

# Serie Temporal Familia ARIMA

Ignacio Fernández Sánchez-Pascuala, Javier Castellano Soria  
*Universidad Complutense de Madrid | Universidad Politécnica de Madrid*

12 de febrero de 2024

## Resumen

Este trabajo se enfoca en la modelización de una serie temporal climática de temperatura media mensual a lo largo de varios años. Se ha utilizado la metodología Box-Jenkins para ajustar un modelo adecuado a la serie.

Se lleva a cabo un análisis detallado que abarca desde la exploración inicial de los datos, la identificación y tratamiento de outliers, hasta la aplicación de transformaciones para estabilizar la serie.

La descomposición de la serie en componentes estacionales, tendencia y residuales se realiza para comprender mejor la estructura subyacente. Se exploran técnicas de suavizado y diferenciación para lograr estacionariedad.

La identificación del modelo ARIMA, su ajuste y pruebas de diagnóstico se presentan, confirmando la idoneidad del modelo para la serie temporal. Se abordan aspectos como la validación de la normalidad, homocedasticidad e independencia de los residuos y la realización de predicciones a futuro.

## 1. Introducción

Este trabajo se centra en el análisis y modelización de una serie temporal climática, específicamente, las temperaturas medias mensuales en las regiones del noreste y este de Inglaterra durante el periodo 2010-2019.

El conjunto de datos proporciona datos mensuales que incluyen la media de precipitación en milímetros cúbicos y la temperatura media en grados centígrados, categorizados por mes, año y estación meteorológica.

## 1.1. Datos Utilizados

En el análisis, se utilizan específicamente los datos de temperatura media mensual. Se ha realizado una partición de los datos reservando los últimos 12 meses para evaluar el rendimiento del modelo posteriormente.

A continuación, se muestra la serie temporal representada gráficamente en la Figura 1, proporcionando una perspectiva inicial de las tendencias y patrones que se abordarán en este trabajo.

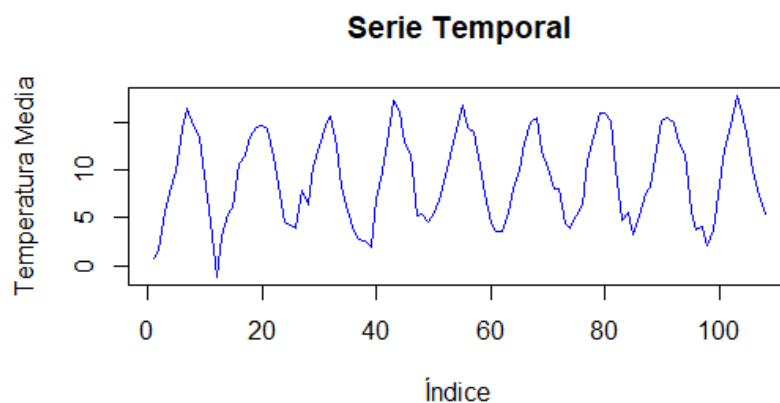


Figura 1: Serie Temporal de Temperatura Media Mensual.

## 1.2. Objetivos del Trabajo

El objetivo principal de este trabajo es aplicar la metodología Box-Jenkins para modelar la serie temporal de temperatura, tratando aspectos como la estacionariedad, identificación y ajuste del modelo de la familia SARIMA, validación del modelo, y generación de predicciones a futuro.

Además, se responderá así al interrogante de si hay una tendencia creciente en los últimos años de la temperatura media en la zona debido al cambio climático global.

En las secciones siguientes, se detallarán los pasos metodológicos, resultados obtenidos y conclusiones derivadas de este estudio.

## 2. Metodología

En este trabajo, se emplea la metodología Box-Jenkins, un procedimiento iterativo que consta de cuatro pasos fundamentales: Identificación, Estimación, Validación y Predicción. Este enfoque no lineal se caracteriza por su naturaleza cíclica, lo que significa que el modelo identificado en una iteración no necesariamente será el más adecuado para obtener predicciones precisas.

Antes de aplicar la metodología Box-Jenkins, se realiza un análisis preliminar de la serie temporal:

- Se realiza un análisis exploratorio de los datos que incluye, la representación de la serie temporal para visualizar patrones y tendencias iniciales, y el tratamiento de valores atípicos.
- La SST se descompone en sus componentes estacionales, de tendencia y residuales.
- Se aplican transformaciones y diferenciaciones necesarias para lograr estacionariedad en la SST. Una serie es estacionaria cuando, sus valores dependen de su posición relativa y no del instante en el que se encuentran. Más detalladamente, el proceso  $X_t$  será estrictamente estacionario si para cualquier vector  $(X_{t_1}, \dots, X_{t_n})$  su distribución no depende de los instantes  $t_1, \dots, t_n$ , pero sí puede depender de la distancia entre estos momentos. Sin embargo, suele considerarse su definición en sentido amplio o débil, esta solo requiere que la serie sea constante en media y que la correlación entre el instante  $t$  y  $t + \tau$  solo dependa de  $\tau$  y no del tiempo.

## 2.1. Análisis exploratorio y tratamiento de valores atípicos

En esta fase, se representa la serie temporal para adquirir una primera noción de su estructura. Esta representación nos ayudará a determinar las futuras transformaciones. Además, se realiza un análisis de la presencia de valores atípicos en la serie.

Este análisis permite determinar si ha habido alguna intervención o si se trata de un valor anómalo por causas desconocidas. Los valores atípicos pueden ser de varios tipos, dependiendo de si su duración es temporal o permanente, y si el efecto es brusco o gradual.

Una vez detectados los valores atípicos, se procederá según la tipología de estos. Este paso es fundamental para un correcto modelaje de los datos y una estimación adecuada de los parámetros del modelo.

## 2.2. Descomposición de la Serie Temporal

La descomposición de una serie temporal es un proceso fundamental para entender su estructura y determinar las transformaciones necesarias para conseguir una serie estacionaria. En este análisis, se descomponen las observaciones temporales en tres componentes principales:

1. **Componente Estacional:** Esta componente representa las fluctuaciones periódicas y recurrentes que ocurren en la serie temporal. En este caso, se busca identificar patrones estacionales que se repiten en los mismos meses del año, generalmente asociados con cambios climáticos predecibles. Hay de dos tipos, aditiva si hay un patrón fijo de amplitud o multiplicativa si la amplitud va aumentando o disminuyendo a lo largo del tiempo.

2. **Componente de Tendencia:** La tendencia refleja la dirección general de la serie temporal a lo largo del tiempo. Se analizan las variaciones a largo plazo, permitiendo la identificación de tendencias ascendentes, descendentes, estables o la presencia de ciclos. En este estudio, se presta especial atención a la tendencia para evaluar su estabilidad y posible influencia de eventos externos como el cambio climático.
3. **Componente Residual:** La componente residual representa la variabilidad no explicada por la componente estacional y de tendencia. Es esencial para identificar patrones no modelados y garantizar que el modelo captura la información completa de la serie.

Matemáticamente, denotando a  $X_t$  la serie temporal,  $S_t$  la componente estacional,  $T_t$  la tendencia y  $E_t$  el residuo, la descomposición puede ser aditiva (1) o multiplicativa (2):

$$X_t = S_t + T_t + E_t, \quad (1)$$

$$X_t = S_t \cdot T_t \cdot E_t. \quad (2)$$

La primera se empleará cuando la estacionalidad sea aditiva y la segunda cuando sea multiplicativa.

### 2.3. Transformaciones y diferenciaciones

A continuación, a partir de los resultados de la descomposición, se realizan, en este orden, las transformaciones, diferenciaciones estacionales y diferenciaciones regulares oportunas para estabilizar la serie y quitar la componente estacional. De este modo, se podrá identificar posteriormente el modelo.

La transformación de la serie se debe considerar cuando esta presenta una estacionalidad multiplicativa. La más usada es la transformación Box-Cox con ecuación (3) donde  $m$  es una constante y el parámetro  $\lambda$  es calculado para obtener la transformación más óptima,

$$T_\lambda(X_t) = \begin{cases} \frac{(X_t+m)^\lambda - 1}{\lambda} & \text{si } \lambda \neq 0 \\ \ln(X_t + m) & \text{si } \lambda = 0 \end{cases} \quad (3)$$

Una vez eliminada la posible estacionalidad multiplicativa, se debe comprobar que la serie no tiene estacionalidad. Para esto nos podemos servir de gráficos como el periodograma. Los picos en el periodograma indican las frecuencias dominantes en la serie temporal. Esto es útil para identificar ciclos estacionales. Si se observa un patrón de estos picos, nos estaría indicando que hay una componente estacional. Para eliminarla se debe hacer una diferenciación estacional de período el observado obteniendo la serie  $\Delta_s X_t$  según:

$$\Delta_s X_t = X_t - X_{t-s} \quad \text{con } t = s + 1, \dots, T, \quad (4)$$

donde  $s$  es el período de la estacionalidad,  $T$  la cantidad total de datos y  $X_t$  la serie tras la posible transformación.

Por último se debe analizar si hay una posible tendencia en nuestra serie. Si ocurriese, se realizarían transformaciones regulares para eliminar la tendencia calculando la serie  $\Delta X_t$ :

$$\Delta X_t = X_t - X_{t-1} \quad \text{con } t = 2, \dots, T, \quad (5)$$

donde  $X_t$  es la serie tras la posible diferenciación anterior.

Cabe destacar que existen contrastes de hipótesis que nos pueden ayudar a determinar la cantidad de diferenciaciones que debemos tomar para estacionarizar la serie.

## 2.4. Identificación

El objetivo es proponer un modelo del tipo  $\text{SARIMA}(p, d, q) \times (P, D, Q)_s$ . La ecuación general del modelo mencionado sería:

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D X_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t, \quad (6)$$

con  $B$  siendo el operador retardo y:

$$\begin{aligned} \phi_p(B) &= 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \\ \Phi_P(B^s) &= 1 - \Phi_{s1} B^s - \Phi_{s2} B^{2s} - \dots - \Phi_{sP} B^{Ps} \\ \theta_q(B) &= 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \\ \Theta_Q(B^s) &= 1 - \Theta_{s1} B^s - \Theta_{s2} B^{2s} - \dots - \Theta_{sQ} B^{Qs}. \end{aligned} \quad (7)$$

Por un lado,  $p$ ,  $d$  y  $q$  son el orden de la componente autoregresiva, la cantidad de diferenciaciones regulares y el orden de la componente de media móvil respectivamente. Por otro,  $P$ ,  $D$  y  $Q$  son análogos a los valores anteriores, pero para la parte estacional.

En función de las transformaciones consideradas en la fase anterior, los valores  $d$  y  $D$  ya estarían determinados. Ahora, tras haber quitado la tendencia y la estacionalidad de la serie, se emplean las funciones de autocorrelación simple y autocorrelación parcial para guiar la elección del resto de hiperparámetros del modelo.

La FAS (Función de Autocorrelación Simple), con ecuación (8), muestra la correlación entre una observación y sus observaciones precedentes a diferentes rezagos. Por otro lado, la FAP (Función de Autocorrelación Parcial), con ecuación (9), representa la correlación entre una observación y sus observaciones pasadas, eliminando las correlaciones de los rezagos intermedios.

$$\text{FAS}(k) = \frac{\sum_{t=k+1}^T ((x_t - \hat{\mu})(x_{t-k} - \hat{\mu}))}{\sum_{t=1}^T (x_t - \hat{\mu})^2}, \quad (8)$$

$$\text{FAP}(k) = \frac{\sum_{t=k+1}^T ((x_t|x_{t-1}, \dots, x_{t-k+1} - \hat{\mu}_1)(x_{t-k}|x_{t-1}, \dots, x_{t-k+1} - \hat{\mu}_2))}{\sqrt{\sum_{t=k+1}^T (x_t|x_{t-1}, \dots, x_{t-k+1} - \hat{\mu}_1)^2 \cdot \sum_{t=k+1}^T (x_{t-k}|x_{t-1}, \dots, x_{t-k+1} - \hat{\mu}_2)^2}}, \quad (9)$$

con  $k$  siendo el rezago,  $x_t$  la serie temporal,  $\hat{\mu}$  la media muestral de la serie,  $T$  la longitud de esta,  $x_t|x_{t-1}, \dots, x_{t-k+1}$  la serie temporal resultante de considerar los residuos de una regresión lineal de  $x_t$  frente a  $x_{t-1}, \dots, x_{t-k+1}$ ,  $x_{t-k}|x_{t-1}, \dots, x_{t-k+1}$  la serie temporal de los residuos de la correspondiente regresión y por último,  $\mu_1$  y  $\mu_2$  las medias de los residuos de las dos regresiones anteriores respectivamente.

Ambas funciones, FAS y FAP, proporcionan información valiosa para la identificación de patrones en la serie temporal. La FAS nos ofrece una visión del posible modelo que subyace y la FAP los órdenes a considerar.

## 2.5. Estimación

Una vez que hemos identificado el modelo SARIMA más apropiado para nuestra serie temporal, se estiman sus parámetros. Dos métodos comunes de estimación de parámetros son el Método de Máxima Verosimilitud (ML) y Mínimos Cuadrados.

En el caso de un modelo con solo una componente autoregresiva, las ecuaciones de Yule-Walker pueden ser utilizadas para la estimación de parámetros. Sin embargo, cuando hay componentes de media móvil, se requieren métodos iterativos de optimización.

## 2.6. Validación

Una vez identificado y estimado el modelo, se deben tener en cuenta varios aspectos para la validación de este. Esto engloba estudiar la posibilidad de simplificar o ampliar el modelo, usar criterios de bondad de ajuste y analizar los residuos.

Para la simplificación se plantean contrastes de hipótesis sobre los parámetros del modelo. Se toma como hipótesis nula que un determinado parámetro sea nulo frente al caso contrario. Si un parámetro tiene un p-valor superior al nivel de significación (0.05 usualmente), se estimaría un modelo sin este parámetro. Este proceso se repite eliminando de uno en uno hasta que todos los parámetros sean significativos. Si ya se ha identificado un modelo significativo hay que evaluar la posible extensión con más parámetros. Se irían añadiendo uno a uno aumentando los órdenes AR y MA probando si se obtiene un modelo significativo que mejore criterios de bondad de ajuste que penalizan el número de parámetros utilizados. Uno de los más usados es el AIC (criterio de información de Akaike):

$$\text{AIC} = 2k - 2\ln(L), \quad (10)$$

con  $k$  el número de parámetros y  $L$  el máximo de la función de verosimilitud. También hay otros como el BIC (criterio de información bayesiano) que penaliza más la incorporación de un nuevo parámetro.

Por último, debemos comprobar que la serie de los residuos es un ruido blanco normal. Debemos verificar la independencia, la normalidad y la homocedasticidad. Para la primera se pueden representar los residuos respecto al tiempo, calcular correlogramas y realizar contrastes de rachas o independencia. Para la normalidad se puede obtener el *Q-Q plot* de los residuos para compararlos con una normal o realizar contrastes como Shapiro-Wilk (menos de 50 observaciones) o Kolmogorov-Smirnov (más de 50 observaciones) donde la hipótesis nula es la normalidad. Para la homocedasticidad se suele ver en que la dispersión de los residuos es constante en el tiempo.

Una prueba estadística importante a considerar es la de Ljung-Box que como hipótesis nula engloba la independencia de los datos (cualquier conjunto de autocorrelaciones es cero), la normalidad y la homocedasticidad. Aunque el p-valor no sea lo suficientemente pequeño como para rechazar la hipótesis nula, se debe realizar de forma adicional la prueba de normalidad.

Si el modelo no cumple una de las hipótesis, se deberá probar con otra propuesta de modelo hasta que se consiga uno que las cumpla todas.

## **2.7. Predicción**

Una vez obtenido un modelo que ha sido validado, podemos calcular predicciones puntuales para tiempos futuros en incluso intervalos de predicción para tener una medida de incertidumbre. Además, si se ha reservado una parte del conjunto de datos que no han intervenido en el ajuste del modelo, se puede calcular una estimación del error de predicción.

## **3. Resultados**

Una vez explicada la metodología utilizada en el estudio de la serie, se exponen los distintos resultados obtenidos en los procesos descritos. Para obtener estos resultados, se ha utilizado el software de R.

### **3.1. Análisis Preliminar**

La serie a tratar se puede ver representada en la Figura 1. Se observa una clara estacionalidad, que será visualizada mejor en la descomposición de la serie.

Se han aplicado una serie de técnicas estadísticas para la detección de outliers. Sin embargo, los resultados indicaron que no existen valores atípicos significativos en la serie. Por lo tanto, no fue necesario realizar correcciones o imputaciones en este aspecto.

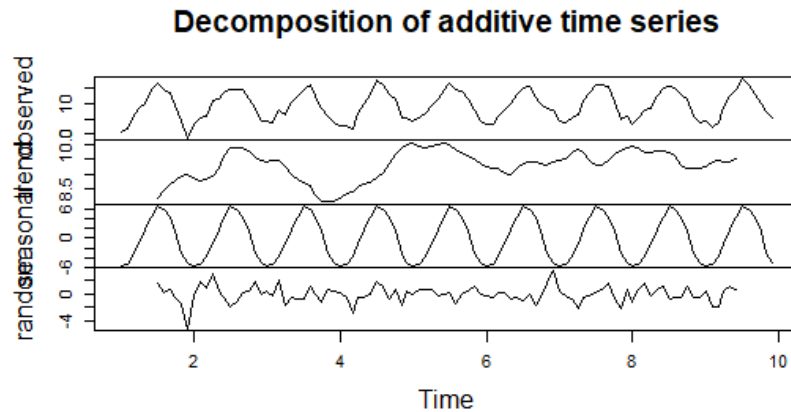


Figura 2: Descomposición Inicial de la Serie Temporal.

La descomposición inicial, presentada en la Figura 2, revela una serie con una clara componente estacional que se repite cada 12 meses. Además, se observa una tendencia con ciertos ciclos que apuntan a la posibilidad de intervenciones externas en la serie. Tras llevar a cabo un análisis más profundo, se observa una intervención de comienzo gradual y duración temporal, con una disminución considerable seguida de una recuperación. Esto ha motivado la decisión de recortar la serie a partir del año 4.

La justificación de este recorte se basa en la búsqueda de una tendencia más estable, menos influenciada por intervenciones externas. Al eliminar el periodo de intervención, se espera que el modelo resultante capture de manera más efectiva los patrones subyacentes y proporcione predicciones más precisas para los años futuros. La nueva serie resultante se presenta en la Figura 3.

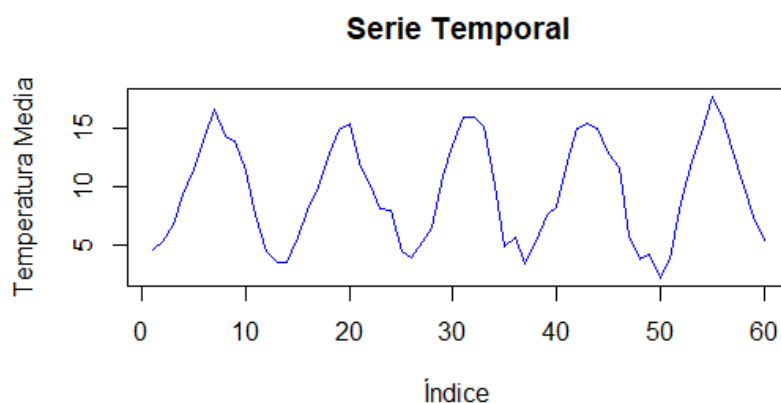


Figura 3: Serie Temporal Recortada a partir del Año 4.

Esta nueva serie revela una tendencia más estable y con menos influencia de intervenciones externas. Además, se ha vuelto a estudiar la posible existencia de outliers, y no se ha detectado ninguno. La descomposición correspondiente se muestra en la Figura 4, donde la componente



estacional sigue siendo evidente cada 12 meses y la tendencia exhibe menos variabilidad en comparación con la serie completa.

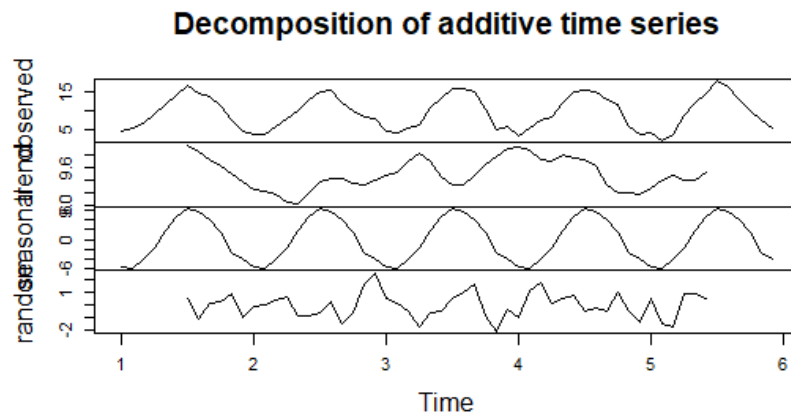


Figura 4: Descomposición de la Serie Temporal Recortada.

Es crucial destacar que la serie no muestra estacionalidad multiplicativa. En consecuencia, no se considera necesario aplicar transformaciones como la de Box-Cox a la serie.

En busca de estacionarizar la serie temporal y eliminar la componente estacional identificada cada 12 meses, se aplicó una diferenciación estacional de primer orden con pauta de repetición 12.

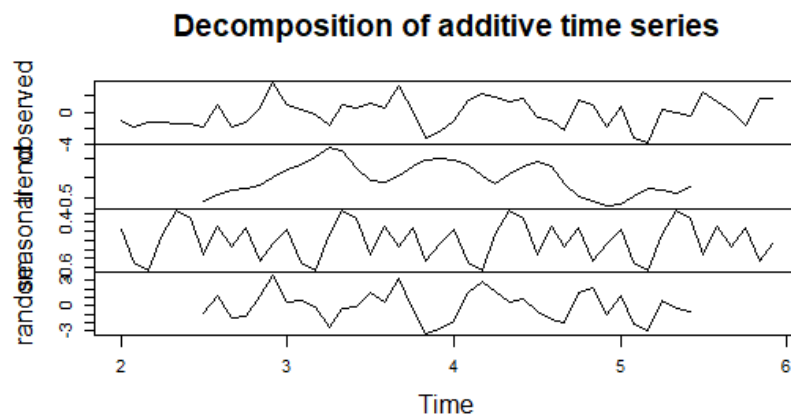


Figura 5: Serie Temporal después de la diferenciación estacional.

La Figura 5 muestra la serie temporal una vez aplicada la diferenciación estacional de primer orden. Se observa que la componente estacional ha desaparecido de la serie.

Esto se puede verificar en los periodogramas de la Figura 6, donde en la serie original se observa un pico de alta frecuencia que indica estacionalidad, sin embargo, después de la diferenciación, los picos son más aleatorios y mucho más bajos en escala.

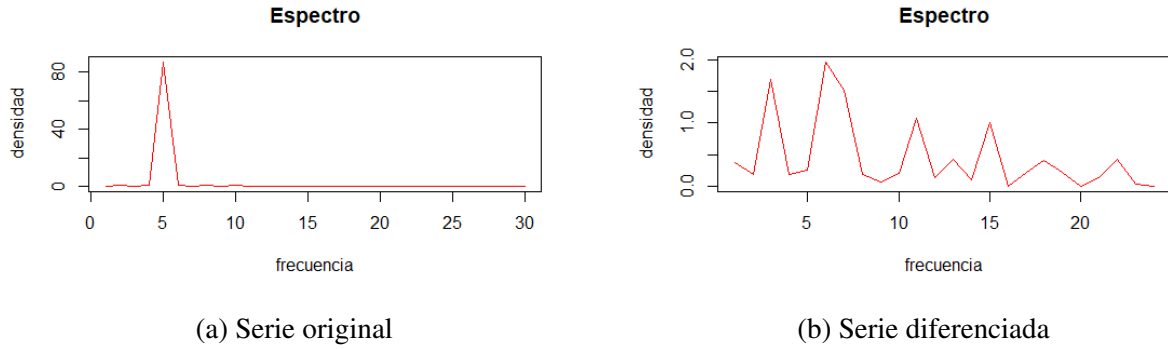


Figura 6: Periodogramas

Cabe destacar que se ha determinado no realizar ninguna diferenciación simple, ya que la serie no presenta una tendencia lineal evidente y el test tampoco lo aconsejaba.

Para comprobar que hemos logrado estacionarizar la serie con las transformaciones mencionadas, se realizó la prueba de Dickey-Fuller. Como el p-valor obtenido fue  $0.022 < 0.05$ , se rechazó la hipótesis nula aceptando la estacionaridad.

### 3.2. Aplicación metodología Box-Jenkins

Para identificar el modelo SARIMA más adecuado a partir de la serie temporal diferenciada estacionalmente, se analiza la Función de Autocorrelación Simple (FAS) y la Función de Autocorrelación Parcial (FAP).

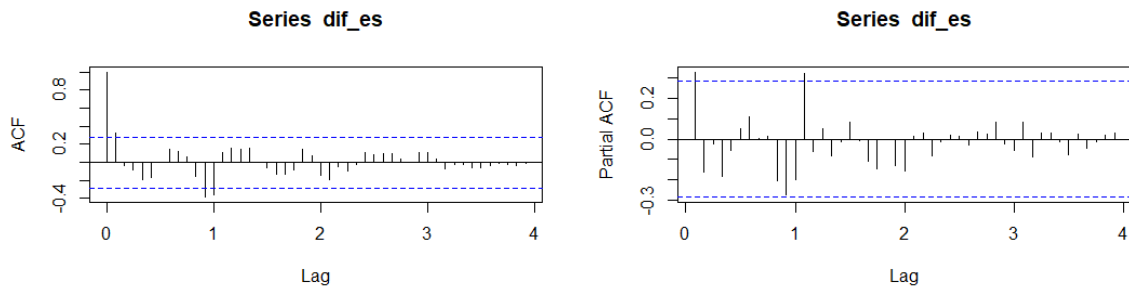


Figura 7: Funciones de Autocorrelación Simple y Parcial respectivamente.

En la FAS de la figura 7 se observa un decaimiento progresivo senoidal de la autocorrelación simple a medida que se toma un rezago mayor. Se tienen dos correlaciones significativas, una con rezago 1 y otra con rezago 12. Esto sugiere la consideración de un modelo  $\text{SARIMA}(p, d, q) \times (P, D, Q)_{12}$ . Los órdenes del modelo nos lo indica la FAP. Se observa una correlación significativa con rezago 1 y otra con rezago 12. Esto indica que la serie puede estar formada por una componente AR o AM de orden 1 y por una AR o AM de orden 1 estacional.

Por lo tanto, consideraremos  $\text{SARIMA}(0, 0, 1) \times (0, 1, 1)_{12}$ ,  $\text{SARIMA}(1, 0, 0) \times (0, 1, 1)_{12}$ ,  $\text{SARIMA}(0, 0, 1) \times (1, 1, 0)_{12}$  o  $\text{SARIMA}(1, 0, 0) \times (1, 1, 0)_{12}$  como propuestas iniciales de

modelo.

Una vez identificado los modelos, se estiman y validan sus parámetros mediante el método de máxima verosimilitud. Además, se comparan su capacidad de ajuste a la serie mediante medidas como el AIC, el error cuadrático medio o la log-verosimilitud.

1. **SARIMA(0, 0, 1)  $\times$  (1, 1, 0)<sub>12</sub> :**

<b>Coefficiente</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt; z )</b>
ma1	0.43558	0.15796	2.7575	0.0058244
sar1	-0.52106	0.13848	-3.7626	0.0001681

Varianza estimada:  $\sigma^2 \approx 1,842$

Log-verosimilitud:  $-84,77$

AIC: 175,55

RMSE: 1,213899

2. **SARIMA(0, 0, 1)  $\times$  (0, 1, 1)<sub>12</sub> :**

<b>Coefficiente</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt; z )</b>
ma1	0.4164	0.1452	2.8680	0.0041309
sma1	-0.9999	0.26496	-3.7741	0.0001606

Varianza estimada:  $\sigma^2 \approx 1,103$

Log-verosimilitud:  $-80,21$

AIC: 166,41

RMSE: 0,9392

3. **SARIMA(1, 0, 0)  $\times$  (1, 1, 0)<sub>12</sub> :**

<b>Coefficiente</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt; z )</b>
ar1	0.3151	0.1355	2.3250	0.0200705
sar1	-0.5020	0.1384	-3.6266	0.0002872

Varianza estimada:  $\sigma^2 \approx 1,923$   
 Log-verosimilitud:  $-85,6$   
 AIC:  $177,19$   
 RMSE:  $1,24031$

#### 4. SARIMA(1, 0, 0) $\times$ (0, 1, 1)<sub>12</sub> :

Coefficiente	Estimate	Std. Error	z value	Pr(> z )
ar1	0.3346	0.1344	2.4899	0.0127770
sma1	-1.0000	0.2686	-3.7218	0.0001978

Varianza estimada:  $\sigma^2 \approx 1,133$   
 Log-verosimilitud:  $-80,81$   
 AIC:  $167,63$   
 RMSE:  $0,9519457$

Todos los modelos cumplen que sus parámetros son significativos a un nivel de significación de 0.01. Sin embargo, el modelo SARIMA(0, 0, 1)  $\times$  (0, 1, 1)<sub>12</sub> tiene un AIC más bajo, lo que indica un mejor ajuste entre los modelos que estamos considerando con la misma cantidad de parámetros. También presenta un menor valor en el error cuadrático medio.

Siguiendo con la metodología descrita, se intentó ampliar este modelo probando a aumentar los órdenes de uno en uno. Todas las posibles combinaciones dieron lugar a modelos con parámetros no significativos descartando así la posibilidad de considerar un modelo más complejo.

Una vez elegido nuestro modelo, se comprueban la independencia, normalidad y homocedasticidad en los residuos. En primer lugar, se representan algunos gráficos sobre los residuos en la Figura 8.

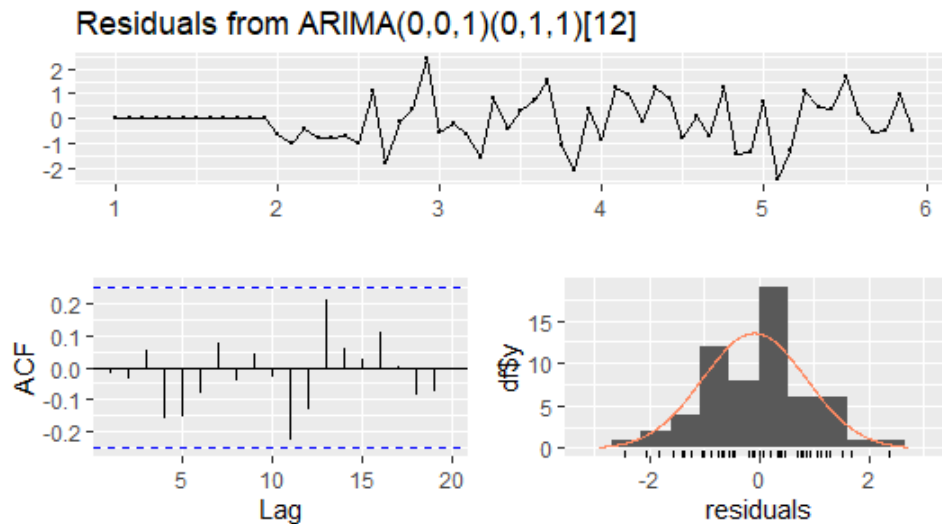


Figura 8: Gráficos de residuos

Se puede observar que los residuos tienen una varianza constante a lo largo del tiempo, por lo que se cumple la homocedasticidad. Además, la serie de los residuos parece asemejarse a un ruido blanco, esto se comprueba teóricamente mediante la prueba de Ljung-Box. Obteniendo un p-valor de  $0,4187 > 0,05$  no se puede rechazar la hipótesis nula de ausencia de autocorrelación, independencia, normalidad y homocedasticidad. A continuación, se comprueba la normalidad con gráficos y contrastes específicos.

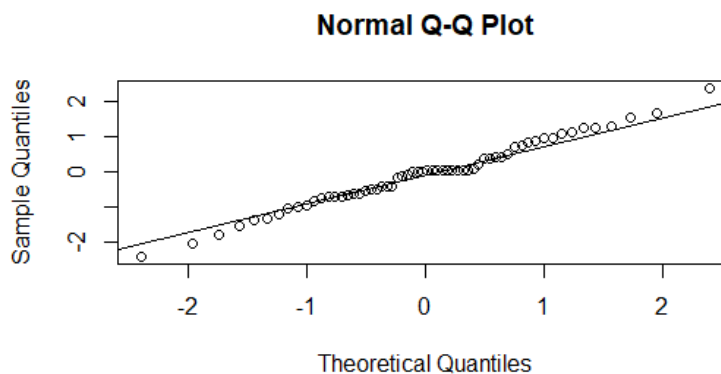


Figura 9: QQPlot residuos

Se observa en el QQPlot de la Figura 8 que los residuos se aproximan a una normal. Esto se verifica con el test de Kolmogorov-Smirnov, ya que hay más de 50 observaciones, donde se obtiene un p-valor de 0.08, por lo que se puede suponer normalidad a un nivel de significación de 0.05.

Tras haber validado nuestro modelo, se predicen los valores de la serie en el año siguiente. Como se observa en la Figura 10, la predicción se ajusta muy bien a los patrones de la serie en los años anteriores. La sombra gris representa los intervalos de predicción en el próximo año.

Como vemos, estos no son muy amplios relativamente, lo cual indica que los residuos no tienen gran variabilidad.

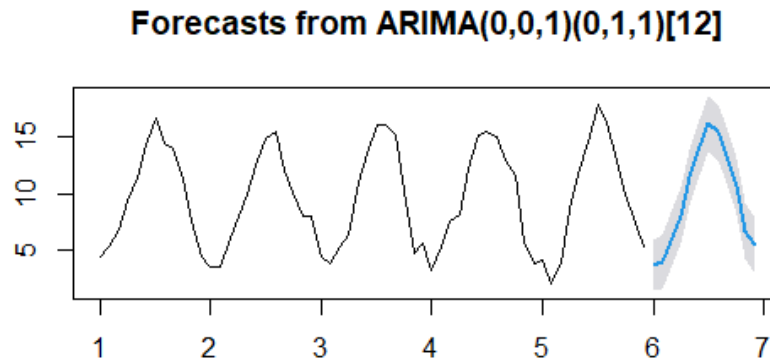


Figura 10: Predicción 1 año modelo

Por último, se comparan los valores predichos con los valores reales en el último año, reservados al principio del estudio, en la Figura 11. Se observa que la predicción se ajusta bastante bien a la serie. Además, se ha estimado el error de predicción mediante la raíz cuadrada del error cuadrático medio, obteniendo un valor de 1.08, lo que indica que de media comete un error de un grado en la temperatura estimada, un valor bastante aceptable.

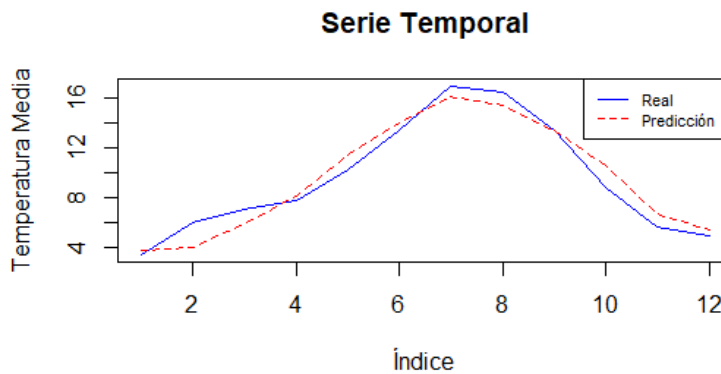


Figura 11: Valores predichos vs reales

## 4. Discusión

Los resultados han mostrado que el modelo  $SARIMA(0, 0, 1) \times (0, 1, 1)_{12}$  es adecuado para la predicción de la temperatura media mensual en la zona. Las predicciones futuras proporcionadas por el modelo sugieren que la temperatura media mensual seguirá patrones estacionales consistentes en los próximos 12 meses.

Sin embargo, es importante tener en cuenta que todas las predicciones están sujetas a incertidumbre y pueden influenciarse por eventos climáticos inesperados o cambios a largo plazo en el clima, como por ejemplo la variabilidad en la tendencia detectada en la zona entre 2010-2014 que desembocó en una tendencia más estable, pero con temperaturas más altas. Las predicciones se han realizado suponiendo esta estabilidad de los últimos años. Variaciones en la tendencia en los próximos años supondrían un inconveniente y sería necesario ajustar otro modelo.

## 5. Conclusiones

En este trabajo, se ha llevado a cabo un análisis exhaustivo de una serie temporal climática, centrándose en la predicción de la temperatura media mensual. Las principales conclusiones son:

1. El modelo  $SARIMA(0, 0, 1) \times (0, 1, 1)_{12}$  se ha identificado como el más adecuado para la predicción, mostrando buen desempeño en las pruebas de diagnóstico.
2. Los residuos del modelo son ruido blanco, independientes y homocedásticos, respaldando la calidad de las predicciones.
3. Las predicciones futuras indican patrones estacionales consistentes, pero con la posibilidad de eventos climáticos inesperados como se ha mostrado en años pasados.

En resumen, este trabajo presenta un enfoque completo para el análisis y modelado de series temporales climáticas, con la aplicación específica a la predicción de la temperatura media mensual. Este tipo de análisis es crucial para tomar decisiones informadas y planificar medidas de adaptación frente al cambio climático.