Surface temperature analysis: variability and trends

Author: Marcel Pagès Ticó, mpagesti7@alumnes.ub.edu
Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Catalonia, Spain.

Advisor: Yolanda Sola Salvatierra, ysola@ub.edu

Abstract: This work analyses global surface temperature anomalies from 1880 to 2024 using the GISS database. ARIMA and Bayesian Structural Time Series models are employed to make predictions. Both approaches reveal a strong warming trend predicting, respectively, an average annual increase of 0.021 K and 0.031 K until 2060. Higher warming over land and, especially in the Bayesian model, Arctic amplification are forecasted. These findings align with IPCC projections and underscore the urgency of climate mitigation strategies.
Keywords: Climate change, Trends, Predictions, ARIMA, Bayesian Structural Time Series.
SDGs: Climate action, Terrestrial life.

I. INTRODUCTION

It is undoubtedly the case that, over recent decades, temperatures have not stopped rising. Even more troubling is the fact that, according to the Intergovernmental Panel on Climate Change (IPCC) projections, in every likelihood they will continue to increase in the next years [1].

This study aims to explore global temperature variability and long-term trends using historical data. By applying time series modeling techniques, the project investigates both temporal and spatial temperature dynamics.

II. DATA AND METHODOLOGY

A. Database description

In order to carry out this study, data provided by the Goddard Institute for Space Studies (GISS) are used [2]. The database contains monthly running mean surface temperature anomalies (A) on a $2° \times 2°$ grid, starting in 1880 and ending in 2024. These anomalies are calculated as a weighted average of nearby stations within 1,200 km, with weights decreasing linearly from unity (at the given location) to zero. Extrapolation by assuming spatial continuity is applied when no station is within 1,200 km. The GISS publication includes data from over 6,300 meteorological stations, while satellite measurements are employed for oceanic areas. To express temperatures of month m and year y as an anomaly, the transformation $A_{m,y} = T_{m,y} - \overline{T}_m^{(1951–1980)}$ is used for every grid point, where $\overline{T}_m^{(1951\ –1980)}$ is the average temperature for month m over the baseline period 1951–1980. Even though temperatures are not in absolute terms, it may be useful to consider that the 1951–1980 global mean surface air temperature is estimated at 287.15 K (14 °C).

This project's analysis is conducted using the R software [3]. The first step consists of importing the corresponding netCDF file, available at the GISS webpage [2]. Some missing values (10.9%) may be present, particularly at extreme latitudes and in the earliest records. For simplicity, these anomalies are removed. Hence, 25,104,721 complete observations are available to conduct the analysis.

Throughout the entire period of study, the global mean temperature anomaly has been of 0.13 K. The warmest complete year to date has been 2024 while the coldest 1909, with global mean temperatures of 288.75 K (15.6 °C) and 286.65 K (13.5 °C), respectively.

B. Time series models

Time series analysis encompasses various statistical models designed to capture patterns and dependencies within data. Two of the most important are as follows.

1. *Autoregressive integrated moving average*

The Autoregressive Integrated Moving Average is a powerful, efficient and widely used statistical model for analyzing and forecasting time series data [4]. It is denoted as ARIMA(p, d, q), where the parameters p, d, and q correspond respectively to the autoregressive order, the degree of differencing, and the moving average order.

In essence, the ARIMA model captures three key dynamics in a time series. First, the autoregressive (AR) component of order p models the dependency between an observation and a number of its previous values. This is expressed as a linear combination of lagged observations:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \epsilon_t, \quad (1)$$

where $\phi_1, \ldots, \phi_p$ are the autoregressive parameters and ϵ_t is a white noise error term with zero mean and constant variance.

Second, the moving average (MA) component of order q incorporates the influence of past forecast errors on the current value, formulated as:

$$X_t = \epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \cdots + \theta_q\epsilon_{t-q}, \tag{2}$$

where $\theta_1, \ldots, \theta_q$ are the moving average parameters.

When combined, these two components form the ARMA(p, q) model:

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t + \theta_1\epsilon_{t-1} + \cdots + \theta_q\epsilon_{t-q}. \tag{3}$$

However, many time series are not stationary -that is, their statistical properties such as mean and variance change over time-. To address this, differencing is applied d times until stationarity is achieved. Using the backshift operator B defined by $BX_t = X_{t-1}$, the differenced series is represented as $Y_t = (1 - B)^d X_t$, where Y_t is assumed to be stationary. Consequently, the full ARIMA(p, d, q) model can be expressed as:

$$Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \epsilon_t^{(Y)} + \theta_1\epsilon_{t-1}^{(Y)} + \cdots + \theta_q\epsilon_{t-q}^{(Y)}, \tag{4}$$

Determining the appropriate d typically involves formal statistical tests for stationarity, such as the augmented Dickey–Fuller test, based on the regression equation:

$$\Delta X_t = \gamma X_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta X_{t-i} + \epsilon_t, \tag{5}$$

where $\Delta X_t = X_t - X_{t-1}$ represents the first difference of the series, γ is the parameter being tested for stationarity, p denotes the number of lagged difference terms included to account for serial correlation, α captures the influence of past differences and ϵ_t is white noise. Thus, the hypotheses to test are:

- H_0: $\gamma = 0$ (non-stationary series)
- H_1: $\gamma < 0$ (stationary series)

Once stationarity is ensured, the AR and MA terms capture the structure of the transformed series.

Given a time series model, it is possible to predict observations at time $t+1$ by replacing in the ARIMA equation data until t. Hence, predictions for $t+2$ can be made using the existing data at t and the prediction for $t+1$, and successively. It has to be taken into account that the further one prediction is, the more error it has.

2. Bayesian Structural Time Series

While ARIMA models are popular due to their simplicity and computational efficiency, they present several limitations. They require stationarity, often demanding transformations, and provide point forecasts with limited uncertainty quantification. In contrast, Bayesian time series models offer a probabilistic framework that incorporates prior knowledge and yields full posterior distributions over latent states and future observations, allowing for more robust forecasts and coherent interval estimates [5]. However, this model is by no means as efficient as the ARIMA, requiring high computation times.

Given a time series $\mathbf{y} = (y_1, \ldots, y_T)$, Bayesian inference treats parameters $\mathbf{\Omega}$ (such as variances or regression coefficients) and latent states $\mathbf{x}_t$ (e.g., underlying trend or seasonality) as random variables. The posterior distribution is derived via Bayes' theorem:

$$p(\mathbf{\Omega}, \mathbf{x}_{1:T} \mid \mathbf{y}_{1:T}) = \frac{p(\mathbf{y}_{1:T} \mid \mathbf{x}_{1:T}, \mathbf{\Omega})\, p(\mathbf{x}_{1:T} \mid \mathbf{\Omega})\, p(\mathbf{\Omega})}{p(\mathbf{y}_{1:T})}. \tag{6}$$

Given the observed data $\mathbf{y}_{1:T}$, posterior inference is typically performed using Markov Chain Monte Carlo (MCMC) methods. Forecasting future values y_{T+h} involves integrating over posterior samples. In practice, this posterior is approximated through MCMC sampling as implemented in the `bsts` package in R, where the initial latent state $\mathbf{x}_1$ is modeled as multivariate normal with mean vector $\mathbf{m}_0 = \mathbf{0}$ and diagonal covariance matrix $\mathbf{C}_0 = \sigma^2\mathbf{I}$, reflecting large initial uncertainty.

Forecasts are then obtained by sampling from the predictive distribution:

$$p(y_{T+h} \mid \mathbf{y}_{1:T}) \approx \frac{1}{S}\sum_{s=1}^{S} p(y_{T+h} \mid \mathbf{\Omega}^{(s)}, \mathbf{x}_{T+h}^{(s)}), \tag{7}$$

where $(\mathbf{\Omega}^{(s)}, \mathbf{x}_{T+h}^{(s)})$ are draws from the posterior distribution, providing full predictive intervals and accounting for both parameter and state uncertainty.

III. FULL TEMPERATURE EVOLUTION

The observed global temperature anomaly data from 1880 to 2024 exhibit distinct phases of variation (Figure 1). Initially, from 1880 to 1910, there is a relatively stable period with a slight but not statistically significant cooling (the Mann-Kendall test -which assumes as a null hypothesis that there is no monotonic trend- presents a p-value of 0.16). This could be attributed to significant volcanic activity, such as the eruption of Krakatoa (1883), minimal solar activity and natural variability [6]. Between 1910 and 1940, a noticeable warming trend emerges (Mann-Kendall p-value of $1 \cdot 10^{-7}$), possibly linked to increased solar activity and reduced volcanic aerosols. However, from 1940 to 1980, temperatures exhibit stagnation (test p-value of 0.34, confirming there is no trend), likely due to high sulfate aerosol emissions from industrial activities, which reflect solar radiation. From 1980 onwards, a sharp and consistent rise in global temperature anomalies is evident (p-value of $3 \cdot 10^{-13}$), driven by anthropogenic greenhouse gas emissions, particularly CO_2 from fossil fuel combustion [1].

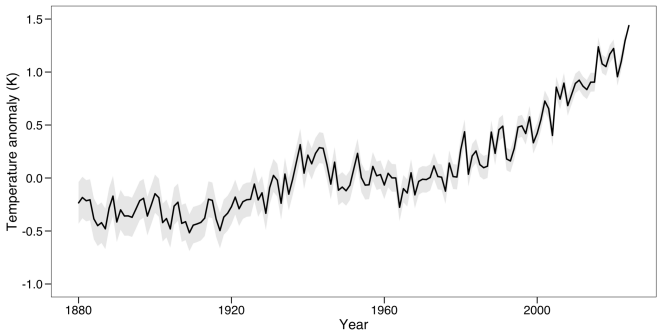

Figure 1: Evolution of annual mean global temperature anomaly, the baseline period being 1951-1980. Shaded areas represent uncertainty.

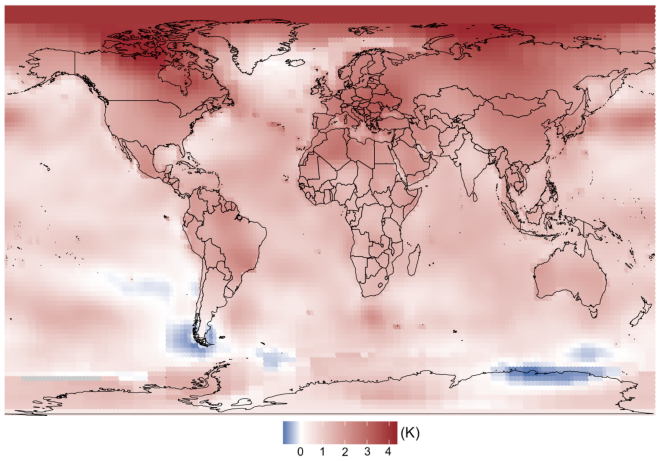

Figure 2: 2024 global temperature anomalies using 1951-1980 as baseline period.

It may also be of interest to analyze how temperatures have evolved in different regions. Temperature anomalies have increased most significantly in high northern latitudes, particularly in the Arctic, Northern Canada, and Siberia (Figure 2). This pattern, known as Arctic amplification, is primarily driven by feedback mechanisms such as ice-albedo reduction, lapse rate increase and greater landmass sensitivity to radiative forcing (difference between absorbed solar radiation and outgoing longwave radiation at tropopause level). Conversely, equatorial regions and parts of the Southern Ocean have experienced comparatively modest warming. These differences arise due to variations in ocean heat capacity, atmospheric circulation patterns, and regional feedbacks. The relatively stable temperatures in these zones are influenced by persistent cloud cover and efficient heat redistribution [7]. This is precisely why interannual fluctuations are high in polar and subpolar regions (Figure 3), particularly in the Arctic, where warming is also intense (Figure 2). Interestingly, while Antarctica shows only modest long-term warming, it is revealed that some Antarctic regions are subject to substantial interannual variability. This suggests that even in the absence of strong long-term trends,

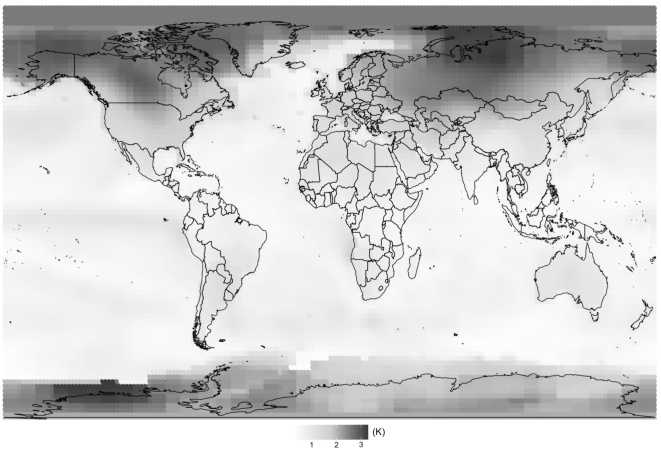

Figure 3: Standard deviation σ with monthly available data since 1880 to capture variability in each region.

local climate processes -such as katabatic winds and sea ice dynamics- can lead to significant short-term fluctuations [8].

IV. TIME SERIES MODELS IMPLEMENTATION

A. Autoregressive integrated moving average

Focusing on the time series of mean global temperature anomaly (Figure 1), it is clear that there is no stationarity because the mean is not constant as it might present a deterministic trend. To make sure, the augmented Dickey-Fuller test is used, giving back $\gamma = 0.10$ with an associated p-value of 0.99, so the series is not stationary. Hence, the series is differenced and the same test is applied, with H_0 being that the series is $d \geq 2$ and H_1 that $d = 1$. Now $\gamma = -7.36$ with an associated p-value smaller than 0.01. In consequence, the one-time differenced series is stationary, $d = 1$.

To determine p and q, the first step consists of plotting the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the differenced time series (Figure 4). It is clear that the ACF is significant until the first lag, while the PACF has a sinusoidal shape, which is an indicator of a MA(1) process [4]. Taking into account all these considerations, the model could arguably be described as an ARIMA(0,1,1). The residuals are constant in mean around 0 and they exhibit neither autocorrelation nor partial autocorrelation (Figure 5). Furthermore, they follow a normal distribution (Shapiro-Wilk normality test gives back a p-value of 0.63 confirming the null hypothesis of normality), which is also a sign that the model might be correct. Notwithstanding, there may be many other models that could adjust properly to this time series. The R function **auto.arima()** included in the **forecast** package returns the best model in terms of the Akaike Information Criteria after testing a great

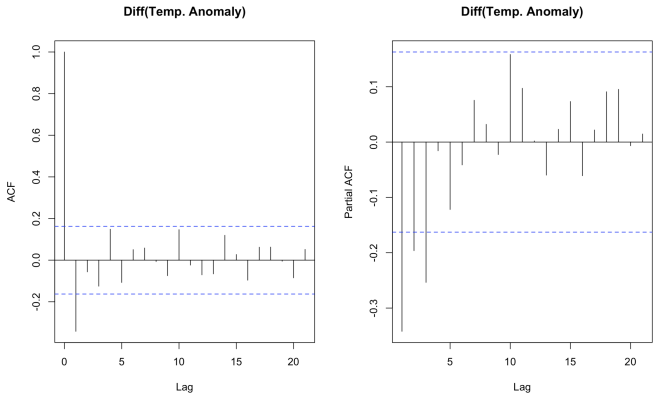

Figure 4: Autocorrelation function and partial autocorrelation fuction of differenced time series.

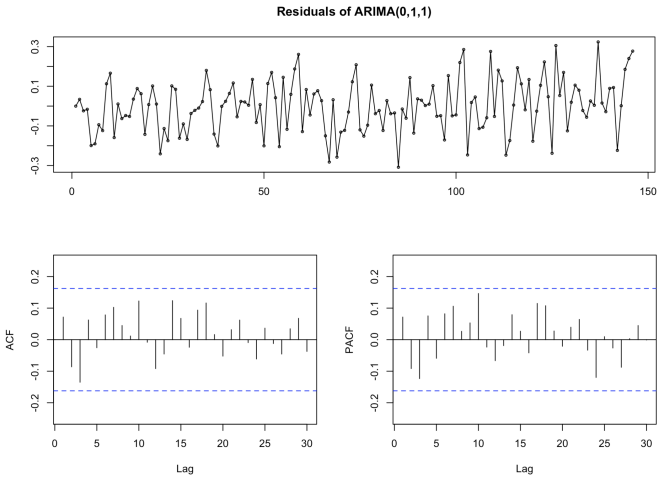

Figure 5: Residuals of ARIMA(0,1,1) model.

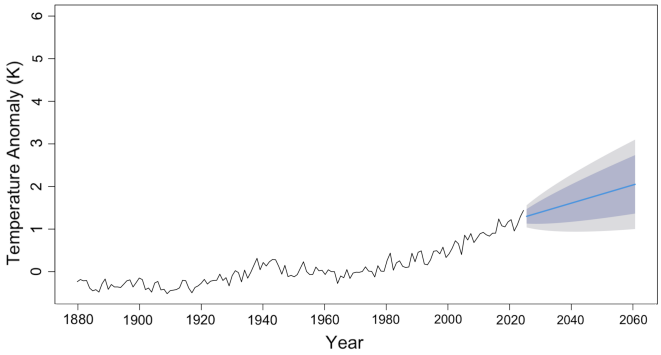

Figure 6: ARIMA's predictions of temperature anomaly from 2025 to 2060. Shaded areas indicate forecast uncertainty: the inner band 80% and the outer one 95%.

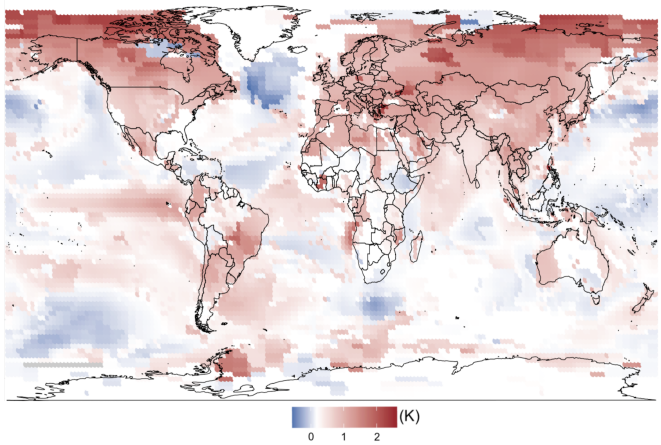

Figure 7: ARIMA's predictions of temperature anomaly difference between 2050 and 2024.

number of combinations of ARIMA models. The chosen model is again the ARIMA(0,1,1).

By using this model, the temperature anomaly until 2060 is predicted (Figure 6). According to this model, in the years to follow the temperature is expected to experience an upward deterministic trend, with a mean global increase of approximately 0.021 K per year.

At this point it may be useful to know how temperatures are expected to evolve in spatial terms. Given the fact that **auto.arima()** is an automated function, a time series with the available data for every location ($2° \times 2°$) is given as input. Therefore, for every pair of longitude and latitude values an ARIMA model is estimated and, with it, predictions are made (Figure 7). It should be emphasized that the validity of the chosen models is not checked because the objective is to gain an understanding of how temperatures are expected to evolve in each region; a comprehensive analysis of each specific location is not intended. A remarkable fact about this map is that the temperatures are expected to increase substantially more in land zones -especially in the northern hemisphere- than across the sea.

B. Bayesian Structural Time Series Forecasting

At this time the aforementioned `bsts` package with 100,000 MCMC draws is implemented to the mean global temperature series. The model predicts from 2025 to 2060 an average increase of 0.031 K per year (Figure 8).

As in the previous section, a model is applied for every location and the expected value for every prediction is extracted (Figure 9). The map suggests an increase of temperatures in most parts of the globe, with northern and terrestrial zones being in the lead.

V. DISCUSSION

The IPCC published on its Sixth Assessment Report (2021) different scenarios of future global temperatures -ranked from 1, best, to 5, worst- depending on Shared Socioeconomic Pathways (SSP), which are linked to the expected radiative forcing levels [1].

The predicted values with the ARIMA model for the mean global temperature are close to the intermediate

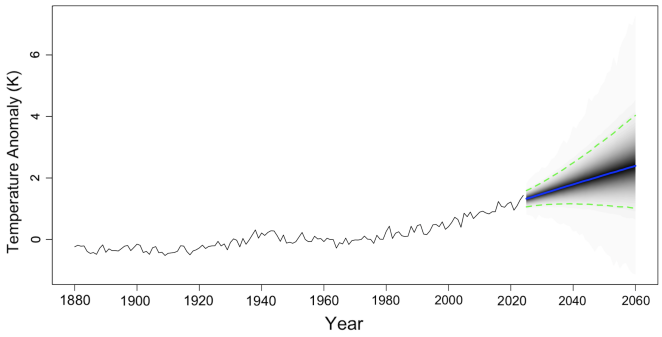

Figure 8: Bayesian predictions of temperature anomaly from 2025 to 2060. The blue line indicates the expected value, and the green band indicates uncertainty of 95%.

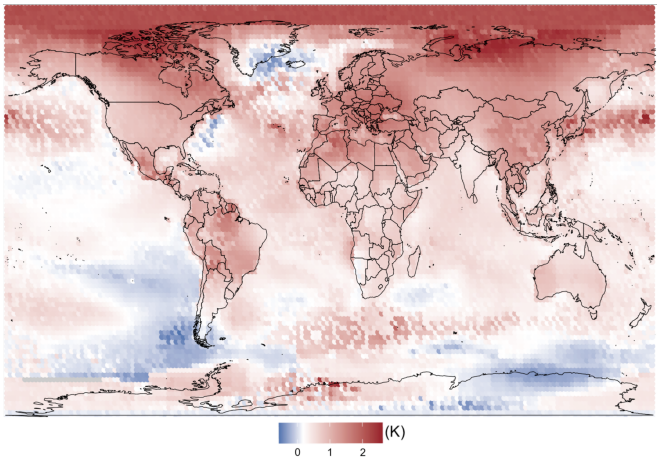

Figure 9: Bayesian predictions of temperature anomaly difference between 2050 and 2024.

scenario, SSP2 (radiative forcing of 4.5 W/m^2), which in the period of study forecasts an increase with an average of approximately 0.023 K per year. On the other hand, the Bayesian Structural Time Series model forecasts more intense warming, resonating with the SSP3 (radiative forcing of 7.0 W/m^2) that predicts an average increase of about 0.028 K per year. Since the Bayesian model yields full posterior distributions over future observations, it is of particular relevance to observe that the predictions that fall within the 95% bear resemblance, in shape and value, to SSP1 and SSP5, respectively (note that the IPCC does not assign any explicit probability to the different SSP). Therefore, it would be fair to say that the posterior distributions given by this model seem to be realistic and in accordance with the IPCC projections.

With regard to spatial predictions, both the IPCC and the ARIMA and Bayesian models agree that continental zones will get hotter than oceanic areas, which could be due to the lower heat capacity of land surfaces, reduced evaporative cooling, and limited thermal redistribution compared to oceans [1]. This being said, the ARIMA predictions seem to underestimate the expected increment in temperatures in various regions, especially in the Artic, while the Bayesian results apparently capture considerably better the projections made by the IPCC.

VI. CONCLUSIONS

In light of the above, there is compelling evidence of global warming, particularly pronounced in the northern hemisphere. Furthermore, like the IPCC, both analyzed time series models suggest temperatures will continue to rise, especially in continental areas.

This underscores the critical importance of immediate and sustained global efforts to mitigate climate change. Without decisive action, the projected increment in temperatures could lead to irreversible impacts on ecosystems, human health, and socioeconomic systems. The evidence calls not for hesitation, but for a unified and science-based response to an escalating global challenge.

Acknowledgments

I would like to thank my tutor Yolanda for her invaluable advice, as well as to my family and friends for their support. I also thank GISS for providing the dataset.

[1] Intergovernmental Panel on Climate Change (IPCC). *Climate Change 2021: The Physical Science Basis*. Contribution of Working Group I to the Sixth Assessment Report of IPCC. Cambridge University Press, 2021. `https://www.ipcc.ch/report/ar6/wg1/`, accessed March 2025.

[2] NASA Goddard Institute for Space Studies (GISS). *GISS Surface Temperature Analysis (GISTEMP)*, `https://data.giss.nasa.gov/gistemp/`, accessed February 2025.

[3] R Core Team. *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2023. `https://www.R-project.org/`

[4] Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

[5] Scott, S. L. (2014). Predicting the present with Bayesian structural time series. *International Journal of Mathematical Modelling and Numerical Optimisation, 5*(1–2), 4–23. `https://doi.org/10.1504/IJMMNO.2014.059942`

[6] Symons, G. (Ed.) (1888). *The Eruption of Krakatoa, and Subsequent Phenomena*, 494 pp., Trübner & Co., London.

[7] Serreze, M. C., & Barry, R. G. (2006) The Arctic amplification debate. *Climatic Change*, 76(3–4), 241–264. `https://doi.org/10.1007/s10584-005-9017-y`

[8] Holland, D. M., & Kwok, R. (1991). Evidence for intensification of the East Antarctic katabatic wind system from satellite image analysis and numerical model experiments. *Journal of Climate*, *4*(4), 619–629.

Anàlisi de temperatura en superfície: variabilitat i tendències

Autor: Marcel Pagès Ticó, mpagesti7@alumnes.ub.edu
Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Catalonia, Spain.

Tutora: Yolanda Sola Salvatierra, ysola@ub.edu

Resum: Aquest treball estudia les anomalies de temperatura global a la superfície entre 1880 i 2024 utilitzant dades proporcionades pel Goddard Institute for Space Studies (GISS). Fent una primera anàlisi queda patent que les últimes dècades hi ha hagut un increment molt pronunciat de temperatures. Per tal de poder fer prediccions es fa ús de dos models de sèries temporals: l'ARIMA i el Bayesià estructural. Ambdós models revelen una tendència clara d'escalfament predint, respectivament, augments mitjans anuals de 0.021 K i 0.031 K fins a l'any 2060. Els resultats també indiquen un escalfament més intens a les zones continentals i, especialment en el model Bayesià, una amplificació àrtica significativa. Aquests patrons coincideixen amb les projeccions de l'Intergovernmental Panel on Climate Change (IPCC) que alerten que, amb pràcticament tota probabilitat, les temperatures seguiran augmentant durant les pròximes dècades. Tot en conjunt posa de manifest la necessitat d'adoptar amb urgència estratègies efectives de mitigació climàtica.
Paraules clau: Canvi climàtic, Tendències, Prediccions, ARIMA, Models Bayesians Estructurals

ODSs: Aquest TFG està relacionat amb els Objectius de Desenvolupament Sostenible (SDGs)

Objectius de Desenvolupament Sostenible (ODSs o SDGs)

1. Fi de les desigualtats	10. Reducció de les desigualtats	
2. Fam zero	11. Ciutats i comunitats sostenibles	
3. Salut i benestar	12. Consum i producció responsables	
4. Educació de qualitat	13. Acció climàtica	X
5. Igualtat de gènere	14. Vida submarina	
6. Aigua neta i sanejament	15. Vida terrestre	X
7. Energia neta i sostenible	16. Pau, justícia i institucions sòlides	
8. Treball digne i creixement econòmic	17. Aliança pels objectius	
9. Indústria, innovació, infraestructures		

El contingut d'aquest TFG es relaciona amb l'ODS 13 (Acció climàtica) perquè posa de manifest l'escalfament que ha patit darrerament el planeta. A més es prediu que les temperatures seguiran en augment. Per tant, això pot servir per conscienciar del greu problema davant del que ens trobem, urgint mesures mitigadores.

També es pot relacionar amb l'ODS 15 (Vida terrestre), ja que el fenomen tractat pot afectar de forma considerable la preservació dels ecosistemes terrestres.

GRAPHICAL ABSTRACT