

TRABAJO PRÁCTICO 1

ACTIVIDAD 1

ANÁLISIS DE SERIES TEMPORALES UNIVERSIDAD DE BUENOS AIRES

Por: Sevann Radhak Triztan

Actividad 1:

Redactar un pequeño informe, justificando cada uno de los pasos que se plantean y las conclusiones que se obtienen.

- 1) Elegir una serie univariadas, pueden tomarla del repositorio GitHub [<https://github.com/jbrownlee/Datasets.git>] que se compartió en la clase 2, de otro repositorio que indiquen o series propias con las que ya trabajen. Describir todos los datos que considere relevantes para un análisis.
- 2) Descomponer dicha serie usando un modelo ETS aditivo, y comparar la tendencia que se obtiene con el filtro de Hodrick-Prescott y realizar modelos de suavizado exponencial de 1ero, 2do y tercer orden. Calcular las funciones ACF y PACF.
- 3) Hacer la autocorrelación para distintos valores de rezago, hacer el correlograma.

SOLUCIÓN

1. Selección y análisis de la serie

Para este análisis, se ha seleccionado la serie univariable de temperaturas máximas diarias en Canberra, Australia, extraída del dataset meteorológico "weatherAUS". La variable analizada es "MaxTemp", que representa las temperaturas máximas diarias registradas en un periodo de tiempo determinado.

Datos relevantes:

- La serie cubre varios años de datos con periodicidad diaria.
- Los datos de temperatura son continuos y se han aplicado técnicas de imputación para manejar los valores faltantes (relleno hacia adelante o "forward fill").
- La temperatura es una variable climatológica que puede exhibir patrones estacionales, tendencias y variaciones cíclicas.

2. Descomposición de la serie

a) Descomposición Error-Trend-Seasonality (ETS)

Utilizamos un modelo ETS aditivo para descomponer la serie en sus tres componentes principales:

- **Tendencia:** refleja cambios a largo plazo en los niveles de la serie.
- **Estacionalidad:** representa patrones repetitivos a lo largo de períodos fijos.
- **Fluctuaciones irregulares:** variaciones aleatorias que no pueden explicarse por los otros componentes.

La descomposición ETS permite observar cómo la temperatura máxima diaria varía no solo por efectos estacionales, sino también por una tendencia general a lo largo del tiempo. El resultado de la descomposición muestra que la serie presenta una tendencia creciente y una fuerte componente estacional.

b) Filtro de Hodrick-Prescott

Se aplicó el filtro de Hodrick-Prescott (HP) para obtener la tendencia a partir de la serie de temperaturas. Se probaron diferentes valores del parámetro λ (100, 1600 y 14400) para evaluar la suavización de la tendencia:

- Para $\lambda = 100$, la tendencia sigue de cerca las fluctuaciones de la serie original.
- Para $\lambda = 1600$ y $\lambda = 14400$, la tendencia se suaviza progresivamente, capturando solo los cambios a largo plazo.

Comparación:

Al comparar la tendencia obtenida del modelo ETS aditivo con la proporcionada por el filtro de HP, se puede observar que ambos métodos capturan la tendencia de manera similar, pero el filtro HP es más flexible dependiendo del valor de λ utilizado (ver Figura 1). El modelo ETS ofrece una descomposición más clara y estructurada en términos de estacionalidad y error.

c) Suavizado exponencial

Se aplicaron tres variantes de suavizado exponencial para modelar la serie:

- Simple Exponential Smoothing (SES), ver Figura 2.
- Double Exponential Smoothing (DES), ver Figura 3.
- Triple Exponential Smoothing (TES), ver Figura 4.

Comparación:

El suavizado SES proporciona una línea suavizada que responde lentamente a cambios en la tendencia. DES es más sensible a la tendencia y proporciona una mejor representación para capturar fluctuaciones a corto plazo. TES, con componentes tanto aditivos como multiplicativos, permite modelar de manera efectiva las fluctuaciones estacionales, ajustándose mejor a series que presentan tanto estacionalidad como tendencia. En los gráficos comparativos, vemos que el suavizado simple sigue de cerca los datos, mientras que el suavizado doble y triple capturan mejor las fluctuaciones estacionales y de tendencia.

3. Funciones de autocorrelación y análisis de rezago

a. Función de autocorrelación (ACF):

Se observó que la autocorrelación disminuye lentamente, lo que sugiere una fuerte tendencia y posible estacionalidad en la serie (ver Figura 5).

b. Función de autocorrelación parcial (PACF):

La PACF sugiere una estacionalidad clara, con patrones repetidos en intervalos de tiempo fijos (ver Figura 6).

4. Estacionaridad

Se realizó la prueba de Dickey-Fuller aumentada (ADF) para evaluar la estacionaridad de la serie de tiempo MaxTemp. Los resultados obtenidos son los siguientes:

- Estadístico ADF: -3.51779
- Valor p: 0.00754
- Número de rezagos utilizados: 1
- Número de observaciones utilizadas: 3507

Valores críticos:

- Nivel de significancia del 1%: -3.4322
- Nivel de significancia del 5%: -2.8624
- Nivel de significancia del 10%: -2.5672
- Estadístico de información de la serie (AIC): 17588.40

Interpretación:

El valor p de la prueba ADF es 0.0075, que es menor que el nivel de significancia común de 0.05. Esto significa que podemos rechazar la hipótesis nula de que la serie es no estacionaria. Además, el estadístico ADF (-3.51779) es menor que los valores críticos para los niveles de 1%, 5% y 10%, lo que refuerza la evidencia de que la serie es estacionaria.

Conclusiones

- **Tendencia y estacionalidad:** el análisis confirma la existencia de una tendencia de largo plazo en la serie de temperaturas máximas diarias, junto con una estacionalidad clara.
- **Descomposición ETS:** el modelo ETS permite capturar bien tanto la tendencia como la estacionalidad.
- **Filtro Hodrick-Prescott:** el filtro HP proporciona una tendencia ajustable dependiendo del valor de lambda, siendo útil para comparar diferentes horizontes temporales.
- **Suavizado exponencial:** el suavizado exponencial de triple orden (TES) fue el más efectivo para capturar la tendencia y estacionalidad, lo que lo convierte en una opción adecuada para el pronóstico de datos meteorológicos.
- **Autocorrelación:** la serie presenta autocorrelación significativa en diferentes rezagos, lo que sugiere la necesidad de incluir componentes estacionales en cualquier modelo predictivo.

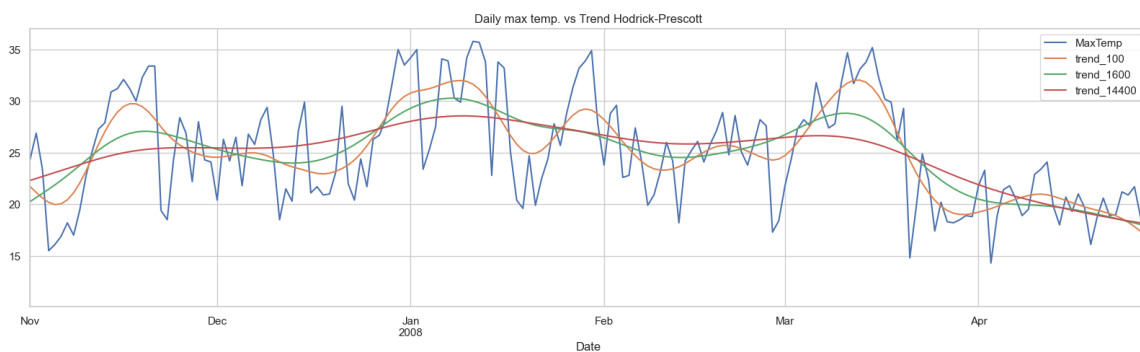


Figura 1: Daily max temp. vs Trend Hodrick-Prescott

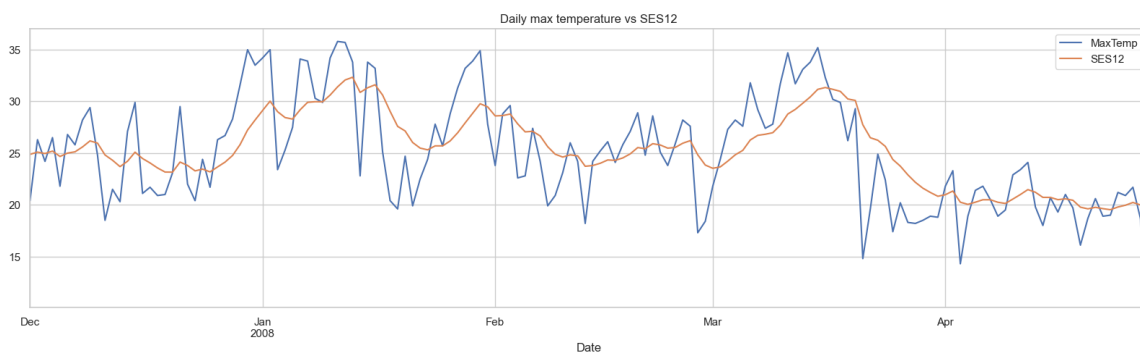


Figura 2: Daily max temperature vs SES12

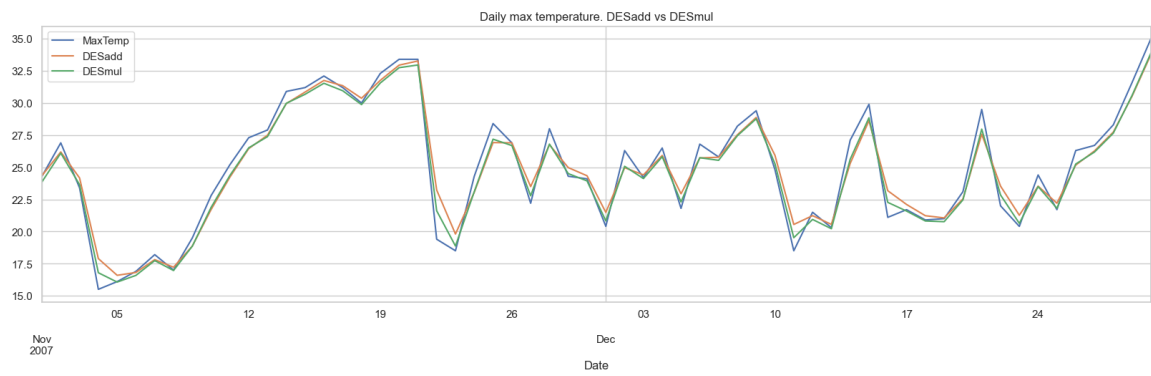


Figura 3: Daily max temperature. DESadd vs DESmul

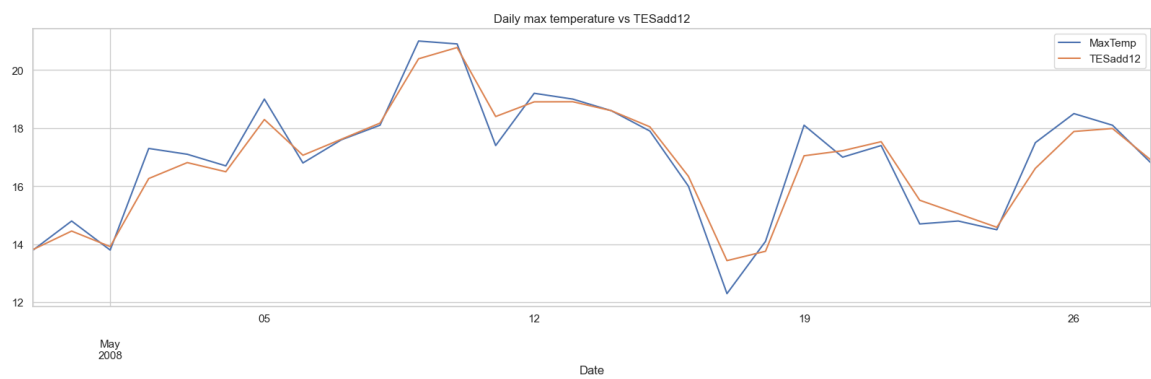


Figura 4: Daily max temperature vs TESadd12

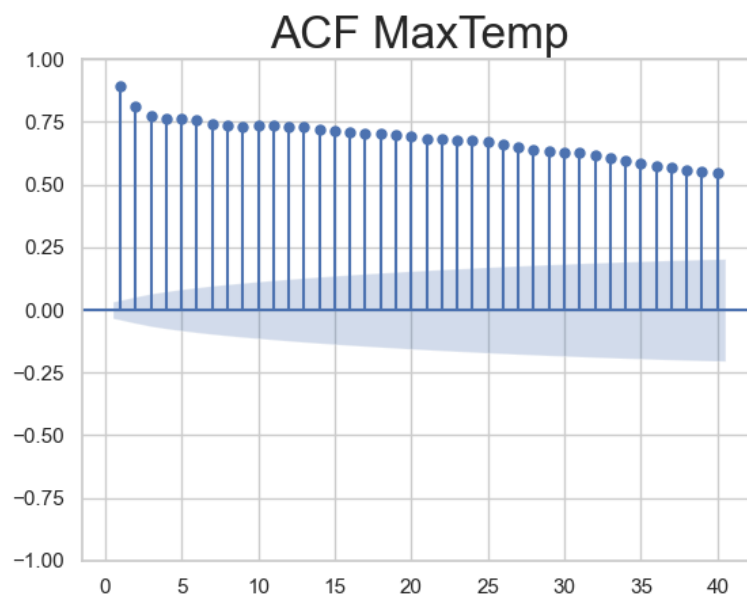


Figura 5: ACF MaxTemp

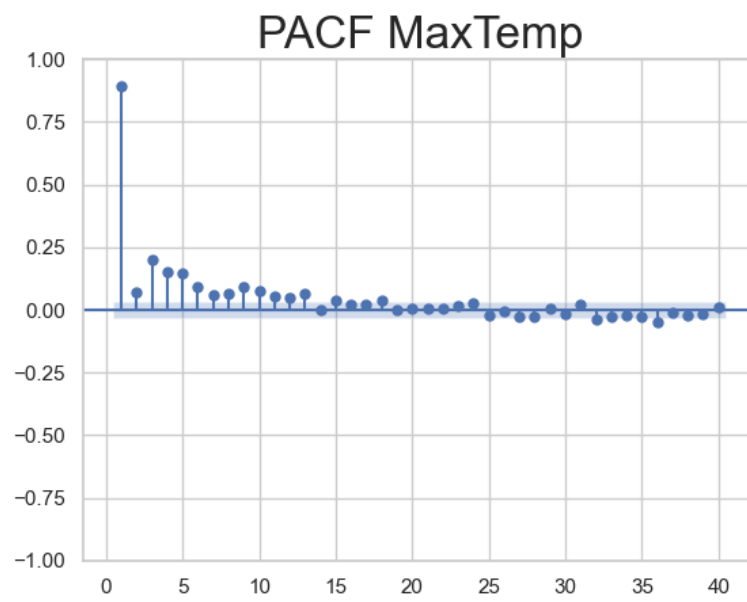


Figura 6: PACF MaxTemp