



UNIVERSIDAD NACIONAL DE COLOMBIA

Una aplicación de redes neuronales y modelos autorregresivos para la estimación de valores de referencia de swaps

Juan Manuel Posada Zuluaga

Universidad Nacional de Colombia
Escuela de Estadística
Medellín, Colombia
2023

Una aplicación de redes neuronales y modelos autorregresivos para la estimación de valores de referencia de swaps

Juan Manuel Posada Zuluaga

Tesis o trabajo de grado presentada(o) como requisito parcial para optar al título de:
Magister en Ciencias-Estadística

Director(a):
Ph.D. César Augusto Gómez Vélez

Universidad Nacional de Colombia
Escuela de Estadística
Medellín, Colombia
2023

A mi familia por siempre brindarme un apoyo incondicional, a mi novia Laura por siempre llenarme de luz y darme fuerzas para cumplir este reto, y a mis compañeros de trabajo por retroalimentarme y brindarme motivación.

Agradecimientos

Agradezco a todas las personas que me apoyaron en el desarrollo de esta investigación.
A todos MUCHAS GRACIAS.

Resumen

Una aplicación de redes neuronales y modelos autorregresivos para la estimación de valores de referencia de swaps.

En este trabajo se realiza una aplicación de redes neuronales y modelos autorregresivos para la estimación del valor de referencia de un swap de tasa de interés teniendo en cuenta el ajuste de valoración por riesgo de crédito de la contraparte (CVA) y el ajuste de valoración de riesgo de crédito de la entidad (DVA) entre dos emisores del sector financiero local. Inicialmente, se utiliza como base del análisis la curva forward IBR, de la cual, se generan diez series de tiempo, cada una relacionada con los periodos de liquidación de los pagos del swap. Para cada serie, se ajusta un modelo de red neuronal y un modelo ARIMA-GARCH, y se evalúan sus respectivas métricas de prueba, con el objetivo de hacer comparaciones entre ellos. Luego, utilizando los modelos obtenidos, se realiza el pronóstico de la curva forward IBR para el siguiente día hábil. Así mismo, se estima la probabilidad de default en cada fecha de liquidación a partir de los respectivos Asset swap spread de los bonos corporativos de los emisores. Posteriormente, se estima el valor razonable del swap a partir del valor libre de riesgo y sus respectivos ajustes. Finalmente, los resultados muestran que para el rango de fechas analizado la valoración estimada a partir del modelo ARIMA-GARCH presenta un menor error de pronóstico en comparación con el modelo de red neuronal, lo que sugiere una mayor precisión en la estimación del valor de referencia del swap.

Palabras clave: Swap, ARIMA, GARCH, RNN, CVA, DVA.

Abstract

A Neural Networks and Autoregressive Models Application for Estimating Reference Values of Swaps.

This research applies neural networks and autoregressive models to estimate the reference value of an interest rate swap, considering the counterparty credit risk adjustment (CVA) and the entity's credit risk adjustment (DVA) between two local financial sector issuers. The analysis starts with the forward IBR curve as the foundation, from which ten time series are generated, each associated with the payment settlement periods of the swap. For each series, a neural network model and an ARIMA-GARCH model are fitted, and their respective test metrics are evaluated for comparison. Subsequently, using the obtained models, the forecast of the forward IBR curve is conducted for the next business day. Additionally, the default probability is estimated for each settlement date based on the respective Asset Swap Spreads of the issuers' corporate bonds. Then, the fair value of the swap is estimated considering the risk-free value and the corresponding adjustments. Finally, the results indicate that, for the analyzed date range, the valuation derived from the ARIMA-GARCH model exhibits lower forecast error compared to the neural network model, suggesting higher precision in estimating the swap's reference value.

Keywords: Swap, ARIMA, GARCH, RNN, CVA, DVA.

Contenido

Agradecimientos	vii
Resumen	ix
1. Introducción	1
2. Objetivos	3
2.1. Objetivo general	3
2.2. Objetivos específicos	3
3. Marco teórico	4
3.1. Swap de tasa de interés	4
3.2. Redes neuronales artificiales	6
3.3. Red neuronal recurrente	10
3.4. Modelos Autorregresivos	12
3.5. Modelos GARCH	13
3.6. Ajustes de valoración (xVA)	14
4. Estudio de antecedentes	16
5. Presentación de resultados	19
5.1. Descripción de los datos y supuestos	19
5.2. Aplicación modelos ARIMA-GARCH	25
5.3. Aplicación redes neuronales recurrentes	33
5.4. Estimación de las probabilidades de default	37
5.5. Estimación Swap libre de riesgo	44
5.6. Estimación CVA/DVA	46
6. Conclusiones y recomendaciones	51
6.1. Conclusiones	51
6.2. Recomendaciones	52
A. Anexo: Modelos ARIMA-GARCH	53
B. Anexo: RNN	63

C. Anexo: Características Emisiones	73
D. Anexo: Código Python RNN	74
Bibliografía	78

1. Introducción

La crisis económica del 2008 deterioró la confianza del público en los mercados de valores, especialmente en los productos derivados, ya que se convirtieron en el mecanismo a través del cual fue posible empaquetar y comercializar hipotecas de baja calidad, propagándolas en gran escala por todo el sistema financiero. Sin embargo, de acuerdo con Sabato et al. (2010), fue la mala gestión del riesgo de crédito lo que contribuyó en mayor medida a la crisis. Por tal motivo y como respuesta a la nueva realidad del mercado, las entidades reguladoras tanto locales como internacionales, introdujeron medidas claras para la estimación del riesgo de crédito en los derivados financieros, dentro de las cuales, se destaca el denominado ajuste de valoración por riesgo de crédito de la contraparte (CVA por sus siglas en inglés) y el ajuste de valoración de riesgo de crédito de la entidad (DVA por sus siglas en inglés).

En el mercado financiero colombiano, se negocian generalmente 4 tipos de productos derivados: opciones, swaps, forwards y futuros. Entre ellos, el swap de tasa de interés IBR (interés bancario de referencia) ha sido especialmente relevante debido a su importancia en el mercado monetario local y su crecimiento en términos de volumen de negociación y frecuencia de transacciones, según lo indicado por (BanRep, 2020). De acuerdo con Asobancaria (2012), este producto refleja de manera adecuada el costo del dinero a corto plazo y, por lo tanto, es una variable clave que influye en las decisiones de inversión y financiamiento de los participantes del mercado.

Ante esta situación, los profesionales financieros encargados de la gestión de activos han emprendido la tarea de modelar el comportamiento de las tasas de interés. Para lograrlo, han utilizado diversas metodologías, entre las que destacan los modelos ARIMA, los modelos ARIMA-GARCH y los modelos de difusión de tasa corta, como los planteados por Vasicek o Brennan-Schwartz. A pesar, de la amplia gama de herramientas disponibles para modelar el comportamiento de la tasa de interés y de los esfuerzos continuos para proponer modelos más sólidos, según Leung et al. (2000), todavía se carece de un modelo que supere significativamente la precisión de un modelo de caminata aleatoria. Debido a esta razón, en los últimos años ha habido un creciente interés en la adopción de la inteligencia artificial de última generación, en particular el uso del aprendizaje profundo, el aprendizaje automático y métodos no paramétricos como las redes neuronales para proyectar este tipo de variables.

En este estudio, se pretende aplicar redes neuronales para predecir la tasa de interés IBR y comparar sus resultados con los de los modelos ARIMA-GARCH. Además, se analizará el impacto de dichas predicciones en la valoración de un swap de tasa de interés, teniendo en cuenta los respectivos ajustes por riesgo crédito CVA/DVA.

2. Objetivos

2.1. Objetivo general

Estimar el valor de referencia de un swap IBR a cinco años entre dos emisores del sector financiero colombiano utilizando redes neuronales y modelos autorregresivos.

2.2. Objetivos específicos

- Predecir las tasas forward semestrales de la curva swap IBR para un periodo de cinco años a través de la aplicación de redes neuronales y modelos autorregresivos.
- Estimar la probabilidad de default para cada uno de los emisores locales a partir de sus correspondientes Asset-Swaps Spread (ASW).
- Estimar el valor de mercado de un swap libre de riesgo a cinco años utilizando la curva swap IBR pronosticada por las redes neuronales y los modelos autorregresivos.
- Calcular la exposición potencial crediticia futura de un swap IBR a cinco años entre dos emisores del sector financiero colombiano (CVA/DVA).

3. Marco teórico

En el ámbito financiero, existen diversos instrumentos con los cuales los inversionistas pueden sacar algún tipo de beneficio económico por medio de su negociación. Entre ellos, se encuentran los títulos de deuda, los títulos patrimoniales, las cuentas de ahorro, los depósitos a plazo fijo y una serie de activos aún más complejos conocidos como derivados. Los derivados se basan en contratos bilaterales, en los cuales dos partes se comprometen a realizar una actividad económica específica. En este estudio, nos centramos en un tipo de derivado en particular: la estimación de los valores de referencia de los swaps. Los swaps son ampliamente utilizados por los inversionistas como cobertura, con el objetivo de estabilizar sus flujos de ingresos y egresos que están naturalmente vinculados a tasas de interés variables o a flujos de dinero pactados en monedas diferentes a la moneda local. Sin embargo, el enfoque de esta investigación se centra en los swaps de tasas de interés y sus respectivos ajustes de valoración, que se definirán más adelante en esta sección. Por lo tanto, consideramos relevante brindar una definición conceptual y matemática del activo base de esta investigación.

3.1. Swap de tasa de interés

Es un instrumento financiero que consiste en el intercambio de una serie de pagos entre dos partes, a partir de distintas tasas de interés para diferentes periodos de tiempo futuros y sobre un valor nominal. En este sentido, se tiene una contraparte que recibe el pago en tasa variable y paga tasa fija y del otro lado se tiene la contraparte que recibe el pago en tasa fija y paga tasa variable. La definición anterior se puede expresar en términos más técnicos de acuerdo con Brigo and Mercurio (2006) de la siguiente forma:

Consideremos un horizonte de tiempo $T > 0$. Donde, para cada instante T_i en un conjunto de fechas previamente pactadas $\mathcal{T} := \{T_\alpha, \dots, T_\beta\}$ la contraparte pagadora de la tasa fija se compromete a desembolsar la cantidad

$$N\tau_i K,$$

que corresponde al resultado de multiplicar la tasa de interés fija K , el valor nominal N , y la porción de año τ_i entre T_{i-1} y T_i . De igual forma la contraparte pagadora de la tasa variable se compromete a desembolsar la cantidad

$$N\tau_i L(T_{i-1}, T_i),$$

que corresponde al resultado de multiplicar la tasa de interés variable $L(T_{i-1}, T_i)$, el valor nominal N , y la porción de año τ_i entre T_{i-1} y T_i . Bajo la descripción anterior asumimos que los pagos son realizados en las mismas fechas y por lo tanto los τ_i son iguales para ambas contrapartes.

La tasa variable cambia constantemente, y por tanto, se actualiza en cada periodo de tiempo, es decir, dado un T_α , la tasa se actualizará en las fechas $T_\alpha, T_{\alpha+1}, \dots, T_{\beta-1}$ y aplicará para los pagos en las fechas $T_{\alpha+1}, \dots, T_\beta$.

La expresión que da el valor para un swap de tasa de interés teniendo en cuenta todos los pagos y el valor del flujo total descontado en un tiempo $t < T_\alpha$ es expresado a continuación:

CONTRAPARTE PAGADORA DE LA TASA FIJA

$$\sum_{i=\alpha+1}^{\beta} D(t, T_i) N\tau_i (L(T_{i-1}, T_i) - K) \quad (3.1.1)$$

CONTRAPARTE PAGADORA DE LA TASA VARIABLE

$$\sum_{i=\alpha+1}^{\beta} D(t, T_i) N\tau_i (K - L(T_{i-1}, T_i)) \quad (3.1.2)$$

Donde el factor de descuento $D(t, T_i)$ es el mismo para cada una de las contrapartes y corresponde al factor de ajuste que hace equivalente una unidad monetaria pagada en T en el tiempo t , con $t < T$.

En el momento en que las partes acuerdan el swap, la tasa fija pactada se determina en función de las condiciones del mercado en ese momento. Se supone que, según la situación económica actual, el valor presente de los pagos fijos y variables es igual. Esto implica que no hay una ventaja inicial para ninguna de las partes y que el equilibrio entre las pérdidas y ganancias es neutral. Sin embargo, una vez que se ha establecido el swap y las condiciones del mercado cambian, una de las contrapartes obtendrá una ventaja sobre la otra dependiendo de cómo se mueva la tasa variable. Esto significa que puede haber una parte ganadora y otra perdedora en función de dichos movimientos.

Por lo anterior, para poder realizar una estimación del precio de un swap en un instante de tiempo t , bastará con estimar el comportamiento futuro de la tasa variable y acorde a dicha estimación aplicar las ecuaciones 3.1.1 si es la parte que paga la tasa fija o 3.1.2 si es

la parte que paga la tasa variable, teniendo en cuenta que los factores de descuento pueden ser conocidos de acuerdo con el panorama actual del mercado.

Ahora bien, estimar el comportamiento futuro de la tasa variable o, dicho de otro modo, el rendimiento esperado de una inversión en distintos periodos de tiempo (Caballero, 2015), ha sido tema de estudio durante muchos años y también se han utilizado una gran cantidad de técnicas estadísticas para ello. Tenemos entre ellas las metodologías propuestas por Nelson and Siegel (1987), Vasicek (1977), Hull and White (1993), entre otros. Sin embargo, algunos autores también han abordado el problema mediante metodologías no paramétricas como Vásquez and Velandia (2005) que utilizaron splines cúbicos, y Xiao (2019) utilizando herramientas de aprendizaje profundo como las redes neuronales.

El enfoque de esta investigación está centrado en una aproximación no paramétrica, más específicamente la aplicación de redes neuronales. De acuerdo con lo anterior se procede a explicar y exponer brevemente la teoría alrededor de dicha metodología.

3.2. Redes neuronales artificiales

Las redes neuronales artificiales son un esfuerzo por replicar, a través de modelos matemáticos, el comportamiento de las redes neuronales presentes en los sistemas biológicos. Como señala Matich (2001), se trata de sistemas computacionales complejos compuestos por unidades simples llamadas neuronas, organizadas en una estructura jerárquica de niveles. Tienen como objetivo procesar información, aprender de ella y detectar patrones complejos para así predecir su comportamiento futuro (Mondragón et al., 2021). En esta sección, examinaremos los conceptos más relevantes relacionados con las redes neuronales y cómo funcionan como sistemas inteligentes para resolver problemas complejos.

Neurona artificial

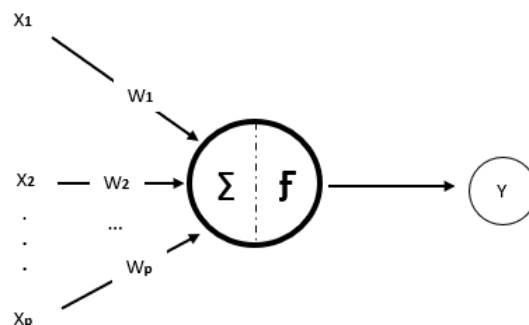


Figura 3.2.1.: Neurona artificial.

En la configuración de una red neuronal artificial, la unidad más básica es la neurona artificial. En la Figura 3.2.1, se muestra la representación gráfica de los elementos que componen esta unidad. Su objetivo principal es procesar la información de entrada $\mathbf{X} = (X_1, X_2, \dots, X_p)$ y generar una salida Y , que puede ser transmitida a otra neurona o ser la solución final del problema en cuestión.

Para lograrlo, la neurona utiliza un vector de pesos $\mathbf{W} = (W_1, W_2, \dots, W_p)$, que son asignados a cada variable de entrada. Estos pesos representan la importancia relativa de cada variable en la determinación del resultado final. La información de entrada \mathbf{X} y el vector de pesos \mathbf{W} se combinan mediante un producto escalar, conocido como entrada neta de la neurona.

$$\text{Entrada neta} = \sum_{i=1}^p x_i w_i \quad (3.2.1)$$

Es importante mencionar que los pesos se definen como parámetros del modelo y se ajustan durante el proceso de entrenamiento de la red neuronal.

Posteriormente, la neurona procesa la información mediante una función de activación \mathbf{F} , que proporciona el resultado final Y . La función de activación introduce no linealidad en la neurona, permitiendo que la red pueda aprender representaciones complejas y no lineales de los datos de entrada (Sierra Ramos, 2022).

$$Y = f(\text{Entrada Neta}) = f\left(\sum_{i=1}^p x_i w_i\right). \quad (3.2.2)$$

Funciones de activación

Según Pose (2009) la función de activación es la encargada de relacionar el estado de actividad de una neurona, es decir, esta función determina el valor de salida de la neurona en función de su entrada y establece si la neurona está completamente inactiva (con una salida de 0 o -1) o activa (con una salida de 1). A continuación, se presentan algunas de las funciones de activación más comunes:

- **Función sigmoide:**

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3.2.3)$$

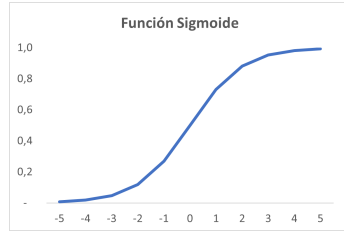


Figura 3.2.2.: Función sigmoide.

- Función tangente hiperbólica:

$$f(x) = \frac{1 - e^{-x}}{1 + e^{-x}} \quad (3.2.4)$$

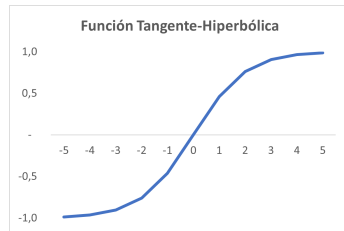


Figura 3.2.3.: Función Tangente hiperbólica.

- Función RELU:

$$f(x) = \max(0, x) \quad (3.2.5)$$

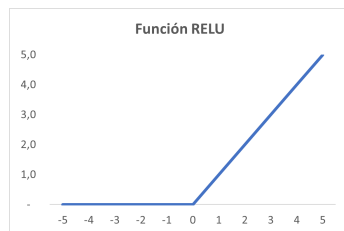


Figura 3.2.4.: Función RELU.

Hasta este punto, se ha proporcionado una explicación sobre la neurona artificial como elemento fundamental en una red neuronal. Sin embargo, es importante destacar que una red neuronal artificial está compuesta por una variedad de estas neuronas, organizadas en capas. La representación gráfica de la configuración más sencilla de una red neuronal se puede observar en la Figura 3.2.5. Allí intervienen los tres elementos mínimos requeridos que se conectan unos con otros en función de obtener un resultado para la variable respuesta Y . Inicialmente, la red neuronal presenta un primer nivel donde se encuentra el vector de entrada de P variables definido como $\mathbf{X} = (X_1, X_2, \dots, X_p)$, en un segundo nivel se encuentran las capas ocultas que se alimentan del vector de entradas previamente ingresado. En este nivel se procesa la información y se realizan los cálculos respectivos a partir de las neuronas o unidades y las funciones de activación previamente especificadas. Por último, se encuentra el nivel de salida donde se procesa la información resultante del nivel anterior y se obtiene una respuesta (James et al., 2021).

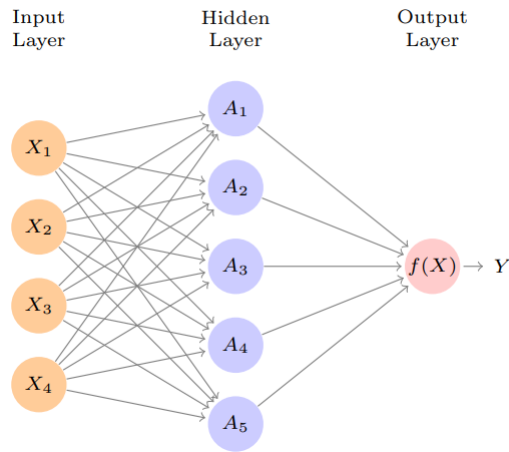


Figura 3.2.5.: Representación gráfica de una red neuronal. Tomado de James et al. (2021).

El diagrama en la Figura 3.2.5 se puede representar matemáticamente como:

$$f(X) = \beta_0 + \sum_{k=1}^K \beta_k h_k(X), \quad (3.2.6)$$

$$h_k(X) = g(w_{k0} + \sum_{j=1}^p w_{kj} X_j), \quad (3.2.7)$$

Donde \mathbf{X} es el vector de variables conocidas, $g(\cdot)$ es una función de activación no lineal, la cual, es definida previo a la construcción de la red (generalmente es una función sigmoide),

K hace referencia al numero de unidades ocultas y los parámetros β y w se estiman a partir de la minimización del error cuadrático medio.

$$\sum_{i=1}^n (y_i - f(x_i))^2. \quad (3.2.8)$$

3.3. Red neuronal recurrente

De la gran variedad de configuraciones existentes, se encuentran las redes neuronales recurrentes, caracterizadas principalmente por ser de naturaleza dinámica y por su capacidad para resolver distintos problemas computacionales como predicciones no lineales, tratamiento de secuencias y modelación de series de tiempo (Cruz et al., 2007), lo cual, es de gran utilidad para nuestro problema de investigación.

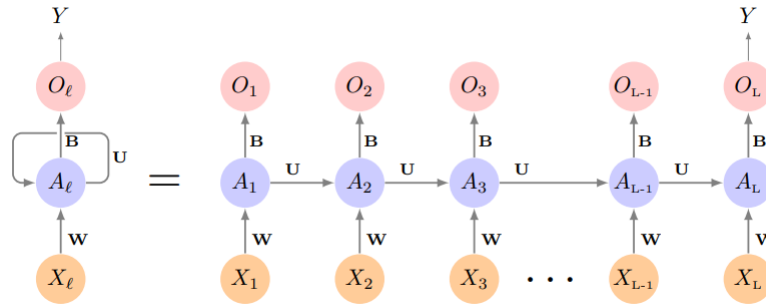


Figura 3.3.1.: Representación grafica de una red neuronal recurrente. Tomado de James et al. (2021).

La representación gráfica de la configuración más sencilla de una red neuronal recurrente se puede observar en la Figura 3.3.1. La configuración de este tipo de red está diseñada para procesarse de forma secuencial. En este caso, tenemos un primer nivel donde se encuentra una secuencia de vectores de entrada, denotados como $\{X_\ell\}_1^L$ con $\mathbf{X} = (X_1, X_2, \dots, X_L)$. En un segundo nivel se encuentran las capas ocultas secuenciales $\{A_\ell\}_1^L, \mathbf{A} = (A_1, A_2, \dots, A_L)$ que se alimentan de cada X_ℓ y del vector de activación $A_{\ell-1}$ del elemento previo de la secuencia. En este nivel se procesa la información y se realizan los cálculos respectivos a partir de funciones de activación previamente especificadas. Por último, se encuentra el nivel de salida donde se procesa la información resultante del nivel anterior y se obtiene una secuencia de predicción O_L (James et al., 2021). El modelo matemático de la red neuronal recurrente se expresa de la siguiente forma:

$$O_\ell = \beta_0 + \sum_{k=1}^K \beta_k A_{\ell k}, \quad (3.3.1)$$

$$A_{\ell k} = g(w_{k0} + \sum_{j=1}^p w_{kj} X_{\ell j} + \sum_{s=1}^K u_{ks} A_{\ell-1,s}), \quad (3.3.2)$$

Donde $X_{\ell j}$ es el vector de variables conocidas, $g(\cdot)$ es una función de activación no lineal, la cual, es definida previamente a la construcción de la red (generalmente es una función sigmoide), K hace referencia al numero de capas ocultas y las matrices de parámetros B , U y W con valores β , μ , y ω se estiman a partir de la minimización del error cuadrático medio.

$$\sum_{i=1}^n (y_i - o_{iL})^2. \quad (3.3.3)$$

RNN-LSTM

Las redes neuronales long-short term memory (RNN-LSTM) propuestas por Hochreiter and Schmidhuber (1997) son una extensión de las redes neuronales recurrentes clásicas. Estas redes están diseñadas con un módulo de memoria más sofisticado, lo que les permite capturar y retener información a corto y largo plazo durante el procesamiento de la información. Debido a esta capacidad, las RNN-LSTM son especialmente adecuadas para abordar tareas que involucran dependencias temporales complejas, como el análisis de las series de tiempo abordadas en esta investigación.

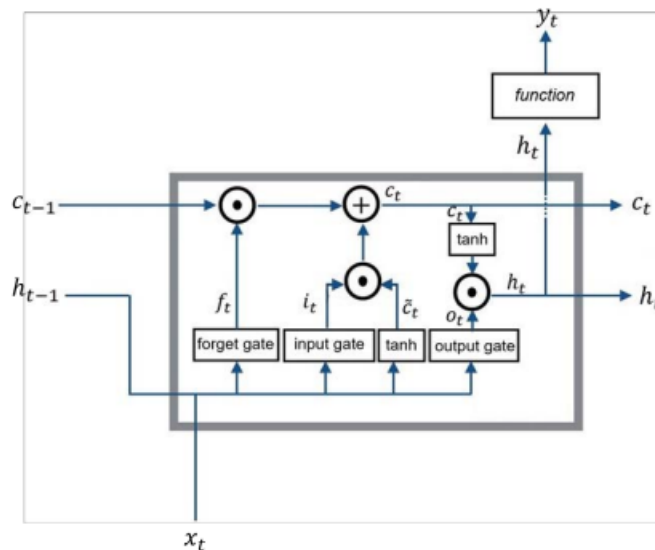


Figura 3.3.2.: Representación grafica RNN-LSTM. Tomado de Cárdenas et al. (2023).

La Figura 3.3.2 ilustra la estructura de una red LSTM. Esta estructura involucra tres puertas esenciales que regulan el flujo de información: la puerta de entrada (input gate), la puerta de olvido (forget gate) y la puerta de salida (output gate). Estas puertas interactúan entre sí de

tal manera que seleccionan y retienen información significativa de la memoria a lo largo del tiempo, a la vez que descartan la información que no contribuye a describir adecuadamente el comportamiento de la serie (Cárdenas et al., 2023). El modelo matemático de la red de acuerdo con Graves (2013) se expresa de la siguiente forma:

$$\begin{aligned}
 f_t &= \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \\
 i_t &= \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \\
 o_t &= \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \\
 \tilde{C}_t &= \tanh(W_C x_t + U_C h_{t-1} + b_C)
 \end{aligned} \tag{3.3.4}$$

Donde x_t representa las variables de entrada. W_f, W_i, W_o son las matrices de pesos de las capas ocultas asignadas a la puerta de olvido, la puerta de entrada y la puerta de salida respectivamente y W_C es la matriz de pesos de la celda estado \tilde{C}_t . Mientras que U_f, U_i, U_o y U_C son las matrices de pesos que conectan el estado de salida de la celda anterior con la información de la celda actual. b_f, b_i, b_o y b_C son cuatro vectores de bias que permiten a la red aprender y adaptarse a patrones más complejos y no lineales en los datos. σ_g es la función de activación. Finalmente, a partir de los resultados obtenidos de las ecuaciones anteriores se obtiene que para cada iteración t la celda estado de salida C_t y la capa de salida h_t puede ser calculada como:

$$\begin{aligned}
 C_t &= f_t C_{t-1} + i_t \tilde{C}_t \\
 h_t &= o_t \tanh(C_t)
 \end{aligned} \tag{3.3.5}$$

3.4. Modelos Autorregresivos

En el campo de las ciencias económicas, numerosas variables macroeconómicas exhiben un comportamiento dinámico a lo largo del tiempo. Entre ellas, se encuentran el producto interno bruto (PIB), la inflación, el consumo de los hogares, la inversión extranjera, entre otras. Estas variables se distinguen por su capacidad de evolucionar de forma temporal, lo que implica que su comportamiento actual puede ser modelado considerando su comportamiento pasado (de Arce, 1998). En este contexto, el uso de los modelos autorregresivos se ha convertido en una herramienta útil, debido a su capacidad para capturar la relación lineal entre el valor actual de una variable y sus valores pasados. Un modelo autorregresivo de orden p , denotado como $AR(p)$, se puede expresar matemáticamente de la siguiente forma:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim RB(0, \sigma^2). \tag{3.4.1}$$

Donde, Y_t representa la serie de tiempo en el momento t ; $\phi_1, \phi_2, \dots, \phi_p$ son los coeficientes autorregresivos que capturan la relación de dependencia de los valores pasados, y ε_t es el

término de error que recoge la variabilidad no explicada en el momento t (Box et al., 2015). La aplicación de este tipo de modelos es bastante útil para la serie de tiempo sujeta de análisis en esta investigación (tasa IBR), partiendo del hecho de que su comportamiento tiende a ser estacionario y por sus características no puede tener un crecimiento sostenido en el tiempo.

En este estudio, además de utilizar los modelos autorregresivos (AR), se emplearán los modelos ARIMA como una herramienta de gran utilidad. Estos modelos combinan el componente autorregresivo y el componente de promedios móviles, permitiendo abordar de manera efectiva las series de tiempo que presentan tendencias o patrones de estacionalidad. Matemáticamente se definen de la siguiente manera:

$$\phi_p(L)(1-L)^d Y_t = \theta_q(L)\varepsilon_t, \quad \varepsilon_t \sim RB(0, \sigma^2). \quad (3.4.2)$$

Donde, $\phi_p(L)$ corresponde al polinomio autorregresivo de orden p , que contiene los coeficientes que multiplican los valores pasados de la serie Y_t ; $(1-L)^d$ es el operador de diferenciación de orden d que se aplica a la serie Y_t para lograr la estacionariedad y $\theta_q(L)$ corresponde al polinomio de promedios móviles de orden q que contiene los coeficientes que multiplican los valores pasados del error ε_t .

3.5. Modelos GARCH

Los modelos GARCH(p, q), también conocidos como modelos autoregresivos condicionalmente heterocedásticos generalizados, fueron introducidos por Bollerslev (1986) y son ampliamente utilizados en el análisis de series de tiempo financieras. Estos modelos se utilizarán como un complemento valioso a los modelos ARIMA que abordaremos en nuestro estudio, ya que nos permiten definir un modelo paramétrico para ε_t distinto a una sucesión independiente idénticamente distribuida (iid) $RB(0, \sigma^2)$. Se dice que un proceso ε_t es GARCH(p, q) si satisface las siguientes especificaciones:

$$\varepsilon_t = \sigma_t z_t, \quad (3.5.1)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \quad t = 0, 1, 2, \dots \quad (3.5.2)$$

donde $z_t \sim iid(0, 1)$, $\omega > 0$, $\alpha_i \geq 0$, $1 \leq i \leq p$, $\beta_j \geq 0$ y $1 \leq j \leq q$.

3.6. Ajustes de valoración (xVA)

Al negociar un contrato bilateral, como lo es un derivado financiero, surgen una serie de riesgos inherentes que, en ciertas circunstancias, pueden materializarse y afectar significativamente a las partes involucradas. Estos riesgos pueden incluir el potencial incumplimiento de una de las contrapartes, la variabilidad de las tasas de financiamiento, la fluctuación del valor de los activos subyacentes y la necesidad de mantener capital adicional para cubrir el riesgo propio. Para abordar esta complejidad y proporcionar una valoración más precisa y detallada del valor razonable de los derivados, surge xVA. El término xVA hace referencia a una variedad de ajustes de valoración aplicados a los derivados financieros que buscan ofrecer una cobertura integral de los distintos riesgos asociados con estos contratos. Entre los ajustes más comunes se tienen el ajuste de riesgo de crédito de la contraparte (CVA), ajuste de riesgo de crédito de la entidad (DVA), ajuste de financiación (FVA), y ajuste de capital (KVA) (Zeitsch, 2017). Bajo el marco regulatorio colombiano actual, las entidades financieras que operen con derivados y otros activos sujetos a riesgo de contraparte están obligadas a reflejar en sus informes financieros el CVA y el DVA. Estos ajustes tienen como objetivo cuantificar y contabilizar los riesgos asociados con el incumplimiento de las contrapartes y las ganancias potenciales que surgirían si la entidad misma incumpliera sus obligaciones. Esta investigación, se centra específicamente en el análisis y comprensión de los dos ajustes mencionados, dado que son de exigible cumplimiento en el ámbito regulatorio del país.

CVA

El CVA se refiere a la exposición de posibles pérdidas relacionadas con el incumplimiento de obligaciones por parte de una contraparte. En otras palabras, representa el monto que se debe ajustar al valor de un derivado libre de riesgo para reflejar el riesgo de que la contraparte no cumpla con sus obligaciones (Aragall, 2013). De esta manera, el valor de un activo ajustado al riesgo estaría dado por:

$$\text{Valor ajustado al riesgo} = \text{Valor libre de riesgo} - \text{CVA} \quad (3.6.1)$$

En consecuencia, el reto se encuentra en medir el valor presente de las potenciales pérdidas futuras ocasionadas ante la materialización del incumplimiento por parte de la contraparte. Dichas pérdidas, se encuentran en función de la exposición esperada presente en las fechas de pago en las que se podría dar el incumplimiento, la probabilidad de default y la proporción neta final a la que se estaría expuestos ante dicho evento. De una manera más técnica Gregory (2020) la define de la siguiente forma:

$$CVA(t) = (1 - R) \sum_{i=1}^T EPE(t, t_i) PD_C(t_{i-1}, t_i)$$

(3.6.2)

donde R es la tasa de recuperación que se asume constante para las dos contrapartes, $(1 - R)$ se define como la pérdida severa dada el incumplimiento de la contraparte, $EPE(t, t_i)$ es el valor esperado descontado en el periodo t de la exposición esperada positiva en un evento de incumplimiento y $PD_C(t_{i-1}, t_i)$ es la probabilidad de incumplimiento de la contraparte en el intervalo $[t_{i-1}, t_i]$.

DVA

Cuando en la negociación de un derivado, las partes involucradas solo consideran el Ajuste por CVA, es posible que no lleguen a un acuerdo en cuanto al precio. Para abordar esta situación, surge la figura del Ajuste por Debito (DVA), que está relacionado con la exposición a posibles pérdidas asociadas al incumplimiento de obligaciones por parte de la entidad misma. En otras palabras, el DVA representa las pérdidas que puede experimentar mi contraparte debido a que yo incumplí mis obligaciones.

La inclusión de este ajuste de riesgo permite que ambas partes contemplen integralmente los posibles incumplimientos, ajustados con sus respectivas probabilidades y exposiciones esperadas. De esta manera, ambas partes pueden llegar a un acuerdo en cuanto a los precios, ya que se considera el riesgo crediticio de ambas entidades en la transacción del derivado. De una manera más técnica Gregory (2020) la define de la siguiente forma:

$$DVA(t) = (1 - R) \sum_{i=1}^T ENE(t, t_i) PD_P(t_{i-1}, t_i), \quad (3.6.3)$$

donde $(1 - R)$ se define como la pérdida severa dada el incumplimiento de la entidad, $ENE(t, t_i)$ es el valor esperado descontado en el periodo t de la exposición esperada negativa en un evento de incumplimiento y $PD_P(t_{i-1}, t_i)$ es la probabilidad de incumplimiento de la entidad en el intervalo $[t_{i-1}, t_i]$.

Posteriormente, luego de haber calculado el DVA, este debe ser incorporado en el valor ajustado de riesgo del activo. Al tener en cuenta este ajuste, se compensa el Ajuste por Crédito (CVA), lo que da como resultado el valor neto de riesgo final.

$$\text{Valor ajustado al riesgo} = \text{Valor libre de riesgo} - \text{CVA} + \text{DVA}. \quad (3.6.4)$$

4. Estudio de antecedentes

El dinero como medio de intercambio entre bienes y servicios ha sido el instrumento de mayor relevancia en la evolución de la sociedad y la economía mundial (de la Fuente Rodríguez, 2018). Su uso es fundamental en el proceso económico y desde luego representa un bien para todo aquel que lo posea (Barrios, 2009). Desde esta perspectiva, el dinero actúa como capital de trabajo y por ende como factor productivo. Por lo tanto, la utilización de dicho factor conlleva un costo a modo de remuneración, el cual es conocido comúnmente como interés (Fernández Navarrete, 2017).

En el mercado colombiano, de acuerdo con Luna Proenza (2012) la mejor herramienta disponible para la medición del dinero es la tasa IBR (indicador bancario de referencia), que por definición es la tasa de interés de corto plazo a la cual están dispuestos a captar o prestar dinero las entidades bancarias locales. De acuerdo con el esquema de formación del indicador, las tasas de interés a 1, 3 y 6 meses se representan a partir de la constitución obligatoria de swaps de tasas de interés a los plazos respectivos y por tanto de allí la importancia de este producto para el mercado monetario local (Asobancaria, 2012).

Desde hace varias décadas, muchos autores han estudiado el comportamiento de las tasas de interés y su aplicación en la valoración de derivados financieros. Algunos de estos estudios son:

Pye (1966) estudió el comportamiento de las tasas de interés a diferentes plazos a través de la construcción de una cadena de Markov con el objetivo de determinar el valor de las inversiones bajo escenarios de incertidumbre. Por otro lado, Ho and Lee (1986) dentro de su investigación propusieron un modelo para predecir los movimientos de las tasas de interés a partir de movimientos estocásticos de manera que se garantice la ausencia de arbitrajes. Finalmente, el modelo puede ser utilizado para estimar el precio de activos atados a tasas de interés como las opciones sobre bonos de deuda.

Hull and White (1993) comparan diferentes enfoques en el desarrollo de modelos de tasas de interés bajo el supuesto de no arbitraje. Además, proponen un procedimiento numérico para la construcción de modelos de un factor a partir de la implementación de árboles trinomiales. Por último, concluyen que el procedimiento es numéricamente eficiente y puede ser utilizado para valorar opciones sobre bonos y derivados financieros atados a tasas de interés.

Adicionalmente, autores como Bicksler and Chen (1986) y Minton (1997) exploran y describen procedimientos adecuados para calcular el valor justo de swaps de tasas de interés. Así mismo, Minton (1997) realiza un acercamiento empírico para el cálculo aproximado de las probabilidades de incumplimiento a partir de la cotización del precio de los swaps.

Es importante resaltar que luego de la crisis financiera del 2008 la forma en que se valoran los derivados financieros ha tenido cambios significativos. Esto se debe principalmente a la inclusión de la gestión activa de los riesgos asociados a la inversión, los cuales tienen un impacto directo en el precio de los activos. García Doffing (2021) describe un marco de referencia de las acciones tomadas por los agentes del mercado en pro de gestionar activamente el riesgo y la implementación de acuerdos de colaterales y ajustes en valoración como el CVA (credit value adjustment), DVA (debit value adjustment) y FVA (funding value adjustment). Además, con el objetivo de controlar y mitigar los efectos del alto apalancamiento en los derivados, las entidades regulatorias han requerido la implementación de ajustes de riesgo de crédito tanto para las contrapartes como para la misma entidad. Estos ajustes deben ser tenidos en cuenta en la divulgación de los informes financieros. (IFRS, 2013).

Bajo esta nueva realidad, algunas investigaciones aplicadas hacia el cálculo del valor adecuado de un swap de tasa de interés son:

Ramírez Amaya (2019) dentro de su investigación analizó el cálculo del valor libre de riesgo con sus respectivos ajustes de valoración (CVA/DVA) para un swap de tasas de interés a cinco años entre Ecopetrol y un banco local colombiano a partir de la aplicación del modelo de tasa corta de Hull-White de un factor y la estimación de la probabilidad de incumplimiento a partir del asset swap spread del mercado de bonos en moneda extranjera.

Por otro lado, Batlle et al. (2020) describen de forma simplificada y de manera general como calcular el valor razonable de un swap de tasas de interés a partir de un modelo binomial de tipos de interés forward y el cálculo de la exposición potencial futura a partir de spreads crediticios de bonos corporativos libres de riesgo.

Para determinar el valor adecuado de los ajustes a valoración que hemos mencionado anteriormente se requiere de un consumo bastante alto en capacidad computacional. Para ello Nisiba (2011) propone un método de expansión asintótica para el cálculo del CVA en swap de tasas de interés. Finalmente concluye que el método empleado presenta resultados numéricos razonables y es mucho más rápido que el método de simulación de Monte Carlo.

Como se puede evidenciar, la parte fundamental en la valoración de un swap de tasas de interés es la estimación del comportamiento futuro del activo subyacente. Ante ello, los investigadores han optado recientemente por implementar herramientas de aprendizaje profundo

para la predicción de este tipo de series de tiempo. A continuación, referenciamos algunos de ellos:

Cornelissen (2021) pronostico la tasa SOFR (Secured Overnight Financing rate) a partir de la implementación de una red neuronal recurrente y comparo el resultado obtenido frente a un modelo autorregresivo. Finalmente concluye que los pronósticos realizados con la red neuronal superan al modelo autorregresivo en una reducción media de 0.45 en el RSME y además estima más eficientemente en el largo plazo.

Mondragón et al. (2021) construyeron una red neuronal recurrente con memoria de largo y corto plazo para la predicción del IBR. Analizaron y determinaron la incidencia de la tasa representativa del mercado y el rendimiento de los bonos del gobierno a diez años en el comportamiento del indicador. Por último, concluye que los resultados obtenidos con la red contemplan un margen de error bajo y por tanto se observa una alta calidad en los pronósticos generados.

Como se puede apreciar, en la literatura se encuentran una variedad de métodos para predecir el comportamiento de las tasas de interés en el corto plazo, sin embargo, hay poca o ninguna literatura que considera la valoración de los nuevos ajustes requeridos para la gestión del riesgo. Es por ello que se hace relevante explorar la aplicación de redes neuronales en la simulación del activo subyacente y en ese sentido cuantificar el valor razonable de un swap de tasas de interés.

5. Presentación de resultados

5.1. Descripción de los datos y supuestos

Los datos utilizados en este estudio se obtuvieron a través de la plataforma Bloomberg, tomando en consideración un rango de fechas de análisis desde el 4 de diciembre de 2017 hasta el 25 de febrero de 2020 (en adelante denominado "hoy"). Estos datos comprenden la curva swap IBR, la curva de precios y rendimientos de los bonos corporativos y gubernamentales, la curva de futuros sobre la TRM, entre otros. Es importante destacar que estos datos provienen del mercado, lo que nos permitió obtener una visión precisa y actualizada de las expectativas y la realidad de condiciones presentes durante dicho periodo de tiempo.

Explicación curva Swap IBR

La curva swap IBR, obtenida de Bloomberg, desempeña un papel fundamental en el desarrollo de este estudio, principalmente porque nos proporciona el punto de partida para construir las series de tiempo sujetas de análisis como lo veremos más adelante en esta sección. Se define como la representación de las tasas de interés correspondientes a la parte fija de los contratos swap a lo largo de diferentes vencimientos o plazos. Es decir, esta curva muestra la relación entre el plazo del swap y la tasa de interés fija pactada. En el desarrollo de este estudio, la denotaremos como $SR(t, T)$, donde t corresponde al momento en el que se considera la tasa y T representa el plazo.

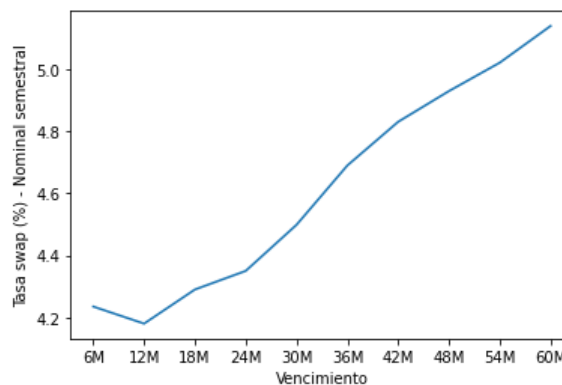


Figura 5.1.1.: Curva $SR(t, T)$; $t = 04/12/2017$. Tomada de Bloomberg.

Ilustración Swap IBR

Con el fin de proporcionar una explicación más clara sobre la información analizada en este estudio, consideramos apropiado ofrecer una representación gráfica del ejercicio financiero objeto de análisis.

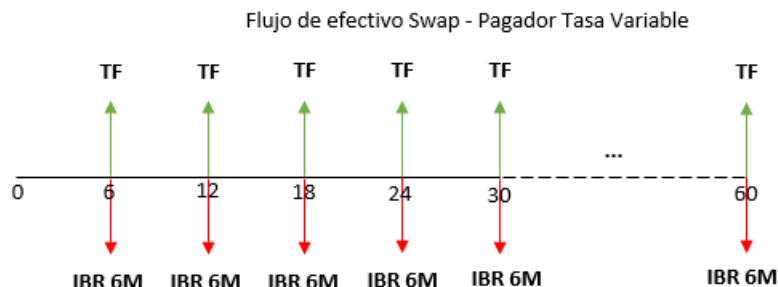


Figura 5.1.2.: Representación grafica swap.

En la Figura 5.1.2, se muestra el flujo de efectivo esperado para un inversionista que ha optado por acordar un swap a cinco años con liquidaciones semestrales, en el cual recibe una tasa fija semestral y paga una tasa variable. Específicamente, tiene la obligación de pagar la IBR a 6 meses. Profundizando un poco más en los detalles, la tasa IBR a 6 meses que se debe pagar, se configura como el promedio geométrico de la tasa diaria IBR durante el último semestre. Por lo tanto, esta transacción se liquida y realiza en cada una de las fechas de liquidación, lo que implica que en cada frecuencia de pago se reinicia el cálculo de la tasa variable. En consecuencia, si el inversionista desea estimar los pagos futuros, tendrá que pronosticar las tasas correspondientes entre las fechas de pago. Es decir, deberá determinar la tasa aplicable de 0 a 6 meses (entre el mes 0 y el mes 6), luego la tasa aplicable de 6 a 12 meses (entre el mes 6 y el mes 12), y así sucesivamente hasta la fecha de finalización del contrato. Estas tasas de interés aplicables entre las fechas intermedias, comúnmente conocidas en el ámbito financiero como tasas forward swap, desempeñan un papel fundamental en el cálculo de la valoración del swap.

Tomando en cuenta la explicación anterior, resulta evidente que la información relevante para este análisis son las tasas forward swap aplicables al ejercicio financiero en cuestión. Esto se debe a que, una vez que se disponga de dicha información, se pueden determinar los flujos de efectivo asociados y, en consecuencia, calcular el valor presente del swap.

Definición técnica tasas forward

Las tasas forward son tasas de interés que se acuerdan en el presente para su aplicación en el futuro. Estas tasas se caracterizan por tres momentos en el tiempo: el tiempo t , que representa el momento en el que se considera la tasa; el tiempo T_α , que marca el inicio de la vigencia de la tasa; y el tiempo T_β , que corresponde al vencimiento de la tasa. Donde, $t \leq T_\alpha < T_\beta$ y se denota por $F(t; T_\alpha, T_\beta)$ (Brigo and Mercurio, 2006).

Transformación curva forward swap

La curva forward swap se puede construir a partir de la curva swap que ha sido definida previamente. Este proceso se ilustrará de manera más detallada utilizando los datos del 04/12/2017 como ejemplo. La metodología propuesta por Smith and Ziemba (2013) guiará el procedimiento, proporcionando una comprensión más profunda de cómo se obtienen las tasas forward swap a partir de la curva swap existente.

Para realizar este análisis, se tomarán en cuenta los siguientes pasos clave:

- **Selección de puntos temporales:** Identificación de puntos temporales en la curva swap del 04/12/2017. Estos puntos serán esenciales para determinar las tasas forward swap en intervalos específicos.
- **Cálculo de tasas cero cupón:** Determinación de las tasas cero cupón asociadas a los momentos temporales identificados. Estas tasas cero cupón representan la rentabilidad al vencimiento sobre los periodos seleccionados.
- **Cálculo de tasas Forward:** Determinación de las tasas forward swap entre los puntos temporales seleccionados. Este cálculo considera la evolución de las tasas de interés a lo largo del tiempo y proporciona una visión más dinámica de las expectativas del mercado.

Para $t = 04/12/2017$ tenemos la siguiente información de la curva swap¹ de mercado:

T	6 Meses	12 Meses	18 Meses	24 Meses	30 Meses
SR(t,T)	4.2350 %	4.1800 %	4.2900 %	4.3500 %	4.4986 %
T	36 Meses	42 Meses	48 Meses	54 Meses	60 Meses
SR(t,T)	4.6900 %	4.8308 %	4.9300 %	5.0216 %	5.1400 %

Tabla 5.1.1.: Curva Swap IBR 04/12/2017. Tomada de Bloomberg.

¹Las tasas se encuentran en Nominal anual semestre vencido (NASV).

En primer lugar, se procede a construir la curva cero cupón implícita, denotada como $ZC(t, T)$, utilizando la técnica estándar de interpolación. Para contextualizar, esta técnica implica considerar un swap como un bono a tasa fija que paga la tasa de interés swap como su tasa de cupón, y se valora utilizando tasas de interés cero cupón. En este sentido, el valor presente resultante de los flujos futuros del bono se iguala a un valor nominal de 100. Este supuesto es comúnmente adoptado para la construcción de la curva, ya que se asume que las tasas swaps se negocian y cotizan a la par².

Entendiendo que en este caso de estudio hemos optado por liquidaciones semestrales, en el primer período de tiempo (6 meses), la tasa $ZC(t, 6) = SR(t, 6) = 4.2350\%$. Luego, Siguiendo la técnica de interpolación definida anteriormente, se tiene que, la tasa implícita cero cupón a 12 meses se determina como la solución para $ZC(t, 12)$ en la siguiente expresión:

$$100 = \frac{\frac{4,18}{2}}{\left(1 + \frac{4,235\%}{2}\right)} + \frac{\frac{4,18}{2} + 100}{\left(1 + \frac{ZC(t, 12)}{2}\right)^2},$$

$$ZC(t, 12) = 4,1794\%$$

Como se puede observar, la ecuación establece que el valor presente del bono a tasa fija por valor de 100 debe ser igual a la suma de los flujos de efectivo de las liquidaciones semestrales. El primer término representa el valor presente del primer pago semestral descontado usando la tasa $ZC(t, 6)$, mientras que el segundo termino representa el valor presente de los pagos restantes y el pago final al vencimiento del bono. Finalmente, se obtiene que la tasa implícita cero cupón a 12 meses, denotada como $ZC(t, 12)$ que satisface la ecuación es igual a 4.1794 %.

Ahora, ambos resultados son utilizados para encontrar la tasa $ZC(t, 18)$:

$$100 = \frac{\frac{4,29}{2}}{\left(1 + \frac{4,235\%}{2}\right)} + \frac{\frac{4,29}{2}}{\left(1 + \frac{4,1794\%}{2}\right)^2} + \frac{\frac{4,29}{2} + 100}{\left(1 + \frac{ZC(t, 18)}{2}\right)^3},$$

$$ZC(t, 18) = 4,2919\%$$

De forma generalizada, la tasa $ZC(t, T)$ con liquidaciones semestrales, a partir del segundo semestre, puede ser determinada usando la siguiente formula:

$$100 = \sum_{i=1}^T \frac{\left(\frac{SR(t, 6T)}{2}\right) \times 100}{\left(1 + \frac{ZC(t, 6i)}{2}\right)^i} + \frac{100}{\left(1 + \frac{ZC(t, 6T)}{2}\right)^T}$$

²Para más detalle véase: A Teaching Note on Pricing and Valuing Interest Rate Swaps Using LIBOR and OIS Discounting (Smith, 2012).

Donde, $SR(t, T)$ es la tasa swap, $ZC(t, T)$ la tasa cero cupón y T el número de semestres sujetos de análisis. Teniendo en cuenta la expresión anterior, procedemos a calcular los demás vencimientos de la curva, obteniendo finalmente la siguiente información:

T	6 Meses	12 Meses	18 Meses	24 Meses	30 Meses
ZC(t, T)	4.2350 %	4.1794 %	4.2919 %	4.3534 %	4.5083 %
T	36 Meses	42 Meses	48 Meses	54 Meses	60 Meses
ZC(t, T)	4.7101 %	4.8593 %	4.9645 %	5.0628 %	5.1929 %

Tabla 5.1.2.: Curva cero cupón 04/12/2017.

Ahora que ya contamos con la curva cero cupón, podemos proceder a construir la curva forward swap implícita, usando el método bootstrap³ que presentamos a continuación:

$$F(t; T_\alpha, T_\beta) = \left(\frac{\left(1 + \frac{ZC(t, T_\beta)}{2}\right)^{\frac{T_\beta}{6}}}{\left(1 + \frac{ZC(t, T_\alpha)}{2}\right)^{\frac{T_\alpha}{6}}} - 1 \right) \times 2$$

$$F(t; 0, 6) = ZC(t, 6) = 4,2350 \%,$$

Donde, $F(t; T_\alpha, T_\beta)$ representa la tasa forward entre el período T_α y el período T_β . $ZC(t, T_\beta)$ es la tasa cero cupón implícita al vencimiento T_β . $ZC(t, T_\alpha)$ es la tasa cero cupón implícita al vencimiento T_α . En este sentido, la fórmula nos permite anticipar la tasa de interés para un periodo entre T_α y T_β al evaluar cómo han cambiado las tasas cero cupón en esos dos periodos. En otras palabras, al conocer el rendimiento desde el inicio hasta T_α y hasta T_β , podemos estimar la rentabilidad esperada en dicho intervalo. A partir de esta fórmula procedemos a calcular los valores de la curva, obteniendo finalmente la siguiente información:

T_α, T_β	0 , 6 Meses	6 , 12 Meses	12 , 18 Meses	18 , 24 Meses	24 , 30 Meses
$F(t; T_\alpha, T_\beta)$	4.2350 %	4.1238 %	4.5172 %	4.5381 %	5.1291 %
T_α, T_β	30 , 36 Meses	36 , 42 Meses	42 , 48 Meses	48 , 54 Meses	54 , 60 Meses
$F(t; T_\alpha, T_\beta)$	5.7217 %	5.7572 %	5.7024 %	5.8507 %	6.3673 %

Tabla 5.1.3.: Curva Forward swap 04/12/2017.

Es importante destacar que los procedimientos descritos anteriormente fueron llevados a cabo para un día específico como parte de una representación metodológica. Sin embargo,

³Para más detalle véase: A Teaching Note on Pricing and Valuing Interest Rate Swaps Using LIBOR and OIS Discounting (Smith, 2012).

este proceso fue replicado para todo el rango de fechas ($t_1 = 04/12/2017$ y $t_n = 25/02/2020$), lo que nos permitió obtener una secuencia de curvas forward swap para para cada uno de los días considerados. Lo cual, visto de forma matricial es la siguiente:

Nodo 1	Nodo 2	...	Nodo 10
$F(t_1; 0, 6)$	$F(t_1; 6, 12)$...	$F(t_1; 54, 60)$
$F(t_2; 0, 6)$	$F(t_2; 6, 12)$...	$F(t_2; 54, 60)$
\vdots	\vdots	\ddots	\vdots
$F(t_n; 0, 6)$	$F(t_n; 6, 12)$...	$F(t_n; 54, 60)$

Como podemos observar, se identifican 10 intervalos clave en la curva. Cada uno de estos intervalos, cuando se considera de forma individual, se convierte en un nodo que representa un punto estratégico en la evolución de la curva. Es por ello, que al examinar cada nodo de manera individual se obtienen series de tiempo separadas para cada vencimiento. Estas series de tiempo son nuestros datos de análisis, y nos proporcionan información detallada sobre la variación de las tasas forward dentro de cada uno de los nodos identificados.

Selección emisores locales

Los emisores seleccionados para la valoración del swap y el cálculo del CVA/DVA fueron Bancolombia y BBVA Colombia. La elección se basó principalmente en el hecho de que ambas entidades presentan emisiones de bonos en dólares a tasa fija, con plazos de vencimiento similares. Otros factores, como la solidez crediticia de los emisores y participación del mercado, también fueron factores influyentes en la elección.

Esta selección nos permitió realizar comparaciones y análisis más precisos al evaluar los riesgos asociados con el swap y su potencial impacto en el valor crediticio como lo veremos posteriormente en el apartado de la estimación de la probabilidad de incumplimiento.

Finalmente, en los cálculos de valoración del swap utilizamos un valor nominal de mil millones de pesos (COP 1.000.000.000). Este valor nominal sirvió como referencia para determinar los flujos de efectivo y evaluar el impacto financiero del swap.

5.2. Aplicación modelos ARIMA-GARCH

En el apartado anterior, se identificaron 10 nodos, cada uno de los cuales representa una serie de tiempo para los diferentes vencimientos. Con esta información como base, en esta sección nos enfocamos en ajustar un modelo ARIMA-GARCH a cada serie. Este enfoque nos permite modelar tanto la estructura de la media como la volatilidad, con el objetivo final de pronosticar el movimiento futuro de cada serie para el próximo día hábil. En otras palabras, buscamos estimar el valor de $F(t_{n+1}; T_\alpha, T_\beta)$.

Para garantizar la consistencia y la comparabilidad en el análisis, es importante destacar que dentro del rango de fechas seleccionado se excluyen los fines de semana y días festivos para evitar distorsiones en los datos. Como resultado de este proceso de selección, cada serie de tiempo tiene un tamaño de 582 registros. Además, para el ajuste de los modelos, se ha establecido un enfoque de validación que implica dividir los datos en conjuntos de entrenamiento y prueba. En este caso, el conjunto de prueba tiene un tamaño de 10 registros, lo que corresponde a un período de 2 semanas. Esta elección se ha realizado teniendo en cuenta que es un horizonte sensible en el sentido de que los datos son diarios, por lo que puede haber suficiente incertidumbre por la naturaleza volátil y cambiante de la tasa de interés. No obstante, si fuera necesario, podríamos considerar otros horizontes de pronóstico. Entre tanto, el conjunto de entrenamiento se utilizó para calibrar y estimar los parámetros del modelo.

Pasos para ajustar el modelo ARIMA-GARCH

Con el propósito de ofrecer una explicación detallada de los procedimientos empleados para seleccionar el mejor modelo en cada uno de los nodos, nos centraremos en la información correspondiente al Nodo 1 (Nodo 0-6M). A continuación, se describirán y explicarán cada uno de los pasos realizados en este proceso:

- **Paso 1:** Dividir la serie en datos de entrenamiento (572 registros) y prueba (10 registros).

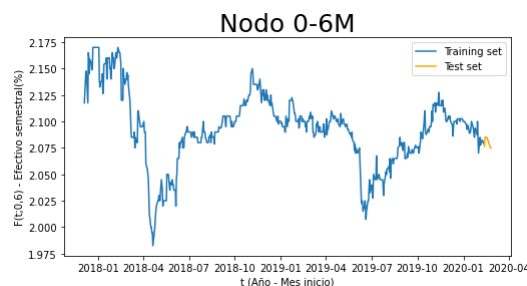


Figura 5.2.1.: Curva $F(t;0,6)$.

Al observar los datos de entrenamiento, se evidencia de manera inicial que el proceso no exhibe estacionariedad. Esto se debe principalmente a la presencia de tendencias tanto crecientes como decrecientes que se mantienen en el tiempo.

- **Paso 2:** Evaluar el comportamiento de la varianza a partir de la gráfica de la serie diferenciada.

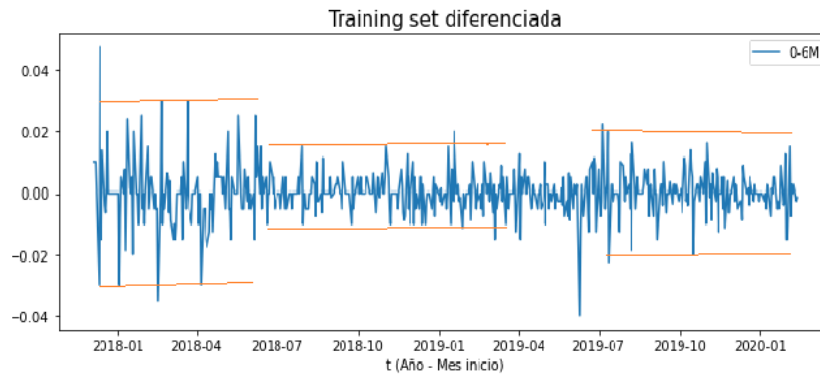


Figura 5.2.2.: Curva $\Delta F(t;0,6)$.

La serie diferenciada exhibe clusters de volatilidad. En particular, podemos notar que al comienzo de la serie, la varianza presenta movimientos dentro de un rango superior al observado en otros momentos. En la parte intermedia, la varianza disminuye y se mantiene de manera homogénea, mientras que al final de la serie se observan pequeños picos nuevamente. Estos patrones sugieren la posible presencia de efectos de heterocedasticidad que se pueden modelar por un ajuste GARCH.

Con el objetivo de lograr una mayor homogeneidad en la varianza, se aplicará el logaritmo natural a la serie. A partir de este punto, trabajaremos con la serie transformada $\text{LN}(F(t;0,6))$.

- **Paso 3:** Graficar ACF y PACF en búsqueda de evidencia de la necesidad de diferenciar la serie.

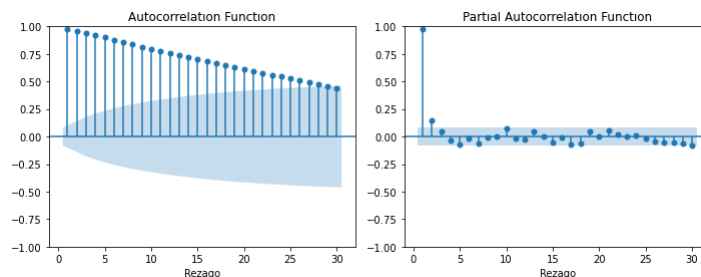


Figura 5.2.3.: Grafica ACF-PACF $\text{LN}(F(t;0,6))$.

La lenta decaída de la función de autocorrelación (ACF) y la presencia de un pico alto en el primer rezago de la función de autocorrelación parcial (PACF) sugieren que puede ser necesario aplicar una diferenciación a la serie para lograr la estacionariedad.

- **Paso 4:** Evaluar la presencia de raíces unitarias a partir del test aumentado de Dickey-Fuller (ADF).

H0: $LN(F(t; 0, 6))$ tiene una raíz unitaria, es decir, no es estacionaria.

H1: $LN(F(t; 0, 6))$ no tiene una raíz unitaria, es decir, es estacionaria.

Resultado	Valor
valor_p	0.1488
estadistico_t	-2.3753
Valor crítico 1 %	-3.442
Valor crítico 5 %	-2.867
Valor crítico 10 %	-2.569

Tabla 5.2.1.: Test ADF LN(F(t;0,6)).

Según los resultados obtenidos, el valor-p = 0.1488, es mayor que el valor de significancia $\alpha = 0.05$. Además, el estadístico t, que en nuestro análisis es -2.3753, al compararlo con el valor crítico del 5 % (-2.867), se encuentra fuera de la región de rechazo. Por lo tanto, hay evidencia estadística con un nivel de confianza del 95 % de no rechazar la hipótesis nula (H0) y en consecuencia, hay existencia de una raíz unitaria en la serie. Ahora, aplicamos nuevamente el test pero a la serie diferenciada para definir si hay presencia de una segunda raíz unitaria.

H0: $\Delta \ln(F(t; 0, 6))$ tiene una raíz unitaria, es decir, no es estacionaria.

H1: $\Delta \ln(F(t; 0, 6))$ no tiene una raíz unitaria, es decir, es estacionaria.

Resultado	Valor
valor_p	< 0.00
estadistico_t	-28.2495
Valor crítico 1 %	-3.442
Valor crítico 5 %	-2.867
Valor crítico 10 %	-2.569

Tabla 5.2.2.: Test ADF $\Delta \text{LN}(F(t;0,6))$.

Según los resultados obtenidos, el valor- $p < 0.00$, es menor que el valor de significancia $\alpha = 0.05$. Además, el estadístico t , que en nuestro análisis es -28.2495 , al compararlo con el valor crítico del 5% (-2.867), se encuentra dentro de la región de rechazo. Por lo tanto, hay evidencia estadística con un nivel de confianza del 95% de rechazar la hipótesis nula (H_0) y en consecuencia, la serie $\Delta \text{LN}(F(t;0,6))$ es estacionaria.

- **Paso 5:** Determinar los ordenes de p y q del modelo $\text{ARMA}(p,q)$, que mejor se ajusten a $\Delta \text{LN}(F(t;0,6))$, a partir de las funciones ACF y PACF.

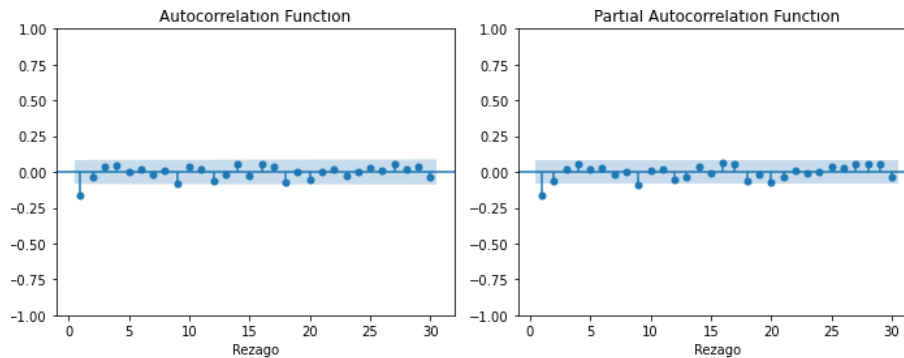


Figura 5.2.4.: Grafica ACF-PACF $\Delta \text{LN}(F(t;0,6))$.

Basándonos en la información obtenida de los correlogramas, se puede observar que tanto la función de autocorrelación (ACF) como la función de autocorrelación parcial (PACF) presentan un corte temprano en el rezago 1. Este patrón sugiere que un posible modelo adecuado sería un $\text{ARMA}(1,1)$.

- **Paso 6:** Realizar la estimación del modelo para verificar la significancia de los parámetros.

	coef	std err	z	$P> z $
ar.L1	0.1303	0.173	0.754	0.451
ma.L1	-0.3140	0.165	-1.898	0.058
sigma2	1.352e-05	4.42e-07	30.614	0.000

Tabla 5.2.3.: Estimación modelo ARIMA(1,1,1).

Según los resultados obtenidos, se encuentra que el Valor-p asignado para ϕ_1 (parámetro que acompaña el componente autorregresivo (AR)), el cual es 0.451 es mayor que el valor de significancia $\alpha = 0.05$ y por lo tanto, consideramos que no se hace relevante. Ante esto, realizamos nuevamente el ajuste pero con el modelo ARIMA(0,1,1).

	coef	std err	z	$P> z $
ma.L1	-0.1837	0.029	-6.389	0.000
sigma2	1.351e-05	4.42e-07	30.590	0.000

Tabla 5.2.4.: Estimación modelo ARIMA(0,1,1).

Finalmente obtenemos que $\text{LN}(F(t;0,6)) \sim \text{ARIMA}(0,1,1)$.

- **Paso 7:** Identificar efectos GARCH en los residuales del modelo.

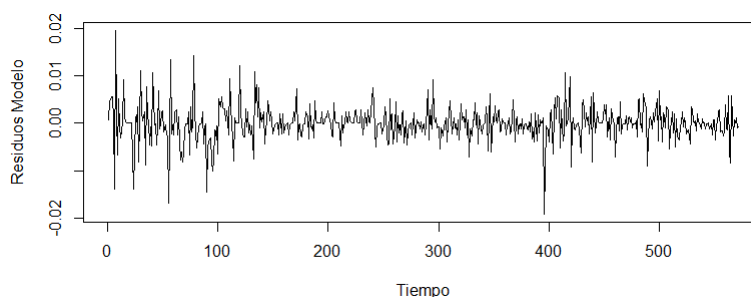


Figura 5.2.5.: Grafica residuales modelo ARIMA(0,1,1).

Al analizar el gráfico de los residuales, podemos notar la presencia de clusters de volatilidad, lo cual sugiere la posible existencia de efectos ARCH. Para determinar de manera más precisa si estos efectos están presentes, recurrimos a la prueba de hipótesis de Engle.

H0: No hay efectos ARCH

H1: Hay efectos ARCH

Resultado	Valor
valor_p	0.04961
estadistico_chi-squared	21.053

Tabla 5.2.5.: Lagrange Multiplier (LM) test.

Como podemos observar en el test LM, el valor-p es igual a 0.04961 y es menor que $\alpha = 0.05$ y por tanto es posible rechazar la hipótesis nula. Del mismo modo, resulta relevante examinar las gráficas ACF y PACF de los residuales del modelo al cuadrado con el fin de complementar el análisis y evaluar la pertinencia de incluir el componente GARCH en la estructura del modelo.

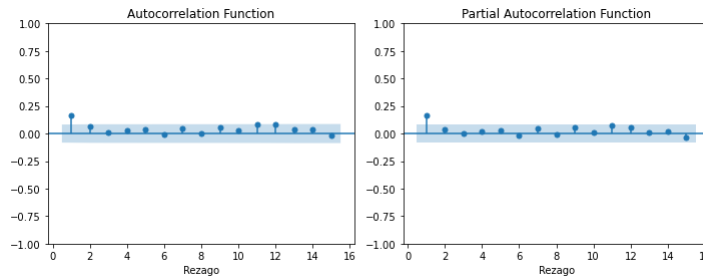


Figura 5.2.6.: Grafica ACF-PACF ϵ_t^2 .

Basándonos en la información obtenida de los correlogramas, se puede observar que tanto la función de autocorrelación (ACF) como la función de autocorrelación parcial (PACF) presentan un corte temprano en el rezago 1. Este patrón sugiere que un posible modelo adecuado para modelar la varianza condicional presente en los datos sería un GARCH(1,1). Finalmente, se continua con el ajuste de un modelo conjunto para media y varianza.

- **Paso 8:** Ajustar un modelo conjunto de media y varianza para $\Delta \text{LN}(F(t;0,6))$.

	Estimate	std err	t value	$Pr(> z)$
mu	0.000001	0.000109	0.013071	0.989571
ma1	-0.149172	0.038467	-3.877881	0.000105
omega	0.000000	0.000001	0.609464	0.542217
alpha1	0.075757	0.018262	4.148354	0.000033
beta1	0.902926	0.013998	64.503571	0.000000
skew	-0.004755	0.087091	-0.054604	0.956454
shape	0.351230	0.079422	4.422328	0.000010

Tabla 5.2.6.: Estimación modelo conjunto.

Considerando los estadísticos brindados por el modelo, podemos evidenciar que los parámetros θ_1 , α_1 y β_1 tienen un valor-p menor al 5 %, lo que indica su significancia. Por lo tanto, el modelo para $\Delta\text{LN}(F(t;0,6))$ se configura como un modelo ARMA(0,1) + GARCH(1,1) + NIG y el modelo para $\text{LN}(F(t;0,6))$ se configura como un modelo ARIMA(0,1,1) + GARCH(1,1) + NIG. Es importante mencionar que, durante el proceso de construcción del modelo, se realizaron diversas simulaciones para evaluar diferentes distribuciones del error (Z_t). Finalmente, luego de analizar los resultados obtenidos, se determinó que la distribución NIG ofrecía pronósticos de mayor calidad en comparación con las otras alternativas consideradas.

- **Paso 9:** Pronosticar los siguientes 10 días para hacer la validación con los datos de prueba.

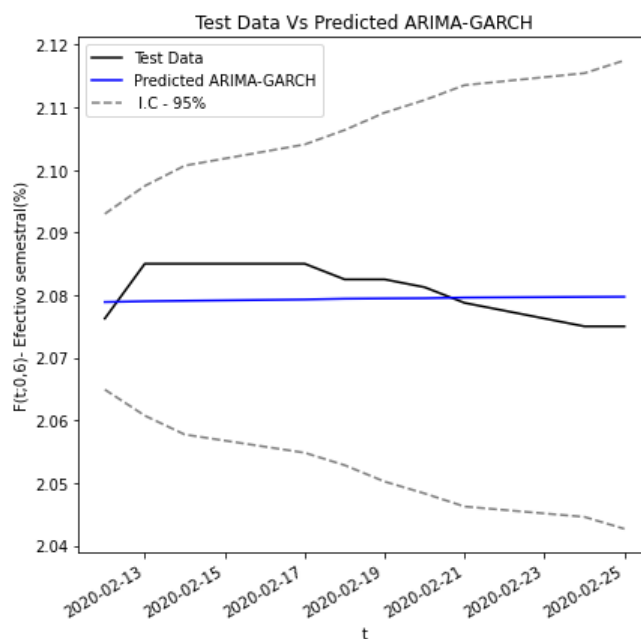


Figura 5.2.7.: Test Data Vs Predicted ARIMA-GARCH .

En este paso de la investigación, utilizamos R studio como nuestra plataforma de programación y nos apoyamos en la versátil librería rugarch para llevar a cabo un proceso integral de estimación y pronóstico. Las predicciones se generan mediante funciones específicas de la librería que aprovechan tanto los datos históricos como la estructura de nuestro modelo ARIMA-GARCH, permitiéndonos anticipar las tendencias y la volatilidad en los próximos 10 días. Además, para cuantificar la incertidumbre en nuestras proyecciones, empleamos la técnica de Bootstrap. Esta técnica nos posibilita la creación de múltiples muestras a partir del modelo obtenido y, mediante la evaluación de percentiles definidos, estimar los intervalos de confianza que enriquecen nuestra comprensión de la variabilidad asociada a los resultados.

- **Paso 10:** Calcular las métricas de desempeño.

En esta etapa de evaluación, utilizamos algunas métricas de desempeño con el objetivo de medir la calidad de nuestras predicciones.

MSE (Error Cuadrático Medio): Esta métrica mide el valor promedio de los errores al cuadrado entre los resultados obtenidos con el modelo y los valores reales. Se calcula como la diferencia entre cada predicción y el valor real, elevando al cuadrado cada diferencia y luego calculando el promedio de estos valores.

MAE (Error Absoluto Medio): Esta métrica mide el valor absoluto promedio de las diferencias entre los resultados obtenidos con el modelo y los valores reales. Se calcula como la diferencia absoluta entre cada predicción y el valor real, y luego calculando el promedio de estas diferencias.

MAPE (Error Porcentual Absoluto Medio): Esta métrica cuantifica el error promedio como un porcentaje del valor real. Se calcula como la diferencia porcentual promedio entre los resultados obtenidos con el modelo y los valores reales, proporcionando una medida de precisión relativa.

Métrica	MSE	MAE	MAPE
Resultado	0.0000178	0.0038421	0.184611

Tabla 5.2.7.: Métricas de desempeño.

Es importante destacar que los pasos descritos anteriormente fueron llevados a cabo para un nodo específico como parte de una representación metodológica. Sin embargo, este proceso fue replicado para todos los nodos. Los resultados obtenidos en cada serie de tiempo, con sus respectivas especificaciones y métricas de prueba se pueden observar en el anexo A

Después de haber seleccionado los modelos más adecuados para cada uno de los nodos, llevamos a cabo el cálculo de la predicción de la respectiva curva⁴ para el siguiente día hábil, denotado como $F(t_{n+1}; T_\alpha, T_\beta)$, la cual se presenta a continuación:

T_α, T_β	0 , 6 Meses	6 , 12 Meses	12 , 18 Meses	18 , 24 Meses	24 , 30 Meses
$F(t_{n+1}; T_\alpha, T_\beta)$	2.075 %	2.126 %	2.233 %	1.898 %	2.058 %
T_α, T_β	30 , 36 Meses	36 , 42 Meses	42 , 48 Meses	48 , 54 Meses	54 , 60 Meses
$F(t_{n+1}; T_\alpha, T_\beta)$	2.360 %	2.352 %	2.286 %	2.303 %	2.484 %

Tabla 5.2.8.: Curva $F(t_{n+1}; T_\alpha, T_\beta)$ ARIMA-GARCH.

⁴Los datos de la predicción están en términos efectivos semestrales.

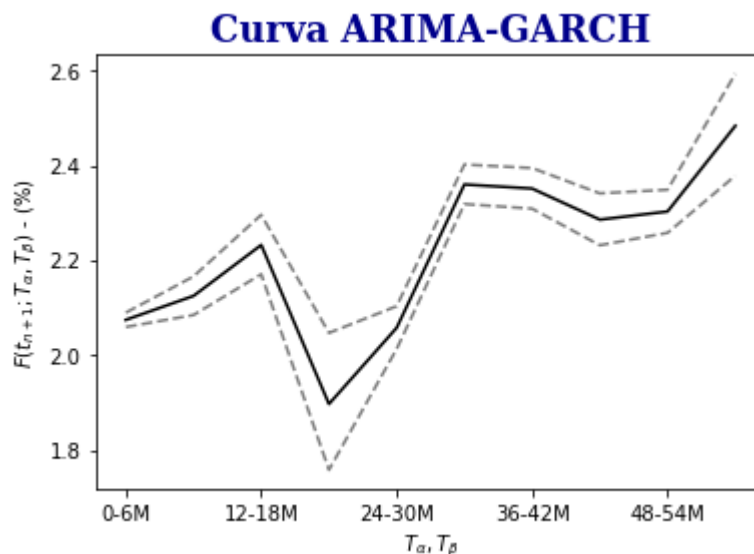


Figura 5.2.8.: Grafica $F(t_{n+1}; T_\alpha, T_\beta)$ ARIMA-GARCH.

La gráfica obtenida muestra la curva estimada de tasas forward para distintos plazos. En esta curva, se han unificado los pronósticos realizados en cada uno de los nodos para el próximo día hábil, t_{n+1} , junto con sus intervalos de confianza correspondientes. Esta información será fundamental en los capítulos posteriores para evaluar el valor del swap sujeto de análisis.

5.3. Aplicación redes neuronales recurrentes

En esta sección, nuestro objetivo es predecir la curva $F(t_{n+1}; T_\alpha, T_\beta)$ mediante el ajuste de una red neuronal recurrente en cada nodo. Para lograr esto, exploramos tres variantes de arquitectura diferentes con el fin de determinar cuál de ellas se ajustaba mejor a los datos. Las variantes utilizadas fueron las siguientes:

- **Vanilla:** Es la arquitectura más simple. Solo tiene dentro de su módulo de activación una función sigmoïdal. El comportamiento de esta red fue explicado en la **sección 3.3**.

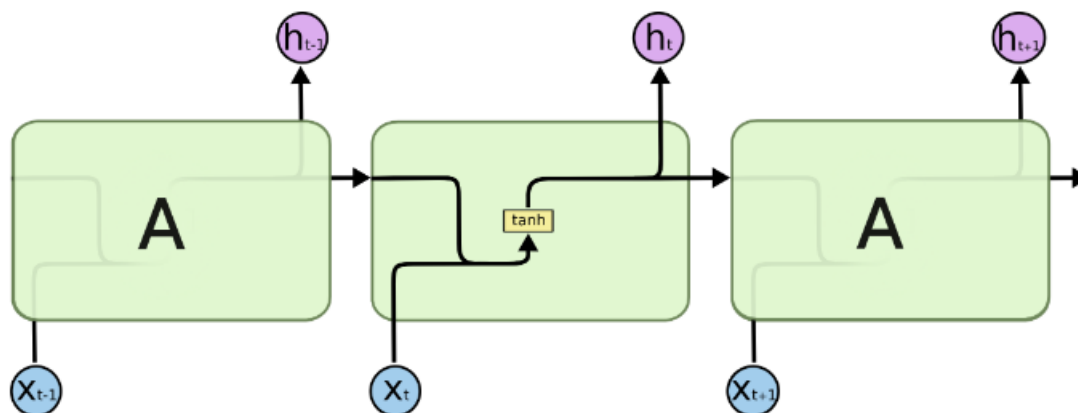


Figura 5.3.1.: Representación grafica RNN Vanilla. Tomado de Olah (2015).

- **LSTM:** Es una arquitectura más compleja. Dentro de su módulo de activación presenta celdas que regulan el flujo de información de corto y largo plazo. Además, cuenta con un conjunto de neuronas que permiten ir liberando la información que ya no es relevante. El comportamiento de esta red fue explicado en la **sección 3.3**.

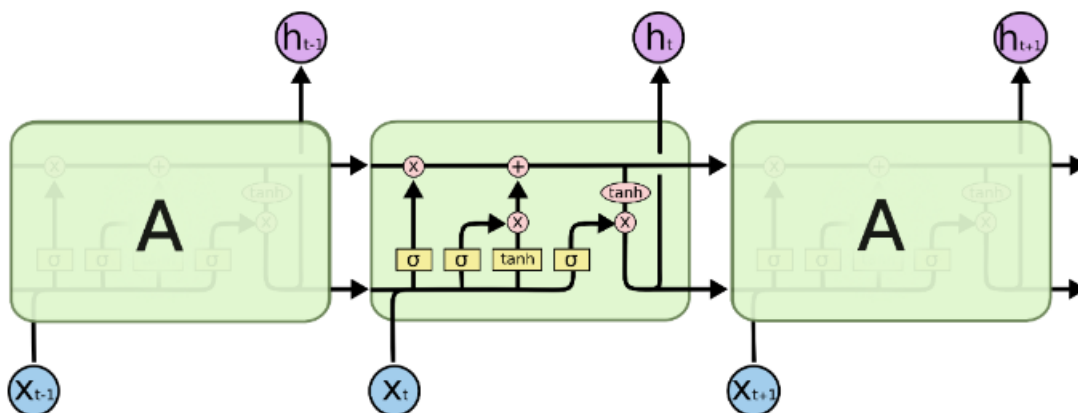


Figura 5.3.2.: Representación grafica RNN LSTM. Tomado de Olah (2015).

- **GRU:** Esta es una modificación a la arquitectura de red anterior, en la que se comprimen y fusionan algunas celdas para hacerla más ligera y mejorar su velocidad de procesamiento.

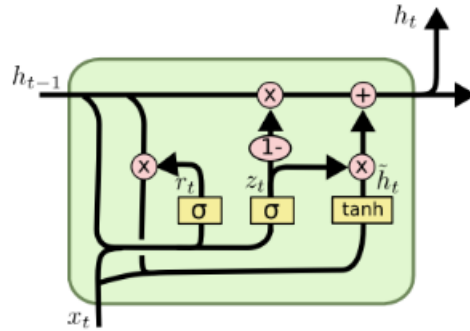


Figura 5.3.3.: Representación grafica RNN GRU. Tomado de Olah (2015).

En la Figura 5.3.3, se presenta el esquema de la red neuronal recurrente GRU. En este diagrama, cada línea facilita el traslado de la información desde la salida de un nodo hasta las entradas de otros nodos. Los círculos de color rosa simbolizan operaciones entre vectores, mientras que, las cajas de color amarillo indican capas de redes neuronales. La unión de líneas denota la concatenación de vectores, y cuando una línea se bifurca indica que su contenido se replica, dirigiéndose a ubicaciones distintas. El modelo matemático de la red de acuerdo con Olah (2015) se expresa de la siguiente forma:

$$\begin{aligned}
 z_t &= \sigma(W_z \cdot [h_{t-1}, x_t]) \\
 r_t &= \sigma(W_r \cdot [h_{t-1}, x_t]) \\
 \tilde{h}_t &= \tanh(W \cdot [r_t * h_{t-1}, x_t]) \\
 h_t &= (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t
 \end{aligned}
 \tag{5.3.1}$$

En las ecuaciones presentadas, x_t representa las variables de entrada en el tiempo t . Las matrices de pesos W_z y W_r están asociadas con las compuertas de actualización (z_t) y reinicio (r_t), respectivamente. La matriz de pesos W está relacionada con la capa candidata \tilde{h}_t . La función de activación σ es aplicada mediante la función sigmoide. El estado oculto en el tiempo t (h_t) se actualiza utilizando la compuerta de actualización z_t , el estado oculto anterior h_{t-1} , y el candidato a estado oculto \tilde{h}_t . Aquí, h_t y h_{t-1} representan el estado oculto en los tiempos t y $t - 1$, respectivamente.

Una vez que hemos explicado los componentes y el funcionamiento de cada una de las estructuras utilizadas en nuestro análisis, ahora procederemos a explicar los hiperparámetros que se utilizaron en cada una de las configuraciones realizadas a través de la librería Darts por medio de la función BlockRNNModel:

- **Epoch:** Determina el número de iteraciones en que la red procesa los datos de entrenamiento para actualizar y calibrar los parámetros.
- **Hidden-dim:** Determina el número de unidades o neuronas en cada capa oculta de la

red.

- **n-rnn-layers:** Determina la cantidad de capas utilizadas en el módulo de activación. Cada capa tiene su propio número de unidades (determinado por "Hidden-dim").
- **Dropout:** Determina el porcentaje de las neuronas que dentro de una capa serán apagadas aleatoriamente en cada paso del entrenamiento, con el fin de evitar problemas de sobreajuste.
- **Optimizer-kwarg:** Determina la tasa de aprendizaje de la red.

Al igual que en la sección anterior, se dividió la serie de tiempo en conjuntos de entrenamiento y prueba. Con el fin de mantener la misma proporción y permitir la comparación de métricas de error, se utilizó un tamaño de 10 registros para el conjunto de prueba. Ahora, luego de tener claridad sobre todos los lineamientos relacionados con las configuraciones y los datos, se realizan las estimaciones pertinentes. En esta fase es importante mencionar, que, para cada uno de los nodos se aplicaron las tres configuraciones de red. Cada configuración se evaluó utilizando métricas de error específicas, cuyos resultados detallados se encuentran disponibles en el Anexo B de este informe.

En referencia a los hiperparámetros presentados en el Anexo B , es importante destacar que estos corresponden a la última configuración implementada que arrojó los resultados más prometedores. Así mismo, es relevante señalar que se llevaron a cabo varias configuraciones en un proceso de prueba y error con el objetivo de optimizar el rendimiento del modelo. Sin embargo, se enfrentó a ciertas limitaciones computacionales, ya que el entrenamiento de estos modelos exigía una capacidad considerable de recursos. A pesar de esto, se realizaron ajustes iterativos en los hiperparámetros con el fin de mejorar el rendimiento predictivo. Finalmente, es importante tener en cuenta que, aunque los resultados actuales son prometedores, existe la posibilidad de explorar otras configuraciones para continuar refinando el modelo y lograr pronósticos con un menor margen de error.

Después de ajustar las tres configuraciones para cada nodo y comparar las métricas de error respectivas, se optó por seleccionar la configuración de red que presentara al menos 2 métricas con los valores más bajos. En otras palabras, se eligió el modelo que demostró tener los errores más bajos en al menos dos medidas de evaluación para cada nodo. Posteriormente, tomando en cuenta que ya se había seleccionado la configuración más adecuada para cada nodo entre las tres opciones disponibles, se procedió a utilizar esta configuración específica para realizar pronósticos en t_{n+1} . Los resultados de estos pronósticos se presentan de manera consolidada a continuación:

T_α, T_β	0 , 6 Meses	6 , 12 Meses	12 , 18 Meses	18 , 24 Meses	24 , 30 Meses
$F(t_{n+1}; T_\alpha, T_\beta)$	2.076 %	2.114 %	2.209 %	1.929 %	2.078 %
T_α, T_β	30 , 36 Meses	36 , 42 Meses	42 , 48 Meses	48 , 54 Meses	54 , 60 Meses
$F(t_{n+1}; T_\alpha, T_\beta)$	2.359 %	2.383 %	2.320 %	2.386 %	2.616 %

Tabla 5.3.1.: Curva $F(t_{n+1}; T_\alpha, T_\beta)$ RNN.

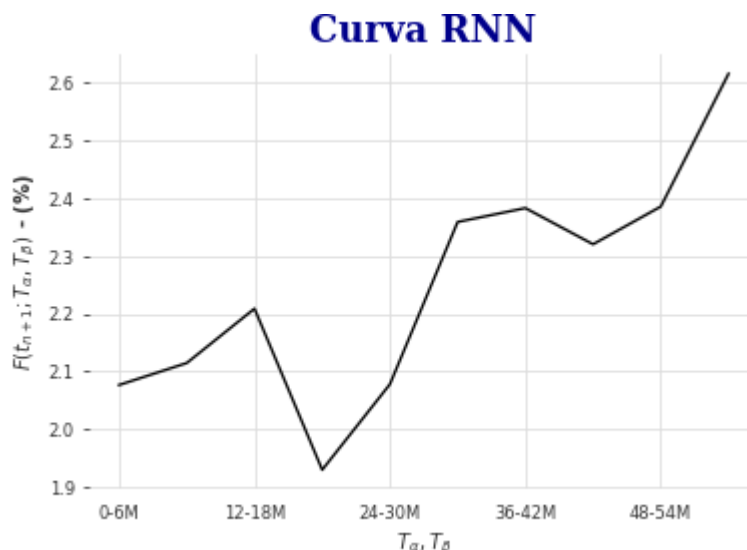


Figura 5.3.4.: Grafica $F(t_{n+1}; T_\alpha, T_\beta)$ RNN.

En esta fase de la investigación, optamos por la programación en Python y nos respaldamos en la librería Darts para llevar a cabo un proceso completo de estimación y pronóstico. Las predicciones se generaron utilizando funciones específicas de la librería, y los detalles más precisos de los códigos utilizados se encuentran disponibles en el Anexo D.

5.4. Estimación de las probabilidades de default

En el cálculo de los ajustes de valoración por riesgo de crédito (CVA/DVA), es fundamental considerar la probabilidad de incumplimiento (PD), tal como se mencionó en las ecuaciones 3.6.2 y 3.6.3. Específicamente, porque nos proporciona información contundente sobre la capacidad de pago de las contrapartes en el ejercicio financiero sujeto de análisis. Es por ello, que en esta sección nos enfocaremos en desarrollar esta métrica para los emisores seleccionados, es decir, BBVA Colombia y Bancolombia.

En principio, es relevante destacar que una manera común de estimar la probabilidad de incumplimiento es mediante el uso de los respectivos Credit Default Swaps (CDS), que, en esencia, no es más que la prima de riesgo que el mercado le cobra a los emisores en caso de

un posible default en sus obligaciones financieras. No obstante, al considerar las limitaciones del mercado local, especialmente en el ámbito corporativo donde no existe un mercado de CDS, resulta necesario buscar alternativas para realizar esta estimación de manera efectiva.

Como solución alternativa, en este estudio adoptamos el enfoque descrito por Hull (2018), el cual plantea que la probabilidad de incumplimiento de una compañía puede ser estimada a partir del precio de sus bonos. Según este enfoque, se asume que la única razón por la cual un bono corporativo se vendería a un precio inferior a un bono similar libre de riesgo es la posibilidad de que la compañía entre en quiebra. Es por ello, que en este estudio utilizamos el asset swap spread (ASW). El asset swap spread es una medida que compara el rendimiento de un bono corporativo con el rendimiento de un instrumento libre de riesgo, en nuestro caso el rendimiento interbancario. Para brindar una explicación detallada sobre cómo funciona el ASW, consideremos la siguiente situación:

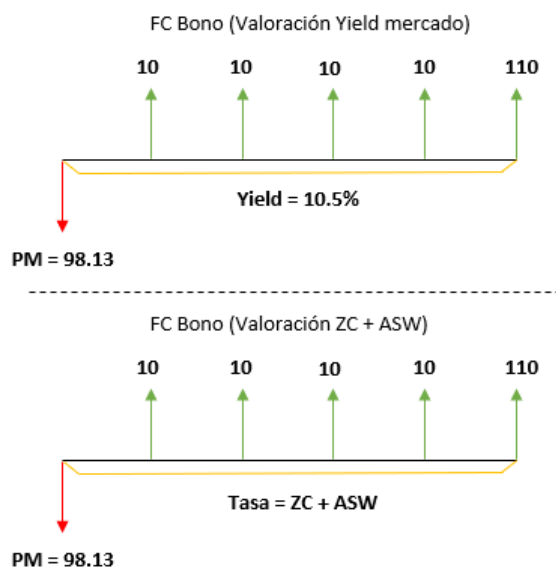


Figura 5.4.1.: Representación gráfica ASW.

En la Figura 5.4.1 se puede observar con claridad el ejercicio financiero involucrado en el cálculo del ASW. Para facilitar la comprensión, consideremos inicialmente la parte superior del gráfico, que ilustra el flujo de caja asociado a la emisión de un bono con una tasa cupón del 10%, vencimiento a cinco años y pagos anuales. Del mismo modo, supongamos una yield de mercado del 10.5%. En este caso, el resultado de calcular el valor presente de los flujos futuros, al que llamaremos PM (precio de mercado), al que se estaría negociando el bono utilizando la tasa de mercado, es de \$98.13. En la segunda parte del gráfico, se presentan nuevamente los mismos flujos de caja del bono. Además, se muestra el precio de mercado

calculado previamente y la tasa implícita cero cupón⁵. Entre tanto, nuestra incógnita es el ASW, que en esencia, es el margen adicional sobre toda la curva cero cupon que el mercado está descontando para tener exposición al emisor del bono.

Finalmente, el efecto observado, es que el ASW es el margen por el cual el rendimiento del bono corporativo esta excediendo el rendimiento de un bono similar libre de riesgo , donde la tasa libre de riesgo esta dada por la curva implícita cero cupón.

Ahora bien, como hemos mencionado previamente el mercado corporativo local presenta ciertas limitaciones, más específicamente, es un mercado poco líquido, no presenta una amplia cantidad de emisores ni emisiones, lo que se traduce en falta de información diaria de precios en el mercado. Esta limitación dificulta el cálculo preciso del ASW, ya que este requiere contar con datos actualizados y representativos de los precios de mercado de los bonos.

Para abordar el problema mencionado, se realizó una búsqueda de las emisiones efectuadas por los agentes en el mercado exterior. En la búsqueda realizada, se buscaron emisiones con características en común entre ambos emisores, como la tasa cupón fija, la moneda de emisión y un vencimiento cercano a cinco años. Estas similitudes se establecieron con el propósito de garantizar que los resultados fueran comparables, ya que las emisiones utilizadas compartían parametrizaciones iguales. En el anexo C se pueden observar las emisiones seleccionadas con las que se realizaron todos los cálculos. Es importante destacar que ambas emisiones comparten vencimientos cercanos y tienen una tasa de rentabilidad fija. Además, dado que los bonos fueron emitidos en un mercado exterior, contamos con información diaria de precios de negociación que utilizaremos como punto de partida para el cálculo del ASW. Sin embargo, los flujos de caja de los bonos están denominados en dólares, por lo que es necesario convertirlos a pesos para mantener la consistencia en el ejercicio financiero.

Para llevar a cabo esta conversión, utilizamos la información relacionada de los contratos futuros sobre la tasa de cambio provenientes de Bloomberg y, por ende, del mercado. Una vez que se ha recopilado y organizado toda la información necesaria, el siguiente paso se centra en abordar la resolución del ejercicio financiero que se presenta en el contexto analizado, el cual se encuentra representado en la Figura 5.4.1. En este proceso, empleamos el solver de Microsoft Excel, que nos permite encontrar de manera precisa el valor de ASW necesario para alcanzar el equilibrio financiero. Obteniendo como resultado los siguientes spreads⁶ en pesos colombianos (COP) para cada uno de los emisores locales desde el 29/01/20 hasta el 25/02/20 .

⁵Hace referencia a la curva cero cupón que calculamos previamente en la sección 5.1.

⁶1pbs es igual a 0.01 %.

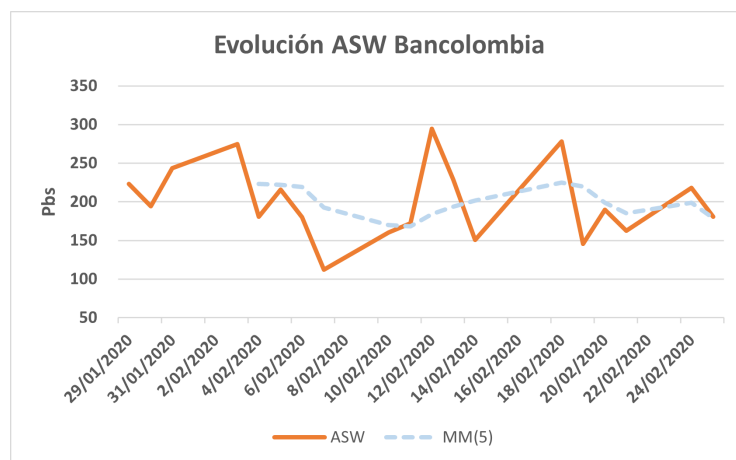


Figura 5.4.2.: ASW Bancolombia.

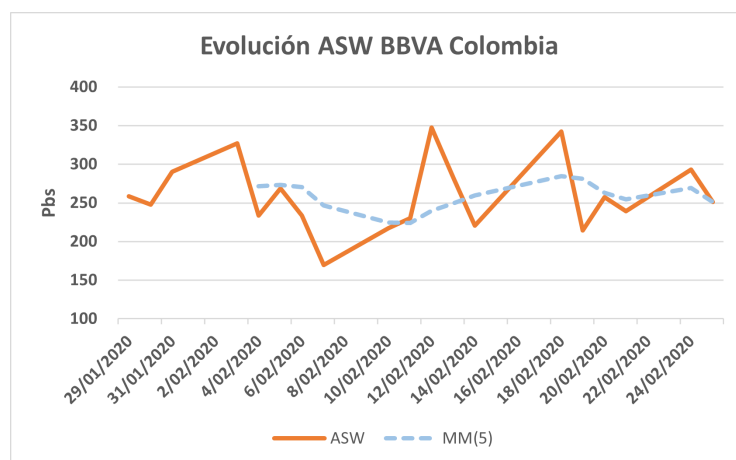


Figura 5.4.3.: ASW BBVA Colombia.

De acuerdo con los resultados obtenidos, podemos observar que el spread para BBVA Colombia se encuentra en un nivel 50 Puntos Básicos (Pbs) o un 0.5% superior al spread de Bancolombia. Este hallazgo indica que, en comparación con las condiciones del mercado, los inversionistas perciben un mayor riesgo al adquirir bonos emitidos por BBVA Colombia en lugar de Bancolombia.

Por otra parte, con el propósito de mitigar el impacto causado por la volatilidad presente en la curva de la tasa de cambio, que se requiere para convertir los flujos a moneda local, hemos optado por emplear una media móvil de cinco días (MM(5)). Esta estrategia nos ha permitido observar una perspectiva acorde con la evolución semanal de las expectativas del riesgo asociado.

Asset-swap spread(Pbs)- Fecha corte	
Emisor	ASW MM(5)
Bancolombia	180
BBVA Colombia	250

Tabla 5.4.1.: ASW IBR.

El cálculo previo nos proporciona una comprensión más profunda del riesgo vinculado a una contraparte, considerando sus emisiones en dólares. Esta métrica que hemos calculado abarca diversos factores, incluyendo aspectos específicos relacionados con la contraparte, como la reducción de su calificación crediticia, la pérdida de reputación y el deterioro de sus utilidades, además de los factores macroeconómicos en el país donde la contraparte concentra sus principales operaciones, tales como la inflación, el tipo de cambio, el desempleo y la tasa de interés, entre otros elementos. Esto se debe a que esta métrica se basa en las emisiones realizadas en los mercados internacionales, proporcionando una visión completa de los riesgos inherentes a la contraparte. Con el objetivo de determinar el riesgo específico de cada emisor, fue necesario realizar la misma búsqueda para la deuda colombiana (TES), considerando tanto las emisiones en dólares como en pesos y de esta forma poder ajustar los spreads descontando el riesgo país asociado.

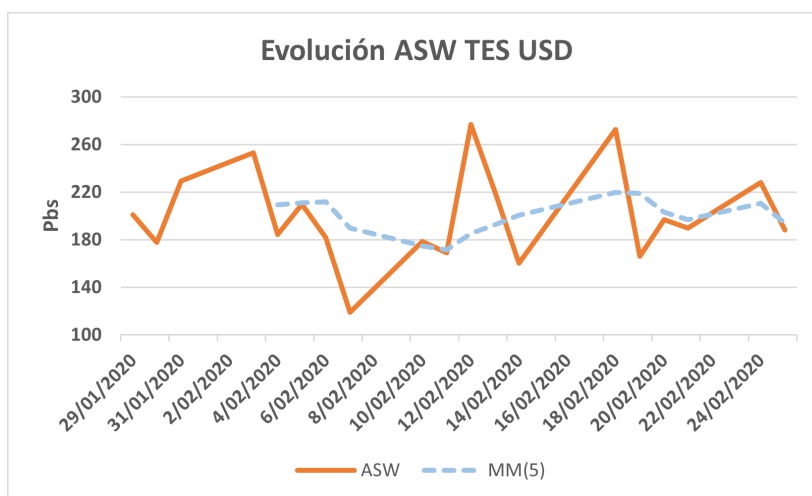


Figura 5.4.4.: ASW TES USD.

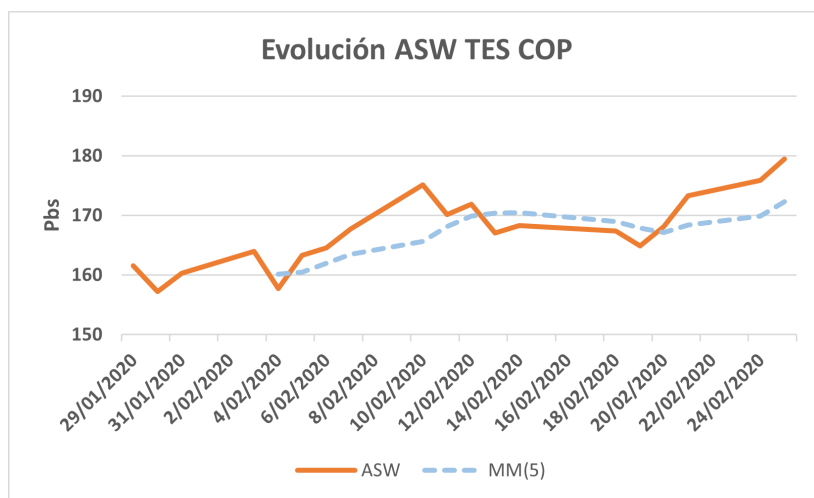


Figura 5.4.5.: ASW TES COP.

Asset-swap spread(Pbs)- Fecha corte	
Emisor	ASW MM(5)
Colombia TES USD	194
Colombia TES COP	172

Tabla 5.4.2.: ASW Colombia.

Podemos notar que la deuda del Estado colombiano con un vencimiento cercano a 5 años presenta un diferencial de aproximadamente 22Pbs ($194-172 = 22$) en comparación con los ASW de las emisiones en dólares y en moneda local. Esta diferencia se puede explicar principalmente como la prima adicional que el mercado le exige al gobierno por tener una deuda denominada en dólares en contraste con una deuda denominada en pesos. Dada esta diferencia, ajustaremos los spreads de los bancos locales descontando los 22 puntos básicos sobre los spreads calculados originalmente. Esto nos permitirá tener exclusivamente el riesgo específico asociado a cada emisor.

Asset-swap spread(Pbs)- Fecha corte	
Emisor	ASW MM(5) Ajustado
Bancolombia	158
BBVA Colombia	228

Tabla 5.4.3.: ASW IBR Ajustado.

Ahora que ya tenemos los ASW ajustados para cada emisor, procedemos a calcular las respectivas probabilidades de incumplimiento en cada uno de los nodos a partir de la siguiente

fórmula⁷:

$$p = \frac{1 - e^{-yt}}{1 - R} \quad (5.4.1)$$

Donde y es el ASW , R^* es la tasa de recuperación y t el tiempo.

Probabilidad de incumplimiento acumulada			Probabilidad entre nodos		
Años	Bancolombia	BBVA Colombia	Nodos	Bancolombia	BBVA Colombia
0.5	1.12 %	1.62 %	1	1.12 %	1.62 %
1.0	2.24 %	3.22 %	2	1.12 %	1.60 %
1.5	3.35 %	4.80 %	3	1.11 %	1.58 %
2.0	4.44 %	6.37 %	4	1.10 %	1.56 %
2.5	5.53 %	7.92 %	5	1.09 %	1.55 %
3.0	6.61 %	9.44 %	6	1.08 %	1.53 %
3.5	7.69 %	10.96 %	7	1.07 %	1.51 %
4.0	8.75 %	12.45 %	8	1.06 %	1.50 %
4.5	9.80 %	13.93 %	9	1.06 %	1.48 %
5.0	10.85 %	15.39 %	10	1.05 %	1.46 %

Tabla 5.4.4.: PD Emisores.

Después de aplicar la fórmula mencionada, se evidencia que las probabilidades acumuladas de incumplimiento a lo largo de los cinco años para Bancolombia y BBVA Colombia son del 10.85 % y 15.39 %, respectivamente. Mientras que, la probabilidad entre nodos, se determina mediante la diferencia entre la probabilidad acumulada del nodo actual y la probabilidad acumulada del nodo anterior.

⁷Para más detalle véase: The Credit Default Swap Basis, Appendix II (Choudry, 2006)

*Hace referencia al porcentaje recuperable de la inversión en caso de que se materialice el incumplimiento de una obligación. Para nuestro ejercicio tomamos el 30 %.

5.5. Estimación Swap libre de riesgo

Para el desarrollo metodológico que nos corresponde de ahora en adelante, supongamos que desde la tesorería de Bancolombia están consolidando los saldos de sus inversiones al cierre del día de hoy. Dentro de su análisis les interesa saber para el siguiente día hábil en cuanto debería estar el valor del swap IBR que tienen contratado con BBVA Colombia. Entre otras cosas, porque quieren estimar las necesidades futuras de liquidez con el objetivo de tomar decisiones a tiempo y de ser necesario hacer las coberturas requeridas para ello. Además, se sabe que Bancolombia es quien recibe la tasa fija, la cual, fue pactada hace 6 meses al 4.25 % nominal semestral. También, que el tiempo restante del swap son 5 años y que los pagos son semestrales.

En consecuencia, el valor del swap libre de riesgo o valor sin default (VND), que nos concierne encontrar se define como el valor presente neto de los flujos futuros esperados en el contrato. Por tanto, el VND para Bancolombia estará dado por la ecuación **3.1.1** y para BBVA Colombia por la ecuación **3.1.2**.

Valoración con curva ARIMA-GARCH

De acuerdo con la nomenclatura referenciada en la ecuación **3.1.1** tenemos la siguiente información:

- N : 1.000.000.000 .
- K^* : 2.125 % .
- L : Curva forward pronosticada en la sección 5.2 .
- D^* : Factores de descuento implícitos de la curva forward.
- t : 10 semestres.

*Tasa efectiva semestral aplicable.

* $D(t, T_i) = \prod_{i=0}^T \frac{1}{1+F(t;T_i,T_{i+1})}$

FC Receptor Tasa fija (Bancolombia)					
Semestre	Derecho	Obligación	Neto	D	VPN
1	\$21.250.000	(\$-20.750.450)	\$499.550	0.9797	\$489.395
2	\$21.250.000	(\$-21.255.460)	(\$-5.460)	0.9593	(\$-5.238)
3	\$21.250.000	(\$-22.328.587)	(\$-1.078.587)	0.9383	(\$-1.012.070)
4	\$21.250.000	(\$-18.978.592)	\$2.271.408	0.9209	\$2.091.633
5	\$21.250.000	(\$-20.583.917)	\$666.083	0.9023	\$600.994
6	\$21.250.000	(\$-23.603.289)	(\$-2.353.289)	0.8815	(\$-2.074.366)
7	\$21.250.000	(\$-23.517.289)	(\$-2.267.289)	0.8612	(\$-1.952.638)
8	\$21.250.000	(\$-22.862.359)	(\$-1.612.359)	0.8420	(\$-1.357.561)
9	\$21.250.000	(\$-23.033.688)	(\$-1.783.688)	0.8230	(\$-1.468.002)
10	\$21.250.000	(\$-24.837.349)	(\$-3.587.349)	0.8031	(\$-2.880.888)
TOTAL					(\$-7.568.741)

Tabla 5.5.1.: VND ARIMA-GARCH.

De acuerdo con la curva ajustada en la sección 5.2, encontramos que para Bancolombia, el valor del swap libre de riesgo es de -7.568.741, lo que representa una pérdida aproximada del 0.76 % del nominal en cuestión. Por otro lado, para BBVA Colombia, el valor del swap libre de riesgo es el mismo pero con signos opuestos, lo que resultaría en una ganancia de 7.568.741.

Ahora vamos a repetir el ejercicio que acabamos de realizar, pero esta vez considerando la curva de tasas forward estimada utilizando la red neuronal.

Valoración con curva RNN

Aplicando una vez más el mismo procedimiento, salvo que ahora utilizaremos la curva pronosticada en la sección 5.3 , se obtienen los siguientes resultados:

FC Receptor Tasa fija (Bancolombia)					
Semestre	Derecho	Obligación	Neto	D	VPN
1	\$21.250.000	(\$-20.762.395)	\$487.606	0.9797	\$477.688
2	\$21.250.000	(\$-21.143.959)	\$106.041	0.9594	\$101.733
3	\$21.250.000	(\$-22.086.901)	(\$-836.901)	0.9386	(\$-785.551)
4	\$21.250.000	(\$-19.291.687)	\$1.958.313	0.9209	\$1.803.367
5	\$21.250.000	(\$-20.776.501)	\$473.499	0.9021	\$427.160
6	\$21.250.000	(\$-23.589.696)	(\$-2.339.696)	0.8813	(\$-2.062.077)
7	\$21.250.000	(\$-23.830.659)	(\$-2.580.659)	0.8608	(\$-2.221.508)
8	\$21.250.000	(\$-23.202.948)	(\$-1.952.948)	0.8413	(\$-1.643.033)
9	\$21.250.000	(\$-23.857.341)	(\$-2.607.341)	0.8217	(\$-2.142.466)
10	\$21.250.000	(\$-26.158.678)	(\$-4.908.678)	0.8008	(\$-3.930.666)
TOTAL					(\$-9.975.354)

Tabla 5.5.2.: VND RNN.

Finalmente, se obtiene que para Bancolombia, el valor del swap libre de riesgo a partir de la curva RNN es de -9.975.354, lo que representa una pérdida aproximada del 1.00 % del nominal en cuestión. Mientras que, para BBVA Colombia, el valor del swap libre de riesgo es el mismo pero con signos opuestos, lo que resultaría en una ganancia de 9.975.354.

5.6. Estimación CVA/DVA

Al examinar detenidamente la ecuación 3.6.2 utilizada para el cálculo del $CVA(t)$, se puede evidenciar que hasta este punto hemos progresado en dos elementos clave para su determinación. En primer lugar, hemos obtenido las probabilidades de default entre los distintos nodos ($PD(t_{i-1}, t_i)$), Las cuales fueron calculadas en la sección 5.4 utilizando los correspondientes Asset Swap Spread. Además, hemos establecido la tasa de recuperación (R) para las contrapartes locales, fijándola en un 30 %.

Del mismo modo, es importante destacar que aún nos falta considerar un término adicional en la ecuación del CVA: $EPE(t, t_i)$. Este término representa el valor presente de la exposición potencial futura en caso de un evento de incumplimiento. Para ser más precisos, podemos expresarlo de la siguiente manera:

$$EPE(t, t_i) = EPF(t, t_i)D(t, t_i), \quad (5.6.1)$$

donde $D(t, t_i)$ corresponde a los factores de descuento utilizados en la sección anterior y se encuentran en función de la estructura de tasas forward de las secciones 5.2 y 5.3. Por otro

lado, el término $EPF(t, t_i)$ se define de la siguiente forma:

$$EPF(t, t_i) = \text{máx}(N(\text{Tasa derecho}(t_i) - \text{Tasa obligación}(t_i)), 0). \quad (5.6.2)$$

La ecuación **5.6.2** representa el cálculo de la exposición potencial futura (EPF) en un momento específico t , en relación con un evento de incumplimiento que podría ocurrir en el momento t_i . Es decir, establece el mecanismo para estimar el valor que tendríamos en riesgo en el momento t_i ante la materialización del incumplimiento.

A partir de las definiciones anteriores procedemos a calcular los respectivos CVA para ambas contrapartes.

Estimación $CVA(t)$ Bancolombia con curva ARIMA-GARCH

CVA Bancolombia/ DVA BBVA Colombia					
Semestre	(1-R)	PD_C	EPF	D	CVA(t)
1	70 %	1.62 %	\$499.550	0.9797	\$5.547
2	70 %	1.60 %	\$0	0.9593	\$0
3	70 %	1.58 %	\$0	0.9383	\$0
4	70 %	1.56 %	\$2.271.408	0.9209	\$22.912
5	70 %	1.55 %	\$666.083	0.9023	\$6.509
6	70 %	1.53 %	\$0	0.8815	\$0
7	70 %	1.51 %	\$0	0.8612	\$0
8	70 %	1.50 %	\$0	0.8420	\$0
9	70 %	1.48 %	\$0	0.8230	\$0
10	70 %	1.46 %	\$0	0.8031	\$0
TOTAL					\$34.968

Tabla 5.6.1.: CVA Bancolombia con curva ARIMA-GARCH.

El cálculo del CVA que acabamos de realizar representa el valor presente de la pérdida esperada en el evento de incumplimiento por parte de BBVA Colombia. Es decir, al estimar el CVA de Bancolombia, simultáneamente estamos estimando el DVA de BBVA Colombia.

Estimación $CVA(t)$ BBVA Colombia con curva ARIMA-GARCH

CVA BBVA Colombia / DVA Bancolombia					
Semestre	(1-R)	PD_C	EPF	D	$CVA(t)$
1	70 %	1.12 %	\$0	0.9797	\$0
2	70 %	1.12 %	\$5.460	0.9593	\$41
3	70 %	1.11 %	\$1.078.587	0.9383	\$7.839
4	70 %	1.10 %	\$0	0.9209	\$0
5	70 %	1.09 %	\$0	0.9023	\$0
6	70 %	1.08 %	\$2.353.289	0.8815	\$15.690
7	70 %	1.07 %	\$2.267.289	0.8612	\$14.654
8	70 %	1.06 %	\$1.612.359	0.8420	\$10.108
9	70 %	1.06 %	\$1.783.688	0.8230	\$10.844
10	70 %	1.05 %	\$3.587.349	0.8031	\$21.114
TOTAL					\$80.290

Tabla 5.6.2.: CVA BBVA Colombia con curva ARIMA-GARCH.

El cálculo del CVA que acabamos de realizar representa el valor presente de la pérdida esperada en el evento de incumplimiento por parte de Bancolombia. Es decir, al estimar el CVA de BBVA Colombia, simultáneamente estamos estimando el DVA de Bancolombia.

Ahora que hemos realizado los ajustes de valoración necesarios, podemos determinar el valor razonable del swap IBR entre Bancolombia y BBVA Colombia. Este valor se define como:

$$\text{Valor Swap} = VND - CVA + DVA. \quad (5.6.3)$$

Tenemos entonces que para el que recibe la tasa fija (Bancolombia):

$$\text{Valor Swap} = -7.568.741 - 34.968 + 80.290 = -7.523.419$$

y para quien recibe la tasa variable (BBVA Colombia):

$$\text{Valor Swap} = 7.568.741 - 80.290 + 34.968 = 7.523.419 .$$

Ahora, procederemos a repetir el ejercicio, teniendo en cuenta la curva de tasas forward estimada con la red neuronal.

Estimación $CVA(t)$ Bancolombia con curva RNN

CVA Bancolombia / DVA BBVA Colombia					
Semestre	(1-R)	PD_C	EPF	D	CVA(t)
1	70 %	1.62 %	\$487.606	0.9797	\$5.415
2	70 %	1.60 %	\$106.041	0.9594	\$1.140
3	70 %	1.58 %	\$0	0.9386	\$0
4	70 %	1.56 %	\$1.958.313	0.9209	\$19.754
5	70 %	1.55 %	\$473.499	0.9021	\$4.626
6	70 %	1.53 %	\$0	0.8813	\$0
7	70 %	1.51 %	\$0	0.8608	\$0
8	70 %	1.50 %	\$0	0.8413	\$0
9	70 %	1.48 %	\$0	0.8217	\$0
10	70 %	1.46 %	\$0	0.8008	\$0
TOTAL					\$30.935

Tabla 5.6.3.: CVA Bancolombia con curva RNN.

Estimación $CVA(t)$ BBVA Colombia con curva RNN

CVA BBVA Colombia / DVA Bancolombia					
Semestre	(1-R)	PD_C	EPF	D	CVA(t)
1	70 %	1.12 %	\$0	0.9797	\$0
2	70 %	1.12 %	\$0	0.9594	\$0
3	70 %	1.11 %	\$836.901	0.9386	\$6.085
4	70 %	1.10 %	\$0	0.9209	\$0
5	70 %	1.09 %	\$0	0.9021	\$0
6	70 %	1.08 %	\$2.339.696	0.8813	\$15.598
7	70 %	1.07 %	\$2.580.659	0.8608	\$16.672
8	70 %	1.06 %	\$1.952.948	0.8413	\$12.233
9	70 %	1.06 %	\$2.607.341	0.8217	\$15.826
10	70 %	1.05 %	\$4.908.678	0.8008	\$28.807
TOTAL					\$95.221

Tabla 5.6.4.: CVA BBVA Colombia con curva RNN.

Finalmente, aplicando la ecuación **5.6.3** tenemos el valor del swap a partir de la curva RNN:

Para el que recibe la tasa fija (Bancolombia):

$$\mathbf{Valor\ Swap} = -9.975.354 - 30.935 + 95.221 = -9.911.068$$

y para quien recibe la tasa variable (BBVA Colombia):

$$\mathbf{Valor\ Swap} = 9.975.354 - 95.221 + 30.935 = 9.911.068 .$$

6. Conclusiones y recomendaciones

6.1. Conclusiones

En este trabajo, se realizaron pronósticos de las tasas forward IBR a 6 meses para un horizonte de cinco años utilizando modelos ARIMA-GARCH y modelos de redes neuronales recurrentes. Para llevar a cabo esta aplicación, se identificaron 10 series de tiempo, correspondientes a cada fecha de liquidación e intercambios de flujos del swap, el cual fue el activo objeto de investigación.

Al analizar las métricas de error utilizadas en el entrenamiento de ambos modelos en las 10 series de tiempo objeto de estudio, se observó que el modelo ARIMA-GARCH exhibió errores menores en las tres métricas evaluadas (MSE, MAE y MAPE). Estos resultados indican que el modelo ARIMA-GARCH es más adecuado para el propósito de nuestra investigación, que consiste en estimar valores de referencia de swaps con vencimiento a 5 años. Sin embargo, aunque nuestros hallazgos respaldan la elección de los modelos autorregresivos, se recomienda una evaluación cuidadosa de los modelos en futuras investigaciones en función de las condiciones cambiantes del mercado.

Así mismo, al analizar la valoración del swap libre de riesgo (VND) en los nodos de la parte corta (0 a 3 años), se puede observar que las diferencias en el valor neto de los flujos entre obligaciones y derechos no presentan una diferencia sustancial, es decir, tanto el modelo ARIMA-GARCH como el modelo RNN generan pronósticos relativamente similares. Sin embargo, las mayores diferencias se presentaron en los nodos 9 y 10, donde el modelo RNN arrojó un pronóstico mayor y, por ende, un mayor valor en la valoración final. Es importante destacar que el rendimiento de la red neuronal es altamente sensible a la elección de la función de activación. Durante el proceso de ajuste de la red, optamos por utilizar la función de activación ReLU. No obstante, existe la posibilidad de que explorando otras opciones de funciones de activación podamos mejorar aún más el rendimiento de la red.

Para incorporar los ajustes por riesgo de crédito en la valoración del swap, se calcularon las probabilidades de default en cada nodo para ambas contrapartes, basándose en sus respectivos asset swap spread en dólares. Este análisis reveló que Bancolombia presenta menores probabilidades de default en comparación con BBVA Colombia, lo cual es comprensible, ya que Bancolombia es un banco más grande y cuenta con un nivel de fondeo mayor en dólares.

Del mismo modo, tras calcular el CVA y DVA e integrarlos en la valoración del swap, se observaron resultados diferentes en función del modelo empleado. Con el modelo ARIMA-GARCH, la valoración del swap se situó en aproximadamente un 0.75 % sobre el valor nominal, lo que puede representar una pérdida o ganancia dependiendo del lado del swap en que se encuentre el inversionista. Por otro lado, el modelo RNN estimó una valoración del 1 % sobre el valor nominal.

En conclusión de nuestra investigación, para la estimación de valores de referencia de swaps, encontramos que tanto las redes neuronales como los modelos autorregresivos son efectivos en vencimientos de 0 a 3 años, ya que ambos métodos producen predicciones adecuadas en este tramo de la curva. Sin embargo, cuando se trata de swaps con vencimientos superiores a 3 años, los modelos autorregresivos presentan un mejor rendimiento, convirtiéndose en la opción más adecuada.

6.2. Recomendaciones

Para futuros análisis, resultaría interesante considerar un swap con un horizonte temporal más corto, por ejemplo, a 3 años. Esto reduciría el tiempo de computación requerido para implementar las redes neuronales y permitiría explorar diferentes configuraciones de hiperparámetros con el objetivo de mejorar la precisión de los pronósticos.

Además, sería beneficioso ampliar el rango de información utilizado en el estudio. Al hacerlo, se podría contar con una base de entrenamiento más extensa para las redes neuronales, lo que potencialmente facilitaría la detección de patrones adicionales y complejos en los datos.

A. Anexo: Modelos ARIMA-GARCH

NODO 0-6M

$$F(t; 0, 6) \sim \text{ARIMA}(0,1,1)\text{-GARCH}(1,1)\text{-NIG}$$

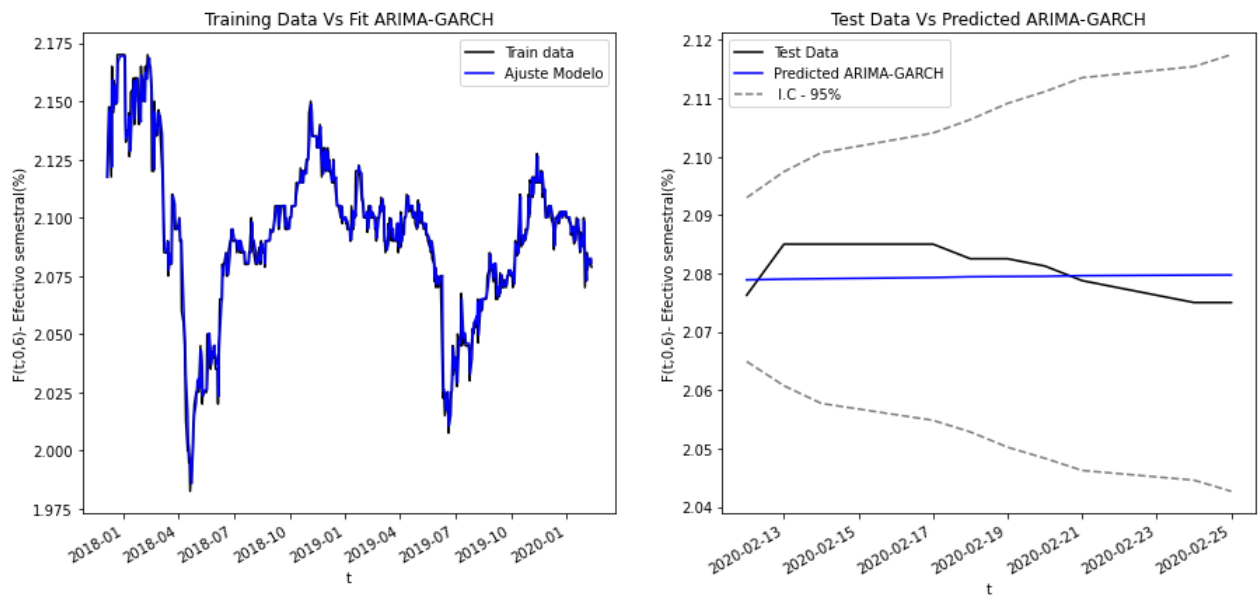


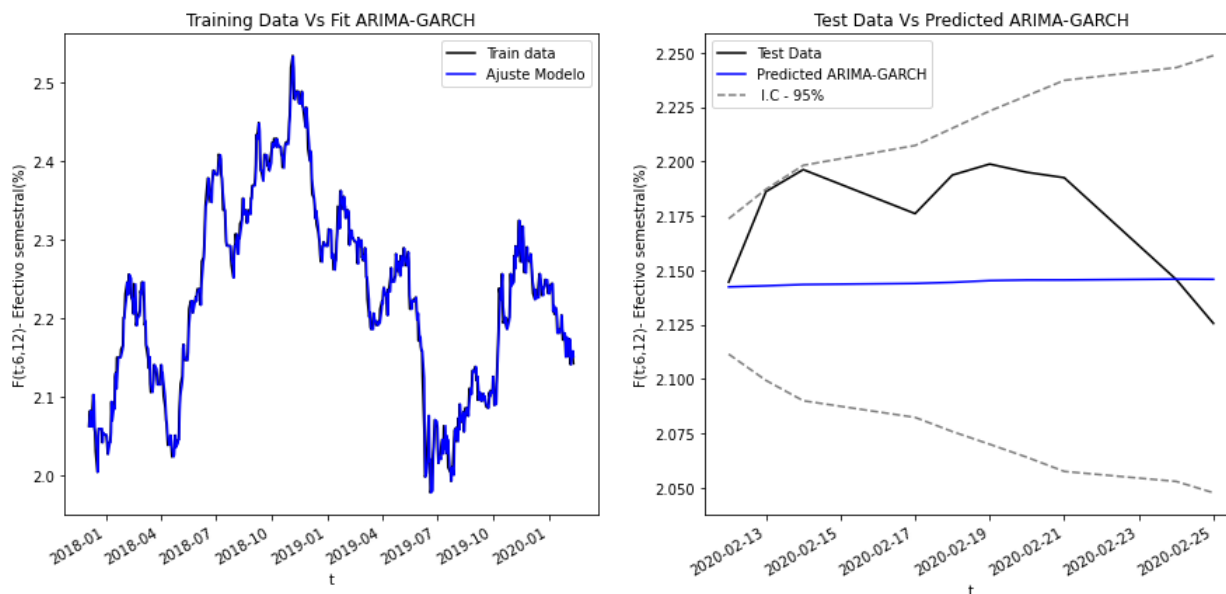
Figura A.0.1.: Nodo 0-6M ARIMA-GARCH.

Métrica	MSE	MAE	MAPE
Resultado	0.0000178	0.0038421	0.00184611

Tabla A.0.1.: Métricas de desempeño.

NODO 6-12M

$$F(t; 6, 12) \sim \text{ARIMA}(0,1,0)\text{-GARCH}(1,1)\text{-NIG}$$

**Figura A.0.2.:** Nodo 6-12M ARIMA-GARCH.

Métrica	MSE	MAE	MAPE
Resultado	0.0016068	0.0350149	0.01600394

Tabla A.0.2.: Métricas de desempeño.

NODO 12-18M

$$F(t; 12, 18) \sim \text{ARIMA}(0,1,1)\text{-GARCH}(1,0)\text{-NIG}$$

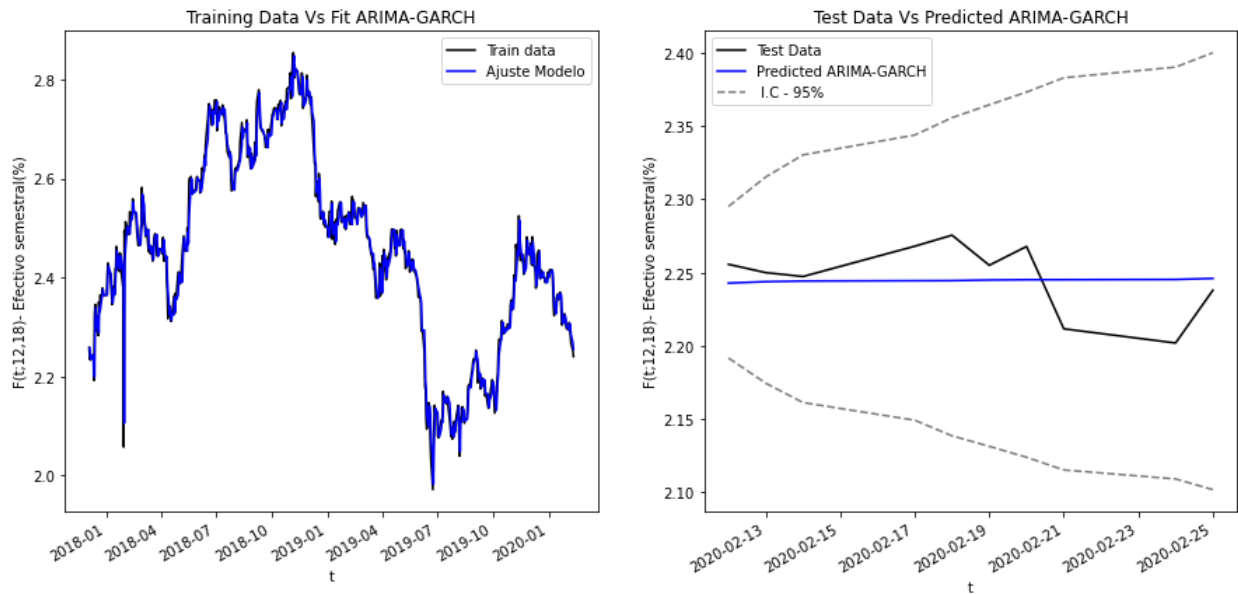


Figura A.0.3.: Nodo 12-18M ARIMA-GARCH.

Métrica	MSE	MAE	MAPE
Resultado	0.0005414	0.0194393	0.00867546

Tabla A.0.3.: Métricas de desempeño.

NODO 18-24M

$$F(t; 18, 24) \sim \text{ARIMA}(0,1,1)\text{-GARCH}(1,0)\text{-NIG}$$

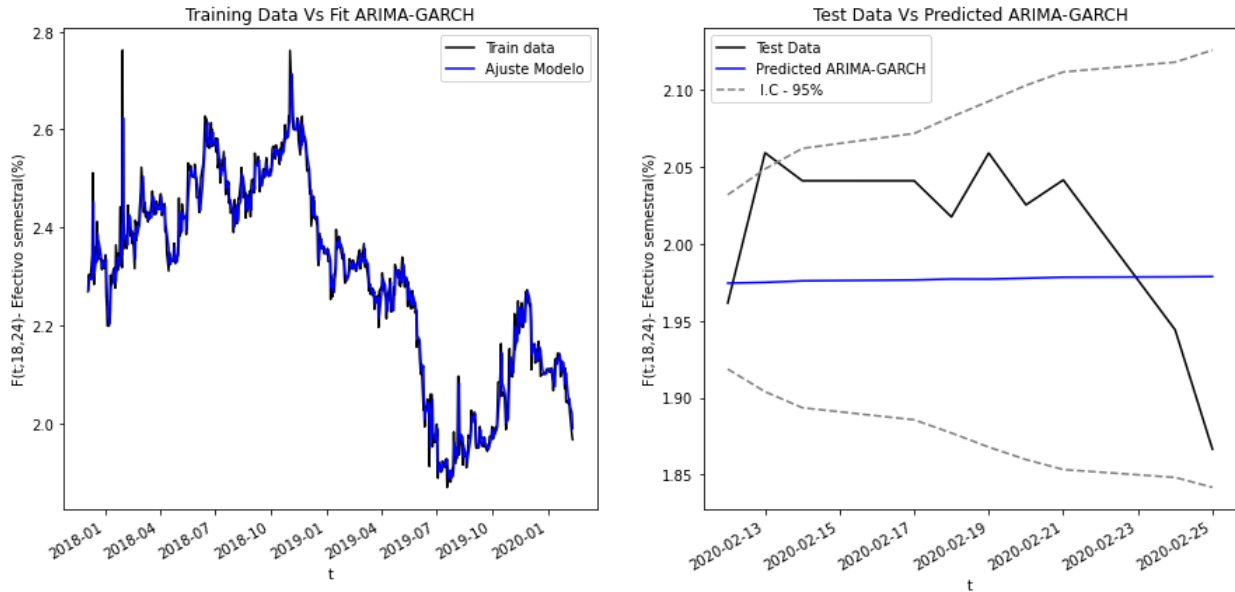


Figura A.0.4.: Nodo 18-24M ARIMA-GARCH.

Métrica	MSE	MAE	MAPE
Resultado	0.0043875	0.0605397	0.03025509

Tabla A.0.4.: Métricas de desempeño.

NODO 24-30M

$$F(t; 24, 30) \sim \text{ARIMA}(2,1,2)\text{-GARCH}(0,1)\text{-NIG}$$

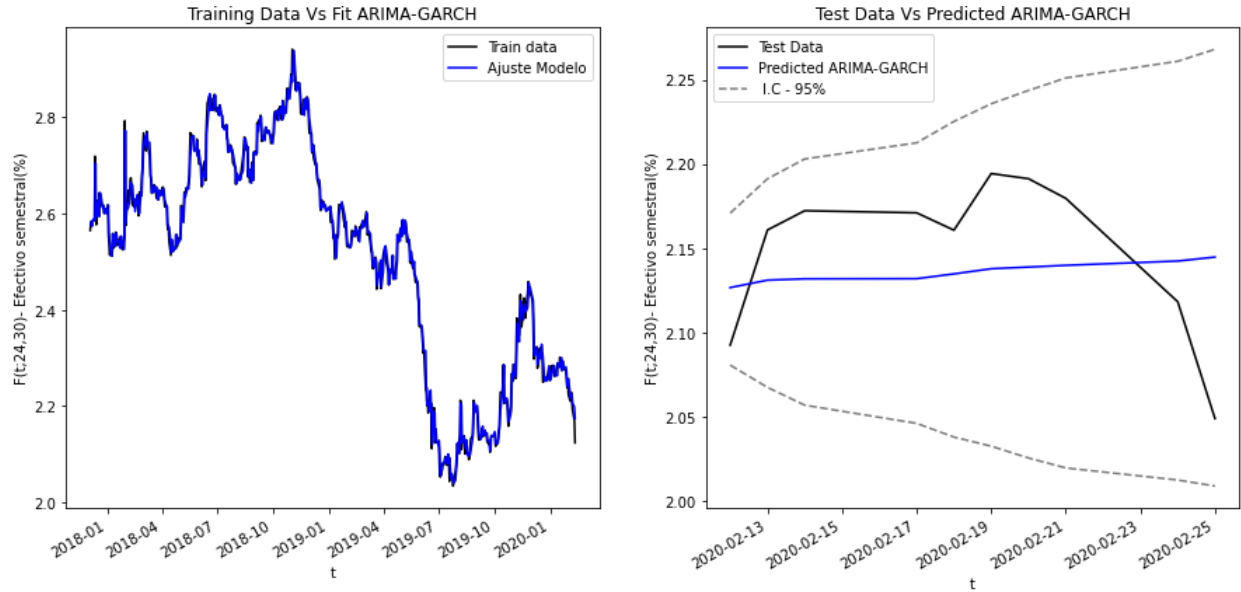


Figura A.0.5.: Nodo 24-30M ARIMA-GARCH.

Métrica	MSE	MAE	MAPE
Resultado	0.0023223	0.0438457	0.020501697

Tabla A.0.5.: Métricas de desempeño.

NODO 30-36M

$$F(t; 30, 36) \sim \text{ARIMA}(0,1,2)\text{-GARCH}(1,1)\text{-NIG}$$

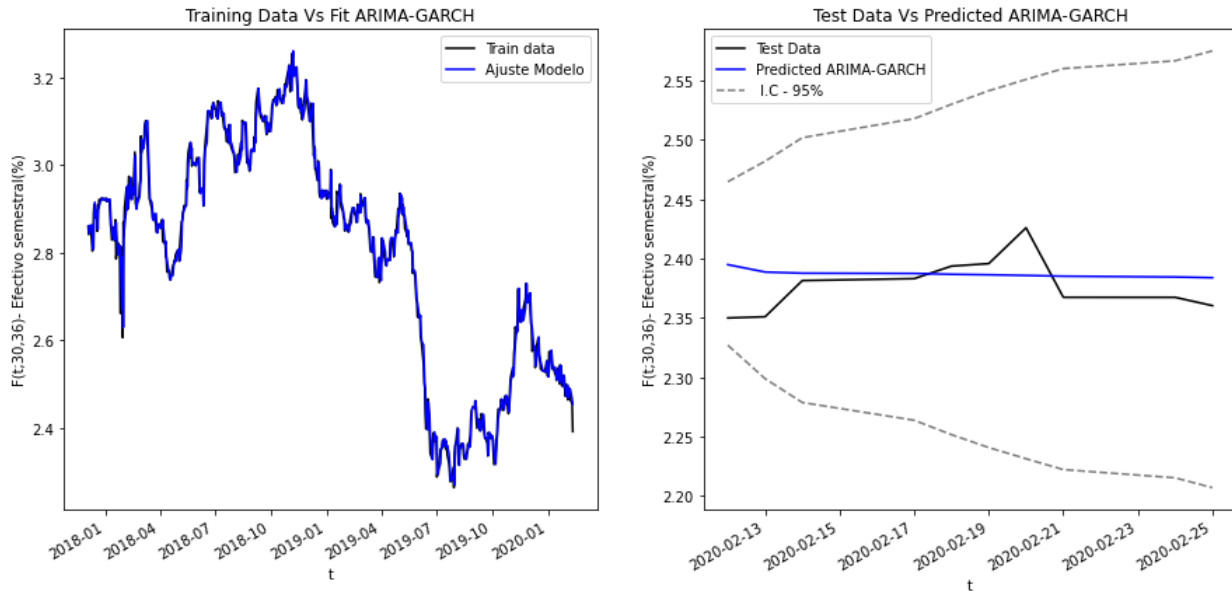


Figura A.0.6.: Nodo 30-36M ARIMA-GARCH.

Métrica	MSE	MAE	MAPE
Resultado	0.0006412	0.0208169	0.008769987

Tabla A.0.6.: Métricas de desempeño.

NODO 36-42M

$$F(t; 36, 42) \sim \text{ARIMA}(1,1,0)\text{-GARCH}(1,0)\text{-NIG}$$

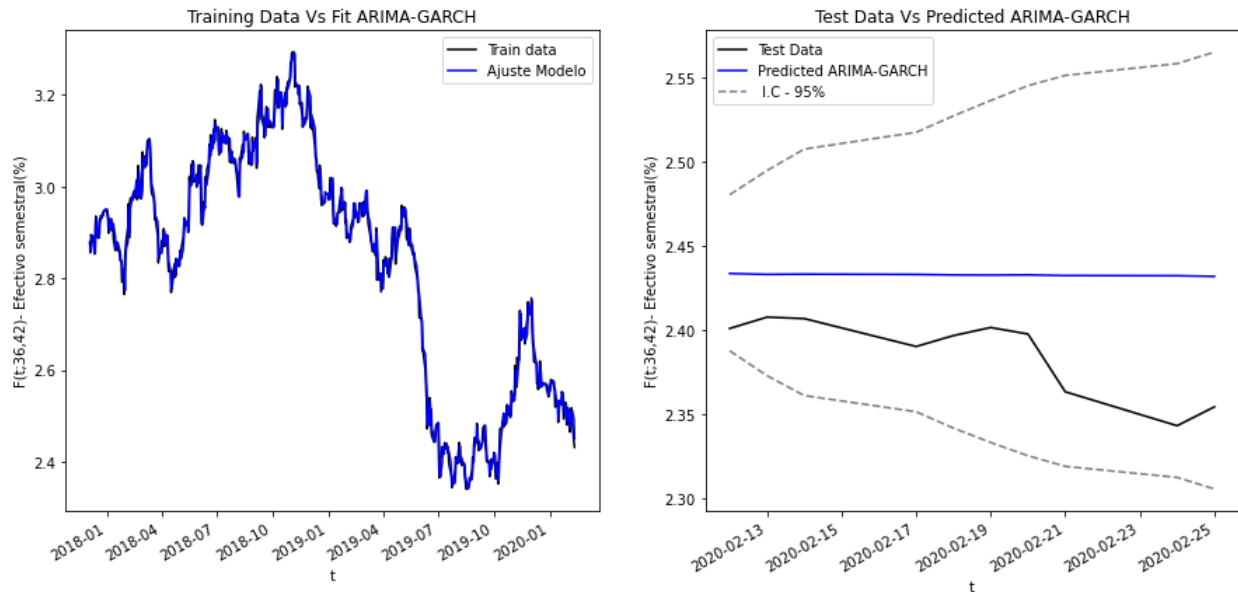


Figura A.0.7.: Nodo 36-42M ARIMA-GARCH.

Métrica	MSE	MAE	MAPE
Resultado	0.0026490	0.0465515	0.01959844

Tabla A.0.7.: Métricas de desempeño.

NODO 42-48M

$$F(t; 42, 48) \sim \text{ARIMA}(0,1,1)\text{-GARCH}(1,0)\text{-NIG}$$

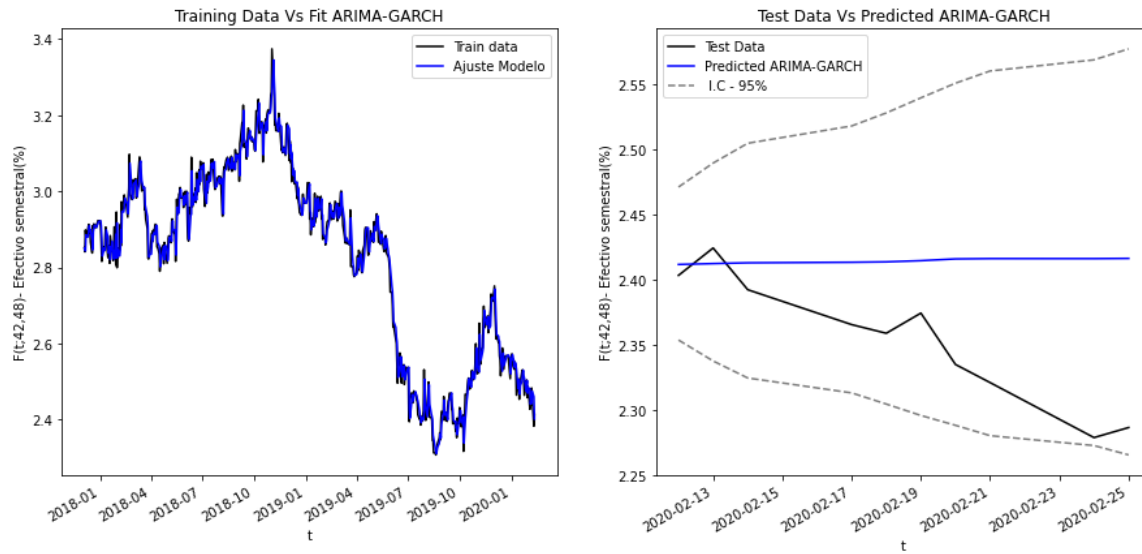


Figura A.0.8.: Nodo 42-48M ARIMA-GARCH.

Métrica	MSE	MAE	MAPE
Resultado	0.0058951	0.0627562	0.02703668

Tabla A.0.8.: Métricas de desempeño.

NODO 48-54M

$$F(t; 48, 54) \sim \text{ARIMA}(0,1,1)\text{-GARCH}(1,0)\text{-NIG}$$

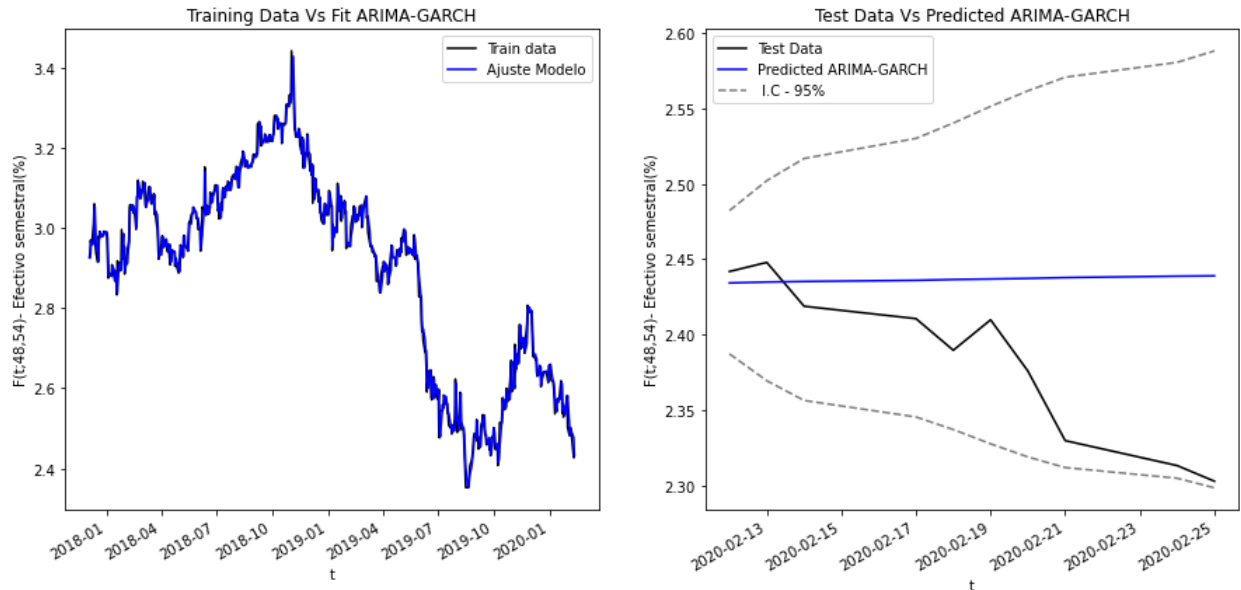


Figura A.0.9.: Nodo 48-54M ARIMA-GARCH.

Métrica	MSE	MAE	MAPE
Resultado	0.0053855	0.0567732	0.02423153

Tabla A.0.9.: Métricas de desempeño.

NODO 54-60M

$$F(t; 54, 60) \sim \text{ARIMA}(0,1,1)\text{-GARCH}(1,0)\text{-NIG}$$

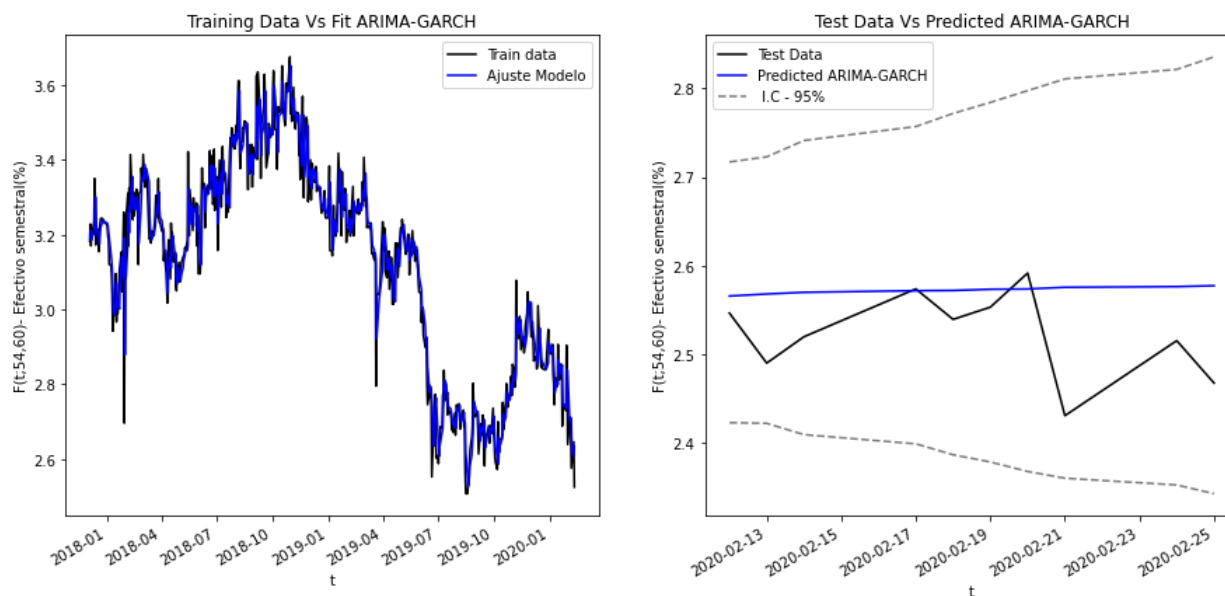


Figura A.0.10.: Nodo 54-60M ARIMA-GARCH.

Métrica	MSE	MAE	MAPE
Resultado	0.0047512	0.0535748	0.02155543

Tabla A.0.10.: Métricas de desempeño.

B. Anexo: RNN

NODO 0-6M

	LSTM	GRU	Vanilla
MAPE	0.042852	0.058499	0.040554
MSE	0.000595	0.001193	0.000515
MAE	0.022054	0.030417	0.021174

Figura B.0.1.: Métricas 0-6M.

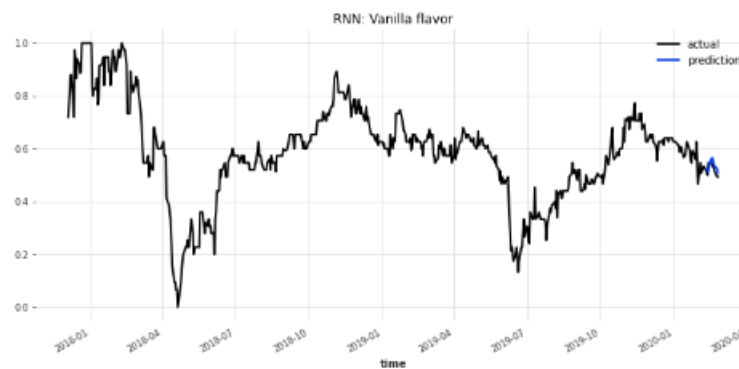


Figura B.0.2.: RNN 0-6M.

- Epoch: 1000
- Hidden-dim: 100
- n-rnn-layers: 10
- Dropout: 0.2
- optimizer-kwarg: 0.001

NODO 6-12M

	LSTM	GRU	Vanilla
MAPE	0.100658	0.166875	0.133528
MSE	0.002144	0.005279	0.003397
MAE	0.034019	0.054270	0.047359

Figura B.0.3.: Métricas 6-12M.



Figura B.0.4.: RNN 6-12M.

- Epoch: 1000
- Hidden-dim: 100
- n-rnn-layers: 10
- Dropout: 0.2
- optimizer-kwarg: 0.001

NODO 12-18M

	LSTM	GRU	Vanilla
MAPE	0.109913	0.104344	0.107593
MSE	0.001499	0.001726	0.001315
MAE	0.033853	0.030330	0.032926

Figura B.0.5.: Métricas 12-18M.

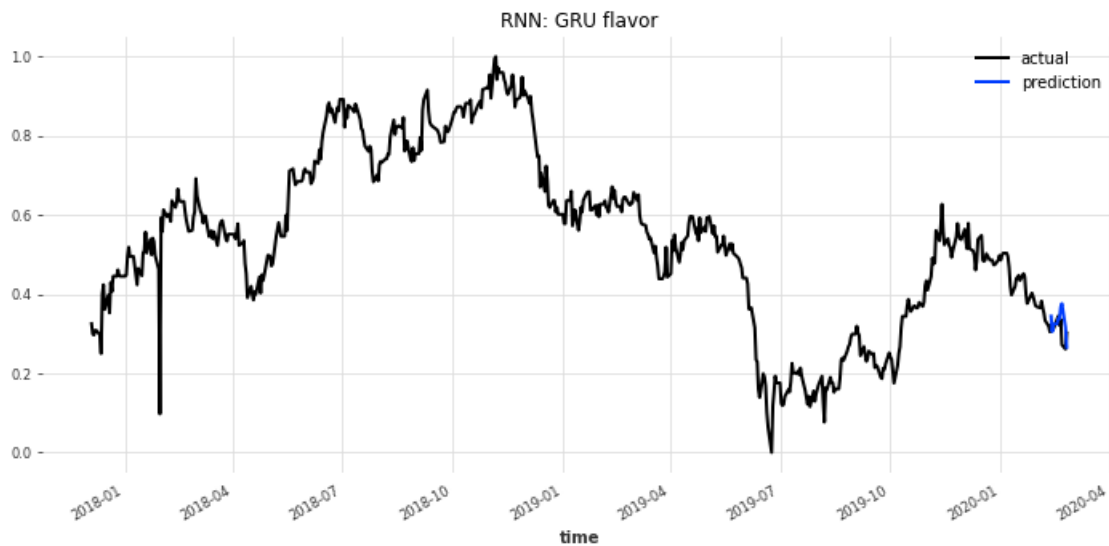


Figura B.0.6.: RNN 12-18M.

- Epoch: 1000
- Hidden-dim: 100
- n-rnn-layers: 10
- Dropout: 0.2
- optimizer-kwarg: 0.001

NODO 18-24M

	LSTM	GRU	Vanilla
MAPE	6.187955	2.810371	3.133283
MSE	0.005608	0.004997	0.006082
MAE	0.052942	0.061734	0.070416

Figura B.0.7.: Métricas 18-24M.



Figura B.0.8.: RNN 18-24M.

- Epoch: 1000
- Hidden-dim: 100
- n-rnn-layers: 10
- Dropout: 0.2
- optimizer-kwarg: 0.001

NODO 24-30M

	LSTM	GRU	Vanilla
MAPE	1.278670	0.680034	0.719505
MSE	0.005552	0.001858	0.001635
MAE	0.062043	0.039582	0.032066

Figura B.0.9.: Métricas 24-30M.

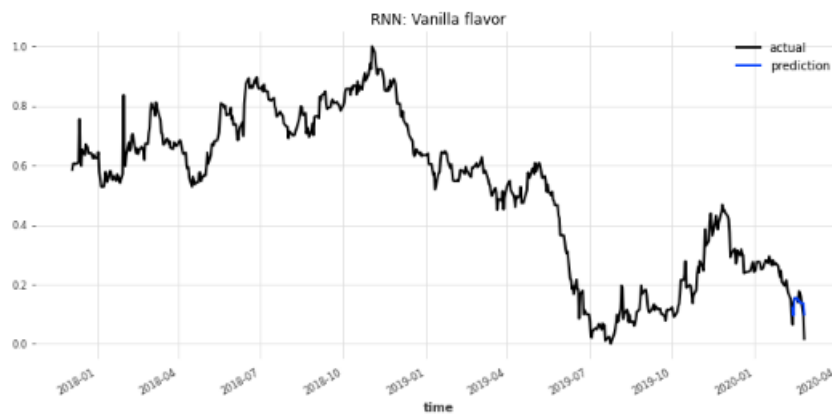


Figura B.0.10.: RNN 24-30M.

- Epoch: 1500
- Hidden-dim: 120
- n-rnn-layers: 5
- Dropout: 0.2
- optimizer-kwargs: 0.001

NODO 30-36M

	LSTM	GRU	Vanilla
MAPE	0.293607	0.143302	0.286797
MSE	0.002441	0.000441	0.001248
MAE	0.028274	0.016994	0.032896

Figura B.0.11.: Métricas 30-36M.

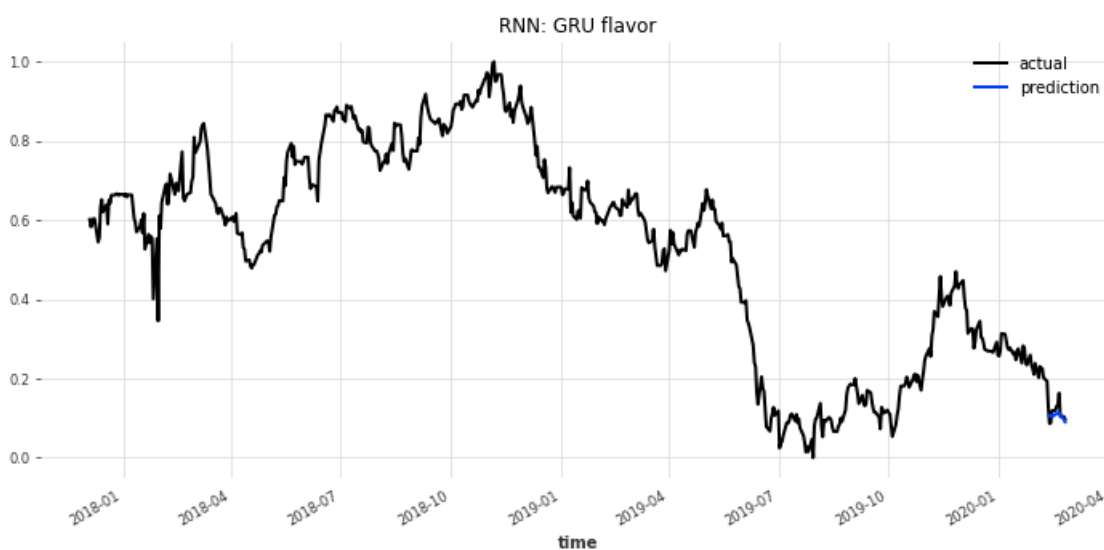


Figura B.0.12.: RNN 30-36M.

- Epoch: 1500
- Hidden-dim: 120
- n-rnn-layers: 5
- Dropout: 0.2
- optimizer-kwarg: 0.001

NODO 36-42M

	LSTM	GRU	Vanilla
MAPE	2.179528	2.861779	1.741358
MSE	0.000788	0.000987	0.000319
MAE	0.024236	0.025479	0.014048

Figura B.0.13.: Métricas 36-42M.

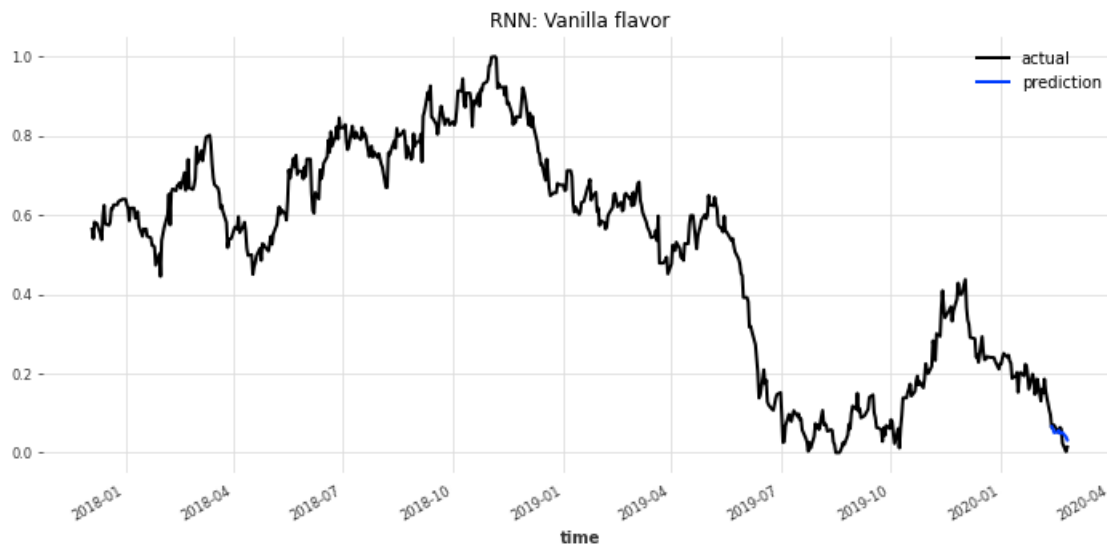


Figura B.0.14.: RNN 36-42M.

- Epoch: 1500
- Hidden-dim: 120
- n-rnn-layers: 5
- Dropout: 0.2
- optimizer-kwarg: 0.001

NODO 42-48M

	LSTM	GRU	Vanilla
MAPE	1.853019	2.618335	0.602609
MSE	0.004684	0.006147	0.000538
MAE	0.053550	0.063593	0.019499

Figura B.0.15.: Métricas 42-48M.



Figura B.0.16.: RNN 42-48M.

- Epoch: 1500
- Hidden-dim: 120
- n-rnn-layers: 5
- Dropout: 0.2
- optimizer-kwarg: 0.001

NODO 48-54M

	LSTM	GRU	Vanilla
MAPE	1.663367	3.471810	1.335593
MSE	0.004801	0.017474	0.002683
MAE	0.059046	0.117892	0.044555

Figura B.0.17.: Métricas 48-54M.



Figura B.0.18.: RNN 48-54M.

- Epoch: 1500
- Hidden-dim: 120
- n-rnn-layers: 5
- Dropout: 0.2
- optimizer-kwarg: 0.001

NODO 54-60M

	LSTM	GRU	Vanilla
MAPE	5.837913	8.386401	4.207344
MSE	0.015246	0.029471	0.008702
MAE	0.112070	0.164618	0.084007

Figura B.0.19.: Métricas 54-60M.

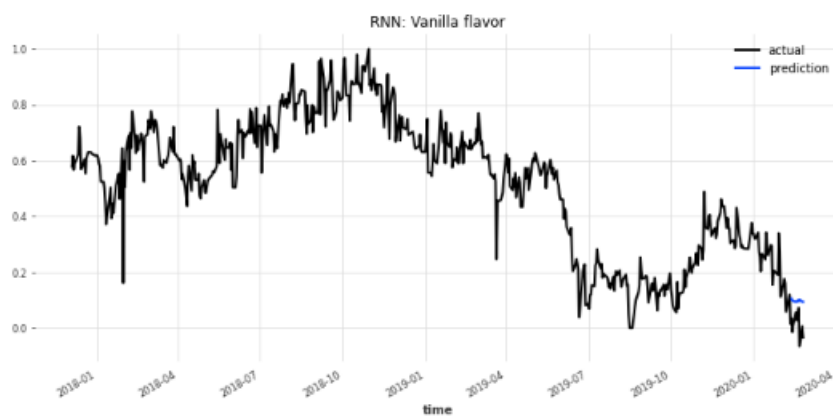


Figura B.0.20.: RNN 54-60M.

- Epoch: 2500
- Hidden-dim: 120
- n-rnn-layers: 5
- Dropout: 0.2
- optimizer-kwarg: 0.001

C. Anexo: Características Emisiones

Emisor	ISIN	Moneda	Vencimiento	Cupón	Frecuencia Cupón
Bancolombia	US05968LAM46	USD	29/01/2025	3.000 %	Semestral
BBVA Colombia	USP1024TAN92	USD	21/04/2025	4.875 %	Semestral
Colombia	US195325BD67	USD	21/05/2024	8.125 %	Semestral
Colombia	COL17CT02385	COP	24/07/2024	10.00 %	Anual

D. Anexo: Código Python RNN

El código adjunto es la representación de la información relacionada al nodo 1, para los demás nodos lo único que cambia es la configuración de los hiperparámetros y la información que ingresa a las funciones. Es decir, la arquitectura del código se conserva.

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import statsmodels as sm
from scipy.stats import shapiro
from scipy import stats
import warnings
import time
import itertools
warnings.filterwarnings("ignore")
import tensorflow as tf
import torch.optim as optim
from torch.utils.tensorboard import SummaryWriter
from darts.dataprocessing.transformers import Scaler
from darts.models import RNNModel, BlockRNNModel
from darts.metrics import mape, rmse, r2_score
from darts.utils.statistics import check_seasonality, plot_acf
from darts.utils.timeseries_generation import datetime_attribute_timeseries
from darts import TimeSeries
```

NODO 0-6M - RNN

```
data = pd.read_excel('./Forward.xlsx')
data.set_index('Fecha', inplace=True)
FC_START = "20191216"
y_t1 = data["0-6M"]
y_t1 = y_t1.squeeze()
y_t1_ts = TimeSeries.from_series(y_t1)
train, val = y_t1_ts.split_after(pd.Timestamp(FC_START))
```

```
EPOCH = 1000
trf = Scaler()
train_trf = trf.fit_transform(train)
val_trf = trf.transform(val)
ts_trf = trf.transform(y_t1_ts)

def run_RNN(flavor, ts, train, val):

    model_RNN = BlockRNNModel(
        model=flavor,
        model_name=flavor + str("RNN"),
        input_chunk_length=30,
        output_chunk_length=1,
        hidden_dim=100,
        n_rnn_layers=10,
        batch_size=64,
        n_epochs=EPOCH,
        dropout=0.2,
        optimizer_kwargs='lr': 1e-3,
        log_tensorboard=True,
        random_state=42,
        force_reset=True)
    if flavor == "RNN": flavor = "Vanilla"

    fit_it(model_RNN, train, val, flavor)
    model_RNN.save(flavor + str(".pt"))
    pred = model_RNN.historical_forecasts(
        ts_trf,
        start=pd.Timestamp("20200211"),
        forecast_horizon=1,
        stride=1,
        retrain=False,
        last_points_only=True,
        verbose=True,)

    plot_fitted(pred, ts, flavor)
    res_acc = accuracy_metrics(pred, ts)
    print(flavor + ": ")
    - = [print(k,":",f'v:.4f') for k,v in res_acc.items()]
```

```

return [pred, res_acc]

def fit_it(model, train, val, flavor):
    t_start = time.perf_counter()
    print("nbeginning the training of the 0 RNN:".format(flavor))

    res = model.fit(train,
                    val_series=val,
                    verbose=True)
    res_time = time.perf_counter() - t_start
    print("training of the 0 RNN has completed:".format(flavor), f'res_time:.2f sec')
    return res

def plot_fitted(pred, act, flavor):
    plt.figure(figsize=(12,5))
    act.plot(label='actual')
    pred.plot(label='prediction')
    plt.title("RNN: 0 flavor".format(flavor))
    plt.legend();

def accuracy_metrics(pred, act):
    act2 = act.slice_intersect(pred)
    pred2 = pred.slice_intersect(act2)
    resid = pred2 - act2
    sr = resid.pd_series()
    sa = act2.pd_series()
    sp = pred2.pd_series()
    res_mape = mape(pred2, act2)
    res_rmse = rmse(pred2, act2)
    mse_rnn = np.mean((sa-sp)**2)
    mae_rnn = np.mean(np.abs(sa-sp))
    mape_rnn = np.mean(np.abs((sa-sp)/sa))
    res_pe = sr / sa
    n_act = len(act2)
    res_rmspe = np.sqrt(np.sum(res_pe**2) / n_act)
    res_std = np.std(sr)
    res_se = res_std / np.sqrt(n_act)
    res_sefc = np.sqrt(res_std + res_se**2)

res_accuracy =

```

```
"MAPE":mape_rnn,"MSE":mse_rnn,  
"MAE": mae_rnn  
return res_accuracy
```

```
flavors = ["LSTM", "GRU", RNN]  
res_flavors = [run_RNN(flv, ts_trf, train_trf, val_trf) for flv in flavors]
```

```
def com_model(flavor, ts):  
    model_RNN = RNNModel.load(str("./Modelos RNN/Modelo 0-6M/") + flavor + str(".pt"))  
    pred = model_RNN.historical_forecasts(  
        ts_trf,  
        start=pd.Timestamp("20200212"),  
        forecast_horizon=1,  
        stride=1,  
        retrain=False  
        last_points_only=True,  
        verbose=True,)  
    plot_fitted(pred, ts, flavor)  
    res_acc = accuracy_metrics(pred, ts)  
    print(flavor + ": ")  
    _ = [print(k,":",f'v:.4f') for k,v in res_acc.items()]  
    return [pred, res_acc]  
res_flavors = [com_model(flv, ts_trf) for flv in flavors]
```

```
model_1 = RNNModel.load("./Modelos RNN/Modelo 0-6M/Vanilla.pt")  
pred_1 = model_1.predict(1,ts_trf)  
pred_1 = trf.inverse_transform(pred_1)
```

Referencias

- Aragall, E. (2013). Cva, dva y fva: Impacto del riesgo de contrapartida en la valoración de los derivados etc. *Observatorio de Divulgación Financiera*, pages 1–17.
- Asobancaria (2012). El indicador bancario de referencia (ibr). *Semana Económica 2012*, 874:1–14.
- BanRep (2020). Determinantes de las dinámicas de los mercados de capitales. *Documentos de trabajo de la Junta directiva del banco de la república*.
- Barrios, A. B. (2009). Los límites del dinero y su importancia como instrumento monetario en la economía. *ECONÓMICAS CUC*, 30(1):259–264.
- Batlle, C. B., Rodríguez, M. G., and Benedicto, T. P. (2020). Valor razonable de un swap: Cva y dva. una aproximación binomial. *Cuadernos de Economía*, 43(122):229–242.
- Bicksler, J. and Chen, A. H. (1986). An economic analysis of interest rate swaps. *The Journal of Finance*, 41(3):645–655.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.
- Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
- Brigo, D. and Mercurio, F. (2006). *Interest rate models-theory and practice: with smile, inflation and credit*, volume 2. Springer.
- Caballero, F. J. H. (2015). Modelación discreta de la estructura de tasas de interés nominales, el caso de Colombia 2004–2013. *ECONÓMICAS CUC*, 36(1):95–112.
- Cárdenas, J.-A., Cristiano-Botia, D. J., and Martínez-Cortés, N. (2023). Colombian inflation forecast using long short-term memory approach. *Borradores de Economía; No. 1241*.
- Choudry, M. (2006). *The Credit Default Swap Basis*. Wiley, Chichester, UK.
- Cornelissen, J. (2021). A study on forecasting sofr with a recurrent neural network using long short-term memory cells. Master's thesis, University of Twente.
- Cruz, I. B., Martínez, S. S., Abed, A. R., Ábalo, R. G., and Lorenzo, M. M. G. (2007). Redes neuronales recurrentes para el análisis de secuencias. *Revista Cubana de Ciencias Informáticas*, 1(4):48–57.
- de Arce, R. (1998). Introducción a los modelos autorregresivos con heterocedasticidad condicional (arch). *Programa de Doctorado en Modelización Económica. Instituto LR Klein. Universidad Autónoma de Madrid. Madrid. Consultado el*, 10.
- de la Fuente Rodríguez, J. (2018). Generalidades del dinero. *Revista de la Facultad de Derecho de México*, 68(271):789–808.

- Fernández Navarrete, J. A. (2017). Módulo valor del dinero en el tiempo: Interés simple e interés compuesto aplicaciones.
- García Doffing, F. L. (2021). Ajustes de valorización financiera, marco teórico y aplicaciones al medio local.
- Graves, A. (2013). Generating sequences with recurrent neural networks. *arXiv preprint arXiv:1308.0850*.
- Gregory, J. (2020). *The XVA Challenge: Counterparty Risk, Funding, Collateral, Capital and Initial Margin*. John Wiley & Sons.
- Ho, T. S. and Lee, S.-B. (1986). Term structure movements and pricing interest rate contingent claims. *the Journal of Finance*, 41(5):1011–1029.
- Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.
- Hull, J. and White, A. (1993). One-factor interest-rate models and the valuation of interest-rate derivative securities. *Journal of financial and quantitative analysis*, 28(2):235–254.
- Hull, J. C. (2018). *Options, Futures, and Other Derivatives*. Pearson, 8th edition.
- IFRS (2013). Norma internacional de información financiera 13: Medición del valor razonable.
- James, G., Witten, D., Hastie, T., and Tibshirani, R. (2021). *An introduction to statistical learning*, volume 112. Springer.
- Leung, M. T., Chen, A.-S., and Daouk, H. (2000). Forecasting exchange rates using general regression neural networks. *Computers & Operations Research*, 27(11-12):1093–1110.
- Luna Proenza, J. M. (2012). Indicador bancario de referencia: desarrollo de un nuevo indicador del mercado monetario en colombia.
- Matich, D. J. (2001). Redes neuronales: Conceptos básicos y aplicaciones. *Universidad Tecnológica Nacional, México*, 41.
- Minton, B. A. (1997). An empirical examination of basic valuation models for plain vanilla us interest rate swaps. *Journal of Financial Economics*, 44(2):251–277.
- Mondragón, G. E. C., Granados, Ó., and Garcia-Bedoya, O. (2021). Predicción de la serie temporal del indicador bancario de referencia (ibr) con redes neuronales. *Revista Mutis*, 11(1).
- Nelson, C. R. and Siegel, A. F. (1987). Parsimonious modeling of yield curves. *Journal of business*, pages 473–489.
- Nisiba, M. (2011). A new approximate cva of interest rate swap in the sabr/libor market model: an asymptotic expansion approach. *Procedia Computer Science*, 4:1412–1421.
- Olah, C. (2015). Understanding lstm networks.
- Pose, M. G. (2009). Introducción a las redes de neuronas artificiales. *Departamento de Tecnologías de la Información y las Comunicaciones. Universidad da Coruña*.
- Pye, G. (1966). A markov model of the term structure. *The Quarterly Journal of Economics*, pages 60–72.
- Ramírez Amaya, F. (2019). Cálculo de cva/dva para swaps de tasa de interés ibr: una aplicación del modelo de tasa corta de hull-white de un factor. Master’s thesis, Universidad

del Rosario.

- Sabato, G. et al. (2010). Financial crisis: where did risk management fail? *International Review of Applied Financial Issues and Economics*, (2):315–327.
- Sierra Ramos, J. M. (2022). Introducción a las redes neuronales artificiales.
- Smith, D. J. (2012). A teaching note on pricing and valuing interest rate swaps using libor and ois discounting. *Boston University School of Management*, 32.
- Smith, D. J. and Ziemba, W. T. (2013). *Bond Math: The Theory Behind the Formulas*. Wiley, Hoboken, NJ.
- Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of financial economics*, 5(2):177–188.
- Vásquez, D. M. and Velandia, L. F. M. (2005). Estimación de la estructura a plazos de las tasas de interés en Colombia por medio del método de funciones b-spline cúbicas. *Revista de economía del Rosario*, 8(1):1–23.
- Xiao, T. (2019). Genetic algorithm based bidirectional recurrent neural network for libor prediction. In *2019 International Conference on contemporary Computing and Informatics (IC3I)*, pages 9–15. IEEE.
- Zeitsch, P. J. (2017). The economics of xva trading. *Journal of Mathematical Finance*, 7(2):239–274.