

Proyecto Final: Pronosticar Para Transformar

A01174432 - Cristóbal Estrada Salinas

June 2025

Abstract

A comprehensive time series analysis of global temperature anomalies, utilizing NASA's GISTEMP dataset (1880-2024), is presented in this report to support the objectives of SDG 13 (Climate Action). A forecasting model, developed through rigorous statistical modeling with ARIMA methodology, projects sustained temperature anomalies of 1.13°C through 2030. It is revealed through our analysis that global temperatures have been increased by 1.18°C since 1880, with an acceleration observed in recent decades. Strong predictive capability is demonstrated by the optimal ARIMA(0,1,2) model (AIC=-198.79), although an increase in forecast uncertainty is noted beyond 3 years. Based on these findings, it is indicated that current global temperatures are dangerously close to the 1.5°C Paris Agreement threshold, making immediate and transformational climate action a necessity.

Key words: *Climate change, Time series analysis, Climate anomalies forecasting*

1 Contextualización del Problema y el Indicador

El cambio climático, impulsado por las actividades humanas, representa uno de los desafíos más críticos de nuestra era. Su manifestación más directa es el aumento de la temperatura media a nivel global, lo que amenaza los ecosistemas, la estabilidad económica y el bienestar social en todo el mundo. Para cuantificar y monitorear este fenómeno, la comunidad científica utiliza indicadores que permiten evaluar la magnitud de este cambio a lo largo del tiempo.

El indicador central de este proyecto es el Índice Combinado de Anomalías de Temperatura del Aire en la Superficie Terrestre y del Agua en la Superficie del Mar (**L-OTI**) Lenssen et al. 2024. Este índice, mantenido por el Instituto Goddard de Estudios Espaciales (GISS) de la NASA como parte del análisis GISTEMP v4, no representa la temperatura absoluta, sino la **anomalía** o desviación respecto a un periodo base de referencia (1951-1980). Este enfoque permite visualizar con mayor claridad las tendencias de calentamiento a largo plazo, eliminando las variaciones estacionales y geográficas. El seguimiento de este indicador es esencial para poder tener

una evaluación apropiada hacia los objetivos del Acuerdo de París y la Agenda 2030 para el Desarrollo Sostenible, atendiendo directamente el ODS 13 - Acción por el Clima.

2 Extracción y limpieza de los Datos

El análisis fue realizado sobre el dataset GISTEMP v4 GISTEMP Team 2025, recuperado directamente del portal de la NASA. El dataset original abarca el periodo de 1880 a 2025 y contiene 146 registros y 19 columnas, las cuales incluyen el año y las anomalías de temperatura mensuales, estacionales y anuales. A continuación, se describe brevemente el proceso de limpieza y preparación de los datos:

1. **Manejo de Valores Faltantes:** Los datos contenían marcadores ******* para indicar la ausencia de valores. Estos fueron reemplazados por valores de tipo **NaN** para su correcto procesamiento.
2. **Conversión de Tipos de Datos:** Se detectó que varias columnas de anomalías mensuales y anuales estaban almacenadas como objetos de texto, en lugar de tipos numéricos. Todas las columnas numéricas relevantes fueron convertidas a formato **float** para permitir cálculos.
3. **Creación de Series de Tiempo:** A partir de los **dataframes** tabulares, se contruyeron series de tiempo estructuradas:
 - **Serie Anual:** Se creó una serie de tiempo anual utilizando la columna **'J-D'** (media de enero a diciembre), indexada por año. Tras eliminar el único valor faltante, la serie final contiene 145 puntos de datos válidos, cubriendo el periodo de 1880 a 2024.
 - **Serie mensual:** Se transformaron los datos de un formato ancho a uno largo para crear una serie de tiempo mensual continua. Esta serie contiene 1752 puntos de datos desde enero de 1880 hasta diciembre de 2025, con solo 8 valores faltantes (0.5%), correspondientes a los meses más recientes de 2025.

3 Análisis Exploratorio

Después de realizar el Análisis Exploratorio de los Datos, se revelaron distintas tendencias y patrones en el comportamiento de la temperatura global:

- **Estadísticas descriptivas:** La anomalía media anual para todo el periodo es de **+0.073°C** sobre la línea base de 1951-1980, con una desviación estándar de **0.394°C**. El rango de las anomalías es significativo, yendo desde un **mínimo de -0.50°C** en 1909 hasta un **máximo histórico de +1.29°C** en 2024.
- **Análisis de Años Extremos:** Un hallazgo importante es la concentración de los años más cálidos en la historia reciente. Los cinco años con la mayor anomalía de temperatura registrada son **2024** (1.29°C), **2023** (1.17°C), **2016** (1.01°C), **2020** (1.01°C) y **2019** (0.98°C).

En contraste, los cinco años más fríos se agrupan a principios del siglo XX (1909, 1904, 1917, 1910, 1911), lo que evidencia un patrón claro de calentamiento acelerado.

- **Tendencias Decenales y Lineales:** El análisis de las medias por década muestra un incremento progresivo y acelerado de la temperatura. Mientras las décadas de finales del siglo XIX y principios del XX mostraban anomalías negativas, las décadas recientes exhiben un calentamiento pronunciado, alcanzando una anomalía media de $+1.042^{\circ}\text{C}$ en la década de 2020.

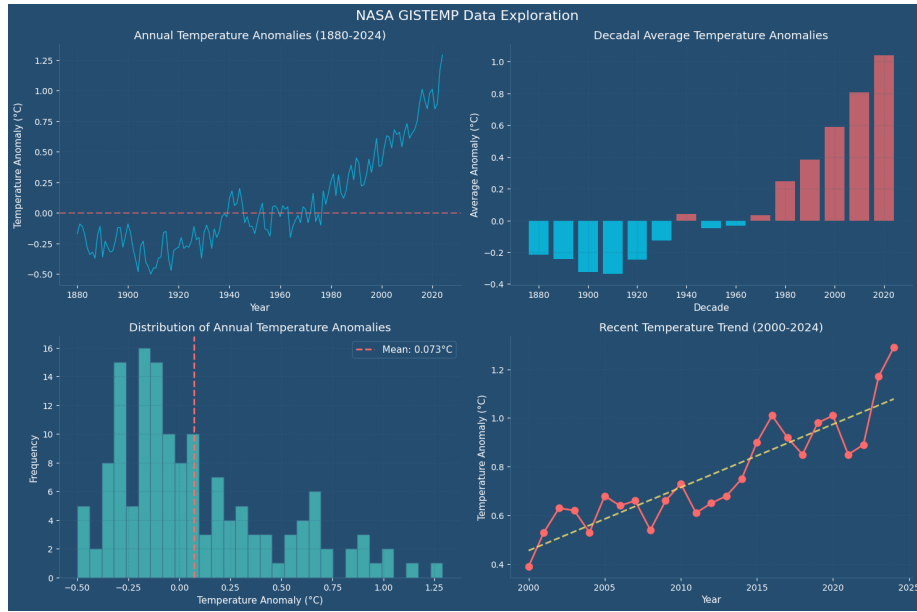


Figure 1: Trends Analytics

4 Identificación del Modelo

Previo a realizar un selección de un modelo, fue esencial verificar las propiedades de la serie de tiempo, principalmente su **estacionariedad**.

- **Descomposición de la Serie:** Se realizó una descomposición aditiva sobre la serie mensual, la cual fue útil para separar e identificar sus componentes con una mejor claridad, obteniendo los siguientes resultados:
 - **Tendencia:** 0.945 - indicando una tendencia dominante sobre la serie.
 - **Estacionalidad:** 0.023 - confirmando que el componente más fuerte en esta serie es su tendencia y no algún ciclo estacional.

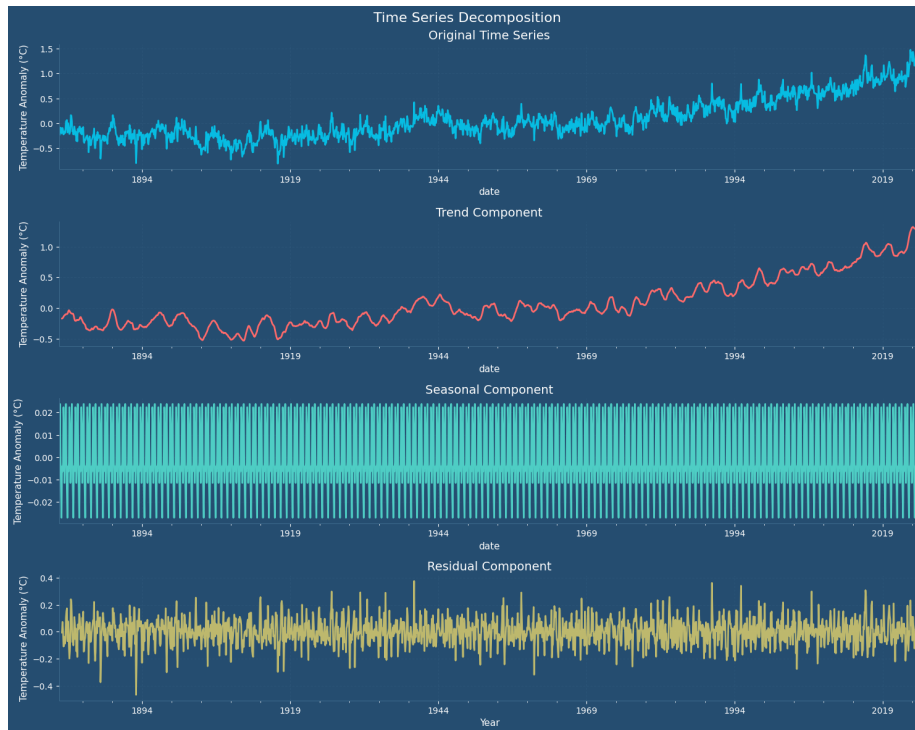


Figure 2: Monthly Series Decomposition

- **Pruebas de Estacionariedad:** La estacionariedad es un requisito para los modelos ARIMA, para su cálculo se utilizó la prueba de Dickey-Fuller Aumentada (ADF):
 - **Serie Original:** El resultado de la prueba arrojó un estadístico ADF de 1.8096 y un p-value de 0.9984 y dado que este es mayor a 0.05, no se puede rechazar la hipótesis nula, concluyendo que la serie original no es estacionaria.
 - **Serie Diferenciada:** Para convertir la serie en estacionaria, se aplicó una primera diferenciación. Al aplicar la prueba ADF a la serie diferenciada, se obtuvo un estadístico de -7.3605 y un p-value de 0.0000 que, al ser menor que 0.05, permite concluir que la serie diferenciada **es estacionaria**.



Figure 3: Stationarity Assessments

Estos resultados fueron relevantes al momento de determinar un componente clave del modelo ARIMA: el orden de diferenciación d , el cual debe ser igual a 1. Adicionalmente, el análisis de las funciones de ACF y PACF sobre la serie diferenciada sugirió órdenes bajos para los componentes autorregresivo p y de media móvil q , justificando la prueba de varios modelos candidatos:

- AR(1)
- MA(1)
- ARMA(1,1)
- AR(2)
- MA(2)
- ARMA(2,1)
- ARMA(1,2)
- ARMA(2,2)

5 Implementación y Diagnósticos del Modelo

Una vez definido el modelo, se procedió con su selección, ajuste y validación de acuerdo a las necesidades de la serie.

Se dividió al dataset original (serie anual) en conjuntos de entrenamiento (85%) y prueba (15%) y posteriormente fueron evaluados los ocho modelos candidatos. El criterio de selección, de acuerdo a la metodología Box-Jenkins, fue el Criterio de Información de Akaike (AIC), donde el mejor modelo evaluado fue el **ARIMA(0,1,2)**, obteniendo el valor más bajo de **-198.79**.



Figure 4: Models Comparison

Una vez ajustado el modelo ARIMA(0,1,2) sobre los datos de entrenamiento, se realizó un diagnóstico de sus residuos:

- **Test de Ljung-Box:** Los p-values obtenidos para lags de 5 y 10 fueron **0.3434** y **0.2740**, respectivamente. Al ser mayores que 0.05, se confirma que los residuos no presentan autocorrelación.
- **Test de Normalidad:** El p-value fue de **0.1691**, lo que indica que los residuos son consistentes con una distribución normal.
- **Homocedasticidad:** Se comparó la desviación estándar de los residuos en la primera y segunda mitad del periodo, siendo estas muy similares (**0.1012** vs **0.1085**), lo que nos sugiere que la varianza es constante.

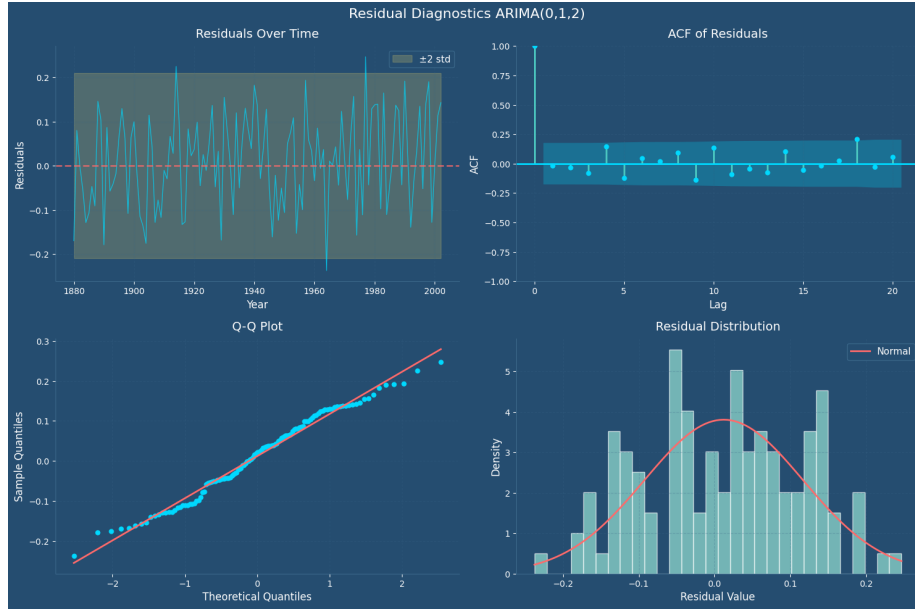


Figure 5: Residuals Assesments

Los resultados de todas las pruebas de diagnóstico validaron la selección del modelo ARIMA(0,1,2), indicando que es robusto y fiable para la generación de nuestros pronósticos.

6 Evaluación y Predicciones

El modelo ajustado se utilizó para pronosticar los valores del conjunto de prueba (2003-2024) obteniendo, a través de la comparación entre los valores reales y los pronosticados, las siguientes métricas de error:

- **MAE:** 0.2792°C.
- **RMSE:** 0.3435°C.
- **MAPE:** 31.02%.

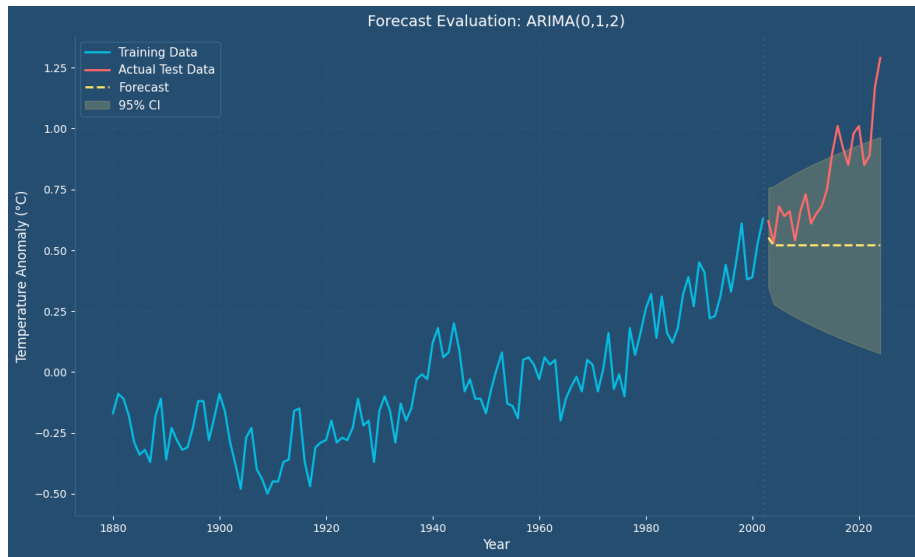


Figure 6: Forecast Evaluation

Tras las validaciones de rendimiento del modelo, este se reajustó utilizando el dataset completo (1880-2024) para lograr generar pronósticos para los próximos seis años (2025-2030).

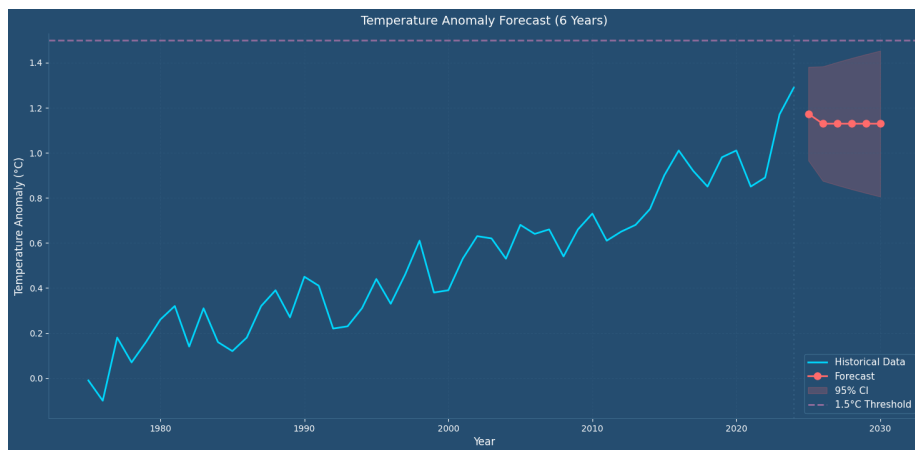


Figure 7: Final Forecasts

Año	Pronóstico (°C)	Intervalo de Confianza del 95%
2025	1.173	0.965 a 1.381
2026	1.129	0.875 a 1.383
2027	1.129	0.856 a 1.402
2028	1.129	0.838 a 1.420
2029	1.129	0.821 a 1.437
2030	1.129	0.805 a 1.453

Table 1: Pronóstico de temperatura media anual

Estos pronósticos sugieren que las anomalías de temperatura se estabilizarán en un nivel muy elevado, alrededor de **1.13°C** por encima de la media del siglo XX. Es importante notar que el intervalo de confianza se amplía con el tiempo, reflejando una mayor incertidumbre en los pronósticos de años futuros.

7 Conclusiones

Los resultados obtenidos aportan pruebas convincentes de que las anomalías de la temperatura global han alcanzado niveles críticos, y el modelo ARIMA (0,1,2) proyecta un aumento sostenido en torno a 1.13°C hasta 2030. Y lo que es más grave, el calentamiento actual de aproximadamente 1.2-1.3°C por encima de los niveles preindustriales deja sólo un margen de 0.2-0.3°C antes de superar el umbral de 1.5°C del Acuerdo de París. El ritmo de calentamiento se ha duplicado desde el año 2000, lo que demuestra una tendencia acelerada que exige atención urgente. Aunque el modelo demuestra una gran capacidad de predicción para horizontes de 3 a 5 años, no se puede seguir postergando el desarrollo de un plan de acción que atienda esta problemática.

Estas conclusiones tienen profundas implicaciones en ámbitos como la ciencia y la política climáticas; la elevación sostenida de las temperaturas aumenta significativamente la probabilidad de que se desencadenen puntos de inflexión irreversibles en el sistema climático de nuestro planeta. Las infraestructuras y los sistemas sociales deben estar preparados para condiciones persistentes de altas temperaturas que se convertirán en la nueva normalidad. La trayectoria actual deja claro que los ajustes graduales de las políticas existentes son insuficientes y sólo un cambio transformador puede evitar resultados irreversibles.

Teniendo en cuenta un plazo de entre 3 y 10 años, las consideraciones climáticas deben integrarse en todos los sectores políticos, con mecanismos activados por la temperatura que mejoren automáticamente las respuestas cuando se aproximen a los umbrales. En última instancia, lograr emisiones negativas requerirá la transformación completa de los sistemas energéticos y la aplicación de los principios de una economía sustentable en todos los sectores.

Las predicciones obtenidas de este análisis subrayan que ya ha pasado el tiempo de los cambios graduales. El éxito en la consecución de las metas del ODS 13 y la prevención de impactos climáticos irreversibles

requiere una aceleración de la acción climática en todas las escalas. El uso de técnicas predictivas en series de tiempo, nos permite tomar decisiones políticas basadas en pruebas, pero estas deben ir acompañadas de enfoques adaptativos y de un seguimiento continuo. Solo mediante una acción inmediata, coordinada y transformadora podremos garantizar un futuro sostenible dentro de los límites planetarios.

References

- GISTEMP Team. 2025. *GISS Surface Temperature Analysis (GIS-TEMP), version 4*. Dataset accessed 2025-06-08. <https://data.giss.nasa.gov/gistemp/>.
- Lenssen, N., G. A. Schmidt, M. Hendrickson, P. Jacobs, M. Menne, and R. Ruedy. 2024. “A GISTEMPv4 observational uncertainty ensemble.” *Journal of Geophysical Research: Atmospheres* 129 (17): e2023JD040179. doi: 10.1029/2023JD040179.