MH3511-project-submission

zhiyi

2023-04-19

## 1. Introduction

Singapore’s climate is affected by many different driving factors, like the El Niño-Southern Oscillation (ENSO), the Indian Ocean Dipole, the Madden-Julian Oscillation, etc., with varying timescales and impacts on Singapore. In this report, we will attempt to understand how ENSO affects Singapore.

We will look at 4 spatial datasets, specifically focusing on monthly data for a region in the Pacific Ocean from 1982-2020, that are used an indicators for the strength of ENSO: Sea Surface Temperature, Subsurface Sea Temperature, Low Level Wind Speed, and Outgoing Longwave Radiation (as a measure of cloudiness). We will also look at Singapore’s rainfall and surface air temperature over the corresponding period, taken from the Changi weather station, to understand how they relate to the ENSO indicators.

## 2. Data Description

Datasets for the ENSO indicators were obtained from the following sources.

* Sea Surface Temperature - OISST.v2, <https://www.cpc.ncep.noaa.gov/data/indices/sstoi.indices>
* Subsurface Sea Temperature - EN4, (<https://climatedataguide.ucar.edu/climate-data/en4-subsurface-temperature-and-salinity-global-oceans>)
* Low Level Wind Speed - ICOADS, (<https://psl.noaa.gov/data/gridded/data.coads.1deg.html>)
* Outgoing Longwave Radiation - CDR (<https://www.ncei.noaa.gov/products/climate-data-records/outgoing-longwave-radiation-monthly>)

These spatial datasets generally contained global data, but of interest to our project is data that lies within the Nino3.4 region (Latitude: 5N-5S, Longitude:170W-120W).

Datasets for Singapore were obtained from data.gov.sg

Rainfall - <https://data.gov.sg/dataset/rainfall-monthly-total> Surface Air Temperature (SAT) - <https://data.gov.sg/dataset/surface-air-temperature-monthly-mean>

Before proceeding to data analysis, we had to clean and process the data. We will outline the general procedure for doing so, though any exceptions will be mentioned in Section 3. The main goal of the processing is to calculate a baseline value for each month based on a reference period (i.e., what is the average value for Jan, Feb, Mar… Dec from Jan 1980 - Dec 2010), which we can use to calculate the anomaly of the data from Jan 2011 - Dec 2022 (i.e. how much the data for a month differs from the average value for that month)

* Data was stored in a .nc file, read using nc\_open from library ncdf4
* 3 dimensions, corresponding to latitude (1 degree), longitude (1 degree), time (monthly)
* For each time value:
* Find all the values with latitude and longitude that fall within the Nino3.4 region
* Check if there are any NA values, if there are, interpolate to fill them
* Calculate the mean of those values
* Extract the means corresponding to the reference period (Jan 1982 - Dec 2010), and calculate the baseline mean for each month (Jan, Feb, Mar… Dec)
* Calculate the difference between the means for the study period (Jan 2011 - Dec 2022) and the baseline means

## 3. Description and Cleaning of Dataset

We obtained data from various sources, spanning a timeframe of 1982-2022. We segmented our data into 2 portions, 1982-2010 as a baseline to evaluate anomalies, and 2011-2023 as a test to be evaluated based on our baseline to check if they are anomalous. We plotted the data through a histogram, and boxplot to determine the suitability of using the data.

## 4. Statistical Analysis

## 5. Conclusion and Discussion

## 6. Appendix