

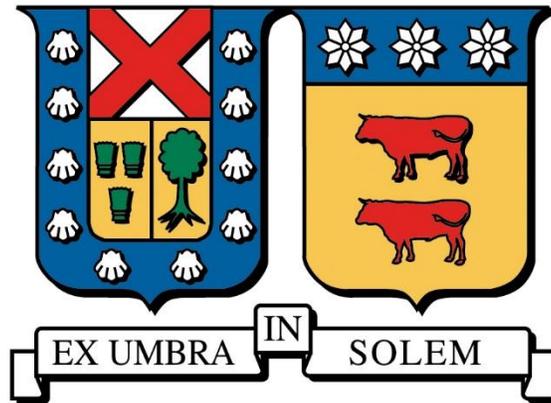
2018

PROPUESTA DE MODELOS ADAPTATIVOS HÍBRIDOS Y NO-HÍBRIDOS PARA EL PRONÓSTICO DE VOLATILIDAD DEL PRECIO MENSUAL DEL COBRE MEDIANTE ALGORITMOS GENÉTICOS

GARCÍA BOZZO, DIEGO IGNACIO

<http://hdl.handle.net/11673/40881>

Repositorio Digital USM, UNIVERSIDAD TECNICA FEDERICO SANTA MARIA



Universidad Técnica Federico Santa María
Departamento de Industrias

**PROPUESTA DE MODELOS ADAPTATIVOS HÍBRIDOS Y NO-
HÍBRIDOS PARA EL PRONÓSTICO DE VOLATILIDAD DEL
PRECIO MENSUAL DEL COBRE MEDIANTE ALGORITMOS
GENÉTICOS**

Tesis para optar al grado de **Ingeniero Civil Industrial**
y de **Magíster en Ciencias de la Ingeniería Industrial**

Autor

DIEGO IGNACIO GARCÍA BOZZO

Profesor Guía

Dr. Werner Kristjanpoller R.

Profesor Correferente Interno

Dr. Javier Scavia D.

Profesor Correferente Externo

Dr. Juan Ignacio Guzmán B.

VALPARAISO, DICIEMBRE 2017

AGRADECIMIENTOS

A mis padres. Por estar siempre ahí. Por creer siempre en mis capacidades. Pero por sobretodo, por confiar en mis decisiones.

A mi hermano y hermana mayor. Estando lejos o estando cerca, me han apoyado siempre que han podido y cuando he necesitado.

A mis amigos cercanos. Parte fundamental de mi vida, nuestro vínculo es de hermandad. Los amigos son la familia que uno elige.

A mis amigos cercanos que recién han partido, de forma inesperada. El Benja. El Pancho. El recuerdo vive más que nunca, y vivirá por siempre.

A mis compañeros de universidad de pregrado, y de magíster. Los momentos vividos estos años no pasan desapercibidos.

A mis profesores. Por la formación entregada. Por confiar en este estudiante.

Gracias a todos ustedes... Por hacer posible esto.

INDICE DE CONTENIDOS

Abstract	5
Resumen Ejecutivo	6
1. Introducción y descripción de la problemática	7
2. Objetivos	15
2.1 Objetivo General.....	15
2.2 Objetivos Específicos	15
3. Revisión de la Literatura	16
4. Marco Teórico	20
4.1 Fundamentos económicos.....	20
4.1.1 Inventarios en Bolsas de Metales.....	21
4.1.2 La actividad especulativa y la participación de inversionistas.....	23
4.1.3 El rol de las curvas de oferta y demanda sobre la volatilidad	28
4.1.4 El tipo de cambio real	29
4.1.5 Los costos de la industria.....	32
4.2 Modelos paramétricos de series de tiempo: ARIMAX y GARCH.....	34
4.3 Modelos no-paramétricos: Sistemas de Inferencia Difusa y Redes Neuronales Artificiales	37
4.3.1 Sistemas de Inferencia Difusa (FIS).....	37
4.3.2 Redes Neuronales Artificiales (ANN)	40
4.4 Algoritmos Genéticos (GAs).....	42
5. Metodología	46
5.1 Fuentes de datos y procesamiento.....	46
5.2 Funcionamiento del algoritmo genético y su contexto en el pronóstico	50
5.3 Funcionamiento adaptativo de los modelos	56
6. Resultados y discusión	65
7. Conclusiones	75
Referencias	78
Apéndice	87

ABSTRACT

This thesis studies volatility forecasting in monthly terms for the copper market, which is of practical interest for different participants such as producers, consumers, governments, and investors.

Using data since 1990 and the main drivers of this specific market, we propose a set of time series models, non-parametric models, and hybrid specifications of both. The adaptability characteristic of these models in exogenous variables, their configuration parameters and window size, simultaneously, are provided by a genetic algorithm in pursuit of achieving the best possible forecasts.

We examine out-of-sample performance based on *Heteroskedasticity-adjusted Mean Squared Error* (HMSE), and we test model superiority using the *Model Confidence Set* (MCS). The results show that making forecasts in an adaptive manner is crucial to obtaining robust and improved performances. The Adaptive-GARCH-FIS specification yields the best forecasting power.

RESUMEN EJECUTIVO

Este trabajo busca realizar pronósticos de volatilidad para el mercado del cobre en frecuencia mensual, lo cual es de interés práctico para diferentes participantes tales como productores, consumidores, gobiernos e inversionistas.

Teniendo en cuenta datos históricos desde 1990 y los principales determinantes de este mercado en específico, se propone un conjunto de modelos de series de tiempo, modelos no-paramétricos, y especificaciones híbridas de ambos. La capacidad de adaptación de estos modelos tanto en variables exógenas, parámetros de configuración y tamaño de ventana, simultáneamente, es entregada mediante un algoritmo genético con el propósito de alcanzar los mejores pronósticos posibles.

El desempeño fuera de muestra examinado se basa en el *Heteroskedasticity-adjusted Mean Squared Error* (HMSE), y se evalúa la superioridad de modelos usando el *Model Confidence Set* (MCS). Los resultados muestran que realizar pronósticos de una forma adaptativa es crucial para obtener desempeños robustos y mejorados. La especificación del tipo Adaptive-GARCH-FIS demuestra poseer el mayor poder de pronóstico.

1. INTRODUCCION Y DESCRIPCIÓN DE LA PROBLEMÁTICA

Los pronósticos económicos están presentes en todas las industrias y mercados, y son la variable principal de interés para realizar planificaciones futuras de empresas y países.

Como se afirma en Armstrong et. al (1986), los pronósticos cumplen con muchas necesidades y son utilizados para planificaciones tanto en el corto como en el largo plazo; ayudan además a tomar decisiones de producción, decisiones de personal, de finanzas y de marketing. Así, son fundamentales para todo participante de un mercado y necesarios para cumplir con sus objetivos en cada una de las industrias respectivas, incluyendo a la industria minera.

En las industrias referidas a minerales comercializados, muchas veces sus participantes tienen bajo poder de influencia sobre los cambios de precios de dichos minerales en el futuro. Esto hace particularmente interesante el hecho de pronosticar el comportamiento del precio del mineral en cuestión, donde se pueden utilizar distintos métodos, criterios, técnicas y teorías, e incluso combinación de estas en determinados casos. Por otro lado, la volatilidad de los precios de los metales hace del pronóstico un ejercicio exigente y una tarea bastante más compleja de lo que comúnmente se podría esperar. Como afirma Humphreys (2011), la volatilidad de los precios de los minerales hace que las evaluaciones futuras sobre el nivel esperado del precio a corto plazo sean algo desafiantes.

La necesidad de proyectar los precios en el futuro exige que los pronósticos tengan el mayor nivel de certeza posible. Tener en cuenta el riesgo asociado a la certeza de predicciones de retornos o precios, hace que estimaciones futuras de volatilidad sean una fuente importante para complementar tales estudios de pronósticos, por ser comúnmente

considerada una medida de nivel de riesgo y por las propiedades estadísticas que permiten relacionar retorno y desviación estándar¹.

La generación de pronósticos en la industria minera va ligada a todo horizonte de tiempo, ya sea en el corto, mediano o largo plazo. El horizonte temporal bajo el cual es estudiada la volatilidad del precio de commodities es muy importante para entender los diferentes alcances que pueden tener sus efectos para diversos participantes. Es así como, en cuanto al mediano plazo, los pronósticos del precio del mineral y de su comportamiento se requieren, por ejemplo, para evaluar factibilidad de nuevos proyectos de expansiones mineras. Por su parte, en el largo plazo se necesitan principalmente para las decisiones de construcción de nuevas operaciones y estudiar la sustentabilidad del negocio minero. En el caso del cobre, esta incertidumbre del precio es lo que tiene a las mayores empresas mineras en espera para autorizar nuevas inversiones (Lagos, 2013).

Con respecto al corto plazo, de acuerdo a los interesados en la volatilidad del precio mensual del cobre, se puede distinguir como un primer actor a las firmas productoras del commodity, a quienes se les dificulta realizar estimaciones realistas de su presupuesto mensual por no tener claridad sobre el comportamiento que tendrá el precio de venta de su producto. Esto ya que, habitualmente, los productores venden a sus clientes sobre la base de promedios mensuales de cotizaciones de bolsa (Radetzki, 2008). Además, desde un punto de vista operacional, el precio mensual del cobre es una variable de entrada crítica a la hora de determinar la política óptima de extracción del metal en aquellos planes mineros de

¹ Por ejemplo, bajo el supuesto de gaussianidad, se pueden utilizar bandas de predicción de la forma $[r_t - z_\alpha \hat{\sigma}_t, r_t + z_\alpha \hat{\sigma}_t]$, donde r_t corresponde al pronóstico del retorno del precio en el tiempo t , z_α corresponde al percentil $100(1-\alpha)$, Y $\hat{\sigma}_t$ a la desviación estándar pronosticada, como se muestra en Shumway y Stoffer (2011).

frecuencia mensual, determinando la mayor proporción del Net Present Value (NPV) percibido por el periodo siguiente. Así, tal volatilidad que este precio puede experimentar permite dimensionar las variaciones que dichos planes podrían sufrir. Por el lado de consumidores, se les obstaculiza alcanzar planeaciones estables al no contar con los posibles cambios que sufriría el precio del producto que desean comprar.

Del mismo modo, otro actor relevante son los gobiernos de naciones en cuyo terreno existe vasta extracción de minerales. Tal como indica David et al. (2009), los shocks de precios de *commodities* son una fuente clave de inestabilidad macroeconómica en países especializados en *commodities*, con reconocibles implicancias de esta sobre el crecimiento de largo plazo. Los royalties ad valorem², los cuales fluctúan según estos precios mensuales, determinarían la magnitud de ingresos que es capaz de percibir el fisco al fin de un periodo dado (Otto et al., 2006). En el contexto nacional, la actividad minera históricamente representa alrededor del 60% de las exportaciones totales. Esto ha implicado que el PIB chileno tenga comportamientos marcados respecto del precio del cobre. Es decir, hay consecuencias macroeconómicas importantes frente a una caída que experimente el precio de este metal, entre ellas el deterioro en la balanza comercial, y el impacto en la recaudación fiscal (Meller, 2013).

Y finalmente, por el lado de los inversores y especuladores también existen repercusiones. A mediados de la década de 1990 se ha observado un aumento del interés en los *commodities* como una clase de activos distinta para los inversores, particularmente

² Royalty sobre el valor. Este impuesto corresponde a un tipo de royalty que se paga independientemente que la mina genere utilidades o pierda dinero, a diferencia de los royalties basados en utilidades. Si bien el valor puede determinarse de varias maneras, una habitual es aquella basada en los ingresos brutos derivados de las ventas, las cuales como se comentó, tienden a concretarse bajo un precio mensual. Estas últimas, al igual que la producción, normalmente varían menos que las ganancias (Radetzki, 2008).

entre los fondos de cobertura, las mesas de negociación propietarias de los principales bancos de inversión y asesores de comercio de commodities (Humphreys, 2010). El caso del cobre no es la excepción, y comprende todas las condiciones para participar en lo que respecta a mercado de futuros³. De un modo más directo, y tal como se muestra en Hull (2015), el nivel de volatilidad es relevante, por ejemplo, para el cálculo de valorización de opciones.

A pesar de la importancia que presentan los pronósticos de precios de minerales, y dejando de lado la gran cantidad de tiempo y recursos que se puede llegar a invertir para efectuarlo, el verdadero valor del pronóstico está en la utilidad que genera según la precisión que demuestre, lo que generalmente no se toma muy en cuenta. Sin esta evaluación se hace imposible cuantificar su precisión y su desempeño, no registrando la verdadera utilidad ni el valor que cada pronóstico genera. Asimismo, se hace imposible también compararlos transversalmente y llegar a resultados que indique cual pronóstico supera al resto, conociendo de esa forma la habilidad relativa de cada una de las estimaciones.

El mercado del cobre, al igual que el de otros commodities, se caracteriza por grandes volatilidades y tendencias persistentes. Lo anterior, señala que el nivel de equilibrio del mercado es cambiante a lo largo del tiempo. Esta volatilidad ha hecho el tema de la predictibilidad de los precios un tema sumamente controversial, y ha despertado interés por determinar alguna habilidad de pronóstico que pueda reducir el riesgo de las decisiones contingentes a los commodities. Un pequeño aumento en la capacidad de predicción o

³Radetzki (2008) señala, entre las más importantes: existencia de muchos compradores y vendedores para proporcionar liquidez; commodity almacenable para permitir una relación entre precios spot y futuros; variabilidad considerable de los precios del activo.

certeza podría mejorar la toma de importantes decisiones dada la gran cantidad de dinero involucrado.

Adicionalmente con respecto a los mercados de commodities, tal como sugiere la teoría económica de mercado de minerales, hay muchos factores que pueden afectar la naturaleza de los fundamentos de mercado (y por tanto, el comportamiento de los precios) que pueden variar mucho de un *commodity* mineral a otro (Tilton y Guzman, 2016). Así, no es adecuado considerar a los commodities en conjunto como una única clase de activos, por lo que tener en cuenta los principales drivers detrás del comportamiento del precio en los mercados de modo específico, y las complejas dinámicas de los procesos de ajuste en estos, resulta indispensable para desarrollar estudios prometedores en materia de pronósticos.

En la consideración de un conjunto de factores explicativos es necesario examinar y hacer uso de cualquier teoría que postule relaciones que determinen la variable dependiente (Labys, 2006). Por tanto, en el caso del comportamiento del precio del cobre, resulta necesario considerar aquella teoría existente que se ha encargado de estudiarlo en el corto plazo. Adicionalmente, Marra (2015) señala que los factores exógenos pueden tener un impacto significativo y no-lineal sobre la volatilidad realizada. Así, tener en cuenta la presencia de variables exógenas y su interrelación en el corto plazo, comprende una fuente relevante para obtener un buen desempeño de los modelos a implementar.

En lo que concierne a técnicas y metodologías utilizadas en materia de pronóstico, modelos paramétricos de series de tiempo han permitido el alcance de importantes resultados por sí solos (como el caso de Dooley y Lenihan, 2005), así como también su

combinación o el uso individual de herramientas de lo que se conoce como *soft computing*, las cuales pertenecen a la categoría de modelos no-paramétricos de inteligencia artificial.

Las técnicas de *soft computing* ofrecen herramientas útiles para predecir entornos ruidosos como los mercados bursátiles, capturando su comportamiento no-lineal (Atsalakis y Valavanis, 2009a). Particularmente, en los últimos años, los sistemas adaptativos⁴ para los problemas de pronóstico han atraído una gran cantidad de atención, destacando las redes neuronales dinámicas (Ghiassi et al., 2005) y la evolución de modelos basados en reglas difusas (Angelov y Filev, 2004). En este contexto, los algoritmos genéticos han sido considerados una herramienta eficiente complementaria en la resolución de dichos problemas.

Este trabajo tiene como objeto obtener pronósticos de volatilidad del precio mensual del cobre, encontrando aquel modelo y especificación que permita alcanzar las mejores predicciones posibles dentro de un conjunto de modelos a proponer. En este, se toman en cuenta aquellos de series de tiempo (desde ahora, modelos paramétricos) del tipo ARMA y GARCH, modelos de *soft computing* (desde ahora, modelos no-paramétricos) tales como Redes Neuronales Artificiales (ANN) y Sistemas de Inferencia Difusa (FIS), y algunas combinaciones híbridas entre estos. Todos ellos teniendo en cuenta el rol que tienen el uso de variables exógenas rezagadas en materia de pronósticos, y todos en un contexto adaptativo. Respecto a lo último, parámetros inherentes a cada tipo de modelo, así como el tamaño de ventana⁵, y variables exógenas incluidas en el set de entrada, son seleccionados

⁴ Un sistema adaptativo, como su nombre lo indica, es uno que se adapta a los cambios en la dinámica del problema. La adaptación también puede incluir actualizaciones de parámetros y/o de estructura inherentes a los modelos comprendidos (Luna y Ballini, 2011).

⁵ Cantidad de observaciones que son consideradas, de forma longitudinal a lo largo del tiempo, para realizar el ajuste, estimación, o entrenamiento de los modelos que se alimentan de series de tiempo (es decir, tanto

bajo un contexto optimizador haciendo uso de un algoritmo genético (GA) que permitirá resolver problemas de gran dimensión que se generarán. El objetivo de predicción en la presente materia es prometedor en términos de factibilidad, al ser la volatilidad más fácil de predecir en comparación a los retornos por contar con características más adecuadas para ser modelada y pronosticada, como lo es, por ejemplo, el clustering (Marra, 2015).

Al menos para el caso del mercado del cobre, pareciera que este es el único estudio referente a pronóstico de volatilidad en un contexto de horizonte de tiempo mensual, a diferencia de aquellos existentes ya sea para el caso de precio o retornos de mercado spot del mismo. El logro del objetivo anterior conduce a que todos los participantes opten a toma de decisiones más informadas bajo incertidumbre.

Los pronósticos se efectúan del modo “one-step ahead out of sample”, donde las estimaciones de los parámetros de todos los modelos son actualizadas en cada paso del procedimiento. Tal actualización en las estimaciones permite captar la dinámica de variación en el tiempo inherente al comportamiento de volatilidad de este tipo de mercados, ya señalada en Pindyck (2004). La medida de volatilidad utilizada será ‘realized variance’ (varianza realizada).

Esta tesis se distribuye del siguiente modo: La Sección 2 da a conocer los objetivos del presente trabajo. Por su parte, la Sección 3 consiste en una revisión de la literatura pertinente asociada a trabajos pasados y similares con respecto al tema en cuestión. Luego, la Sección 4 consiste en un marco conceptual que repasa los fundamentos económicos detrás del comportamiento del precio del cobre en el corto plazo, así como las

modelos paramétricos como no paramétricos). Por tanto, es un valor que es posible de configurar en cualquier modelo seleccionado.

descripciones de modelos y técnicas de *soft computing* a utilizar más adelante.

Posteriormente, la Sección 4 detalla las fuentes de datos y procesamiento de estos para efecto de la metodología, junto a la contextualización del problema y cómo se aborda. A continuación, la Sección 5 da a conocer los resultados de los modelos propuestos en la sección previa y la discusión de estos. Finalmente, la Sección 6 presenta las conclusiones del estudio y posibles trabajos futuros a explorar.

2. OBJETIVOS

2.1 Objetivo General

Pronosticar la volatilidad del precio mensual del cobre, y encontrar el modelo de mejor desempeño para realizar aquello dentro un conjunto de modelos propuestos de tipo paramétrico, no-paramétrico y combinaciones híbridas entre estos.

2.2 Objetivos Específicos

- Evidenciar la importancia que tiene la capacidad de adaptación en los modelos involucrados a fin de obtener mejores pronósticos, tanto en el tamaño de ventana utilizado como en los parámetros inherentes a dichos modelos y las variables explicativas a considerar.
- Corroborar la mejora en el poder predictivo que puede entregar la consideración de técnicas de *soft computing* sobre los modelos paramétricos, mediante la combinación híbrida entre ambos.
- Comparar los diferentes modelos implementados según métricas de desempeño objetivas, y deducir las ventajas y/o desventajas que pueden manifestarse entre estos.
- Comprobar cómo el uso de algoritmos genéticos permite resolver problemas de gran dimensión de forma eficiente en contextos de aplicación práctica.
- Mostrar cómo la sinergia entre el uso de fundamentos teóricos asociados a la variable de interés, y el empleo de una metodología acorde, permite alcanzar importantes resultados en materia de pronóstico.

3. REVISIÓN DE LA LITERATURA

A pesar de la importancia práctica que tiene para los participantes mencionados el pronosticar dicha volatilidad, la investigación sobre la volatilidad de retornos de *commodities* y su riesgo ha sido menos abarcada que su contrapartida en precios y retornos de los mismos. Sin embargo, los estudios centrados en la volatilidad de *commodities* han ido ganando importancia debido a la creciente volatilidad y al papel cada vez más importante que desempeñan estos en los mercados de activos internacionales. (Chkili et al., 2014).

En este sentido, muchos estudios anteriores han demostrado que modelos econométricos complejos no necesariamente generan pronósticos más precisos que los modelos más simples (Makridakis et al., 1979). Esto ha incentivado a ciertos autores a estudiar variaciones del precio del *commodity* a partir de variaciones en alguna variable explicativa, mediante cálculos aproximados de elasticidad, como en Sánchez (2007). Si bien dicho enfoque permite comprender en forma aproximada la relación existente entre dos variables, no asegura causalidad y por tanto no tiene un valor predictivo. Más aun, manteniendo este objetivo de sólo estudiar relaciones, se han realizados análisis más complejos donde también se incluye la influencia de variables a diferentes escalas de tiempo. En este sentido, Kristjanpoller y Sierra (2014) se restringen a determinar relaciones entre el tipo de cambio, precio del cobre, y el mercado accionario chileno a distintas escalas de tiempo; esto lo llevan a cabo mediante técnicas más completas como el enfoque wavelet y la aplicación de causalidad de Granger, obteniendo conclusiones relevantes para el ámbito de la economía de Chile.

De un modo general, en términos de modelación y valorización, varios estudios se pueden mencionar independientemente de la categoría específica del commodity, la metodología empleada, y el tipo de mercado ya sea spot o de derivados (e.g., Dahl y Iglesias, 2009; Trolle y Schwartz, 2009; Choi y Hammoudeh, 2010; Larsson y Nossman, 2011; Arouri et al., 2012a; Vivian y Wohar, 2012; Creti et al., 2013; Thuraissamy et al., 2013; Cortazar et al. 2015; Cortazar et al. 2016; Bentes, 2016; Kang et al., 2017). En el contexto de realizar pronósticos, la investigación también es diversa en ese sentido (ver por ejemplo: Trück y Liang, 2012; Arouri et al., 2012b; Kang y Yoon, 2013; Bentes, 2015; Triantafyllou, 2015; Basher et al, 2015; Degiannakis y Potamia, 2017; Bollerslev et al., 2017). Sin embargo, todo lo anterior se enmarca en un análisis basado en frecuencia intradiaria, diaria, o semanal.

En el contexto de técnicas y metodologías utilizadas en materia de pronóstico, en lo que respecta ya sea mercados de commodities o mercados de acciones, se pueden destacar estudios como los de Olson y Mossman (2003), Hamid y Iqbal (2004) y Sánchez et al. (2015) haciendo uso de Artificial Neural Networks (ANN); Tay y Cao (2002) y Huang y Tsai (2009) en el caso de Support Vector Machines (SVR); Sheta (2006), Chang y Liu (2008), Atsalakis y Valavanis (2009b) y Boyacioglu y Avci (2010), que emplean Fuzzy Inference System (FIS) o Adaptive Neuro Fuzzy Inference System (ANFIS).

En lo que respecta a configuraciones híbridas entre modelos paramétricos y no-paramétricos se puede mencionar a Donaldson y Kamstra (1997) que aplican un modelo de red neuronal en combinación con un modelo GARCH para capturar los efectos de la volatilidad en los retornos de las acciones, o también Fiordaliso (1998) que emplea un FIS para combinar un conjunto de pronósticos individuales. Por su parte, Guresen et al. (2011)

usa modelos híbridos GARCH-ANN para predecir los tipos de cambio diarios de NASDAQ, obteniendo resultados que no muestran mejoras significativas para el modelo híbrido en comparación con los modelos tradicionales. Hajizadeh et al. (2012) utilizan dos modelos híbridos de redes neuronales con el objetivo de mejorar el poder predictivo de los modelos de la familia GARCH, utilizando variables explicativas para predecir la volatilidad de los retornos del S&P500. Kristjanpoller y Minutolo (2016), y Kristjanpoller y Hernandez (2017) combinan ANN y modelos tipo GARCH para predecir volatilidad en mercados spot de *commodities* tales como de petróleo, oro, plata y cobre, en frecuencia diaria; de modo similar, Zhang (2003) emplea modelos híbridos del tipo ARIMA-ANN. Finalmente, cabe mencionar los estudios de Panapakidis y Athanasios (2017) que combina transformación wavelet, ANFIS, algoritmos genéticos (GAs) y ANN para el caso de gas natural; y Chang y Fuan (2008) que usan un modelo híbrido del tipo Wavelet-FIS.

En el marco de modelos adaptativos y su utilización, donde los algoritmos genéticos han sido considerados una herramienta eficiente, destacan los trabajos de Goonatilake et al. (1995) que proponen el uso de sistemas difusos para tomar decisiones de *trading*, y usan un GA para encontrar la configuración óptima de conjuntos difusos. Wagner et al. (2007) incluye un enfoque adaptable de ventana, donde se elige el mejor tamaño de esta para que los modelos de series de tiempo alcancen la mejor predicción de los datos actuales, aplicándolos para estimar el Producto Interno Bruto de EE. UU. Por su parte, Hung (2011a) y Hung (2011b) combina los sistemas difusos y el modelo GARCH para pronosticar de manera adaptativa la volatilidad del mercado de valores, basado en GA y en la optimización de enjambre de partículas (PSO) respectivamente, para estimar parámetros propios de los modelos FIS y GARCH. Wei (2013) emplea GA para refinar los pesos de las

reglas conjuntas en un modelo ANFIS para pronosticar el índice de acciones de Taiwán. Finalmente, Singh y Mohanty (2017) ocupan GA junto con ajuste difuso para entrenar los parámetros libres en su modelo de neuronas generalizadas adaptativas, bajo el contexto de pronóstico de precios de la electricidad.

A pesar de la vasta literatura mencionada anteriormente, en términos de analizar volatilidad mensual, no son muchos los estudios existentes para mercados spot de *commodities* minerales (especialmente para el cobre), a pesar de la importancia práctica que esto posee. Entre los pocos desarrollados sólo en términos de modelación, se puede señalar que varía también la metodología empleada (e.g. Davidson et al., 1998; Brunetti y Gilbert, 1995; Tully y Lucey, 2007; Labys et al., 2000). En el caso de pronósticos en este mismo horizonte temporal, destacan los trabajos de Labys, 2006; García-Cicco y Montero, 2011; Kriechbaumer et al., 2014 para el caso del cobre, existiendo también para el caso de otros *commodities* minerales (ver por ejemplo: Dooley y Lenihan, 2005; Chen et al., 2016).

No obstante todo lo anterior, en términos de predicción de volatilidad, no existen aquellos que apunten a pronósticos mensuales en mercados spot (al menos para nuestro conocimiento).

4. MARCO TEÓRICO

La combinación de un amplio sustento teórico, y el empleo de herramientas de modelación con características que se adecúen al comportamiento del fenómeno en estudio, posibilitaría el alcanzar los objetivos propuestos. Dicho esto, la presente sección se divide en las siguientes partes: fundamentos económicos relacionados al comportamiento del precio del cobre y los cambios en este, en el corto plazo; descripción de los modelos matemáticos paramétricos y no-paramétricos que se utilizarán para pronosticar, ya sea de forma individual o híbrida; y descripción de los algoritmos genéticos, como herramienta complementaria de gran relevancia para los modelos matemáticos anteriores.

4.1 Fundamentos Económicos

Un paso esencial para modelar mercados de minerales (y por consecuencia tener mayor posibilidad de alcanzar pronósticos certeros) es que el modelo cuente con componentes que reflejen varios aspectos de demanda, oferta, comercio y determinación del precio (Labys, 1999). Dado lo anterior, el enfoque está en centrarse en aquellos factores que determinan el precio y que determinan los cambios de este en el corto plazo.

En el contexto de variables explicativas en un ambiente de condiciones de equilibrio⁶, la idea de una única (o singular) variable exógena siendo causa de los valores es engañosa dado que todas las otras variables exógenas también juegan un rol (Boland,

⁶ Si bien la idea de un estado de equilibrio es ampliamente aceptada en lo que respecta a mercados, algunos autores afirman que, en lugar de preocuparse por el equilibrio alcanzado, deberíamos centrarnos más en el fenómeno del desequilibrio. Si bien el precio puede cambiar en la dirección correcta, puede ser que el equilibrio nunca se alcance en tiempo real (Boland, 2017).

2017). Para el caso particular del mercado del cobre, la teoría económica por detrás es relativamente numerosa.

4.1.1 Inventarios en Bolsas de Metales

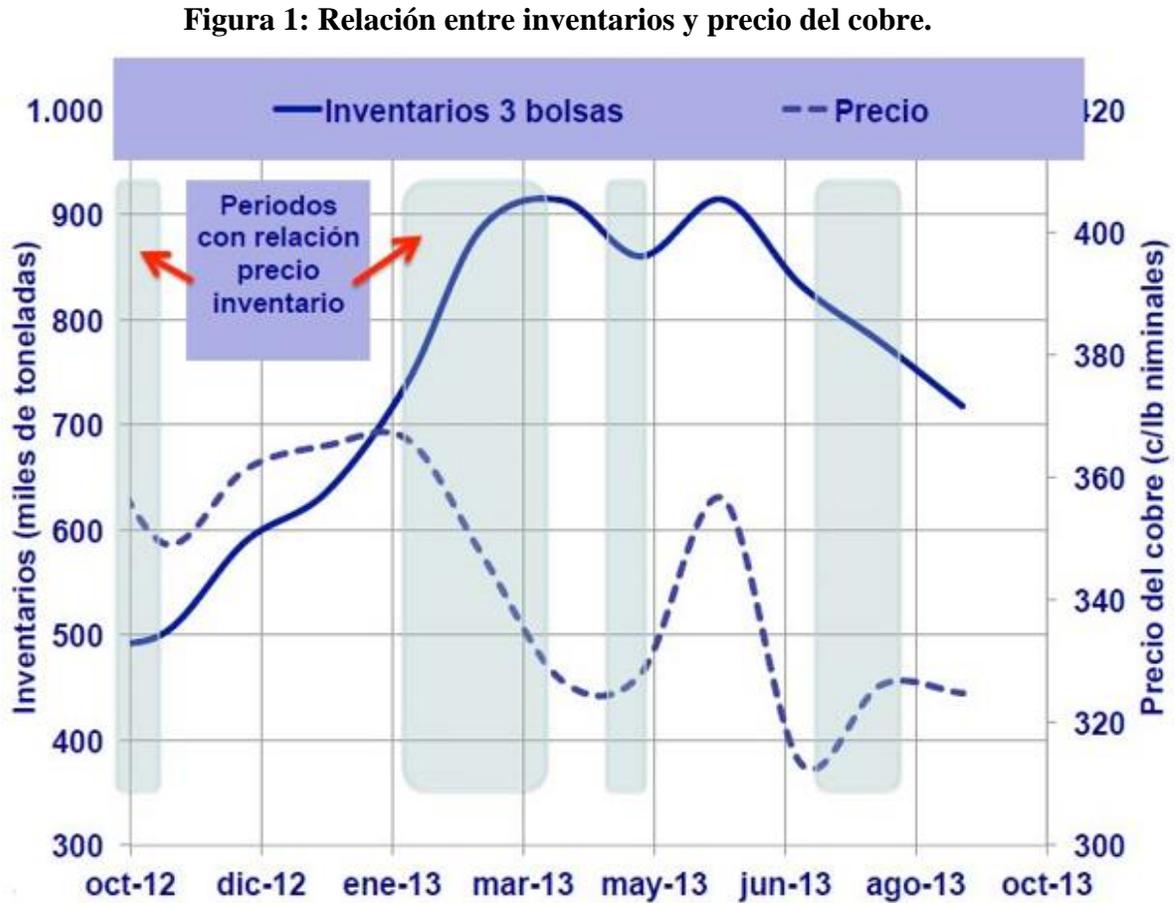
En primer lugar, ampliamente es aceptada la existencia de una relación inversa entre el precio del cobre y los inventarios del mismo en las bolsas de metales. Brunetti y Gilbert (1995) ofrecen una explicación basándose en cómo los *fundamentals* determinan la volatilidad de precios de metales, presentando el rol que tienen los inventarios de bolsa. La idea central de los inventarios de cobre como determinante del precio es que, si la demanda de cobre aumenta más rápido que la oferta de cobre, entonces los inventarios disminuyen, y el precio debiera subir (y viceversa).

Entre estudios destacados se reconoce también el realizado por Guzmán (2004), el cual relaciona el precio del cobre con el stock existente (tanto en las bolsas de metales como en las bodegas de los consumidores), así como con el inventario manejado por empresas productoras. Una aplicación directa la atribuye a determinar efectos de recortes de producción llevados a cabo durante el 2003, en el precio evidenciado por el cobre en el mismo año, obteniendo también un pronóstico de frecuencia trimestral. Por consecuente, este trabajo aporta al estudio de la relación precio-stock en el corto plazo, y contribuyendo con un gran sustento teórico para estudios futuros.

Sin embargo, estudios como Pincheira (2000) también demuestran que dicha relación inversa se desvanece en ciertos periodos, donde ambas variables se muestran positivamente correlacionadas. Esto último se debería a la participación activa de especuladores y grandes inversionistas institucionales a nivel global en el mercado no físico

en ciertos periodos (Lagos, 2013), como se puede evidenciar en las zonas no oscuras de la

Fig. 1:



Fuente: Curso Gestión del negocio minero, 2015, Prof Gustavo Lagos, Clase Ejecutiva UC.

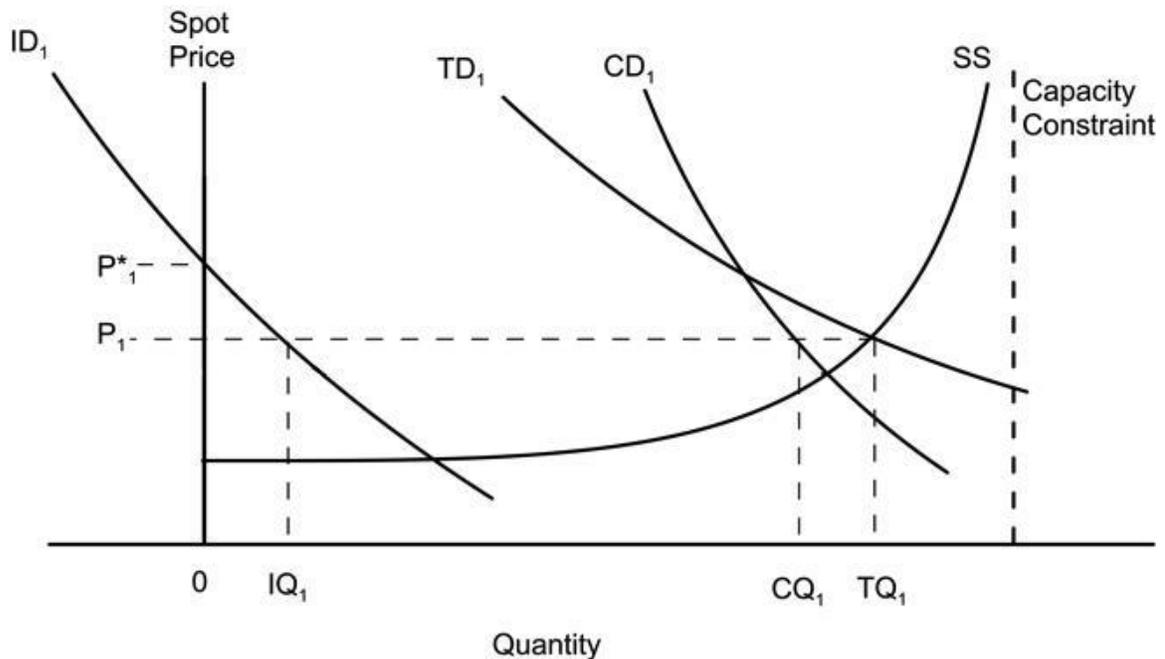
Entre las explicaciones para la débil o nula relación existente entre el precio del cobre y los inventarios en ciertos periodos, se encuentra la incertidumbre económica. Cuando esta aumenta, los inversionistas retiran sus fondos de los *commodities* y se refugian en instrumentos de renta fija. Justamente esto fue lo que ocurrió en 2011 con la crisis europea cuando el precio del cobre bajó cerca de un 30% sin que hubiese aumento de inventarios.

4.1.2 La actividad especulativa y la participación de inversionistas⁷

La influencia que tendrían especuladores e inversionistas ya se ha discutido previamente en algunos estudios (e.g. Jaramillo y Selaive, 2006; Radetzki, 2008; Humphreys, 2015; Brunetti et al., 2016). En el caso particular del cobre, Tilton et al. (2011) profundiza en dicha relación de una manera económica y gráfica.

La **Fig. 2a** muestra que el precio de mercado P_1 – el precio *spot*, por ejemplo – de un *commodity* mineral depende de su curva de oferta de mercado (SS) y su curva de demanda total de mercado (TD_1). Esta última, a su vez, depende de la curva de demanda de consumo (CD_1) y la curva de demanda de inversionistas (ID_1).

Figura 2a. Periodo 1: Equilibrio de mercado inicial.



Fuente: Minerals Economics and Policy. John Tilton y Juan Ignacio Guzmán (2016)

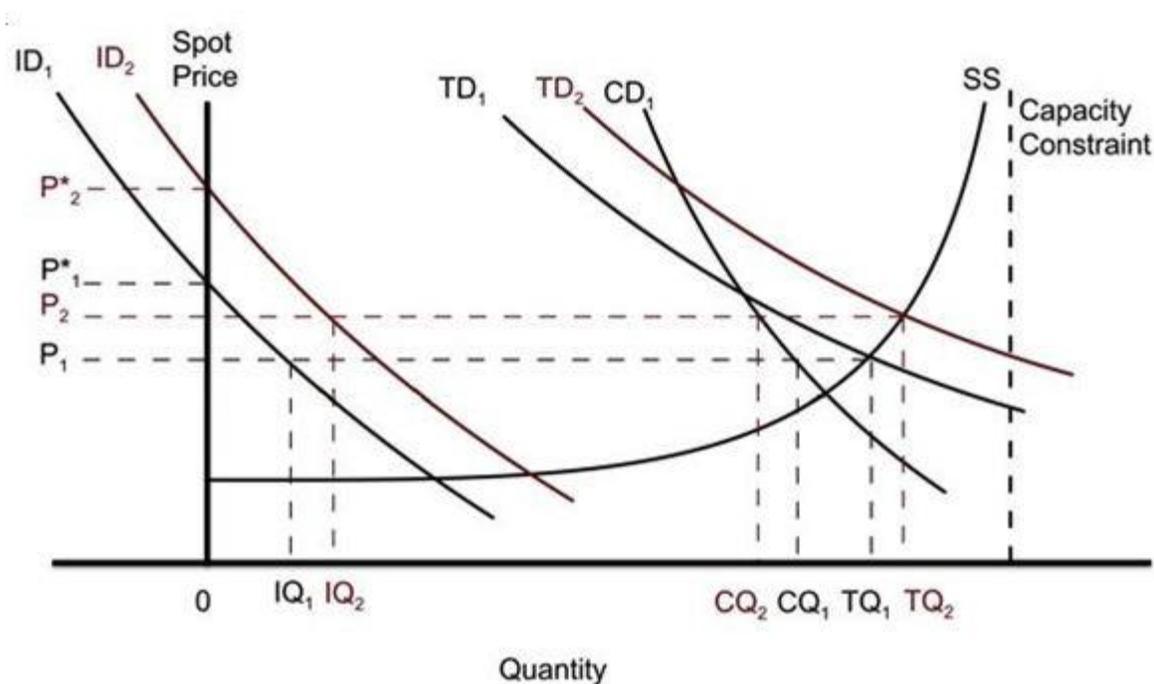
⁷ Esta sección se basa fuertemente en Tilton y Guzmán (2016).

Al precio de mercado P_1 , consumidores demandarán una cantidad igual a OCQ_1 , y los inversionistas una cantidad OIQ_1 (la cual agregarán a sus inventarios existentes). La demanda total es la suma de las dos, o también, OTQ_1 . La demanda de inversionistas, vale la pena notar, puede ser negativa. Mientras el precio aumente por sobre P^*_1 , los inversionistas en equilibrio empiezan a agotar sus inventarios y aumentar el suministro.

Como se muestra, ambas curvas de demanda de consumidores e inversores tienen pendientes negativas. Esto parece razonable para la curva de demanda de consumidores, pero la curva de demanda de inversores podría no tener una pendiente negativa sobre su rango entero. Esto ya que la demanda de inversores, entre otras cosas, depende tanto del precio en la actualidad y del precio que los inversores esperan que prevalezca en el futuro. Cuando el precio actual aumenta, esto puede causar expectativas acerca de los precios en el futuro que aumenten también. De ser así, un aumento en el precio actual puede causar que incrementemente la demanda de inversores. Sin embargo, mientras los precios continúen aumentando, a un cierto punto, la pendiente de la curva de demanda de inversores se volverá negativa. Esto dado que a mayores y mayores precios en la actualidad, se vuelve incrementadamente difícil para los inversores de mantener la expectativa de que los precios en el futuro continúen aumentando. Así, para algún precio, la pendiente de la curva de demanda de inversores se volverá negativa como lo muestran **Fig 2a** y **Fig 2b**.

Los precios de commodities pueden aumentar en el tiempo como resultado de un desplazamiento hacia adentro de la curva de oferta de mercado, un desplazamiento hacia afuera de la curva de demanda total (debido a un desplazamiento hacia afuera tanto en la curva de demanda del consumidor como del inversor), o ambas. Como se muestra en la **Fig. 2b**, un desplazamiento hacia afuera en la curva total de demanda, causada por un

Figura 2b. Periodo 2: Equilibrio de mercado después de un incremento en la demanda de inversores.



Fuente: Minerals Economics and Policy (Tilton y Guzmán, 2016).

desplazamiento hacia afuera en la curva de inversores, lleva a un incremento en el precio de mercado. Una figura similar podría mostrar el mismo aumento en el precio causado por un desplazamiento comparable en la curva de demanda de consumidores en vez de la curva de inversores. Lo que no es claro es hasta qué punto los cambios en la curva de demanda total que están detrás de las alzas en los precios han sido impulsados por cambios en la curva de demanda de inversores, y hasta qué punto lo son por aquellos de la curva de demanda de consumidores. Los estudios empíricos realizados para resolver este asunto han producido resultados que dan soporte a ambas posturas de forma exclusiva, conduciendo así a conclusiones conflictivas, en general, para esta situación.

La influencia de los inversores descrita, hasta ahora discutida en esta sección, se ha centrado cuando estos como el precio del commodity se refieren a un mismo mercado exclusivo (por ejemplo: el mercado spot, el mercado de futuros de 30 días, el mercado de futuros de 90 días). Pero otro asunto relevante comprende la influencia que tienen los inversores de los mercados de futuros sobre el precio spot.

Dado que la mayoría de inversores tradicionalmente no han querido recibir existencias físicas, estos han invertido en gran medida adquiriendo contratos de futuros y luego vendiendo estos contratos a una fecha de maduración (o antes de esta). La forma en que esas compras afectan el precio spot depende de la relación entre este último y el precio de futuros.

Cuando el precio por un commodity es menor que su precio futuro, el mercado se dice estar en *contango*. Cuando lo opuesto es el caso, esto es, cuando el precio spot es mayor que su precio futuro, el mercado está en *backwardation*. Los mercados en contango pueden separarse más aún en aquellos en *strong contango* (donde el diferencial entre precio spot y futuro es suficiente para cubrir los costos de comprar y almacenar existencias físicas, o también llamado, *convenience yield*), y aquellos en *weak contango* (donde el diferencial no es suficiente para cubrir tales costos).

Ahora, cuando los mercados están en strong contango, cualquier incremento en la demanda de inversores que aumente el precio futuro de un commodity debiera tener un efecto inmediato y comparable en su precio spot. Esto dado que un incremento en los precios futuros incentivará a los inversores a comprar en el mercado spot y, al mismo

tiempo, vender en corto en el mercado de futuros. Luego estos almacenan los inventarios adquiridos para cubrir su compromiso de entregar el producto en el futuro.

Esas transacciones causarían que el precio spot se incremente y el precio futuro disminuya hasta que el diferencial entre ambos ya no incentive más arbitraje intertemporal. Dada esta situación, los precios spot y futuros debieran estar cercanamente correlacionados, y por tanto la demanda de inversores (o cualquier cosa que provoque aumento del precio futuro) debiese impactar similarmente al precio *spot*.

Sin embargo, cuando los mercados están en *weak contango* o *backwardation*, los inversores no tienen un incentivo para comprar en el mercado spot y vender en corto en el mercado de futuros cuando la demanda de inversores incrementa el precio en el mercado de futuros. Como resultado, los precios en el mercado spot probablemente estén menos correlacionados con los movimientos de los precios futuros. Bajo estas condiciones de mercado la demanda de inversores en el mercado de futuros puede tener baja influencia en el precio spot. En resumen, la respuesta a la pregunta de si los inversores en el mercado de futuros afectan el mercado spot dependerá de si los mercados *spot* y de futuros, en el instante, están en *strong contango*⁸. Finalmente, tal como indican Tilton y Guzmán (2016), entre otros autores, se cree que los inversores y los especuladores acentúan la volatilidad de los precios en las tendencias a corto y largo plazo.

⁸ Una investigación reciente (Gulley y Tilton, 2014) estima las correlaciones entre los precios *spot* y futuro del cobre de la Bolsa de Metales de Londres durante periodos de *strong contango*, *weak contango* y *backwardation*. Si bien se descubre que tales correlaciones son mayores durante periodos de *strong contango*, estas además son bastante altas durante periodos de *weak contango* y *backwardation*.

4.1.3 El rol de las curvas de oferta y demanda sobre la volatilidad

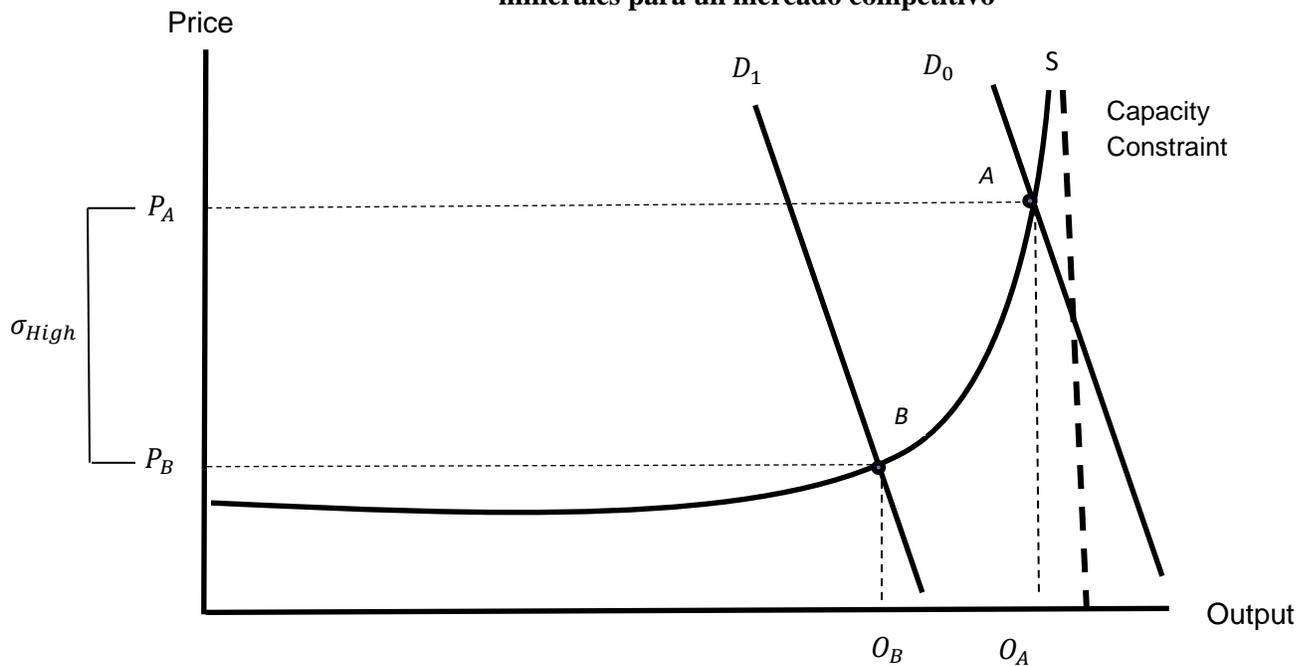
Por otro lado, la teoría tradicional de la economía de minerales establece que la volatilidad de corto plazo de los commodities se debe principalmente a variaciones de la demanda, dado que la oferta no tiene capacidad de ajuste en ese periodo, como señala Radetzki (2008) or Tilton y Guzmán (2016). Dichas variaciones en la demanda se deben a las condiciones del ciclo económico que imperen en el mercado en ese momento, de modo que la actividad económica en el corto plazo resulta relevante de considerar.

La demanda en el corto plazo, en ese sentido, es muy receptiva o elástica con respecto al ingreso. A su vez, tal como suponen los economistas de minerales, una vez que la capacidad total de la industria esta próxima de ser alcanzada, la curva de oferta no es muy receptiva a los cambios en el precio en el corto plazo. Esto se muestra en la **Fig. 3** para un mercado competitivo⁹.

En la **Fig. 3** la curva D_0 representa la demanda cuando la economía está en auge, y la curva D_1 en el punto medio del ciclo económico. Los precios de commodities minerales vendidos en mercados competitivos podrían variar bastante a lo largo del ciclo económico. La producción que los productores suministran al mercado también cambia marcadamente, desde una gran cantidad Q_A a una baja Q_B . Como resultado, grandes cambios en el precio desde P_A a P_B son necesarios para mantener tanto la curva de oferta y demanda en balance sobre el ciclo económico a medida que la curva de demanda se desplaza de un lado a otro.

⁹ En un mercado de productores, la volatilidad del precio no debiera ser tan significativa como en mercados competitivos (Tilton y Guzmán, 2016).

Figura 3. Explicación para la volatilidad del precio bajo la visión de la economía de minerales para un mercado competitivo



Fuente: Minerals Economics and Policy (Tilton y Guzmán, 2016).

Así, este comportamiento de los fundamentos de mercado genera una transición desde un equilibrio inicial A a un estado de equilibrio B.

En este sentido, en que la demanda resulta protagonista en la volatilidad de los precios, Radetzki (2008) señala que en los últimos 20 años hubo un cambio del centro de gravedad de la economía mundial y del consumo de *commodities* primarios desde las prósperas naciones de la OCDE hacia la populosa Asia en desarrollo.

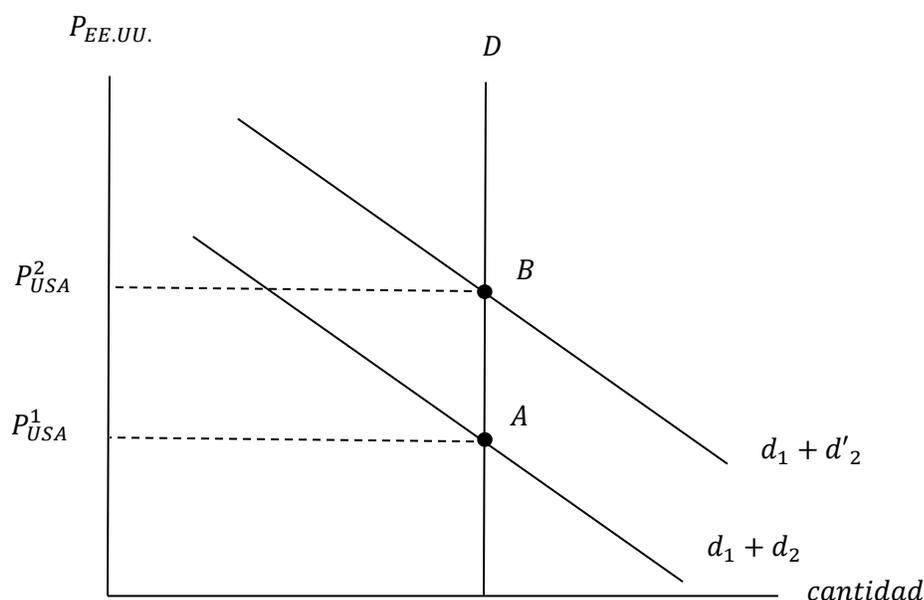
4.1.4 El Tipo de cambio real

En última instancia, el tipo de cambio real es otro de los determinantes clásicos utilizados en la determinación del precio del cobre. Un aumento de esta variable corresponde a una apreciación real del dólar respecto de una canasta de monedas. De

acuerdo a De Gregorio et al. (2005), tal apreciación real del dólar estadounidense genera una disminución del precio del cobre, y viceversa.

La **Fig. 4** profundiza la idea anterior. Supóngase que se comienza desde una situación de equilibrio, *A*, donde la demanda mundial es igual a la oferta. Si el tipo de cambio real de EE.UU. se deprecia, para un nivel de precio dado del *commodity* P_{USA}^1 , la

Figura 4. Precio de Equilibrio y Depreciación Real



Fuente: Elaboración Propia, basada en Fluctuaciones del dólar, precio del cobre, y términos de intercambio (De Gregorio et al., 2005).

demanda de este país no cambia. Sin embargo a ese P_{USA}^1 , el precio relativo del *commodity* en el resto de los países (los cuales conforman la canasta de monedas) caerá, con lo cual la demanda en estos aumentará de d_2 a d'_2 , y el equilibrio se alcanzará en el punto *B* con un consecuente aumento del precio P_{USA}^1 a P_{USA}^2 . Así, el aumento de la demanda del resto de

los países conlleva a un aumento del precio del *commodity* en términos de la moneda en la cual se transa (en este caso, dólares).

Lo anterior se refiere al efecto que tiene el tipo de cambio real sobre la demanda (y luego, por consecuencia, en el precio del metal). No obstante, la teoría tradicional de la economía de minerales afirma que la demanda en el corto plazo es inelástica con respecto al precio (sobretudo en aquellos con demanda derivada), y por tanto, cuestiona en cierto grado lo establecido en los párrafos anteriores (para más detalle, revisar Tilton y Guzmán, 2016). Si bien el grado del efecto que pueden tener los cambios del precio sobre la demanda es discutible, al menos la existencia es posible de afirmar, y por tanto es necesario de considerar.

Por otro lado, y de un modo menos discutible en la literatura, el tipo de cambio puede tener una gran influencia sobre la oferta de un *commodity* mineral. Lo anterior se sustenta al recordar la curva de *cash cost* (C1), la cual se aproxima a lo que los economistas llaman la curva de costo variable promedio, y que a menudo refleja la curva de oferta de mercado en el corto plazo en los mercados de minerales¹⁰. Luego, una depreciación del dólar respecto a la moneda local generaría una reducción en los insumos valorizados en dólares necesarios para la producción de cobre (por ejemplo, insumos energéticos), llevando a una alteración en la curva de oferta de corto plazo, y por ende, en el precio del *commodity*. Si bien tal movimiento en la curva de costos de corto plazo es cuestionable

¹⁰ No obstante, existen condiciones restrictivas que deben cumplirse para suponer lo anterior. Algunas de estas son: mercado competitivo (en el sentido en que ningún recorte de producción por parte de algún productor altera el precio de mercado); los costos marginales de corto plazo de las minas permanecen constantes mientras se expande la producción hasta que la capacidad es alcanzada; las compañías mientras cierran sus minas cuando el precio está por debajo de su *cash cost* y reabren cuando los precios están por encima de este.

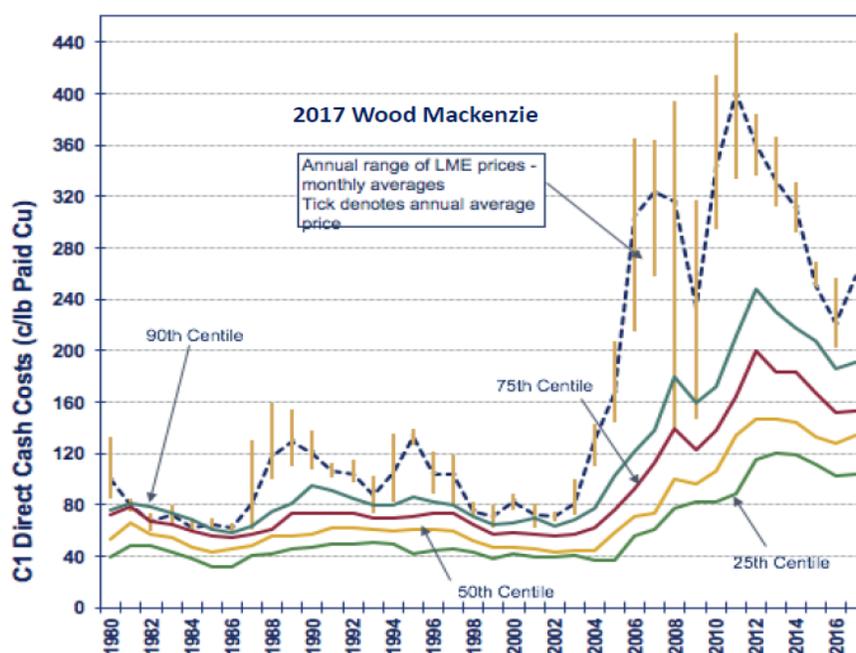
según lo establecido por la teoría de economía de minerales, tal hecho es posible también al existir una dependencia intrínseca entre los desplazamientos de las curvas de los fundamentos de mercado (para profundizar en esto, revisar García y Guzmán, 2017).

De este modo, un empeoramiento o mejoramiento significativo de la economía de EE.UU. resulta determinante en el precio del cobre al estar asociada a un fortalecimiento o debilitamiento del dólar. Finalmente, existen otros estudios que también tratan la relación entre el tipo de cambio y precios de commodities (ver por ejemplo: Ridler Y Yandle, 1972; Dornbusch et al, 1985; Gilbert, 1989; Reinhart, 1991).

4.1.5 Los costos de la industria

Si bien en la sub sección anterior se ha expuesto la influencia de los costos sobre el precio mediante las variaciones del tipo de cambio, existen motivos adicionales para relacionar ambas variables. La **Fig. 5** ilustra una curva de costos propia de la industria del

Figura 5. Curva de costos C1 para la industria del cobre.



Fuente: Wood Mackenzie

cobre, la cual se puede apreciar para los diferentes cuartiles.

Ya a simple vista se aprecia un comportamiento similar (aunque no necesariamente simultáneo, más bien retardado) entre ambas variables. En el corto plazo, los movimientos de la curva de oferta pueden depender de aquellos de la curva de demanda dado que los productores, en su propósito de maximizar ganancias, alteran su comportamiento más allá de sólo cambiar la cantidad que suministran al mercado. En este horizonte temporal, existen otras razones tales como cierre de minas o cambios en planes mineros que pueden generar desplazamientos en la curva de oferta. Incluso, proyectos de tipo *redfield*¹¹ se implementan en el corto plazo, conduciendo a cambios del programa de producción en tal periodo, y por ende, contribuyendo a movimientos en la curva de oferta.

De un punto de vista más sólido, la relación entre precios y costos también ha sido discutida en estudios académicos como en Tilton (2014), la cual se basa en las relaciones de causalidad entre precios y productividad. Cuando el precio de un mercado de minerales está deprimido, lo que sucede durante un ciclo bajo, la industria se encuentra operando en el segmento plano de la curva de oferta de corto plazo. Así, una caída en el precio (por ejemplo, debido a una recesión en la economía) causa que la productividad laboral se incremente, conduciendo a una reducción en los costos de producción, y por tanto, un desplazamiento hacia abajo en la curva de oferta de corto plazo. Esto en definitiva conlleva a que el precio de mercado decline más aún.

De modo contrario, durante un ciclo alto, la dirección de la causalidad entre precio y productividad opera en modo contrario. Si el precio está alto, la renta obtenida aumenta,

¹¹ Estos proyectos apuntan a iniciativas que buscan optimizaciones en la operación actual de la mina (típicamente involucrando inversiones marginales para eliminar cuellos de botella).

conduciendo a descuidos en la productividad laboral. Esto por consecuencia conlleva a que la curva de oferta de corto plazo se desplace hacia arriba, y luego, los efectos posteriormente en el precio de mercado también existen en este caso.

Finalmente, resulta necesario reconocer también la existencia de relaciones dinámicas y complejas entre los factores mencionados, así también el cómo estos determinan los cambios en los precios. La oferta y la demanda pueden no estar siempre en equilibrio, y algunos mercados pueden no equilibrarse rápidamente cuando las condiciones cambian repentinamente (Pindyck y Rubinfeld, 2017), lo cual implica que ajustes dinámicos a través del tiempo podrían determinar cambios en los precios. Respecto a lo anterior, García y Guzmán (2017) profundizan en la explicación de tales dinámicas otorgándole un rol clave al concepto de velocidad de reversión de los mercados de minerales, y la relación intrínseca que posee con la volatilidad de estos. Si bien profundizar en estas dinámicas no es el objeto de estudio¹² en esta tesis, podrían influir (ya sea de modo positivo o negativo) en el objetivo de realizar pronósticos.

Lo que es sí cierto, es que para efectos de este estudio, las variables mencionadas se consideraran como aquellos determinantes principales del precio del cobre y de sus cambios.

4.2 Modelos paramétricos de series de tiempo: ARIMAX y GARCH

Los modelos Autoregresivos Integrados de Media Móvil (ARIMA) son particularmente adecuados cuando se trata de la predicción de variables que se observan a

¹² De hecho, desde un punto de vista meramente económico, Boland (2017) indica que quienes construyen modelos de equilibrio deben lidiar con qué tan rápido y en qué dirección cambia el precio, refiriéndose en este contexto a los conceptos de velocidad y de dirección de ajuste.

corto plazo, es decir, trimestralmente o mensualmente (Labys, 2006), lo cual justifica su uso en el contexto de este estudio, además de ser ampliamente utilizados como *benchmark*. El uso de modelos ARIMA tiene ciertas limitaciones dado que sólo se enfoca en la tendencia durante el tiempo y la utiliza para extrapolar en el futuro sin la consideración de factores externos que pueden afectar la fluctuación de precios en el tiempo. Es por esto que su extensión, la cual permite la adición de una o más variables exógenas (refiriéndose como ARIMAX), ha sido capaz de mejorar el desempeño de pronósticos.

La parte autoregresiva del modelo comprende rezagos de la misma variable dependiente, mientras que la parte de media móvil considera valores rezagados del error de pronóstico. Por su parte, las variables exógenas comprenden rezagos de los factores externos para modelar la variable dependiente. Si bien la definición convencional de los modelos ARIMA (p,q) se refiere a los parámetros p y q como órdenes máximos de los rezagos (empezando desde el 1 hasta p o q, respectivamente), esto no es necesario, pudiéndose estimar modelos de este tipo considerando sólo ciertos rezagos de forma selectiva. Esto da origen a diferentes combinaciones posibles para especificar este tipo de modelo, las cuales aumentan en gran medida si también se meditan aquellos rezagos de variables exógenas de este mismo modo. Tal idea de considerar sólo determinados rezagos de modo independiente permite captar posibles efectos estacionales.

La varianza condicional de las innovaciones es típicamente supuesta como constante en el modelo ARIMA o ARIMAX (lo que se conoce como homocedasticidad). Sin embargo, tal idea es inapropiada en aquellas series de tiempo, habitualmente económicas o financieras, cuyos periodos de alta volatilidad son seguidos por otros de relativa estabilidad, o viceversa. De este modo, con el objeto de considerar estos cambios

en la varianza condicional (fenómeno conocido como heteroscedasticidad), es que son introducidos los modelos Autoregresivos de Heteroscedasticidad Condicionada (ARCH) y su generalización (GARCH) por Engle (1982) y Bollerslev (1986), los cuales introducen la volatilidad condicional pasada como variable explicativa para predecir volatilidad, en adición a noticias de volatilidad de periodos anteriores; y desde entonces se ha convertido en el modelo heteroscedástico autorregresivo utilizado más ampliamente (Labys, 2006). Su origen se encuentra en el intento por obtener una predicción adecuada para la inflación en el Reino Unido, sujeta a fuerte volatilidad y con periodos de especial calma o de especial agitación. Adicionalmente, estudios como Bentes (2015b) han mostrado que la volatilidad pronosticada por modelos GARCH supera a aquellos de volatilidad implícita al producir pronósticos fuera de muestra.

De un modo similar al caso de modelos ARIMA, los modelos GARCH tienen su extensión en el mismo sentido. Tanto la ecuación de media como de varianza pueden ser modeladas como función de los valores pasados de las variables dependiente y de aquellas independientes o exógenas. En términos matemáticos, un modelo GARCH (p,q) con variables explicativas es de la forma:

$$r_t = \alpha + \sum_{i=1}^I \beta_i r_{t-i} + \sum_{j=1}^J \gamma_j \varepsilon_{t-j} + \sum_{k=1}^K \delta_k x_{t-k} + \varepsilon_t \quad (1a)$$

$$\varepsilon_t \sim N(0, \sigma_t) \quad (1b)$$

$$\sigma_t^2 = c + \sum_{p=1}^P \lambda_p \varepsilon_{t-p}^2 + \sum_{q=1}^Q \theta_q \sigma_{t-q}^2 + \sum_{h=1}^H \phi_h y_{t-h} \quad (1c)$$

donde x_{t-k} denota el rezago k de la variable exógena x en la ecuación de la media, y y_{t-h} denota el rezago h de la variable exógena y en la ecuación de la varianza. Notar que, al igual que en los modelos ARMA, también es posible de agregar términos de media móvil en la ecuación de la media, como muestra la Eq. (1a). Además, los modelos GARCH también permiten la especificación selectiva y singular de determinados rezagos en sus variables explicativas, para propósitos de estacionalidad como en cualquier modelo autoregresivo.

4.3 Modelos no paramétricos: Sistemas de Inferencia Difusa y Redes Neuronales Artificiales

4.3.1 Sistemas de Inferencia Difusa (FIS)

Los sistemas basados en reglas difusas comprenden una aproximación con habilidades de aprendizaje, y consisten principalmente en un conjunto de reglas difusas del tipo Sugeno de primer orden (Takagi y Sugeno, 1985). La premisa de los sistemas de inferencia difusa (FIS), en el contexto de predicción, es que los valores previos de las series de tiempo pueden ser usadas para estimar los valores futuros. Lo que hace de FIS una herramienta de uso eficiente, en comparación con otros métodos, es lidiar con la vaguedad de predictores y las relaciones no-lineales entre estos y variables dependientes (Shahab, 2014).

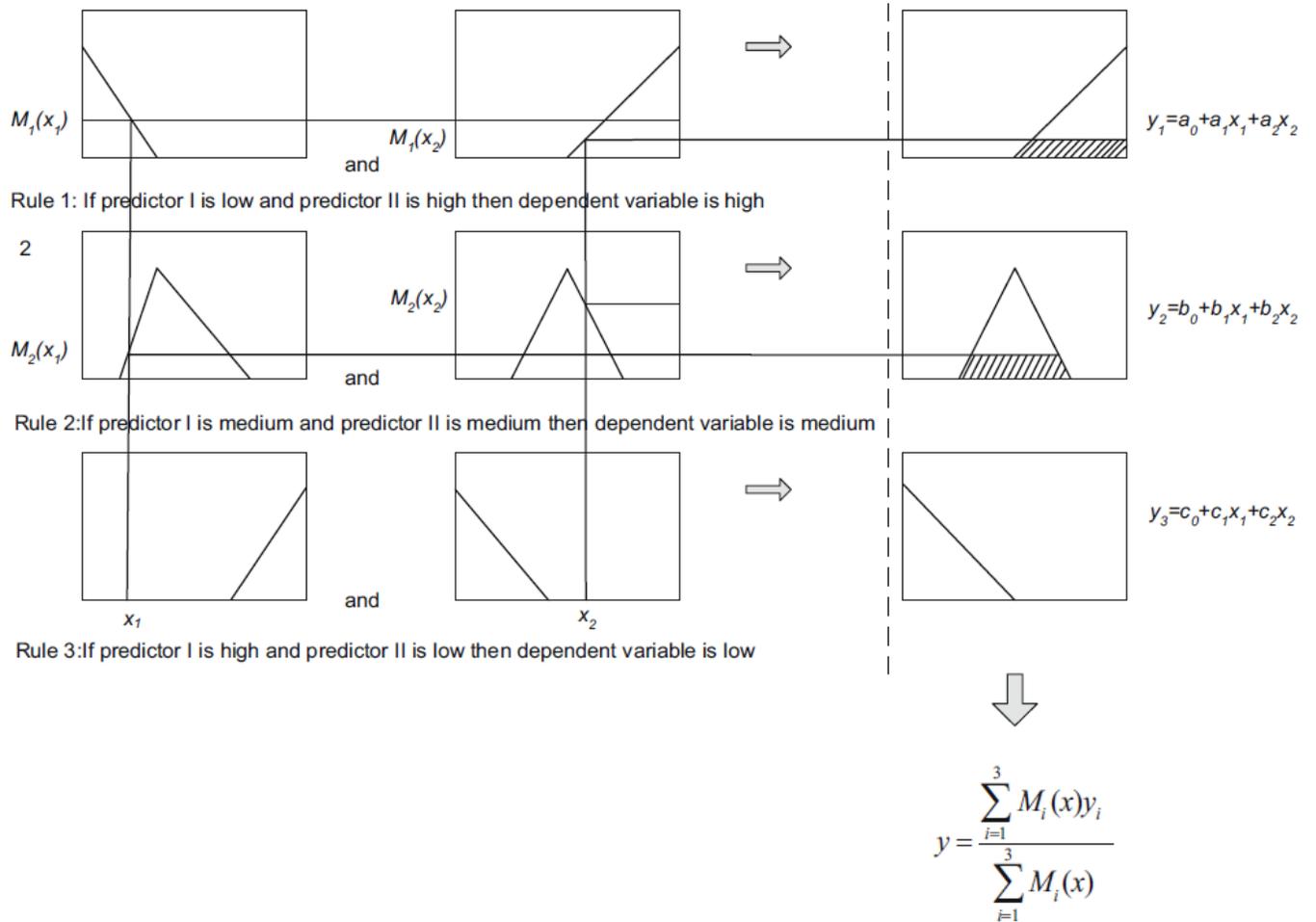
En términos metódicos, estos sistemas de un modo uniforme estiman funciones no-lineales con una precisión arbitraria, donde interpretan los valores en el vector de entrada y , basado en determinados conjuntos de reglas (del tipo *if-then*), asigna valores a un vector de salida.

Interpretar una regla *if-then* comprende distintas partes: primero evaluar el antecedente, y luego aplicar ese resultado a la consecuencia (más conocida como implicancia). Tal antecedente de una regla puede tener múltiples partes, las cuales son calculadas simultáneamente y resueltas a un número singular usando operadores lógicos (habitualmente del tipo *AND* y *OR*). Esto se conoce como la acción de fuzificación de la entrada.

Por su parte, la implicancia especifica que se asignará un conjunto difuso a la salida. Luego, una función de implicancia modifica ese conjunto difuso en el grado especificado por el antecedente. Así, la salida de cada regla es un conjunto difuso. En el caso específico de un sistema de inferencia de tipo Sugeno, se considera un modelo de regresión lineal para cada regla y se agrega la salida de cada modelo mediante promedio ponderado. La **Fig. 6** resume el funcionamiento de este tipo de sistemas a través de un ejemplo, donde $M_i(x)$ es el valor calculado de membresía asociado con la regla difusa i -ésima (un total de tres), componiéndose el sistema de dos entradas y una salida.

De un modo más general, tal como indica Luna y Ballini (2011), un sistema basado en reglas difusas también se puede ver como un conjunto de modelos locales que representan un mapeo local entre variables independientes y dependientes. El nivel de activación de cada modelo local está determinado por el grado de pertenencia del patrón de entrada real a cada subregión. En términos más detallados, cada patrón de entrada tendrá un grado de membresía o pertenencia asociado con cada región del espacio de partición de entrada, y se calcula a través de funciones de membresía gaussianas. En el propósito de identificar grupos de datos que en conjunto produzcan una representación concisa del

Figura 6. Ejemplo de un sistema de inferencia difusa de tipo Sugeno.



Fuente: Data-driving modelling: using MATLAB in water resources and environmental engineering.

comportamiento de un sistema, el uso del algoritmo de *subtracting clustering*, descrito en Chiu (1994) es acertado en el uso de sistemas difusos del tipo Sugeno. De este modo, se agrupan datos de entrada con un rango específico de influencia de *cluster* (conocido como radio, cuyo valor puede tomar entre 0 y 1 de forma continua), lo que permite determinar la cantidad de estos para la data de entrada. Cada *cluster* tendrá asignada su propia regla difusa, y cada variable de entrada y salida tendrá una función de membresía por *cluster*.

Así, utilizar un menor rango de influencia crea usualmente *clusters* más pequeños y de mayor cantidad, conduciendo a un mayor número de reglas difusas. Una vez hecho esto, se tiene una estructura del tipo FIS que posteriormente puede ser evaluada, en este caso, para realizar pronósticos.

En conclusión, para especificar tal estructura FIS descrita antes, se requieren 3 tipos de entrada: vector de variables de entrada, vector de variables de salida (que guardan relación con aquellas variables de entrada), y un determinado radio.

4.3.2 Redes Neuronales Artificiales (ANN)

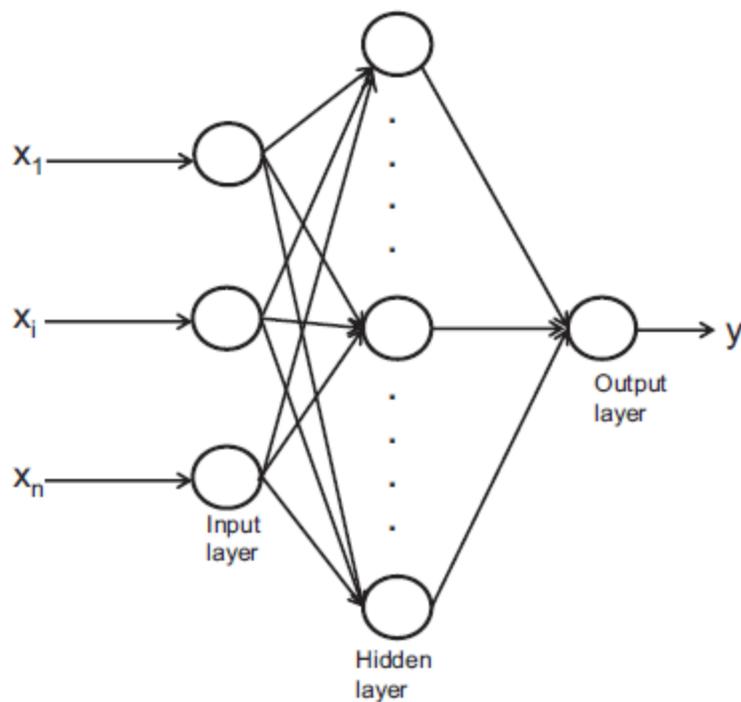
Las redes neuronales artificiales son una poderosa herramienta no paramétrica utilizada para el filtrado de señales, el reconocimiento de patrones y la interpolación, entre muchas otras aplicaciones. Dentro de sus características, también pueden tolerar datos con errores y encontrar asociaciones no-lineales entre los parámetros del modelo. Además, una de sus principales ventajas sobre otros métodos econométricos es que no es necesario tomar la funcionalidad del modelo, lo que significa que no necesita hacer una suposición sobre la relación funcional entre las variables; sin embargo, es necesario incorporar las variables apropiadas para poder hacer una buena estimación.

Cada red neuronal conecta un grupo de variables de entrada $\{x_i\}$, $i = 1, \dots, k$, con un grupo de una o más variables de salida $\{y_j\}$, $j = 1, \dots, k$, y cero, una o más capas ocultas. Las neuronas están conectadas entre capas para hacer conexiones que se activan al alcanzar un umbral, dado que la evaluación de la función de transferencia se basa en los parámetros de entrada. Cada capa puede tener un número diferente de neuronas. La entrada

y la salida pueden comprender variables binarias continuas, discretas o una combinación de todas ellas.

El modelado tradicional de ANN utiliza el enfoque *feed-forward backpropagation*, introducido por primera vez por Rumelhart et al (1986a) y Rumelhart et al. (1986b). El algoritmo de *backpropagation* es un algoritmo para aprendizaje supervisado que busca minimizar el error cuadrático por gradiente máximo de descenso. Este se basa sobre la “propagación posterior” de los errores. La **Fig. 7** ilustra gráficamente este tipo de red neuronal:

Figura 7. Topología de red neuronal multicapa del tipo *feed forward backpropagation*.



Fuente: Forecasting the COMEX copper spot price by means of neural networks and ARIMA models (Sánchez et al., 2015).

De una manera más detallada, la forma general de este tipo de entrada de ANN ingresa valores (uno a la vez), y dado un conjunto de ponderaciones y una función de activación, calcula los valores de salida asociados. En el contexto de predicción, sólo una variable de salida estará presente, la cual corresponderá al valor pronosticado. Este valor se compara con la salida de destino para calcular el error de este registro. El enfoque de gradiente máximo de descenso y el conjunto de parámetros definidos por el usuario (como la tasa de aprendizaje) se utilizan para propagar este error a través de la red ajustando los pesos para minimizar el error total acumulado. El proceso termina y la red está entrenada cuando el error se estabiliza a un nivel aceptable (Guiassi et al., 2005).

En síntesis, para estimar el modelo de red neuronal, es necesario definir las variables de entrada y los parámetros característicos de la red.

4.4 Algoritmos Genéticos (GAs)

Los algoritmos genéticos han sido utilizados por muchos investigadores como herramienta de búsqueda y optimización. Fueron introducidos por Holland en 1975, y a pesar de que existen otros algoritmos evolutivos de la misma lógica, GAs han seguido siendo el enfoque más popular en la literatura (Aguilar et al., 2015), siendo utilizados ampliamente en aplicaciones financieras.

La razón fundamental para el uso de esta herramienta radica en la resolución de problemas de optimización donde la complejidad central está en la gran cantidad de combinaciones posibles que pueden dar solución a este. En estos casos, un objetivo importante a menudo es encontrar rápidamente una solución satisfactoria, aunque no necesariamente óptima. Aquí, la convergencia al óptimo global es algo que sólo podemos

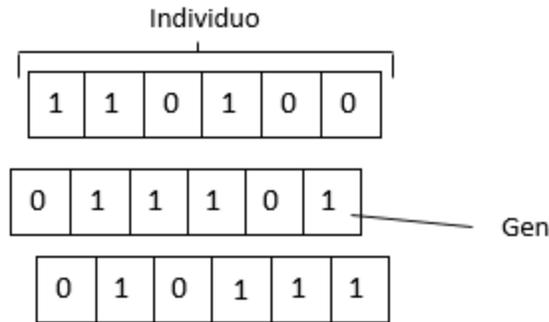
esperar, pero que a menudo es imposible de garantizar dado el tamaño del espacio de búsqueda y el tiempo disponible para el algoritmo de optimización, que lo aleja de propósitos prácticos. En este sentido, los algoritmos evolutivos tienden a encontrar buenas soluciones a muchos problemas con bastante rapidez (Lankhorst, 1996). Tal como el mismo Lankhorst indica, encontrar los puntos óptimos de un gran espacio de búsqueda es a menudo un problema del tipo *NP-hard*¹³, y requiere un número exponencialmente creciente de puntos de búsqueda de espacio para ser investigados a medida que el problema se amplía.

El espacio de búsqueda, Ω , corresponde al conjunto finito de posibles soluciones para el problema. Cada solución de este conjunto se conoce como “individuo”, los que a su vez están compuestos de cromosomas o genes que permiten distinguir a uno de otro. Además, al igual que un problema de optimización, existe una función objetivo a maximizar o minimizar. En el contexto de este algoritmo, dicha función es denominada función de aptitud, y por tanto, el individuo más apto corresponderá a la mejor solución que se quiere encontrar para el problema de optimización. De este modo, GAs proporcionan una forma probabilística de realizar una búsqueda en Ω para una variable arbitraria, dada una codificación¹⁴ de individuos o instancias en cadenas de símbolos (Schmitt, 2001), las cuales habitualmente son de tipo binarias como muestra la **Fig. 8**:

¹³ Este tipo de problemas aparecen comúnmente en lo que se conoce como Combinatorial Optimization. Un ejemplo de esto corresponde al famoso Travelling Salesman Problem (TSP). Para mayor discusión sobre ellos, revisar Jünger et al. (2001).

¹⁴ La codificación adecuada de soluciones mejora la eficacia y eficiencia del algoritmo, sobretodo cuando las variables del problema no son fáciles de codificar como binarias (tales como enteras o continuas).

Figura 8. Población compuesta por tres individuos codificados binariamente.



Fuente: Elaboración propia.

Dicho algoritmo comienza generando una población inicial aleatoria de individuos pertenecientes a Ω , que funcionará como punto de partida. En cada iteración, el algoritmo genético realiza una serie de operaciones en cada población para producir así una nueva generación. Para crear esta última, se seleccionan ciertos individuos en la población actual (llamados “padres”) y se utilizan para crear individuos nuevos que conformarán la siguiente población (conocidos como “hijos”).

Tal como Schmitt señala, el algoritmo comprende tres operaciones: mutación, entrecruzamiento y selección. La primera genera hijos mediante cambios aleatorios en los genes de un padre singular, lo cual persigue visitar la mayor cantidad de puntos posibles en el espacio de búsqueda mediante la creación de diversidad en los individuos (concepto conocido como “exploración”); la segunda combina genes entre un par de padres dando origen a nuevos hijos, lo que en fin, apunta a explotar la vecindad de soluciones relativamente buenas (concepto conocido como “explotación”); y la tercera, modela el éxito reproductivo según la aptitud haciendo que individuos padres sean copiados directamente a la siguiente generación. Estas operaciones se aplican cíclicamente y de manera repetitiva a

poblaciones finitas de tamaño fijo en cada generación, hasta que se satisface una condición de saturación u otra condición de límite.

Así como se mencionó en la introducción, el uso de GAs ha sido protagonista en materia de pronósticos bajo un concepto adaptativo. Tal como Panapakidis y Dagoumas (2017) señalan, un estudio crucial sobre el éxito general de una tarea de pronóstico es la selección adecuada del número y tipo de entradas. En este sentido, el algoritmo genético formará parte complementaria de los modelos de pronóstico ya mencionados en esta sección.

5. METODOLOGÍA

En esta sección se darán a conocer cuáles serán las variables que se utilizarán para la modelación posterior, discerniendo aquellas más aptas de utilizar respecto a otras según algún criterio.

5.1 Fuentes de datos y procesamiento

A este punto de modelar, lo ideal es disponer con bases de datos históricas que permitan reflejar lo mejor posible los fundamentos económicos expuestos previamente, y que determinan los precios del cobre junto a los movimientos detrás de estos. Si bien no es posible contar con estas siempre, los esfuerzos se concentran en obtener aquellas series observables y que, a la vez, se publiquen oportunamente para llevar a cabo estudios de esta naturaleza.

En primer lugar, se obtienen los precios spot nominales del cobre en frecuencia diaria cotizados en la London Metal Exchange. Con respecto a las variables exógenas, se toma en primer lugar los precios futuros a 3 meses diarios y nominales, a fin de representar la actividad especulativa y rol de los inversionistas¹⁵. Por el lado de los inventarios, se toma aquella suma total de inventarios diarios de las 3 principales bolsas de metales (LME,

¹⁵ Se realiza la distinción sólo entre periodos de *contango* y *backwardation* (y no entre *strong* o *weak-contango* también) por motivos de simplicidad, recordando de igual modo que la relación entre precios spot y futuros incluso se mantiene, aunque de manera más débil, en periodos de *weak-contango* o *backwardation*. Relajamos este supuesto de modo de no afectar significativamente nuestros resultados (aunque no necesariamente cierto), al igual que otros estudios como en Levine (2016). Para lo anterior, se emplea una variable dicotómica para modelar tales situaciones. Dado que se quiere utilizar esta información para motivos de pronóstico, para la situación de *contango* (o *backwardation*) del periodo t se usará como proxy aquella situación de *contango* (o *backwardation*) que existió en el periodo inmediatamente anterior, $t-1$. Si bien hay casos en que situaciones de *contango* o *backwardation* se mantienen por un periodo de 1 sólo mes, en la gran mayoría de los casos tales situaciones se manifiestan por varios meses consecutivos (incluso, años).

COMEX and Shanghai). Las 3 series de tiempo anteriores fueron obtenidas desde COCHILCO.

Para reflejar la actividad económica, se decide utilizar el índice de producción industrial de China¹⁶ obtenido del Ministerio de Asuntos Económicos de Taiwan, cuyo reporte está en frecuencia mensual. Por el lado del tipo de cambio real, se decide utilizar aquel tipo de cambio multilateral Major Currency Index, que en su construcción considera a los 7 principales socios comerciales de USA respecto al dólar estadounidense. Este se obtiene desde la U.S Federal Reserve.

Desafortunadamente, no se cuenta con datos históricos mensuales para los costos de la industria (al menos para el conocimiento del autor, no existen reportes para esta frecuencia). Por lo cual, se espera que el resto de las variables pueda captar de alguna forma, incluso si no es en su totalidad, los efectos de los costos sobre la variable de interés. Todos los precios fueron deflactados por el Producer Price Index mensual (también obtenido de la fuente anterior), con base Nov 2016 = 100.

En lo que concierne al procesamiento de los datos, en primer lugar, se toman the daily series in first level, y luego son calculados sus promedios mensuales. A continuación, estas series de promedios mensuales son transformadas en su primera diferencia de

¹⁶ Tal como muestran los resultados de Labys et al. (1999), las fluctuaciones del índice de producción industrial influyen los precios de metales (incluidos el cobre) de un modo razonablemente fuerte, en comparación con otros factores macroeconómicos. Lo anterior, basado en data mensual. Escogemos aquel correspondiente a China debido a que su impulso industrializador ha sido un fenómeno que ha fijado marcas claras en el crecimiento del consumo de una serie de *commodities* que suelen utilizarse intensivamente en la fase de acelerada industrialización, de modo que en el período de 15 años entre 1990 y 2005, la demanda de cobre aumentó entre seis y ocho veces junto a la de aluminio y níquel. La participación de China en el crecimiento de la demanda mundial, al menos para el periodo 2000-2005, fue de 95% para cobre (Radetzki, 2008).

logaritmo natural, resultando en su rentabilidad o variación porcentual mensual, r_t . Esta es calculada específicamente de acuerdo a la Eq. (2), como en Tsay (2010):

$$r_t = \ln\left(\frac{S_t}{S_{t-1}}\right) \quad (2)$$

donde S_t es el promedio mensual de cada serie en el mes t , $S_t = \frac{1}{N} \sum_{i=1}^N P_i$, y P_i corresponde a la cotización del día i (del precio u otro tipo de serie) perteneciente al mes t compuesto de N días, según sea el caso.

En segundo lugar, se toman nuevamente las series diarias en primer nivel, y estas se transforman en su primera diferencia mediante logaritmo natural, resultando en su retorno o variación porcentual diaria, r_i , de modo similar al caso anterior:

$$r_i = \ln\left(\frac{P_i}{P_{i-1}}\right) \quad (3)$$

Las series obtenidas a partir de la Eq. (3) se calculan de ese modo para construir las series de volatilidad en frecuencia mensual para cada variable¹⁷. Así, quedarán en coherencia con aquellas obtenidas a partir de la Eq. (2), sirviendo como entrada directa para la Eq. (1c) y (1a) del modelo GARCH, respectivamente.

La medida de volatilidad a utilizar será la varianza realizada, RV, que se utilizará como proxy para la varianza condicional poblacional σ^2 , de un modo similar como se muestra en Fuertes (2009). Esta se computa como:

¹⁷ Notar que no se contará con una correspondiente al índice de producción industrial chino, dado que como se mencionó anteriormente, los reportes asociados al índice son de frecuencia mensual (no existen reportes diarios). Por tanto, no se puede construir la serie de volatilidad mensual para esta serie, en particular.

$$RV = \frac{1}{N} \sum_{i=1}^N (r_i - \bar{r})^2 \quad (4)$$

donde $\bar{r} = \frac{1}{N} \sum_{i=1}^N r_i$, es decir, el promedio mensual de los retornos diarios.

Finalmente, la varianza realizada es mensualizada multiplicando el lado derecho de la Eq. (4) por el número de días N correspondiente al mes t , resultando:

$$RV_t = \sum_{i=1}^N (r_i - \bar{r})^2 \quad (5)$$

De este modo, se dispone en primera instancia una cantidad de 10 series de tiempo diferentes: 5 para retornos, y 4 para volatilidades, ambas asociadas al precio *spot* del cobre, precio futuro del cobre, inventarios, y tipo de cambio real, cubriendo el periodo Feb 1990¹⁸ hasta Nov 2016, alcanzando un total de $Q = 322$ observaciones por serie. Además de los argumentos prácticos mencionados en la introducción, estimamos que llevar a cabo el estudio en frecuencia mensual resulta adecuado, en primer lugar, para evitar series ruidosas como aquellas habituales de frecuencia diaria o intradiaria; y también, dado que esta frecuencia refleja mejor las fuerzas de demanda y oferta, y por tanto revelará ciclos que incorporan alguna forma de dinámica microeconómica o macroeconómica (Labys et al., 2000).

¹⁸ Entre los diferentes argumentos para comenzar la data en esta fecha, destacan: lo mencionado por Radetzki (2008), en que el impulso industrializador de China evolucionó seriamente a partir de los 90, siendo el ritmo tan asombroso como el que tuvo Japón entre 1950 y 1975; el comienzo del uso intensivo de commodities en los fondos de cobertura desde mediados de 1990 (Humphreys, 2010); el eludir periodos donde sucedieron cambios estructurales, como por ejemplo, la recesión económica prolongada que siguió los shocks de petróleo en 1970 y el colapso de la Unión Soviética, la cual suprimió la demanda global de minerales (Humphreys, 2015); y por último, debido a dificultades de obtener bases de datos en periodos previos a este.

Por su parte, todas las series resultan ser estacionarias mediante la prueba ADF, para un nivel de significancia del 1%. Por tanto, es adecuado utilizarlas en los modelos paramétricos mencionados previamente.

5.2 Especificación del algoritmo genético y su contexto en el pronóstico

Dado el set de variables explicativas que se tiene, la pregunta a responder es: ¿cuál o cuáles de ellas, en forma conjunta, permiten alcanzar mejores pronósticos mensuales de volatilidad? En el caso del mercado del cobre, la relativamente nula capacidad de ajuste de la oferta en el corto plazo podría llevar a procesos de ajustes dinámicos en el tiempo, como explica García y Guzmán (2017), debido a los cambios que presenta la demanda que son principales responsables de volatilidad en el corto plazo. En este sentido, entonces, modelaremos lo anterior utilizando la idea de que la información contenida en el pasado reciente (dentro del corto plazo) tiene influencia en los precios percibidos en el presente.

Para esto, en los modelos se consideran los rezagos de todas las series de tiempo anteriores, tomando aquellos desde el periodo 1 hacia atrás hasta un máximo de 24, de modo de captar posibles patrones de comportamiento en ciertos periodos¹⁹ en este horizonte temporal. La hipótesis detrás de esto es que la variabilidad en los precios del cobre observada en un mes t , está determinada por aquella información pasada reciente que está contenida en la variabilidad de variables exógenas, de un periodo máximo de hasta 2 años. Es posible que, sólo algunos meses dentro de estos 24 mencionados, posean mayor

¹⁹Al menos, en términos de ciclicidad, Labys et al. (2000) encontró periodicidad menor que 12 y 24 meses, relacionado a influencias especulativas, y ciclos macroeconómicos de corto plazo en la economía estadounidense, respectivamente.

influencia en el valor actual, y por tanto, contienen la información que posee mayor capacidad predictiva.

A este punto, dado este conjunto de series de tiempo que sirven de entrada a cualquier modelo, la pregunta que cabe plantearse es: ¿cuántas especificaciones distintas pueden crearse, por ejemplo, por el modelo de tipo ARIMAX? Como se puede notar, lo anterior da origen a diferentes combinaciones posibles, dependiendo del tipo de modelo y su especificación exacta. Lo único que es claro es que, alguna de estas dentro de las existentes, será la que entregue los mejores pronósticos. Notar que para el caso específico de los modelos GARCH, la cantidad de combinaciones posibles abarcan las series de tiempo tanto de retorno mensual como de varianza, dado que el desempeño de la ecuación de esta última depende de cómo se especifique la ecuación de la media debido al término de las innovaciones, como se puede deducir de la Eq. (1a) y la Eq. (1c). En conclusión, esto conduce a que, en este tipo de modelos en particular, el espacio de búsqueda sea considerablemente mayor que en el caso ARIMAX, por ejemplo.

El Apéndice ilustra la forma de estimar la cantidad de combinaciones posibles para el caso de modelos GARCH hasta un máximo de 24 rezagos, a fin de dimensionar el tamaño de los problemas que se generan. Así, para un modelo GARCH con 198 variables de entrada²⁰, la cantidad de combinaciones es de aproximadamente $U = 9.64 \times 10^{60}$,

²⁰ Para la ecuación de la media: 24 rezagos de retornos del cobre, 24 de medias móviles, 24 del *Major Currency Index*, 24 de Inventarios, 24 del Índice de Producción Industrial chino, 1 de retornos de futuros del cobre, 1 de variable dicotómica, y 1 de una constante. Esto implica una cantidad de 123 variables disponibles para esta ecuación.

Para la ecuación de la varianza: 24 rezagos de varianza de retornos del cobre, 24 del *Major Currency Index*, 24 de Inventarios, 1 de varianza de retornos de futuros del cobre, 1 de variable dicotómica, y 1 de una constante, resultando una cantidad de 75 variables para esta ecuación (en el caso de retornos y varianzas de futuros del cobre, se consideró 1 sólo rezago, puesto que al cambiar la situación de *contango* a *backwardation* después de muchos meses consecutivos, no sería coherente considerar información de muchos meses hacia

mientras que en el caso de un modelo ARIMAX con la misma cantidad de rezagos la cantidad desciende a $U = 5.31 \times 10^{36}$ aprox. Notar que la forma de cuantificar las combinaciones en el caso de los modelos no-paramétricos es igual a las del modelo ARIMAX.

Dado el contexto anterior, y de un modo más amplio, el objetivo de encontrar los pronósticos más certeros en este artículo se sustenta en contestar las siguientes preguntas: ¿cuál o cuáles modelos (si es el caso, de forma híbrida) se deben seleccionar? ¿cuál o cuáles variables (en conjunto), y en qué rezagos, se debe(n) considerar como entrada en los modelos a utilizar? ¿cuál es el tamaño de ventana de datos a emplear para estimar/entrenar los modelos? ¿qué valor deben tomar los parámetros inherentes a las configuraciones de los modelos a utilizar? De este modo, el uso de algoritmos genéticos, como herramienta de optimización, se empleará para dar cabida a las tres últimas interrogantes, mientras que la primera se abordará mediante simple selección y combinación de modelos de forma arbitraria a la hora de experimentar. Con este medio, quisiéramos que las complejas dinámicas que determinan los precios y cambios en estos comentadas previamente sean captadas de alguna forma.

Para medir el desempeño de los modelos de pronóstico, se emplearán las funciones de pérdida HMSE, Theil-U, QLike y MSE. Las dos primeras por el hecho de ser más informativas (Fuertes, 2009). Por su parte, Qlike and MSE son reconocidas por ser funciones de pérdida robustas (Patton, 2011).

atrás para la situación presente). Así, el número total de variables para especificar un modelo GARCH, de la forma que sea, es de 198.

No obstante, la selección del mejor modelo (en el contexto del problema de optimización que será resuelto por el algoritmo genético) estará basada en la función de pérdida HMSE dado que, como señala Fuertes, está ajustada por heteroscedasticidad:

$$HMSE = \frac{1}{n} \sum_{i=1}^n \left(1 - \frac{\tilde{\sigma}_t^2}{RV_t}\right)^2 \quad (6)$$

donde $\tilde{\sigma}_t^2$ corresponde a la varianza pronosticada para el mes t , y n a la cantidad de periodos a predecir. Notar que la no-linealidad de la Eq. (6) justifica también el uso de el algoritmo genético como herramienta de optimización, siendo parte de la función objetivo (o más bien, función de aptitud). Así, el resto de las funciones de pérdida sólo serán informativas sobre cada modelo, pero no para comparar entre diferentes modelos.

En este sentido, de forma particular, la metodología es parecida a la utilizada en Hung (2011a). El horizonte de tiempo a pronosticar escogido, y que es relevante en el cálculo de la función de pérdida, es aquel correspondiente a 5 años, lo que implica predecir un total de 60 observaciones que comprende el periodo Dic 2011 hasta Nov 2016. De este modo, en la Eq. (6), $n=60$.

Con respecto a la GA, la codificación de individuos es especificada mediante strings cuyos cromosomas serán del tipo binario en los modelos paramétricos (tanto para las variables exógenas de entrada como en los parámetros inherentes de los modelos, y el tamaño de ventana). De este modo, por ejemplo, un valor igual a “1” en un cromosoma referente a un determinado slot de un individuo significará: “utilizar esta específica serie de tiempo” para el conjunto de variables que serán entrada en ese tipo de modelo (for instance, para el modelo GARCH); mientras que un valor igual a “0” significará: “omitir esta serie

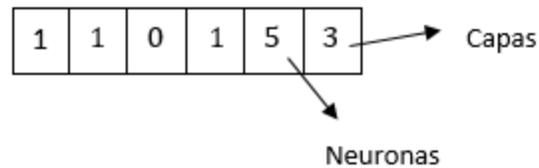
de tiempo” para el conjunto de variables serán entrada en ese tipo de modelo. Por tanto, cada slot está asociado a una variable singular y rezago único. Cada individuo representa, entonces, una especificación única de un mismo tipo de modelo (ya sea ARIMAX o GARCH).

En el caso de los modelos no-paramétricos, la codificación binaria sólo se utiliza para las variables explicativas de entrada del mismo modo que antes. Por el lado de los parámetros inherentes a los modelos y el tamaño de ventana, estos son usados por sí mismos²¹. Es decir, los tamaños de ventana serán variables enteras no binarias, al igual que la cantidad de neuronas y capas ocultas en el caso de los modelos que involucren ANN. En el caso de aquellos que comprendan FIS, el valor del radio a utilizar será una variable de tipo continua. Así entonces, en el contexto del problema de optimización a resolver por la GA, en aquellos modelos paramétricos, ANN, o combinaciones entre estos, se resolverán Problemas Enteros; mientras que en aquellos que consideren sólo modelos FIS, o híbridos de paramétricos con FIS, se resolverán Problemas Entero-Mixtos²². **Fig. 9** ejemplifica el caso de codificación para un modelo del tipo ANN:

²¹ La razón de esta diferencia se basa en la eficiencia de los softwares utilizados en cada modelo, en lo que refiere a tiempos computacionales de procesamiento de los programas. La capacidad para codificar y relacionar de forma eficiente los individuos propuestos, con los modelos seleccionados dentro de estos programas, varía de un software a otro.

²² Tal distinción será fundamental al estar relacionado con la complejidad de los problemas que resolverá el algoritmo. Los Problemas Enteros comprenden un espacio de búsqueda finito al estar asociado a variables de tipo discretas, mientras que los Problemas Entero-Mixtos conciernen un espacio de búsqueda infinito, por involucrar variables de tipo continua, aparte de las discretas. Siendo la GA una heurística, las probabilidades de alcanzar óptimos globales en estos últimos problemas disminuyen en comparación a los primeros.

Figura 9. Ejemplo de una cadena reducida de *strings* para un modelo ANN.



Fuente: Elaboración Propia

Notar que, con lo descrito anteriormente, se puede deducir que la cantidad de combinaciones posibles a la hora de estimar modelos (independientemente del tipo) se incrementa aún más, al permitir la posibilidad de que sus parámetros inherentes así como el tamaño de ventana varíen. Esto, por ende, conduce a enfrentar espacios de búsqueda aún más grandes que los comentados previamente²³. Para lidiar con esto, es que se toman en cuenta las siguientes consideraciones: la primera, disminuir la cantidad de rezagos desde 24 hasta 12 sólo en el caso de estimar con los modelos paramétricos; en segundo lugar, y también sólo en aquellos modelos paramétricos, considerar sólo 13 ventanas posibles; y en tercera instancia, discernir cuáles variables son aquellas más aptas para ser modeladas de forma lineal en reemplazo de una no-lineal. Para esto último, se opta por utilizar el *test* de Teraesvirta (Teraesvirta et al., 1993), el cual asesora si un grupo de series de tiempo están relacionadas no-linealmente. Esto permite identificar relaciones del tipo lineal y no-lineal entre variables explicativas y variables explicadas, de modo de orientar mejor la entrada de datos para motivos de pronóstico. Por tanto, en aquellos modelos paramétricos, se quisiera priorizar algunas variables explicativas en la entrada, seleccionando sólo aquellas que

²³ Esto podría generar que el algoritmo genético converja a óptimos locales al depender su desempeño, en gran medida, de la diversidad que presente la población inicial aleatoria generada. Tal como indica Lankorst (1996), el tamaño de la población es además influyente en el comportamiento de la convergencia. Si esta es demasiado pequeña, podría ocurrir convergencia prematura al igual que en el caso de una tasa de mutación muy baja. En este sentido, se pueden tomar las siguientes consideraciones: reducir el espacio de búsqueda, utilizar tasas de mutación relativamente altas, o utilizar gran tamaño de población.

resulten tener una relación lineal con la variable a explicar. A modo de ilustración, en el mismo ejemplo del modelo GARCH, esto permite reducir el espacio de búsqueda²⁴ desde más de $U = 9.64 \times 10^{60}$ combinaciones a otras del orden de $U = 3.86 \times 10^{29}$ combinaciones.

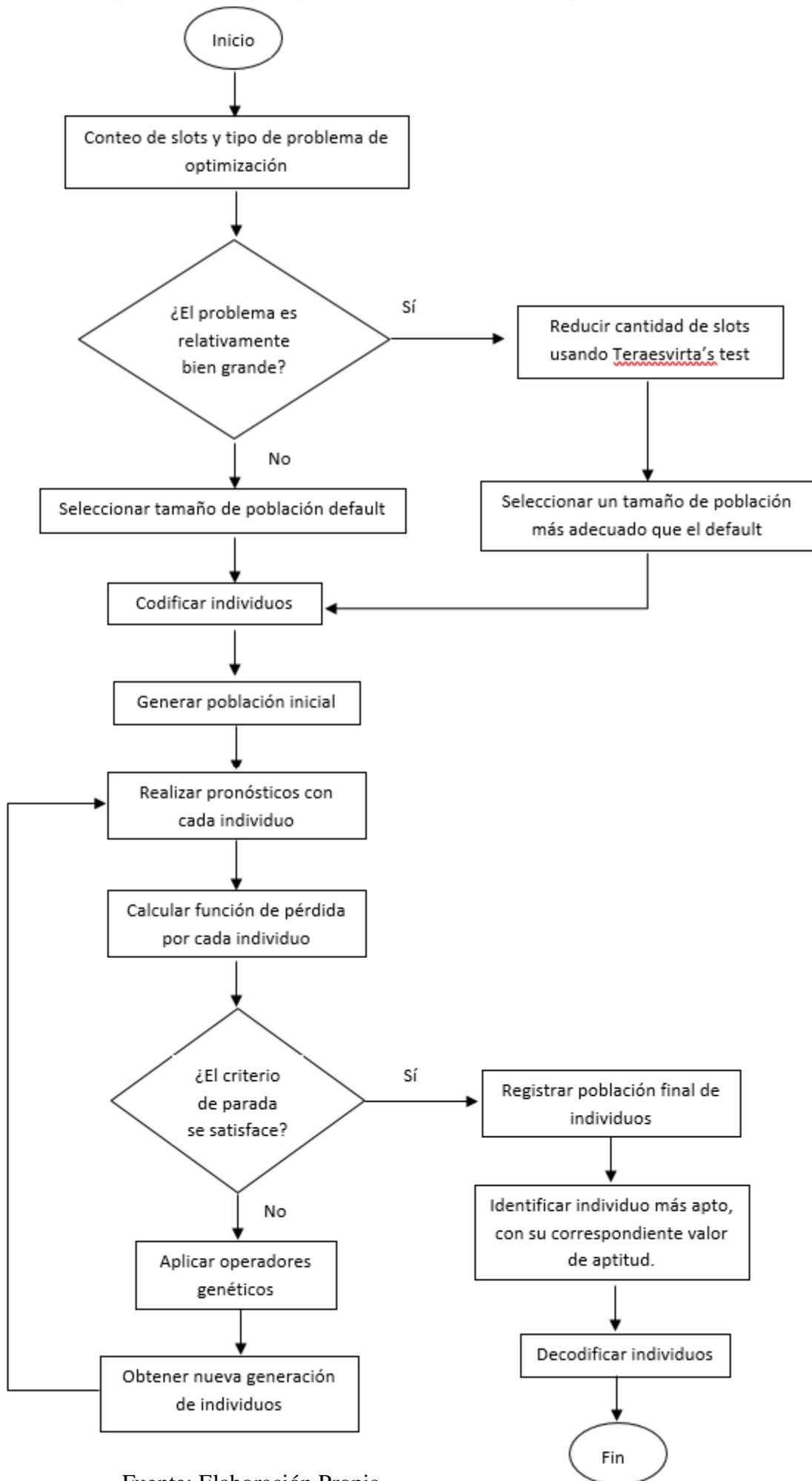
5.3 Funcionamiento adaptativo de los modelos

Fig. 10 permite representar gráficamente el funcionamiento de todos los modelos adaptativos usados en este estudio. Como se puede apreciar, la identificación inicial de la complejidad del problema de optimización es necesaria para decidir la utilización de la herramienta con el objeto simplificarlo. En este caso, el *test* de Teraesvirta. En problemas relativamente grandes (que en este estudio consideramos arbitrariamente aquellos del tipo Entero-Mixtos, y/o aquellos con individuos con cantidad de *slots* (o variables) igual o superior a 80), decidimos utilizar grandes tamaños de población de individuos, de un orden de 900 por generación. En el caso contrario, tamaños de población de 400 individuos por generación, al cual se refiere como valor default en el esquema mostrado.

En cada generación, cada individuo (o especificación de modelo) efectúa exactamente 60 pronósticos fuera de muestra y del tipo *one-step* hacia adelante, para un tamaño de ventana fijo, pero variable a lo largo de todos los que conforman la población. En otras palabras, cada modelo no cambia su especificación ni tamaño de ventana al pronosticar periodos distintos; pero dos modelos pertenecientes a una misma población podrían utilizar tamaños de ventana diferentes. Luego los pronósticos son utilizados para

²⁴ Para la ecuación de la media, mediante Teraesvirta's test, la cantidad de variables se redujo a 63, mientras que para la ecuación de la varianza se redujo a 27.

Figura 10. Diagrama de Flujo. Operación de un modelo adaptativo.



Fuente: Elaboración Propia

computar la aptitud de cada individuo, o función de pérdida HMSE, de la Eq. (6). En el caso de los modelos de paramétricos, se puede desprender que los parámetros estimados de sus ecuaciones se reajustan de un periodo a otro, al considerar distintas observaciones de la muestra. De un modo similar, los modelos que involucran ANN utilizan un set de observaciones de entrenamiento y test cuya suma equivale al tamaño de ventana, para cada periodo antes de realizar cada pronóstico; en el caso de los modelos que comprenden FIS, se crea una estructura de tipo Sugeno que contiene el entrenamiento en cada paso del proceso de *rolling window*, para después evaluarla en el pronóstico del periodo siguiente. De este modo, el número de reglas difusas codificadas en la estructura varía en el tiempo, de acuerdo a los cambios o variaciones identificadas en las series de tiempo, mientras que la nueva información se procesa en cada instante de tiempo.

En la transición de una generación a otra, se aplican los operadores genéticos de selección, mutación, y *crossover*²⁵. De este modo, para el ejemplo de 900 individuos, la siguiente generación está compuesta por los mismos 90 mejores individuos de la generación anterior, por 81 individuos hijos provenientes de mutación, y 729 individuos hijos obtenidos mediante *crossover*. Este proceso iterativo del GA continua hasta que el cambio porcentual acumulado en el valor de la función de aptitud sobre un valor límite especificado de número de generaciones es menor que la tolerancia deseada; con este criterio de parada, todos los individuos en la población final son, en cierto modo, similares. Finalmente, habiéndose detenido la ejecución del algoritmo, se obtienen y registran los

²⁵ Se utilizó un valor de 10% para Selección; adicionalmente, 90% *crossover*, y por tanto, 10% para mutación. Esto dado que, aparte de llevar a mejores resultados por medio de pruebas experimentales, *crossover* es el principal driver de búsqueda de algoritmos genéticos (Lankhorst, 1996). Mutación principalmente sirve para mantener un cierto nivel de diversidad genética en la población para prevenir convergencia prematura de la población a valores subóptimas.

individuos que componen la población final, incluyendo tanto sus composiciones en cromosomas como sus valores de aptitud.

En lo que respecta a modelos híbridos, como se vio en la introducción, su uso comprende una alternativa potencial para alcanzar mejores pronósticos que aquellos alcanzados por modelos de forma independiente. **Fig. 11** ilustra el proceso de conformación de estos, los cuales se obtienen realizando dos pasos. En primer lugar, una vez seleccionado el tipo de modelo paramétrico y la obtención de la población final de individuos de forma adaptativa, se rankean aquellas especificaciones de modelos según el valor de su HMSE, desde menor a mayor. De esta población total, se identifican y seleccionan aquellas 12 mejores especificaciones²⁶, y se identifica el valor de tamaños de ventana T óptimos de estos²⁷, para luego predecir los últimos $C = Q - T$ observaciones. Esto es, el número total de observaciones que se disponen (las cuales comprenden out and in-sample), menos la cantidad óptima de observaciones que resultaron necesitar esos individuos para estimarse y predecir.

De este modo, la data de entrada para efectuar el Paso 2 se conforma de: 12 series de tiempo, $\{Y_i\}_{i=1}^{12}$, donde cada una comprende C pronósticos realizados por las mejores especificaciones del modelo paramétrico en cuestión²⁸; y el set de variables explicativas,

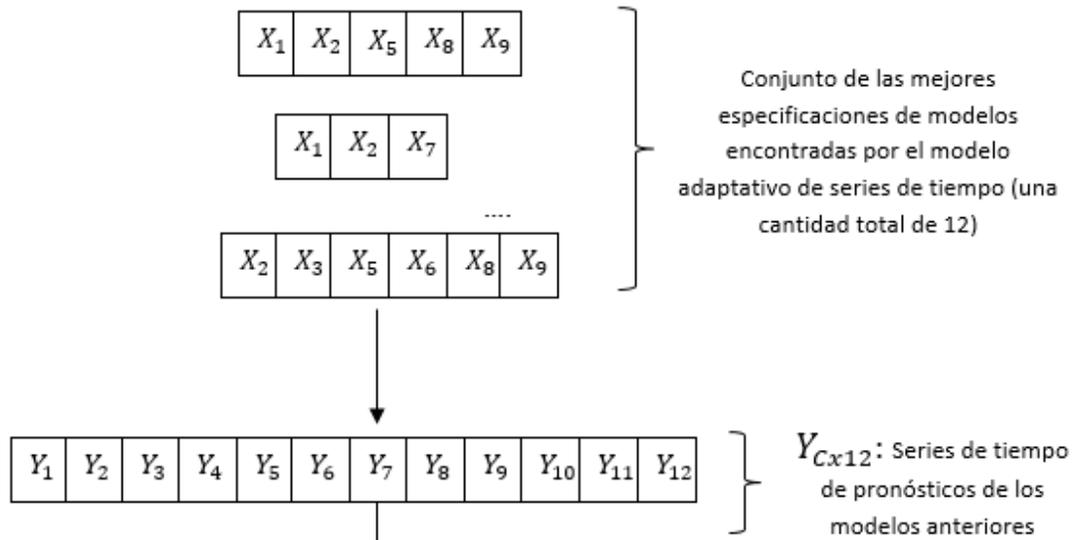
²⁶ Esta cantidad se determinó de forma arbitraria, basada en el propósito de que los valores de aptitud no se alejen considerablemente de aquel mejor obtenido.

²⁷ Todas las pruebas, independiente del modelo, arrojaron tamaños de ventana iguales para casi todos los individuos en la población final. Aquellos pocos casos en que no fue así, el valor de aptitud era peor.

²⁸ Sólo por propósitos de alcance, todos los modelos híbridos creados se componen sólo de modelos paramétricos en su Paso 1, y sólo de modelos no-paramétricos en su Paso 2.

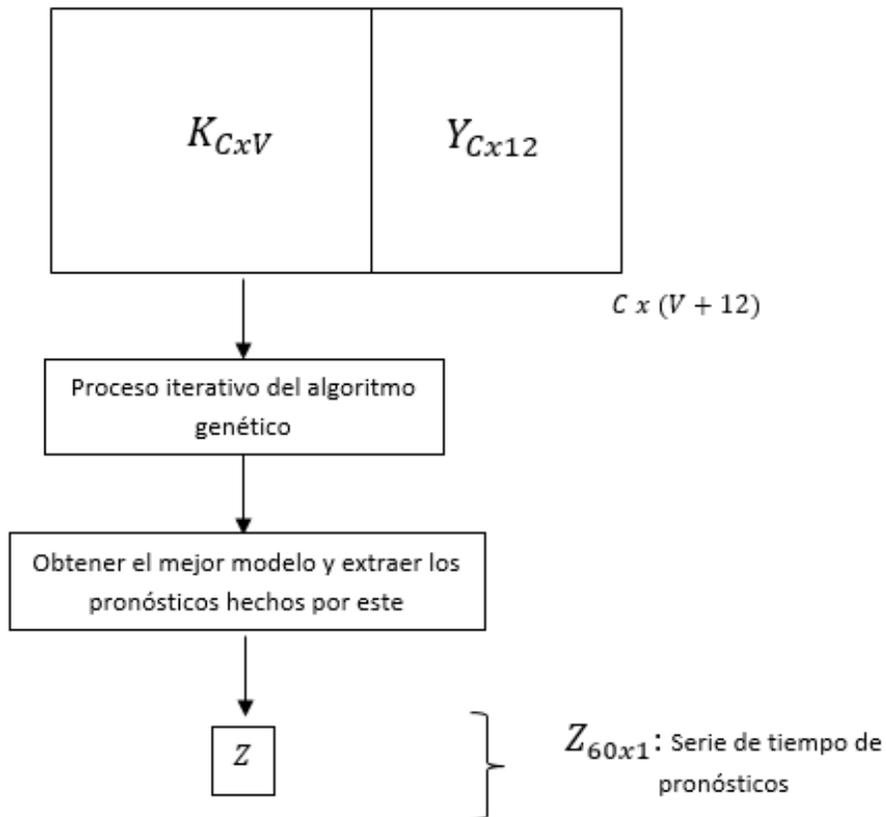
Figura 11. Proceso de formación de un modelo híbrido.

Paso 1:



Codificar las series de tiempo de Y_{Cx12}

Paso 2:



que consiste en varianzas rezagadas hasta un máximo de 24 veces²⁹. Lo primero apunta al objetivo de aumentar la probabilidad de que el modelo no-paramétrico reconozca un patrón de predicción, al considerar pronósticos realizados por modelos paramétricos similares, y al detectar que tales pronósticos deben superarse. Esto último se busca que se logre complementando tales predicciones mediante alguna combinación de dichos modelos en conjunto y con la información de variables explicativas, como en el Paso 1. Bajo el mismo propósito anterior, es que se consideró esa cantidad de pronósticos para conformar el modelo híbrido, es decir C , puesto que cada una de estas observaciones es generada por el mismo tamaño de muestra, que corresponde a T .

Para efectos de probar la posible sinergia, las series de tiempo asociadas a los pronósticos de los mejores modelos son codificadas, al igual que las variables explicativas, mediante *string* binarios como antes. Así, el Paso 2 también corresponde a un modelo adaptativo de la misma especie como el explicado en la **Fig. 11**. Sin embargo, aquella serie de mejores pronósticos (correspondiente a aquella especificación mejor) se configuró como un input de entrada fijo. Es decir, queremos que la información de variables explicativas, sumada a la de pronósticos del resto de los 11 modelos restantes, complementen (de alguna forma conjunta) los pronósticos hechos por la mejor especificación del modelo de series de tiempo, mediante alguna combinación. Esto conlleva a una matriz de entrada resultante dada por la combinación de dos matrices: K , que en el esquema representa aquella que sólo contiene información de rezagos de las variables explicativas (donde V equivale al número de series de tiempo de dichas variables), mientras que la matriz Y a aquella matriz que

²⁹ Para este caso de modelos híbridos se decidió aumentar esta cantidad ya que el tamaño de los problemas que se generan en este paso resulta ser de menor dimensión, que en un caso GARCH por ejemplo. Esto pues no involucra la especificación de una ecuación de media, o la integración de medias móviles (aunque esto último podría hacerse).

contiene las series de tiempo que conciernen los pronósticos de los mejores 12 modelos.

Por su parte, del mismo modo que en los modelos no-híbridos, este paso del modelo adaptativo considera también como variables a los parámetros inherentes a la configuración de los modelos no-paramétricos. Se proponen así los siguientes modelos:

Tabla 1. Modelos híbridos y no-híbridos propuestos.

Tipo de modelo		Cantidad de variables en GA	Parámetros de configuración variables. [LB, UB] ^b	Tamaño de Ventana [LB,UB]	Terasvirta's test	Tamaño de población	Tipo de problema
No-híbrido	1 A ^a -ARIMAX	64	Ninguno	[96,168]	No	400	Entero
	2 A-GARCH	99	Órdenes p y q : [1,4] ^f	[96,168]	Sí	900	Entero
	3 A-GARCH ^c	58	Órdenes p y q : [1,4]	[96,168]	No	400	Entero
	4. A-ANN	77	Neuronas: [3,78], Capas: [2,30]	[48,124]	No	400	Entero
	5 GARCH(1,1)	1	-	[96,168]	-	-	-
	6. A-FIS	76	Radio:]0, 0.3] ^e	[40,60] ^e	No	900	Entero Mixto
Híbrido	7. A-GARCH-ANN	88	Neuronas: [3,78], Capas: [2,30]	[48,124]	Sí ^d	900	Entero
	8. A-GARCH-FIS	87	Radio:]0, 0.3]	[40,60]	Sí ^d	900	Entero Mixto
	9. A-ARIMAX-FIS	76	Radio:]0, 0.3]	[40,106]	No	900	Entero Mixto

a. "A-" denota "Adaptativo".

b. LB denota Lower Bound (Cota Inferior). UB denota Upper Bound (Cota Superior).

c. Este corresponde a un A-GARCH auto-explicado, es decir, utiliza solo variables relacionadas a rezagos de la variable dependiente (tanto retornos como varianzas).

d. Las pruebas sólo fueron aplicadas en el Paso 1 en estos casos, ya que estos son modelos híbridos.

e. En experimentos previos, todos los modelos que involucran FIS entregaron mejores resultados en esos rangos de radios y tamaños de ventana.

f. En experimentos previos, todos los modelos que involucran GARCH entregaron mejores resultados en esos rangos de ordines de p y q , principalmente debido a asuntos de convergencia.

Los modelos híbridos que presenta la **Tabla 1** se conforman, en el Paso 1, a partir de los modelos no-híbridos mostrados ahí mismo. Por ejemplo, el modelo híbrido A-ARIMAX-FIS se confecciona utilizando, en su primera parte, el modelo de series de tiempo A-ARIMAX. Por tanto, el tamaño de ventana óptimo resultante del primero es determinante para las dimensiones de la entrada de datos del segundo. Otra consecuencia de

esto es que la segunda columna, para el caso de los modelos híbridos, se refiere a los parámetros variables asociados al modelo no-paramétrico utilizado en el Paso 2.

La proposición de tales modelos radica, principalmente, en las ventajas que ofrece cada tipo de modelo de forma separada, y cómo se pueden complementar estos para conducir a mejores resultados³⁰. Sin embargo, algunos modelos tienen un propósito específico: el modelo GARCH (1,1), el cual no es del tipo adaptativo (y que está contenido dentro de la gama de especificaciones que comprende el A-GARCH), se decidió proponer de forma independiente por ser ampliamente utilizado (Hull, 2015); por su parte, el modelo A-GARCH autoexplicado, el cual sólo comprende rezagos (tanto de retornos como varianzas) correspondientes a la variable dependiente, se sugiere para dimensionar el rol que tienen los rezagos de variables propias la dependiente a la hora de pronosticar. Finalmente, el modelo A-ARIMAX-FIS, de forma excepcional, sólo utilizó el mejor modelo proveniente de aquel de series de tiempo, en reemplazo de los 12 mejores³¹.

Para efectos de probar la superioridad estadística de los modelos propuestos, el test de Model Confidence Set (MCS) es aplicado (Hansen et al., 2003, 2011). Este se construye a partir de un criterio de evaluación y de diferentes modelos, distinguiendo el mejor de ellos de acuerdo a su precisión en el pronóstico.

Los pronósticos de los modelos de series de tiempo se obtuvieron mediante el software Eviews, mientras que el algoritmo genético y los pronósticos de aquellos modelos

³⁰ Por ejemplo, Kristjanpoller y Minutolo (2015) explican que una de las ventajas que un modelo ANN–GARCH tiene sobre otras técnicas es la habilidad que la ANN tiene de aprender de los errores pronosticados por el GARCH.

³¹ La razón radica en que los 12 mejores modelos se diferenciaban muy pocos unos de otros en términos de aptitud (diferencias se reflejaban en la quinta cifra significativa), lo cual refleja que los pronósticos que ofrecen estos son muy similares.

no-paramétricos o que incluyen inteligencia artificial se implementaron mediante distintas *toolbox* de Matlab³². Para proveer la capacidad de adaptación en los modelos estimados en Eviews mediante el algoritmo genético se implementó una interfaz simultánea entre los softwares correspondientes (Eviews y Matlab). El software *R* fue utilizado para la aplicación del test de Teraesvirta, y para un procesamiento de datos más expedito en la entrada de estos a los diferentes programas.

³² Específicamente: *Global Optimization toolbox*, *Neural Network toolbox*, y *Fuzzy Logic toolbox*.

6. RESULTADOS Y DISCUSION

La **Tabla 2** resume los diferentes desempeños obtenidos por los modelos propuestos, destacando también aquellos parámetros de configuración óptimos para cada caso, así como el mejor tamaño de ventana que debe considerarse para efectuar pronósticos en cada uno.

Tabla 2. Resultados de los modelos híbridos y no-híbridos propuestos.

Tipo de modelo		Funciones de pérdida				Tamaño de ventana óptimo	Parámetros de configuración óptimos	MCS valor-p
		HMSE	THEIL-U	Q-LIKE	MSE			
No-híbrido	1 A ^a -ARIMAX	0.272	0.321	-4.966	4.53E-06	156	-	0.144
	2 A-GARCH	0.165	0.401	-5.280	5.09E-06	162	$p=0, q=2$	0.068
	3 A-GARCH ^b	0.163	0.413	-5.246	5.30E-06	168	$p=0, q=2$	0.068
	4. A-ANN	0.263	0.546	-5.415	7.23E-06	118	Neuronas: 34, Capas: 22	0.008
	5 GARCH(1,1)	0.376	0.366	-4.954	4.77E-06	168	-	0.067
	6. A-FIS	0.173	0.328	-5.046	4.65E-06	47	Radio: 0.2820	0.244
Híbrido	7. A-GARCH-ANN	0.263	0.54	-5.416	7.24E-06	118	Neuronas: 34, Capas: 12	0.024
	8. A-GARCH-FIS	0.130	0.279	-5.086	3.34E-06	45	Radio: 0.276	1
	9. A-ARIMAX-FIS	0.164	0.288	-5.069	3.51E-06	47	Radio: 0.2825	0.684

a. “A-“ denota “Adaptativo”.

b. Este corresponde a un A-GARCH auto-explicado, es decir, utiliza solo variables relacionadas a rezagos de la variable dependiente (tanto retornos como varianzas).

Si bien se exponen diferentes funciones de pérdida para comparar resultados, la única relevante para determinar cuál es el mejor modelo (y comparar desempeños entre modelos) corresponde al HMSE. Esto dado que, como se explicó en la sección anterior, esta corresponde a la función de aptitud a optimizar mediante el algoritmo genético. Por tanto, podría ser que si dicha función objetivo se asigna a otra función de pérdida en el GA, tanto la especificación del modelo óptimo así como su tipo sea diferente. En otras palabras, la selección de la función de pérdida que se utilizará como aptitud es inherente al mejor modelo que resultará obtenerse.

A primera vista, se puede observar que el tamaño de ventana óptimo resultante no siempre equivale al mayor que se dispone, como en el caso A-GARCH, que alcanza mejores pronósticos con un tamaño de ventana de 162 en reemplazo de 168.

Al considerar sólo los modelos paramétricos, se aprecia que los del tipo A-GARCH resultaron tener mejor desempeño que aquel A-ARIMAX al observar el HMSE, dejando en evidencia dos cosas: en primer lugar, la mejora en los pronósticos debido a la capacidad que tienen los modelos GARCH de considerar la heteroscedasticidad, a diferencia de los ARIMAX cuya varianza condicional de las innovaciones es supuesta como constante; y en segundo lugar, la importancia de considerar la capacidad de adaptación en modelos paramétricos, la cual queda reflejada en el mejor desempeño de modelos menos eficientes y adaptativos en comparación con aquellos más eficientes pero no-adaptativos (como en el caso A-ARIMAX against GARCH(1,1)). Adicionalmente, en este mismo sentido, el hecho de considerar la capacidad de adaptación también resulta interesante al observar cómo modelos sólo paramétricos pero adaptativos logran superar, incluso, a modelos híbridos adaptativos competitivos, como en el caso A-GARCH contra A-ARIMAX-FIS.

Más aun, basta inferir que, dado que los modelos adaptativos seleccionan determinados rezagos de variables explicativas (ya sea de forma consecutiva, o intermitente para captar posibles fenómenos estacionales), estos tienen gran valor al contener en ellos aquellas especificaciones que comprenden rezagos consecutivos hasta un cierto límite arbitrario, por ejemplo, como en el tipo GARCH. Estas últimas son muy comunes, como por ejemplo, en Kristjanpoller y Hernandez (2017) que emplean regresores de hasta 1 sólo periodo hacia atrás, o incluso hasta 2 como en Gargano y Timmermann (2014). Es decir, fijando un mismo tipo de modelo, la naturaleza de adaptación comprende una generalización para los modelos que no tienen esta capacidad.

Con respecto al caso particular de los modelos A-GARCH, se aprecia que estos alcanzan mejores desempeños cuando no consideran the squared innovations from previous periods, lo cual podría suceder debido a que tal información es captada por variables exógenas de esos periodos anteriores o más lejanos, comprendiendo así una decisión determinante el hecho de haberlas incluido. Esto sería posiblemente, en un primer intento, una demostración del rol relevante que juegan estas variables en el modelo. De modo contrario, la varianza condicional de periodos pasados propia de este tipo de modelos sí resulta relevante en estos hasta al menos los 2 meses anteriores, permitiendo capturar fenómenos como *volatility clustering*.

Por otro lado, también queda en evidencia que la inclusión de modelos no-paramétricos en aquellos paramétricos con el objeto de hacerlos híbridos, no siempre lleva a mejores resultados, algo que ya se había reportado en la introducción al mencionar resultados de otras investigaciones. Lo anterior se deduce al observar el desempeño de modelos como A-GARCH y A-GARCH-ANN, donde la mayor inclusión de inteligencia artificial e innovación en el segundo no generó mejoras de los pronósticos del primero. Si bien se quiso reducir la posibilidad de lo anterior mediante la inclusión de pronósticos de las 12 mejores especificaciones del modelo del Paso 1, no siempre se pueden reconocer esos patrones dentro del procesamiento de información en tales modelos. El motivo de esto último, al menos en este caso, se encontraría en el tipo de modelo no-paramétrico utilizado (redes neuronales), el cual pareciera no rendir bien por no contar con suficiente cantidad de datos para entrenar, validar, y luego pronosticar como se quisiera. Esto no sucede, por ejemplo, en los trabajos de Kristjanpoller y Hernandez (2017) o Hamid y Iqbal (2004) que

usan datos en frecuencia diaria³³. Por el contrario, como se puede apreciar en la misma tabla, todos los modelos que involucran FIS necesitaron un tamaño de ventana menor, de entre 45 y 50 observaciones. Debido a esto, podemos aseverar que realizar pronósticos basado en modelos de Fuzzy Inference System comprende una mejor alternativa en aquellos modelos que involucran ANN para bases de datos que comprenden periodos relativamente reducidos.

No obstante lo anterior, el modelo de mejor desempeño encontrado resultó ser uno de especificación híbrida. El tipo de modelo, que resultó ser de la forma A-GARCH-FIS, es una muestra clara de sinergia entre los mejores modelos adaptativos encontrados, de forma independiente, para el caso paramétrico y no-paramétrico: el modelo A-GARCH, y el modelo A-FIS, respectivamente. El valor de $HMSE = 0.1304$ comprende una mejora de un 22% en el desempeño, comparado al valor de $HMSE = 0.1652$ del mejor modelo paramétrico. Adicionalmente, este es significativamente diferente y mejor que el resto de los modelos dado que el valor-p del MCS es 1. Por tanto, en este caso al menos, se puede afirmar que la inclusión de inteligencia artificial propia de modelos no-paramétricos (como los considerados en este estudio) permite mejorar pronósticos. Y además, el valor óptimo para el parámetro del mejor modelo, con un radio = 0.2825 (un valor muy específico), resalta la importancia de estimar este modelo adaptativamente al no ser posible de obtener tal valor de otro modo.

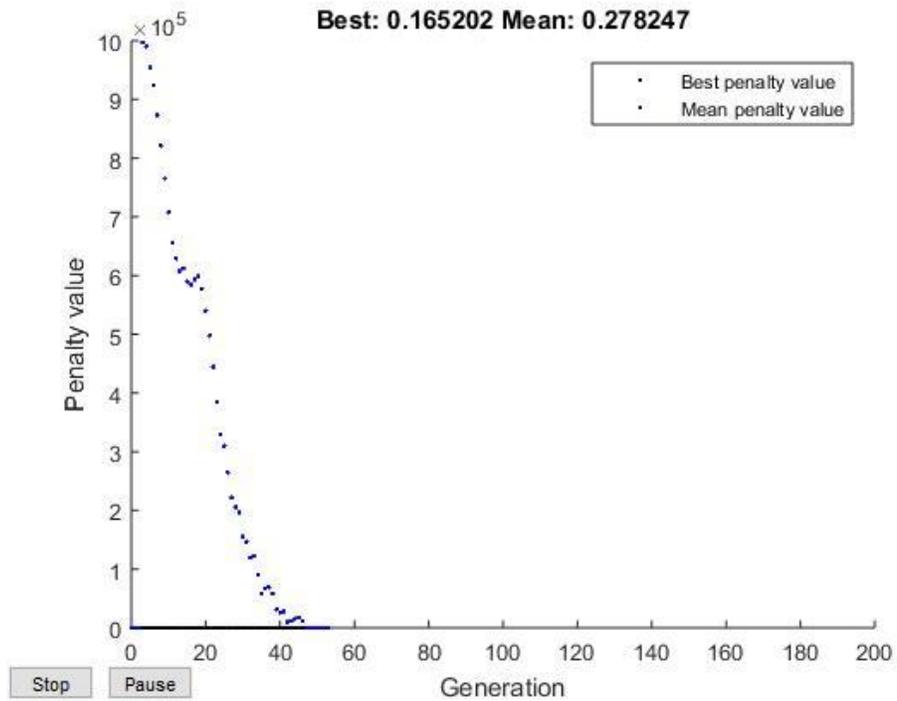
A nivel general, se puede afirmar que aquellos modelos que involucraron el uso de FIS resultaron ser estadísticamente mejores que el resto según MCS, para un nivel de confianza del 90%. En tal categoría, el mayor nivel de mejora se percibe en el desempeño

³³ A modo de comparación, por ejemplo, Hamid y Iqbal (2004) dispone de 2531 observaciones para efectuar pronósticos con ANN.

del modelo A-ARIMAX-FIS, reportando alrededor de un 40%. Su MCS p-value de 0.684 también lo ubica en una categoría diferente al resto de los modelos en términos de desempeño, aunque sin poder superar al modelo A-GARCH-FIS. Por su parte, los peores niveles de mejora se apreciaron en el modelo A-GARCH-ANN, resultando insignificantes.

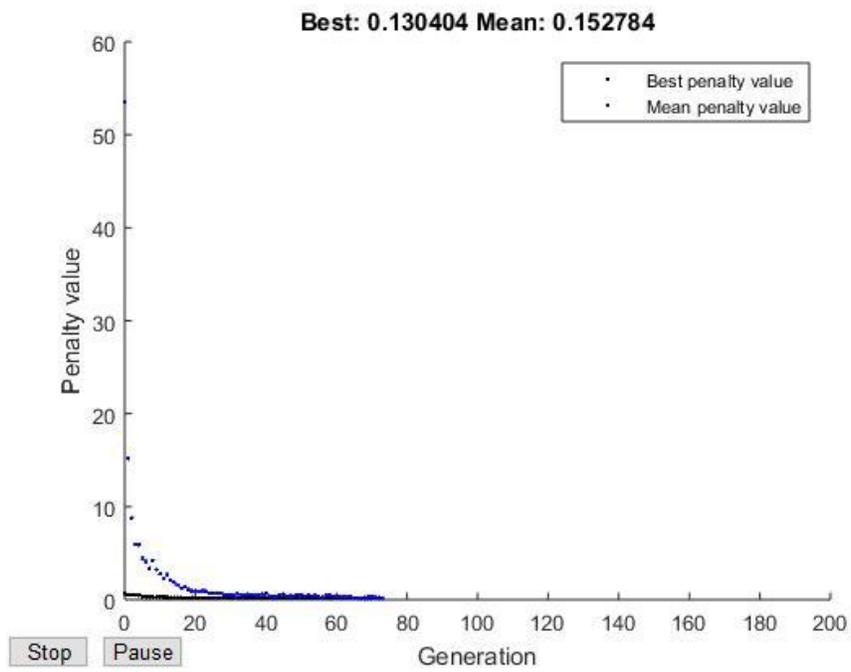
La eficiencia que posee el GA para alcanzar soluciones satisfactorias en el abordaje de problemas complejos, como los mostrados en este estudio, no puede pasar desapercibida. En el modelo ganador, la cantidad de generaciones que se necesitó para alcanzar la especificación óptima fue de 53 generaciones, donde cada una se conformó de 900 individuos (en el Paso 1). Es decir, se necesitó estimar, a lo menos, alrededor de 47.000 especificaciones distintas, y pronosticar 60 periodos hacia adelante con cada una de ellas, para hallar la óptima. Si bien esta cantidad pareciera ser alta, no tiene comparación con aquella que se necesitaría para estimar la cantidad total de posibles especificaciones, que es del orden de 9.64×10^{60} (para el mismo Paso 1). En este sentido, el GA permitió reducir la cantidad de modelos de manera abismante, siendo así fundamental para que el problema tenga una solución óptima en la práctica. La **Fig. 12a** y **Fig 12b** ilustra el desempeño del algoritmo genético a lo largo de las generaciones, tanto para el Paso 1 como Paso 2, respectivamente:

Figura 12a. Desempeño del algoritmo genético en el Paso 1.



Fuente: Elaboración propia.

Figura 12b. Desempeño del algoritmo genético en el Paso 2.

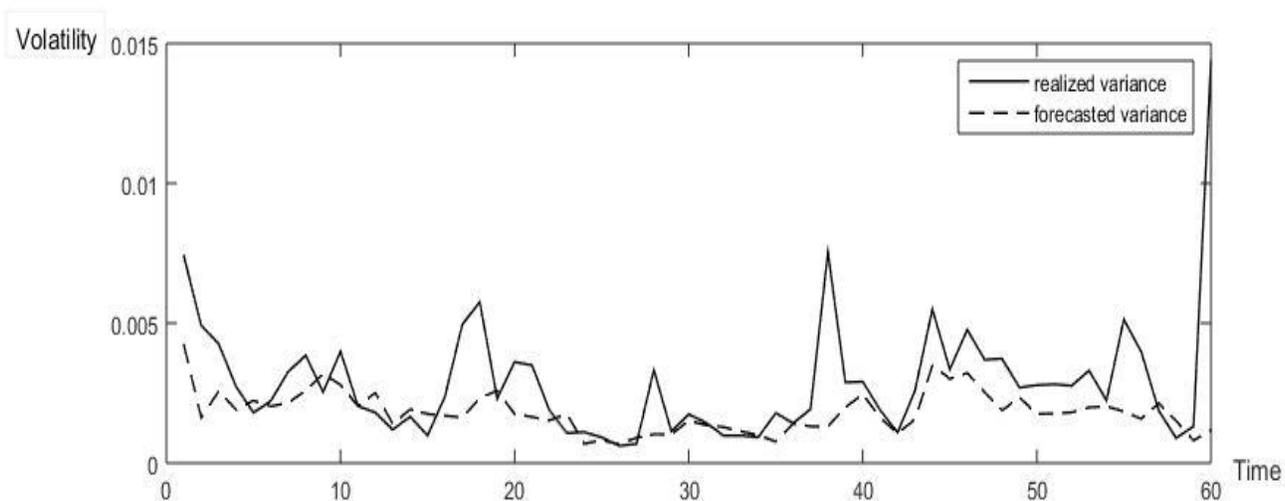


Fuente: Elaboración propia.

En su Paso 2, el algoritmo genético converge después de 73 generaciones lo que, dado un tamaño de población de 900 individuos por generación, implica estimar una cantidad de individuos asociados a 65.700 modelos.

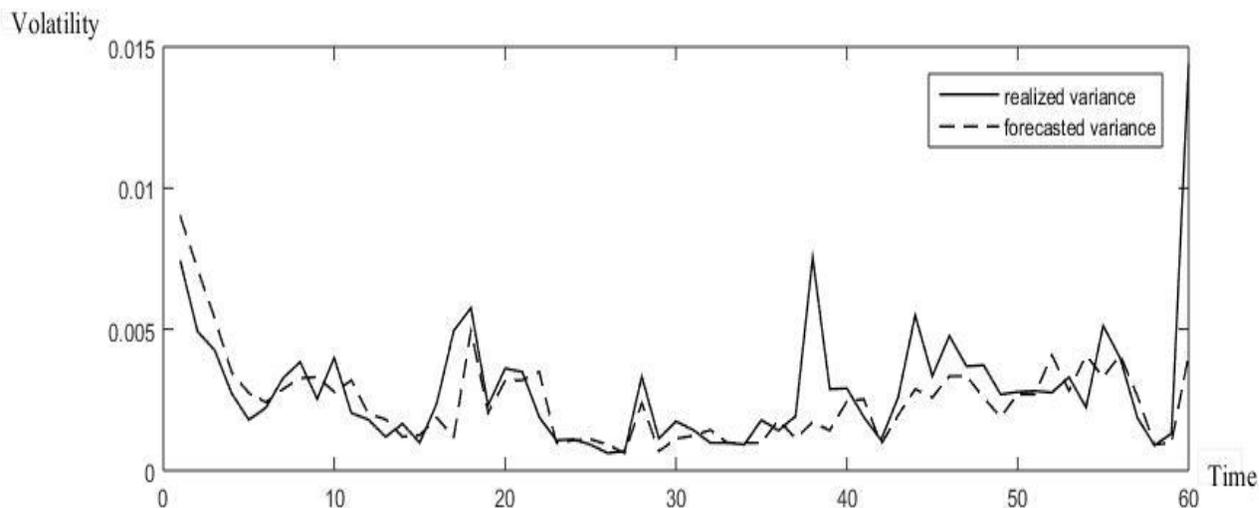
Respecto a los pronósticos en sí, la **Fig. 13a** y **Fig. 13b** muestran gráficamente los valores reales versus los pronosticados por el mejor modelo encontrado, en primer lugar para su Paso 1, y luego para su Paso 2 respectivamente (recordar que el modelo número 2 corresponde a aquel modelo utilizado en el Paso 1 del mejor modelo encontrado).

Figura 13a. Comparación entre volatilidad observada, y volatilidad pronosticada por el modelo A-GARCH.



Fuente: Elaboración Propia.

Figura. 13b. Comparación entre volatilidad observada, y volatilidad pronosticada por el modelo A-GARCH-FIS.



Fuente: Elaboración Propia.

A partir de las figuras anteriores, se puede apreciar gráficamente, en general, cómo el modelo híbrido permite captar de mejor manera el comportamiento observado de la volatilidad, destacándose el seguimiento de cambios abruptos en diversos periodos de tiempo. Esto principalmente se basa en las cualidades de modelación no-lineal que ofrece, en este caso, el sistema de inferencia difusa.

Finalmente, con respecto al rol que tienen los rezagos de las variables explicativas, las cuales sirvieron como input para alimentar los modelos, efectivamente existe una capacidad de predicción relevante en ellas, al menos en el contexto de modelos híbridos cuando se consideran de un modo adaptativo. Si bien en modelos paramétricos pareciera también existir un poder predictivo en variables explicativas (reflejado en el buen desempeño del modelo A-GARCH), no resulta ser determinante dicha capacidad cuando se compara con aquel modelo A-GARCH auto-explicado, pues ambos desempeños son similares (con un valor de HMSE de ambos alrededor de 0.16). La **Tabla 3** muestra qué

variables, de modo específico, resultaron ser relevantes según los resultados del mejor modelo adaptativo, A-GARCH-FIS. Es necesario mencionar que la inexistencia de algunas variables en la **Tabla 3a** (que sólo está relacionada al modelo en Paso1), se debe ya sea por no ser consideradas por los resultados del Teraesvirta's test, o porque el modelo adaptativo arrojó que tales variables no tienen valor adicional para efectos de pronóstico. De modo contrario, la inexistencia de variables en la **Tabla 3b** se debe sólo a la segunda causa, pues como se explicó en la metodología, el Teraesvirta's test no es utilizado para realizar el Paso 2 para al modelo no-paramétrico.

Si bien no se desea asesorar una explicación económica sobre por qué esas variables fueron seleccionadas por el modelo adaptativo, sí podemos afirmar que esta selección específica de variables (y no consecutiva en rezagos) permite alcanzar los mejores pronósticos, lo cual es posible gracias a la característica de adaptación de los modelos adquirida resaltando nuevamente su importancia. Por ejemplo, si el modelo ganador hubiese considerado sólo 2 rezagos de cada variable explicativa para ser estimado, tal modelo especificado (que es no-adaptativo) no llevaría a los mejores pronósticos. Para el modelo comprendido en el Paso 1, un total de 38 variables explicativas son seleccionadas, de las 78 posibles consideradas en la alimentación de ese modelo. Por el lado del modelo comprendido en el Paso 2, 45 variables explicativas fueron seleccionadas de un total de 85. En este último, cabe notar también que 8 de los 11 mejores modelos fueron seleccionados, a fin de complementar los pronósticos del mejor modelo encontrado en Paso 1; lo anterior, permite inferir que las especificaciones detrás de esos modelos contienen poder predictivo que no contemplaba el primero.

Tabla 3a. Variables explicativas óptimas resultantes del mejor modelo (retornos y varianzas)

VARIABLES DE RETORNOS	VARIABLES DE VARIANZA
1 AR ^a : 1,8,10,12	1 Constante
2 MA ^b : 1,2,8,9,10,12	2 Rezagos VCU: 4,5
3 Rezagos RMCI: 4,7,8,10,12	3 Rezagos VMCI: 4,5,6,8,10,11,12
4 Rezagos RINV: 2,4,5,8	4 Rezagos VINV: 5,9
5 Rezagos RIPICH: 2,3,5,7,9,12	

a. Términos autoregresivos.

b. Términos de media móvil.

RMCI: Retornos del Major Currency Index, RINV: Retornos de Inventarios, RIPICH: Retornos del índice de producción industrial de China, VCU: Varianza de los retornos del cobre spot, VMCI: Varianza de los retornos del Major Currency Index, VINV: Varianza de los retornos de Inventarios.

Tabla 3b. Variable explicativas óptimas resultantes del mejor modelo (sólo varianzas)

VARIABLES DE VARIANZA
1. ^a M: M2, M5, M6, M7, M9, M10, M11, M12
2. Rezagos VCU: 1,2,4,6,8,10,11,13,15,16,18,19,20
3. Rezagos VMCI: 2,3,5,6,7,9,11,14,16,18,19,21,23
4. Rezagos VINV: 3,5,7,8,11,12,15,16,17,19
5. Rezagos VFUT: 1

a. M se refiere a los mejores modelos entregados por el Paso 1. El número acompañante corresponde al orden del ranking que pertenece.

VCU: Varianza de retornos spot del cobre, VMCI: Varianza de los retornos del Major Currency Index, VINV: Varianza de los retornos de Inventarios, VFUT: Varianza de los retornos de futuros del cobre.

7. CONCLUSIONES

El objetivo principal abordado en este paper fue realizar pronósticos mensuales para la volatilidad del precio del cobre, algo que, si bien posee importantes consecuencias prácticas para diversos participantes de mercado, es un tópico que no ha sido asesorado previamente ni siquiera en otros mercados spot de *commodities*. Si bien existen numerosas investigaciones que abordan el objetivo anterior para otras frecuencias temporales, la menor cantidad de agentes interesados en aquello limita las consecuencias prácticas que esos estudios poseen.

Para efecto de lograr el objetivo anterior, se propuso un set de modelos conformado por aquellos de tipo paramétrico (en particular, modelos de series de tiempo), otros de tipo no-paramétrico (específicamente, utilizados en el área de soft computing), y combinaciones híbridas entre estos. Todos ellos han sido estimados bajo un contexto adaptativo, a fin de determinar qué especificación dentro de un modelo lleva a la obtención de mejor desempeño en sus pronósticos. Se ha demostrado que la capacidad de adaptación en estos modelos tanto en variables explicativas, como en el tamaño de ventana y en parámetros de configuración inherentes, es indispensable para alcanzar robustez y resultados satisfactorios. Mas aún, cuando estos parámetros muchas veces son asumidos comúnmente de forma fija.

Por su parte, el uso de inteligencia artificial ha sido determinante para la obtención de tales resultados. En primer lugar, para mejorar los pronósticos efectuados por otros modelos, llevando a aumentos de 20% y hasta 40% en el poder predictivo de estos en el caso de Fuzzy Inference Systems (superando otros modelos ampliamente utilizados); y por otro lado, como herramienta complementaria para resolver problemas complejos, como el uso del Teraesvirta's test para detectar relaciones no-lineales y el empleo de GAs para la

resolución de problemas de gran tamaño. En particular, el rol que el algoritmo genético ha tenido en esta investigación en materia de pronóstico ha sido clave para explotar el contexto adaptativo en los modelos propuestos de un modo eficiente.

Relacionado a lo anterior, la consideración de fundamentos teóricos también ha comprendido una fuente determinante en el propósito de alcanzar el objetivo mencionado. En este sentido, la correcta selección de variables que podrían funcionar de proxy para reflejar tales fundamentos, complementada con una metodología coherente, es indispensable más aún cuando existen limitaciones respecto a la data disponible. Sin embargo, profundizar en la asesoría de por qué determinadas variables explicativas conducen a mejores desempeños en el pronóstico, en comparación con otras, no es un asunto simple de realizar. Esto más aun considerando la característica de “blackbox” que tienden a poseer las herramientas de soft computing encargadas de efectuar pronósticos. De acuerdo a estas últimas ocupadas en esta investigación, los Fuzzy Inference Systems resultaron comprender una fuente de gran consideración para efectuar problemas de pronóstico en aquellas bases de datos no muy numerosas.

En materia de estudios futuros, podrían abordarse distintas opciones entre las posibles de mencionar. En primer lugar, desde un punto de modelos a utilizar, puede considerarse el utilizar una especificación diferente para la distribución de probabilidad de las innovaciones presentes en el modelo GARCH, las cuales se han supuesto gaussianas para esta investigación. Otras alternativas (como la distribución *t-student*) podrían abarcarse también en este contexto adaptativo, bajo la mismo fundamento de codificación utilizado en el presente trabajo y aprovechando la plataforma del mismo software encargado de ajustar tales modelos. Esta misma ventaja puede utilizarse (con el mismo propósito de modelo adaptativo) para estimar diferentes familias de modelos GARCH, en

especial el modelo GARCH-M, lo cual genera expectativa para la mejora de resultados al existir una gran gama de estos y aumentar las posibilidades de encontrar un mejor modelo, más aún cuando existen diferentes especificaciones cuyas cualidades podrían explotarse y explicar mejor el fenómeno de interés.

Otras especificaciones, como los MGARCH, también son prometedores de explorar (específicamente la especificación BEKK) al permitir la posibilidad de inclusión de covarianzas entre variables, en la ecuación de varianza condicional. No obstante, se debe resguardar tanto la complejidad en la implementación como la cantidad de variables a incluir de modo de eludir problemas de convergencia en la estimación de parámetros del modelo. Por otro lado, en materia de técnicas de inteligencia artificial, se puede mencionar la implementación de otro tipo de redes neuronales (como las recurrentes o convolucionales) para revisar si esta metodología puede mantenerse competitiva para pronósticos que involucran reducida cantidad de observaciones, como lo fue en este estudio. A su vez, en el propósito de profundizar en la eficiencia de la capacidad de adaptación de los modelos, la utilización de la herramienta *Particle Swam Optimization* (PSO) es curiosa para fines de comparación con el desempeño de los algoritmos genéticos.

Por último, y desde un punto de vista más conceptual, explorar la aplicación de este trabajo para otros commodities minerales y otras frecuencias resulta interesante por las prometedores similitudes que pueden existir con las utilizadas aquí. En este sentido, tener en cuenta la disponibilidad y grado de influencia de las variables utilizadas en este estudio pero para otro commodity, y el mantener el enfoque en el corto plazo (como frecuencias trimestrales por ejemplo), resultan ser aspectos de gran consideración para tal efecto.

REFERENCIAS

- Aguilar, R., Valenzuela, M. & Rodríguez, J. (2015). Genetic algorithms and Darwinian approaches in financial applications: A survey. *Expert Systems with Applications*, 42, 7684 – 7697.
- Angelov, P. P., & Filev, D. P. (2004). An approach to online identification of Takagi–Sugeno fuzzy models. *IEEE Transactions on Systems, Man and Cybernetics—part B*, 34(1), 484–498.
- Armstrong, J., Makridakis, Schultz, Gardney & Fischhoff (1986). The Ombudsman: Research on Forecasting: A quarter-Century Review, 1960-1984. *Interfaces*, 16 (January, -February). 89-109.
- Arouri, M., Lahiani, A., Lévy, A. & Nguyen, D. (2012). Forecasting the conditional volatility of oil spot and futures prices with structural breaks and long memory models. *Energy Economics*, 34, 283 – 293.
- Arratia, A. (2014). *Computational Finance. Atlantis Studies in Computational Finance and Financial Engineering*, vol. 1. Paris: Atlantis Press.
- Atsalakis, G. & Valavanis, K. (2009a). Surveying stock market forecasting techniques – Part II: Soft computing methods. *Expert Systems with Applications*, 36, 5932 – 5941.
- Atsalakis, G. & Valavanis, H. (2009b). Forecasting stock market short-term trends using a neuro-fuzzy based methodology. *Expert Systems with Applications*, 36, 10696 – 10707.
- Basher, S. & Sadorsky, P. (2016). Hedging emerging market stock prices with oil, gold, VIX, and bonds: A comparison between DCC, ADCC and GO-GARCH. *Energy Economics*, 54, 235 – 247.
- Bentes, S. (2015a). Forecasting volatility in gold returns under the GARCH, IGARCH and FIGARCH frameworks: New evidence. *Physica A*, 438, 355 – 364.
- Bentes, S. (2015b). A comparative analysis of the predictive power of implied volatility indices and GARCH forecasted volatility. *Physica A*, 424, 105 – 112.
- Boland, L. (2017). *Equilibrium models in economics*. New York: Oxford University Press.
- Bollerslev, T. (1986). Generalized autoregressive heteroskedasticity. *Journal of Econometrics*, 52, 307 – 327.

- Bollerslev, T., Hood, B., Huss, J. & Pedersen, L. (2017). Risk Everywhere: Modeling and Managing Volatility.
- Boyacioglu, M. & Avci, D. (2010). An Adaptive Network-Based Fuzzy Inference System (ANFIS) for the prediction of stock market return: The case of the Istanbul Stock Exchange. *Expert Systems with Applications*, 37, 7908 – 7912.
- Brunetti, C. & Gilbert, C. (1995). Metals price volatility, 1972-95. *Resources Policy*, 21, 237 – 254.
- Brunetti, C., Büyüksahin, B. & Harrys, J. (2016). Speculators, prices and market volatility. *Journal of Financial and Quantitative Analysis*, 51, 1545 – 1574.
- Cashin, P. & McDermott, C. IMF Econ Rev (2002) 49: 175.
<https://doi.org/10.2307/3872481>
- Chang, P. & Fan, C. (2008). A hybrid system integrating a wavelet and TSK fuzzy rules for stock price forecasting. *IEEE Transactions on Systems, Man and Cybernetics – Part C: Applications and Reviews*, 38, No. 6.
- Chang, P. & Liu, C. (2008). A TSK type fuzzy rule based system for stock price prediction. *Expert Systems with Applications*, 34, 135 – 144.
- Chen, Y., He, K. & Chuan, Z. (2016). A novel grey wave forecasting method for predicting metal prices. *Resources Policy*, 49, 323 – 331.
- Chkili, W., Hammoudeh, S. & Nguyen, D. (2014). Volatility forecasting and risk management for commodity markets in the presence of asymmetry and long memory. *Energy Economics*, 41, 1 – 18.
- Chiu, S. (1994). A cluster estimation method with extension to fuzzy model identification. In *Proceedings of the IEEE international conference on fuzzy systems*, vol. 2, 1240–1245.
- Comisión Chilena del Cobre. <https://www.cochilco.cl> . Accessed 20 November 2017.
- Cortazar, G., Gutierrez, S., & Ortega, H. (2015). Empirical performance of commodity pricing models: When is it worth to use a stochastic volatility specification? *Journal of Futures Markets*. Forthcoming.
- David, S., Kevin H. & Jaffrey, G. (2009). Commodity price volatility and world market integration since 1700. *The Review of Economics and Statistics*. MIT Press, vol. 93(3), 800-813.

- Davidson, R., Labys, W. & Lesourd, J. (1998). Wavelet analysis of commodity price behavior. *Computational Economics*, 11, 103 – 128.
- Degiannakis, S. & Potamia, A. (2017). Multiple-days-ahead value-at-risk and expected shortfall forecasting for stock indices, commodities and exchange rates: Inter-day versus intra-day data. *International Review of Financial Analysis*, 49, 176 – 190.
- De Gregorio, J., González, H. & Jaque, F. (2005). Fluctuaciones del dólar, precio del cobre y términos de intercambio. Central Bank of Chile Working Papers, No 310.
- Dooley, G. & Lenihan, H. (2005). An assessment of time series methods in metal price forecasting. *Resources Policy*, 30, 208 – 217.
- Donaldson, R. & Kamstra, M. (1997). An artificial neural network-GARCH model for international stock return volatility. *Journal of Empirical Finance*, 4, 17 – 46.
- Dornbusch, R. (1985). Policy and performance links between LDC debtors and industrial nations. *Brooking Papers on Economic Activity*, 2, 303 – 368.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 50, 987–1007.
- Fiordaliso, A. (1998). A nonlinear forecasts combination method based on Takagi–Sugeno fuzzy systems. *International Journal of Forecasting*, 14, 367 – 379.
- Fuertes, A., Izzeldin, M. & Kalotychou, E. (2009). On forecasting daily stock volatility: The role of intraday information and market conditions. *International Journal of Forecasting*, 25, 259 – 281.
- Garcia, D. & Guzman, J. (2017). Price volatility and reversión rate in commodity markets. Working Paper, Mining Department, Pontifical Catholic University of Chile.
- Garcia-Cicco, J. & Montero, R. (2011). Modeling copper price: a regime-switching approach. Central Bank of Chile Working Papers, No. 613.
- Gargano, A. & Timmermann, A. (2014). Forecasting commodity price indexes using macroeconomic and financial predictors. *International Journal of Forecasting*, 30, 825 – 843.
- Gilbert, C. (1989). The impact of exchange rates and developing country debt on commodity prices. *The Economic Journal* 99, 773-784.

- Guiassi, M., Saidane, H. & Zimbra, D. (2005). A dynamic artificial neural network model for forecasting time series events. *International Journal of Forecasting*, 21, 341 – 362.
- Guzman, J.I. (2004). Precio e inventarios de cobre y recortes de produccion. ¿Minería para siempre? Foro en Economía de Minerales - Vol I. Santiago, Chile: Ediciones Universidad Católica de Chile.
- Goonatilake, S., Campbell, J., & Ahmad, N. (1995). Genetic-fuzzy systems for financial decision making. In *Advances in fuzzy logic, Neural networks and genetic algorithms*. Berlin: Springer, 202–223.
- Guresen, E., Kayakutlu, G. & Daim, T. (2011). Using artificial neural network models in stock market index prediction. *Expert systems with Applications*, 38, 10389 – 10397.
- Hamid, S. and Iqbal, Z. (2004). Using neural networks for forecasting volatility of S&P 500 Index future prices. *Journal of Business Research*, 57, 1116 – 1125.
- Hansen, P., Lunde, A. & Nason, J. (2003). Choosing the best volatility models: The model confidence set approach. *Oxford Bulletin of Economics and Statistics*, 65, 839 – 861.
- Hansen, P., Lunde, A. & Nason, J. (2011). The model confidence set. *Econometrica*, 79, 453–497.
- Hajizadeh, E., Seifi, A., Fazel, M. & Turksen, I. (2012). A hybrid modeling approach for forecasting the volatility of S&P 500 index return. *Expert Systems with Applications*, 39, 431 – 436.
- Huang, C. & Tsai, C. (2009). A hybrid SOFM-SVR with a filter-based feature selection for stock market forecasting. *Expert Systems with Applications*, 36, 1529 – 1539.
- Hull, J. (2015). *Options, futures and other derivatives*, 9th edn. New Jersey: Pearson Education.
- Humphreys, D. (2010). The great metals boom: A retrospective. *Resources Policy*, 35, 1 – 13.
- Humphreys, D. (2011). Pricing and Trading in Metals and Minerals. In: P. Darling (Ed). *SME mining engineering handbook*, 3rd edn. United States: Society for Mining, Metallurgy and Exploration, Inc.
- Humphreys, D. (2015). *The remaking of the mining industry*. New York: Palgrave Macmillan.

- Hung, J. (2011a). Applying a combined fuzzy systems and GARCH model to adaptively forecast stock market volatility. *Applied Soft Computing*, 11, 3938 – 3945.
- Hung, J. (2011b). Adaptive Fuzzy-GARCH model applied to forecasting the volatility of stock markets using particle swarm optimization. *Information Sciences*, 181, 4673 – 4683.
- Jaramillo, P. & Selaive, J. (2006). Actividad especulativa y precio del cobre. Central Bank of Chile Working Papers, No. 384.
- Jünger, M., Reinelt, G., & Rinaldi, G. (2001). Combinatorial Optimization – Eureka, You Shrink!. New York: Springer.
- Kang, S. & Yoon, S. (2013). Modeling and forecasting the volatility of petroleum futures prices. *Energy Economics*, 36, 354 – 362.
- Kriechbaumer, T., Angus, A., Parsons, D. & Rivas, M. (2014). An improved wavelet–ARIMA approach for forecasting metal prices. *Resources Policy*, 39, 32 – 41.
- Kristjanpoller, W. & Sierra, A. (2014). Relación entre el dólar, el precio del cobre y el IPSA en distintas escalas de tiempo: una aproximación a través de wavelet. Central Bank of Chile.
- Kristjanpoller, W. & Minutolo, M. (2015). Gold price volatility: A forecasting approach using the Artificial Neural Network–GARCH model. *Expert Systems with Applications*, 42, 7245 – 7251.
- Kristjanpoller, W. & Minutolo, M. (2016). Forecasting volatility of oil price using an Artificial Neural Network-GARCH model. *Expert Systems with Applications*, 65, 233 – 241.
- Kristjanpoller, W. & Hernandez, E. (2017). Volatility of main metals forecasted by a hybrid ANN-GARCH model with regressors. *Expert Systems with Applications*, 84, 290 – 300.
- Labys, W. (1999). Modeling energy and mineral markets. New York: Springer.
- Labys, W., Kouassi, E. & Terraza, M. (2000). Short-term cycles in primary commodity prices. *The Developing Economies*, 38, 330 – 342.
- Labys, W. (2016). Modeling and forecasting primary commodity prices. New York: Routledge.
- Lagos, G. (2013). El incierto precio del cobre, <http://www.gustavolagos.cl/publicaciones.html> . Accessed on 27 November 2017.

- Lankhorst, M. M. (1996). Genetic algorithms in data analysis Groningen: s.n.
- Larsson, K., & Nossman, M. (2011). Jumps and stochastic volatility in oil prices: Time series evidence. *Energy Economics*, 33(3), 504-514.
- Levine, A., Hua, Y. & Richardson, M. (2016). Commodities for the long run. NBER Working Paper Series, No. 22793.
- Luna, I. & Ballini, R. (2011). Top-down strategies based on adaptive fuzzy rule-based systems for daily time series forecasting. *International Journal of Forecasting*, 27, 708 – 724.
- Makridakis, S., Hibon, M., & Moser, C. (1979). Accuracy of Forecasting: An Empirical Investigation. *Journal of the Royal Statistical Society*, 142(2), 97-145.
- Marra, S. (2015). Predicting volatility. Investment Research. Lazard Asset Management LLC, https://www.lazardassetmanagement.com/us/en_us/research-insights/investment-research/22430-Predicting-Volatility . Accessed on 27 November 2017.
- Matlab (2016). Version R2016a. The MathWorks Inc., Natick, Massachusetts, USA.
- Meller, P. (2013). La viga maestra y el sueldo de Chile. Santiago: Uqbar Editores.
- Ministry of Economic Affairs of Taiwan. Industrial Statistics. <http://dmz9.moea.gov.tw/gmweb/investigate/InvestigateDB.aspx?lang=E> . Accessed 27 November 2017.
- Olson, D. & Mossman, C. (2003). Neural network forecasts of Canadian stock returns using accounting ratios. *International Journal of Forecasting*, 19, 453 – 465.
- Otto, J., Andrews, C., Cawood, C., Dogget, M., Guj, P., Stermole, F., Stermole, J. & Tilton, J. (2006). Mineral royalty instruments. In: J. Otto et al. (Eds). *Mining Royalties. A Global Study of Their Impact on Investors, Government, and Civil Society*. Washington: The World Bank.
- Panapakidis, I. & Dagoumas, A. (2017). Day-ahead natural gas demand forecasting based on the combination of wavelet transform and ANFIS/genetic algorithm/neural network model. *Energy*, 118, 231 – 245.
- Patton, A. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160, 246 – 256.
- Pindyck, R. (2004). Volatility and commodity price dynamics. *Journal of Futures Markets*, 24(11), 1029-1047.

- Pindyck, R. & Rubinfeld, D. (2017). *Microeconomics*, 8th edn. London: Pearson Education.
- Pincheira, P.M. (2000). The price-stock relationship in the copper market: a surprising approach. Cochilco, Santiago, Chile.
- R Development Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>
- Radetzki, M. (2008). *A handbook of primary commodities in the global economy*. Cambridge University Press, New York.
- Reinhart, C. *IMF Econ Rev* (1991) 38: 506. doi.org/10.2307/3867156
- Ridler, D., & Yandle, C. (1972). A Simplified Method for Analyzing the Effects of Exchange Rate Changes on Exports of a Primary Commodity. *Staff Papers (International Monetary Fund)*, 19(3), 559-578. doi:10.2307/3866417
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986a). Learning internal representation by back-propagating errors. In: D. E. Rumelhart, & J. L. McClelland. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 1. Boston: MIT Press.
- Rumelhart, D., Hinton, G., & Williams, R. (1986b). Learning representations by back-propagating errors. *Nature*, 323, 533– 536.
- Sánchez, D. (2007). Precio del cobre, tipo de cambio y costos de producción. Tesis de Magíster. Santiago, Pontificia Universidad Católica de Chile.
- Sánchez, F., de Cos, F., Suárez, A., Krzemien, A. & Riesgo, P. (2015). Forecasting the COMEX copper spot price by means of neural networks and ARIMA models. *Resources Policy*, 45, 37 – 43.
- Schmitt, L. (2001). Theory of genetic algorithms. *Theoretical Computer Science*, 251, 1 – 61.
- Shahab, A. (2014). *Data-driven modeling: Using MATLAB in Water Resources and Environmental Engineering*. New York: Springer.
- Sheta, A. (2006). Software effort estimation and stock market prediction using Takagi-Sugeno fuzzy models. *IEEE International Conference on Fuzzy Systems*.
- Shumway, R. & Stoffer, D. (2011). *Time series analysis and its applications*, 3rd edn. New York: Springer.

- Singh, N. & Mohanty S. (2017) Short term price forecasting using adaptive generalized neuron model. In: S. Bhatia et al. (Eds). *Advances in Computer and Computational Sciences. Advances in Intelligent Systems and Computing*, vol 553. Singapore: Springer.
- Takagi, T. & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man and Cybernetics, SMC-15*, 116 – 132.
- Tay, F. & Cao, L. (2002). Modified support vector machines in financial time series forecasting. *Neurocomputing*, 48, 847 – 861.
- Teraesvirta, T., Lin, C. & Granger, C. (1993). Power of the neural network linearity test. *Journal of Time Series Analysis*, 14, 209–220.
- Tilton, J., Humphreys, D. & Radetzki, M. (2011). Investor demand and spot commodity prices. *Resources Policy*, 36, 187 – 95.
- Tilton, J. (2014). Cyclical and secular determinants of productivity in the copper, aluminum, iron ore, and coal industries. *Mineral Economics*, 27, 1 – 19.
- Tilton, J. & Guzmán, J. (2016). *Mineral Economics and Policy*. New York: Resources for the Future Press.
- Triantafyllou, A., Dotsis, G. & Sarris, A. (2015). Volatility forecasting and time-varying variance risk premiums in grains commodity markets. *Journal of Agricultural Economics*, 66, 329 – 357.
- Trolle, A. & Schwartz, E. (2009). Unspanned stochastic volatility and the pricing of commodity derivatives. *Reviews of Financial Studies*, 22(11), 4423 – 4461.
- Truck, S. & Liang, K. (2012). Modelling and forecasting volatility in the gold market. *International Journal of Banking and Finance*, vol. 9.
- Tsay, R. (2005). *Analysis of financial time series*. Chicago: John Wiley.
- Tully, E. & Lucey, B. (2007). A power GARCH examination of the gold market. *Research in International Business and Finance*, 21, 316 – 325.
- U.S Federal Reserve Bank of St. Louis. Economic Data. <https://fred.stlouisfed.org/series> . Accessed November 27 2017
- Wagner, N., Michalewicz, Z., Khouja, M. & McGregor, R. (2007). Time Series Forecasting for Dynamic Environments: The DyFor Genetic Program Model. *IEEE Transactions on Evolutionary Computation*, 11, No. 4.

- Wei, L. (2013). A GA-weighted ANFIS model based on multiple stock market volatility causality for TAIEX forecasting. *Applied soft computing*, 13, 911 – 920.
- Zhang, G. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159 – 175.

APÉNDICE. Cantidad de especificaciones posibles para modelo GARCH con variables explicativas

Sea V el conjunto de variables posibles que pueden utilizarse de entrada en un modelo GARCH con variables exógenas, tanto en su ecuación de media como de varianza. Denotemos $V^m \subseteq V$ el conjunto de variables que sólo sirven de entrada para la Eq. (1a), y $V^v \subseteq V$ conjunto de variables de entrada para la Eq. (1c). Luego, $V = V^m \cup V^v$.

Según lo explicado en la sección 3.2, se desprende que la cantidad de elementos (variables) que poseen los subconjuntos anteriores, es decir la cardinalidad de cada uno, está dada por: $|V^m| = 123$, $|V^v| = 75$. A su vez, la cantidad de combinaciones sin repetición con sólo i elementos, con $i < n$, donde n es la cantidad total de elementos del conjunto, está dada por $\binom{n}{k} = n!/(k!(n-k)!)$.

Luego, si consideramos todas las combinaciones posibles para cada elemento de n en los conjuntos V^v y V^m anteriores, que llamaremos U_v y U_m respectivamente, se tiene:

$$U_v = \sum_{i=1}^{75} \binom{75}{i} = 2^{75} \quad , \quad U_m = \sum_{i=1}^{123} \binom{123}{i} = 2^{123}$$

las cuales equivalen a la cantidad de distintas maneras con las que se pueden especificar las ecuaciones Eq. (1a) y Eq. (1c), de modo independiente respectivamente. Entonces, combinando las especificaciones de cada ecuación, y sólo considerando variables exógenas, cada modelo puede especificarse una cantidad igual a $U_x = (2^{75})(2^{123})$. Finalmente, dado que el modelo GARCH puede ajustar sus órdenes p y q hasta sólo un máximo de 4 (GARCH(1,1), GARCH(3,1), ..., GARCH(4,4)), lo cual da origen a 24 formas distintas, se tiene una cantidad total de $U = (2^{75})(2^{123})(24) = \mathbf{9.64 \times 10^{60}}$ formas posibles.