The relationship between annual temperatures in successive years in Florida

Cool Coatis

January 2022

1 Introduction

Earth's warming climate is a critical issue in today's society which requires an urgent understanding. Constant monitoring and the analysis of long term trends in temperatures are needed to achieve this. This will allow us to form valid predictions about the future climate, and provide robust arguments to policy makers about how to mitigate these effects.

In this study, we analysed whether temperatures of one year are significantly correlated with the next year (successive years), across years in a given location, using an annual temperature data set from Key West in Florida, USA for the 20th Century.

2 Materials & Methods

2.1 The data

The annual temperature dataset from Key West in Florida, USA between 1901-2000, was used for this study, which consisted of 100 observations. The data analysed consists of two columns: column 1 = the temperature in each year and column 2 = the temperature in the successive year to that in column 1.

2.2 Permutation analysis

As the data is a time series (and hence the measurements of climatic variables are not independent), we used a permutation analysis to assess whether the relationship between temperature in successive years is due to random chance.

For each permutation, the order of annual temperature was 'shuffled' and the permutated temperatures were compared with the original temperatures to calculate a correlation coefficient (method=Pearson's). 10,000 permutation was conducted to generate a distribution of random correlation coefficients of the relationship between annual temperatures of pairs of successive years.

2.3 Estimation of p-values

The probability that a relationship between successive years' temperatures could be due to chance was then assessed. We did this by calculating the proportion of random correlation coefficients which are greater than the observed correlation coefficient.

In order to account for the random nature of our results due to 'shuffling' the data, we opted to include a function in our code which would run the entire permutation analysis 100 times, producing 100 randomly generated p-values. We then plotted the distribution of these coefficients and took the mean of this distribution to be the overall p-value. Whilst a reproducible result could also have been achieved using a (set seed) function in R, we decided that this approach is more statistically sound. We report the results as statistically significant if p = or ¡ 0.05.

2.4 Computing tools

All analyses and plotting were carried out in RStudio, using the packages 'ggplot2' and 'ggthemes' to aid in data visualisation.

3 Results and Discussion

Our permutation analysis showed that there was a significant positive correlation between temperatures in successive years in Florida between the years 1901-2000 (correlation coefficient=0.326; Fig.1). The mean p-value calculated over 100 repeats of this analysis was p = 0.0005 (Fig.2).

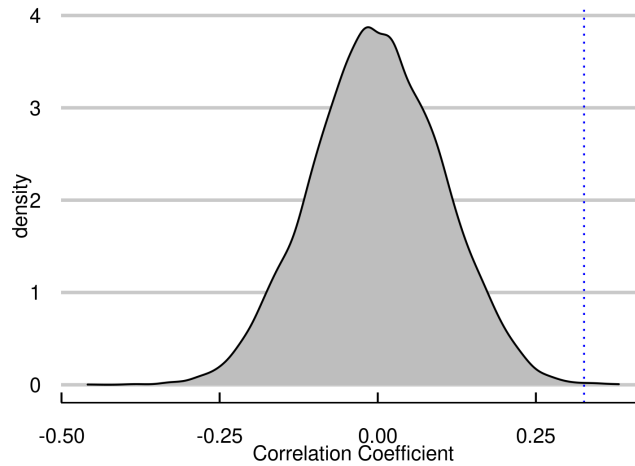

Figure 1: Distribution of random correlation coefficients between temperatures of successive years in Florida 1901-2000. N=10000. Line represents the observed correlation coefficient of 0.326.
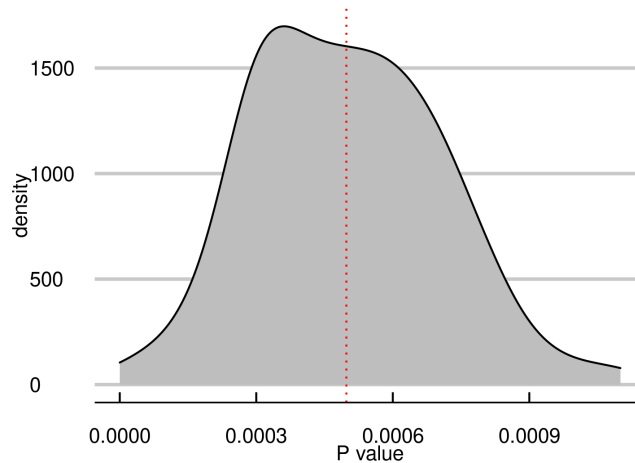

Figure 2: Density plot of p-values assessing significance of relationship between temperatures of successive years in Florida 1901-2000. N = 100. Mean = 0.0005

This can be taken as initial evidence for a warming climate. Further investigations could focus on assessing climate in a wider geographical range, as well as using models to predict future changes in global temperatures.