

Actividad 3: Series Temporales

Pablo Sánchez, Katherine Carbonell, Matías Dávila

Problema nº1

a)

Periodo dentro de la muestra: Se refiere al periodo para el que se dispone tanto de los valores reales (Y_t) como de las predicciones de ambos métodos (A y B). Según la tabla, el periodo de la muestra va de 2003 a 2009.

Para evaluar cuál de los métodos funciona mejor, normalmente nos fijamos en lo cerca que coinciden las predicciones (tanto del método A como del método B) con los valores reales (Y_t). Entre las métricas de error más utilizadas en el análisis predictivo se incluyen:

- Error medio absoluto (MAE): Mide la diferencia absoluta media entre los valores predichos y los valores reales.
- Error cuadrático medio (ECM): Se centra en la diferencia al cuadrado entre los valores predichos y los reales, dando más peso a los errores mayores.
- Error cuadrático medio (RMSE): Es la raíz cuadrada del MSE, proporcionando una medida en la misma unidad que los valores reales.

Para el periodo de la muestra (2003-2009), las métricas de error para cada método son:

- Error medio absoluto (MAE):
 - Método A: 0.66
 - Método B: 1,41
- Error cuadrático medio (RMSE):
 - Método A 0.87
 - Método B: 1,71

Basándonos tanto en el MAE como en el RMSE, el Método A presenta valores de error inferiores en comparación con el Método B, lo que significa que el Método A proporciona predicciones más precisas durante el periodo de la muestra. Así, el método A muestra una mejor capacidad de predicción para el periodo analizado.

b)

El periodo extra-muestra va de 2010 a 2012. Calculemos ahora las métricas de error (MAE y RMSE) para este periodo.

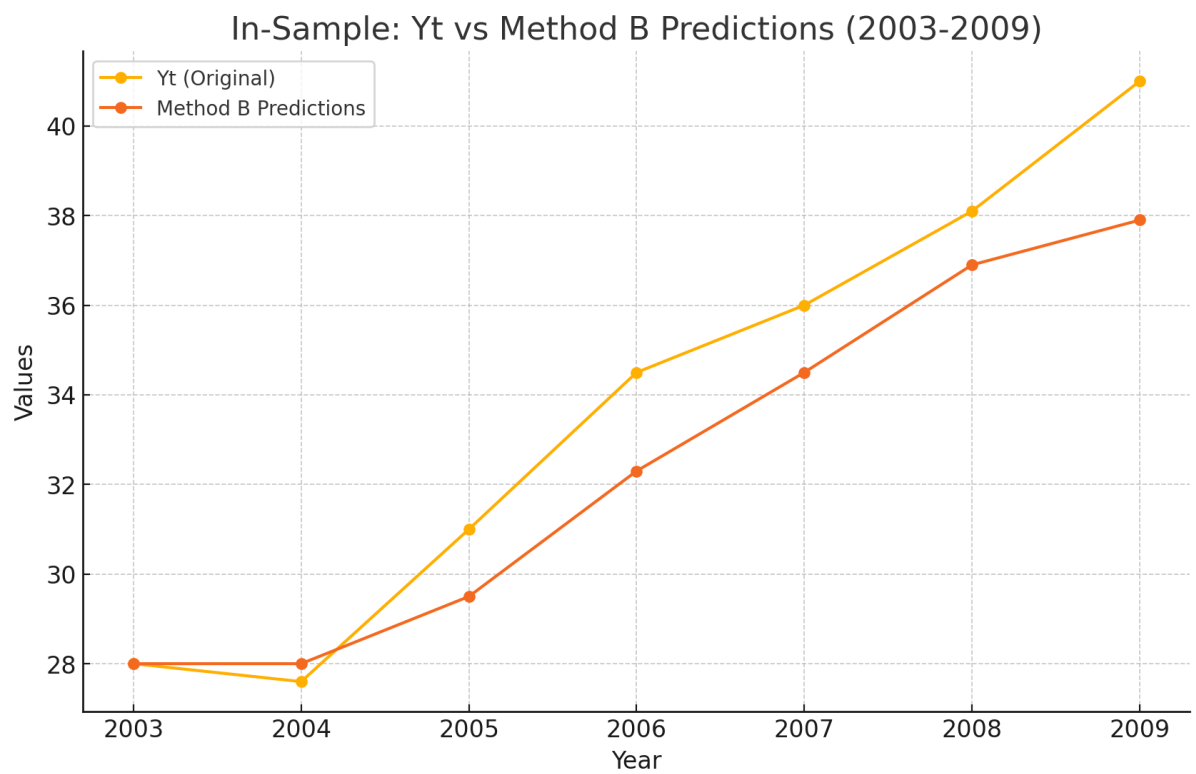
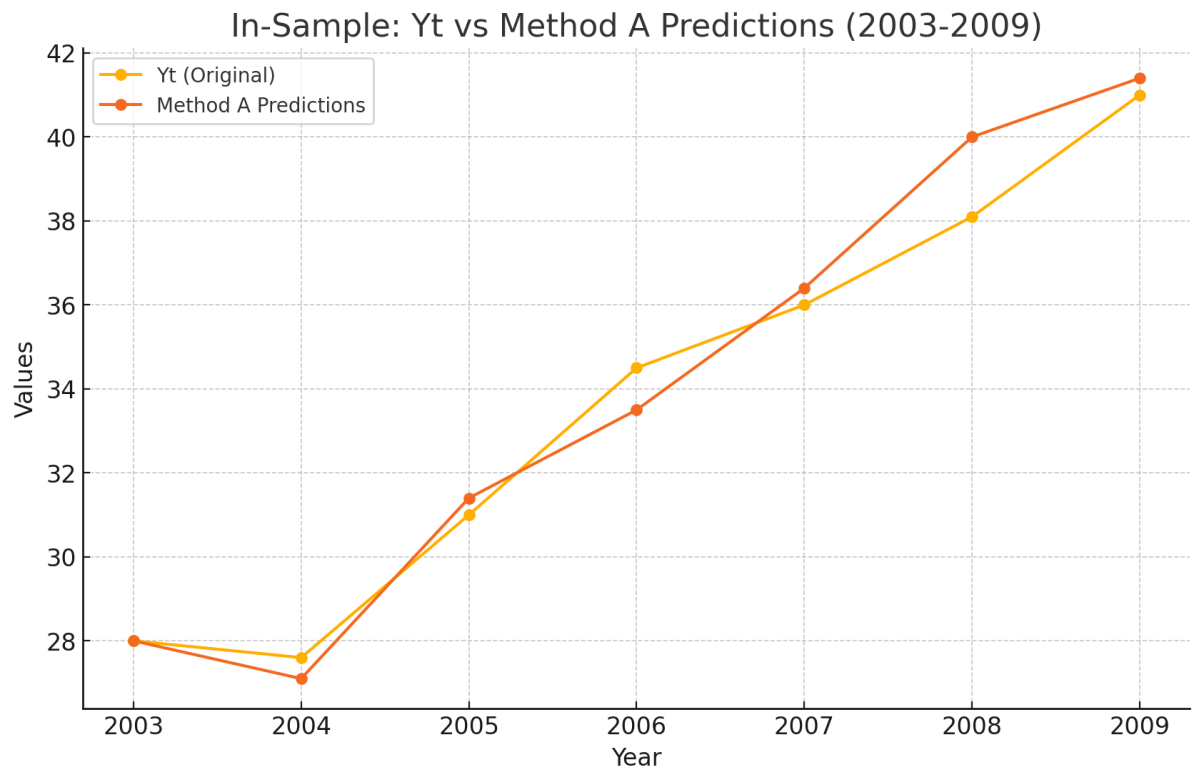
Para el periodo extra-muestra (2010-2012), las métricas de error para cada método son:

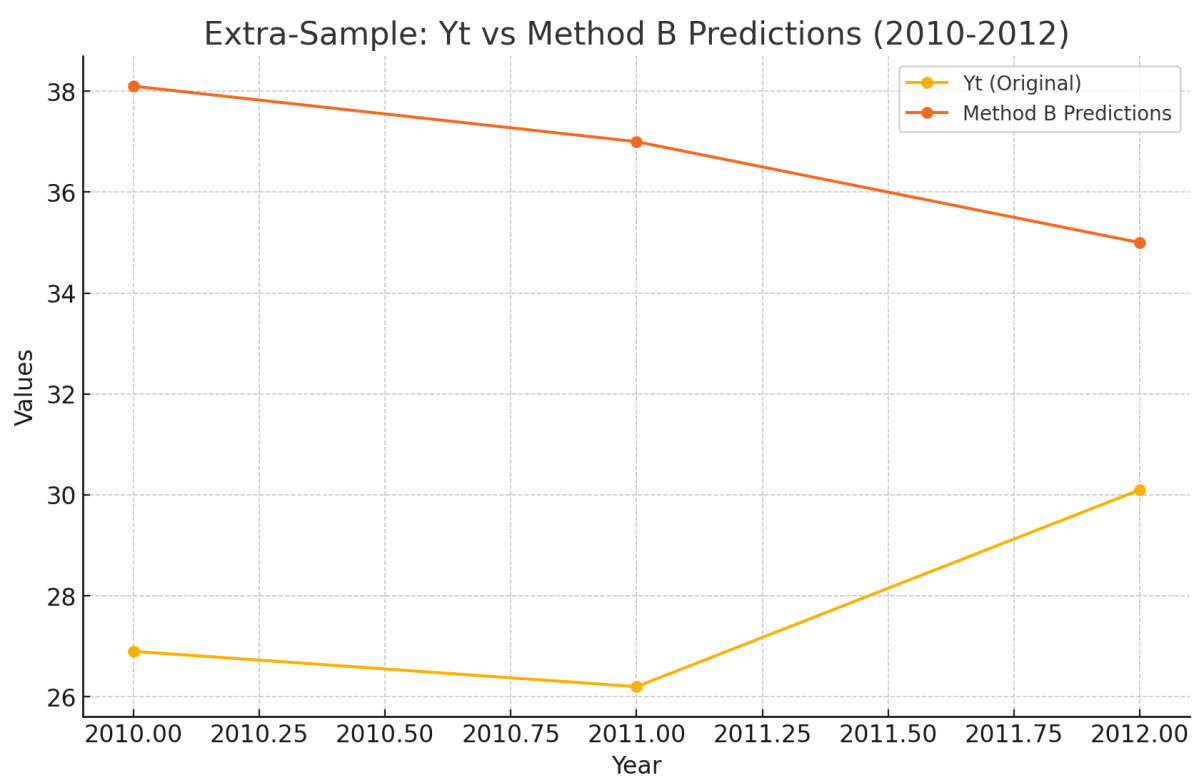
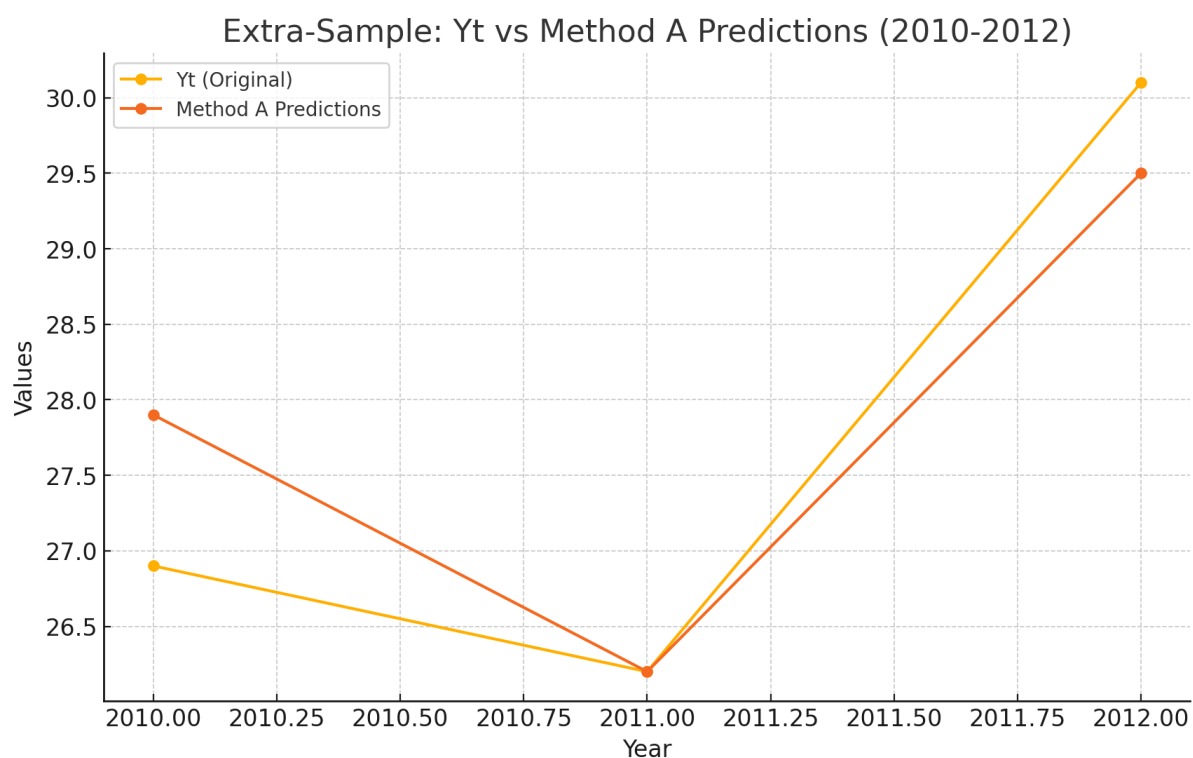
- Error medio absoluto (MAE):
 - Método A: 0.53
 - Método B: 8,97
- Error cuadrático medio (RMSE):
 - Método A 0.67
 - Método B: 9,42

En este periodo, el Método A sigue superando al Método B por un margen significativo, mostrando errores mucho menores. Las clasificaciones siguen siendo las mismas, siendo el Método A el mejor predictor tanto para el periodo dentro de la muestra como para el periodo fuera de la muestra.

Así pues, el método A tiene sistemáticamente una mejor capacidad de predicción en ambos periodos, mientras que el método B obtiene malos resultados, especialmente en el periodo extra-muestra.

c)





Los gráficos ilustran la comparación entre los valores reales (Y_t) y las predicciones del método A y el método B, tanto para el periodo dentro de la muestra (2003-2009) como para el periodo fuera de la muestra (2010-2012):

1. Periodo dentro de la muestra (2003-2009):

- Las predicciones del método A siguen de cerca los valores reales, con pequeñas desviaciones.
- Las predicciones del método B se sitúan sistemáticamente por debajo de los valores reales, lo que indica un rendimiento inferior.

2. Periodo extra-muestra (2010-2012):

- El método A sigue ajustándose a los valores reales, con pequeños errores de predicción.
- El método B muestra grandes desviaciones respecto a los valores reales, especialmente para los años 2010-2012, lo que indica una escasa capacidad de predicción.

Conclusiones:

- En ambos periodos, el Método A proporciona un mejor rendimiento predictivo, mientras que el Método B muestra errores significativos, particularmente en el periodo extra-muestra.
- Las clasificaciones predictivas se mantienen constantes en ambos periodos, siendo el método A el superior.

Problema nº2

a)

Las predicciones utilizando el Suavizado Exponencial (AES) son las siguientes:

- AES(0,3):
 - Año 1: 80,2
 - Año 2: 80,2
 - Año 3: 79,54
 - Año 4: 77,94
 - Año 5: 77,51
 - Año 6: 77,77

- AES(0,6):
 - Año 1: 80,2
 - Año 2: 80,2
 - Año 3: 78,88
 - Año 4: 76,07
 - Año 5: 76,33
 - 6º año: 77,57

Aquí tienes los errores para el método de Suavización Exponencial (AES):

- AES(0.3):
 - Período muestral (primeros 4 años):
 - MAE: 2.24
 - RMSE: 2.98
 - Período extra-muestral (últimos 2 años):
 - MAE: 3.03
 - RMSE: 3.71
- AES(0.6):
 - Período muestral:
 - MAE: 1.83
 - RMSE: 2.59
 - Período extra-muestral:
 - MAE: 3.52
 - RMSE: 3.81

En el período muestral, el método AES(0.6) tiene mejores resultados (menores errores) en comparación con AES(0.3). Sin embargo, en el período extra-muestral, ambos métodos tienen errores similares, aunque AES(0.3) es ligeramente mejor.

b)

Las predicciones utilizando medias móviles (MM) son las siguientes:

- MM(2) (media móvil de 2 periodos):

- Año 3: 79,10
- Año 4: 76,10
- Año 5: 75,35
- Año 6: 77,45

- MM(3) (media móvil de 3 periodos):

- Año 4: 77,47
- Año 5: 76,23
- Año 6: 76,37

En cuanto a la parte (b), podemos observar que ambos métodos proporcionan predicciones relativamente similares, con pequeñas variaciones.

Aquí tienes los errores para el método de Media Móvil (MM):

- MM(2):

- Período muestral:
 - MAE: 2.65
 - RMSE: 3.48
- Período extra-muestral:
 - MAE: 3.95
 - RMSE: 4.05

- MM(3):

- Período muestral:
 - MAE: 0.97
 - RMSE: 0.97
- Período extra-muestral:
 - MAE: 2.97
 - RMSE: 3.07

En este caso, MM(3) tiene un rendimiento significativamente mejor en el período muestral (con errores más bajos) en comparación con MM(2). Además, en el período extra-muestral, MM(3) también es mejor que MM(2).

c)

Las predicciones con el método Naive y el método de la media simple son las siguientes:

- Método Naive (valor anterior como predicción):
 - Año 2: 80,2
 - Año 3: 78,0
 - Año 4: 74,2
 - Año 5: 76,5
 - Año 6: 78,4
- Método de la media simple (media de todas las observaciones anteriores):
 - Año 2: 80,2
 - Año 3: 79,10
 - Año 4: 77,47
 - Año 5: 77,23
 - Año 6: 77,46

Aquí los errores para el método ingenuo y el método de media simple:

- Método ingenuo:
 - Período muestral:
 - MAE: 2.77
 - RMSE: 2.86
 - Período extra-muestral:
 - MAE: 3.85
 - RMSE: 4.32
- Método de media simple:
 - Período muestral:
 - MAE: 2.69
 - RMSE: 3.15
 - Período extra-muestral:
 - MAE: 3.02
 - RMSE: 3.54

Entre el método ingenuo y el método de media simple, el método de media simple tiene mejor capacidad predictiva tanto en el período muestral como en el extra-muestral, mostrando consistentemente errores más bajos. Sin embargo, ambos métodos son menos precisos que los métodos basados en suavización exponencial (AES) o media móvil (MM), que proporcionan mejores resultados globalmente.

Problema nº3

1. Introducción: Contextualización de los datos

Vamos a trabajar con temperaturas de una ciudad real. Elijo una ciudad de las disponibles, por ejemplo, San Francisco, y describo el contexto climático para esta ciudad.

2. Aplicación empírica

a. Determinación del tipo de serie temporal

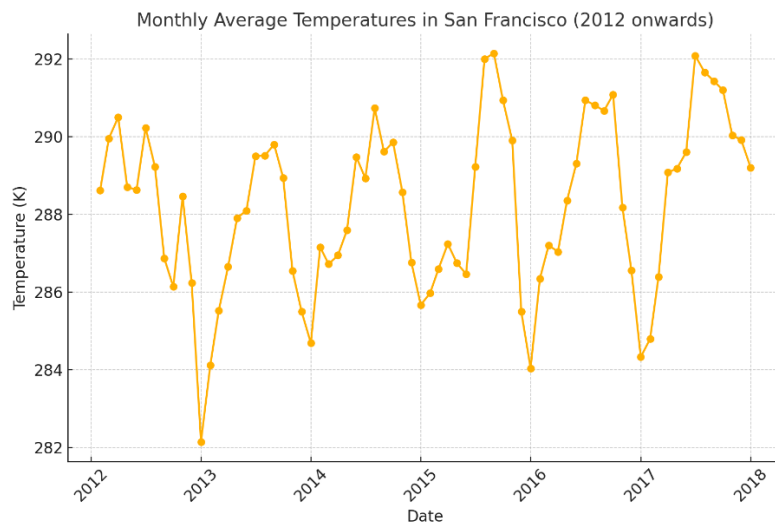
Para determinar el tipo de serie temporal, primero analizamos el gráfico de las temperaturas mensuales. El gráfico mostró una clara estacionalidad, con fluctuaciones anuales regulares en las temperaturas. Además, aplicamos la prueba de Dickey-Fuller para verificar si la serie era estacionaria o no.

Resultados de la prueba Dickey-Fuller:

- Estadístico de prueba: -4.016
- p-valor: 0.0013
- Valores críticos:
 - 1%: -3.529
 - 5%: -2.904
 - 10%: -2.590

Conclusión: Con base en estos resultados, podemos concluir que la serie es **estacionaria**, ya que el p-valor es menor que 0.05 y el estadístico de prueba es menor que los valores críticos. Esto indica que no es necesario realizar diferenciaciones adicionales para hacer la serie estacionaria.

b. Representación gráfica de la serie temporal



Este gráfico confirma la **estacionalidad** anual, con picos regulares que coinciden con los veranos y mínimos en los inviernos.

c. Contraste (Contraste de Daniel y de Kruskal-Wallis):

Resultados del contraste de Kruskal-Wallis:

- Estadístico de prueba: 39.71
- p-valor: 0.00004

Interpretación:

Dado que el p-valor es muy bajo (significativamente menor que 0.05), podemos rechazar la hipótesis nula de que las temperaturas de los diferentes meses provienen de la misma distribución. Esto sugiere que existen diferencias significativas entre las temperaturas de los distintos meses, lo que es coherente con la estacionalidad observada en la serie.

d. Conclusiones a partir de los contrastes

A partir de los resultados del contraste de Kruskal-Wallis, podemos concluir que las diferencias entre los meses son estadísticamente significativas. Esto refuerza la idea de que la serie tiene un comportamiento estacional, con fluctuaciones claramente diferenciadas según el mes.

e. Determinación del esquema de integración

Dado que la prueba de Dickey-Fuller indicó que la serie ya es estacionaria, no es necesario realizar diferenciaciones adicionales. Por lo tanto, no necesitamos aplicar ninguna transformación a la serie para fines de análisis.

Bibliografía

Dataset con el que hemos trabajado: [Temperature \(kaggle.com\)](https://www.kaggle.com/datasets/rajatprasad/temperature)