Enviforecasting Co
2-39-17, Tsurumaki, Setagaya-ku, Tokyo 154-0016, Japan
☎ (+81) (90) 8876 5118
✉ admin@enviforecasting.com

2^{nd} **round Candidates** June 24, 2022

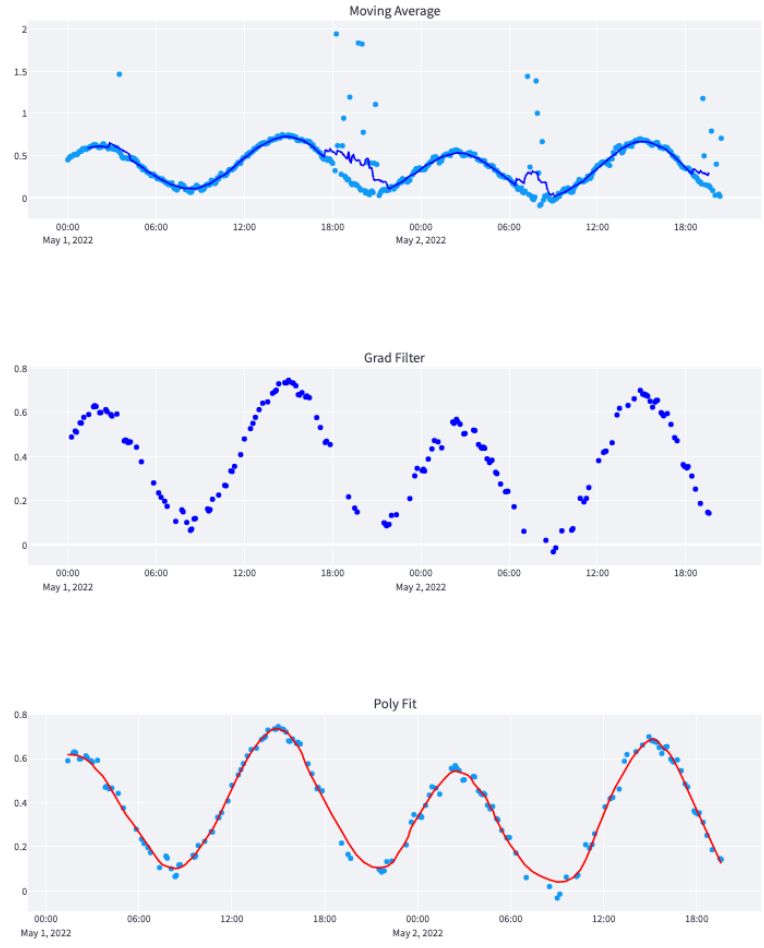
Recruitment Assignment for Shortlisted Candidates

There are three questions in this assignment. Each question contains basic, intermediate and advanced components. You must answer at least one component of any question to be considered for a position. Each question and component will carry a weight. Thus, try to answer as many questions as possible, since those who can answer more will have a better opportunity to receive an offer. Some components will require coding. Python is the preferred language for coding. If you are not able to make a program, you may describe the methodology you would adopt citing methods, modules or frameworks that can be used.

1. Accessing IoT weather station data using provided API.
   http://diffuse.cc:5000 has several APIs that provide access to weather station data. The url "http://iot.diffuse.cc:5000/wsdata/recent.php?station=station&hours=hrs&minutes=mins" will provide a html table of current weather data whereas.
   "http://iot.diffuse.cc:5000/getwsdata_recent.php?station=station&hours=hrs&minutes=mins" will provide json of the same data. (ex.
   "http://iot.diffuse.cc:5000/wsdatarecent.php?station=ITOKYO91&hours=1&minutes=0"
   Or
   "http://iot.diffuse.cc:5000/getwsdata_recent.php?station=ITOKYO91&hours=1&minutes=0"
) Using either of these outputs make the following:
   Basic Make a summary of weather information consisting of the following weather daily summaries for the past week for the weather station=ITOKYO91:
   Total Rainfall, average temperature, maximum temperature, minimum temperature, average wind speed, maximum wind gust
   Intermediate Make a program to create a plot of the weather parameters that can be retrieved for a given past time interval (ex. Last 30 min, last 1 hr. etc). Provide a working code.
   Advanced How would you create a streaming weather data plot that can be made available through a web server? Provide a working code. If difficult, outline the program. (The 'Realtime Streaming' menu item available at the http://apn.enviforecasting.com/ provide a streaming data sample)

2. Filtering IoT device input
   Often it is necessary to filter out noise from data retrieved from monitoring devices. The water levels measured at two stations are given in the data files station-1.csv and station-2.csv. It is necessary to create a filter to smooth these data.
   Basic: Create a program to filter the outliers and make smooth output for station 1 data given in station-1.csv. A sample of filtering approach is shown in the figure(2-1) for station-1.
   Intermediate Apply the program to data in station 2. Does it provide acceptable results?
   Advanced If not provide a modified version of the filtering program. The performance of the filtering approach adopted here is shown in the figure(2-2) for station-2.

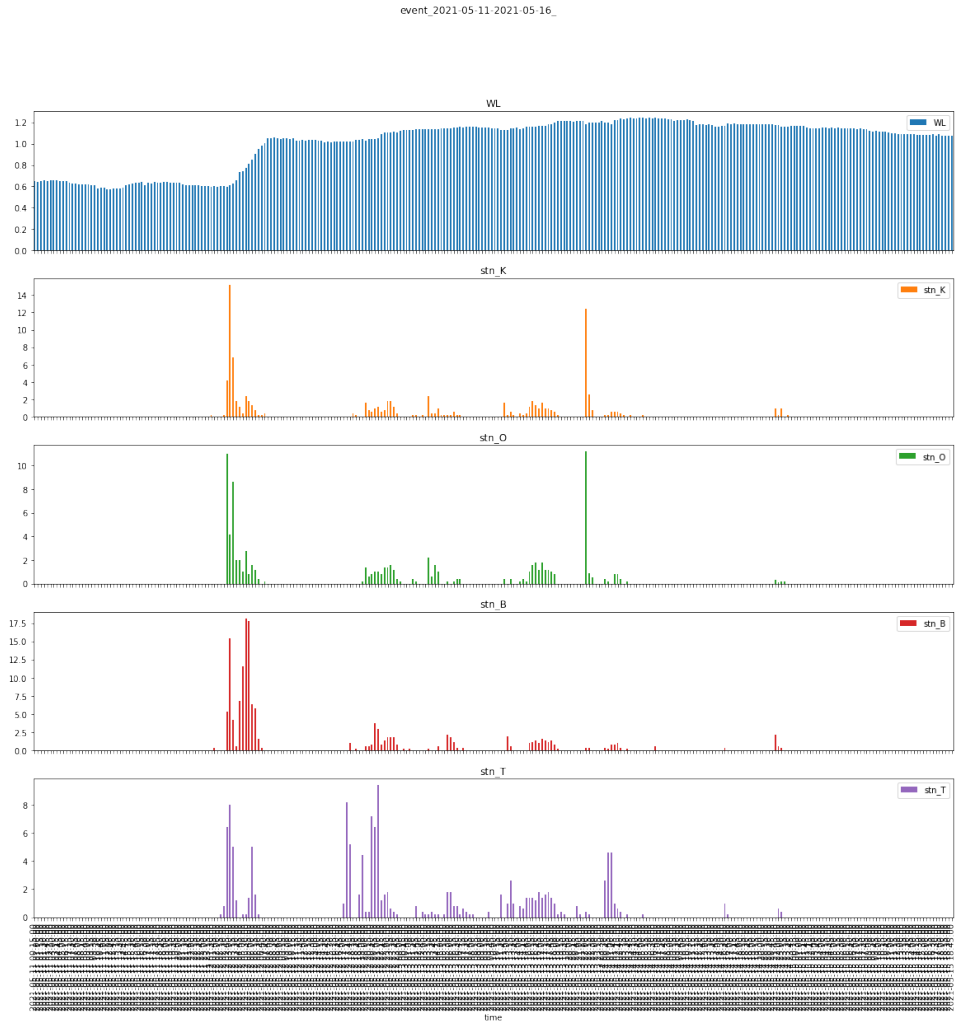
2-1 Filtering Observations of Station-1



2-2 Filtering Observations of Station-2

3. Using Machine Learning for forecasting

Basic The figure below shows the water level of a lake and the rainfall recorded in 4 weather stations nearby during one heavy rain event. If you are asked to build a ML model to estimate the lake water level based on the rainfall from the weather stations what would be the approach you would use. List the types of models that can be used, there drawbacks and for what practical use we can use them. For the practical use discuss also the training and verification process of an ML model. List the libraries and/or programs you would use to build the ML model.



3-1 Water Level and Rainfall during event-1

Intermediate LSTM (Long Short Term Memory) is a class of models that are suited to model time series information. They can also be used to estimate future value of the variable and are thus suited for forecasting. The figure 3-2 shows the result of a LSTM model trained for the event shown in figure 3-1. It uses a look-back window of size 24 (6 hrs) and forecasts 6 hrs. ahead. Looking at the forecast values at different future lead times, discuss how we can use this result. The model used for this analysis is

```
tensorflow.keras,Sequential([
    LSTM(64, input_shape=[n_input_steps, n_features]),
    Dense(n_output_steps)])
```

Improve upon the above model or build your own model to reduce forecasting error and can be used to forecast 8 steps (2 hrs) to 24 steps (6 hrs) ahead and show results. You may try/suggest other possibilities such as resampling data from 15 min to 30 min, etc.

Enviforecasting Co
2-39-17, Tsurumaki, Setagaya-ku, Tokyo 154-0016, Japan
☎ (+81) (90) 8876 5118
✉ admin@enviforecasting.com

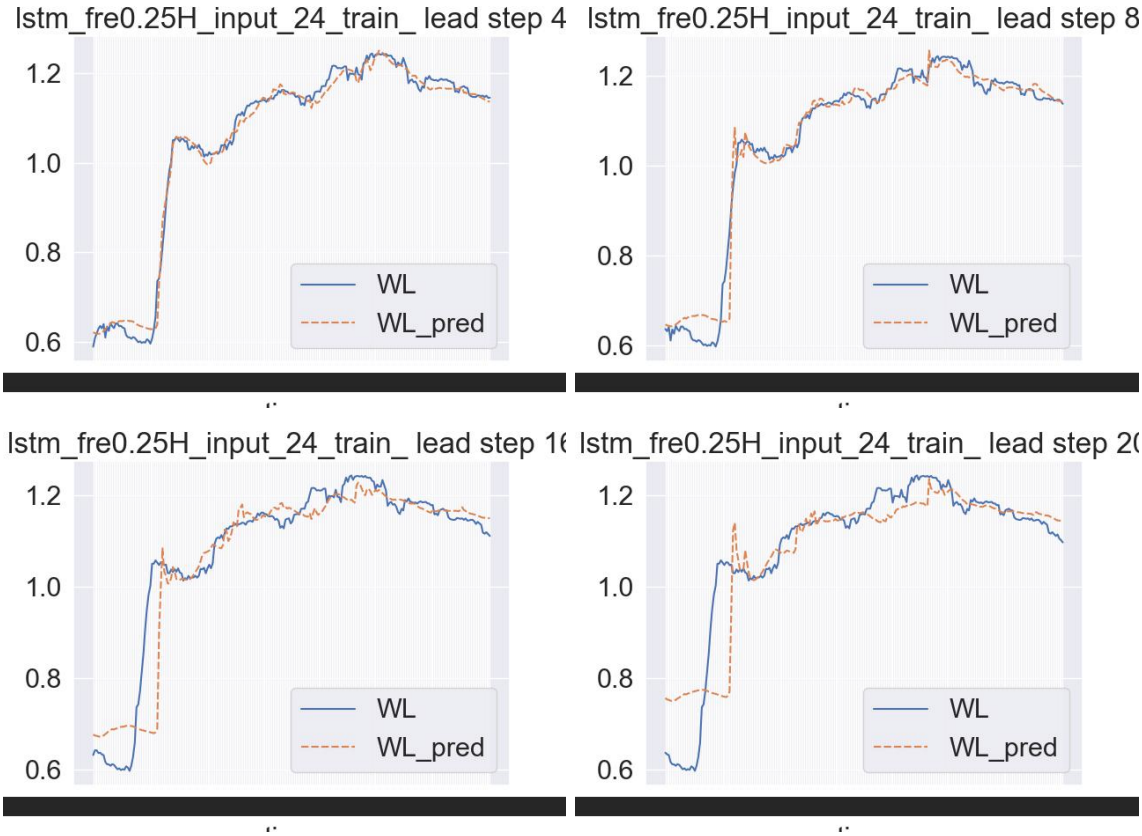
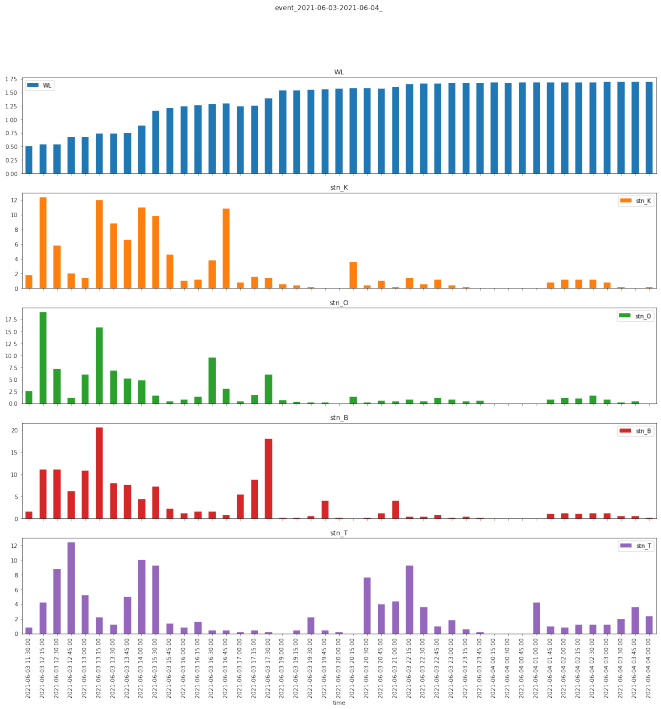


Figure 3-2 Forecast with 1, 2, 4 and 5 hrs lead time

Advanced The Figure 3-3 shows the water level and rainfall distribution for another heavy rain event (event-2). Would the model derived for event-1 work well for the event-2? Apply the model for event-2 and if not satisfactory derive a model that can perform well for both events.



3-1 Water Level and Rainfall during event-2

..