Climate Project Report

David Stanko

DATA 3320

The Problem

I analyzed time series data containing the sea surface temperature, wind speed, and wave height over eighty years from the seven major ocean basins. My goal was to determine climate trends over time. The data was taken from the ERA5 Reanalysis Product. There is a separate dataset for each ocean basin. Each of those datasets contains the columns time, wave\_height, wind\_speed, and sst (which is sea surface temperature). The time ranges from 01/01/1940 to 02/29/2024. The wave height is measured in meters, the wind speed is measured in meters per second, and the sea surface temperature is measured in Kelvins.

My Analysis

I wanted to extract and predict the global warming trend in the South Pacific Ocean, using the sea surface temperature (SST) for that ocean. First, I calculated and plotted a ten-year moving average of the SST in the South Pacific over the entire time period. Based on the moving average in the figure below, trend is that the SST in the South Pacific Ocean has been drastically increasing between the 1980s and now. While there are small fluctuations in the moving average between the 1980s and now, the general trend is that the SST is increasing. So, there is a clear global warming trend in the last 40 years.

A graph of a number of years

Description automatically generated

Next, I used lagged autoregression with a ten-year lag to forecast the warming trend ten years into the future. The following figure shows the observed SST from 01/01/1940 to 02/29/2024 in blue. It also shows the predicted SST between 03/03/2021 and 02/28/2034 in orange. Note that the forecast ends ten years after the time series ends. According to the forecast, the sea surface temperature will continue to increase in the future. There are some slight fluctuations in the future SST prediction, but the general trend is increasing. We can trust this forecast since the root mean squared error is only 0.25 K, which is small compared to the range of the SST moving average (about 1.5 K). The implications are that, based on the rate of global warming in the past, global warming will continue.

A graph showing the growth of the stock market

Description automatically generated

Additional Analysis Question

I wanted to answer the following additional questions about the dataset: In the North Pacific and South Pacific ocean basins, what is the relationship between sea surface temperature and wave height? What is the relationship between sea surface temperature and wind?

First, I ran a linear regression model with wind\_speed and wave\_height as the predictors and SST as the target, for each of the two ocean basins. For each ocean basin, I found the coefficients of the linear regression problem. In the North Pacific Ocean, the coefficient of wind\_speed was about -0.28, and the coefficient of wave height was about -1.73. That means that when the wave height increases by 1 m, the SST decreases by 1.73 K. When the wind speed increases by 1 m/s, the SST decreases by .28 K. So, in the North Pacific, wave height has a much stronger relationship with SST than wind does. A relationship between SST and wind still exists, but it’s weaker. In the South Pacific, the coefficient of wind\_speed was -0.44, and the coefficient of wave\_height was -0.29. So, in the South Pacific, In the South Pacific, using the same interpretation of the linear regression coefficients, wind has a slightly stronger relationship with SST than wave height. Both of these are also negative relationships.

To further confirm my conclusions from the linear regression model, I ran a random forest regressor on each dataset and looked at the feature importances. For each random forest regressor, I conducted a parameter sweep for the number of trees and the minimum number of samples. For the South Pacific, I used a regressor with 128 trees and 16 minimum samples. For the North Pacific, I used a regressor with 8 trees and 16 samples. For the North Pacific, the feature importance of the wind speed was about 0.29, and the feature importance of the wave height was about 0.71. That indicates that wave height is much more important than wind for predicting SST. For the South Pacific, the feature importance of wind\_speed was about 0.52, and the feature importance of wave\_height was about 0.48. So, wind is only slightly more important than wave height for predicting SST.

Conclusion

The sea surface temperature in the South Pacific Ocean has been drastically increasing between the 1980s and now. Based on the lagged autoregression forecast, the SST in the South Pacific Ocean will continue to increase between now and 2034. So, we should expect global warming to continue. Furthermore, in the North Pacific Ocean, there is a strong negative relationship between wave height and SST. There is a negative relationship between wind and SST too, but it's much weaker. Furthermore, wave height is much more important than wind for predicting SST in the North Pacific. In the South Pacific Ocean, there is also a negative relationship between wave height and SST. There is a slightly stronger negative relationship between wind speed and SST. Furthermore, the feature importances indicate that wind is more important than wave height for predicting SST, but only slightly. So, as global warming continues in the South Pacific Ocean in the next ten years, we expect the wave height to decrease, and we expect the wind speed to decrease at a slightly faster rate.