Nangga Nan 933004912,
Yuncheng Yu 224004002,
Gage Howe 632007308,
Juan Gomez Sandoval 933006119,
Ya-Ru Yang, 833003041

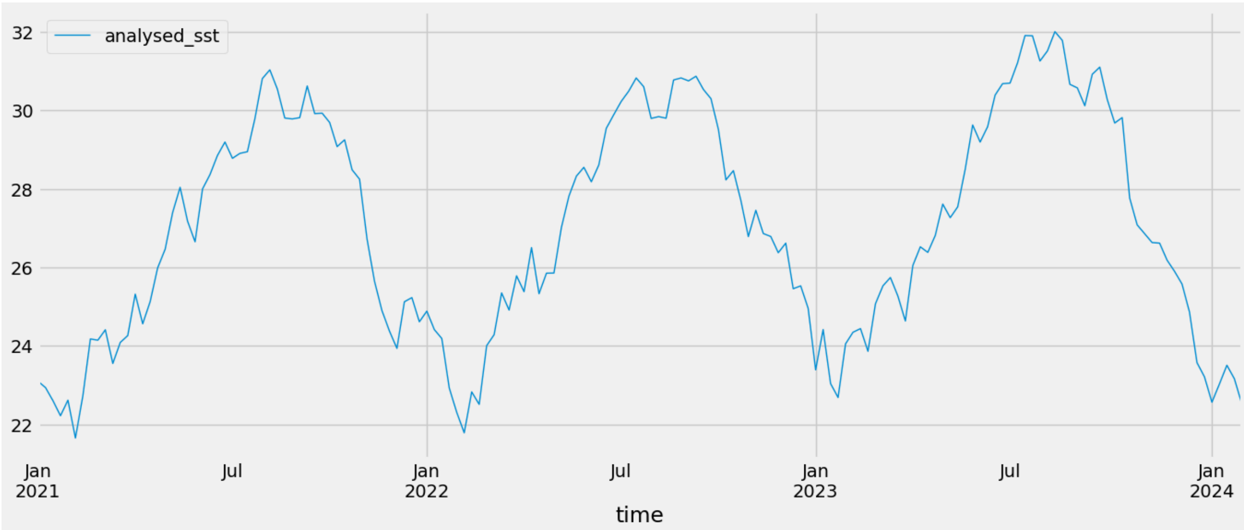
Final Report

Summary: use the Holt-Winters and ARIMA models to predict sea surface temperature in the Florida Key West region. Sea surface temperature is a good indicator of El Nino, which in turn can cause increased rainfall and temperature drop during winter. While the ARIMA model performed badly, the Holt-Winters model performed okay, which can be used to predict sea surface temperature in the future.

Problem Statement: predict sea surface temperature in the Florida region to predict El Nino.

Datasets: Copernicus “cmems_obs-sst_glo_phy_nrt_l4_P1D-m” dataset for the bounding box region of W79-88, N24-31 between 2021-01-01 and 2024-02-04. To use the dataset, run the Python script and rename the file to “copernicus_florida_sst.nc” after downloading. You may have to change to your directory to run it. Or, you can upload the renamed file to Google Colab, which will probably be much easier. Data processing techniques include converting .nc to .csv, limiting results to only N24.55 W81.75 (Florida Key West Latitude Longitude), converting to datetime type of the ‘time’ column, setting ‘time’ column values as row index, trimming down to only ‘time’ and ‘analysed_sst’ columns, converting from Kelvin to Celcius, and taking weekly average of the daily values.

Data Exploration: visualizations of weekly averaged SST values show that the weekly averaged SST is slowly increasing. For example, in the following graph, compare the SST value of July 2022 with that of July 2023. A small increase in SST could mean a lot to the global environment.

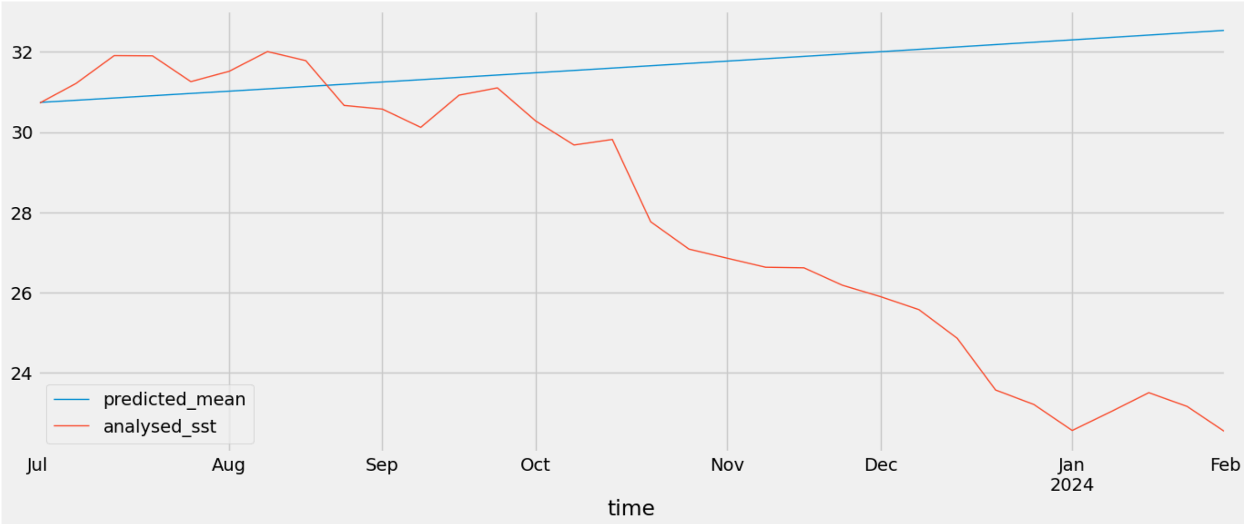


weekly averaged SST

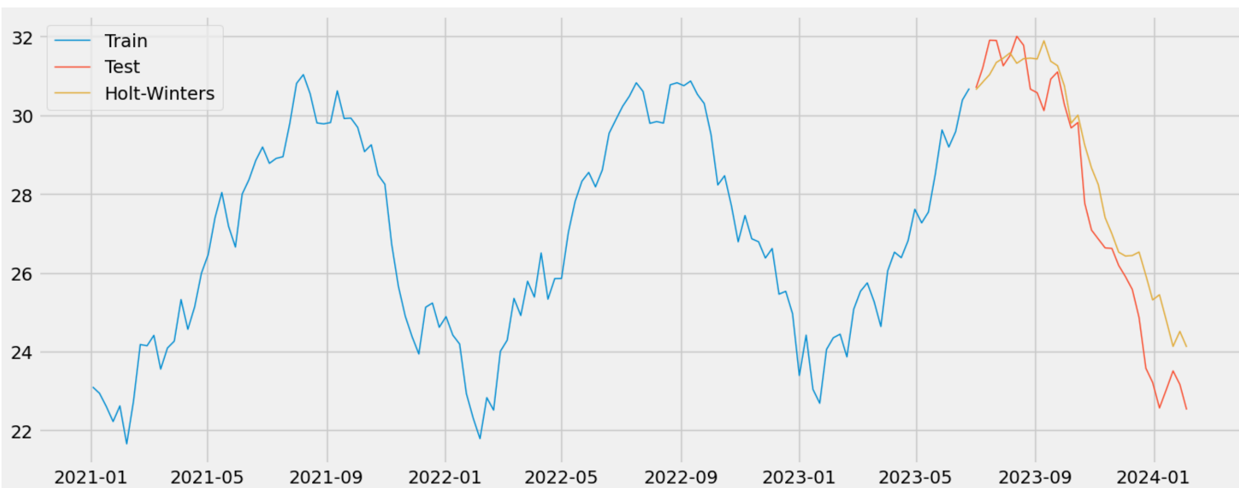
Code: this code is developed on top of the wonderful git repository https://github.com/yashdeep01/Time-Series-Forecasting/tree/main!

Methodology: following the Holt-Winters and ARIMA from https://github.com/yashdeep01/Time-Series-Forecasting/tree/main. This is intended as a starting point, but ran out of time to explore modern Neural Network models.

Modeling and Analysis & Visualization and Interpretation: due to time constraints, did not quite understand the Holt-Winters and ARIMA models. However, as can be seen in the following two graphs, it is strange that the value 52 works for the Holt-Winters but not for the ARIMA. In https://github.com/yashdeep01/Time-Series-Forecasting/tree/main, the value 52 worked for both. Tried some values for the ‘m’ parameter of auto_arima, such as 12, 60, 162, but all did not work very well, not sure of the causes. The predictions of the Holt-Winters are okay, while those of the ARIMA are pretty bad. Future work should focus on optimizing these two models and trying modern NN models.



ARIMA m=52



Holt-Winters seasonal_periods=52

Conclusions and Recommendations: attempted to use the Holt-Winters and ARIMA models to predict sea surface temperature in the Florida Key West region. Future work should focus on optimizing these two models and trying modern NN models.

References:

1. https://github.com/yashdeep01/Time-Series-Forecasting/tree/main
2. https://github.com/Unidata/netcdf4-python/blob/master/examples/reading_netCDF.ipynb
3. ODYSSEA Global Sea Surface Temperature Gridded Level 4 Daily Multi-Sensor Observations | Copernicus Marine Service