



INSTITUTO POLITECNICO NACIONAL

Escuela Superior de Cómputo
(ESCOM)

Licenciatura en Ciencias de Datos.

Nombre de la unidad de aprendizaje:

Sistemas de Información Geográfico.

Grupo: 6AM1.

Nombre de la Actividad:

“Tarea 11.- Extracción de señales”.

Nombre del alumno(a):

Arteaga Gonzalez Edwin Yahir.

Juárez Gaona Erick Rafael.

Rico Gaytán Diana Andrea.

Ruiz Merino Wendy Ivonne.

Fecha:

25/06/2025.

Introducción

El cambio climático se ha convertido en uno de los mayores desafíos que enfrenta la humanidad en el siglo XXI. Uno de los principales indicadores del impacto del cambio climático es el aumento de las temperaturas globales, que afecta tanto a la vida en el planeta como a los ecosistemas y las actividades humanas. Los datos de temperatura global permiten analizar cómo ha cambiado la temperatura promedio de la Tierra a lo largo del tiempo, ofreciendo una base fundamental para la toma de decisiones en políticas públicas y en la ciencia del clima.

En este análisis, se utiliza un dataset histórico de temperaturas globales, que abarca desde el año 1750 hasta la fecha, para estudiar cómo ha evolucionado la temperatura media global de la Tierra. El dataset incluye no solo las temperaturas medias de la tierra, sino también las temperaturas máximas y mínimas, así como las incertidumbres asociadas a las mediciones.

El objetivo principal de este reporte es aplicar modelos de predicción y filtros de serie de tiempo para analizar la serie temporal de temperaturas globales. A través de este análisis, se busca:

- Identificar las tendencias subyacentes en la serie de tiempo.
- Aplicar modelos estadísticos como ARIMA para predecir las temperaturas futuras.
- Implementar filtros como Wiener, EWMA, Kalman y Hodrick-Prescott para suavizar la serie temporal, eliminar el ruido y extraer las tendencias.

El enfoque Box-Jenkins se seguirá a lo largo del proceso para ajustar el modelo ARIMA y hacer pronósticos fiables basados en los datos históricos.

Problemática

El cambio climático es uno de los problemas más urgentes a nivel global, con implicaciones profundas para los ecosistemas, la salud humana y la economía. Entre los indicadores más críticos para evaluar el cambio climático se encuentra el aumento de las temperaturas globales. Las mediciones de temperatura global a lo largo del tiempo proporcionan una base fundamental para comprender cómo ha cambiado el clima de la Tierra y predecir los impactos futuros.

Sin embargo, los datos de temperatura pueden estar afectados por una variedad de factores, como fluctuaciones estacionales, errores de medición y ruido aleatorio en las observaciones. Estos ruidos pueden dificultar la identificación de patrones claros y las tendencias subyacentes que reflejan el cambio climático real. Además, los datos de temperatura, en muchos casos, incluyen incertidumbres asociadas con las mediciones y las estimaciones de temperatura, lo que añade un desafío adicional en su interpretación.

El análisis de series temporales de temperatura global debe ser capaz de eliminar el ruido presente en los datos, suavizar las fluctuaciones no significativas y extraer tendencias claras que permitan comprender la evolución de las temperaturas globales a lo largo del tiempo. El reto radica en utilizar herramientas estadísticas y computacionales adecuadas para filtrar los datos, identificar modelos de predicción adecuados y proporcionar un marco fiable para la toma de decisiones sobre políticas ambientales y la prevención del cambio climático.

Este reporte tiene como objetivo aplicar modelos de serie de tiempo y filtros de suavizado que permitan obtener predicciones fiables de la temperatura futura y detectar las tendencias a largo plazo, mitigando así los efectos del ruido y las fluctuaciones.

Modelo Estadístico

El análisis de series temporales es una herramienta fundamental para estudiar datos que están correlacionados en el tiempo, como es el caso de las temperaturas globales. Para abordar este tipo de datos, se emplea el modelo ARIMA (AutoRegressive Integrated Moving Average), que es uno de los más utilizados en el análisis de series temporales con tendencias y estacionalidades.

El modelo ARIMA es adecuado cuando los datos no son estacionarios, es decir, cuando muestran una tendencia o estacionalidad que debe ser modelada antes de hacer predicciones. Este modelo se define mediante tres parámetros:

- **p (AutoRegresivo - AR):** El número de términos autoregresivos en el modelo. Estos términos representan la relación entre una observación y las observaciones pasadas.
- **d (Diferenciación - I):** El número de veces que la serie temporal debe diferenciarse para hacerla estacionaria, es decir, para eliminar la tendencia a largo plazo.

- q (Media Móvil - MA): El número de términos de media móvil que modelan el error de las observaciones pasadas.

Modelo Computacional

El modelo computacional que se utiliza en este análisis está basado en el proceso de modelado de series temporales, específicamente el modelo ARIMA (AutoRegressive Integrated Moving Average), que es adecuado para predecir series temporales no estacionarias con tendencias y estacionalidades. Este modelo es ampliamente utilizado para el análisis de series temporales que no siguen una estructura predecible y exhiben tendencias complejas.

Además del modelo ARIMA, se aplicarán filtros de serie temporal para suavizar la serie y extraer las tendencias subyacentes, minimizando el ruido presente en los datos. Los filtros aplicados son:

1. Filtro de Wiener: Para reducir el ruido blanco.
2. Filtro EWMA (Exponential Weighted Moving Average): Para dar más peso a las observaciones recientes.
3. Filtro de Kalman: Para obtener una estimación continua de la serie temporal.
4. Filtro Hodrick-Prescott: Para separar la tendencia y el ciclo de los datos.

Metodología: Enfoque Box-Jenkins

El enfoque Box-Jenkins es un procedimiento sistemático para identificar, estimar y diagnosticar modelos de series temporales ARIMA. Se basa en tres pasos clave:

1. Identificación del modelo: Determinar los parámetros adecuados del modelo ARIMA utilizando herramientas como ACF (Autocorrelation Function) y PACF (Partial Autocorrelation Function).
2. Estimación del modelo: Ajustar el modelo ARIMA a los datos de la serie temporal y obtener los parámetros óptimos.
3. Diagnóstico del modelo: Evaluar los residuos del modelo para asegurarse de que no quedan patrones no capturados y verificar que los residuos son ruido blanco.

1. Identificación del Modelo:

En la identificación del modelo, el objetivo es elegir los valores adecuados de los parámetros p , d y q , que corresponden a:

- p : El número de términos autorregresivos (AR).
- d : El número de diferencias necesarias para hacer la serie estacionaria.
- q : El número de términos de media móvil (MA).

Para identificar los valores de estos parámetros, se utilizan las gráficas de ACF y PACF.

2. Estimación del Modelo ARIMA:

Una vez que hemos identificado los valores de p , d y q , procedemos a ajustar el modelo ARIMA utilizando los parámetros seleccionados. Esto se hace utilizando la función de ajuste de ARIMA en statsmodels.

3. Diagnóstico del Modelo:

Después de ajustar el modelo ARIMA, es esencial verificar si el modelo es adecuado para los datos. Para ello, debemos analizar los residuos del modelo. Los residuos son las diferencias entre los valores observados y los valores predichos por el modelo. Si los residuos muestran autocorrelación significativa, significa que el modelo no ha capturado toda la estructura de los datos.

Interpretación de los Resultados:

- Si ACF de los residuos muestra que los residuos son cercanos a cero en todos los lags, entonces los residuos son ruido blanco, lo que indica que el modelo ARIMA ha capturado bien la estructura de la serie temporal.
- Si los residuos muestran autocorrelación significativa, se debe revisar el modelo y ajustar los parámetros.

Propuesta de solución para el uso de filtros de serie de tiempo óptimos

En el análisis de series temporales, especialmente cuando se trabaja con datos que contienen fluctuaciones estacionales, errores de medición o ruido aleatorio (como las temperaturas globales), es fundamental aplicar filtros adecuados para suavizar las series

y extraer las tendencias subyacentes. Los filtros de serie de tiempo optimizados permiten mejorar la precisión de las predicciones y las inferencias realizadas a partir de los datos.

A continuación, se presentan los filtros óptimos que se proponen para el tratamiento de la serie temporal de las temperaturas globales:

1. Filtro de Wiener

El filtro de Wiener es un filtro adaptativo que se utiliza para reducir el ruido en las series temporales. Es especialmente útil para eliminar ruido blanco (ruido aleatorio sin patrón), preservando la estructura de la señal subyacente.

- **Aplicación:** El filtro de Wiener es ideal cuando se tiene un ruido blanco que perturba la serie temporal pero no afecta significativamente a las tendencias principales. Es particularmente útil cuando se tienen datos con incertidumbres en las mediciones, como ocurre con las temperaturas globales.
- **Ventajas:** Permite suavizar los datos sin perder características importantes de la serie, como picos y valles significativos.

2. Filtro de Media Móvil Exponencialmente Ponderada (EWMA)

El filtro EWMA asigna más peso a los datos más recientes, lo que lo hace adecuado para series temporales con tendencias y donde las observaciones más recientes tienen mayor relevancia para las predicciones futuras.

- **Aplicación:** Este filtro es útil cuando se desea capturar tendencias recientes en los datos, y es particularmente efectivo cuando la serie está sujeta a cambios abruptos o variaciones recientes de gran relevancia.
- **Ventajas:** Proporciona un suavizado flexible en tiempo real que da mayor importancia a las observaciones más recientes sin perder la capacidad de modelar las tendencias.

3. Filtro de Kalman

El filtro de Kalman es un filtro recursivo que permite hacer estimaciones óptimas a partir de datos ruidosos, proporcionando un conjunto de estimaciones del estado de un sistema dinámico. Se utiliza ampliamente para hacer estimaciones de series temporales ruidosas.

- **Aplicación:** El filtro de Kalman es adecuado cuando los datos son ruidosos, pero se desea estimación continua de la serie temporal. En este caso, ayuda a suavizar la serie de temperaturas globales sin perder la precisión de los cambios continuos.
- **Ventajas:** Proporciona una estimación más precisa de la serie, considerando tanto el estado actual como las mediciones previas y la incertidumbre.

4. Filtro Hodrick-Prescott

El filtro Hodrick-Prescott es utilizado para descomponer una serie temporal en dos componentes: la tendencia y el ciclo. Es ampliamente utilizado en economía y análisis de series temporales de largo plazo.

- **Aplicación:** Es ideal para separar las fluctuaciones cíclicas (cortos cambios) de las tendencias a largo plazo en los datos. En el caso de las temperaturas globales, este filtro puede ayudar a separar las variaciones estacionales de las tendencias subyacentes.
- **Ventajas:** Permite un análisis más detallado de las tendencias a largo plazo, lo que es útil para la predicción futura de datos que exhiben cambios cíclicos.

Conclusiones por Integrante

Arteaga González Edwin Yahir.

En este análisis, mi enfoque principal fue aplicar y ajustar el modelo ARIMA para capturar las tendencias de las temperaturas globales a lo largo del tiempo. A través de la identificación de los parámetros p , d y q , logré ajustar un modelo ARIMA(5, 1, 0), el cual fue capaz de proporcionar pronósticos razonables para las temperaturas futuras. Sin embargo, los residuos del modelo mostraron que aún hay autocorrelación, lo que indica que algunos patrones importantes no fueron capturados completamente.

- **Fortalezas:** El modelo ARIMA funcionó bien para capturar la tendencia general y proporcionar pronósticos a futuro, pero la precisión del modelo puede mejorar con un ajuste de los parámetros.
- **Áreas de mejora:** Se requiere un ajuste más fino del modelo para eliminar la autocorrelación en los residuos y mejorar la predicción a largo plazo.

Juárez Gaona Erick Rafael.

Mi principal tarea fue evaluar la aplicación del filtro de Wiener, que es especialmente útil para eliminar el ruido blanco en la serie temporal de las temperaturas globales. Este filtro fue eficaz para suavizar las fluctuaciones aleatorias y permitir una visualización más clara de las tendencias subyacentes. Sin embargo, el filtro de Wiener suavizó demasiado las fluctuaciones, lo que puede haber resultado en la pérdida de detalles importantes en la serie temporal, como los picos de temperatura.

- **Fortalezas:** El filtro de Wiener fue útil para reducir el ruido sin eliminar las tendencias generales, proporcionando una visión más limpia de la serie.
- **Áreas de mejora:** El filtro debe usarse con cautela, ya que puede suavizar demasiado y perder detalles importantes de la serie temporal.

Rico Gaytán Diana Andrea.

En mi parte del proyecto, trabajé con el filtro EWMA (Exponential Weighted Moving Average), que da más peso a los datos más recientes. Este filtro fue muy útil para capturar las tendencias recientes de las temperaturas globales. Sin embargo, debido a su enfoque en las observaciones más recientes, el filtro no logró captar bien las fluctuaciones más lejanamente relacionadas con el comportamiento de la serie a largo plazo. Esto puede ser útil para datos más cortoplacistas, pero limita su efectividad cuando se analizan tendencias de largo plazo.

- **Fortalezas:** El filtro EWMA permitió capturar rápidamente las tendencias recientes, lo que es ideal para escenarios donde se desean predicciones de corto plazo.
- **Áreas de mejora:** No es adecuado para capturar fluctuaciones a largo plazo, por lo que debe combinarse con otros métodos para obtener una visión más completa de la serie temporal.

Ruiz Merino Wendy Ivonne.

Mi contribución al proyecto fue evaluar los filtros de Kalman y Hodrick-Prescott, que son muy efectivos para trabajar con series temporales ruidosas y separar las componentes de tendencia y ciclo. El filtro de Kalman proporcionó estimaciones precisas de las temperaturas globales incluso en presencia de ruido, mientras que el filtro Hodrick-Prescott permitió separar claramente las fluctuaciones cíclicas de la tendencia.

subyacente. Este último es especialmente útil para comprender las variaciones a largo plazo en los datos.

- **Fortalezas:** Los filtros de Kalman y Hodrick-Prescott fueron excelentes para proporcionar estimaciones continuas y separar las tendencias cíclicas de las tendencias a largo plazo.
- **Áreas de mejora:** Aunque ambos filtros fueron eficaces, deben usarse en combinación con otros filtros o modelos para obtener una perspectiva completa del comportamiento de la serie temporal.

Conclusión General

El análisis de la serie temporal de temperaturas globales utilizando técnicas avanzadas de modelado y filtrado ha permitido extraer valiosas tendencias y realizar pronósticos sobre las temperaturas futuras. A través de la aplicación del enfoque Box-Jenkins con el modelo ARIMA y varios filtros de serie temporal (Wiener, EWMA, Kalman, y Hodrick-Prescott), hemos logrado una comprensión más clara de las variaciones a largo plazo y las fluctuaciones estacionales en los datos de temperatura.

El modelo ARIMA, aunque eficaz para capturar la tendencia general de las temperaturas, mostró la necesidad de ajustes finos debido a los residuos que indican cierta autocorrelación no capturada. Los filtros aplicados (Wiener, EWMA, Kalman y Hodrick-Prescott) han demostrado ser complementarios al modelo ARIMA, proporcionando una forma de suavizar los datos y mejorar las estimaciones.

- El filtro de Wiener fue efectivo para reducir el ruido blanco, aunque suavizó en exceso algunas fluctuaciones, lo que podría no ser deseable para mantener detalles finos en los datos.
- El filtro EWMA resultó útil para capturar las tendencias recientes, pero su enfoque en las observaciones más cercanas limita su capacidad para modelar variaciones a largo plazo.
- El filtro de Kalman proporcionó una estimación precisa de la serie temporal en presencia de ruido, mostrando su capacidad para realizar un seguimiento continuo de los datos.
- El filtro Hodrick-Prescott permitió separar las componentes cíclicas de las tendencias subyacentes, lo que es útil para entender mejor las fluctuaciones a

largo plazo de las temperaturas globales.

En conjunto, los resultados obtenidos demuestran que una combinación de modelos ARIMA con filtros de serie temporal es una herramienta poderosa para predecir las tendencias futuras y suavizar las fluctuaciones en los datos de temperatura. La aplicación de estos métodos proporciona una mejor comprensión del cambio climático y de cómo las temperaturas globales evolucionarán en el futuro, lo que resulta crucial para la toma de decisiones en políticas ambientales y el desarrollo de estrategias para mitigar los efectos del calentamiento global.

Este análisis también resalta la importancia de seleccionar el modelo y filtro adecuado según el comportamiento de los datos, y sugiere que futuras investigaciones pueden centrarse en ajustar aún más los modelos y explorar otros métodos para mejorar la precisión de las predicciones.