resulta bidireccional, es decir, los precios de las fibras no serían de utilidad en la proyección de dichos indicadores. Al igual que en la correlación dinámica, los peores resultados son obtenidos para el ratio entre las fibras, lo que indica que podrían no obtenerse buenos resultados en su proyección usando los indicadores seleccionados.

Las Tablas 23 a 25 muestran las estimaciones realizadas mediante modelos autorregresivos con variables explicativas para fibra corta, fibra larga y ratio, respectivamente:

Tabla 23: Estimación modelo de indicadores macroeconómicos para fibra corta

Muestra: 01-2000 a 12-2012			
Variable	Coeficiente		t-estadístico
$\Delta \log(p_{t-1}^{corta})$		0,35	3,72
$\Delta \log(CLI\ G7_{t-2})$		4,34	3,55
$\Delta \log(p_{t-2}^{larga})$		-0,36	-3,80
$\Delta \log(WTI_{t-1})$		0,08	3,31
$\Delta \log(p_{t-1}^{larga})$		0,47	3,49
$\Delta \log(WTI_{t-3})$		0,07	2,72
$\Delta Infl\ USA_{t-2}$		-0,01	-2,25
Estadísticos			
R-cuadrado ajustado	0,61	AIC	-4,46
DW	2,15	SIC	-4,33

Fuente: Estimaciones obtenidas usando EViews 9 SV

Tabla 24: Estimaciones modelo de indicadores macroeconómicos para fibra larga

Muestra: 01-2000 a 12-2012			
Variable	Coeficiente		t-estadístico
$\Delta \log(p_{t-1}^{larga})$		0,50	5,87
$\Delta \log(CLI\ G7_{t-1})$		2,65	3,06
$\Delta \log(p_{t-3}^{corta})$		-0,21	-4,58
$\Delta \log(p_{t-1}^{corta})$		0,22	3,28
$\Delta \log(WTI_{t-1})$		0,05	3,11
$\Delta \log(p_{t-9}^{corta})$		0,15	3,20
$\Delta \log(ACWI_{t-7})$		-0,10	-3,12
$\Delta \log(CLI\ G7_{t-6})$		2,57	2,70
$\Delta \log(p_{t-10}^{larga})$		-0,11	-1,90
$\Delta Infl\ USA_{t-4}$		-0,01	-1,66
Estadísticos			
R-cuadrado ajustado	0,75	AIC	-5,26
DW	2,01	SIC	-5,04

Fuente: Estimaciones obtenidas usando EViews 9 SV

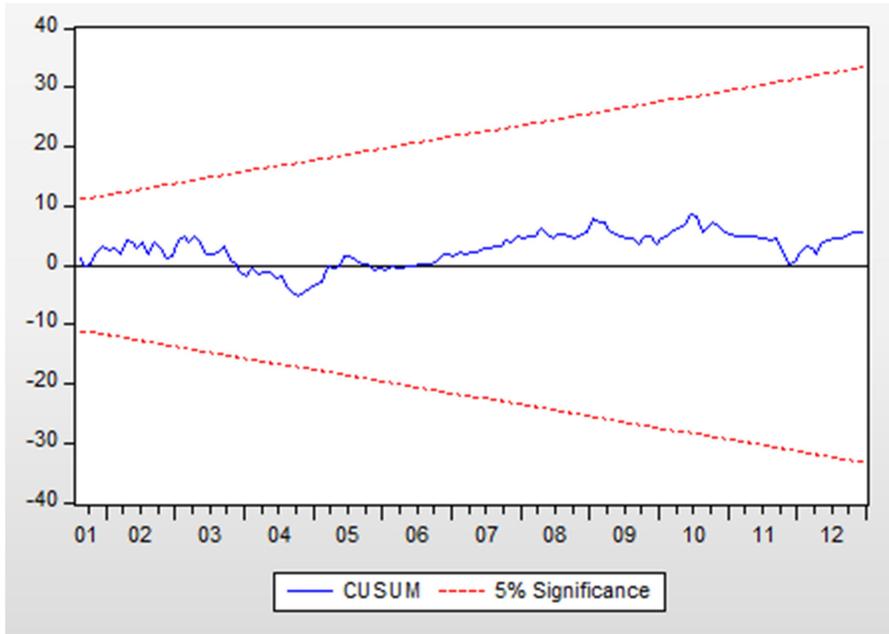
Tabla 25: Estimaciones modelo de indicadores macroeconómicos para ratio

Muestra: 01-2000 a 12-2012			
Variable	Coeficiente		t-estadístico
$\Delta \log(p_{t-10}^{corta})$		0,14	3,39
$\Delta \log(WTI_{t-3})$		-0,06	-2,99
$\Delta ratio_{t-2}$		0,15	2,03
$\Delta Infl\ USA_{t-2}$		0,01	3,09
$\Delta \log(WTI_{t-2})$		-0,06	-2,70
$\Delta ratio_{t-11}$		-0,17	-2,18
$\Delta \log(WTI_{t-8})$		-0,05	-2,27
$\Delta \log(p_{t-6}^{larga})$		0,12	2,18
$\Delta Infl\ USA_{t-6}$		-0,01	-2,00
Estadísticos			
R-cuadrado ajustado	0,26	AIC	-4,96
DW	1,97	SIC	-4,78

Fuente: Estimaciones obtenidas usando EViews 9 SV

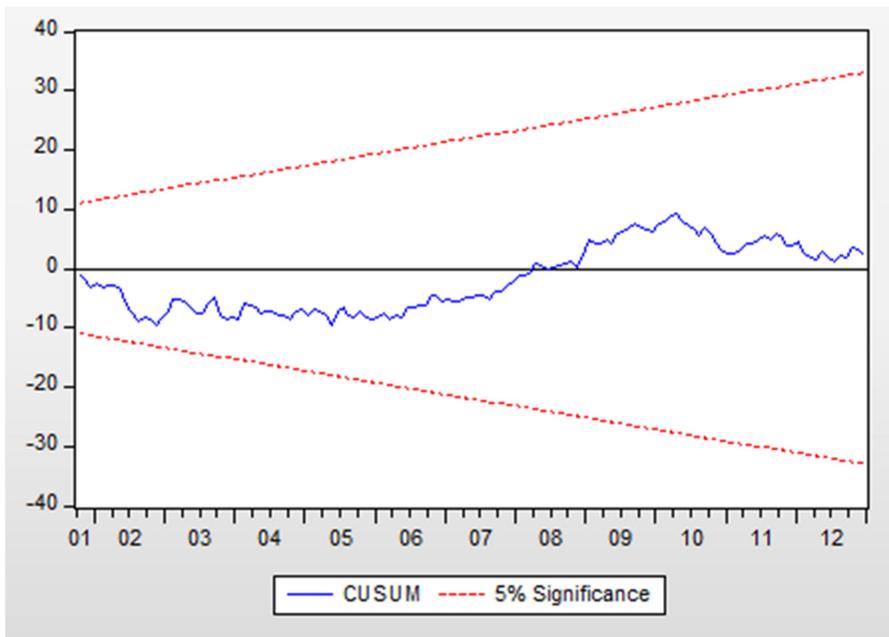
Se obtiene un modelo con un buen ajuste y sin problemas aparentes de autocorrelación según el estadístico Durbin-Watson. Para verificar su estabilidad se obtienen las sumas acumuladas de las desviaciones o CUSUM; en este, dado que las desviaciones se acumulan periodo a periodo, una leve desviación con respecto a la media del proceso conduce a valores en constante aumento o disminución, por esto, si este se encuentra dentro del intervalo de confianza para un nivel de confianza de 95%, delimitado por las líneas rojas, el modelo estaría bien especificado.

Gráfico 49: CUSUM modelo de indicadores macroeconómicos fibra corta



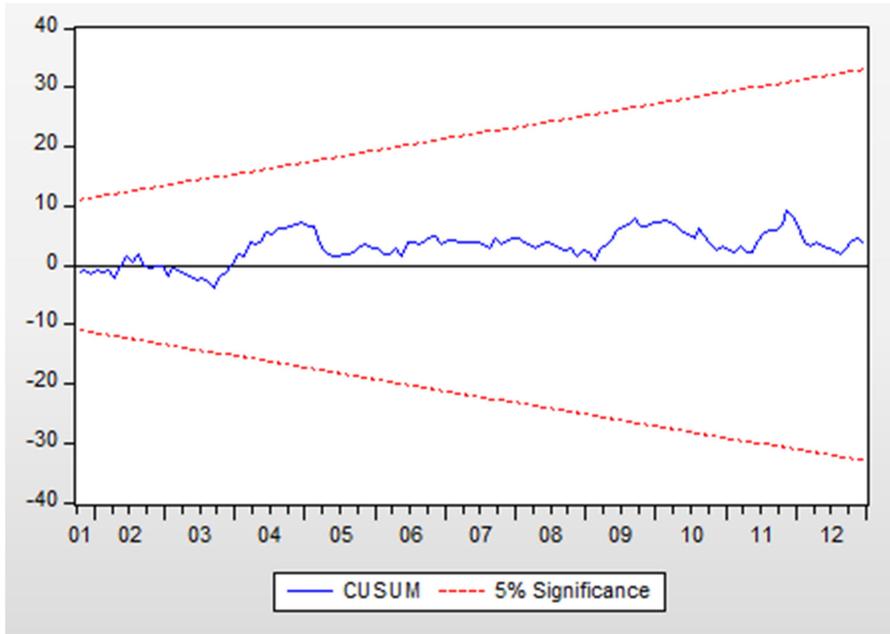
Fuente: Estimaciones obtenidas usando EViews 9 SV

Gráfico 50: CUSUM modelo indicadores macroeconómicos fibra larga



Fuente: Estimaciones obtenidas usando EViews 9 SV

Gráfico 51: CUSUM modelo indicadores macroeconómicos ratio



Fuente: Estimaciones obtenidas usando EViews 9 SV

Anexo C: Estimaciones modelo corrección de errores

En el estudio se consideran los datos de días de inventario de oferta y demanda descritos en la sección 5.3. En el caso de los días de inventario de oferta, las series reportadas por Hawkins Wright Ltd. y Pulp and Paper Products Council son prácticamente idénticas, por lo que se empalman a fin de obtener una historia de datos más larga. De esta forma se obtienen las siguientes series de inventarios a estudiar:

- Oferta fibra corta
- Oferta fibra larga
- Oferta BCP
- Demanda fibra corta
- Demanda fibra larga
- Demanda BCP

En primer lugar se debe determinar el grado de integración de las variables a utilizar a fin de verificar si el modelo es aplicable. La Tabla 26 reporta los test de raíz unitaria descritos en la sección 4.1, aplicados a los datos de inventario.

Tabla 26: Test de raíz unitaria sobre series de inventario

Muestra: 01-2000 a 12-2012					
Variable	Dickey-Fuller Aumentado			Phillips-Perron	
	t-estadístico	k	Tendencia	Intercepto	t-estadístico
Nivel					
Demanda fibra corta (log)	-3,70 *	0	Sí	Sí	-3,42
Demanda fibra larga (log)	-3,59 *	0	Sí	Sí	-3,60 *
Demanda BCP (log)	-3,90 *	0	Sí	Sí	-3,66 *
Oferta fibra corta (log)	-2,77	0	No	Sí	-3,22 *
Oferta fibra larga (log)	-2,66	0	No	Sí	-2,66
Oferta BCP (log)	-3,75 *	6	No	Sí	-3,04 *
Primera diferencia					
Demanda fibra corta (log)	-10,36 *	1	No	No	-13,53 *
Demanda fibra larga (log)	-13,55 *	0	No	No	-13,55 *
Demanda BCP (log)	-10,46 *	1	No	No	-13,44 *
Oferta fibra corta (log)	-8,68 *	0	No	No	-8,70 *
Oferta fibra larga (log)	-9,36 *	0	No	No	-9,34 *
Oferta BCP (log)	-7,43 *	3	No	No	-7,85 *

Fuente: Estimaciones obtenidas usando EViews 9 SV

Se asume que todas las variables en nivel resultan no estacionarias, ya que en todos los casos se obtiene que el intercepto resulta significativo y el t-estadístico se encuentra demasiado cerca de la zona de aceptación, por ende no se puede rechazar la hipótesis nula de raíz unitaria.

Para la primera diferencia de todas las variables en estudio se obtiene rechazo de la hipótesis nula de raíz unitaria, por ende se obtiene que todas las variables resultan integradas de primer orden y se evaluará la existencia de cointegración entre estas y el precio de la fibra.

Una vez determinado el orden de integración de las variables, se procede a verificar la existencia de cointegración mediante el procedimiento de Engle-Granger, el cual consiste en estimar la relación de largo plazo mediante mínimos cuadrados ordinarios y testear estacionariedad de los residuos. Para esto se considera la siguiente relación de largo plazo:

$$\log(p_t^i) = \alpha + \beta \log(inv_t^i) + \varepsilon_t$$

Donde p_t^i corresponde al precio de la fibra i en el periodo t , e inv_t^i corresponde al logaritmo del inventario considerado para la fibra i en el periodo t . Se espera una relación negativa entre el precio de la fibra y los días de inventario, ya que un mayor nivel de precios produce estrechez en el mercado reduciendo los inventarios, mientras que un menor nivel de precios produce holguras en el mercado, aumentando los inventarios.

La Tabla 27 muestra los resultados del procedimiento de Engle-Granger para las funciones estimadas:

Tabla 27: Resultados procedimiento de Engle-Granger sobre precios e inventarios

Variable independiente	Coefficiente	t-estadístico	R-cuadrado	ADF sobre residuos (t-estadístico)
Variable dependiente: fibra corta				
Demanda fibra corta (log)	-0,93	-14,88	0,61	-3,46 **
Demanda BCP (log)	-1,02	-13,81	0,57	-3,53 **
Oferta fibra corta (log)	0,28	2,71	0,07	-2,52
Oferta BCP (log)	0,02	0,17	0,00	-3,27 *
Variable dependiente: fibra larga				
Demanda fibra larga (log)	-1,03	-13,67	0,57	-2,94
Demanda BCP (log)	-1,15	-15,90	0,64	-3,23 *
Oferta fibra larga (log)	-0,25	-2,44	0,06	-2,00
Oferta BCP (log)	-0,03	-0,25	0,00	-2,71

Fuente: Estimaciones obtenidas usando EViews 9 SV

Todos los casos destacados en la Tabla 27 muestran existencia de cointegración al rechazar la hipótesis nula de raíz unitaria sobre los residuos, ya que el t-estadístico obtenido con el test ADF resulta mayor que el valor crítico tabulado por Engle y Yoo, y tanto intercepto como tendencia resultan no significativos. Se decide no considerar los inventarios de oferta de BCP para estimar las variaciones en el precio de la fibra corta, ya que tanto el t-estadístico como el R-cuadrado reportados indican que la relación de largo plazo propuesta no es significativa y no debe ser considerada para proyectar el precio de la fibra corta.

En general se observa que los inventarios de demanda tienen una relación más consistente con el nivel de precios, tanto en orden de magnitud del coeficiente como en significancia estadística según el t-estadístico y R-cuadrado reportado. A pesar de que estos no reflejan las perturbaciones producidas por crisis o efectos externos (ver sección 1.1) podrían ser útiles en la proyección de largo plazo.

Dado que los inventarios agregados (BCP) consideran los otros dos tipos utilizados y podría resultar una regresión lineal de estos, se decide evaluar el efecto de cada tipo de inventario sobre las fibras en estudio.

Tabla 28: Evaluación de colinealidad de fibras

Muestra: 01-2000 a 12-2012				
Variable dependiente	Fibra corta		Fibra larga	
Variable independiente	Coefficiente	t-estadístico	Coefficiente	t-estadístico
Demanda fibra corta (log)	-0,88	-3,77	-	-
Demanda fibra larga (log)	-	-	0,22	0,92
Demanda BCP (log)	-0,05	-0,20	-1,37	-5,41

Fuente: Estimaciones obtenidas usando EViews 9 SV

Al no obtener una mejora con respecto a la relación considerada anteriormente (los t-estadísticos de la fibra complementaria no resultan significativos), sólo se considera el inventario de fibra corta para la proyección del precio de fibra corta y el inventario de BCP para la proyección del precio de fibra larga.

De forma análoga al Anexo B: Estimaciones modelo de indicadores macroeconómicos, se analiza la existencia de causalidad de Granger entre los inventarios a utilizar y las fibras bajo estudio.

Tabla 29: Test de causalidad de Granger sobre inventarios

Hipótesis nula	p-valor
Fibra larga no causa demanda fibra corta	0,00
Demanda fibra corta no causa fibra larga	0,00
Oferta fibra larga no causa fibra larga	0,00
Fibra larga no causa oferta fibra larga	0,01
Oferta BCP no causa fibra larga	0,00
Fibra larga no causa oferta BCP	0,00
Demanda fibra larga no causa fibra larga	0,46
Fibra larga no causa demanda fibra larga	0,01

Fuente: Elaboración propia

Se obtiene una causalidad fuerte y bidireccional en todos los casos, a excepción de los inventarios de demanda de fibra larga, lo que hace intuir que podrían no obtenerse buenos resultados al usarlos en la proyección del precio de fibra larga.

Se procede a estimar el modelo propuesto:

$$\Delta \log(p_t^i) = \delta \Delta \log(p_{t-1}^i) + \varphi \Delta \log(inv_{t-1}^i) - \alpha (\log(p_{t-1}^i) - \beta \log(inv_{t-1}^i)) + c + \varepsilon_t$$

Donde p_t^i corresponde al precio de la fibra i en el periodo t , inv_t^i corresponde al logaritmo del inventario considerado para la fibra i en el periodo t y $(\log(p_{t-1}^i) - \beta \log(inv_{t-1}^i))$ corresponde a la relación de cointegración.

Se realizan proyecciones para ambas fibras utilizando las variables con las que se obtienen relaciones de cointegración. La Tabla 30 resume las métricas de ajuste de cada proyección sobre el periodo 01-2013 a 10-2014.

Tabla 30: Ajuste de proyecciones para modelo de corrección de errores

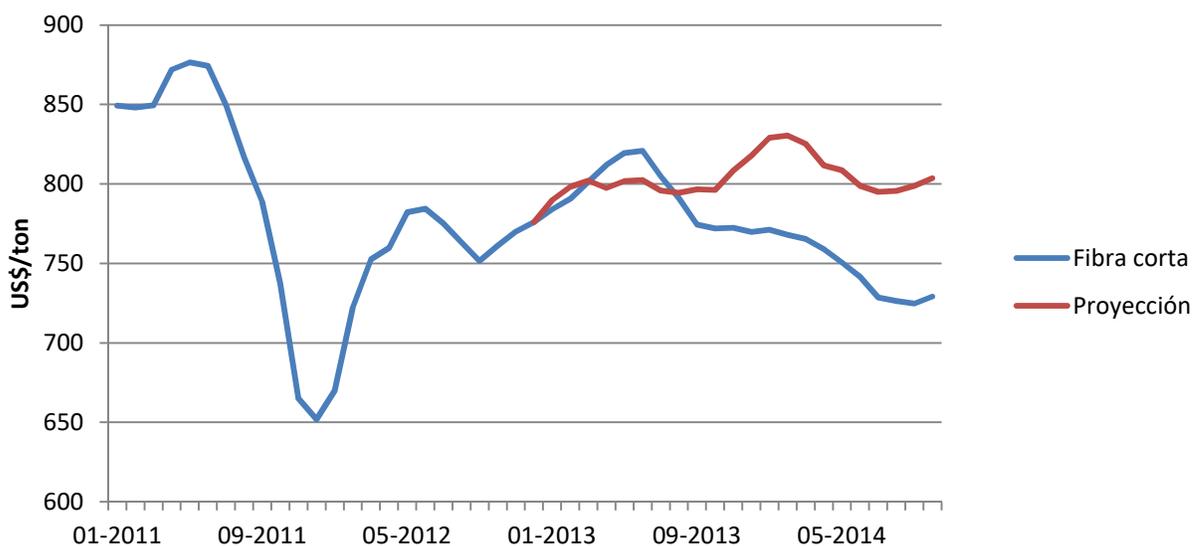
Muestra: 01-2000 a 12-2012			
Variable independiente	RMSE	MAE	MAPE
Variable dependiente: fibra corta			
Demanda fibra corta (log)	35,22	28,65	3,80
Demanda BCP (log)	64,20	50,79	6,77
Variable dependiente: fibra larga			
Demanda BCP (log)	46,56	41,66	4,65

Fuente: Estimaciones obtenidas usando EViews 9 SV

En el caso de la fibra corta se obtienen mejores resultados al utilizar los inventarios de demanda de fibra corta, mientras que en la fibra larga sólo se puede utilizar los inventarios de demanda agregados, con los cuales se obtienen errores similares a los reportados para fibra corta.

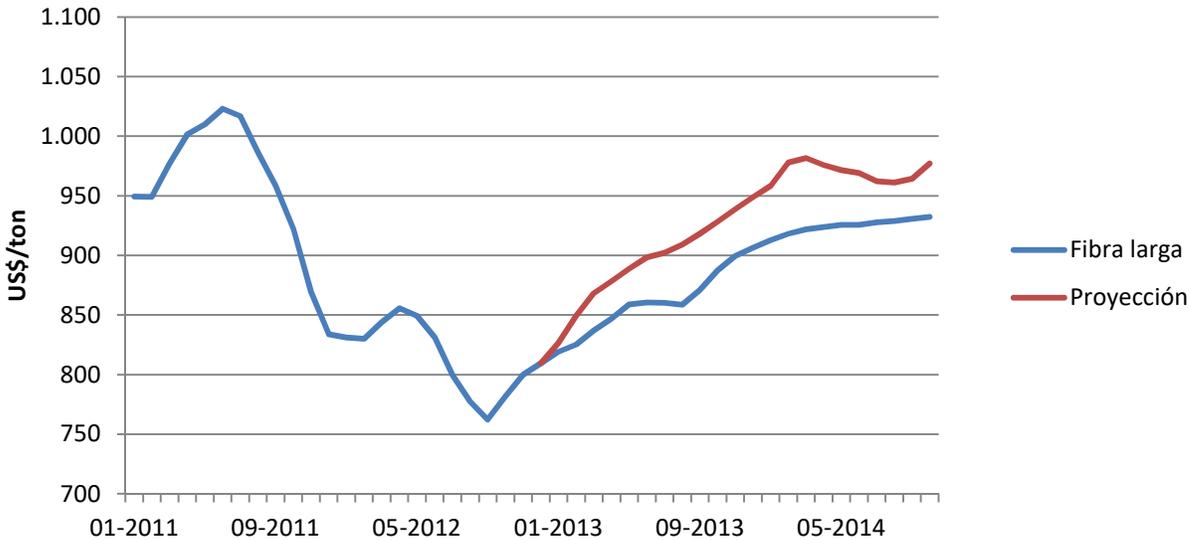
Los Gráficos 52 y 53 muestran las proyecciones dinámicas para el periodo enero 2013 a octubre 2014 (sólo se cuenta con inventarios de demanda hasta octubre 2014) realizadas para ambos tipos de fibra usando los datos reales de inventario. Para proyectar fibra corta se utilizan días de inventario de demanda de fibra corta y para fibra larga se utilizan días de inventario de demanda de BCP.

Gráfico 52: Proyección fibra corta con datos originales de inventario



Fuente: Elaboración propia

Gráfico 53: Proyección fibra larga con datos originales de inventario

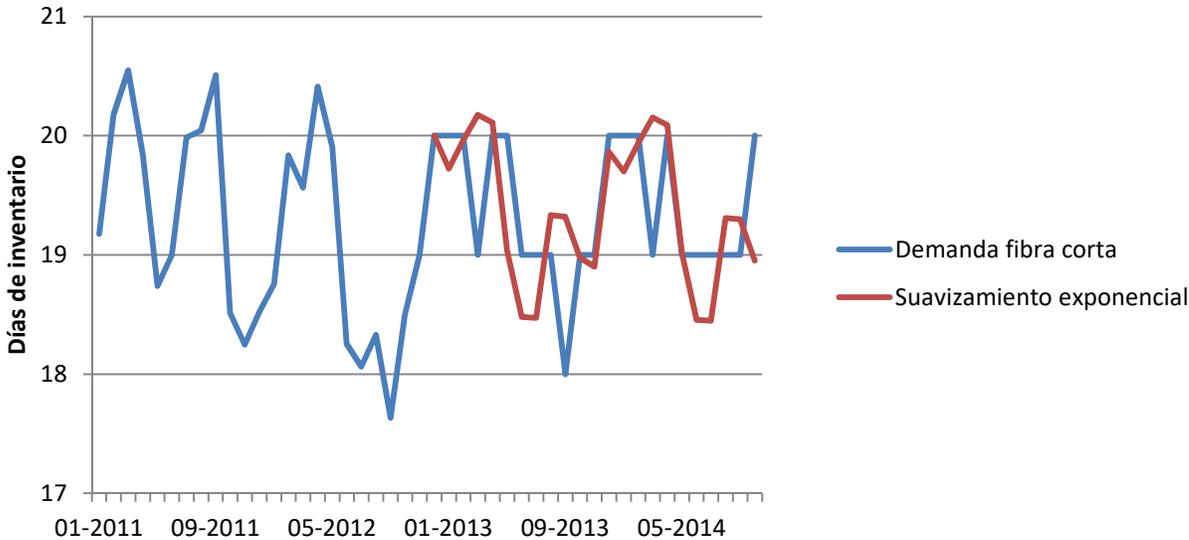


Fuente: Elaboración propia

A pesar de que se obtiene una proyección con métricas de error bajas (MAPE menor a 10%), es evidente que para proyectar el precio en el largo plazo es necesario también realizar una proyección sobre los días de inventario. Para esto se realiza una proyección de los dos tipos de inventario a utilizar mediante suavizamiento exponencial.

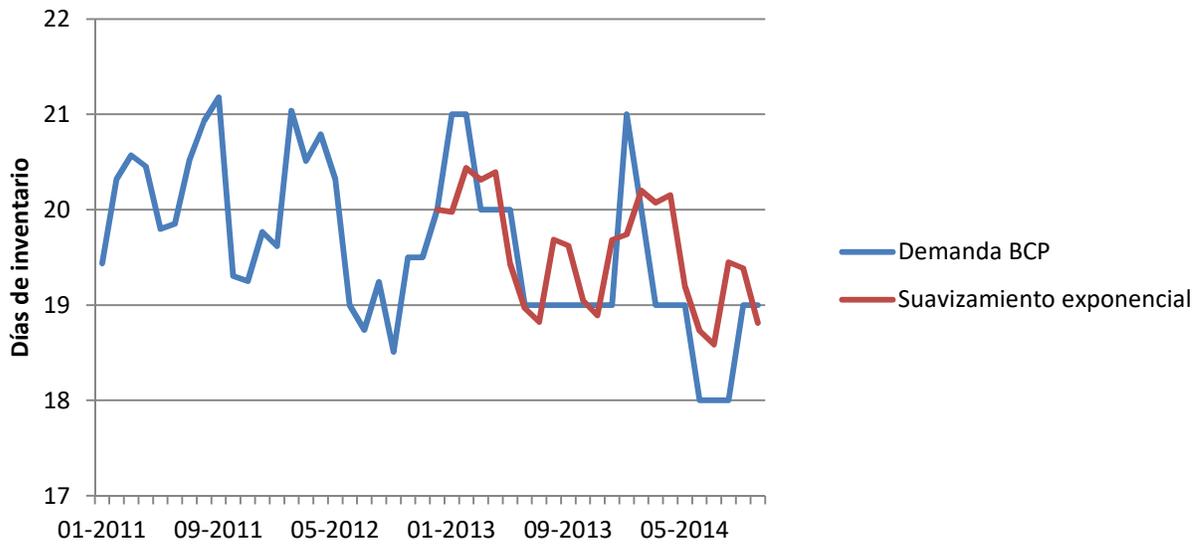
Los Gráficos 54 y 55 muestran las proyecciones realizadas sobre los días de inventario de fibra corta y BCP, respectivamente, usando la muestra 01-2010 a 12-2012, dado que en este periodo se identifica un cambio de tendencia.

Gráfico 54: Proyección sobre inventarios de demanda fibra corta



Fuente: Elaboración propia

Gráfico 55: Proyección sobre inventarios de demanda BCP

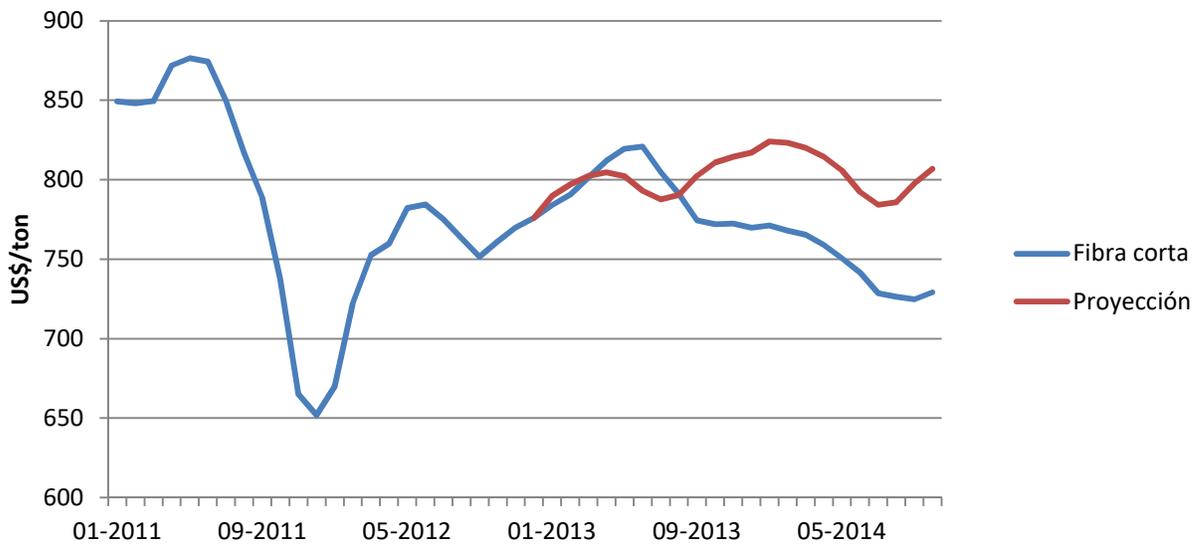


Fuente: Elaboración propia

En ambos casos se aprecia que la proyección mediante suavizamiento exponencial mantiene la tendencia y estacionalidad que presentan estos en los periodos anteriores.

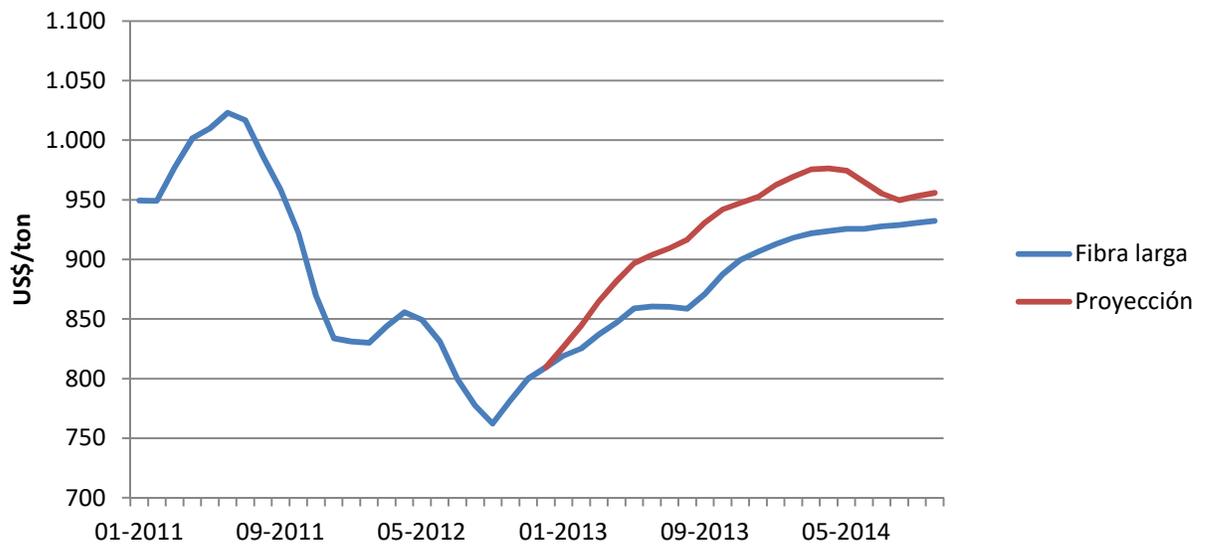
Los Gráficos 56 y 57 muestran las proyecciones realizadas para el precio de ambas fibras para el periodo enero 2013 a abril 2016 usando la proyección auxiliar de datos de inventarios.

Gráfico 56: Proyección fibra corta con datos estimados de inventarios



Fuente: Estimación propia

Gráfico 57: Proyección fibra larga con datos estimados de inventarios



Fuente: Estimación propia

En ambos casos el modelo ofrece una proyección estable basada en la relación estacionaria entre precios e inventarios; si bien el precio puede verse afectado por efectos externos en el corto plazo, este debería retornar al equilibrio mencionado.

Las Tablas 31 y 32 presentan las estimaciones realizadas para el modelo propuesto:

$$\Delta p_t^i = \delta \Delta p_{t-1}^i + \varphi \Delta f(inv_{t-1}^j) - \alpha (p_{t-1}^i - \beta f(inv_{t-1}^j)) + c + \varepsilon_t$$

Tabla 31: Estimaciones modelo de corrección de errores para fibra corta

Muestra: 01-2000 a 12-2012			
Parámetro	Coeficiente		t-estadístico
δ		0,68	11,71
φ		0,21	4,23
α		0,11	5,42
β		0,91	8,31
c		1,07	5,33
Estadísticos			
R-cuadrado ajustado	0,57	AIC	-4,35
DW	2,23	SIC	-4,20

Fuente: Estimaciones obtenidas usando EViews 9 SV

Tabla 32: Estimaciones modelo de corrección de errores para fibra larga

Muestra: 01-2000 a 12-2012			
Parámetro	Coeficiente		t-estadístico
δ		0,74	15,19
φ		0,17	4,04
α		0,05	4,05
β		1,11	5,96
c		0,55	3,91
Estadísticos			
R-cuadrado ajustado	0,69	AIC	-5,05
DW	1,94	SIC	-4,88

Fuente: Estimaciones obtenidas usando EViews 9 SV