

Climate Change Analysis from Buoy Station 44013

Anna, Bruce, and Jenna

9/25/2020

Research Question

The purpose of this project is to investigate evidence of global warming from weather buoy data. Meteorological data has been collected from NOAA National Data Buoy Center Station 44013, which is located about 11 miles offshore from Boston, MA. The specific question being asked is as follows: Can data from a single weather buoy provide evidence to support the conclusion that global warming is a real phenomenon?

Approach

While there are several different approaches one could take to look for evidence of global warming or climate change, we chose to focus on air temperature as our outcome variable of interest. We are interested in analyzing air temperature patterns to determine whether there is a significant upward trend over the course of several years. Specifically, we chose to analyze the average air temperature for each month from 2000 to 2018. Our primary analysis consisted of running a series of 12 linear regressions—one for each month—with average monthly air temperature as the outcome variable and the year as the predictor. By observing the coefficients of the predictor for each fitted regression model, we were able to determine the general trend of the air temperature for each month over the course of 18 years, with a positive coefficient indicating a general rise in temperatures and a negative coefficient indicating a general fall in temperatures.

Code Organization

The R code is organized into several chunks with comments for clarity. To start, the following R packages are loaded: tidyverse, stringr, rstanarm, lubridate, and gridExtra. Then, the code reads the meteorological data for a specified range of years from the NOAA website using a series of urls. This produces a separate data frame for each year in the specified range of 2000-2018.

The next goal is to create a single data frame containing the variables of interest for each of every year in the range. This is accomplished by creating a loop within a loop of code:

- The inner loop: Contains the columns that specified years, months, and air temperatures to create one smaller dataset. It then filters out the NAs¹ which is uncollected data.
- The outer loop: Creates a new variable for average temperature (AvgTMP) by taking the mean of the air temperature (ATMP) for every month of every year within the set.

Finally, a loop was added to run a series of linear regressions to analyze the change in average temperatures over time. For each month across every year in the range, plots of the data points with regression lines are created using ggplot2. This was the final step once a tidy data frame was created with only the outcome and predictor variables.

Conclusions

Insert text here about conclusions

¹In the data, there were values with “NA” that designated an unknown or untracked value. This is because NOAA has the occasional lack of funding which leaves empty spaces in their data. So, we had to ignore those data points in our research.

