



Pronósticos de variables climatológicas mediante los modelos de punto de cambio y Holt-Winters.

Forecasts of climatological variables using the change point and Holt-Winters models.

^{a*}Juan Camilo Valderrama-Balaguera, ^bHugo Fernando Castro-Silva, ^cCésar Augusto Dávila Carrillo.

 ^a Magíster en Ingeniería, jcvalderrama@uniboyaca.edu.co, Universidad de Boyacá, Tunja, Colombia.

 ^b Doctor en Gerencia de Proyectos, hugofernando.castro@uptc.edu.co, Universidad Pedagógica y Tecnológica de Colombia, Sogamoso, Colombia.

 ^c Maestrante en Educación Matemática, cesaraugustode@ufps.edu.co, Universidad Francisco de Paula Santander, Cúcuta, Colombia.

Recibido: Mayo 22 de 2021 **Aceptado:** Agosto 27 de 2021

Forma de citar: J.C. Valderrama-Balaguera, H.F. Castro-Silva, C.A. Dávila-Carrillo, "Pronósticos de variables climatológicas mediante los modelos de punto de cambio y Holt-Winters." *Mundo Fesc*, vol 11, no. S2 pp. 337-352, 2021.

Resumen: _____

En este estudio se analiza una serie de tiempo con datos históricos diarios desde enero 1 de 1989 hasta el 31 de diciembre del año 2021 de la variable precipitación con un total de 12053 observaciones, estos datos son obtenidos a partir de la estación climatológica Tunjuelito. Para la investigación se tuvieron en cuenta los registros de la variable "precipitación", el objetivo fue analizar las tendencias, utilizar los datos hasta el 31 de diciembre de 2020 para estimar un pronóstico para el año 2021 método de Holt-Winters y el modelo de punto de cambio, se comparan los datos observados con los pronosticados. Por último, se realizan pruebas estadísticas para contrastar el grado de similitud de los datos obtenidos a partir de los pronósticos con los datos observados arrojados por la estación. Los resultados demuestran que los pronósticos obtenidos con el modelo de punto de cambio evidencian una mayor precisión y se ajusta relativamente bien a los datos observados. Sin embargo este estudio se considera preliminar y para que los resultados puedan ser considerados como concluyentes de deben aplicar a una cantidad significativa de series de tiempo de variables meteorológicas.

Palabras clave: Holt-Winters, precipitación, pronósticos, punto de cambio.

Autor para correspondencia:

*Correo electrónico: jcvalderrama@uniboyaca.edu.co



© 2021. Fundación de Estudios Superiores Comfanorte.

Abstract:

This study analyzes a time series with daily historical data from January 1, 1989 to December 31, 2021 of the precipitation variable with a total of 12053 observations, these data are obtained from the Tunjuelito climatological station. For the research the records of the variable "precipitation" were taken into account, the objective was to analyze the trends, use the data up to December 31, 2020 to estimate a forecast for the year 2021 Holt-Winters method and the change point model, the observed data are compared with the predicted data. Finally, statistical tests are performed to contrast the degree of similarity of the data obtained from the forecasts with the observed data provided by the station. The results show that the data obtained from the change point model show higher accuracy and fit relatively well with the observed data. However, this study is considered preliminary and for the results to be considered conclusive they must be applied to a significant number of time series of meteorological variables.

Keywords: Holt-Winters, precipitation, forecasts, change point.

Introducción

El análisis de series de tiempo ha sido ampliamente utilizado para predecir eventos ambientales. Las condiciones climáticas pronosticadas en el futuro juegan un papel importante en el proceso de toma de decisiones de muchas organizaciones, para predecir eventos futuros de variables climatológicas, se necesita confiar en la información obtenida de eventos pasados. El pronóstico de variables climatológicas resulta de numerosas interacciones que crean sistemas complejos, es una tarea difícil pero importante. El modelado de series de tiempo para generar datos y pronosticar la precipitación es un paso importante en el diseño y análisis de los recursos hídricos.

Los pronósticos de variables ambientales es una herramienta muy útil en la gestión de los recursos naturales, la importancia de los pronósticos en temas medioambientales hace que utilicen métodos estadísticos más precisos para estudiar el tiempo y el cambio climático, el objetivo principal de este estudio es investigar el uso de formas aditivas y multiplicativas del modelo de series temporales de Holt-Winters y punto de cambio para predecir variables ambientales como la precipitación.

La precipitación es producida directamente

por la microfísica de las nubes, pero la ocurrencia de precipitación está asociada con dinámica ambiental y termodinámica de eventos climáticos [1]. Los pronósticos utilizados en variables climatológicas benefician al sector de la agricultura, ya que, en muchas regiones, la variabilidad climática interanual puede afectar la producción agrícola, disminuir los ingresos de los agricultores y aumentar los precios del mercado [2] *interannual climate variability can disrupt agricultural production, lower farmers' incomes and increase market prices. Increasing demand thus exists for agro-climatic services in the country. Fulfilling such demand requires robust and consistent approaches for seasonal climate forecasting. Here, we assess seasonal precipitation predictability and forecast skill at agriculturally-relevant timescales for five departments that represent key growing areas of major staple crops (rice, maize, and beans).*

Los recursos hídricos limitados para suministro de agua en el sector agrícola e industrial, han causado problemas importantes, particularmente en áreas áridas y semiáridas. Por lo tanto, el pronóstico y la estimación de la precipitación son uno de los parámetros climáticos más importantes para utilizar los recursos hídricos de manera eficiente [3]. En general,

se requiere la medición y predicción de esta variable para estudiar su comportamiento. Además, la predicción de la precipitación para cualquier tipo de territorio se considera uno de los parámetros climáticos más importantes para el uso óptimo de los recursos hídricos [4] *even in life loss, because of causing adverse impacts, such as flash floods, in urban and sometimes in rural areas. Thus, long-term forecast of such events is of great importance for the preparation of local authorities in order to confront and mitigate the adverse consequences. The objective of this study is to estimate the possibility of forecasting the maximum daily precipitation for the next coming year. For this reason, appropriate prognostic models, such as Artificial Neural Networks (ANNs).*

La predicción climática es el proceso de predecir la probable tendencia del desarrollo climático en el futuro en función de las leyes cambiantes del clima pasado. En los últimos años, se ha reconocido cada vez más el papel progresista de la predicción climática en la prevención y mitigación de desastres, con las necesidades del desarrollo social y económico que se están presentando en el mundo, la investigación sobre la predicción del clima debe mejorarse con urgencia [5].

A continuación, se estudian los pronósticos utilizando el modelo bayesiano junto a la teoría de puntos de cambio frente al método tradicional de Holt-Winters. El estudio utiliza un enfoque cuantitativo y compara los resultados del modelo de punto de cambio con el modelo de Holt-Winters utilizados para pronosticar el comportamiento de la precipitación durante un año (2018), como resultado se comparan los valores obtenidos en los pronósticos con los datos observados de la serie original.

Según lo descrito por [6] los algoritmos comúnmente para hacer estudios de pronósticos de variables meteorológicas son: Suavizado exponencial único (SES), Ingenio estacional (Snaive), ARIMA

estacional (SARIMA), Red neuronal de avance (FFNN), Imputación basada en deformación de tiempo dinámica (DTWBI), Series temporales estructurales bayesianas (BST). Hay otros métodos que son utilizados por las estaciones meteorológicas de los aeropuertos como la afirma [7] donde se utilizan datos de predicción numérica del tiempo (NWP) los cuales proporcionan una calidad aceptable de los pronósticos, o como lo mencionan [8] *histograms provide a summary of the data that is more informative than those provided by other aggregates such as the mean. Some fields where HTS are useful include economy, official statistics and environmental science. This article adapts the k-Nearest Neighbours (k-NN donde mediante la aplicación el algoritmo de vecinos más cercanos (KNN) tiene la capacidad de arrojar resultados prometedores en los pronósticos de variables meteorológicas, y de acuerdo a [9] el modelo de Holt-Winters es un método efectivo para la predicción y modelamiento de parámetros climáticos.*

El modelo de punto de cambio también es una herramienta que se ha utilizado para analizar datos en series de tiempo en áreas como el mercado bursátil [10], [11], datos geológicos [12]–[14], climatológicos [10]–[11], análisis de datos genéticos [17], y para predecir catástrofes como terremotos [18] entre otros.

Este artículo abarca el marco teórico para contextualizar sobre los métodos utilizados; la metodología empleada para el logro de los objetivos; el análisis y la interpretación de los resultados para identificar el comportamiento de los modelos calculando el pronóstico de la precipitación; y por último se presentan las conclusiones sintetizando los aspectos más relevantes como alcance y limitaciones de la investigación.

Quienes buscan un punto de cambio investigan principalmente un problema de dos maneras diferentes: (a) Detectar una falla que no afecta la media de las

observaciones, pero cambia su estructura de covarianza. Cuando sucede esto recomiendan considerar procedimientos que tengan en cuenta perturbaciones suaves o profundas, realizar una prueba de tipo no paramétrica como la prueba de Kolmogorov-Smirnov. Y (b) Detectar una falla y estimar los parámetros de cambio para una falla que ocurre al mismo tiempo en la media y la covarianza de un proceso auto regresivo de orden acotado. Se recomienda aplicar una prueba de razón de verosimilitud y hacer las normalizaciones apropiadas [19].

La inferencia estadística sobre los puntos de cambio tiene dos aspectos, el primero es detectar si hay algún cambio en la secuencia de variables aleatorias observadas. Y el segundo es estimar el número de cambios y sus ubicaciones correspondientes. Los métodos para la estimación de PC son principalmente la razón de verosimilitud, no paramétrico y bayesiano. Algunos autores también consideraron el problema del punto de cambio en otras configuraciones del modelo, como gamma y exponencial [17].

El modelo verosimilitud se usa para la ubicación del punto de cambio, así como un criterio de información desarrollado, para el caso de varianza conocida [20] g“Este modelo es una prueba de hipótesis que ayuda a elegir el “mejor” modelo entre dos modelos anidados. “Modelos anidados” significa que uno es un caso especial del otro. El mejor modelo es el que hace que los datos sean más probables, o maximiza la función de verosimilitud. Sin embargo aunque el concepto es relativamente fácil de comprender, los cálculos para encontrar las entradas para el procedimiento no lo son” [21]. Las pruebas de razón de verosimilitud usan funciones de logaritmo, que son difíciles y largas de calcular a mano, la mayoría de los paquetes de software estadísticos tienen funciones integradas para manejarlos [22].

Mientras que los métodos no paramétricos están diseñados sin una suposición paramétrica particular [23]. el método no

paramétrico es una herramienta flexible que mecaniza la construcción de modelos de predicción, seleccionando variables relevantes, transformando las variables predictoras, tratando valores perdidos y previniendo sobreajustes mediante un auto test. También permite predecir tomando en cuenta factores estructurales que pudieran tener influencia sobre la variable respuesta, generando modelos hipotéticos [24]. Los métodos no paramétricos han sido tradicionalmente recomendados para conjuntos de datos más pequeños y / o datos no distribuidos normalmente [25]. Sin embargo estos métodos, tienden a ser menos precisos en problemas de gran dimensión debido a la llamada efecto de la dimensionalidad [15].

Otro método de CP es el método Bayesiano (MB) para problemas de puntos de cambio, presentado por Barry y Hartigan en 1993. Con este método basado en modelos, se estima una partición desconocida de bloques para un conjunto de datos. Por lo tanto, la probabilidad de una partición se calcula utilizando la llamada cohesión [26]“Una secuencia de observaciones sufre cambios repentinos en tiempos desconocidos. Modelamos el proceso suponiendo que hay una secuencia subyacente de parámetros divididos en bloques contiguos de valores de parámetros iguales; el comienzo de cada bloque se dice que es un punto de cambio. Luego se asume que las observaciones son independientes en diferentes bloques dada la secuencia de parámetros. En un análisis bayesiano, es necesario dar distribuciones de probabilidad tanto a los puntos de cambio como a los parámetros” [12]

El método bayesiano se ha generalizado particularmente porque es útil para la solución de problemas en la toma de decisiones. La utilidad de este método consiste básicamente en el uso de situaciones en las que existe información limitada acerca de un gran número de variables o cuando la información proviene de diferentes fuentes. El método primordialmente incorpora

conocimiento previo para poder estimar modelos útiles dentro de un espacio maestro y de este modo poder estimar parámetros que provengan de la experiencia o de una teoría probabilística [27].

Los pronósticos son una herramienta que proporciona un estimado cuantitativo de la probabilidad de eventos futuros [28]. El objetivo de un pronóstico es posibilitar las decisiones sobre el futuro y proporcionar una estimación del riesgo involucrado en una decisión [29].

Los pronósticos se dividen en tres métodos estandarizados: cualitativos, de proyección histórica y causales. Los métodos cualitativos se enfocan en determinar cuáles son aquellos factores intangibles y subjetivos que afectan el pronóstico a partir del juicio, la intuición, las encuestas o técnicas comparativas para generar estimados cuantitativos acerca del futuro [30]. Los métodos de proyección histórica se desarrollan a partir de datos del pasado y sirven para predecir comportamientos futuros [31]. Y por último los métodos causales, por su parte, “asumen que el factor que va a ser pronosticado exhibe una relación causa-efecto con una o más variables independientes. El propósito de los modelos causales es describir la forma de relación entre las variables y usarla para predecir valores futuros de la variable dependiente. Dentro de los métodos causales más utilizados se encuentran las técnicas de regresión y las técnicas econométricas” [28].

Los métodos cuantitativos normalmente utilizados para ejecutar los pronósticos son: regresión lineal, medias móviles y suavizamiento exponencial entre otros. En cuanto al método de regresión lineal, es una herramienta importante y útil en muchos análisis estadísticos para estudiar la relación entre variables, el cual se usa principalmente para “predecir valores de la variable de respuesta a valores interesantes de las variables predictoras, descubrir los predictores asociados con la variable de

respuesta y estimar cómo los cambios en las variables predictoras afectan la variable de respuesta” [32].

Para esta investigación se utiliza el modelo de Holt-Winters el cual es una extensión del método de suavizamiento exponencial, es considerado una herramienta poderosa para las series tiempo [33]. Este método es un método de vital importancia para que haya una buena predicción, mediante este método se logra una confiabilidad alta para producir predicciones robustas y precisas dentro de los procesos de una empresa [34]. Finalmente en [35] se explica brevemente la metodología para ésta técnica de la siguiente manera, se debe asignar pesos exponencialmente decrecientes a medida que las observaciones son más distantes del estado presente, es decir, las observaciones recientes tienen un peso relativamente mayor en la predicción que las anteriores.

La exactitud de los modelos mencionados se puede determinar comparando los valores observados utilizados con los valores pronosticados, así mismo lo aseguran [36] “el error de pronóstico nos dice qué tan buen desempeño tiene el modelo al compararlo consigo mismo usando datos históricos”, por lo tanto analizar los errores de predicción es una tarea que permite una comparación de los modelos de pronóstico y la selección del modelo más adecuado [37]. Dentro de la literatura se encuentran varios procedimientos para calcular el error global de un pronóstico los cuales se utilizan para medir el desempeño, comparar varios modelos de pronóstico y hacer un monitoreo a los pronósticos. Los más comunes son la desviación media absoluta (MAD; mean absolute deviation), error porcentual (MAPE; mean absolute percent error) y el error medio cuadrático (MSE; mean squared error), entre otros [38].

La precipitación es un componente fundamental del ciclo global del agua y una variable hidrológica clave del ciclo del agua para la meteorología, la climatología y la

hidrología. Las observaciones precisas de la precipitación y sus distribuciones regionales y globales han sido durante mucho tiempo desafíos científicos, los pronósticos de la precipitación han madurado durante cuatro décadas de desarrollo [39]. Desde el punto de vista físico, la precipitación se define como “todas las partículas acuosas en fase líquida o sólida que se originan en la atmósfera y caen a la superficie de la Tierra” [40].

Es un componente fundamental del ciclo global del agua, la precipitación son los flujos de humedad de la atmósfera a la superficie de la tierra, y es una variable hidrológica clave del ciclo del agua para la meteorología, la climatología y la hidrología, la atmósfera obtiene aproximadamente las tres cuartas partes de su energía calorífica de la liberación de calor latente por la precipitación [41]. La precipitación varía de un año a otro y durante décadas, y los cambios en cantidad, intensidad, frecuencia y tipo afectan el medio ambiente y la sociedad. Las lluvias moderadas constantes penetran el suelo y benefician a las plantas, mientras que la misma cantidad de lluvia en un corto período de tiempo puede causar inundaciones, dejando los suelos mucho más secos [42].

Materiales y métodos

El procedimiento realizado para el cálculo del pronóstico incluye una serie de seis pasos consecutivos. Una vez depurados los datos de la precipitación para la serie de tiempo bajo estudio se procede a ejecutar el primer paso que consiste en graficar los datos con el fin de ilustrar su comportamiento en la ventana de tiempo; luego en el segundo paso de la metodología se identifican componentes de la serie tales como tendencia, estacionalidad y aleatoriedad; el tercer paso se dedica al cálculo del pronóstico mediante Holt-Winters; mientras que en el cuarto paso se calculan los puntos de cambio de la serie en las dimensiones de la media y de la varianza; el quinto paso de la metodología utiliza los resultados obtenidos en el paso

anterior para pronosticar la precipitación mediante el método del punto de cambio; finalmente en el paso sexto se comparan los pronósticos obtenidos por los dos métodos, aplicando métodos estadísticos inferenciales tales como pruebas de dependencia y el análisis de residuales. En el siguiente aparatado se desarrollan cada uno de estos pasos de la metodología empleada para esta investigación.

Los datos utilizados fueron suministrados por una estación meteorológica de Tunjuelito – Colombia, encargada de hacer monitoreo a la zona de estudio, lo cual garantiza tanto la rigurosidad científica y metodología para la recolección de los datos. Los datos de esta variable meteorológica de precipitación constituyen una serie de tiempo diaria recopilada desde el año 1989 hasta el 31 de diciembre de 2020. Con estos datos diarios históricos se realizan los pronósticos para el año 2021 considerando dos métodos de cálculo como son el Holt-Winters y el de punto de cambio.

Para el desarrollo de los pasos de la metodología de este estudio, se utilizó R el cual es “un lenguaje y entorno para computación estadística y gráficos. R proporciona una amplia variedad de técnicas estadísticas (modelos lineales y no lineales, pruebas estadísticas clásicas, análisis de series de tiempo, clasificación, agrupación, etc.) técnicas gráficas, y es altamente extensible. Una de las fortalezas de R es la facilidad con la que se pueden producir gráficos de calidad de publicación bien diseñados, incluyendo símbolos matemáticos y fórmulas donde sea necesario. “R está disponible como software libre bajo los términos de la Free Software Foundation General Public License (GNU) en forma de código fuente. Compila y se ejecuta en una amplia variedad de plataformas UNIX y sistemas similares” [43].

La precipitación de acuerdo a [44] se clasifica dependiendo de la intensidad, cuando es menor a 2 mm/hora es débil,

entre 2 mm/hora y 15 mm/hora es moderada, entre 15 mm/hora y 30 mm/hora es fuerte, muy fuerte si es entre 30 mm/hora y 60 mm/hora y cuando son mayores a 60 mm/hora se catalogan como torrenciales. Teniendo en cuenta lo anterior los datos de acuerdo los datos de esta investigación se agrupan como se ilustra en la figura 1, el 69.78% de los datos son precipitación débil, el 28.31% a precipitación moderada, el 1.76% y 0.14% a precipitación fuerte y muy fuerte respectivamente. Debido a que el modelo de punto de cambio a resultado ser útil para la predicción de variables del mercado de valores, hidrologías o de fenómenos naturales como en los estudios de estudios realizados por [16], [45], [46] y [47], uno de los aportes de este trabajo es verificar que el modelo sea robusto para el pronóstico de variables atmosféricas.

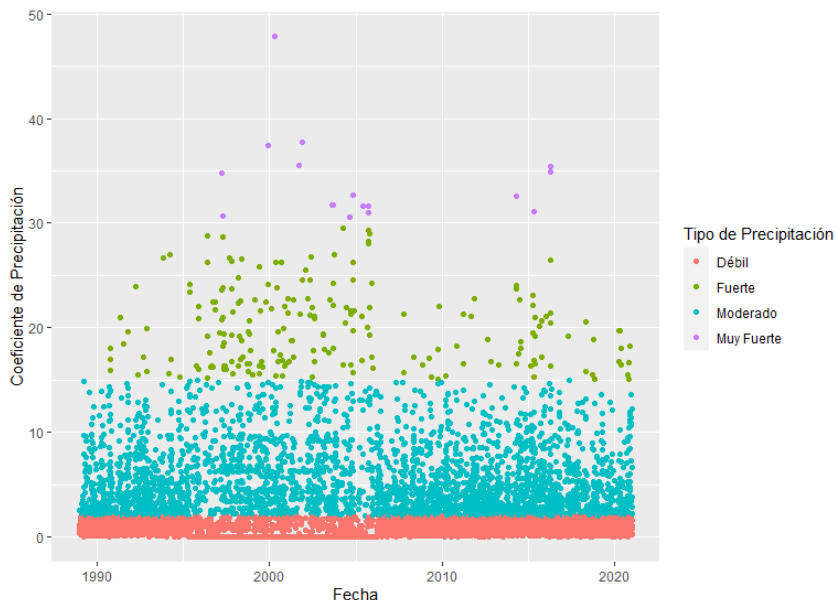


Figura 1. Agrupación de los datos según intensidad de la precipitación.

Resultados y discusión

Paso 1. Graficar la serie de tiempo. Graficar una serie de tiempo permitió visualizar tendencias de valores numéricos en función del tiempo, cada punto en el gráfico corresponde tanto a un tiempo como una cantidad que se está midiendo para este estudio la precipitación dado en milímetros (mm). En la Figura 1, se puede establecer que la mayor parte del tiempo donde la precipitación obtuvo valores altos es en la década de 1995 al año 2005, constatando la información brindada por la estación climatológica.

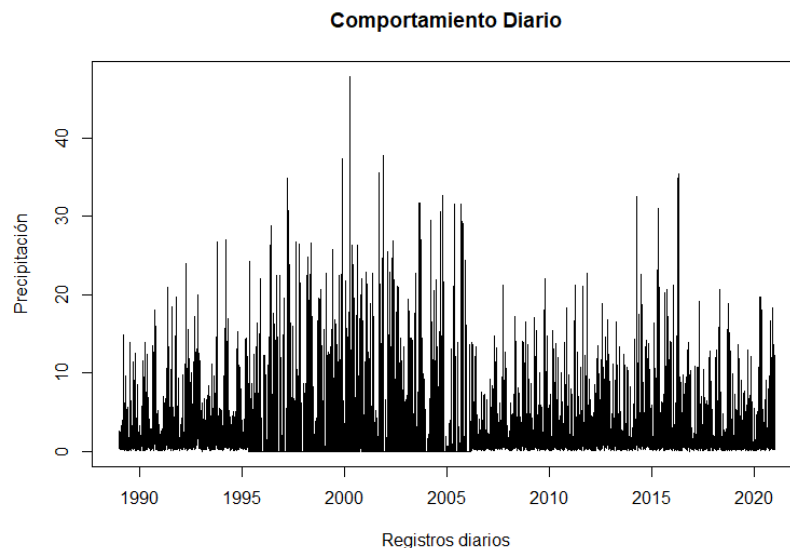


Figura 2. Serie de tiempo datos de precipitación desde enero de 1989 hasta diciembre de 2020.

El paso 2 de la metodología permite identificar los componentes de la serie, tales como tendencia, estacionalidad y aleatoriedad.

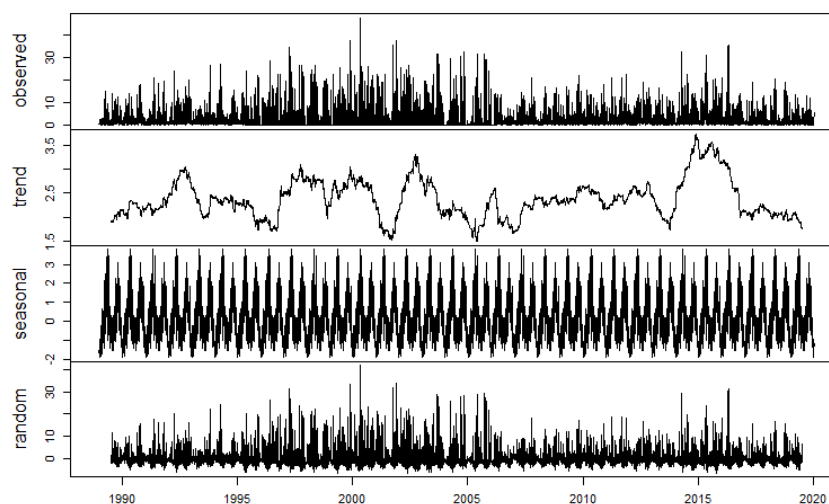


Figura 3. Componentes de la serie tiempo 1989 a 2020.

La serie presenta fluctuaciones movimientos que ocurren años tras año más o menos con la misma intensidad, este fenómeno permite afirmar que la serie de tiempo de la variable de precipitación de los datos recogidos por la estación Tunjuelito desde el año 1989 hasta 2020 es estacional (Figura 2). Se evidencia que la tendencia permite predecir los estados futuros de la serie a lo largo del tiempo donde hay tendencias versátiles (aumento, disminución y estabilidad del coeficiente de precipitación), el comportamiento de la precipitación sufre cambios que no son cíclicos es decir que su comportamiento es irregular y causando un alto grado de aleatoriedad.

En el paso 3. Se realiza el cálculo del pronóstico mediante el modelo Holt-Winters. En la Figura 3 se ilustra el pronóstico mediante el modelo de Holt-Winters dado que utiliza un promedio ponderado de los valores pasados y actuales (1989-2020) para proporcionar

predicciones (2021), para este cálculo se les dio más peso a los datos más recientes para aumentar la importancia de estos datos en comparación con los datos más antiguos.

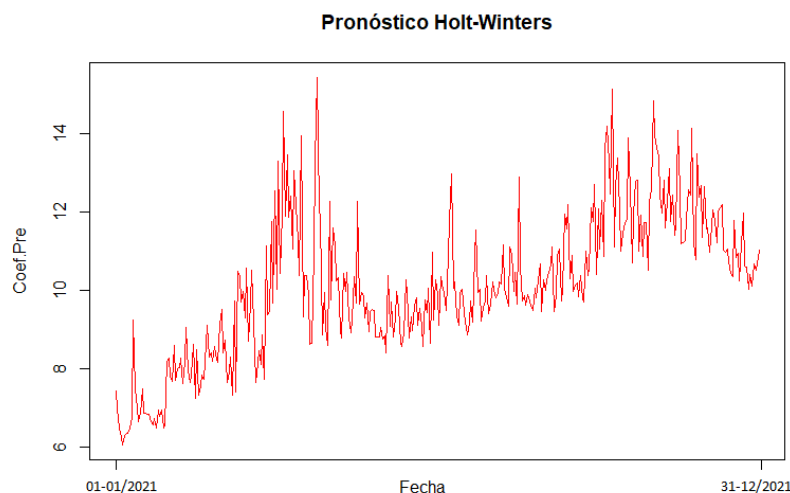


Figura 3. Resultados pronóstico para el año 2021 mediante el método Holt-Winters.

Para realizar el cálculo de los pronósticos mediante el modelo de punto de cambio, previamente se deben identificar y estimar los puntos de cambio tal como se propone en el paso 4 de la metodología establecida para esta investigación.

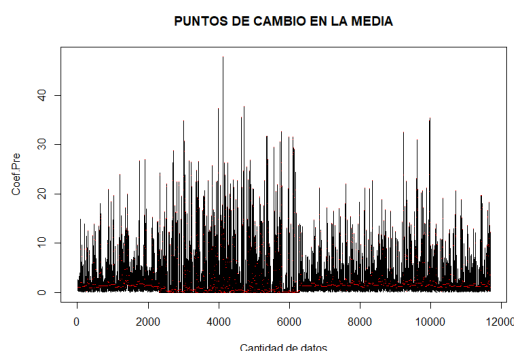


Figura 4. Estimación puntos de cambio en la media.

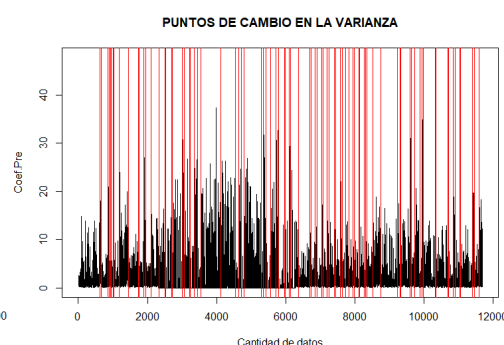


Figura 5. Estimación puntos de cambio en la Varianza.

La mayoría de puntos de cambio calculados en la media de la serie de tiempo original se reflejan en los días donde la variable precipitación tuvo valores menores a 10 mm (figura 4), mientras que los puntos de cambio situados en la varianza toman valores superiores a 10 mm (figura 5).

Una vez estimados los puntos de cambio se procede al paso 5 de la metodología, el cual consiste en calcular el pronóstico mediante el modelo de punto de cambio. El modelo de punto de cambio muestra una tendencia a la media, es decir, que considera los días sin lluvia para su proyección, pero en la serie final representa la tendencia, en la cual, no se visibilizan los días sin lluvia ni días de eventos extremos, los cuales, son de gran importancia para la toma de decisiones de los sectores que se pueden ver afectados o beneficiados con las fluctuaciones de la variable de estudio (Figura 6).

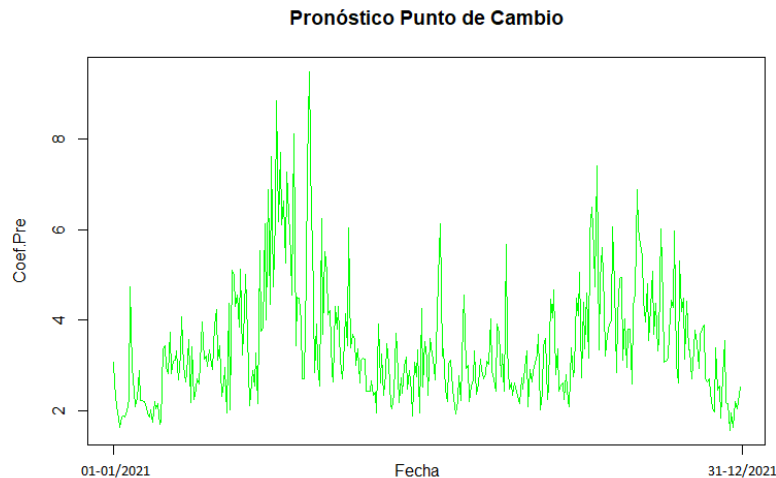


Figura 6. Resultados pronóstico para el año 2021 mediante el método de punto de cambio.

Finalmente, en el paso 6 se realiza una comparación de los resultados de pronóstico obtenidos con la aplicación de los dos modelos. Se graficaron las series con mayor significancia según la figura 7, en donde, se evidencia que las dos series representan la misma tendencia de régimen bimodal, pero se puede ver que el modelo de punto de cambio no representa los eventos extremos, es decir, no representa los días secos ni de alta precipitación, y modelo Holt-Winters toma los valores extremos, esta información puede ofrecer una solución viable y razonablemente precisa los estudios del comportamiento de la precipitación.

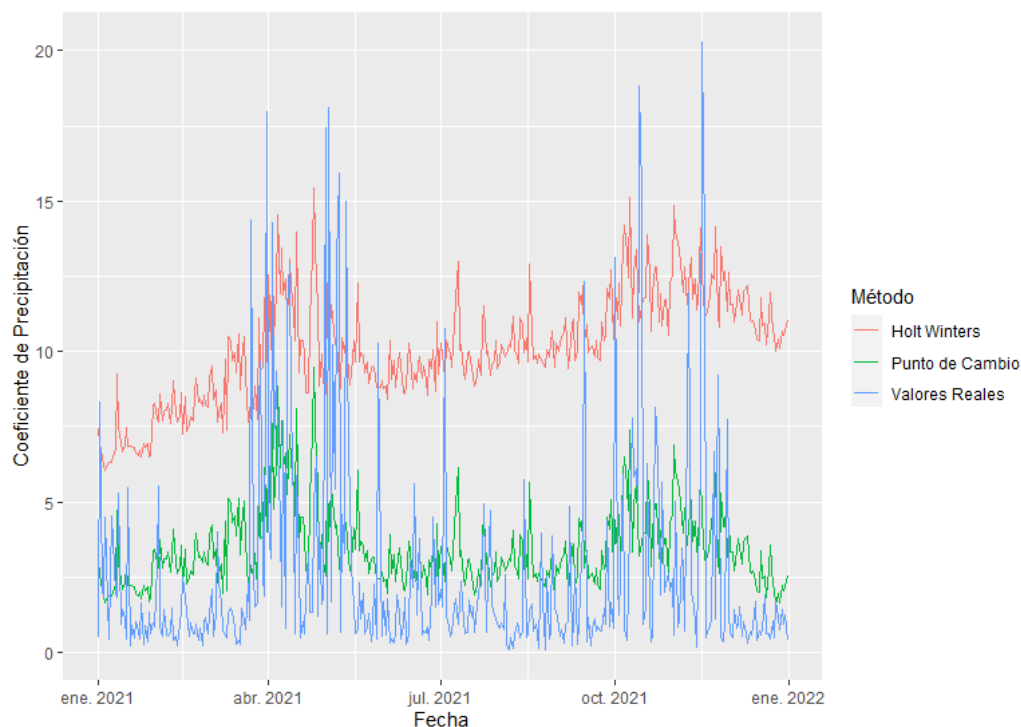


Figura 7. Comparación datos observados 2021 con datos pronosticados 2021 desde el primero de enero al 31 de diciembre.

Para obtener estos resultados de los pronósticos para el año 2021 se utilizaron los datos de la serie de tiempo hasta el 31-diciembre/2020 con el objetivo de analizar los datos calculados con los datos observados obtenidos para la estación Tunjuelito durante el año 2021. Para estimar la validez de los resultados arrojados mediante la aplicación del modelo Holt-

Winters y del modelo de punto de cambio se compararon con los datos observados mediante pruebas estadísticas para calcular el grado de similitud de los pronósticos con los datos reales como se observa en la Tabla I.

Tabla I. Pruebas de dependencia estadística de los datos observados con los datos pronosticados.

Prueba	Valor del estadístico	Confiabilidad	p-value	Planteamiento de las Hipótesis	Hipótesis que NO se acepta
Kruskal-Wallis	364	95%	0,4901	H_0 : Las medias del pronóstico HW son estadísticamente iguales a la de los datos observados.	H_1
				H_1 : Las medias del pronóstico HW estadísticamente no son iguales a los datos observados.	
Kruskal-Wallis	364	95%	0,4901	H_0 : Las medias del pronóstico PC estadísticamente son iguales a los datos observados.	H_1
				H_1 : Las medias del pronóstico PC estadísticamente no son iguales a los datos observados.	
Fisher	0,14	95%	2,2E-16	H_0 : es homogénea la varianza del pronóstico HW a la varianza de los datos observados.	H_0
				H_1 : No es homogénea la varianza del pronóstico HW a la varianza de los datos observados.	
Fisher	0,27	95%	2,2E-16	H_0 : es homogénea la varianza del pronóstico PC a la varianza de los datos observados.	H_0
				H_1 : No es homogénea la varianza del pronóstico PC a la varianza de los datos observados.	
Wilcoxon	51339	95%	2,2E-16	H_0 : Los datos del pronóstico HW son dependientes de los datos observados.	H_0
				H_1 : Los datos del pronóstico HW son independientes de los datos observados.	
Wilcoxon	6554	95%	2,2E-16	H_0 : Los datos del pronóstico PC son dependientes de los datos observados.	H_0
				H_1 : Los datos del pronóstico PC son independientes de los datos observados.	
T	4,443	95%	1,11E-05	H_0 : es homogénea la varianza del pronóstico HW a la varianza de los datos observados.	H_0
				H_1 : No es homogénea la varianza del pronóstico HW a la varianza de los datos observados.	
T	36,602	95%	2,20E-16	H_0 : es homogénea la varianza del pronóstico PC a la varianza de los datos observados.	H_0
				H_1 : No es homogénea la varianza del pronóstico PC a la varianza de los datos observados.	

H_0 : Hipótesis Nula HW: Holt-Winters

H_1 : Hipótesis Alterna PC: Punto de Cambio

De acuerdo a [40] la predicción de variables meteorológicas generalmente implica que

la probabilidad será más alta de lo normal. Tal afirmación requiere conocimiento del “comportamiento normal”, es decir, requiere una hipótesis nula. Las hipótesis calculadas en la Tabla 1 se comprobaron de tres maneras: comparando los datos observados con los pronosticados, comparar la independencia estadística de los datos observados con la distribución de los datos pronosticados, y comparar la razón de probabilidad con la formulación de las hipótesis nulas. Para el caso de la prueba de Kruskal-Wallis al aceptar la H_1 se demuestra que no hay diferencia significativa en la media de los datos observados con los pronosticados.

Con los resultados de la prueba de Kruskal-Wallis presentados en la Tabla 1, se puede comprobar que hay un alto grado de supeditación de los valores pronosticados con los valores observados, se puede afirmar que los datos obtenidos de los métodos Holt-Winters y Punto de cambio son estadísticamente independientes de la serie real a pesar de que provienen de la misma muestra. Además, con el ánimo de obtener un índice de comparación entre los pronósticos de los dos modelos, se calculó el Error Medio Absoluto (*MAE*), que es una medida robusta al considerar el alto volumen de datos utilizados para pronosticar, el MAE para los pronósticos obtenidos con el modelo de Holt Winters es de 7.92 mientras que el MAE para los resultados del modelo de punto de cambio fue de 2.56, lo que permite evidenciar que el pronóstico arrojado por este último modelo se ajusta mejor a los datos observados, corroborando lo ilustrado en la figura 7.

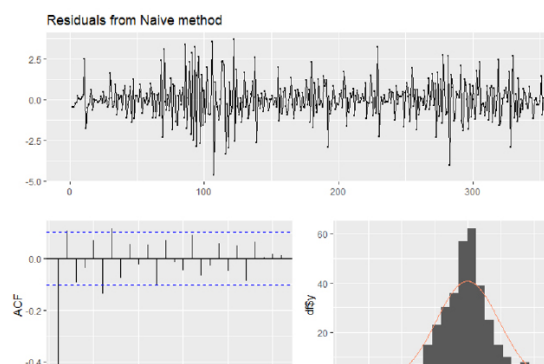


Figura 8. Residuales de datos obtenido mediante el método Holt-Winters.

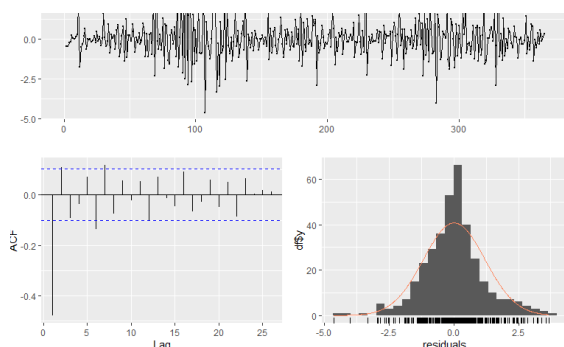


Figura 9. Residuales de datos obtenidos mediante el método Punto de Cambio.

Otra manera de constatar los resultados de las pruebas de la Tabla 1, es el análisis de residuales como se muestra en las figuras 8 y 9 estos gráficos muestran que el método Naive analiza que pronósticos parecen tener en cuenta toda la información

disponible. Los residuales fueron útiles para comprobar que la predicción de la precipitación para el año 2021 mediante los modelos de Holt-Winters y punto de cambio son un buen pronóstico, teniendo en cuenta que los resultados cumplen con las siguientes propiedades: la media de los residuos es cercana a cero y no existe una correlación significativa en la serie de residuos como se constata en la prueba de Kruskal-Wallis (Tabla 1). La gráfica de tiempo de los residuos muestra que la variación de los residuos se mantiene prácticamente igual a través de los datos históricos (1987-2020), aparte de los valores atípicos, y, por lo tanto, la variación residual puede tratarse como constante. Esto también se puede ver en el histograma de los residuos, el histograma indica que los residuos pueden no ser normales, considerando que la cola derecha de las Figuras 8 y 9 parecen demasiado larga para los dos casos, incluso cuando se ignoran los valores atípicos.

Conclusiones

Los resultados de los pronósticos obtenidos mediante la aplicación del modelo Holt-Winters se ajustan en menor proporción a los datos observados, sin embargo, de acuerdo al análisis de residuales se considera como un buen pronóstico porque se encontró que no hay ninguna relación entre los datos observados y los calculados.

Mediante la aplicación del modelo de punto de cambio para calcular el pronóstico de la precipitación del año 2021 el pronóstico obtenido muestra una tendencia a la media, es decir, que considera los días con precipitación débil para su proyección, pero en la serie final representa la tendencia, en la cual no se visibilizan los datos con precipitación débil ni datos con precipitación fuerte o muy fuerte.

Al comparar los pronósticos obtenidos con el modelo de Holt-Winters frente a los resultados del modelo de punto de cambio se puede comprobar que las dos series representan la misma tendencia de régimen bimodal, considerando que el estudio tiene un volumen de datos alto, un parámetro para concluir cuál de los pronósticos es mejor fue el cálculo del error medio absoluto (MAE), obteniendo un valor de 2.56 y 7.92 para el pronóstico de punto de cambio y Holt-Winters respectivamente, por lo tanto, se evidencia que el pronóstico calculado mediante la estimación de puntos de cambio se ajusta mejor a los datos observados.

El modelo de punto de cambio se ha utilizado para predecir el comportamiento de variables relacionadas con el mercado bursátil, la predicción de fenómenos naturales como terremotos y la predicción de comportamientos de cuencas hidrológicas, en esta investigación se busca generar interés en la comunidad científica para ampliar el espectro de uso de este modelo, iniciando con variables de tipo meteorológicas tales como la precipitación. Los resultados de esta investigación demuestran un uso potencial positivo del modelo de cambio para el pronóstico de variables atmosféricas.

En este estudio, los datos obtenidos del modelo de punto de cambio muestran una mayor precisión y se ajusta relativamente bien a los datos observados. Sin embargo, puede parecer que el procedimiento de optimización tiene un papel importante en la convergencia del modelo. Sin embargo, este estudio se considera preliminar y para que los resultados puedan ser considerados como concluyentes de deben aplicar a una cantidad significativa de series de tiempo de variables meteorológicas. Es recomendable probar también con la utilización de algoritmos enfocados en el pronóstico de variables climatológicas con el fin de verificar la obtención de mejores resultados.

Referencias

- [1] C. H. Sui, X. Li, and M. J. Yang, "On the definition of precipitation efficiency," *J. Atmos. Sci.*, vol. 64, no. 12, pp. 4506–4513, 2007, doi: 10.1175/2007JAS2332.1
- [2] A. Esquivel *et al.*, "Predictability of seasonal precipitation across major crop growing areas in Colombia," *Clim. Serv.*, vol. 12, no. March, pp. 36–47, 2018, doi: 10.1016/j.cliser.2018.09.001.
- [3] H. Feidas, C. Noullopoulou, T. Makrogiannis, and E. Bora-Senta, "Trend analysis of precipitation time series in Greece and their relationship with circulation using surface and satellite data: 1955-2001," *Theor. Appl. Climatol.*, vol. 87, no. 1–4, pp. 155–177, 2007, doi: 10.1007/s00704-006-0200-5.
- [4] P. T. Nastos, A. G. Paliatsos, K. V. Koukouletsos, I. K. Larissi, and K. P. Moustris, "Artificial neural networks modeling for forecasting the maximum daily total precipitation at Athens, Greece," *Atmos. Res.*, vol. 144, pp. 141–150, 2014, doi: 10.1016/j.atmosres.2013.11.013
- [5] T. Li, C. Qiao, L. Wang, J. Chen, and Y. Ren, "An Algorithm for Precipitation Correction in Flood Season Based on Dendritic Neural Network," *Front. Plant Sci.*, vol. 13, no. May, pp. 1–13, 2022, doi: 10.3389/fpls.2022.862558
- [6] T. T. H. Phan, É. Poisson Caillault, and A. Bigand, "Comparative study on univariate forecasting methods for meteorological time series," *Eur. Signal Process. Conf.*, vol. 2018-Septe, pp. 2380–2384, 2018, doi: 10.23919/EUSIPCO.2018.8553576
- [7] A. J. M. Jacobs and N. Maat, "Numerical guidance methods for decision support in aviation meteorological forecasting," *Weather Forecast.*, vol. 20, no. 1, pp. 82–100, 2005, doi: 10.1175/WAF-827.1
- [8] J. Arroyo and C. Maté, "Forecasting histogram time series with k-nearest neighbours methods," *Int. J. Forecast.*, vol. 25, no. 1, pp. 192–207, 2009, doi: 10.1016/j.ijforecast.2008.07.003.
- [9] M. Heydari, H. B. Ghadim, M. Rashidi, and M. Noori, "Application of holt-winters time series models for predicting climatic parameters (Case study: Robat Garah-Bil station, Iran)," *Polish J. Environ. Stud.*, vol. 29, no. 1, pp. 617–627, 2020, doi: 10.15244/pjoes/100496
- [10] M. V. Zhitlukhin and W. T. Ziemba, "Exit strategies in bubble-like markets using a changepoint model," *Quant. Financ. Lett.*, vol. 4, no. 1, pp. 47–52, 2016, doi: 10.1080/21649502.2015.1165918
- [11] M. J. Lenardon and A. Amirdjanova, "Interaction between stock indices via changepoint analysis," pp. 573–586, 2006, doi: 10.1002/asmb
- [12] D. Barry and J. A. Hartigan, "A Bayesian Analysis for Change Point Problems," *J. Am. Stat. Assoc.*, vol. 88, no. 421, p. 309, 1993, doi: 10.2307/2290726
- [13] R. Cmejla, J. Rusz, P. Bergl, and J. Vokral, "Bayesian changepoint detection for the automatic assessment of fluency and articulatory disorders," *Speech Commun.*, vol. 55, no. 1, pp. 178–189, 2013, doi: <https://doi.org/10.1016/j.specom.2012.08.003>
- [14] C. Jeong and J. Kim, "Bayesian multiple structural change-points estimation in time series models with

- genetic algorithm,” *J. Korean Stat. Soc.*, vol. 42, no. 4, pp. 459–468, Dec. 2013, doi: 10.1016/j.jkss.2013.02.001
- [15] S. Liu, M. Yamada, N. Collier, and M. Sugiyama, “Change-point detection in time-series data by relative density-ratio estimation,” *Neural Netw.*, vol. 43, pp. 72–83, Jul. 2013, doi: 10.1016/j.neunet.2013.01.012
- [16] S. Li and R. Lund, “Multiple Changepoint Detection via Genetic Algorithms,” *J. Clim.*, vol. 25, no. 2, pp. 674–686, Jan. 2012, doi: 10.1175/2011JCLI4055.1
- [17] J. Chen and K. Gupta, *Parametric Statistical Change Point Analysis*, 2nd ed. Boston: Birkhäuser, Springer, 2012
- [18] F. A. Alawadhi and D. Alhulail, “Bayesian change points analysis for earthquakes body wave magnitude,” *J. Appl. Stat.*, vol. 43, no. 9, pp. 1567–1582, 2016, doi: 10.1080/02664763.2015.1117585
- [19] D. Picard, “Testing and estimating change-points in time series,” vol. 17, no. 4, pp. 841–867, 2013
- [20] P. J. Plummer and B.S.E., “Decting Change-Points in a compound Poisson,” 2012
- [21] E. L. Lehmann, “On likelihood ratio tests,” *Optimality*, vol. 49, no. 2006, pp. 1–8, 2006, doi: 10.1214/074921706000000356
- [22] T. Wang, W. Tian, and W. Ning, “Likelihood ratio test change-point detection in the skew slash distribution,” *Commun. Stat. Simul. Comput.*, vol. 0, no. 0, pp. 1–13, 2020, doi: 10.1080/03610918.2020.1755869
- [23] Y. Zhou, L. Fu, and B. Zhang, “Two non parametric methods for change-point detection in distribution,” *Commun. Stat. - Theory Methods*, vol. 46, no. 6, pp. 2801–2815, 2017, doi: 10.1080/03610926.2015.1048891
- [24] J. Vanegas and F. Vásquez, “[Multivariate Adaptive Regression Splines (MARS), an alternative for the analysis of time series],” *Gac. Sanit.*, no. xx, pp. 4–6, Dec. 2016, doi: 10.1016/j.gaceta.2016.10.003.
- [25] V. Grech and N. Calleja, “WASP (Write a Scientific Paper): Parametric vs. non-parametric tests,” *Early Hum. Dev.*, no. xxxx, pp. 1–2, 2018, doi: 10.1016/j.earlhumdev.2018.04.014
- [26] S. Herrmann, H. Schwender, K. Ickstadt, and P. Müller, “A Bayesian changepoint analysis of ChIP-Seq data of Lamin B,” *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1844, no. 1 PART A, pp. 138–144, 2014, doi: 10.1016/j.bbapap.2013.09.001
- [27] L. O. Mesa, M. Rivera, and J. A. Romero, “Descripción general de la Inferencia Bayesiana y sus aplicaciones en los procesos de gestión,” *La Simulación al Serv. la Acad.*, vol. 2, pp. 1–28, 2011, doi: ISSN2027-7709
- [28] A. Contreras Juárez, C. Atziry Zuñiga, J. L. Martínez Flores, and D. Sánchez Partida, “Análisis de series de tiempo en el pronóstico de la demanda de almacenamiento de productos perecederos,” *Estud. Gerenciales*, vol. 32, no. 141, pp. 387–396, Oct. 2016, doi: 10.1016/j.estger.2016.11.002
- [29] G. Masterton, “What to do with a forecast?,” *Synthese*, vol. 191, no. 8, pp. 1881–1907, 2014, doi: 10.1007/s11229-013-0384-z
- [30] R. Ballou, *Logística: Administración de la cadena de suministro*, 5ta Ed.

- México, 2004.
- [31] M. Mas-Machuca, M. Sainz, and C. Martínez-Costa, "A review of forecasting models for new products," *Intang. Cap.*, vol. 10, no. 1, pp. 1–25, 2014, doi: 10.3926/ic.482
- [32] D. J. Eck, "Bootstrapping for multivariate linear regression models," *Stat. Probab. Lett.*, vol. 134, pp. 141–149, 2018, doi: 10.1016/j.spl.2017.11.001
- [33] L. Ferbar Tratar, B. Mojšker, and A. Toman, "Demand forecasting with four-parameter exponential smoothing," *Int. J. Prod. Econ.*, vol. 181, pp. 162–173, 2016, doi: 10.1016/j.ijpe.2016.08.004
- [34] J. W. Taylor, "Multi-item sales forecasting with total and split exponential smoothing," *J. Oper. Res. Soc.*, vol. 62, no. 3, pp. 555–563, 2011, doi: 10.1057/jors.2010.95
- [35] E. M. de Oliveira and F. L. Cyrino Oliveira, "Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods," *Energy*, vol. 144, pp. 776–788, 2018, doi: 10.1016/j.energy.2017.12.049
- [36] B. Heizer, J., & Render, *Principios de administración de operaciones*. Pearson Educación., 7ma Ed. Pearson, 2004
- [37] L. Frías-Paredes, F. Mallor, M. Gastón-Romeo, and T. León, "Dynamic mean absolute error as new measure for assessing forecasting errors," *Energy Convers. Manag.*, vol. 162, no. December 2017, pp. 176–188, 2018, doi: 10.1016/j.enconman.2018.02.030
- [38] M. V. Shcherbakov, A. Brebels, N. L. Shcherbakova, A. P. Tyukov, T. A. Janovsky, and V. A. evich Kamaev, "A survey of forecast error measures," *World Appl. Sci. J.*, vol. 24, no. 24, pp. 171–176, 2013, doi: 10.5829/idosi.wasj.2013.24.itmies.80032
- [39] "Precipitation," *Adv. Remote Sens.*, pp. 621–647, Jan. 2020, doi: 10.1016/B978-0-12-815826-5.00016-7
- [40] S. Michaelides, V. Levizzani, E. Anagnostou, P. Bauer, T. Kasparis, and J. E. Lane, "Precipitation: Measurement, remote sensing, climatology and modeling," *Atmos. Res.*, vol. 94, no. 4, pp. 512–533, 2009, doi: 10.1016/j.atmosres.2009.08.017
- [41] C. Kummerow, W. Barnes, T. Kozu, J. Shiue, and J. Simpson, "The Tropical Rainfall Measuring Mission (TRMM) sensor package," *J. Atmos. Ocean. Technol.*, vol. 15, no. 3, pp. 809–817, 1998