Elaboración y comparación del pronóstico de la temperatura mínima mediante un modelo ARIMA y redes neuronales de un cultivo de rosa de la sabana de Bogotá.

Development and comparison of the minimum temperature forecast using an ARIMA model and neural networks for a rose crop in the Bogotá savannah.

Duarte Sandoval C. (2022)[1]

RESUMEN

El sector floricultor es muy importante en la sabana de Bogotá gracias a la alta productividad y competencia a nivel económico en el país. Al ser uno de los sectores agrícolas más desarrollados, requiere de tecnologías que ayuden a la mejora continua del proceso. Cambios negativos en el clima contribuyen a la afectación en la productividad, por lo que es de gran importancia conocer estos posibles cambios en variables climáticas como lo son la precipitación, humedad y temperatura. Se pretende obtener un buen modelo con el que se pueda pronosticar cambios en las temperaturas mínimas para el correcto desarrollo fisiológico de la rosa. El pronóstico fue realizado con una base de datos de una compañía floricultora de la Sabana de Bogotá y consistió en la obtención de un pronóstico de temperaturas mínimas para el mes de febrero del año 2022. Se obtuvieron tres pronósticos, uno con un modelo ARIMA, uno con redes neuronales de retroalimentación y otro con redes neuronales recurrentes, los cuales se compararon con los valores reales. Los resultados indican menor error en el modelo ARIMA pero mayor variación en las temperaturas mínimas con las redes neuronales recurrentes, por lo que se considera necesario elaborar un modelo de series multivariadas en el que se incluyan variables como temperatura máxima y humedad relativa con para que el modelo sea más acertado a los datos reales.

PALABRAS CLAVE: Pronóstico, clima, rosa, temperatura mínima.

ABSTRACT

The floriculture sector is very important in the Bogotá´s Savannah due to the high productivity and economic competition in the country. As one of the most developed agricultural sectors, it requires technologies that help to continuously improve the process. Negative changes in the climate contribute to the affectation in productivity, so it is of great importance to know these possible changes in climatic variables such as precipitation, humidity, and temperature. The aim is to obtain a good model with which to forecast changes in minimum temperatures for the correct physiological development of the rose. The forecast was carried out with a database of a floricultural company in the Bogotá´s Savannah and consisted of obtaining a forecast of minimum temperatures for the month of February 2022. Three forecasts were obtained, one with an ARIMA model, one with feedback neural networks and one with recurrent neural networks, which were compared with the real values. The results indicate lower error in the ARIMA model but greater variation in the minimum temperatures with the recurrent neural networks, so it is considered necessary to develop a multivariate series model in which variables such as maximum temperature and relative humidity are included so that the model is more accurate to the real data. KEYWORDS: Forecast, weather, rose, minimum temperatura.

[1] Estudiante de la especialización en estadística aplicada. Fundación universitaria Los Libertadores. E-mail: cduartes@libertadores.edu.co

REFERENTES TEÓRICOS

El municipio del Rosal se encuentra ubicado en la sabana de occidente, su principal actividad económica es la agrícola, donde el 69% del suelo es utilizado para la siembra de diferentes cultivos y el clima es principalmente de tipo bosque montañoso, con temperatura promedio de 13 grados que puede descender hasta -1 grados (Alcaldía del Rosal, 2018).

Según Asocolflores, para este 2022 únicamente en la temporada de San Valentín se exportaron cerca de 650 millones de tallos. Así mismo, el sector floricultor se encuentra en el segundo lugar de exportaciones a nivel nacional y Colombia se encuentra en el segundo puesto como exportador de flores a nivel mundial, siendo el departamento de Cundinamarca el lugar con la mayor concentración de la exportación de flores (66%) (Ministerio de Agricultura y Desarrollo Rural, 2020).

El IDEAM define una helada como la ocurrencia de una temperatura igual o menor a 0°C a un nivel de 2 metros sobre el nivel del suelo, en lugares cuyas alturas sean desde los 2500 m.s.n.m en adelante; temperatura en la que los tejidos de una planta pueden llegar a sufrir daños irreversibles. Estos eventos ocurren con mayor frecuencia en las temporadas secas de diciembre a enero de cada año. Existen diferentes tipos de helada, como heladas por advección, por evaporación, radiación, siendo esta última típica de regiones tropicales y se caracteriza principalmente por la pérdida de temperatura que tienen las plantas en horas de la noche al ceder ésta misma al ambiente por medio de la radiación. Según los efectos que producen las heladas en las plantas, existen las heladas blancas determinada porque las gotas de rocío se congelan sobre las hojas por la alta humedad, y las heladas negras en las que no hay formación de hielo debido a la baja humedad pero que si ocasiona quemadura en el tejido de la planta. Como consecuencia, se producen cambios fisiológicos en las plantas como la respiración, actividad fotosintética, absorción de agua, destrucción de tejidos y la muerte celular (IDEAM, 2012).

Para el caso de la rosa, el tema de las diferentes variables climáticas, como temperatura (indispensable para el tiempo de floración), humedad relativa (que tiene gran importancia en la sanidad de los cultivos), radiación (contribuye a la temperatura bajo invernadero) y precipitación (disponibilidad de agua), son de gran importancia para una producción de calidad. En el caso de las heladas, pese a que los cultivos se encuentran bajo invernadero para la protección, la diferencia entre la temperatura ambiente en horas de la noche y la temperatura interna no es muy marcada (Monroy, Pérez, & Cure, 2001).

Con el fin de obtener pronósticos para las diferentes variables climatológicas, se han realizado varios modelos, como Villazón Bustillos, y otros (2016) que por medio de series de tiempo elaboraron un pronóstico de sequía para la región de Chihuahua, México. Los autores muestran la implementación de dos modelos, el primero es un modelo autorregresivo integrado de medias móviles (ARIMA) y el segundo consiste en modelos no lineales autorregresivos con entrada exógena (NARX), donde el primero emplea series de tiempo y el segundo redes neuronales. Por otra parte, Vanegas Ramos, (2011) implementó modelos locales en el espacio de fase tomando inicialmente 1582 series de tiempo de estaciones cercanas, reduciéndolas a 2 y evaluando 8 variables hidrometereológicas, obteniendo un modelo de regresión lineal con resultados aceptables pero que se ajustan únicamente para dichas estaciones; en este modelo también se resalta que valores no tomados de eventos extremos como temperaturas muy bajas o precipitaciones muy altas no se pueden predecir.

Para la predicción de precipitaciones en la sabana de Bogotá Ramírez Forero & Malagón Márquez, (2018) emplearon el modelo de la interpolación determinística espacio temporal y también un modelo ARIMA, pero en ambos métodos se empleó la función multicuadrática inversa. En esta investigación cabe resaltar que a pesar de que el modelo ARIMA tuvo menor error en la predicción,

éste emplea mayor tiempo en el procesamiento de la información.

Las series de tiempo son conjuntos de datos que se observan durante un periodo de tiempo, se puede definir como un proceso estocástico, es decir un conjunto de variables aleatorias clasificadas según un determinado parámetro. Estas variables pueden ser aleatorias, discretas o continuas y la variable tiempo puede tomar valores reales o discretos. (Reyes Polanco, 2021).

En los modelos de series de tiempo, es importante que las series no tengan datos faltantes, para rellenar esta información Alfaro & Soley (2009), emplean dos métodos para la obtención de datos faltantes en estaciones meteorológicas, donde la primera consiste en la obtención de datos de estaciones cercanas y mediante el análisis multivariado de componentes principales se complementa la información faltante; el segundo método se emplea cuando no hay información de estaciones cercanas y consiste en emplear modelos predictivos autorregresivos AR(p) con la información existente de la estación meteorológica en estudio.

El análisis de componentes principales toma una cantidad de las variables originales y las combina de forma lineal con las nuevas combinaciones (Guisande Golzalez, y otros, 2006). Al realizar una serie de combinaciones lineales de las variables, el ACP analiza la estructura de las matrices de varianzas y covarianzas, así como la matriz de correlación. Para determinar el número de componentes principales a utilizar se debe definir si se utilizará la matriz de varianzas y covarianzas o la matriz de correlaciones. En el primer caso, hay dos métodos, el primero consiste en dejar solo los componentes cuya raíces sean diferentes de cero, el segundo es emplear el porcentaje de varianza total y retener los valores menores a p componentes; en el segundo caso, hay dos opciones a tener en cuenta si se emplea la matriz de correlación, primero se retienen las componentes cuyas raíces sean mayores a 1 y segundo, por una prueba gráfica en la que se observará una curva descendente y una vez ésta se empiece a volver recta, se visualizará el número de componentes principales a usar (Hernández Rodríguez, 1998).

Alfaro & Soley (2009), sostienen en el análisis de componentes principales que las matrices de covarianzas se utilizan cuando se desea conservar la diferencia de la varianza de las estaciones climatológicas, mientras que si lo importante es analizar la forma de las curvas de las estaciones es mejor emplear la matriz de correlaciones. En ambos casos, entre mayores sean tanto los valores de covarianza como los de correlación, mayor será la relación entre cada estación estudiada.

Para el pronóstico de series de tiempo existen diferentes métodos, como la suavización exponencial lineal donde se toma la serie de tiempo y se emplea un "mecanismo de autocorrección que busca ajustar los pronósticos en dirección opuesta a las desviaciones del pasado mediante una corrección que se ve afectada por un coeficiente de suavización", aquí se debe tener en cuenta que solo se obtiene del modelo la predicción del último periodo (Salazar López, 2019), la suavización exponencial de Holt, que también emplea una ecuación lineal pero necesita de dos constantes de suavización; la suavización exponencial de Winter la cual se emplea cuando los datos son estacionales y usa tres constantes de suavización, los modelos autorregresivos integrados y de promedio móvil (ARIMA) que toma valores de la serie de tiempo y se usan como valores independientes en modelos de regresión (Toledo Vega, 2015). Los modelos ARIMA autorregresivos reflejan la información de valores de periodos del pasado (Aguirre Jaime, 1994). La ecuación de los modelos autorregresivos AR(p) está dada por:

$$y_t = \delta_0 + \theta_1 y_{t-1} + \delta_2 y_{t-2} + \cdots \delta_p y_{t-p} + \omega_t$$

Donde ωt es el error o ruido blanco y debe cumplir con 3 condiciones: media 0, varianza constante y covarianza nula.

Un modelo de media móvil muestra el valor de una variable a lo largo de un periodo de tiempo y en sucesión de errores en dicho tiempo, la ecuación es:

$$y_t = \mu + \omega_t + \omega_1 y_{t-1} + \omega_2 y_{t-2} + \cdots \omega_q y_{t-q}$$

Fernández indica que "un modelo ARIMA permite describir un valor como una función lineal de datos anteriores y errores debidos al azar, además, puede incluir un componente cíclico o estacional. Es decir, debe contener todos los elementos necesarios para describir el fenómeno". El método Box y Jenkins permite evaluar cada modelo por medio de una autocorrelación parcial (Fournies Pallavicini, 2015).

Para emplear los modelos ARIMA es importante que los procesos estocásticos sean estacionarios, es decir que no pueden tener una tendencia creciente o decreciente y en caso de que eso no ocurra, se debe hacer la transformación de la serie. Para determinar si una serie es o no estacionaria, generalmente basta con observar una gráfica de los datos (de Arce & Mahia, Año no especificado).

El modelo lineal emplea tres representaciones, si se conoce el modelo se puede emplear la representación de puramente medias móviles, puramente autorregresiva o finita. Cuando del modelo es lineal y depende de la parte autorregresiva y de la parte de medias móviles, recibe el nombre de ARMA (p,q). (González Casimiro, Año no especificado). A los modelos ARMA se les aplica una diferenciación estacional previamente, a diferencia de un modelo ARIMA, un a un modelo ARMA se le realiza un ajuste y una estimación sobre la serie original, mientras que el ARMA lo hace en la serie ya diferenciada (Aguirre Jaime, 1994).

La ecuación de un modelo ARMA es la sumatoria de un modelo AR(p) y un modelo MA(q):

$$\begin{aligned} y_t = \mu + \omega_t + \omega_1 y_{t-1} + \omega_2 y_{t-2} + \cdots \omega_q y_{t-q} \\ + \delta_0 + \theta_1 y_{t-1} + \delta_2 y_{t-2} \\ + \cdots \delta_p y_{t-p} \end{aligned}$$

METODOLOGÍA

Inicialmente se tomó una base de datos de la empresa de flores con temperaturas mínimas desde enero de 2019 hasta enero de 2022, se verificó que no hubiera datos faltantes y se procedió a obtener el pronóstico del mes de febrero de 2022 mediante un modelo ARIMA y mediante el modelo de redes neuronales de retroalimentación nntear y el modelo de redes neuronales recurrentes Elman. Luego de hacer un análisis descriptivo, para el modelo ARIMA se empleó la metodología de Box y Jenkins, en el cual se inició con:

- **Identificación del posible modelo:** Se determinó si la serie era o no estacionaria mediante las pruebas de dickey Fuller y Phillips Perron.

- **Selección de la transformación:** Para convertir la serie en estacionaria, en este caso fue necesario aplicar únicamente una diferencia ordinaria.

- **Determinación modelo ARIMA:** Mediante las gráficas ACF y PACF, se propuso 23 posibles modelos.

- **Estimación:** Se aplicó las pruebas de autocorrelación de Liun Box, de aleatoriedad de Run test y de normalidad de Jarque Bera, así como la obtención del BIC.

- **Diagnóstico:** Luego de aplicar las pruebas se seleccionó el modelo con menor BIC y que cumpliera con las pruebas de autocorrelación y de aleatoriedad.

- **Predicción:** Se procedió a obtener el pronóstico con el modelo seleccionado, es decir el ARIMA (5,1,1).

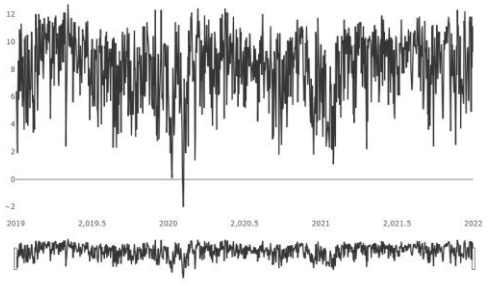
Para el caso de las redes neuronales, en el modelo de retroalimentación nntear, en el que se hace la red con una capa oculta y las entradas con un retraso, se obtienen los residuales y el modelo. Finalmente se hace un ajuste con los valores históricos.

Por último, en el modelo de redes neuronales recurrentes Elman, se tomó la serie, se normalizó y se dividieron los datos en un conjunto de

entrenamiento (80%) y un conjunto de prueba (20%). Luego se creó un dataframe para adelantar los valores de la serie en un futuro de 28 días, se eliminaron los NA y se determinaron los valores de entrada y salida de la red. La red creada fue de tres capas ocultas con una de 7 neuronas y 2 de 1, el ritmo de aprendizaje fue de 0.1 y un máximo de 15000 iteraciones. Luego se obtuvo el pronóstico y se desnormalizaron los datos. Finalmente se calculó el error de los tres pronósticos.

ANALISIS DE RESULTADOS

Al crear la serie de tiempo con los valores de temperaturas mínimas obtenidos en la finca, (Gráfica 1) se obtiene una media de 8,13 °C, siendo el valor mínimo de -2°C y el máximo de 12,7°. Al elaborar un gráfico de la temperatura mínima con cada año de la serie de tiempo, (Gráfica 2) se observa mayor variación en las temperaturas mínimas en los meses de diciembre, enero, febrero, agosto y septiembre y los meses de mayo, junio y julio son los que menos variabilidad en la temperatura tienen, información que concuerda de acuerdo a la información proporcionada por el IDEAM, la cual indica que los meses donde más probabilidad hay de heladas son los meses secos, especialmente de diciembre a febrero

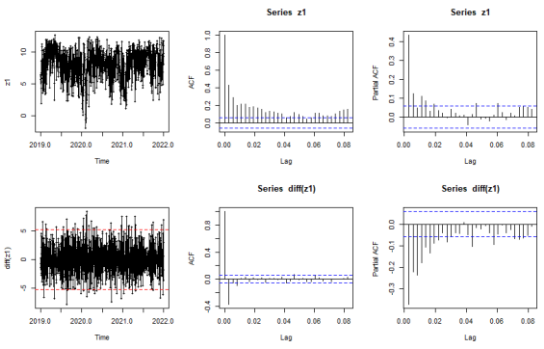


Gráfica 1: Serie de tiempo con datos de temperaturas mínimas de enero 1 de 2019 a enero 31 de 2022, con frecuencia diaria para un total de 1096 datos.

Gráfica 2: Temperaturas mínimas para cada año de la serie de tiempo.

Modelo ARIMA

Como primera parte de la metodología se identifica el modelo. Al observar la gráfica 3, se deduce que los datos no tienen una media ni varianza constante, por lo que se considera que los datos no son estacionarios. Adicionalmente, al observar la gráfica ACF se ve que los valores tienden a decrecer muy lentamente por lo que se considera que no es estacionaria. Ya que se debe cumplir que los datos sean estacionarios, es decir de media y varianza constante se realiza una diferencia ordianaria para convertir la serie en estacionaria., por lo que se hace la prueba de dickey Fuller y Phillips perron para evaluar estacionariedad (Tabla 1).



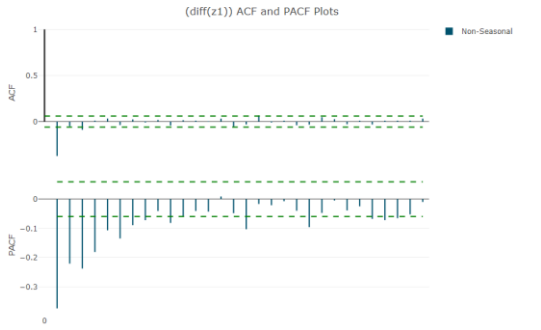
Gráfica 3: Gráficas ACF y PACF para la serie original y con una diferencia ordinaria.

Si se hace una diferencia a la serie original al observar la gráfica 3 pareciera que esta se volviera estacionaria, pues parece que se vuelve la media más constante y los valores en el ACF decrecen rápidamente. A la serie original se le aplican las pruebas de estacionariedad de Dickey- Fuller y de Phillips Perron y ambas indican que la serie es

estacionaria. Sin embargo, al observar la gráfica de descomposición, se nota una tendencia, por lo que se aplica nuevamente la prueba de Dickey – Fuller con el paquete URCA y el resultado del estadístico de prueba es 0,9683, el cual cae fuera de la región de rechazo por lo que se aprueba la hipótesis nula y se determina que la serie original es no estacionaria (Tabla 1). Luego de hacer una diferencia con Dickey- Fuller y Phillips Perron, ambas indican que la serie es estacionaria, y al aplicar Dickey – Fuller con el paquete URCA en esta ocasión, el valor del estadístico (-17.9855) cae en la región de rechazo, por lo que la serie con una diferencia ordinaria es estacionaria.

		Dickey Fuller	**Phillips Perron**	**Dickey Fuller URCA**
Serie sin diferenciar		0,01	0,01	-0,97
Serie diferenciada		0,01	0,01	-17,99

Tabla 1: valores del estadístico de prueba de las pruebas de estacionariedad para la serie sin y con diferencia.



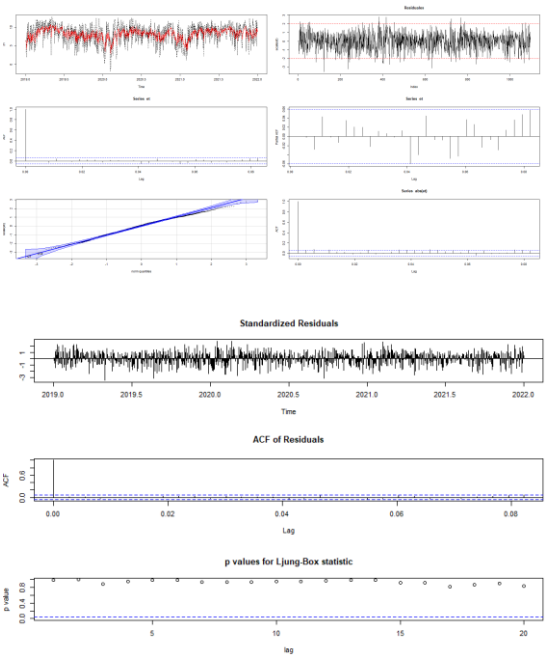
Gráfica 4: valores de ACF y PACF

AL observar el gráfico ACF y PACF, los modelos propuestos se presentan en la Tabla 2.

MODELO	BIC	TEST DE AUTOCORRELACIÓN LIUNG BOX	TEST DE ALEATORIEDAD RUN TEST	TEST DE NORMALIDAD JARQUE BERA	OBSERVACIONES
ARIMA(7,1,0)	4905,33	0.01829	0.04452	0.7383	
ARIMA(6,1,0)	4906,91	0.006278	0.5219	0.04636	
ARIMA(5,1,0)	4919,74	0.0007622	0.5994	0.1261	
ARIMA(4,1,0)	4925,22	0.0002567	0.8491	0.09774	
ARIMA(3,1,0)	4954,63	0,00	0.8195	0.1391	
ARIMA(2,1,0)	5010,93	0,00	0.3924	0.6411	
ARIMA(1,1,0)	5058,31	2.2e-16		0.6059	
ARIMA(0,1,1)	4887,80	0,00	0.001505	0.005666	
ARIMA(0,1,3)	4837,18	0.2301	0.3688	0.001439	
ARIMA(7,1,1)	4840,45	0.5149	0.9407	0.0004847	LAGS AR 1,2,5 Y MA1
ARIMA(6,1,1)	4840,45	0.515	0.9408	0.000485	LAGS AR 1,2,5 Y MA1
ARIMA(5,1,1)	4840,45	0.515	0.9407	0.000485	LAGS AR 1,2,5 Y MA1
ARIMA(4,1,1)	4841,10	0.3751	0.3387	0.0003088	LAGS AR 1,2,4 Y MA1
ARIMA(3,1,1)	4833,35	0.1536	0.6735	0.001345	LAG 1 Y MA 1
ARIMA(2,1,1)	4833,35	0.1536	0.6735	0.001345	LAG 1 Y MA1
ARIMA(1,1,1)	4834,35	0.1536	0.6735	0.001345	
ARIMA(7,1,3)	4861,49	0.4409	0.224	0.0004727	LAGS AR 1,2,3,4 Y MA1
ARIMA(6,1,3)	4861,49	0.4409	0.224	0.0004727	LAGS AR 1,2,3,4 Y MA1
ARIMA(5,1,3)	4854,26	0.5215	0.4322	0,00002	LAGS AR 1,2,3,5 Y MA1
ARIMA(4,1,3)	4859,33	0.2659	0.667	0.0005601	
ARIMA(3,1,3)	4841,34	0.3896	0.9548	0.0008627	LAGS AR 1,2,3 Y MA 3
ARIMA(2,1,3)	4887,80	0,00	0.001505	0.005665	LAG MA1
ARIMA(1,1,3)	4887802,00	0,00	0.001505	0.005665	LAG MA1

Tabla 2: Resultados de la estimación y diagnóstico del modelo ARIMA.

Inicialmente se seleccionan los modelos de menor BIC y que cumplieran con las pruebas de autocorrelación y de aleatoriedad, en este caso se selecciona el modelo ARIMA (5,1,1)



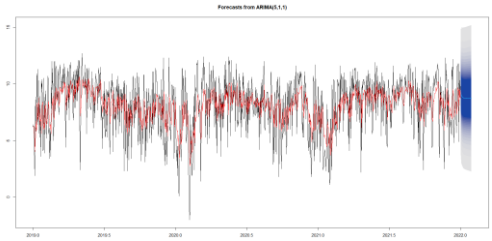
Gráfica 5: Residuales del modelo ARIMA (5,1,1)

Los residuales del modelo indican ruido blanco con media 0 y que los datos no son normales, así como los valores del estadístico de prueba del test de autocorrelación, que se alejan bastante del mínimo requerido. (Gráfica 5)

AÑO	MES	DIA	FORECAST	REAL
2022	Febrero	1	9,31	2,8
2022	Febrero	2	8,94	6,2
2022	Febrero	3	8,99	3,4
2022	Febrero	4	8,86	7,2
2022	Febrero	5	8,86	10,7
2022	Febrero	6	8,81	11
2022	Febrero	7	8,78	10,3
2022	Febrero	8	8,77	11,4
2022	Febrero	9	8,75	8,7
2022	Febrero	10	8,75	6,7
2022	Febrero	11	8,74	8
2022	Febrero	12	8,74	8,2
2022	Febrero	13	8,73	11,9
2022	Febrero	14	8,73	9,1
2022	Febrero	15	8,73	11,4
2022	Febrero	16	8,73	11,6
2022	Febrero	17	8,73	11,2
2022	Febrero	18	8,73	11
2022	Febrero	19	8,73	11,4
2022	Febrero	20	8,73	10,01
2022	Febrero	21	8,73	10,07
2022	Febrero	22	8,73	6,5
2022	Febrero	23	8,73	8,3
2022	Febrero	24	8,73	10,9
2022	Febrero	25	8,73	10,2
2022	Febrero	26	8,73	12
2022	Febrero	27	8,73	12,1
2022	Febrero	28	8,73	12,1

Tabla 3: Resultados del pronóstico vs valores reales para el mes de febrero modelo ARIMA

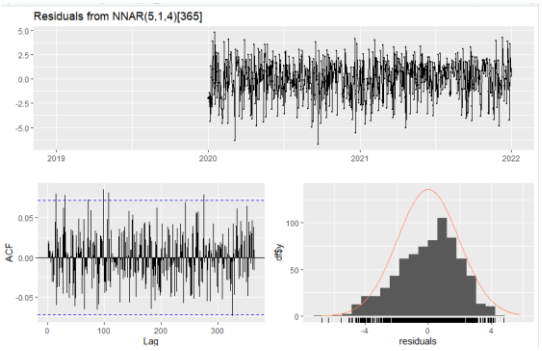
Loa valores del pronóstico muestran que éste tiende a arrojar datos cercanos a la media (Gráfica 6), sin embargo, en días de valores cercanos a los mínimos o máximos históricos, el pronóstico no es acertado. También se observa en la tabla 3, que después de los primeros 13 días de febrero el valor del pronóstico no varía.



Gráfica 6: Pronóstico con el modelo ARIMA (5,1,1) y sus coeficientes.

Redes neuronales de retroalimentación

Debido a que se considera que con un modelo ARIMA el pronóstico no es muy bueno, se procede a obtener un nuevo pronóstico mediante redes neuronales de retroalimentación (nntear). En este caso se elaboró la red con una capa oculta y los residuales indican que los valores de la media no se ajustan a 0 (Gráfica 7), del mismo modo, al imprimir los valores del pronóstico con un nivel de significancia del 95% para el mes de febrero (Tabla 4), aunque estos son más cercanos a los reales en el caso de las temperaturas más bajas, tampoco se considera un buen modelo.

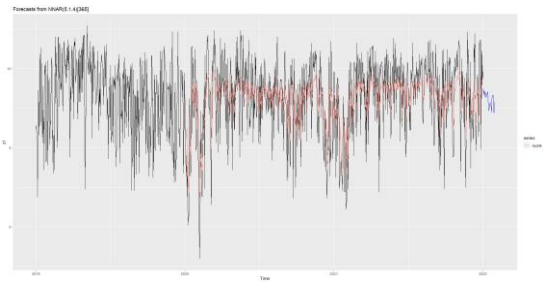


Gráfica 7: Residuales para una red neuronal de retroalimentación.

AÑO	MES	DIA	FORECAST	REAL
2022	Febrero	1	8,7	2,8
2022	Febrero	2	8,3	6,2
2022	Febrero	3	8,6	3,4
2022	Febrero	4	8,5	7,2
2022	Febrero	5	8,2	10,7
2022	Febrero	6	8,5	11
2022	Febrero	7	8,4	10,3
2022	Febrero	8	8,2	11,4
2022	Febrero	9	8,4	8,7
2022	Febrero	10	8,4	6,7
2022	Febrero	11	8,5	8
2022	Febrero	12	8,5	8,2
2022	Febrero	13	7,6	11,9
2022	Febrero	14	7,3	9,1
2022	Febrero	15	7,4	11,4
2022	Febrero	16	7,6	11,6
2022	Febrero	17	7,8	11,2
2022	Febrero	18	7,8	11
2022	Febrero	19	7,7	11,4
2022	Febrero	20	7,5	10,01
2022	Febrero	21	7,2	10,07
2022	Febrero	22	8,0	6,5
2022	Febrero	23	8,3	8,3
2022	Febrero	24	8,2	10,9
2022	Febrero	25	8,3	10,2
2022	Febrero	26	7,3	12
2022	Febrero	27	7,2	12,1
2022	Febrero	28	7,9	12,1

Tabla 4: Resultados del pronóstico vs valores reales para el mes de febrero con redes de retroalimentación.

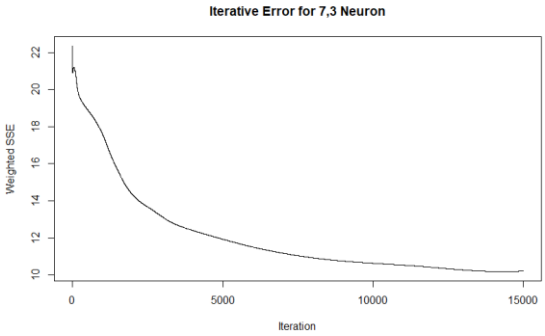
Si se observa la gráfica 8, los valores arrojados por el modelo tienden a variar más y acercarse a los reales, pero al arrojar el pronóstico para el mes de febrero, el modelo de redes neuronales recurrentes se centra en valores cercanos a la media.



Gráfica 8: Pronóstico con el modelo de redes neuronales de retroalimentación.

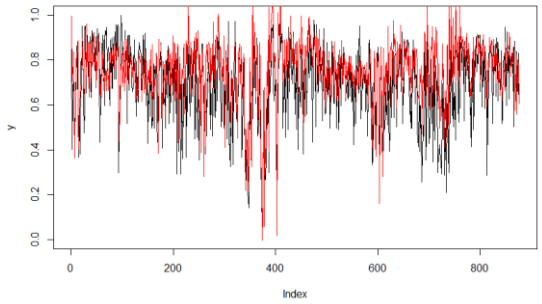
Redes neuronales de recurrentes

Los 1096 datos se dividieron en un conjunto de entrenamiento (80%) y un conjunto de prueba (20%). La creación del dataframe se empleó para adelantar los valores de la serie en un futuro de 28 días, se eliminaron los NA y se determinaron los valores de entrada y salida de la red. En la gráfica 9 se observa que llegando a las 15000 iteraciones el error tiende a ser cercano a cero.



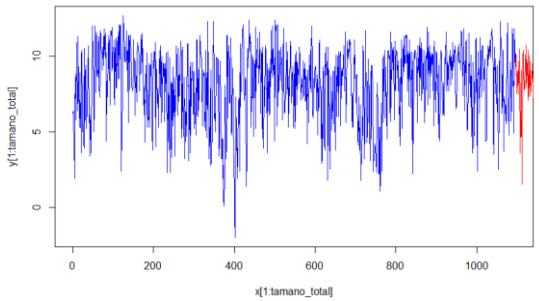
Gráfica 9: Error para 15000 iteraciones.

Cuando se obtienen los valores del modelo en la gráfica 10 se puede notar que la predicción es más cercana a los datos reales que con los modelos anteriormente propuestos, incluso hay picos mucho más marcados que los datos de la serie original.



Gráfica 10: Predicción de los valores con el método de Elman

Finalmente, al predecir los datos, se observa que éstos son mucho más similares a los valores originales (Gráfica 11) y al elaborar el pronóstico para el mes de febrero, los valores tienden a tener más picos similares a los datos reales que a un valor promedio de temperatura. (Tabla 5)



Gráfica 11: Pronóstico para el mes de febrero con el método Elman

AÑO	MES	DIA	FORECAST	REAL
2022	Febrero	1	8,7	2,8
2022	Febrero	2	10,1	6,2
2022	Febrero	3	7,5	3,4
2022	Febrero	4	9,0	7,2
2022	Febrero	5	8,0	10,7
2022	Febrero	6	9,1	11
2022	Febrero	7	7,5	10,3
2022	Febrero	8	8,8	11,4
2022	Febrero	9	10,5	8,7
2022	Febrero	10	9,2	6,7
2022	Febrero	11	3,6	8
2022	Febrero	12	5,7	8,2
2022	Febrero	13	7,3	11,9
2022	Febrero	14	9,1	9,1
2022	Febrero	15	7,9	11,4
2022	Febrero	16	1,5	11,6
2022	Febrero	17	10,2	11,2
2022	Febrero	18	7,3	11
2022	Febrero	19	10,4	11,4
2022	Febrero	20	9,1	10,01
2022	Febrero	21	7,8	10,07
2022	Febrero	22	7,8	6,5
2022	Febrero	23	8,3	8,3
2022	Febrero	24	9,5	10,9
2022	Febrero	25	10,7	10,2
2022	Febrero	26	7,9	12
2022	Febrero	27	8,5	12,1
2022	Febrero	28	9,8	12,1

Tabla 5: Resultados del pronóstico vs valores reales para el mes de febrero con redes recurrentes.

Por último, al observar los resultados de los errores cuadráticos medios de cada pronóstico con los valores reales, se observa que el modelo ARIMA tiene menor error, aunque los valores del pronóstico de Elman y nntear son más cercanos a los reales, esto ocurre debido a que el error cuadrático medio trabaja mejor con resultados de la media de los datos que con grandes errores respecto a los reales. En el caso del error MAPE, este mide el error en porcentaje y se debe tener en cuenta para casos de variables cercanas a 0 que el calculo del error no será muy bueno.

MODELO	RMSE	MAPE
ARIMA	14,21	32,73
NNTEAR	16,68	35,70
ELMAN	18,08	36,37

Tabla 6: Errores cuadráticos medios para cada modelo

CONCLUSIONES

Se considera importante crear modelos de alta precisión debido a que temperaturas muy bajas pueden ocasionar problemas fisiológicos a las plantas que en condiciones extremas pueden generar daños irreparables. A pesar de que los valores de temperaturas mínimas con el pronóstico del modelo Elman son más cercanas a las reales y varían más, el error cuadrático medio indica que el modelo ARIMA es el mejor respecto a los valores reales, sin embargo, estos datos tienden a ser muy similares a la media, por lo que en meses donde las mínimas no varíen mucho como julio y agosto se podría realizar un buen pronóstico, pero en temporadas de mayor variación, o en meses con probabilidades de heladas, como diciembre y enero, cuando las mínimas tiendan a cero grados, el pronóstico no será bueno.

Teniendo en cuenta lo anterior, se considera que para que el pronóstico sea mucho mejor, es necesario incluir más variables climáticas, como humedad relativa y quizá temperatura máxima.

REFERENCIAS BIBLIOGRÁFICAS

Aguirre Jaime, A. (1994). Introducción al tratamiento de series temporales: aplicación a las ciencias de la salud. En A. Aguirre Jaime, Introducción al tratamiento de series temporales: aplicación a las ciencias de la salud (págs. 231-233). Madrid: Ediciones Díaz de Santos.

Alcaldía del Rosal. (15 de Marzo de 2018). El Rosal Cundinamarca. Obtenido de El Rosal Cundinamarca: http://www.elrosal-cundinamarca.gov.co/municipio/generalidades-el-rosal-cundinamarca

Alfaro, E. J., & Soley, F. J. (2009). Descropción de dos métodos de rellenado dedatos ausentes en series de tiempometeorológicas. Revista de Matemática: Teoría y aplicaciones, 60-75.

Asocolflores. (Febrero de 2022). Asocolflores. Obtenido de Asocolflores: https://asocolflores.org/es/asi-registraron-los-medios-la-temporada-de-san-valentin-2022/

de Arce, R., & Mahia, R. (Año no especificado). Modelos ARIMA Programa Citius.- Técnicas de Previsión de variables financieras. Obtenido de Modelos ARIMA Programa Citius.- Técnicas de Previsión de variables financieras: https://d1wqtxts1xzle7.cloudfront.net/53321017/Box-Jenkins-with-cover-page-v2.PDF?Expires=1645989030&Signature=HQPxH4a4k11YcmQCVYnfCfExYPkVdxZOWocx71I7CgOOHWbGa1S-1LSqSvDhiyN4zS1ojp6QPDzseHCRyee4Zw8lqx-l1lHVt2fZQDzSaWPT1NDq4~TdD3CzM6HHOijEf6JtJpQdbKEJYjaj

Fenández, S. d. (Año no especificado). Portal estadistica aplicada. Obtenido de Portal estadistica aplicada: https://www.estadistica.net/ECONOMETRIA/SERIES-TEMPORALES/modelo-arima.pdf

Fournies Pallavicini, A. (Abril de 2015). ResearchGate. Obtenido de ResearchGate: https://www.researchgate.net/publication/274961095_Modelos_ARMA_y_Box_and_Jenkins

González Casimiro, M. P. (Año no especificado). Análisis de series temporales: Modelos ARIMA. Bilbao: Euskal Herriko Unibertsitatea.

Guisande Golzalez, C., Barreiro Felpeto, A., Maneiro Estraviz, I., Riveiro Alarcón, I., Vergara Castaño, A. R., & Vaamonde Liste, A. (2006). Tratamiento de dados. En C. Guisande Golzalez, A. Barreiro Felpeto, I. Maneiro Estraviz, I. Riveiro Alarcón, A. R. Vergara Castaño, & A. Vaamonde Liste, Tratamiento de dados (pág. 235). España: Diaz de Santos.

Hernández Rodríguez, O. (1998). Temas de análisis estadístico multivariado. En O. Hernández Rodríguez, Temas de análisis estadístico multivariado (págs. 7-19). San josé: Editorial de la universidad de Costa Rica.

IDEAM. (2012). Actualización notas técnicas helada 2012. Obtenido de Actualización notas técnicas helada 2012: http://www.ideam.gov.co/documents/21021/21147/Documento+FINAL+actualizacion+nota+tecnica+heladas.pdf/e10a0183-62e6-410a-8e96-7e0739f6f06b

Ministerio de Agricultura y Desarrollo Rural . (Febrero de 2020). Ministerio de Agricultura y Desarrollo Rural . Obtenido de Ministerio de Agricultura y Desarrollo Rural : https://sioc.minagricultura.gov.co/Flores/Documentos/2019-12-30%20Cifras%20Sectoriales.pdf

Monroy, N., Pérez, I., & Cure, J. R. (2001). Estudio de la variabilidad en el clima y la producción de rosas en la sabana de bogotá. Revista de Ingeniería Universidad de los Andes, 38-43.

Ramirez Forero, D. D., & Malagón Márquez, D. A. (2018). Repositorio Universidad Distrital. Interpoladores determinísticos espacio-temporales, series de tiempo y analisis de datos funcionales para el estudio y predicción de la precipitación en Cundinamara y Bogotá D.C. Bogotá, Colombia: Repositorio Universidad Distrital. Obtenido de Repositorio Universidad Distrital: https://repository.udistrital.edu.co/bitstream/handle/11349/14699/Ram%c3%adrezForeroDilsonDavid2018.pdf?sequence=1&isAllowed=y

Reyes Polanco, A. E. (2021). SERIES DE TIEMPO: Conceptos básicos de análisis de series de tiempo en el dominio del tiempo. En A. E. Reyes Polanco, SERIES DE TIEMPO: Conceptos básicos de análisis de series de tiempo en el dominio del tiempo (págs. 21-25). Madrid: Edicion Kindle.

Salazar López, B. (30 de Junio de 2019). Ingeniería Industrial online. Obtenido de Ingeniería Industrial online: https://www.ingenieriaindustrialonline.com/pronostico-de-la-demanda/suavizacion-exponencial-simple/

Toledo Vega, L. C. (1 de Marzo de 2015). Universidad Autónoma del Estado de México. Obtenido de Universidad Autónoma del Estado de México: http://ri.uaemex.mx/oca/bitstream/20.500.11799/32339/1/secme-11680.pdf

Vanegas Ramos, M. Á. (2011). Implementación de modelos locales en el espacio de fase para el pronóstico de variables hidrometeorológicas a partir de series de tiempo. Implementación de modelos locales en el espacio de fase para el pronóstico de variables hidrometeorológicas a partir de series de tiempo. Bogotá, Colombia: Departamento de Ingeniería Civil y Agrícola.

Villazón Bustillos, D., Rubio Arias, H. O., Ortega Gutiérrez, J. Á., Rentería Villalobos, M., Gonzáles Gurrola, L. C., & Pinales Munguia, A. (2016). Análisis en series de tiempo para el pronóstico de sequía en la región noroeste del estado de Chihuahua. Ecosistemas y Recursos Agropecuarios, 207-315.