Air Quality

REQUEST FOR PROPOSAL
RFP #: TS – F2.P

TITLE: Ozone Forecast
CLOSING DATE AND TIME: OCTOBER 22, 2020 @ 5:00 PM

Ozone Forecast: TS – F2.P

Background and Purpose

By responding to this Request for Proposal (RFP), the Proposer agrees that s/he has read and understood all documents within this RFP package.

Submission Details

Responders to this RFP should supply:

- A business report up to 10 pages (not including cover page or table of contents), including any supporting plots and tables.
- The commented code (in a separate file) used to produce the results.
- Both files should be uploaded to Moodle by October 22, 2020 @ 5:00 PM.

The report should address **all points described in the "Objective" section** below.

Objective

A local medical facility (hereafter the "client") is seeking proposals for understanding and predicting Daily Max 8-hour Ozone Concentration at the Millbrook School station in Wake County, NC. The goal of the model is to accurately predict the Daily Max 8-hour Ozone Concentration at the Millbrook School station in Wake County, NC for two weeks. With this in mind, the client requests that you have a test data set of the last two weeks in the data and use a validation data set of 28 days. The scope of services includes the following:

- Explore and provide information on the following potential models:
    - Best ARIMA model
    - Best ARIMAX model
    - Best UCM model
    - Neural Network model
    - Ensemble model
- Provide a measure of accuracy on the test data set (such as MAPE and MAE) for each of the above models and provide feedback on potential gains and/or challenges for each method.
- Make a recommendation which model should be used by the client moving forward.
- Provide graphics of best model(s) forecast for the test data overlaid with the observed test data.
- Provide predictions for June 1,2020-June 14, 2020 on the form https://docs.google.com/forms/d/e/1FAIpQLScVY9DTvRanulQC-y_mcSpPlkiu2Lw7tw8gb0J778q6ttc-qQ/viewform?vc=0&c=0&w=1

Data Provided

- There are several pollution data sets provided in a csv format (Carbon monoxide, Nitrogen dioxide, Sulfur dioxide). Each data file is formatted similarly containing the following variables:

Date - Date of the observation (in days)
Source -usually AQS (Air Quality System)
Site ID – unique ID for the given monitoring station
POC – Parameter Occurrence Code (used to distinguish different instruments that measure the same parameter at the same site)
Daily Measurements – The measurement of interest
Units – Units in which daily measurements are recorded
Daily AQI – The daily Air Quality measurement
Site – Name of the site in which data was recorded
State Code – State code (37)
State – State (NC)
County Code – County Code (183)
County – County (Wake)
Site Latitude and Site Longitude of where the measurement was taken

- In addition to the pollution data sets, a data set containing weather information is also provided with the following variables:

PRCP - Precipitation
SNWD - Snow depth
TMAX - Maximum temperature
WSF2 - Fastest 2-minute wind speed
WSF5 - Fastest 5-second wind speed
SNOW - Snowfall
AWND - Average wind speed
WT01 - Fog, ice fog, or freezing fog (may include heavy fog)
TAVG - Average Temperature.
TMIN - Minimum temperature