Buoy Project

Jason Huang

2020/9/25

1. Abstract

This report seeks to use 20-years of data NDBC Station 44013 to check whether there is global warming. After data collection and cleaning, time series model and linear regression model were used to fit the data. Conclusions were drawn based on the results of models.

2. Data Collection

Data was abstracted from the NDBC website. 30-years of data was collected from 1987 to 2016. Six variables were selected namely, year(YY), month(MM), day(DD), hour(hh), air temperature(ATMP) and sea surface temperature(WTMP). In order to see how the trend of temperature change, the only two related variables ATMP and WTMP were selected.

3. Data Cleaning

Variable ATMP and WTMP were converted to *double* and a new variable 'YYYY_MM_DD' was added to the dataset. The new variable is *date* type. Data with temperatures higher than 90℃ were deleted. The Average daily temperature was first calculated by taking the mean of 24 hours temperature. Then, the average monthly temperature was calculated by taking the mean of the average daily temperature per month. The reason why the temperature were calculated in this way was that all data was taken into consideration.

4. Model Fitting

Data was first plotted to see if any model could be considered. The time series model was considered given the shape of the plot. Also, the linear regression model was considered since it was a very simple model and the sign of the slope can be used to check the trend of climate change.
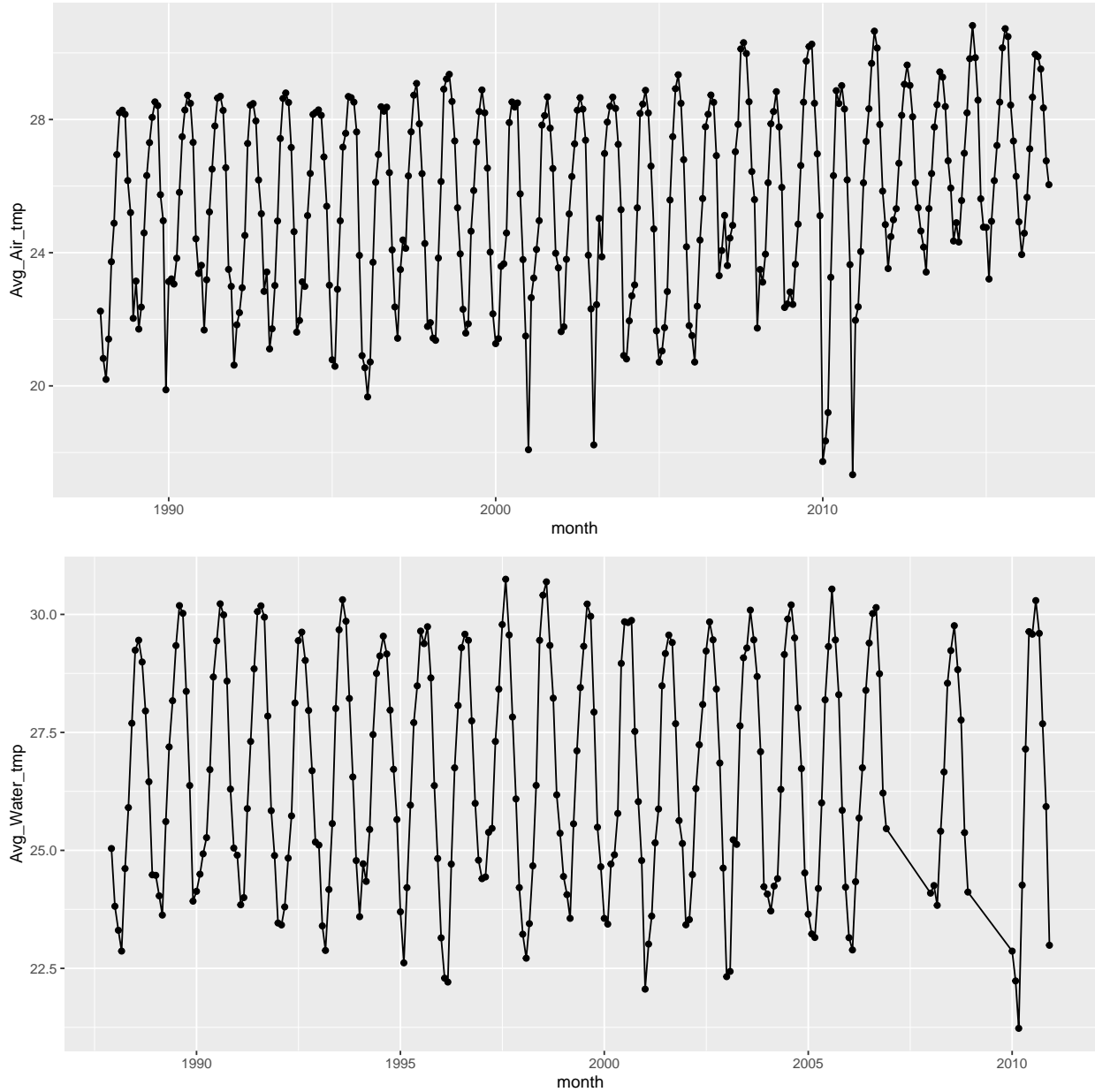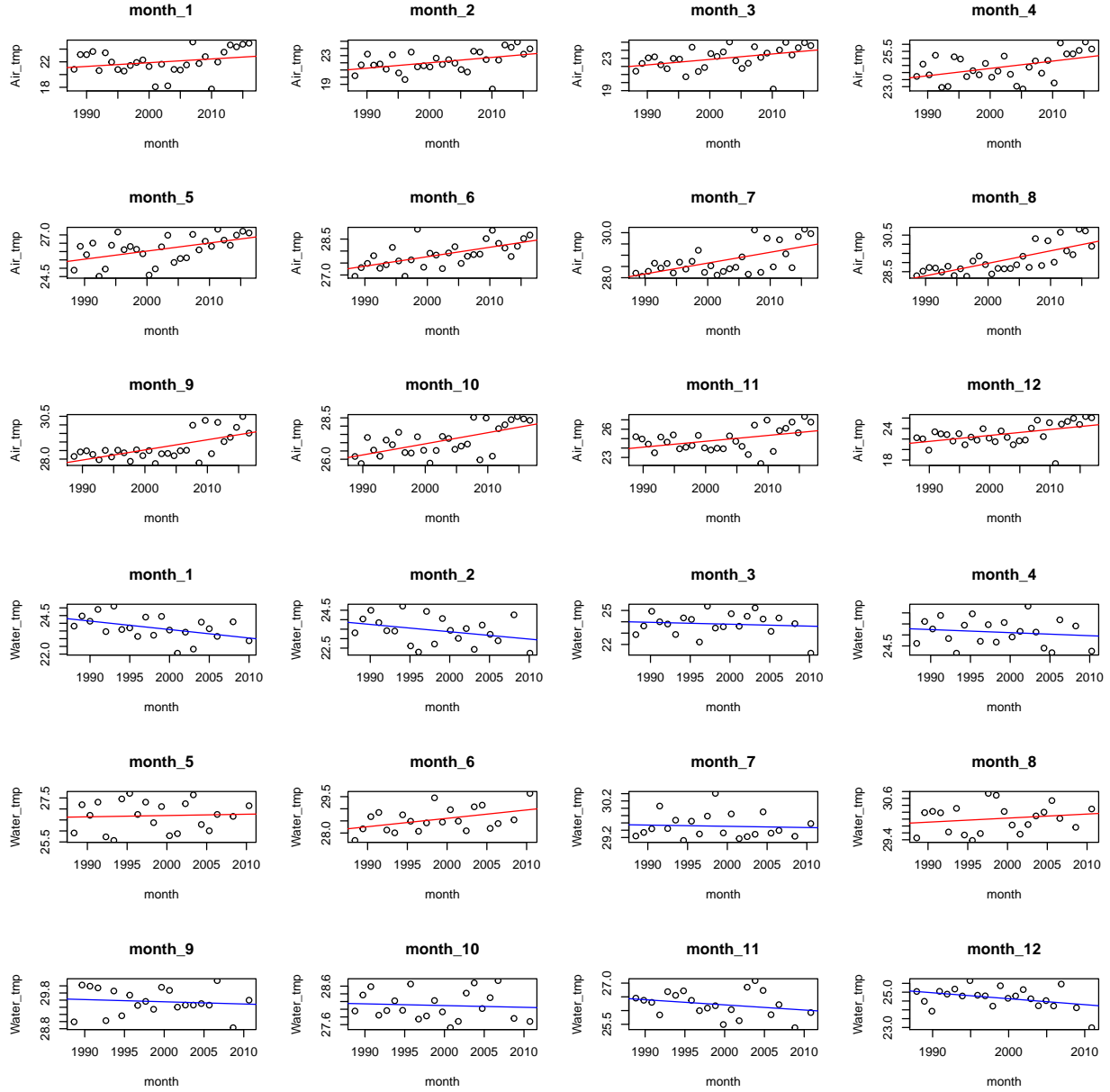
4.1 Time Series Model

Given it was a seasonal data, seasonal ARIMA model was considered. The trend of data within a year was check. After setting the seasonal period s to be 12, ACF and PACF were plotted. By checking the plot, model $ARIMA(1, 0, 0) \times (0, 1, 1)_{12}$ was considered. The predicted data in the next 20 years was plotted. From this model, conclusion about global warming can't be drawn. No strong sign of rising or dropping in temperature was shown from the model. However, signs of a rise in air temperature and a drop in sea surface temperature were shown in the predicted data plot.

Series: diff12_air_tmp

Series: diff12_water_tmp
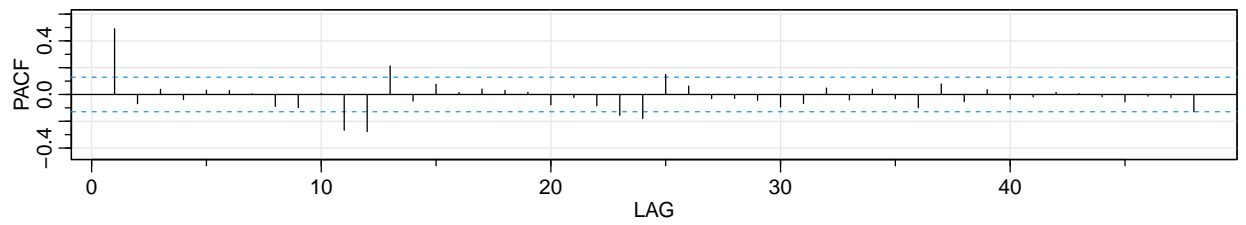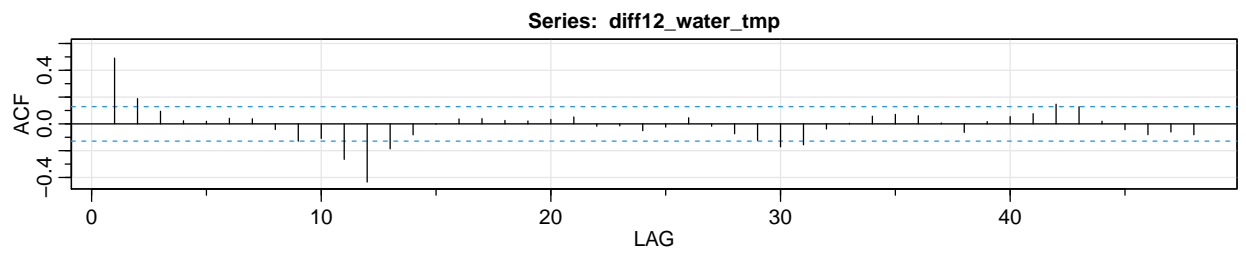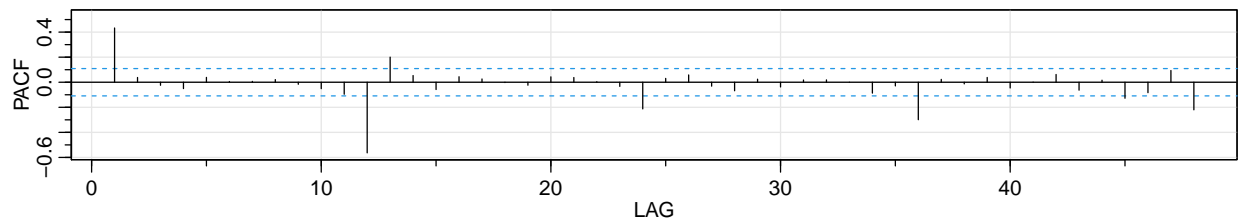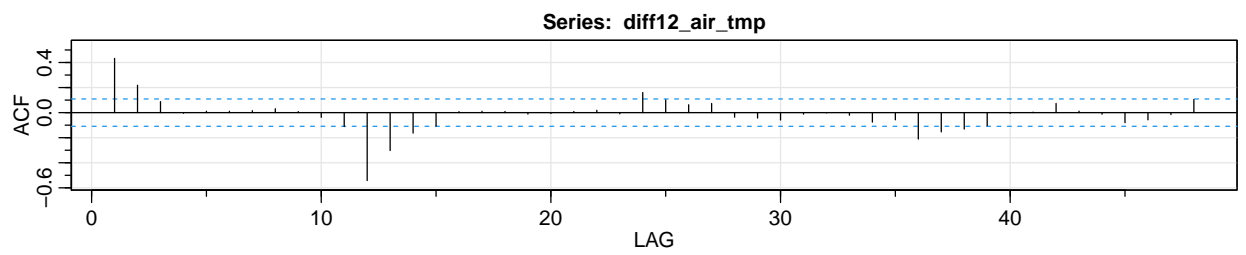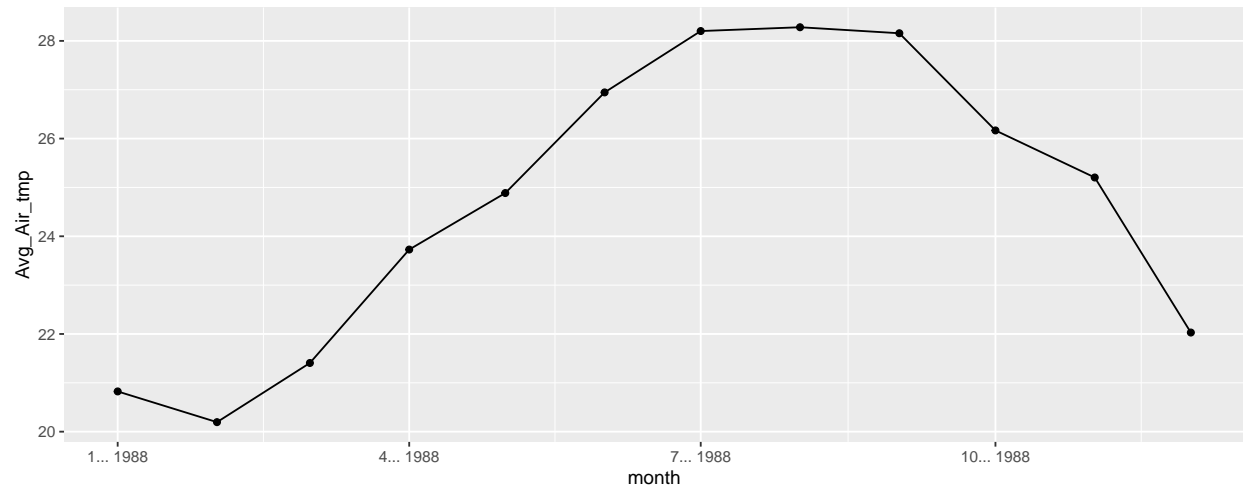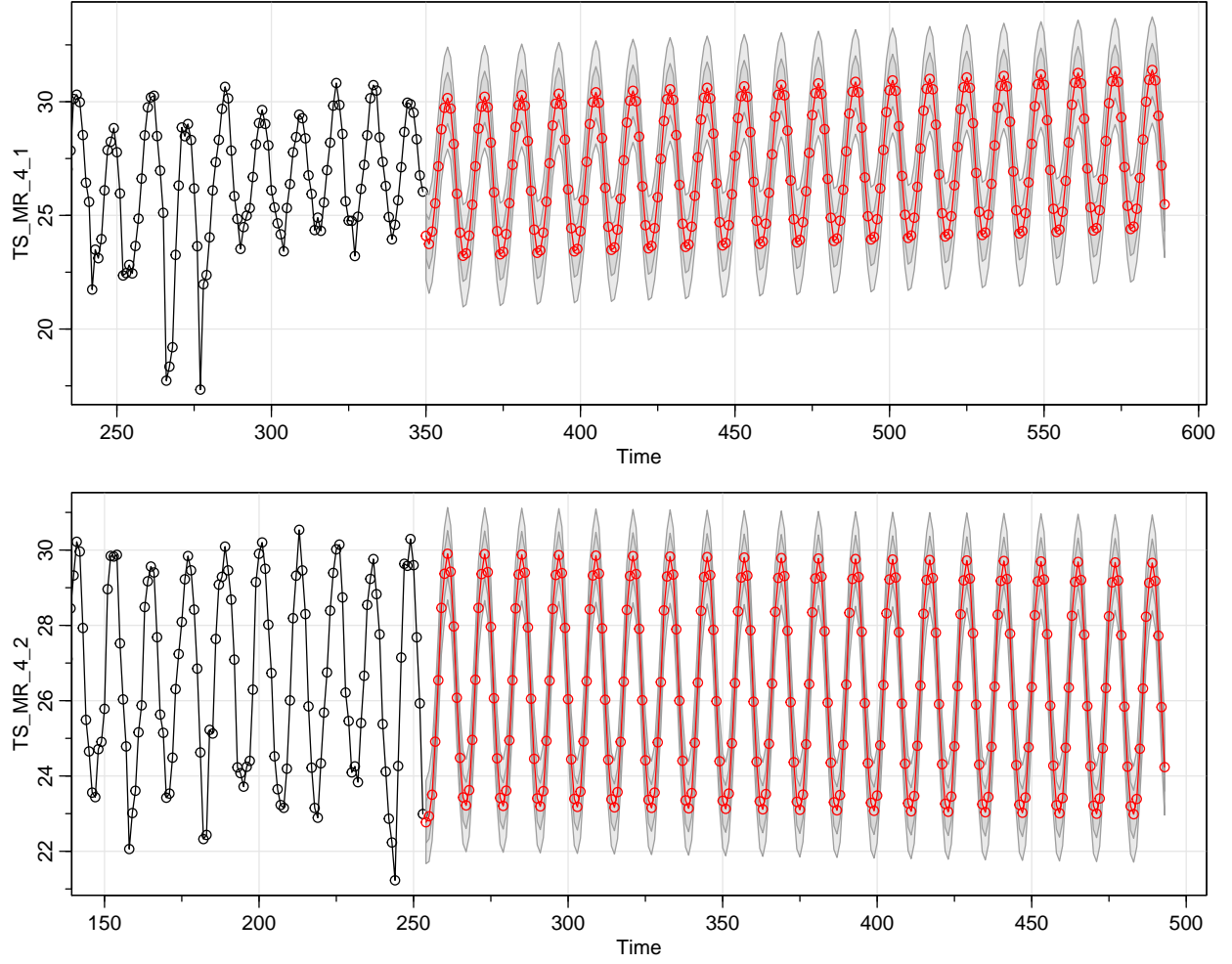
4.2 Linear Regression Model

For linear regression model, data was divided into 12 months. Data for each month was fitted to a linear regression model. All fitted lines and data were plotted on the same page. The red color indicated a upward trend and blue color indicated a downward trend. For air temperature, fitted lines from all months have a positive slope. For sea surface temperature, fitted lines from May to June and August have a positive slope. It is suggested that air temperature is increasing in these 30 years and sea surface temperature is increasing in the summer season.

5. Conclusion

From the time series model, both predicted dropping and rising trend of temperature were shown. For the linear regression model, a strong rising trend of air temperature was shown. Based on the data of air temperature, a sign of global warming was indicated. Based on the data of sea surface temperature, the sign of global warming is not that intense. However, no conclusion can be drawn regarding whether global warming caused the rising of air temperature in this area. The time series model used in this report is rather simple. A more complex model is likely to give a more accurate prediction.