

Modelo para la predicción de la variable Tasa Representativa del Mercado colombiano basado en un análisis de Box Jenkins versus la técnica de Kalman Filters

Daniela Arciniegas Hernández

Alexander Castaño Arévalo

Trabajo de Grado para Optar al Título de Ingeniero Industrial

Director

Vlakxmir Robles Marín

Magister en Ingeniería

Codirector

Henry Lamos Díaz

Ph.D en Matemáticas y Física

Universidad Industrial de Santander

Facultad de Ingenierías Físico – Mecánicas

Escuela de Estudios Industriales y Empresariales

Bucaramanga

2022

Dedicatoria

Le dedico el resultado de todo el esfuerzo y empeño que le di a este proyecto a toda mi familia, sin la ayuda de ellos mi formación académica no hubiese sido posible, su cariño y comprensión siempre estuvieron presentes. En especial, a mi mamá, ya que ha sido y es mi pilar y apoyo durante todo este tiempo, me ha enseñado a ser la persona que soy hoy, mis principios, mis valores, mi perseverancia y a ser una mujer resiliente.

Daniela Arciniegas Hernández

Le dedico este logro a mis padres y mi tío, quienes a lo largo de todo mi proceso y formación universitaria han velado por mi bienestar y educación siendo mi apoyo incondicional, depositando su confianza en cada reto que se presentaba sin dudar de mí. Es por ellos que me he forjado como alguien perseverante y responsable.

Alexander Castaño Arevalo

Agradecimientos

Agradecemos a Dios por guiarnos y darnos la capacidad de superar cada reto, para poder llegar hasta este momento en el que finalizamos esta etapa tan importante en nuestras vidas.

A nuestra familia por su apoyo y de dedicación en nuestra formación, nuestros amigos por acompañarnos en este recorrido de aprendizaje y experiencias que llevaremos en nuestra memoria y corazón.

Nuestra admiración y agradecimiento especial a nuestro director de proyecto Vlaxmir Robles Marín, por su confianza, apoyo y paciencia, le agradecemos por guiarnos en el camino de ser excelentes profesionales desde el inicio de nuestra carrera (introductorio).

Finalmente, a nuestra alma mater la Universidad Industrial de Santander, le agradecemos por darnos la oportunidad de aprender y formarnos como personas y profesionales excepcionales.

Tabla de Contenido

Introducción 16

1. Generalidades de la Investigación 19

1.1. Objetivos 19

1.1.1. Objetivo general 19

1.1.2. Objetivos específicos 19

1.2. Metodología 20

1.2.1. Fase I. Realizar una revisión bibliográfica en general sobre el problema a tratar 21

1.2.2. Revisión de literatura sobre los métodos de solución 21

1.2.3. Fase II. Selección, análisis y extracción de datos 21

1.2.4. Fase III. Definir y desarrollar los algoritmos a usar 21

1.2.5. Fase IV. Evaluar los algoritmos 21

1.2.6. Fase V. Documentación 21

1.2.7. Fase VI. Evaluación de los modelos planteados 22

1.2.8. Fase VII. Conclusión y recomendaciones 22

2. Revisión de literatura 23

2.1. Análisis Bibliométrico 23

2.1.1. Scopus 24

2.1.2. Web of Science 32

2.1.3. Análisis preliminar de la literatura 39

3. Planteamiento del problema 43

3.1. Pregunta de investigación 45

4. Marco de antecedentes 46

4.1.	Marco teórico	48
4.1.1.	Métodos de predicción	48
4.1.2.	Tasa representativa del mercado colombiano (TRM).....	54
4.1.3.	Python	56
5.	Proceso de obtención de datos	59
5.1.	Selección de la fuente de información	59
5.2.	Selección y delimitación de datos.....	59
5.3.	Descarga de los datos.....	59
6.	Análisis exploratorio de datos.....	62
7.	Limpieza y procesamiento de datos.....	64
7.1.	Cargue de bibliotecas.....	64
7.2.	Cargue y depuración de información.....	65
7.3.	Gráfica de la información de interés en Jupyter Notebook.....	70
8.	Modelo Box Jenkins	71
8.1.	Paso 1: Identificación del modelo.....	73
8.2.	Paso 2: Estimación de parámetros iniciales	74
8.3.	Paso 3: Diagnóstico del modelo.....	76
8.4.	Paso 4: Predicción del modelo.....	78
9.	Modelo Kalman Filters	80
9.1.	Código del modelo del Filtro de Kalman.....	80
9.2.	Implementación del Filtro de Kalman para la estimación de la TRM.....	82
10.	Comparación de resultados	86
11.	Conclusiones	90

12.	Recomendaciones	91
	Referencias Bibliográficas	92
	Lista de anexos.....	96

Lista de Tablas

Tabla 1. Tabla de cumplimiento de objetivos.....	18
Tabla 2. Pasos para la implementación del modelo Box Jenkins	71
Tabla 3. RMSE para el Box Jenkins y Kalman Filters	86
Tabla 4. CFE, MAD y MAPE para el Box Jenkins y Kalman Filters	89

Lista de Figuras

Figura 1. Metodología general del estudio 20

Figura 2. Ecuación de búsqueda en Scopus 24

Figura 3. Desarrollo del tema con el pasar de los años, Adaptador de Scopus (2022)..... 25

Figura 4. Área temática, adaptada de Scopus (2022)..... 26

Figura 5. Autores del documento. Adaptado de Scopus (2022) 27

Figura 6. Países tema de investigación. Adaptado de Scopus (2022)..... 28

Figura 7. Concurrencia palabras clave en Scopus. Elaboración propia en VOSviewer (2022)... 30

Figura 8. Autores más importantes teniendo en cuentas las citaciones en Scopus. Elaboración propia en VOSviewer (2022). 31

Figura 9. Ecuación de búsqueda en Web of Science 32

Figura 10. Desarrollo del tema con el pasar de los años. Adaptado de Web of Science (2022) . 33

Figura 11. Área temática. Adaptado de Web of Science (2022). 34

Figura 12. Autores más relevantes en Web of Science..... 35

Figura 13. Países destacados en tema de investigación. Adaptado de Web of Science (2022)... 36

Figura 14. Concurrencia palabras clave en Web of Science. Elaboración propia en VOSviewer (2022). 37

Figura 15. Autores más importantes teniendo en cuenta las citaciones en Web of Science. Elaboración propia en VOSviewer (2022)..... 38

Figura 16. Ecuación del modelo Box-Jenkins 49

Figura 17. Ecuación del modelo Box Jenkins ARIMA (p, q)..... 50

Figura 18. Ecuación de regresión lineal múltiple o simple..... 50

Figura 19. Ecuación para el Cálculo de β_0 51

Figura 20. Ecuación para el Cálculo de β_1 51

Figura 21. Ecuación para el Cálculo de β_2 51

Figura 22. Ecuación de transición..... 52

Figura 23. Ecuación de medidas 52

Figura 24. Ecuación de propagación para $x_{t+1}|t$ 53

Figura 25. Ecuación de propagación para $P_{t+1}|t$ 53

Figura 26. Ecuación de actualización para K_t 53

Figura 27. Ecuación de actualización $X_t|t$ 54

Figura 28. Ecuación de actualización para $P_t|t$ 54

Figura 29. Banco de la República de Colombia TRM (2022) 60

Figura 30. Datos obtenidos del banco de la República de Colombia TRM (2022)..... 61

Figura 31. Comportamiento de la TRM trimestral 2013-2018..... 62

Figura 32. Bibliotecas cargadas en Python para la implementación de los modelos 64

Figura 33. Cargue de información en Python. Elaboración propia en Jupyter Notebook 66

Figura 34. Archivo descargado del Banco de la república con todo el Histórico de TRM (1991 – 2022)..... 67

Figura 35. Tasa representativa del mercado Periodo 2013-2018..... 68

Figura 36. Depuración de la información entre el 2013 y el 2018 para aplicación de modelos. Elaborado en Jupyter Notebook..... 69

Figura 37. Gráfica de la TRM entre el 2013-2018 en Jupyter Notebook. Elaboración propia.... 70

Figura 38. Autocorrelación y autocorrelación parcial del modelo 74

Figura 39. Entrenamiento y resumen de ARIMA para la serie 2013-2018. Elaboración propia en Jupyter Notebook 75

Figura 40. Cálculo del RMSE para el modelo ARIMA. Elaboración propia en Jupyter Notebook 77

Figura 46. 78

Figura 47. Instalación de tssltools_labs2 para la implementación del modelo de Kalman Filters 80

Figura 48. Código de implementación del Filtro de Kalman. Elaboración propia en Jupyter Notebook - Colab 81

Figura 49. Implementación del Filtro de Kalman para la estimación de la TRM 83

Figura 50. Gráfica de predicción de la TRM por el filtro de Kalman VS TRM Real. Elaboración propia Excel 84

Figura 51. Ecuación de CFE 87

Figura 52. Ecuación de MAD 88

Figura 53. Ecuación de MAPE 88

Lista de Apéndices

Los apéndices están adjuntos y puede visualizarlos en el Drive compartido

https://drive.google.com/drive/folders/1OP_us0Mn1C17r5OQI_WQ-h4FXZBIQGcZ?usp=sharing

Apéndice A. Artículo científico publicable

Apéndice B. Scripts con la programación en el lenguaje de Python

Apéndice C. Comparación de los modelos aplicados a la predicción de la TRM

Glosario

TRM: la tasa de cambio representativa del mercado (TRM) es la cantidad de pesos colombianos por un dólar de los Estados Unidos. La TRM se calcula con base en las operaciones de compra y venta de divisas entre intermediarios financieros que transan en el mercado cambiario colombiano, con cumplimiento el mismo día cuando se realiza la negociación de las divisas.

Actualmente la Superintendencia Financiera de Colombia es la que calcula y certifica diariamente la TRM con base en las operaciones registradas el día hábil inmediatamente anterior. (BANREP, 2021)

Fenómenos económicos: los fenómenos económicos son fenómenos sociales. Ambos conceptos se encuentran en la relación de la especie (fenómeno económico) al género (fenómeno social). El concepto de fenómeno social tiene mayor extensión, pero menor comprensión que el concepto de fenómeno económico. Los fenómenos económicos presentan las cualidades propias de los fenómenos sociales más aquellas que, por su naturaleza, les son peculiares.

Los fenómenos económicos son fenómenos sociales engendrados por la actividad que desarrollan los hombres dentro de la sociedad para satisfacer sus necesidades. (UNIVERSIDAD DE CHILE, 1937)

Kalman Filters (Filtro de Kalman): el filtro de Kalman es un algoritmo que estima el estado de un sistema a partir de datos medidos. Fue desarrollado originalmente por el ingeniero húngaro Rudolf Kalman, de quien toma su nombre. El algoritmo del filtro es un proceso de dos pasos: el

primer paso predice el estado del sistema, mientras que el segundo utiliza las mediciones de ruido para ajustar la estimación del estado del sistema.

En la actualidad existen diversas variantes del filtro de Kalman original. Estos filtros son muy utilizados para aplicaciones que dependen de la estimación, como la visión artificial, los sistemas de guía y navegación, la econometría y el procesamiento de señales. (MATHWORKS, 2020)

Metodología Box Jenkins: el acrónimo ARIMA significa media móvil integrado autorregresiva. Los retrasos de las series estacionarias en la ecuación de pronóstico se denominan términos “autorregresivos”, los retrasos de los errores de pronóstico se denominan términos de “promedio móvil” y se dice que una serie de tiempo que debe diferenciarse para hacerse estacionaria es “integrada” versión de una serie estacionaria. Los modelos de caminata aleatoria y de tendencia aleatoria, los modelos autorregresivos y los modelos de suavizado exponencial son casos especiales de los modelos ARIMA. (Gómez, 2020).

Resumen

Título: Modelo para la predicción de la variable Tasa Representativa del Mercado colombiano basado en un análisis de Box Jenkins versus la técnica de Kalman Filters*

Autor: Daniela Arciniegas Hernández y Alexander Castaño Arévalo**

Director: Vlaxmir Robles Marín

Codirector: Henry Lamos Díaz

Palabras Clave: Economía, pronóstico, TRM, Kalman Filters, Box Jenkins, modelo matemático.

Descripción: Las variables monetarias y cambiarias son una parte relevante en el desempeño real de la economía. Si bien es cierto que algunos de los fenómenos económicos no se pueden cuantificar o son muy complejos para su pleno entendimiento, la cuantificación de las realidades económicas ayuda en el estudio, la gestión y planificación del desarrollo de producción de bienes y servicios, además afectan directamente la calidad de vida de las personas, entre estas variables económicas se encuentra la Tasa Representativa del Mercado (TRM) la cual será el principal objeto de estudio del presente trabajo, se pondrá bajo análisis y medición de pronósticos económicos como la metodología de Box Jenkins y el modelo de Kalman Filters para acercarnos a determinar cuál es el modelo más apropiado y confiable para estimar dicha variable.

* Trabajo de grado

** Facultad de Ingenierías Físico-Mecánicas. Escuela de Estudios Industriales y Empresariales.

Director: Vlaxmir Robles Marín, Magister en Ingeniería. Codirector: Henry Lamos Díaz, Ph.D. en Física, Matemática.

Abstract

Title: Model for the prediction of the variable Representative Rate of the Colombian Market based on an analysis of Box Jenkins versus the Kalman Filters technique*

Author(s): Daniela Arciniegas Hernández and Alexander Castaño Arevalo**

Director: Vlakxmir Robles Marín

Co-director: Henry Lamos Díaz

Key Words: Economics, forecast, TRM, Kalman Filters, Box Jenkins, mathematical model.

Description: Monetary and exchange variables are a relevant part of the real performance of the economy. Although it is true that some of the economic phenomena cannot be quantified or are too complex for their full understanding, the quantification of economic realities helps in the study, management and planning of the development of production of goods and services, and also directly affects the quality of life of people, among these economic variables is the Market Representative Rate (TRM) which will be the main object of study of this work, it will be put under analysis and measurement of economic forecasts such as the methodology of Box Jenkins and the Kalman Filters model to get closer to determining which is the most appropriate and reliable model to estimate said variable.

* Bachelor's degree

** Faculty of Physical Mechanical Engineering. School of Industrial and Business Studies.

Director: Vlakxmir Robles Marín, Master's degree in engineering. Co-director: Henry Lamos Díaz, Ph.D. in Mathematical Physics.

Introducción

La incertidumbre, la previsión y la prevención, son factores que representan la necesidad del hombre de conocer todo a su alrededor, poder implementar estrategias a su conveniencia, y que mejor manera de controlar lo que vendrá, los escenarios que se presentarán y todo lo que pueda ocurrir, mediante el entendimiento de la naturaleza que lo rodea. En este sentido, el ámbito de la economía y las finanzas no son la excepción, y en particular las tasas de cambio son ejemplos palpables de esta necesidad de conocer y pronosticar su comportamiento dada su importancia en el desempeño económico.

Actualmente existen muchas teorías que intentan explicar los determinantes de los tipos de cambio, estos abarcan tanto a factores del ámbito financiero, componentes de expectativas y variables de la esfera real de la economía. La búsqueda por tratar de revelar las condicionantes del tipo de cambio no carece de sentido si se recuerda que se trata de una variable financiera clave en todo país. De ahí que existan múltiples modelos (monetarios, de portafolio, diferenciales de productividad) cuyo objeto de estudio se centre en la evolución del tipo de cambio nominal y real. La importancia de la predicción de la tasa de cambio resulta evidente. A modo de ejemplo, tenemos el cambio entre eurodólar (EUR/USD), ya que, estas dos divisas son parte fundamental de la economía internacional, por lo que sus fluctuaciones impactan en aquellos países que utilizan estas dos monedas como medio de intercambio para transacciones internacionales y como medio de reservas de divisas. Sin embargo, entender y predecir las fluctuaciones del tipo de cambio no es en lo absoluto una tarea sencilla.

Al respecto, el presente trabajo constituye un aporte adicional en este campo dentro de las investigaciones desarrolladas en el país, pues propone la utilización de métodos econométricos

para el tratamiento de series temporales, con el fin de identificar el patrón de comportamiento pasado, para la predicción de la variable tasa representativa del mercado colombiano basado en un análisis de Box Jenkins versus la técnica de Kalman Filters.

La finalidad de este trabajo es facilitar y difundir las herramientas para una mejor comprensión acerca del mercado de pronósticos en las variables económicas y las ventajas que se presentan al estar preparados a las posibles fluctuaciones. Mediante el uso de la información disponible al público nacional se pretende identificar una metodología apropiada para el mejoramiento de la confiabilidad de los pronósticos de la tasa de cambio, de tal manera que se pueda reducir la incertidumbre en la toma de decisiones de los diferentes actores económicos tales como el banco central, la empresas importadoras y exportadoras, bancos privados y comerciales, entre otros.

Para el desarrollo de esta investigación, se ha indagado en diferentes fuentes de carácter instrumental, bibliográfico, histórico, científico y académico, ya que no es un tema que se profundiza dentro de un plan de estudios de ingeniería industrial; por lo que nos hemos apoyado en campos como el de la economía y las finanzas. Este proyecto tiene como propósito informar a futuras investigaciones de modelos predictivos de variables económicas y la correcta aplicación de técnicas como la de Kalman Filters, especialmente en el sector económico empresarial.

Tabla 1.

Tabla de cumplimiento de objetivos.

Objetivos	Capítulos y subcapítulos
<p>Realizar una revisión bibliográfica sobre la aplicación de modelos relacionados a la técnica Box Jenkins, Kalman Filters en la predicción de las tasas de cambio, sobre el sistema monetario y cambiario colombiano.</p> <ul style="list-style-type: none"> • Construir la ecuación de búsqueda • Realizar el análisis bibliométrico de los resultados de la ecuación de búsqueda 	<p>Capítulo 2 Subcapítulo 2.1</p>
<p>Aplicar las técnicas de Box Jenkins y Kalman Filters a los datos históricos de la variable Tasa Representativa del Mercado (TRM) para un periodo de cinco años.</p>	<p>Capítulo 5,6 y 7</p>
<p>Validar los modelos mediante métricas de desempeño.</p>	<p>Capítulo 8,9 y 10</p>
<p>Elaborar un artículo de carácter publicable en base a la investigación realizada.</p>	<p>Apéndice A</p>

1. Generalidades de la Investigación

1.1. Objetivos

1.1.1. Objetivo general

Analizar la eficiencia en los pronósticos de series de tiempo de la variable económica Tasa Representativa del Mercado (TRM) Colombia mediante la técnica tradicional Box Jenkins versus la técnica de Kalman Filters.

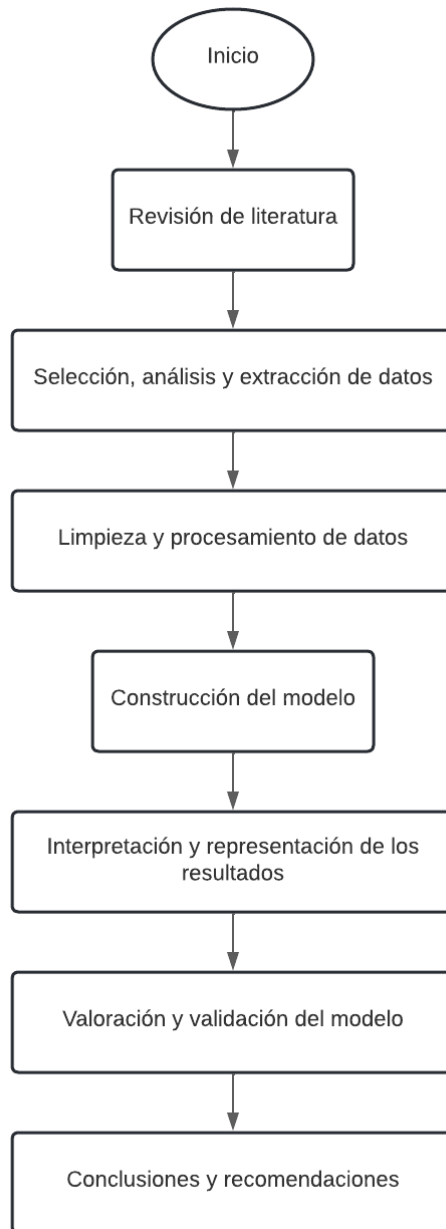
1.1.2. Objetivos específicos

- Realizar una revisión bibliográfica sobre la aplicación de modelos relacionados a la técnica Box Jenkins, Kalman Filters en la predicción de las tasas de cambio, sobre el sistema monetario y cambiario colombiano.
- Aplicar las técnicas de Box Jenkins y Kalman Filters a los datos históricos de la variable Tasa Representativa del Mercado (TRM) para un periodo de cinco años.
- Validar los modelos mediante métricas de desempeño. Elaborar un artículo de carácter publicable en base a la investigación realizada.
- Elaborar un artículo de carácter publicable en base a la investigación realizada.

1.2. Metodología

Figura 1.

Metodología general del estudio



Fuente: Elaboración propia

1.2.1. Fase I. Realizar una revisión bibliográfica en general sobre el problema a tratar

- Revisión de literatura
- Análisis bibliométrico, elaborar una ecuación de búsqueda en la base de datos SCOPUS y Web of Science.

1.2.2. Revisión de literatura sobre los métodos de solución

- Revisar e identificar como se ha abordado el tema de estudio en los últimos años
- Señalar la naturaleza de los datos de entrada de los modelos

1.2.3. Fase II. Selección, análisis y extracción de datos

- Determinar sobre cual conjunto de datos se va a trabajar
- Seleccionar el método para la extracción de datos

1.2.4. Fase III. Definir y desarrollar los algoritmos a usar

- Método Box Jenkins y Kalman Filters aplicado en la predicción de la tasa interna de retorno.

1.2.5. Fase IV. Evaluar los algoritmos

- Probar el algoritmo por medio del software de programación Phyton.
- Comparar los resultados generados en la investigación.

1.2.6. Fase V. Documentación

- Elaboración del documento de proyecto de grado
- Escritura del artículo de carácter publicable con la información destacada de la investigación y los resultados obtenidos.

1.2.7. Fase VI. Evaluación de los modelos planteados

- Verificar que cada modelo se adapta estructural y funcionalmente, esto implica corregir errores en el modelo si los hay.
- Validar el modelo seleccionado mediante un conjunto de métricas de desempeño.

1.2.8. Fase VII. Conclusión y recomendaciones

- Concluir a cerca del procedimiento, hallazgos y resultados obtenidos en la investigación.
- Realizar un grupo de recomendaciones para trabajos y estudios futuros.

2. Revisión de literatura

2.1. Análisis Bibliométrico

Se puede identificar el interés por el tema en especial en los años recientes debido al crecimiento exponencial de las predicciones económicas, esto tiene una estrecha relación con el surgimiento de nuevas tecnologías aptas para la comercialización, toma de decisiones de los productos o servicios en el mercado, de esta forma, se crea la necesidad de realizar investigaciones con el fin de obtener nuevas ideas para la adaptación y actualización de los cambios sociales que se van presentando.

Teniendo en cuenta el gran interés que se presenta por investigar este tipo de temas se evidencia un amplio volumen de información que se ha generado en su mayoría en los últimos años, por lo tanto ha resultado atractivo para la comunidad interesada en investigar los temas en cuestión, en consecuencia se crea la disciplina conocida como bibliometría la cual emplea conocimientos estadísticos y matemáticos con la finalidad de analizar el funcionamiento científico, por ello cuenta con mecanismos utilizados para evaluar los aspectos de este fenómeno llamados indicadores bibliométricos como documentos publicados por año, autores, países, área temática, palabras clave y citación de autores. de lo anterior se deriva la importancia que tiene la bibliometría en el campo de la investigación por consiguiente se han creado opciones en las bases de datos como SCOPUS y Web of Science que permiten extraer y analizar la información adquirida.

Para obtener un buen criterio de investigación de la predicción económica en el país, se utilizaron dos de las principales bases de datos, Scopus y Web of Science, en segundo lugar, se hizo un análisis tipo bola de nieve de citas en los artículos más relevante de la investigación.

Considerando que cada año se registra gran cantidad de artículos es necesario acudir al análisis bibliométrico de tal forma que se puedan obtener los documentos relevantes para el tema de investigación en estudio. Los indicadores básicos se tendrán en cuenta para mostrar los documentos publicados por año, autores, países, área temática, para los indicadores de relación como palabras clave y citación de autores se utilizará la herramienta VOSviewer que sirve para construir y visualizar redes bibliométricas como palabras claves y citación de autores entre otras.

A continuación, se mostrará la ecuación de búsqueda y los análisis que se realizaron en las bases de datos Scopus y Web of Science.

2.1.1. Scopus

Para realizar la búsqueda en Scopus se utilizó la siguiente ecuación:

Figura 1.

Ecuación de búsqueda en Scopus

finanzas AND (kalman and filter)AND (box and Jenkins)AND forecast (1)

Nota: Al momento de realizar la búsqueda en Scopus, se fueron agregando elementos que se acercan al fin de este documento.

Dicha consulta se realizó desde el día 10 de abril del año 2022, con el fin de limitar los documentos con enfoque en el proyecto propuesto.

2.1.1.1. Indicadores básicos

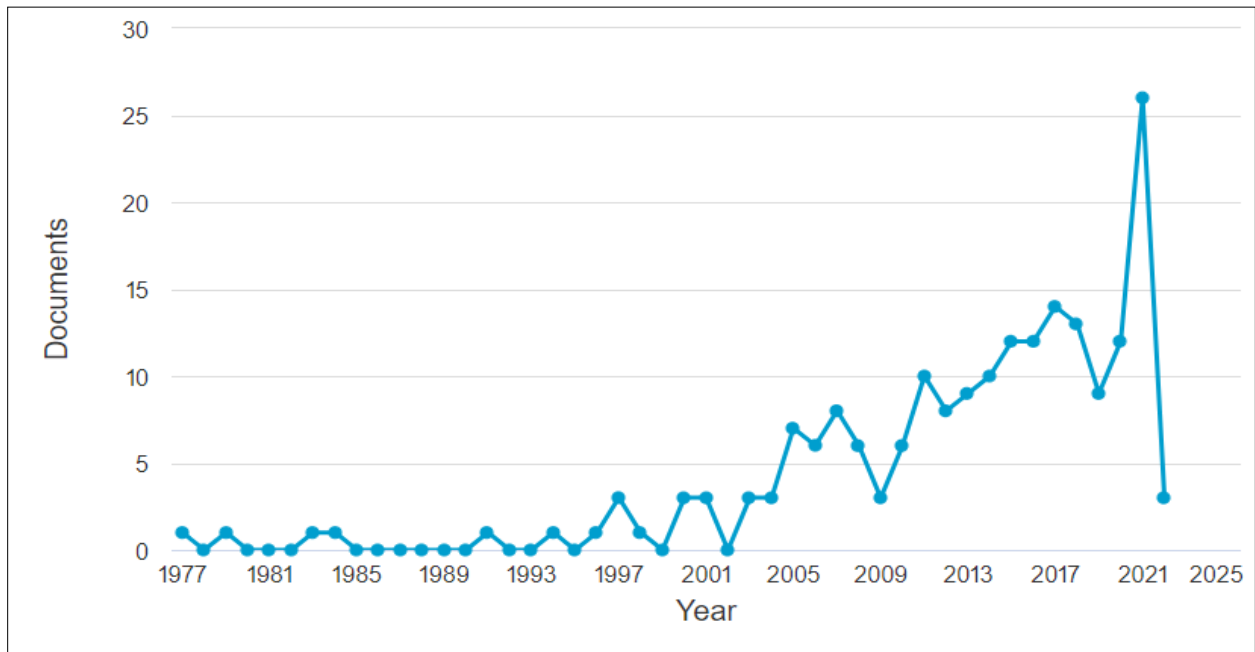
Al analizar los resultados de la búsqueda, se obtuvieron los siguientes gráficos estadísticos.

2.1.1.1.1. Documentos por año

En la figura 1 se logra evidenciar el interés de investigar sobre este tema con el pasar de los años por ejemplo del 2009 hasta la actualidad se ve un crecimiento positivo, en el año 2019 disminuye un poco, sin embargo, el año 2021 es el más llamativo debido a que se publicaron 26 documentos, en la actualidad del 2022 ya hay 3 documentos publicados.

Figura 2.

Desarrollo del tema con el pasar de los años, Adaptador de Scopus (2022)



Nota: Esta figura es extraída de Scopus, teniendo en cuenta la consulta realizada según el número de documentos tratados por año, con una serie histórica desde 1977 hasta el presente año.

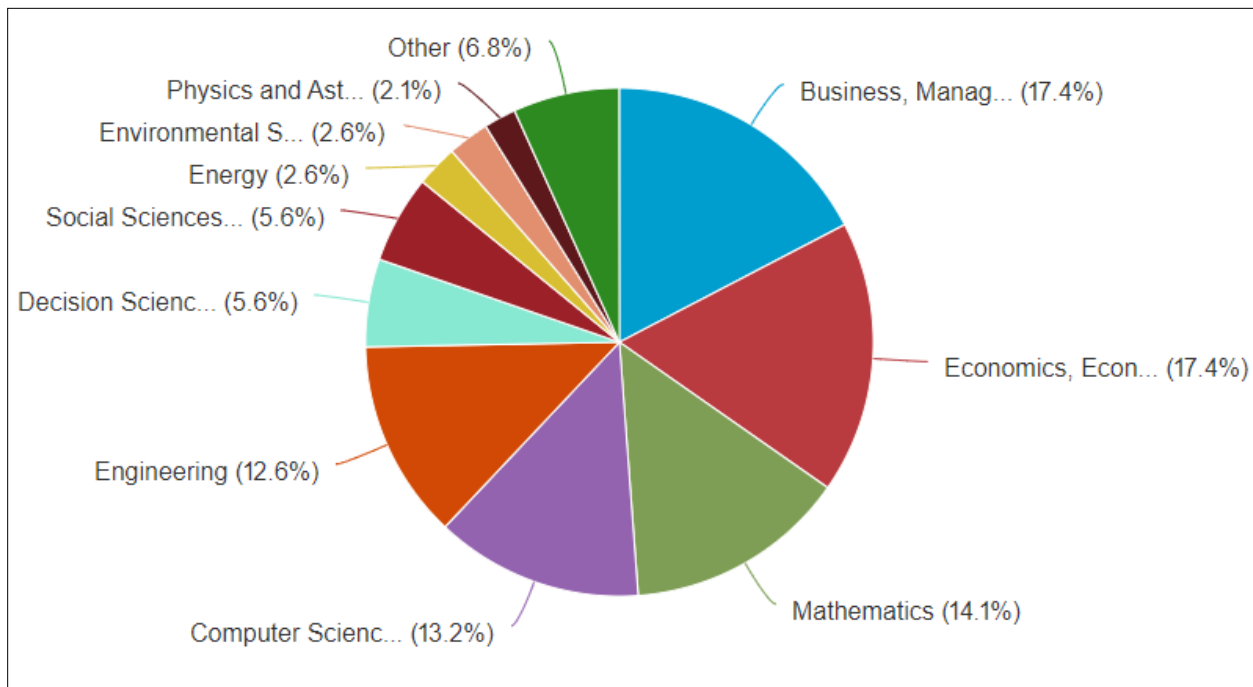
Fuente: Scopus, búsqueda realizada según la Ecuación 1. (Scopus, 2022)

2.1.1.1.2. Área temática

En la figura 2 las áreas de investigación más relevantes en esta temática teniendo en cuenta los documentos publicados son Negocios, gestión y contabilidad (17,4%); Economía, econometría y finanzas (17,4%); Matemáticas (14,1%), seguido de ciencias de computación (13,2%) e ingeniería (12,6%). Se puede evidenciar que estas son las 5 áreas que predominan con un 74,7% de los documentos publicados para el tema de interés en esta investigación, lo que sugiere que es un tema propio de nuestra disciplina de estudio ingeniería.

Figura 3.

Área temática, adaptada de Scopus (2022)



Nota: Esta figura es extraída de Scopus, teniendo en cuenta la consulta realizada según el porcentaje de documentos que existen de cada tema.

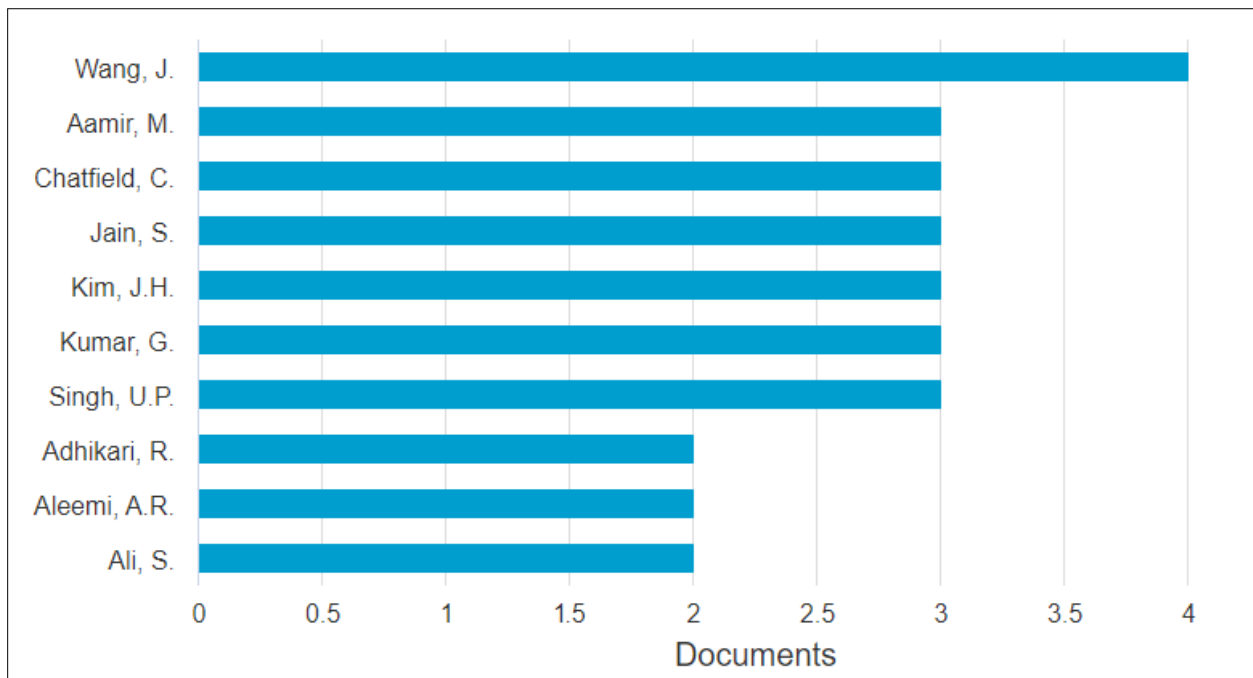
Fuente: Scopus, búsqueda realizada según la Ecuación 1. (Scopus, 2022)

2.1.1.1.3. Autores

La figura 3 muestra los 10 autores más relevantes para la investigación teniendo en cuenta el número de documentos publicados, cabe señalar que los autores con mayor número de publicaciones son Wang, J. teniendo un total de 4 documentos, luego Aamir, M., Chatfield, C., Jain, S., Kim, J. H. cada uno con 3 documentos, se refleja buen material de diferentes autores que pueden servir de referencia para la investigación.

Figura 4.

Autores del documento. Adaptado de Scopus (2022)



Nota: Esta figura es extraída de Scopus, teniendo en cuenta la consulta realizada según el número de documentos por autor.

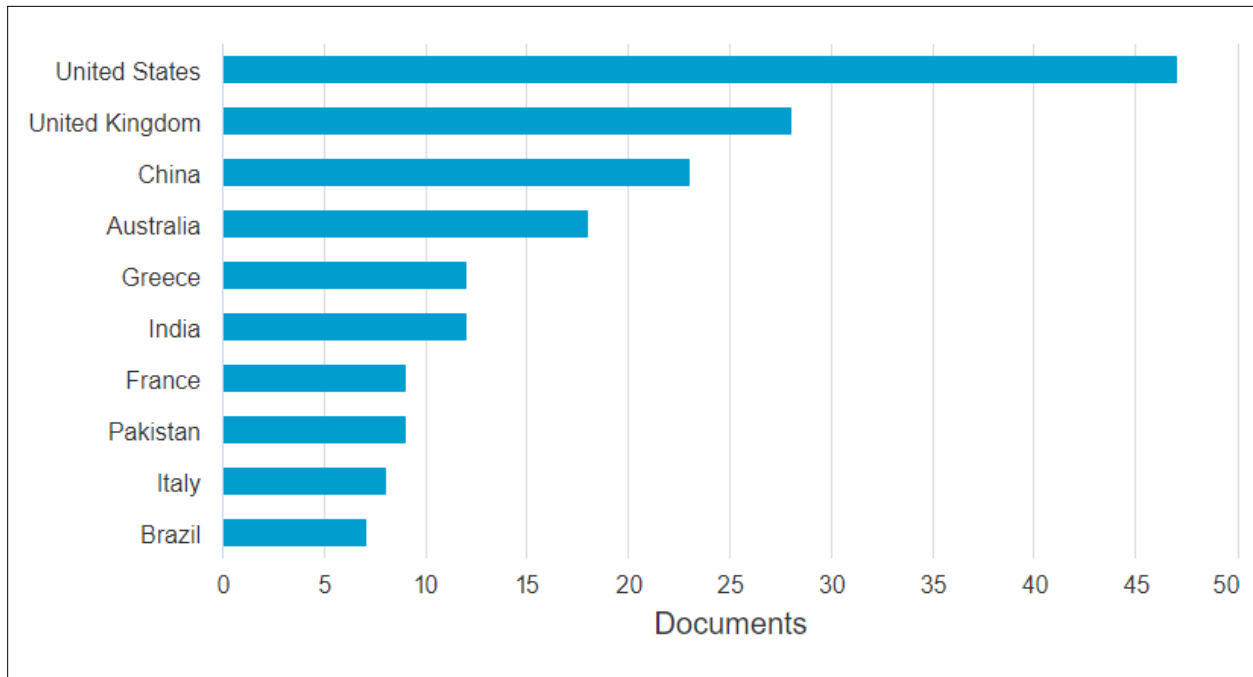
Fuente: Scopus, búsqueda realizada según la Ecuación 1. (Scopus, 2022)

2.1.1.1.4. Países

En la Figura 4 muestra los 10 países con mayor número de documentos publicados, se puede observar que EE. UU. está liderando con un total de 47 documentos relacionados con el tema de investigación, luego está Reino Unido con 28 documentos, después China con 23 documentos, también podemos observar que Brasil es el único país de Sudamérica que aparece en este top con un total de 7 documentos, con respecto a la cantidad de documentos publicados se refleja interés por parte de estos países en investigar estos temas.

Figura 5.

Países tema de investigación. Adaptado de Scopus (2022)



Nota: Esta figura es extraída de Scopus, teniendo en cuenta la consulta realizada según el número de documentos por país.

Fuente: Scopus, búsqueda realizada según la Ecuación 1. (Scopus, 2022)

Finalmente, en la base de datos Scopus a partir de la ecuación de búsqueda utilizada genera resultados favorables para el tema de investigación debido a la cantidad de documentos encontrados, además se observa el crecimiento año tras año de artículos publicados, lo cual demuestra interés en investigar estos temas, en consecuencia, la investigación toma valor puesto que es clara la importancia de continuar indagando respecto a dicho tema en cuestión.

2.1.1.2. Indicadores de relación

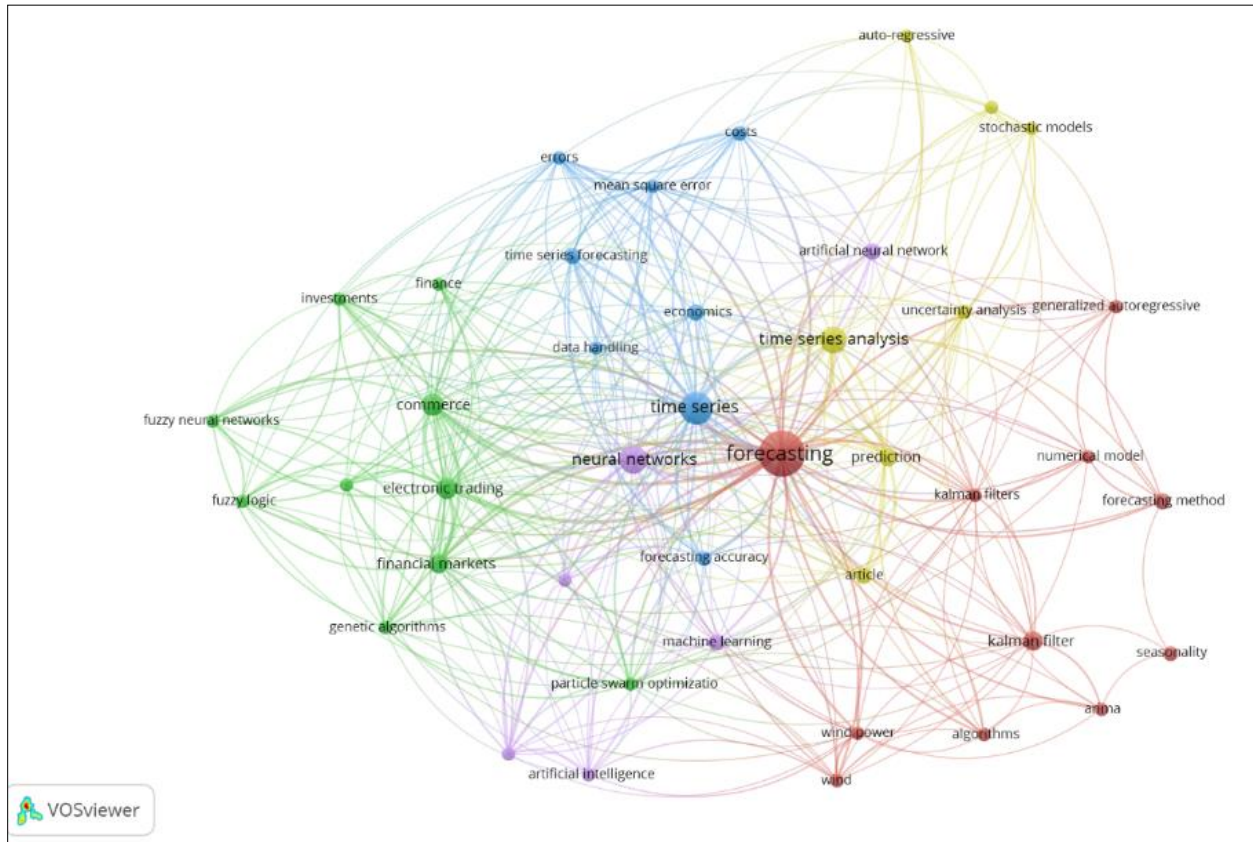
Para la realización de los indicadores de relación se utilizó el software de VOSviewer, el cual es una herramienta de software que sirve para construir y visualizar redes bibliométricas. VOSviewer también ofrece funciones de extracción de texto que se pueden utilizar en redes de concurrencia de términos importantes extraídos de un cuerpo de literatura científica. En este caso se tomó la información de Scopus en formato CSV para traerlo a VOSviewer.

2.1.1.2.1. Palabras claves

En la figura 5 se observa la concurrencia de palabras claves reunidas por colores en 5 clústeres los cuales representan algoritmos de agrupamiento por temáticas relevantes. En los colores que mejor se puede evidenciar es en el rojo, azul y morado en donde la palabra más usada es *forecasting*, seguida de *time series*, *neural networks* respectivamente. Además, es importante destacar que los círculos que tienen mayor tamaño representan más frecuencia de aparición de la palabra en los artículos.

Figura 6.

Concurrencia palabras clave en Scopus. Elaboración propia en VOSviewer (2022)



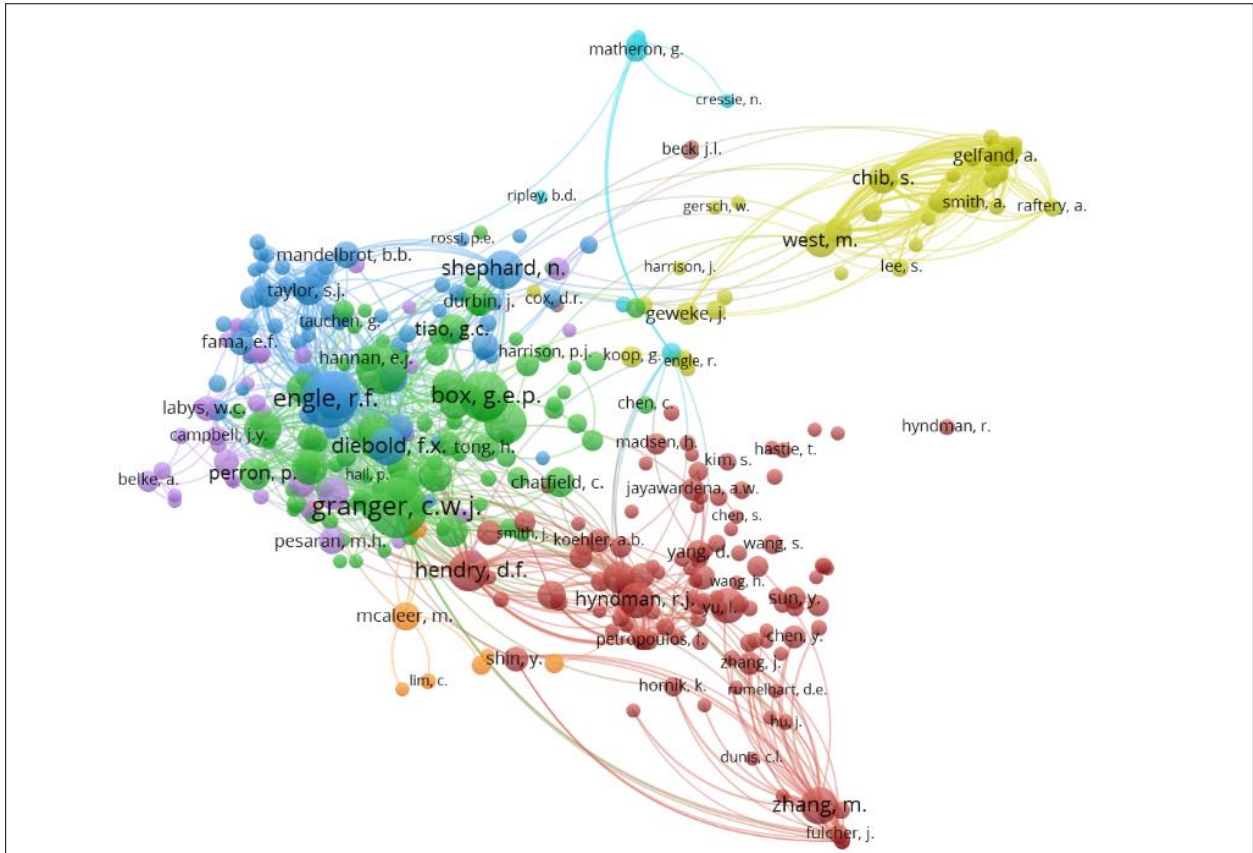
Nota: Esta figura es extraída de VOSviewer, con elaboración propia teniendo en cuenta las palabras clave ya mencionadas. (Visualizing scientific landscapes, 2022)

2.1.1.2.2. Citación de autores

En la figura 6 se observa 7 clústeres los cuales representan grupos constituidos por los autores más importantes, teniendo en cuenta las citaciones a lo largo del tiempo. En este caso se evidencia que los círculos con mayor tamaño es el Granjer, C.W.J; Engle, R.F. y Hendry D.F. lo que significa que estos autores son los que más citaciones tienen al pasar los años.

Figura 7.

Autores más importantes teniendo en cuenta las citas en Scopus. Elaboración propia en VOSviewer (2022).



Nota: Esta figura es extraída de VOSviewer, con elaboración propia teniendo en cuenta las palabras clave ya mencionadas. (Visualizing scientific landscapes, 2022)

2.1.2. *Web of Science*

Para realizar la búsqueda en web of science se utilizó la siguiente ecuación:

Figura 9. *Ecuación de búsqueda en Web of Science*

$$((kalman\ filter)AND\ (forecast)AND\ (finance)) \quad (2)$$

Teniendo en cuenta la ecuación evidenciada anteriormente, se obtuvieron 92 documentos al día 12 de marzo del 2022. Se limitó la información teniendo en cuenta los documentos con enfoque al proyecto propuesto.

2.1.2.1. **Indicadores básicos**

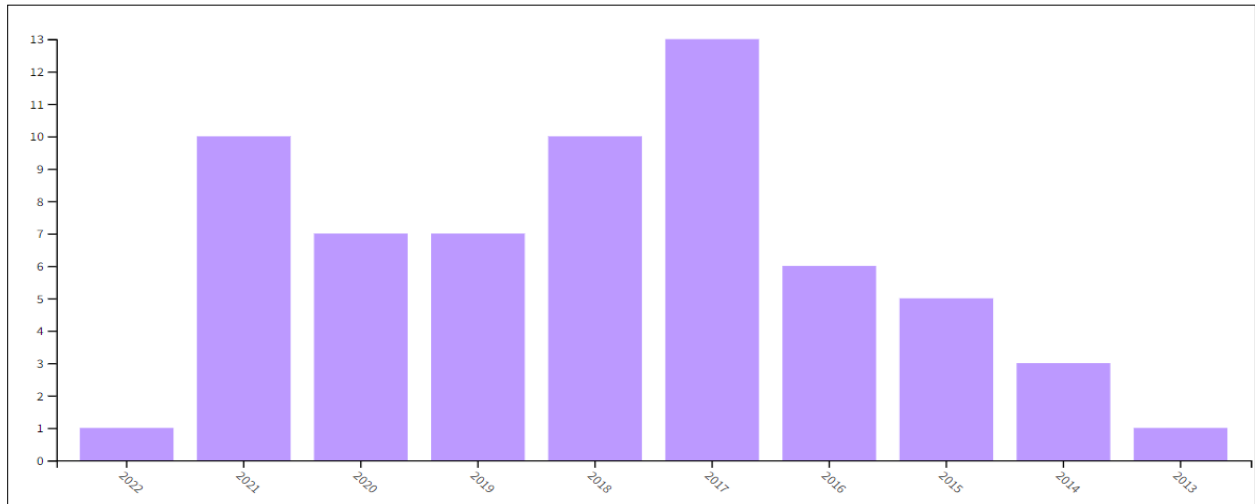
Se limitó la información teniendo en cuenta los documentos con enfoque al proyecto propuesto.

2.1.2.1.1. *Documentos por año*

En la figura 7 se puede observar el aumento gradual sobre el tema de la investigación con el pasar de los años por ejemplo a partir del año 2013 al 2017 se evidencia un crecimiento positivo, mientras que en los años siguientes disminuye un poco, pero logra mantenerse hasta el 2021, y en lo que va del año 2022 ya hay un documento publicado a cerca de este tema. Al igual que en la base de datos Scopus se logra evidenciar el crecimiento progresivo por parte de la comunidad académica por abordar este tipo de investigaciones.

Figura 8.

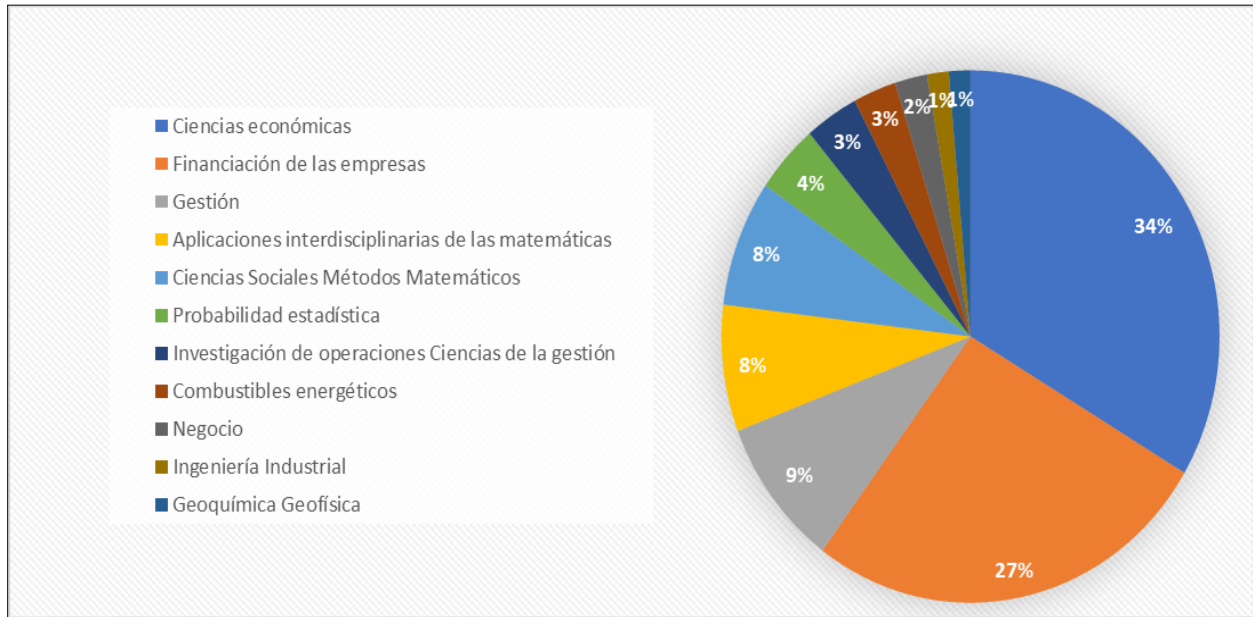
Desarrollo del tema con el pasar de los años. Adaptado de Web of Science (2022)

**2.1.2.1.2. Área temática**

En la figura 8 las áreas de investigación más relevantes en esta temática teniendo en cuenta los documentos publicados con mayor relevancia son: Ciencias económicas con un 34% de publicaciones, Financiación de las empresas 27%, y Gestión con 9%. Las anteriores áreas tienen el 70% de publicaciones respecto al tema en cuestión de la investigación, al igual que en la base de datos de Scopus se logra percibir la fuerza de investigación en economía.

Figura 9.

Área temática. Adaptado de Web of Science (2022).

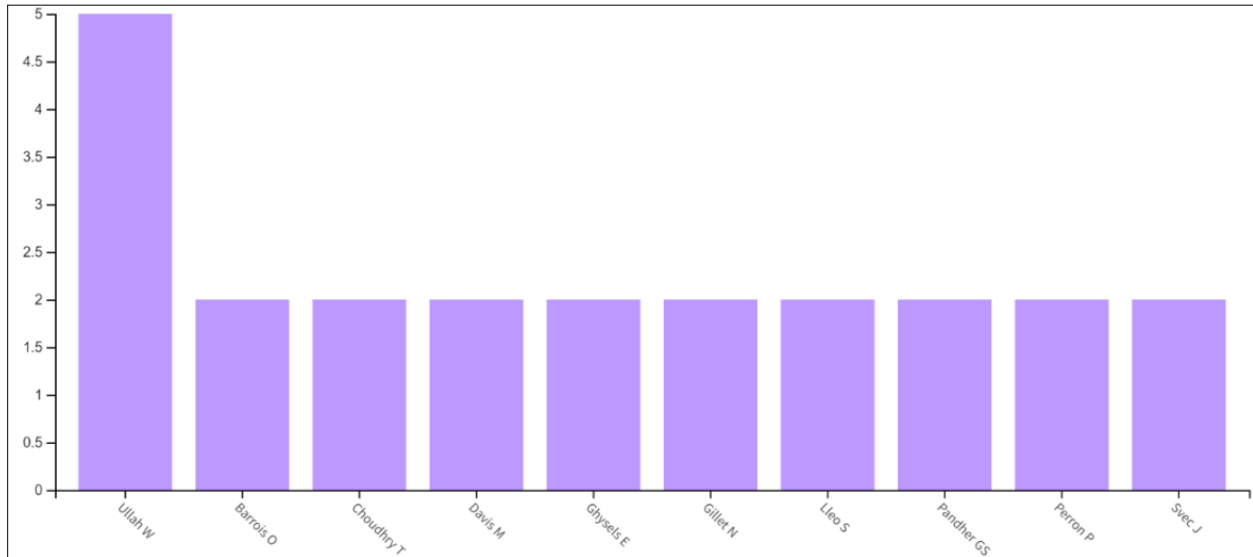


2.1.2.1.3. Autores

En la figura 9 muestra los 10 autores más relevantes para la investigación teniendo en cuenta el número de documentos publicados, cabe señalar que el autor con mayor número de publicaciones registradas en la base de datos Web Of Science es Ullah W. con 5 publicados, los demas autores como Barrois O., Choudhry T., Ghysel E., entre otros, tienen 2 documentos publicados en esta base de datos, se refleja buen material que pueden servir de referencia para la investigación ya que se encontraron autores diferentes a los de Scopus.

Figura 10.

Autores más relevantes en Web of Science



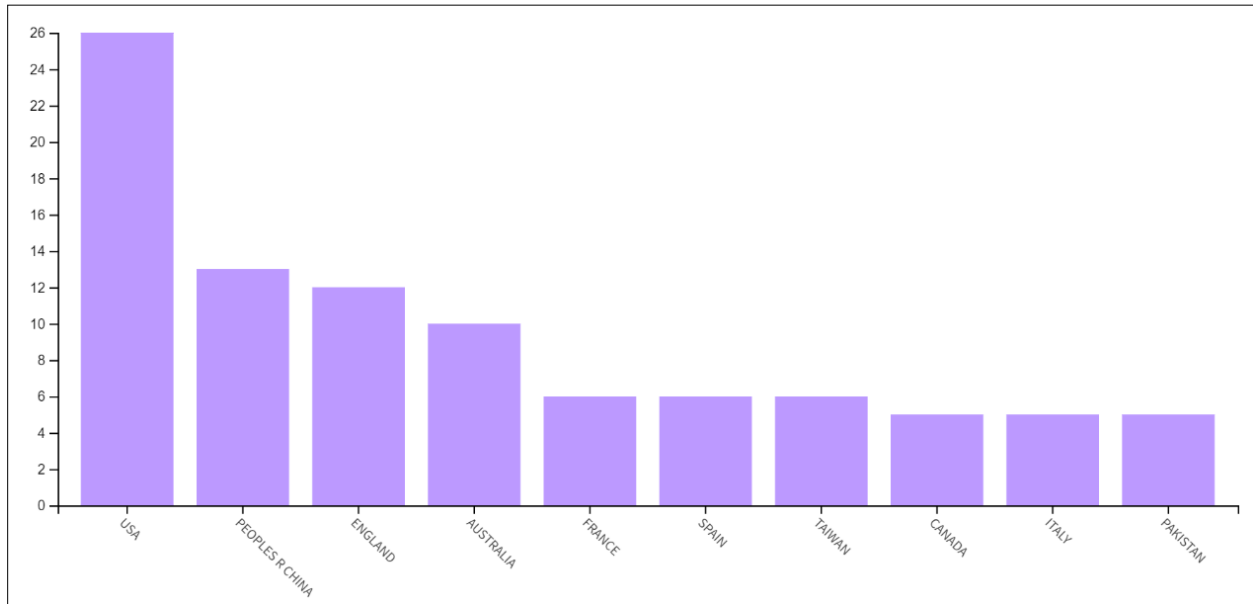
Nota: Teniendo en cuenta la búsqueda realizada en Web of Science esta gráfica representa a los autores más relevantes dentro del tema de investigación de interés de este proyecto.

2.1.2.1.4. Países

En la Figura 10 muestra los 10 países con mayor número de documentos publicados en la respectiva base de datos, se puede observar que Estados Unidos está liderando con un total de veintiséis documentos relacionados con el tema en cuestión de la investigación, luego está China con trece registros, después Inglaterra con doce publicaciones, también se puede evidenciar que no hay ningún país de Sudamérica, cabe resaltar que los países desarrollados tienen gran interés por investigar sobre el tema en estudio, por lo cual sugieren continuar indagando sobre la temática propuesta en la investigación.

Figura 11.

Países destacados en tema de investigación. Adaptado de Web of Science (2022)



Nota: Teniendo en cuenta la búsqueda realizada en Web of Science esta gráfica representa a los países más relevantes dentro del tema de investigación de interés de este proyecto.

2.1.2.2. Indicadores de relación

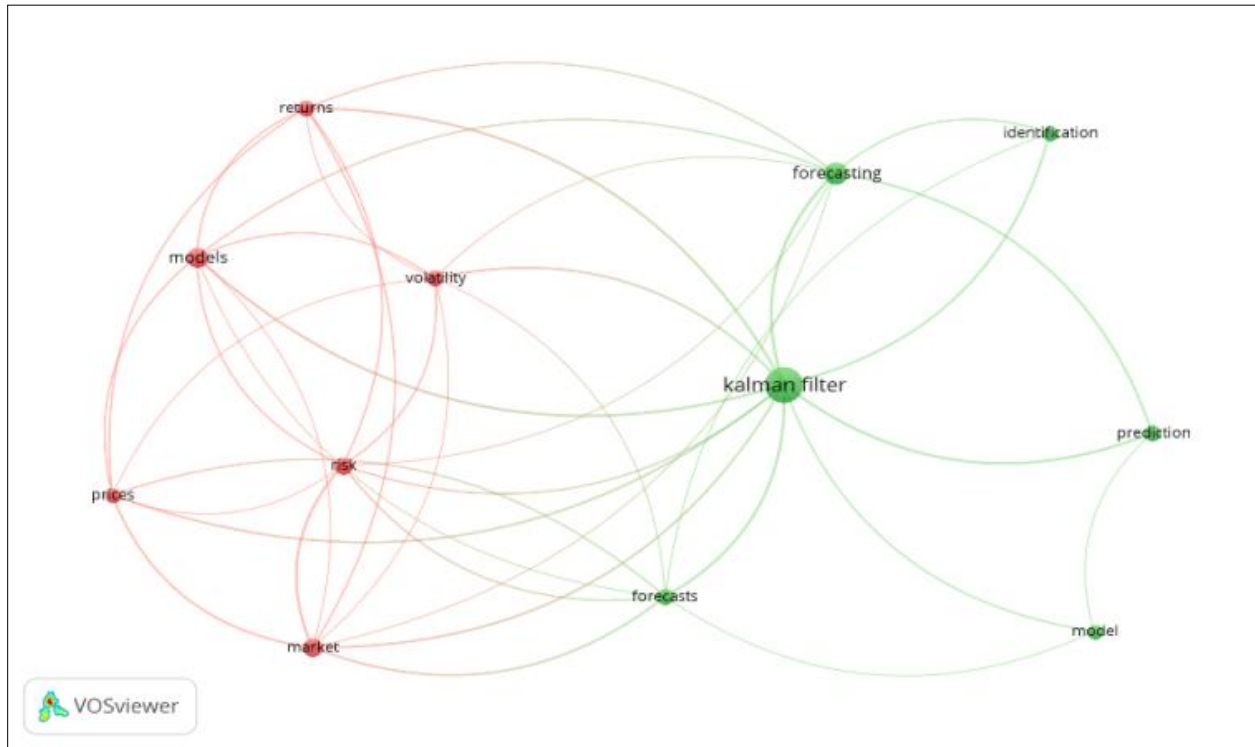
Al igual que con Scopus, para la realización de los indicadores de relación se utilizó el software de VOSviewer.

2.1.2.2.1. Palabras claves

En la figura 11 se evidencian dos clústeres, donde las palabras que más se repiten son Kalman Filters y forecast, teniendo en cuenta que las esferas de mayor tamaño es proporcional a la relevancia de las palabras claves.

Figura 12.

Concurrencia palabras clave en Web of Science. Elaboración propia en VOSviewer (2022).



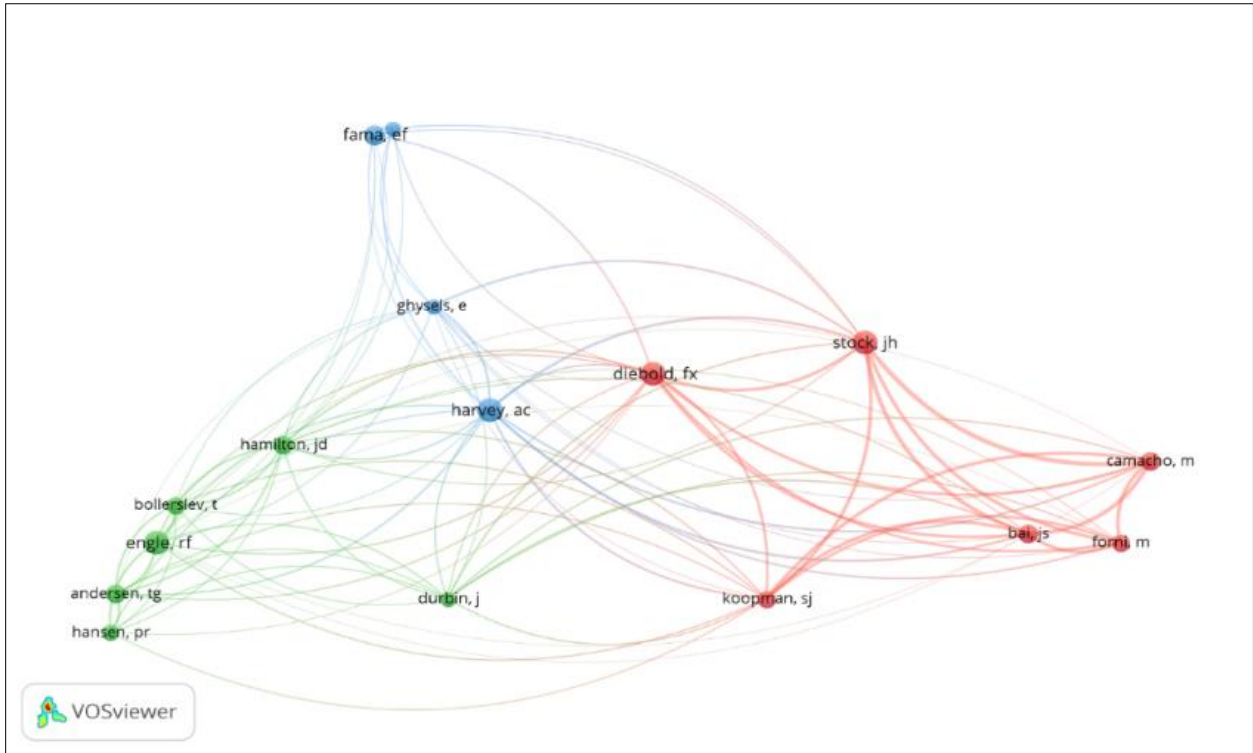
Nota: Esta figura es extraída de VOSviewer, con elaboración propia teniendo en cuenta las palabras clave ya mencionadas. (Visualizing scientific landscapes, 2022)

2.1.2.2.2. Citación de autores

En la figura 12 se observa la concurrencia de citas distribuidas por colores en 3 clústeres de algoritmos de agrupamiento que representa las citas de los autores más importantes a lo largo del tiempo. En este caso se evidencia que las esferas con mayor tamaño son las de Engle, R.F., Harvey, AC., y Diebold, FX., lo que significa que es el autor que más relevancia tiene con el pasar del tiempo.

Figura 13.

Autores más importantes teniendo en cuenta las citaciones en Web of Science. Elaboración propia en VOSviewer (2022)



Nota: Esta figura es extraída de VOSviewer, con elaboración propia teniendo en cuenta las palabras clave ya mencionadas. (Visualizing scientific landscapes, 2022)

2.1.3. Análisis preliminar de la literatura

El paradigma central de las finanzas consiste en asumir ciertos riesgos para obtener beneficios, pero no todos los riesgos aportan las mismas compensaciones. Tanto los riesgos como los beneficios pertenecen al futuro. Es así como se busca minimizar los riesgos de las pérdidas previstas al tiempo que se desea maximizar los rendimientos esperados.

La metodología Box-Jenkins para analizar series de tiempo con problemas de nivel y el análisis de variables latentes, presenta varias limitantes tales como eliminar observaciones en el proceso de diferenciación. Sin embargo, la metodología que emplea el modelo de estados resuelve algunos de estos inconvenientes. (Gómez, 2020)

La representación en el modelo de estados de la serie permite analizar el comportamiento de variables latentes y a su vez el Filtro Kalman permite analizar procesos en los cuales la función de máximo-verosimilitud no es conocida, además permite encontrar estimaciones para procesos que no poseen normalidad en los residuales.

Dentro de los autores más destacados encontramos a Engle y Granger ganadores del premio nobel de economía en 2003, Engle desarrolló un método que permite estimar la duración y magnitud de las fluctuaciones, lo cual es crucial para entender si las predicciones sobre los precios son confiables; las predicciones con alta volatilidad (varianza) son menos confiables que las predicciones con baja volatilidad. Un ejemplo de esto es que se encontró que, en las negociaciones salariales, a mayor aumento salarial, mayor incertidumbre asociada a la inflación futura. Mientras que Granger realizó aportes al tema de la causalidad. En términos filosóficos, la causalidad se define cuando tiene poder predictivo, pero está respaldada por la teoría. En otras palabras, no es suficiente que exista una correlación entre las variables, también es necesario explicar el

mecanismo por el cual el movimiento de una variable provoca el movimiento de otra variable. (H. Stock & M. Watson , 2020)

Estos dos autores cambiaron el rostro del pronóstico econométrico, ya que era la simple correlación entre variables, a veces por casualidad, pero ahora es la detección de una relación económica relevante entre variables, cuyo uso ha mejorado significativamente las técnicas de pronóstico. Todo esto gracias al estudio y desarrollo de "métodos de analizar las series temporales con volatilidad variante en el tiempo (ARCH)". (EcuRed, 2010)

Otro autor que contribuyó significativamente en la econometría es David F Hendry, en su artículo "Econometría: ¿Alquimia o ciencia?" en este menciona que la disputa metodológica viene desde tiempo atrás y no se puede reclamar un solo método válido, pero los resultados recientes sobre la teoría y la práctica de la selección del modelo fue justa para resolver muchos de los temas polémicos. Explica la evolución de un enfoque de modelado econométrico basado en análisis estadísticos cuidadosos de los datos disponibles, utilizando la teoría económica para guiar la especificación general del modelo. A partir de una base sólida en la teoría de la reducción, a través de una variedad de estudios aplicados y de simulación, demuestra que en general los procedimientos específicos tienen excelentes propiedades. Hendry, DF (1993). (H. Stock & M. Watson , 2020)

En una búsqueda más cercana nos encontramos con Aguirre Ariza & Castañeda Rueda (2019), en su trabajo de grado titulado *análisis de pronósticos de precios mediante series temporales aplicados al caso de la bolsa de valores colombiana*, a través del cual los autores establecen que la necesidad de estabilidad y seguridad no es ajena a las bolsas de valores, es por eso que a través del trabajo pretenden realizar el análisis a través de un modelo de pronóstico la

efectividad de la predicción de los precios de las acciones para la toma de decisiones de los inversores en el mercado accionario colombiano, mediante el uso de herramientas como series temporales y el algoritmo del filtro de Kalman. (Luis Edwin & Danny, 2010)

Comenzando la selección de la variable económica, se recopiló y se estructuró el historial necesario para realizar el pronóstico a partir del uso de varias herramientas como lo son el análisis fundamental, el análisis técnico y entra ellas el filtro de Kalman. Al final, se obtiene como resultado un mayor acierto en la predicción a comparación de diversos métodos más sencillos y populares entre los inversionistas no especializados en la predicción, sin embargo, el uso de este método de análisis demanda mayor dedicación, observación y tiempo para obtener mejores resultados.

Los autores concluyen entonces que, en las situaciones actuales de economías tan abiertas y entrelazadas, los eventos diversos que pudieran afectar un sector económico en específico no son los únicos a tomar en cuenta y los que influyen en el estado de estas economías no son tan evidentes y predecibles como se deseara; es por esto que para modelar de manera correcta una variable económica resulta necesario tomar una perspectiva más probabilística.

Por otro lado, tenemos que los autores Vargas Vigoya & Mier Santander (2020), en su trabajo titulado *modelo de integración entre expectativas de empresarios y método estadístico para pronosticar el crecimiento de las ventas del sector industrial en Colombia, una aplicación a la planeación de las empresas*, establecen que el propósito de esta investigación es evaluar la metodología Box Jenkins y Filtro de Kalman para pronosticar el Crecimiento de las Ventas del Sector Industrial en Colombia, utilizando la serie de tiempo y las expectativas de empresarios, con el fin de presentar a las empresas información precisa del comportamiento futuro de su entorno de

tal forma que sea útil para implementar dentro de su planeación estratégica. (Vargas Vigoya & Mier Santander, 2012)

Los autores concluyeron entonces que el modelo estado-espacio estimado a través del Filtro de Kalman es una metodología que permite pronosticar una variable económica no observada en función de variables explicativas observadas que influyen en la predicción. Esta metodología es aplicada al caso colombiano con el fin de pronosticar el crecimiento de las ventas del sector industrial mediante la integración de la información contenida en las expectativas de empresarios presentadas por el Banco de la República a través de la EMEE y la serie de tiempo mensual de indicadores presentada por el DANE a través de la MMM, demostrando que las expectativas contienen información útil para mejorar la precisión del pronóstico para el periodo de Octubre de 2006 a Enero de 2012. Comparando con el pronóstico ingenuo la integración de las expectativas y la serie de tiempo mejora el pronóstico en el 52%, mientras que una metodología Box-Jenkins que utiliza únicamente la serie de tiempo mejora el pronóstico en el 34% sobre la base de prueba. Este resultado permite concluir que el juicio de los empresarios logra capturar anticipadamente variables exógenas que afectan el comportamiento futuro de la serie en cada momento del tiempo. (Dorado & Ruiz, 2013)

Aunque la investigación es concluyente en afirmar que las expectativas de los empresarios contienen información útil que ayuda a mejorar la precisión de los pronósticos, estas no son acertadas en presencia de puntos de inflexión o momentos de alta volatilidad, en los que se percibe que el Filtro de Kalman al integrar las expectativas con la serie de tiempo no logra anticipar los cambios bruscos de tendencia.

3. Planteamiento del problema

Tradicionalmente, el funcionamiento de las tasas de cambio en una economía ha sido una temática que ha generado interés en los economistas académicos y autoridades monetarias de diferentes naciones a lo largo de la historia. Esta preocupación se deriva de las repercusiones que el desempeño de la tasa de cambio puede tener sobre la economía real, en términos de desempleo, producción, etc. y en general sobre la calidad de vida de los habitantes de una sociedad. En este sentido encontrar las herramientas o los métodos ideales que permitan realizar pronósticos confiables de sus fluctuaciones se ha convertido en un desafío no solo desde el punto de vista académico sino también para los responsables de la política económica.

Desde el punto de vista empírico, son múltiples los modelos y técnicas utilizadas con fines predictivos, donde resalta la diversidad de herramientas econométricas empleadas para tal fin. Sin embargo, desde décadas atrás en autores como Meese y Rogoff (1983), que destacaban el pobre desempeño de una variedad de modelos de tipo de cambio estructural como el modelo monetario o el modelo de portafolio. Específicamente, estos autores revelan que, a partir del período de tasas flotantes posterior a Bretton Woods, las predicciones fuera de la muestra de las tasas de cambio son superadas, especialmente en el corto plazo, por un simple modelo de paseo aleatorio, el cual no utiliza información basada en las variables fundamentales de la economía. Debido a esto se desarrolló una extensa literatura durante el período post-Bretton Woods que muestra resultados sólidos a partir del uso de técnicas econométricas no lineales, diferentes divisas, frecuencia de los datos y muestras. (Cheung, Chinn, & Garcia Pascual, 2005)

En Colombia la tasa de cambio más importante es la relación peso/dólar y se mide a través de la Tasa representativa de mercado (TRM), monitoreada por el Banco de la República, bajo libre

fluctuación de las fuerzas del mercado. Las variaciones de la divisa colombiana generan preocupación en los efectos que puede tener sobre la actividad del sector empresarial y los consumidores, impactando el funcionamiento económico mediante su efecto directo sobre las importaciones y exportaciones, ya que si se presentan alzas en este indicador los precios de los productos importados también lo harán, un claro ejemplo en relación a los empresarios es que para los que exportan, una mayor tasa de cambio representa una ganancia superior a la acostumbrada; por otra parte en el caso de los que importan, cuando la tasa de cambio aumenta las ganancias disminuyen. Otros que se benefician con el alza del dólar son las compañías y empresarios que tienen activos con la moneda extranjera.

En la actualidad, factores como el precio del dólar, la emergencia sanitaria, la oferta y demanda agregadas, la baja del petróleo, y los cambios imprevistos en la bolsa de valores, dan paso a un decrecimiento de la economía, el alza de este indicador altera el desarrollo de todo el mercado colombiano. Teniendo en cuenta que los constantes cambios dentro de la economía mundial y más específicamente dentro de la economía nacional, hacen de las inversiones dentro del mercado bursátil un negocio cada vez más riesgoso; un pequeño cambio en alguna de las variables que influyen en las negociaciones. Debido a esto, la importancia de una correcta predicción de las variables económicas (TRM) dentro de este mercado es el motor que impulsa la investigación de métodos más acertados y ajustados a la realidad.

La técnica Box Jenkins ha sido ampliamente utilizada durante las últimas décadas para la predicción de variables económicas con resultados relativamente aceptables; por otra parte, la metodología de Kalman Filters se han implementado en diferentes campos del conocimiento científico tales como la orientación, navegación y control de vehículos y cohetes espaciales,

pronósticos del clima, etc. Con resultados prometedores, pero hasta ahora es muy incipiente en el mundo económico y financiero. Razón por la que estudiar el comportamiento de las variables económicas, partiendo de la variable tasa representativa del mercado, basado en un análisis de Box Jenkins vs. la técnica de Kalman Filters es fundamental para determinar la eficiencia de los pronósticos de series de tiempo de la variable TRM, contribuyendo así con el crecimiento del mercado bursátil colombiano.

3.1. Pregunta de investigación

De este análisis del problema entonces nos planteamos la pregunta a resolver con la presente investigación. ¿Es más eficiente y confiable la metodología Kalman Filters en comparación al método Box Jenkins para el pronóstico de la variable económica Tasa Representativa del Mercado Colombiano (TRM)?

4. Marco de antecedentes

Desde la publicación por parte del señor Rudolph E. Kalman de su algoritmo, se han llevado a cabo la aplicación de este filtro para diversos campos de la ciencia; sistemas de navegación, aeronáutica, seguimiento de objetos, oceanografía, hidrología, sistemas de control, sistemas de automatización y entre otras, la más concerniente a esta investigación, la econometría.

A escala un poco global el uso del filtro para la mejora de variables en la econometría se ha extendido al estudio de elementos como proyecciones de tasas de interés, caso de *The application of the Kalman filter to the Fisher equation: Italian and German term structure of interest rates* (Panseri, Urga, & Cristini, 1999); o ya un poco más general, la tesis doctoral *Aplicaciones econométricas del filtro de Kalman y algunas variaciones numéricas: el filtro de Chandrasekhar*. (Sotoca López, 1992)

En Colombia han incursionado en temas como la economía subterránea (Arango A, Misas, & López, 2011), En el año 2011 en la Universidad de Medellín se publicó un artículo “Estimación de la relación entre el nivel de precios y la tasa de cambio para Colombia (1991-2006), mediante el filtro de kalman” (Chavarro Miranda Fernando; Grautoff Laverde Manfred; Mancipe Moncada, Erika Viviana). La predicción de betas y VaR de portafolios de acciones (Giraldo Gomez, 2005), la estimación de la tasa de interés natural en Colombia (Echavarría Soto, López Enciso, Misas Arango, Téllez Corredor, & Parra Alvarez, 2018) la construcción de un modelo para las series del PIB trimestral (Galvis, 1991), por mencionar algunos.

En un marco más cercano, un estudio realizado en la Universidad Industrial de Santander, hace referencia a la utilización de las cadenas de Márkov para el análisis de pronósticos y series de tiempo, en donde el resultado obtenido fue que el modelo de Márkov no se considera un método

confiable para realizar pronósticos a largo plazo- Acevedo Beltrán, Serrano Acevedo, M. E., Diaz Bohórquez, C. E., & Universidad Industrial de Santander. Escuela de Estudios Industriales y Empresariales. Tesis. (2011). Aplicación de cadenas de Markov para el análisis y pronóstico de series de tiempo [recurso electrónico]. UIS.

Siguiendo esta línea de proyectos de investigación de la universidad Industrial de Santander, en este se concluye que ningún modelo resultó significativamente valido para la predicción de precio de las acciones usando modelos de predicción logística- Mendez Pineda, Talero Sarmiento, L. H., & Lamos Díaz, H. (2021). Un modelo para la predicción del movimiento del precio de las acciones del mercado bursátil basado en un análisis de sentimiento y datos históricos de la BVC [recurso electrónico]. UIS.

4.1. Marco teórico

4.1.1. Métodos de predicción

Los métodos de predicción pueden ser clasificados en dos categorías: métodos cualitativos y métodos cuantitativos (Uriel & Muñiz, 1998)

El primero es un enfoque cualitativo, utilizado para hacer predicciones en contextos menos conocidos, como la innovación tecnológica, social, política, etc. Estos modelos se basan en las opiniones de expertos que, en base a su conocimiento y experiencia, emiten juicios sobre las preguntas planteadas para realizar predicciones, y generalmente predicciones a largo plazo. Algunos métodos cualitativos incluyen: pronósticos prospectivos, analogías históricas, panel de consenso y el método Delphi.

Por otro lado, los métodos cuantitativos se utilizan para predecir datos futuros basados en datos pasados. Estos métodos son generalmente adecuados para el corto o mediano plazo cuando se dispone de datos numéricos anteriores y es razonable suponer que se espera que ciertos patrones en los datos persistan en el futuro. Hay dos grandes categorías de métodos cuantitativos: métodos de series temporales univariados (incluidos los métodos de descomposición y los modelos ARIMA) y métodos causales

4.1.1.1. Series de tiempo y modelos causales

Una serie de tiempo es una secuencia de datos numéricos, cada uno asociado con un punto específico en el tiempo. Podemos dar ejemplos de series temporales como la tasa de inflación mensual, el tipo de cambio diario, el PIB trimestral, la tasa de desempleo anual, etc. Una característica de estas series es que el paso del tiempo es uniforme para cada una de ellas, es decir,

si se utiliza una serie con datos semanales, se debe mantener el mismo tipo de datos (con la misma frecuencia) en toda la serie. (Casas Tragadora, 2002)

El análisis de series de datos se denomina análisis de series de tiempo univariadas. Si se está estudiando un grupo de series de la misma serie temporal, el análisis se denomina análisis de series temporales múltiples. El propósito de este análisis es estudiar la estructura dinámica o temporal de la información. A través del estudio y comprensión del comportamiento de las series económicas, se pueden realizar predicciones de extrapolación utilizando la información disponible a la fecha. (Casas Tragadora, 2002)

4.1.1.2. Metodología Box-Jenkins

Es un método cuantitativo que utiliza series de tiempo univariados, conocido como modelos ARIMA o modelo de Box-Jenkins. Se basa en el hecho de que una serie temporal que se busca pronosticar es generada mediante un proceso aleatorio, cuya naturaleza puede representarse mediante modelo. El modelo ARIMA univariados busca predecir valores futuros de series temporales a partir de datos pasados de la serie y errores de pronósticos anteriores. La notación compacta para los modelos ARIMA es la siguiente:

Figura 16.

Ecuación del modelo Box-Jenkins

$$ARIMA(p, d, q) \quad (3)$$

Donde p es el número de parámetros autorregresivos, d es el número de diferenciaciones para la serie estacionaria, y q es el número de parámetros de las medias móviles. El modelo Box Jenkins ARIMA (p, q) viene representado por la siguiente ecuación:

Figura 17.

Ecuación del modelo Box Jenkins ARIMA (p, q)

$$Y_t = \phi_0 + \phi_1 y_{t-1} + L + \phi_p y_{t-p} + \alpha_t - \theta_1 a_{t-1} - L - \theta_q a_{t-q} \quad (4)$$

La parte autorregresiva (AR) del modelo es: $\phi_1 y_{t-1} + L + \phi_p y_{t-p}$, mientras que la parte de medias móviles del modelo (MA) es: $-\theta_1 a_{t-1} - L - \theta_q a_{t-q}$. Los coeficientes de los parámetros $\phi_0, \phi_1, L, \phi_p, \theta_1, \theta_q, \theta_1, \theta_q$ son determinados a partir de los datos, a través de cualquier método estadístico consistente. El método de Box-Jenkins hace predicciones sin condiciones previas y, una vez que se encuentra el modelo, está listo para hacer predicciones y comparar datos reales y estimados para el modelo dado. Pero además de requerir un número significativo de observaciones, estimar e interpretar sus coeficientes es difícil y, a menudo, da como resultado resultados deficientes para el pronóstico a largo plazo.

El modelo causal se basa en el supuesto de que la variable pronosticada (dependiente) depende de uno o más factores (la variable independiente), de modo que cuando estos cambien, causaran variaciones en la primera, de ahí el nombre entre estos modelos. Para relacionar estas variables se utilizan métodos como la regresión lineal múltiple o la regresión lineal simple.

Figura 18.

Ecuación de regresión lineal múltiple o simple

$$y = \beta_0 + \beta_1 x + \beta_2 z \quad (5)$$

Donde y es la variable pronosticada, X y Z son variables independientes y b_0, b_1 y b_2 son coeficientes de ajuste. Así mismo para obtener estos coeficientes, las fórmulas son las siguientes:

Figura 19.

Ecuación para el Cálculo de β_0

$$\beta_0 = \frac{D_0}{D} \quad (6)$$

Figura 20.

Ecuación para el Cálculo de β_1

$$\beta_1 = \frac{D_1}{D} \quad (7)$$

Figura 21.

Ecuación para el Cálculo de β_2

$$\beta_2 = \frac{D_2}{D} \quad (8)$$

Nota: Las ecuaciones 6,7 y 8, representan las ecuaciones para el cálculo de β_0, β_1 y β_2 donde los D_0, D_1, D_2 y D son determinantes de tercer orden.

4.1.1.3. Modelo Kalman Filters

El Filtro de Kalman es una herramienta desarrollada por Rudolf E. Kalman en 1960 empleada para controlar los sistemas ruidosos, específicamente para identificar el problema lineal cuadrático, cuando este es sometido a ruido blanco aditivo. Según Andrews (2001), el estimador resultante es estadísticamente óptimo con respecto a cualquier función cuadrática de estimación del error.

Consiste en un estimador recursivo basado en la teoría de sistemas dinámicos, cuyo propósito es optimizar el vector de estado en un momento determinado a través de un algoritmo basado en la información disponible por las observaciones realizadas hasta ese momento. A diferencia de otros métodos de estimación, permite la representación de procesos no estacionarios

sin requerir información a priori sobre la historia del proceso o la observación de la variable de estado en estudio, que puede no ser directa.

La estructura de un modelo en forma de espacio de estados viene determinada por el conjunto de dos ecuaciones, denominadas como ecuación de transición y ecuación de medidas.

4.1.1.3.1. Primera ecuación: Ecuación de transición

Figura 22.

Ecuación de transición

$$x_{t+1} = F_t x_t + G_t u_t + S_t w_t \quad (9)$$

Donde: F_t es la matriz de transición de estados de dimensión $(n \times n)$, G_t es la matriz de control o de input de dimensión $(n \times p)$, S_t es la matriz de ponderación de la perturbación de dimensión $(n \times s)$, x_t es el vector de estado de dimensión $(n \times 1)$, u_t es el vector de control de dimensión $(p \times 1)$, w_t es el vector de la perturbación del sistema de dimensión $(s \times 1)$

4.1.1.3.2. Segunda ecuación: Ecuación de medidas

Puesto que el vector de estado no puede ser observado directamente, se le va a relacionar con una sucesión de observaciones perturbadas por un ruido mediante la siguiente expresión, que recibe el nombre de ecuación del output o de las observaciones:

Figura 23.

Ecuación de medidas

$$Y_t = H_t x_t + y_t \quad (10)$$

Donde: y_t es el vector de medidas u observaciones de dimensión $(m \times 1)$, H_t es la matriz de medidas u observaciones de dimensión $(m \times n)$, v_t es el vector de error de medidas de dimensión $(m \times 1)$.

4.1.1.3.3. Tercera ecuación: Ecuación del filtro de Kalman

Las primeras serán las responsables de proyectar hacia adelante en el tiempo, el estado actual, por una parte y la covarianza del error por otra, a fin de obtener las estimaciones “a priori” correspondientes al siguiente periodo de tiempo, es decir, sin incorporar la información asociada a la nueva observación del siguiente periodo. Las segundas son el vehículo a través del cual se produce la retroalimentación en el modelo al incorporar una nueva medida a la estimación a priori se obtendrá la estimación “a posteriori” mejorada.

A continuación, se muestran las ecuaciones de propagación y actualización, adaptadas del filtro de Kalman al análisis y predicción de entornos productivos empresariales

Figura 24.

Ecuación de propagación para $x_{t+1|t}$

$$\hat{x}_{t+1|t} = F_t \hat{X}_{t|t} + G_t u_t \tag{11}$$

Figura 25.

Ecuación de propagación para $P_{t+1|t}$

$$P_{t+1|t} = F_t P_{t|t} F_t^1 + S_t Q_t S_t^1 \tag{12}$$

Figura 26.

Ecuación de actualización para K_t

$$K_t = P_{t|t-1} H_t' (H_t P_{t|t-1} H_t' + R_t)^{-1} \tag{13}$$

Figura 27.*Ecuación de actualización $\widehat{X}_{t|t}$*

$$\widehat{X}_{t|t} = \widehat{X}_{t|t-1} + K(y_t - H_t \widehat{X}_{t|t-1}) \quad (14)$$

Figura 28.*Ecuación de actualización para $P_{t|t}$*

$$P_{t|t} = (I - K_t H_t) P_{t|t-1} \quad (15)$$

La esencia del filtro de Kalman es obtener una estimación óptima del proceso que se irá mejorando paulatinamente, gracias a la retroalimentación obtenida en cada uno de los periodos. En otras palabras, dados unos valores iniciales, el filtro estimara recursivamente el estado del proceso en un momento de tiempo. Este valor a su vez será necesario para evaluar las observaciones. A través del error de medida se obtiene la retroalimentación necesaria para refinar el modelo según el cual se estimará un nuevo valor de estado, aunque este nuevo valor corresponderá al valor del siguiente periodo. La ventaja de este algoritmo respecto a otros es que el filtro de Kalman es capaz de escoger la forma óptima cuando se conocen las varianzas de los ruidos que afectan al sistema, además es el mejor estimador lineal si sus vectores tienen estadísticas arbitrarias y si éstas son variables aleatorias normales.

4.1.2. Tasa representativa del mercado colombiano (TRM)

La Tasa de Cambio Representativa del Mercado–TRM corresponde al promedio ponderado de las operaciones de compra y venta de contado de dólares de los Estados Unidos de América a cambio de moneda legal colombiana. Para que este cálculo sea representativo, se ponderan las transacciones según su monto y se realiza un promedio aritmético simple.

Entre las 5:30 y 6:00 de la tarde se publica la tasa oficial con vigencia hasta el próximo día hábil y en base al promedio de las operaciones de ese día, lo quiere decir que cuando un día es festivo en Colombia, festivo en Estados Unidos, o es sábado o domingo, la tasa oficial (TRM) seguirá siendo la calculada en el último día que operó el mercado, y tendrá vigencia hasta el día siguiente. La entidad encargada de calcular y validar la TRM diaria es la Superintendencia Financiera de Colombia. Sin embargo, el Banco de la República es responsable de definir la metodología para el cálculo de la TRM a través del manual del Departamento de Operaciones y Desarrollo de Mercados. Los cambios en los métodos de cálculo se definen mediante resoluciones posteriores publicadas a través circulares reglamentarias. (¿Cómo se calcula la TRM del dólar en Colombia?, 2022).

Se consideran únicamente las negociaciones directas de compra y venta de divisas con un monto superior a los 5.000 dólares, realizadas por bancos comerciales, sociedades comisionistas de bolsa, compañías de financiamiento, la Financiera de Desarrollo Nacional -FDN- y el Banco de Comercio Exterior de Colombia -BANCOLDEX, acordadas para el cumplimiento del mismo día de su negociación. Excluye OTC y derivados, las operaciones extrabursátiles pueden ser: canje de billetes de alta denominación, canje de billetes o monedas de baja denominación, canje de billetes deteriorados, consignación a cuentas de depósito, etc. Mientras que un derivado es un instrumento financiero cuyo valor depende de otro precio (en este caso, un instrumento cuyo precio depende del dólar estadounidense), negociado en el mercado principalmente para cubrir el riesgo de los tenedores de activos.

4.1.3. Python

Python fue creado por el científico informático Guido van Rossum, que había estado usando un lenguaje llamado ABC mientras trabajaba en Centrum Wiskunde & Informatics (CWI), el Instituto Nacional Holandés de Matemáticas e Informática. Aunque le gustaban algunos aspectos de ABC, estaba frustrado por la dificultad de difundir el idioma.

Durante las vacaciones de Navidad de 1989, van Rosen decidió intentar crear su propio lenguaje. Más de un año después, en febrero de 1991, subió la primera versión de su creación a USENET.

Por otro lado, disfruta leyendo los textos de los episodios de Monty Python's Flying Circus, la famosa compañía de comedia británica. En busca de un nombre que fuera "corto, único y ligeramente críptico", eligió llamarlo Python.

En términos técnicos, Python es un lenguaje de programación orientado a objetos de alto nivel con semántica dinámica incorporada, que se utiliza principalmente para el desarrollo de aplicaciones informáticas y web.

Es muy atractivo en el campo del desarrollo rápido de aplicaciones (RAD) porque proporciona opciones dinámicas de escritura y enlace dinámico. Python es relativamente simple y, por lo tanto, fácil de aprender, ya que requiere una sintaxis única que se centre en la legibilidad. Los desarrolladores pueden leer y traducir el código de Python más fácilmente que otros lenguajes.

Por lo tanto, esto reduce el costo de mantenimiento y desarrollo del programa, ya que permite que los equipos trabajen juntos sin barreras significativas de lenguaje y experimentación.

Además, admite el uso de módulos y paquetes, lo que significa que los programas se pueden diseñar en un estilo modular y el código se puede reutilizar en varios proyectos. Una vez

que se desarrolla un módulo o paquete, se puede ampliar para su uso en otros proyectos y se puede importar o exportar fácilmente.

Uno de los beneficios más importantes de Python es que tanto la biblioteca estándar como el intérprete están disponibles gratuitamente, tanto en formato binario como fuente.

Tampoco hay exclusividad, ya que Python y todas las herramientas necesarias están disponibles en todas las plataformas principales. Por tanto, es una opción multiplataforma bastante atractiva para los desarrolladores que no quieren pagar altos costes de desarrollo.

En resumen, es un lenguaje de programación relativamente fácil de aprender, y las herramientas necesarias también se proporcionan de forma gratuita. Esto lo hace accesible para casi todos. Si tienes tiempo para aprender, podrás crear esos proyectos que tienes en mente.

Python es un lenguaje de programación de propósito general, en otras palabras, puede usarse para casi cualquier cosa. Lo que es más importante, es un lenguaje interpretado, lo que significa que el código escrito no se traduce realmente a un formato legible por computadora en tiempo de ejecución.

Este lenguaje también se conoce como "lenguaje de secuencias de comandos" porque originalmente estaba destinado a ser utilizado en proyectos simples.

El concepto de un "lenguaje de secuencias de comandos" ha cambiado mucho desde sus inicios, ya que Python ahora se usa para escribir grandes aplicaciones de estilo empresarial, no solo simples.

Una encuesta de 2019 de usuarios de Python indicó que los usos más populares fueron el desarrollo web y el análisis de datos. Solo alrededor del 6% de los encuestados lo usan para el desarrollo de juegos o desarrollo de aplicaciones.

4.1.3.1. ¿Cómo funciona Python?

El lenguaje de programación Python utiliza módulos de código mutables en lugar de la larga lista de instrucciones de los lenguajes de programación funcionales. La implementación predeterminada de Python se llama "cpython". En resumen, no traduce su código a lenguaje de máquina o código de máquina que el hardware entienda. Esto en realidad lo convierte en un código de bytes. La CPU no puede entender este código de bytes. Entonces, necesitamos un intérprete llamado Python Virtual Machine (PVM) para ejecutar el código de bytes.

El intérprete de Python realiza las siguientes tareas para ejecutar programas:

4.1.3.2. Pasos de interpretación de Python

Paso 1: el intérprete lee el código o las instrucciones de Python. Luego verifica que las instrucciones tengan el formato correcto, es decir, comprueba la sintaxis de cada línea. Si se encuentran errores, detiene inmediatamente la traducción y muestra un mensaje de error.

Paso 2. Si no hay errores, es decir, si la instrucción o código Python está bien formado, el intérprete lo traduce a una forma equivalente en un lenguaje intermedio llamado "bytecode". De esta manera, después de ejecutar el script o el código Python, se traduce completamente a código de bytes.

Paso 3: el código de bytes se envía a la máquina virtual de Python y el código de bytes se ejecuta nuevamente en PVN. Si ocurre un error durante esta ejecución, se detiene con un mensaje de error.

5. Proceso de obtención de datos

5.1. Selección de la fuente de información

Con base en lo previamente establecido para este estudio, los tipos de datos de entrada que se utilizaron fue; información teórica de diferentes fuentes como páginas web y libros virtuales y datos numéricos que se obtuvieron de la página oficial del banco de la república en referencia a la TRM histórica. Para el primer tipo de datos, datos de texto no estructurados, se seleccionó la suficiente cantidad de información en los principales portales web y datos históricos colombianos relacionados con las predicciones de variables económicas y los modelos que se aplican, en donde evidenciamos la carente información disponible en español sobre el modelo de Kalman Filters.

5.2. Selección y delimitación de datos

En esta fase se observa la estructura en que las fuentes proveen los datos. En el caso de las predicciones históricas de las variables económicas (TRM), estas vienen en periodo de días y contienen los atributos “Fecha” y “TRM (COP/USD)”. Para los datos textuales, se selecciona el cuerpo de los diferentes libros y fuentes de información, el título y la fecha de publicación, depurando elementos como imágenes, publicidad y otros elementos que son propios de la fuente de información, pero no del tema a estudiar en sí.

5.3. Descarga de los datos

Luego de escogidas las fuentes de información y seleccionados los datos que se van a usar se procede con la obtención de los datos, para esto se escoge un horizonte de tiempo de 30 años que comprende el periodo desde el 27 noviembre de 1991 hasta el 15 de enero 2022 para los datos numéricos. A continuación, se descargan manualmente de la página oficial del Banco de la

República de Colombia (TRM Peso/dólar) de los cuales estudiaremos un período de 5 años (2013 a 2015).

Figura 14.

Banco de la República de Colombia TRM (2022)



Nota: Dicha imagen es extraída de la página oficial del banco de la república, teniendo en cuenta la información necesaria para realizar el modelo respecto a la tasa representativa de mercado (TRM). Fuente: (BANREP, 2021)

Figura 15.

Datos obtenidos del banco de la República de Colombia TRM (2022)

Tasa de cambio representativa del mercado (TRM)
 1.1.2. Serie histórica para un rango de fechas dado periodicidad diaria

Información disponible desde el 27 de noviembre de 1991.

Fecha (dd/mm/aaaa)	TRM (COP/USD)
15/01/2022	\$ 3.993,65
14/01/2022	\$ 3.950,40
13/01/2022	\$ 3.970,08
12/01/2022	\$ 4.011,65
11/01/2022	\$ 4.043,46
10/01/2022	\$ 4.043,46
09/01/2022	\$ 4.043,46
08/01/2022	\$ 4.043,46
07/01/2022	\$ 4.039,31
06/01/2022	\$ 4.042,36
05/01/2022	\$ 4.084,11
04/01/2022	\$ 4.082,75

Sheet1 (+)

Nota: Dicha imagen es extraída del archivo de Excel extraído directamente de la página oficial del banco de la república, teniendo en cuenta la información necesaria para realizar el modelo respecto a la tasa representativa de mercado (TRM). Fuente: (BANREP, 2021)

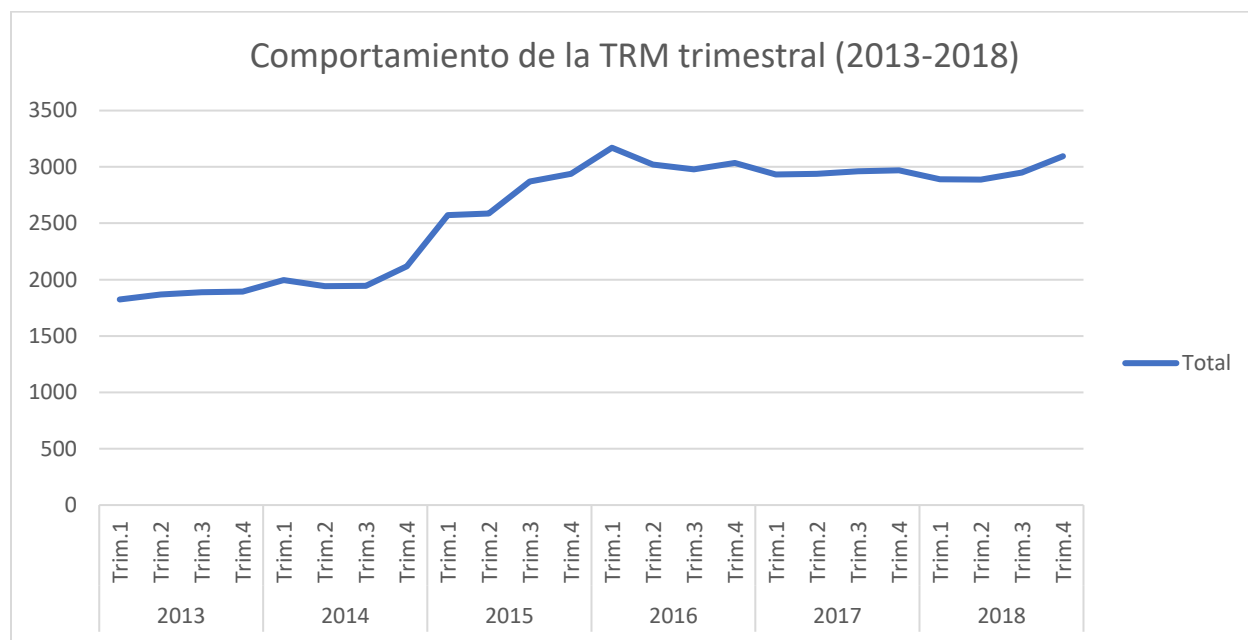
6. Análisis exploratorio de datos

Teniendo en cuenta la información recolectada en el ítem anterior, es posible realizar un análisis preliminar de la información, con el fin de evidenciar el comportamiento de la tasa representativa del mercado (TRM).

Para ello, se evaluará el comportamiento de este indicador en un periodo comprendido entre el 2013 al 2018 como se muestra en la siguiente figura:

Figura 16.

Comportamiento de la TRM trimestral 2013-2018



Nota: La figura anterior comprara los datos de la TRM, teniendo en cuenta un periodo de 5 años, dado el enfoque y limitación de este problema de investigación. Para mejor lectura del gráfico se calculó el valor de la TRM trimestral, teniendo en cuenta los datos extraídos del banco de la república. Fuente: (BANREP, 2021)

Como se puede evidenciar en la figura anterior, el comportamiento de la TRM tiene una tendencia ascendente positiva, mostrando un decrecimiento en su pendiente desde el año 2016, estabilizándose alrededor de los 2700-3000.

Para nuestro estudio en el periodo que comprende de 2013 a 2018 la tasa de cambio inició con un valor de 2392,46 pesos colombianos el primero de enero de 2013, experimento un alza continua hasta el año 2016 en donde su punto máximo fue de 3170,20 pesos colombianos, esto muy seguramente por el modelo económico del momento, para los años siguientes 2017 y 2018 experimentó un declive gradual quedándose en un rango de 2933,95 y 3095,794 pesos colombianos. Su volatilidad se sintió en el transcurso de los años, debido a que estuvo marcada por coyunturas económicas como los precios del petróleo, las decisiones de los bancos centrales, incluido el colombiano, y otros factores que le pusieron presiones bajistas y alcistas a la moneda estadounidense. En el 2017 la variación en los tipos de cambio del mercado local fue impulsado por un petróleo más fuerte debido a la confianza en los recortes de producción de la OPEP.

El primer día en el cargo de Donald Trump también ejerció presión a la baja e incertidumbre en los mercados debido a las perspectivas de seguridad global bajo el presidente de EE. UU.

Adicionalmente la variable TRM en el periodo establecido de 5 años registró una media de 2925,615, una mediana de 2937,63 y una moda de 2897,83 en donde el punto mínimo registrado fue de 2362,42 y el punto máximo registrado fue de 3274,47.

7. Limpieza y procesamiento de datos

Antes de la implementación y análisis del Modelo Box Jenkins y el modelo Kalman Filters, surge la necesidad de realizar una limpieza y procesamiento de la información recopilada del banco de la república, así como lo es: la depuración de los datos, la generación de una tabla con únicamente los datos de interés teniendo en cuenta el periodo descrito dentro del proyecto (2013 – 2018), así como la implementación del modelo en Python, mediante Jupyter, un entorno de trabajo interactivo que permite desarrollar código basado en lenguaje Python.

Según lo anterior, basados en lo que trabajamos en Jupyter con el fin de preparar y depurar la información para la implementación de los modelos.

7.1. Cargue de bibliotecas

Dentro de Python, para la implementación del código, existen algunas bibliotecas que dentro de ellas contienen módulos de programación que brindan información al sistema de entrada y salida de archivos, lo cual, garantiza el funcionamiento de las herramientas de dicha biblioteca.

Según esto, se requieren instalar las siguientes bibliotecas:

Figura 17.

Bibliotecas cargadas en Python para la implementación de los modelos

```
import csv
import datetime
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np

import warnings
warnings.filterwarnings('ignore')
#EL ARCHIVO TSSLTOOLS DEBE ESTAR EN LA CARPETA DONDE CORRE EL NOTEBOOK
from tssltools_lab2 import kfs_res
from tssltools_lab2 import LGSS
```

Fuente: Elaboración propia en Jupyter / Anaconda (Anaconda, 2022)

Cabe resaltar, que bibliotecas como: *csv*, *datetime*, *numpy*, *pandas*, entre otras, son bibliotecas que ya se encuentran instaladas por defecto dentro del sistema de Python, teniendo en cuenta un intérprete de Python 3.10. No obstante, la biblioteca *tssltools_labs2* funciona para la implementación del modelo de Kalman Filters, pero esta no se encuentra instalada dentro del intérprete; es por esto que cabe la necesidad que en el momento que se requiere la utilización de esta biblioteca, se debe instalar el Anexo N°1: *tssltools_lab2.py*, archivo el cuál fue tomado del repositorio de *jrakhshanda*, desarrolladora de Software de la Universidad de Linköping (*jrakhshanda*, 2021)

7.2. Cargue y depuración de información

Luego de la instalación de bibliotecas según el ítem anterior, se procede a realizar el cargue de la información, considerando que el Anexo N°1 se encuentra en la misma instalación del intérprete. Para este paso, es necesario tener el archivo descargado del banco de la república en la misma ubicación que este archivo, para así, no tener inconvenientes en su lectura, tal como se muestra en la siguiente figura:

Figura 18.

Cargue de información en Python. Elaboración propia en Jupyter Notebook

```
In [47]: data_trm = pd.read_csv('TRM.csv')
```

```
In [48]: data_trm.head()
```

Out[48]:

	FECHA	TRM
0	27/11/1991	693.32
1	28/11/1991	693.99
2	29/11/1991	694.70
3	30/11/1991	694.70
4	1/12/1991	643.42

Nota: En esta figura, tan solo aparecen los primeros 5 datos, dado que se utilizó el comando “head ()”

Como se puede evidenciar en la figura anterior, la TRM del 27 de noviembre de 1991 tenía un valor de \$693.32, tal y como se puede confirmar directamente en la página del banco de la república como se muestra a continuación:

Figura 19.

Archivo descargado del Banco de la república con todo el Histórico de TRM (1991 – 2022)

Fecha (dd/mm/yyyy)	TRM (COP/USD)
27/11/1991	693,32
28/11/1991	693,99
29/11/1991	694,7
30/11/1991	694,7
1/12/1991	643,42
2/12/1991	643,42
3/12/1991	639,22
4/12/1991	635,7
5/12/1991	631,51
6/12/1991	627,16
7/12/1991	638,06
8/12/1991	638,06
9/12/1991	638,06
10/12/1991	622,92

Nota: Tal como se muestra en la figura, desde TRM del 27/11/1991 concuerda exactamente con los datos colocados en el programa Python, mediante Jupyter Notebook.

Luego de verificar que la información ingresada a Jupyter Notebook es la correcta, se procede a: dar a la columna de fecha un formato específico de fecha, dar a la columna un formato numérico, y, no menos importante, filtrar la información a interés de este proyecto entre el 1 de enero del 2013 hasta el 31 de diciembre del 2018, como se evidencia en la siguiente figura:

Figura 20.

Tasa representativa del mercado Periodo 2013-2018

```
In [49]: data_fecha = pd.to_datetime(data_trm['FECHA'])

In [50]: df = pd.DataFrame() #SE REQUIRE CREAR DATAFRAME VACIO PARA PODER INCLUIR DATOS
df['Fecha']=data_fecha[7707:9897]

In [51]: data_TRM = data_trm['TRM']

In [52]: data_TRM

Out[52]: 0      693.32
1      693.99
2      694.70
3      694.70
4      643.42
...
10937   3881.76
10938   3881.76
10939   3881.76
10940   3874.41
10941   3876.86
Name: TRM, Length: 10942, dtype: float64
```

Nota: Se toma el valor de 7707:9897, teniendo en cuenta el histórico de datos para así filtrar únicamente las tasas entre el 2013 al 2018

Adicional a esto, es necesario, dado ambos modelos, dejar únicos los valores de la TRM, y dado el caso que en dos o más fechas se repita la misma TRM, es necesario eliminarlo; según esto, se realizó la depuración como se muestra en la siguiente figura:

Figura 21.

Depuración de la información entre el 2013 y el 2018 para aplicación de modelos. Elaborado en Jupyter Notebook

```
In [65]: #CUENTA CUANTOS VALORES UNICOS EXITES
fecha.value_counts()
```

```
Out[65]: 2013-02-01    1
         2017-03-01    1
         2016-12-28    1
         2016-12-29    1
         2016-12-30    1
         ..
         2014-12-27    1
         2014-12-26    1
         2014-12-25    1
         2014-12-24    1
         2018-12-31    1
         Name: Fecha, Length: 2190, dtype: int64
```

```
In [66]: #CUENTA CUANTAS VECES SE REPITE UN DATO
(fecha.str.replace('-', '').astype(int)/10000).value_counts()
```

```
Out[66]: 2013.0201    1
         2017.0301    1
         2016.1228    1
         2016.1229    1
         2016.1230    1
         ..
         2014.1227    1
         2014.1226    1
         2014.1225    1
         2014.1224    1
         2018.1231    1
         Name: Fecha, Length: 2190, dtype: int64
```

```
In [67]: #REDONDEAR Y MUESTRA VALORES UNICOS
AÑO=(fecha.str.replace('-', '').astype(int)/10000).round().unique()
print(AÑO)
```

```
[2013. 2014. 2015. 2016. 2017. 2018.]
```

7.3. Gráfica de la información de interés en Jupyter Notebook

Ya conociendo la tasa representativa del mercado a interés de este trabajo, que es entre el 2013 al 2018, se realiza la gráfica en Python con el fin de garantizar el correcto funcionamiento de los datos como se muestra a continuación.

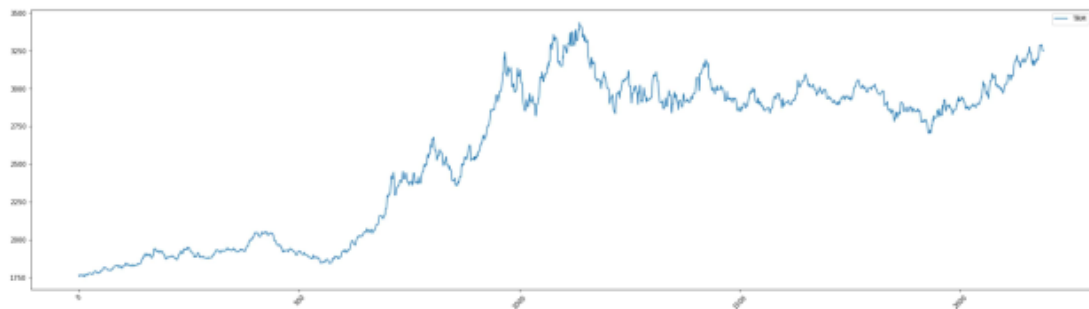
Figura 22.

Gráfica de la TRM entre el 2013-2018 en Jupyter Notebook. Elaboración propia

```
In [56]: from pandas import read_csv
from pandas import datetime
from matplotlib import pyplot

plt.rcParams["figure.figsize"] = (30,8)
series = read_csv('TRM_2013-2018.csv')
print(series.head())
series.plot()
pyplot.xticks(rotation=45)
pyplot.show()
```

	Fecha	TRM
0	2013-02-01	1768.23
1	2013-03-01	1759.97
2	2013-04-01	1760.83
3	2013-05-01	1767.54
4	2013-06-01	1767.54



Al momento de realizar la gráfica de la TRM, desde el 2013 hasta el 2018, es notorio un aumento paulatino de este indicador, teniendo en cuenta factores como inflación, deuda externa, producto interno bruto, nivel de importaciones y exportaciones, entre otros factores que llegan a afectar la tasa.

8. Modelo Box Jenkins

Para realizar la estimación del modelo Box Jenkins se realiza la utilización de modelos autorregresivos integrados de promedios móviles (En sus siglas, ARIMA).

Para su implementación, se debe seguir los pasos como lo indica en la siguiente tabla:

Tabla 2.

Pasos para la implementación del modelo Box Jenkins

#	Paso	Descripción
Paso 1	Identificación del modelo	En esta etapa determinamos el modelo y el orden de estos que mejor se ajustan a estos datos, es decir, encontramos valores apropiados de p, d y q. En esta etapa utilizamos métodos gráficos, el correlograma y el correlograma parcial.
Paso 2	Estimación de parámetros del método escogido	Estimar el modelo seleccionado, es decir estimar los parámetros de los términos autorregresivos y de promedios móviles incluidos en el modelo. Este cálculo puede efectuarse mediante mínimos cuadrados simples, pero puede recurrirse a métodos de estimación no lineal en los parámetros.

Continuación Tabla 2

Pasos para la implementación del modelo Box Jenkins

Paso 3	Diagnóstico del modelo	<p>Después de seleccionar un modelo ARIMA en particular y de estimar sus parámetros, veremos si el modelo seleccionado se ajusta adecuadamente a nuestros datos ya que en caso de que no, puede que exista otro modelo ARIMA que se ajuste mejor a los datos, para esto existen dos métodos: **El primero consiste en crear un modelo con más rezagos para identificar que no sean estadísticamente significativos. **El segundo consiste en analizar los residuos del modelo seleccionado. Esta etapa es fundamental para continuar con el pronóstico, ya que en caso contrario es necesario volver al paso uno y replantear la metodología, esto lo comprobaremos mediante la prueba Ljung-Box.</p>
Paso 4	Predicción del modelo	<p>La popularidad de estos modelos, está en la capacidad de éxito en el pronóstico, que en muchas ocasiones resulta mejor que modelos tradicionales de econometría y otros modelos normativos de teoría financiera. En esta etapa final realizaremos inferencias y pronósticos con el modelo estimado para la toma de decisiones de inversión a corto plazo.</p>

Nota: Esta tabla está elaborada basada en la información del documento “SERIES TEMPORALES, MODELO ARIMA – METODOLOGÍA DE BOX – JENKINS” del profesor Santiago de la Fuente Fernández de la Universidad Autónoma de Madrid. Fuente: (de la Fuente Fernández, 2020)

8.1. Paso 1: Identificación del modelo

Para llegar a cabo la identificación del modelo, se calculó la autocorrelación y autocorrelación parcial, dado que estas son medidas de relación entre los valores de series actuales y los valores anteriores, lo cual, define cuál de las dos es más útil para predecir los valores futuros. Para dicho proceso, en Python, se utilizó la librería de “*pandas.plotting*” que contiene la gráfica de dichas correlación, tal como se muestra en la siguiente figura:

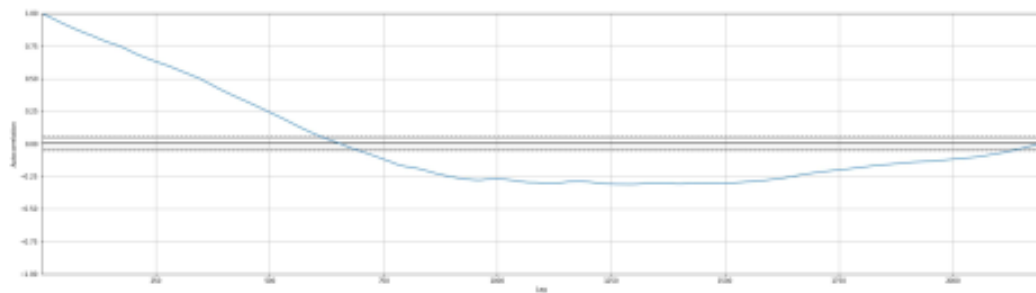
Figura 23.

Autocorrelación y autocorrelación parcial del modelo

```
In [66]: from pandas import read_csv
from pandas import datetime
from matplotlib import pyplot
from pandas.plotting import autocorrelation_plot

...
La autocorrelación y la autocorrelación parcial son medidas de la relación entre
y pasadas e indican qué valores de series pasadas son más útiles para predecir va
...

series = read_csv('TRM_2013-2018.csv')
autocorrelation_plot(series['TRM'])
pyplot.show()
```



```
In [67]: series = read_csv('TRM_2013-2018.csv')
```

Al momento de realizar un análisis de autocorrelación de las series de la TRM, desde el 2013 hasta el 2018, se evidencian en las líneas negras, los valores que mejor funcionan para así pronosticar futuramente este indicador.

8.2. Paso 2: Estimación de parámetros iniciales

Para realizar la estimación de parámetros iniciales del modelo, es necesario la instalación de “ARIMA” en Python, función que se encuentra en la biblioteca “*statsmodels.tsa.arima.model*”, hacer el cargue de la información, realizar el entrenamiento del modelo y calcular la tabla resumen de la ejecución del modelo, tal como se muestra en la siguiente figura:

Figura 24.

Entrenamiento y resumen de ARIMA para la serie 2013-2018. Elaboración propia en Jupyter

Notebook

```
In [75]: # fit an ARIMA model and plot residual errors
from pandas import datetime
from pandas import read_csv
from pandas import DataFrame
from statsmodels.tsa.arima.model import ARIMA
from matplotlib import pyplot
# Load dataset

series = read_csv('TRM_2013-2018.csv', header=0, index_col=0, parse_dates=True)
series.index = result
# fit model
model = ARIMA(series, order=(5,1,0))
model_fit = model.fit()
# summary of fit model
print(model_fit.summary())
# line plot of residuals
residuals = DataFrame(model_fit.resid)
residuals.plot()
pyplot.show()
# density plot of residuals
residuals.plot(kind='kde')
pyplot.show()
# summary stats of residuals
print(residuals.describe())
```

SARIMAX Results

```
=====
=
Dep. Variable:          TRM    No. Observations:          219
1
Model:                ARIMA(5, 1, 0)  Log Likelihood            -9352.44
0
Date:                 Fri, 13 Jan 2023  AIC                        18716.88
1
Time:                 21:35:17      BIC                        18751.03
1
Sample:                0      HQIC                       18729.36
2
- 2191
```

Nota: Teniendo en cuenta los resultados del ARIMA, se puede conocerlos indicadores de entrada del modelo; cabe resaltar que el valor log likelihood es una medida de bondad de ajuste

para cualquier modelo, con el fin de conocer, a manera comparativa con otro modelo, cuál es mejor. Para este modelo, da un buen valor de 9352.44.

8.3. Paso 3: Diagnóstico del modelo

Para realizar el diagnóstico del modelo, se realizó el cálculo de la raíz del error cuadrático medio (RMSE, por sus siglas en inglés), el cuál mide la cantidad de error que hay entre dos conjuntos de datos; teniendo en cuenta la documentación del IBM respecto al RMSE, se puede afirmar que: este indicador muestra la diferencia entre los valores previstos y los valores observados del modelo. (IBM Documentation, 2022)

Por consiguiente, este indicador compara la estimación predicha de un modelo con el valor observado ya que, entre menor sea el valor del RMSE, más cercana será la estimación con el valor observado.

El indicador mencionado anteriormente se evidencia en la siguiente figura:

Figura 25.

Cálculo del RMSE para el modelo ARIMA. Elaboración propia en Jupyter Notebook

```
In [76]: # evaluate an ARIMA model using a walk-forward validation
from pandas import read_csv
from pandas import datetime
from matplotlib import pyplot
from statsmodels.tsa.arima.model import ARIMA
from sklearn.metrics import mean_squared_error
from math import sqrt
# Load dataset
series = read_csv('TRM_2013-2018.csv', header=0, index_col=0, parse_dates=True, s
series.index = result
# split into train and test sets
X = series.values
size = int(len(X) * 0.66)
train, test = X[0:size], X[size:len(X)]
history = [x for x in train]
predictions = list()
# walk-forward validation
for t in range(len(test)):
    model = ARIMA(history, order=(5,1,0))
    model_fit = model.fit()
    output = model_fit.forecast()
    yhat = output[0]
    predictions.append(yhat)
    obs = test[t]
    history.append(obs)
    print('Pronostico=%f, Esperado=%f' % (yhat, obs))
# evaluate forecasts
rmse = sqrt(mean_squared_error(test, predictions))
print('Test RMSE: %.3f' % rmse)
# plot forecasts against actual outcomes
pyplot.rcParams["figure.figsize"] = (25,8)
pyplot.plot(test)
pyplot.plot(predictions, color='orange')
pyplot.show()
```

Test RMSE: 14.737

Nota: El indicador RMSE, es un valor porcentual, por lo tanto, está indicando que los datos que se calcularán de la predicción del modelo, se parecerán en un 20.487% respecto a los

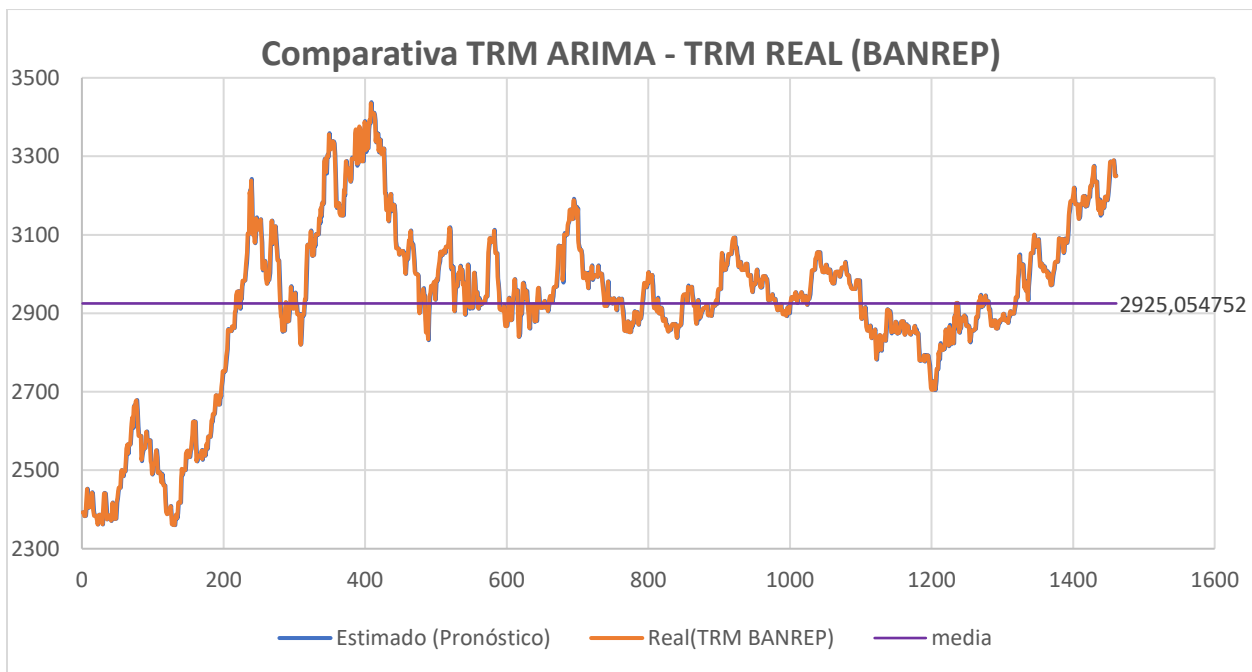
datos o, también significa que la predicción está equivocada en un 20.487% respecto a los valores tomados.

8.4. Paso 4: Predicción del modelo

Teniendo en cuenta el diagnóstico del modelo, se procede a realizar la estimación de la predicción, tomando como referencia los parámetros del ítem 8.2, dando como resultado la siguiente gráfica:

Figura 26.

Gráfica de predicción de la TRM por el modelo ARIMA CS TRM Real. Elaboración propia Excel



Nota: En esta gráfica se descargaron los datos suministrados por la implementación en Jupyter, para, mediante el uso de Excel, obtener una gráfica comparativa de este. La línea azul representa la predicción del modelo ARIMA, mientras que la línea naranja representa la gráfica de los datos extraídos del banco de la república.

Cómo se puede observar en la figura anterior, las líneas en azul presentan la predicción del Modelo ARIMA, y las naranjas la representación gráfica de los datos extraídos del banco de la república; adicional, como se mencionó anteriormente, el RMSE es un valor aceptable, y el pronóstico lo comprueba dado que el modelo representa los datos reales con un error del 20.487%, o lo que es lo mismo, los datos pronosticados por el modelo ARIMA, se parecen en un 79.513% a los datos reales de la TRM dados por la base de datos del banco de la república y la media obtenida es de 2925,054.

9. Modelo Kalman Filters

Para la implementación del Filtro de Kalman, luego de realizar el cargue y depuración de la información, es importante validar que la biblioteca del Anexo N°1 se encuentra activo, para lo cual es necesario probar la línea de código que se muestra en la siguiente figura:

Figura 27.

Instalación de tssltools_lab2 para la implementación del modelo de Kalman Filters

```
In [77]: #EL ARCHIVO TSSLTOOLS DEBE ESTAR EN LA CARPETA DONDE CORRE EL NOTEBOOK
from tssltools_lab2 import LGSS
model = LGSS(T, R, Q, Z, sigma_eps**2, a1, P1)
help(model.get_params)
```

```
Help on method get_params in module tssltools_lab2:
```

```
get_params() method of tssltools_lab2.LGSS instance
Return all model parameters.
```

```
T, R, Q, Z, H, a1, P1 = model.get_params()
```

Nota: Como se muestra en la figura anterior, se realiza una prueba de la función LGSS contenida dentro del módulo instalado del Anexo N°1. La función “help” de Python se utiliza con el fin de obtener ayuda de esta función, y en caso de arrojar una solución indicará que ha sido instalada correctamente y ya podemos utilizarla.

9.1. Código del modelo del Filtro de Kalman

Teniendo en cuenta que ya contamos con la biblioteca del Anexo N°1 instalada y los módulos importados en nuestro IDE (Entorno integrado de desarrollo) ya es posible implementar los códigos de la siguiente figura:

Figura 28. Código de implementación del Filtro de Kalman. Elaboración propia en Jupyter

Notebook - Colab

```

[49] def kalman_filter(y, model: LGSS):
    """Kalman filter for LGSS model with one-dimensional observation.

    :param y: (n,) array of observations. May contain nan, which encodes missing observations.
    :param model: LGSS object with the model specification.

    :return kfs_res: Container class with member variables,
        alpha_pred: (d,1,n) array of predicted state means. alpha_hat t|t-1
        P_pred: (d,d,n) array of predicted state covariances. predict t|t-1
        alpha_filt: (d,1,n) array of filtered state means. alpha_hat t-1|t-1
        P_filt: (d,d,n) array of filtered state covariances. predict t-1|t-1
        y_pred: (n,) array of means of p(y_t | y_{1:t-1}) y_hat t|t-1
        F_pred: (n,) array of variances of p(y_t | y_{1:t-1}) F t|t-1
    """

    n = len(y)
    d = model.d # State dimension
    alpha_pred = np.zeros((d, 1, n))
    P_pred = np.zeros((d, d, n))
    alpha_filt = np.zeros((d, 1, n))
    P_filt = np.zeros((d, d, n))
    y_pred = np.zeros(n)
    F_pred = np.zeros(n)

    T, R, Q, Z, H, a1, P1 = model.get_params() # Get all model parameters (for brevity)

```

```

for t in range(n):
    # Time update (predict)
    # ADD CODE HERE
    #predict alpha_t in lecture 5 slide 15
    # for alpha_hat 1|0 = a1 and p1|0 = p1
    if t == 0 :
        alpha_pred[:, :, t] = a1
        P_pred[:, :, t] = P1
    if t != 0 :
        alpha_pred[:, :, t] = T@alpha_filt[:, :, t-1]
        #P_pred[:, :, t]= T.dot(P_filt[:, :, t-1].dot(T.transpose())) + R.dot(Q.dot(R.transpose()))
        P_pred[:, :, t] = T@P_filt[:, :, t-1]@(T.transpose()) + R@Q@(R.transpose())

    # Compute prediction of current output
    # ADD CODE HERE
    #predict yt in lecture 5 slide 15

    y_pred[t] = Z@alpha_pred[:, :, t]
    F_pred[t] = Z@P_pred[:, :, t]@(Z.transpose())+ H

```

```

if np.isnan(y[t]):
    alpha_filt[:, :, t] = alpha_pred[:, :, t].copy()
    P_filt[:, :, t] = P_pred[:, :, t].copy()
else:
    K = P_pred[:, :, t]@(Z.transpose())*(1/F_pred[t])
    alpha_filt[:, :, t] = alpha_pred[:, :, t] + (K*((y[t] - y_pred[t])))

    I = np.identity(d)
    P_filt[:, :, t] = (I - (K@Z))@P_pred[:, :, t]

kf = kfs_res(alpha_pred, P_pred, alpha_filt, P_filt, y_pred, F_pred)
return kf

```

Nota: Se puede apreciar el código completo de implementación del filtro de Kalman, donde, iteración a iteración según los parámetros del modelo, va realizando las iteraciones necesarias para llegar a una buena estimación.

9.2. Implementación del Filtro de Kalman para la estimación de la TRM

Considerando la figura 28, es posible realizar la implementación del filtro de Kalman para el problema propuesto en este proyecto, tal como se muestra en la siguiente figura:

Figura 29.*Implementación del Filtro de Kalman para la estimación de la TRM*

```

✓ 1 h ▶ n = 2000
train_val = y.copy()
train_val[n:] = np.nan
kalman = kalman_filter(train_val,model)
sd = np.sqrt(kalman.F_pred)
plt.plot(u,y, label = "Data TRM")

plt.plot(u,kalman.y_pred,'y', label = "Y_Predicción FILTRO DE KALMAN")
plt.plot(u,kalman.y_pred+sd,'c--' )
plt.plot(u,kalman.y_pred-sd,'c--' )

plt.axvline(x=u[n])
plt.xlabel("FECHA")
plt.ylabel("TRM")
plt.title("Kalman Filter")
plt.legend()

```

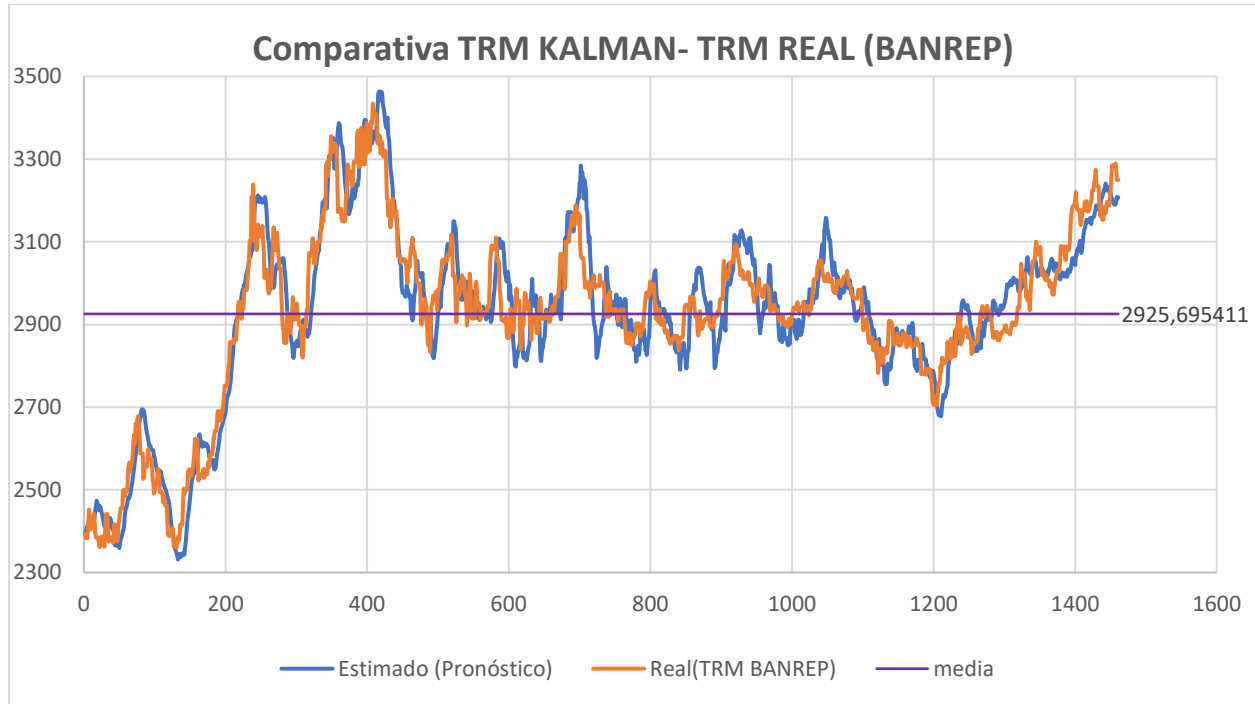
Nota: El número de iteraciones puede variar según la exactitud del modelo, es decir, entre mayor sea el número de iteraciones a realizar (Variable $n = 2000$) mayor será la exactitud del modelo, por lo tanto, menor será el RMSE (Error cuadrático medio), el cual, nos ayudará de manera comparativa en los modelos implementados en este trabajo.

Adicional, en la figura se evidencia, en la parte superior izquierda que la implementación tardó aproximadamente una hora en ejecutarse, dando como resultado la siguiente gráfica comparativa de la TRM real con la estimada de este modelo.

Figura 30.

Gráfica de predicción de la TRM por el filtro de Kalman VS TRM Real. Elaboración propia

Excel



Nota: En esta gráfica se descargaron los datos suministrados por la implementación en Jupyter, para, mediante el uso de Excel, obtener una gráfica comparativa de este. La línea azul representa la predicción del filtro de KALMAN, mientras que la línea naranja representa la gráfica de los datos extraídos del banco de la república.

Cómo se puede observar en la figura anterior, las líneas en azul presentan la predicción del filtro KALMAN, y las naranjas la representación gráfica de los datos extraídos del banco de la república; adicional, como se mencionó anteriormente, el RMSE es un valor aceptable, y el pronóstico lo comprueba dado que el modelo representa los datos reales con un error del 72.990%, o lo que es lo mismo, los datos pronosticados por el filtro KALMAN, se parecen en un 27.01% a

los datos reales de la TRM dados por la base de datos del banco de la república y la media obtenida es de 2925,695.

10. Comparación de resultados

Para realizar una comparación directa de los modelos, existe la necesidad de calcular el error cuadrático medio (RMSE, por sus siglas en inglés “*root mean square error*”), el cual indica que tan lejos, porcentualmente, se encuentra nuestra estimación de los datos reales.

Considerando lo anterior, se muestra la tabla con los valores RMSE calculados para los modelos de Box Jenkins (ARIMA) y Kalman Filters.

Tabla 3.

RMSE para el Box Jenkins y Kalman Filters

Modelo	MSE	RMSE
Box Jenkins (ARIMA)	419,720	20,487
Kalman Filters	5327,635	72,990

Nota: Para calcular el valor del RMSE, primero es necesario calcular el valor del MSE, el cuál mide el promedio de errores elevados al cuadrado. Este valor, es estrictamente positivo, ya que se calcula del cuadrado de la resta entre el valor real y el valor estimado. El RMSE se determina de la raíz cuadrada del MSE.

Como se evidencia en la tabla anterior, el modelo ARIMA (Box Jenkins) tiene un valor de RMSE del 20,487%, mientras que, para el modelo de Kalman Filters, dicho indicador es del 72,990%, lo que da a entender que, con los parámetros dados al modelo, el que da una mejor proximidad respecto al comportamiento real de la TRM es el modelo de Box Jenkins.

Una medida de error de pronóstico calculada para un solo método durante un solo período de tiempo no tiene sentido. Aquí es donde la métrica de error es útil cuando la comparamos con métricas de otros métodos de pronóstico o con métricas de otros períodos de tiempo. Por esto otros indicadores que respaldan el indicador anteriormente mencionado RMSE son el CFE, MAD y MAPE.

Suma acumulada de errores de pronóstico (CFE): Es la suma acumulada de los errores de pronóstico. Nos permite evaluar el sesgo de pronóstico. Por ejemplo, si el valor real de la demanda a lo largo del tiempo siempre es mayor que el valor pronosticado, el CFE será mayor, lo que indica que existe un error sistemático en el cálculo de la demanda.

- Aplicando la suma acumulada de errores de pronóstico obtenemos que para el modelo ARIMA da un resultado de 818,89 y para el filtro de KALMAN un resultado de -117,12 dando así que el de mejor resultado es KALMAN.

Figura 51.

Ecuación de CFE

$$CFE = \sum Error\ de\ pronóstico \quad (16)$$

Desviación media absoluta (MAD): Mide la dispersión de los errores de pronóstico, es decir, el tamaño de la unidad de error. Es el valor absoluto de la diferencia entre la demanda real y la prevista dividida por el número de períodos.

- Aplicando la desviación media absoluta obtenemos que para el modelo ARIMA da un resultado de 12,86 y para el filtro de KALMAN un resultado de 58,03 dando así que el de mejor resultado es el modelo ARIMA.

Figura 52.

Ecuación de MAD

$$MAD = \frac{\sum |Real - Pronóstico|}{n} \tag{17}$$

Error porcentual medio absoluto (MAPE): Es la desviación en porcentaje, no en unidades como las medidas anteriores. Es el promedio del error absoluto o diferencia entre la demanda real y la prevista, expresado como porcentaje de la real.

- Aplicando el error porcentual medio absoluto obtenemos que para el modelo ARIMA da un resultado de 0,439% y para el filtro de KALMAN un resultado de 1,981% dando así que el de mejor resultado es el modelo ARIMA.

Figura 53.

Ecuación de MAPE

$$MAPE = \frac{\sum_i^n 100 |Real_i - Pronóstico_i|}{\frac{Real}{n}} \tag{18}$$

Tabla 4.*CFE, MAD y MAPE para el Box Jenkins y Kalman Filters*

	ARIMA	KALMAN
CFE	818,89	-117,12
MAD	12,86	58,03
MAPE	0,439%	1,981%

Nota: si bien el filtro de KALMAN obtiene el mejor resultado en CFE, termina siendo descartado por las demás medidas. ARIMA presenta los mejores resultados en 3 de las 4 medidas.

11. Conclusiones

Se analizó la eficiencia de los modelos de pronóstico Box Jenkins y Kalman Filter, llegando a ultimar que la tasa representativa del mercado (TRM), teniendo en cuenta los parámetros dados para los modelos, es más cercana la estimación del modelo Box Jenkins (ARIMA) respecto al modelo de Kalman Filters, dado que el valor del error cuadrático medio (RMSE) es mayor para este último modelo.

Se realizó una revisión bibliográfica sobre la aplicación de modelos y técnicas relacionadas con la estimación de indicadores económicos en diferentes países, llegando a definir óptima la solución del modelo ARIMA, teniendo en consideración que el modelo de Kalman Filters puede llegar a ser más eficiente con unos parámetros diferentes, pero, utilizando más recursos computacionales en su ejecución.

Por último, este trabajo destaca la importancia de estimar correctamente el pronóstico de la Tasa representativa del mercado (TRM), valor de conversión entre los pesos colombianos y el dólar, dado que es totalmente necesario para la toma de decisiones financieras como: inversiones en bolsa, ETF's, CDT's, compra y venta de dólares, importaciones y/o exportaciones, y demás transacciones y movimientos que involucren directamente el peso colombiano y el dólar.

12. Recomendaciones

Para futuras investigaciones respecto a este tema, se recomienda:

1. Optimizar el modelo de Kalman Filter con el fin de mejorar los recursos computacionales que utiliza para la ejecución del modelo, y así, llegar a modificar los parámetros fácilmente y obtener mejores resultados de este.
2. Este trabajo enfatiza un periodo de evaluación de la TRM entre el 2013 y el 2018, lo cual, para futuras actualizaciones, es necesario utilizar periodos de estudio más actuales, para así, conseguir resultados más allegados a la realidad de la TRM.
3. Al contar con un modelo que minimiza el error cuadrático medio (RMSE), y en general, todos los errores que pueden llegar a tener estos respecto a su proximidad con el comportamiento real de los datos, se puede pensar en aplicar en diferentes indicadores económicos que causan hoy en día gran impacto al país.
4. Resulta pertinente y recomendado basar próximas investigaciones con fuertes bases estadísticas que permitan tener una confiabilidad alta en los resultados obtenidos.

Referencias Bibliográficas

- Anaconda. (2022). *Jupyter, Anaconda*. Obtenido de Anaconda:
<https://anaconda.org/anaconda/jupyter>
- Arango A, C., Misas, M., & López, E. (2011). *Economía Subterránea en Colombia 1976-2003: Una medición a partir de la Demanda de Efectivo*. Obtenido de Economía Subterránea en Colombia 1976-2003: Una medición a partir de la Demanda de Efectivo:
<https://ideas.repec.org/p/bdr/borrec/335.html>
- BANREP. (2021). *Tasa Representativa del Mercado (TRM - Pesos por Dolar)*. Obtenido de
<https://www.banrep.gov.co/es/estadisticas/trm>
- Casas Tragadora, C. (2002). *Universidad del Pacífico*. Obtenido de Econometría moderna:
<https://faculty.up.edu.pe/es/publications/econometr%C4%B1a-moderna>
- Cheung, Y.-W., Chinn, M., & Garcia Pascual, A. (2005). *Modelos empíricos de tipo de cambio de los noventa: ¿alguno es apto para sobrevivir?* Obtenido de Sciencedirect:
<https://www.sciencedirect.com/science/article/abs/pii/S0261560605000719>
- de la Fuente Fernández, S. (2020). *Series Temporales. Modelo ARIMA - Metodología Box - Jenkins*. Obtenido de Series Temporales. Modelo ARIMA - Metodología Box - Jenkins:
<https://www.estadistica.net/ECONOMETRIA/SERIES-TEMPORALES/modelo-arima.pdf>
- Dorado , J., & Ruiz, J. (2013). *Implementación de filtros de Kalman como método de ajuste a los modelos de pronóstico (GFS) de temperaturas máximas y mínimas para algunas ciudades de Colombia*. Obtenido de Documentos alojados en el IDEAM:

http://www.ideam.gov.co/documents/21021/21132/Filtro_de_Kalman.pdf/e98277fe-a85c-49b6-b11b-e415f262c066

Echavarría Soto, J., López Enciso, E., Misas Arango, M., Téllez Corredor, J., & Parra Alvarez, J. (2018). *La Tasa de Interés Natual en Colombia*. Obtenido de Banco de la república: <https://www.banrep.gov.co/docum/ftp/borra412.pdf>

EcuRed. (Noviembre de 2010). *EcuRed - Robert Fry Engle*. Obtenido de EcuRed - Robert Fry Engle: https://www.ecured.cu/Robert_Fry_Engle

Galvis, C. (1991). *Construcción de un modelo para las series del PIB trimestral usando filtros de Kalman*. Obtenido de <https://scienti.minciencias.gov.co/gruplac/jsp/visualiza/visualizagr.jsp?nro=00000000000453>

Giraldo Gomez, N. (29 de Junio de 2005). *Predicción de betas y VaR de portafolios de acciones mediante el filtro de Kalman y los modelos GARCH*. Obtenido de http://www.scielo.org.co/scielo.php?pid=S0120-35922005000100006&script=sci_arttext&tlng=es

Gómez, B. (26 de Enero de 2020). *Rpubs by Rstudio*. Obtenido de Rpubs by Rstudio: <https://rpubs.com/brianz0r/574656>

H. Stock, J., & M. Watson, M. (2020). *Introducción a la Econometría*. España: Pearson Educación S.A.

IBM Documentation. (01 de 12 de 2022). *Raíz del error cuadrático medio*. Obtenido de <https://www.ibm.com/docs/es/cloud-paks/cp-data/4.5.x?topic=overview-root-mean-squared-error>

- jrakhshanda. (2021). *Time-Series-Sequence-Learning*. Obtenido de Time-Series-Sequence-Learning: <https://github.com/jrakhshanda/Time-Series-Sequence-Learning/tree/master/Lab01>
- Luis Edwin, A. A., & Danny, C. R. (2010). *Repositorio UTP*. Obtenido de titulado análisis de pronósticos de precios mediante series temporales aplicados al caso de la bolsa de valores colombiana: <https://repositorio.utp.edu.co/server/api/core/bitstreams/9363be09-ed03-43a8-b513-be1abb7d65c0/content>
- MATHWORKS. (2020). *Diseño y utilización de filtros de Kalman en MATLAB y Simulink*. Obtenido de Diseño y utilización de filtros de Kalman en MATLAB y Simulink: <https://la.mathworks.com/discovery/kalman-filter.html>
- Panseri, C., Urga, G., & Cristini, A. (1999). *The Application of the Kalman Filter to the Fisher Equation: Italian and German Term Structure of Interest Rates*. Obtenido de Ideas - Repec: <https://www.semanticscholar.org/paper/The-Application-of-the-Kalman-Filter-to-the-Fisher-Panseri-Urga/6941a51b018865f969e18fb6175417d1e3b6ab3c>
- Scopus. (Diciembre de 2022). *Scopus*. Obtenido de Scopus: <https://www.scopus.com/freelookup/form/author.uri?zone=TopNavBar&origin=NO%20ORIGIN%20DEFINED>
- Sotoca López, S. (1992). *Aplicaciones econométricas del filtro de Kalman y algunas variaciones numéricas: el Filtro de Chandrasekhar*. Obtenido de Aplicaciones econométricas del filtro de Kalman y algunas variaciones numéricas: el Filtro de Chandrasekhar: <https://eprints.ucm.es/id/eprint/3445/1/T17765.pdf>

UNIVERSIDAD DE CHILE. (1937). *FACULTAD DE DERECHO - CANALES DE LA FACULTAD DE CIENCIAS JURIDICAS Y SOCIALES*. Obtenido de LA CIENCIA DE LA ECONOMÍA:

https://web.uchile.cl/vignette/analesderecho/CDA/an_der_simple/0,1362,SCID%253D227%2526ISID%253D11%2526PRT%253D214,00.html

Uriel, E., & Muñiz, m. (1998). *Estadística Económica y empresarial - Teoría y ejercicios*.

Vargas Vigoya, P., & Mier Santander, M. (2012). *Repositorio Universidad Javeriana*. Obtenido de Modelo de integración entre expectativas de empresarios y método estadístico para pronosticar el crecimiento de las ventas del sector industrial en colombia, una aplicación a la planeación de las empresas: <http://hdl.handle.net/10554/13679>

Visualizing scientific landscapes. (2022). *Visualizing scientific landscapes*. Obtenido de Visualizing scientific landscapes: <https://www.vosviewer.com/>

Lista de anexos

Anexo N°1: Archivo de tssltools_lab2.py



tssltools_lab2.py