

# Time Series Analysis on the Evolution of Global Temperature

First Name: Arpitha, Last Name: Jayarama, CWID: A20380321, Class: ITMD 527-04

First Name: Roshni, Last Name: Jariwala, CWID: A20352616, Class: ITMD 527-04

First Name: Manthan, Last Name: Kapadia, CWID: A20376738, Class: ITMD 527-04

**Each group must submit ONLY one copy by a single team member!**

**Type of Your Projects:** Time-series models

## 1. Introduction

The aim of this project is to perform time series analysis on the global temperature evolution. Time series is one of the interesting topics in data analytics, we are motivated towards exploring this method of analysis where we are also able to perform future predictions. The approach is to work on two models based on land-ocean and land-air data which involves analyzing both the models and together coming up with the best model which can predict overall temperature.

## 2. Data Sets

Our data set is regarding the temperature anomaly that is recoded monthly over a span of 137 years i.e. 1880 to 2017. We collected the data set from National Centers for Environmental Information, it contains about 1646 observations including features like temperature anomaly, total error variance, high and low frequency error variance, bias error variance and diagnostic variable.

Where:

Temperature anomaly: It is the deviation of the temperature from the reference value or long term average. positive anomaly means that the temperature is higher/warmer than the reference value and negative anomaly means the temperature is lesser/cooler than the reference value.

Total error variance: It is the error generated while recording temperature. This is nothing but mean squared error.

## 3. Research Problems

We chose the subject as global temperature changes as it is one of the major concern in the world right now, there have been scenario's where people were not prepared for this abrupt change in the climatic conditions hence we think it is important to understand the behavior in the temperature changes and come up with predictions which can help people be prepared for the worst.

Assume that today is to a great degree warm, at some area. At that point, there is an inclination for tomorrow to be hotter than normal. So, on general basis, whatever happens at some time in near future is influenced by what is happening or occurring right now.

So, the problem lies in analyzing the minute trends in temperature change which contains lot of influential information which helps to draw significant results about predicting the temperature.

#### **4. Potential Solutions**

A time series is any series of measurements taken at regular time intervals. Examples include the following: prices on the New York Stock Exchange at the close of each business day; the maximum temperature in London each day; the total wheat harvest in Canada each year. Time series is the best available statistical methodology which can help us to analyze temperature behavior. So, we are trying to closely look in to anomalies in temperature change and come up with a trend which will help in prediction of the exact future climatic conditions or near to it.

#### **5. Evaluations**

Evaluation is basically a research which might be about deciding how compelling the model is, which parts of it are functioning admirably and which require modifying, or whether a few parameters react to specific strategies or conditions uniquely in contrast to others. In the event that the models are dependable, we need to give the evaluation a structure that will reveal us what we need to know. We move ahead with hold-out evaluation as the dataset is very large enough. In this, we divide the dataset into "training set" and "test set" i.e. 80% and 20% and perform evaluation operation. We are going to build the model based on "training set" and use "test set" for evaluation. We will be taking into consideration one of the metrics like AIC, BIC or CP mallow statistics.

#### **6. Expected Outcomes**

We are creating a model based on land-ocean and land-air data. The potential outcome is a model which produces an optimum design which can be used in operational monitoring and climate assessment activities as a future scope of this project.