Univariate Multi-Step Ahead Prediction results

Outlined below are results from the test done.

Algorithms

- Averaging baseline
- Naïve baseline
- LSTM
- MLP regressor
- XGBoost regressor

Parameters

- k = 5 – prediction horizon
- p = a variable from 1 to 15 times steps to evaluate the optimal.

The Naïve baseline

The last column of the p vector is repeated across the k vector for baseline evaluation. In these case iterating p values is not needed hence the value of k is equal to the value of p. K was taken to be 5 and p was also set at 5. Table 1 shows the prediction error across the prediction horizon (k). Figure 1 shown the test-prediction comparison at t+1 whereas Figure 2 shows the comparison at t+5.

```
pred_column = np.column_stack([x_test[:,x_test.shape[1]-1]])
y_pred = np.repeat(pred_array, k, axis=1)
```

Table 1: RMSE across k

DAY	Prediction Error (RMSE)
1	0.0455
2	0.0615
3	0.0649
4	0.0716
5	0.0796
Overall prediction error	0.0656

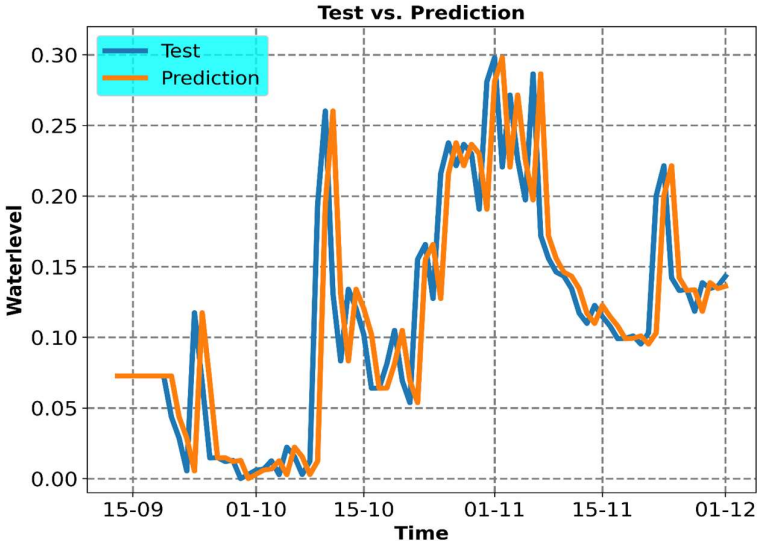

Figure 1: Prediction vs Test at (t+1)

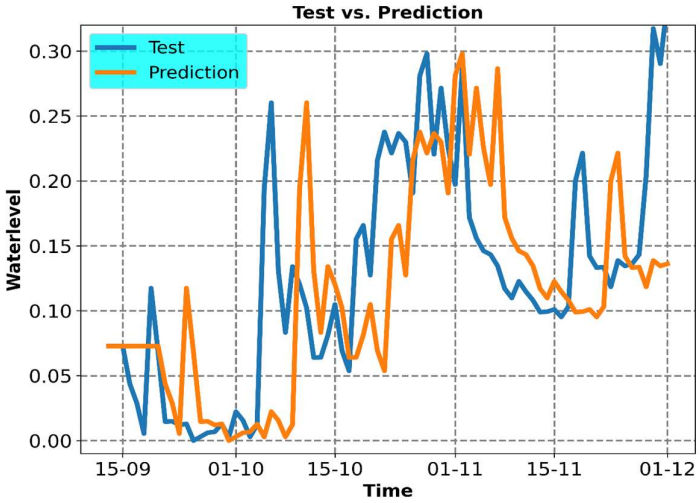

Figure 2: Prediction vs Test at (t+5)

The Averaging baseline

In this case the average of the optimal p is repeated across k for the baseline evaluation. The optimal p is realized by iterating over the range outlined under parameters. The threshold metric used to determine the optimality was the RMSE. K was set to 5 The optimal p was found to be 4 hence p was set to 4 during test. Table 2 shows the error result after experimentation and Figure 3 and 4 show some test-prediction comparison plots.

Table 2: RMSE across k

DAY	Prediction Error (RMSE)
1	0.0479
2	0.0562
3	0.0621
4	0.0692
5	0.0778
Overall prediction error	0.0635

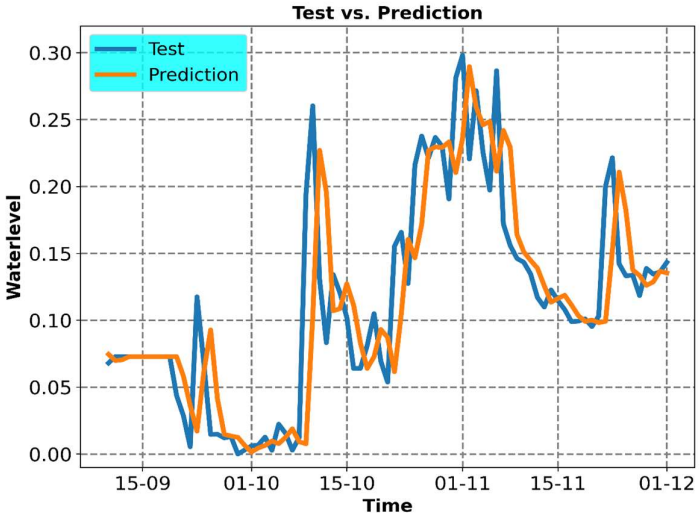

Figure 3: Prediction vs Test at (t+1)

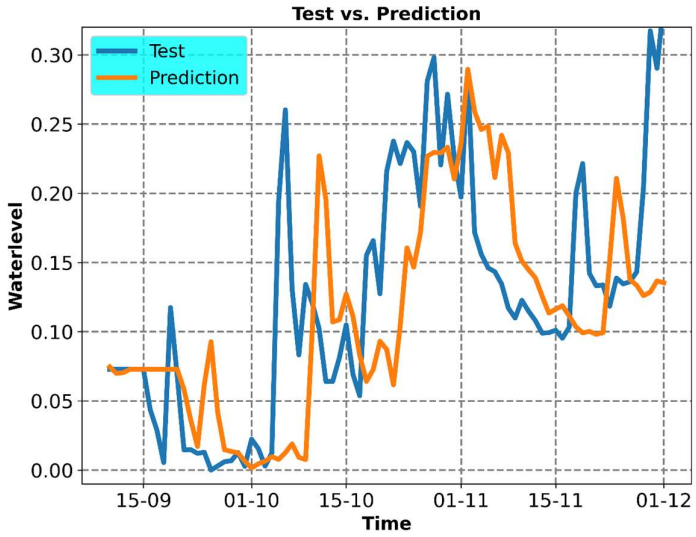

Figure 4: Prediction vs Test at (t+5)

LSTM

The optimum p was found to be 4.

Table 3: RMSE across k

DAY	Prediction Error (RMSE)
1	0.0465
2	0.0537
3	0.0604
4	0.0674
5	0.0741
Overall prediction error	0.0615

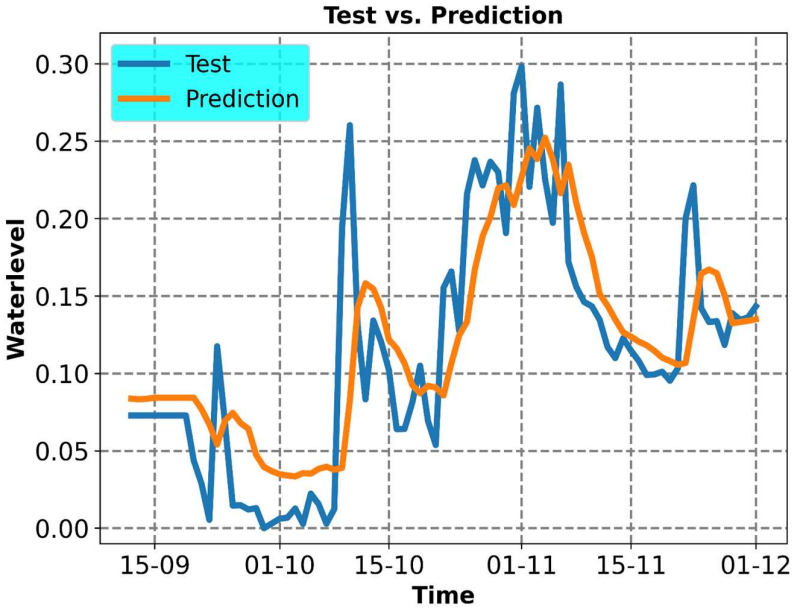

Figure 5: Prediction vs Test at (t+1)

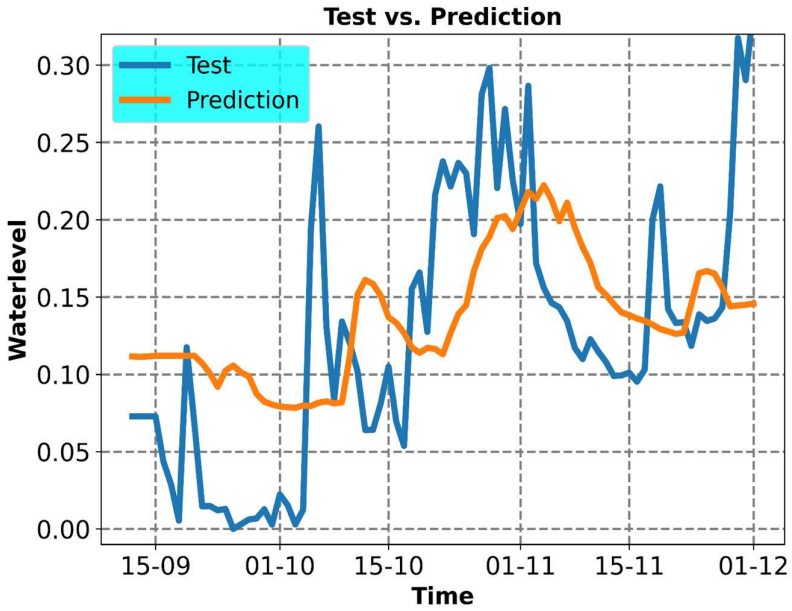

Figure 6: Prediction vs Test at (t+5)

MLPRegressor

The optimum p was found to be 7.

DAY	Prediction Error (RMSE)
1	**0.0466**
2	0.0571
3	0.0605
4	0.0677
5	0.0733
Overall prediction error	**0.0627**

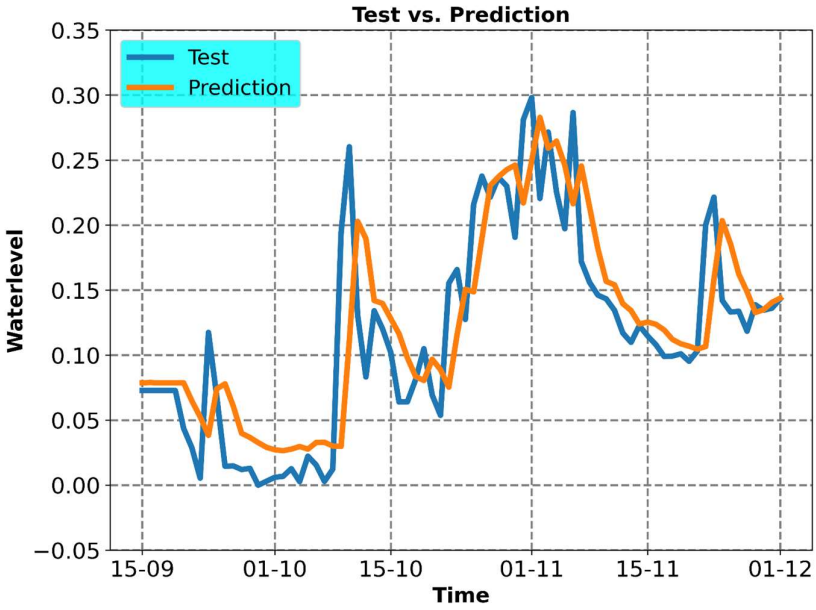

Figure 7: Prediction vs Test at (t+1)

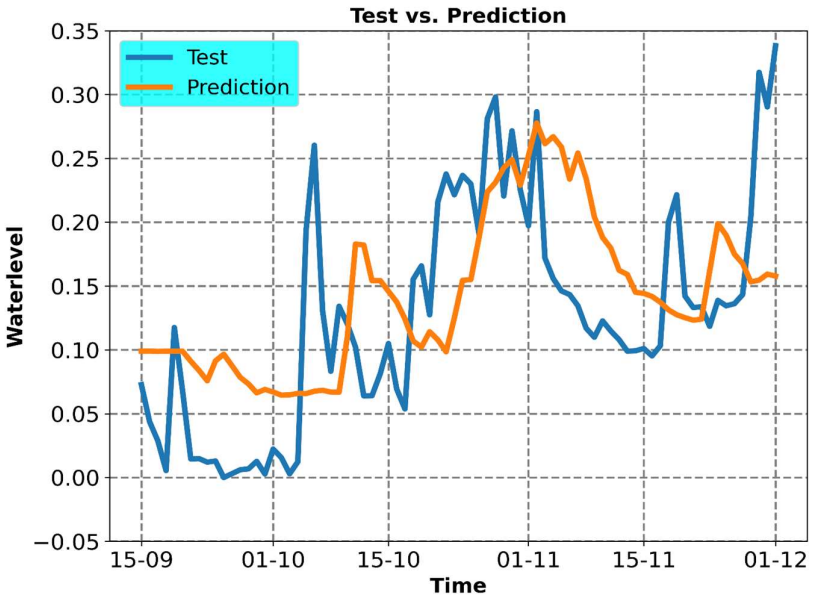

Figure 8: Prediction vs Test at (t+5)

XGBoost Regressor

The optimum p was found to be 4.

DAY	Prediction Error (RMSE)
1	**0.0420**
2	0.0567
3	0.0764
4	0.0877
5	0.0901
Overall prediction error	**0.0730**

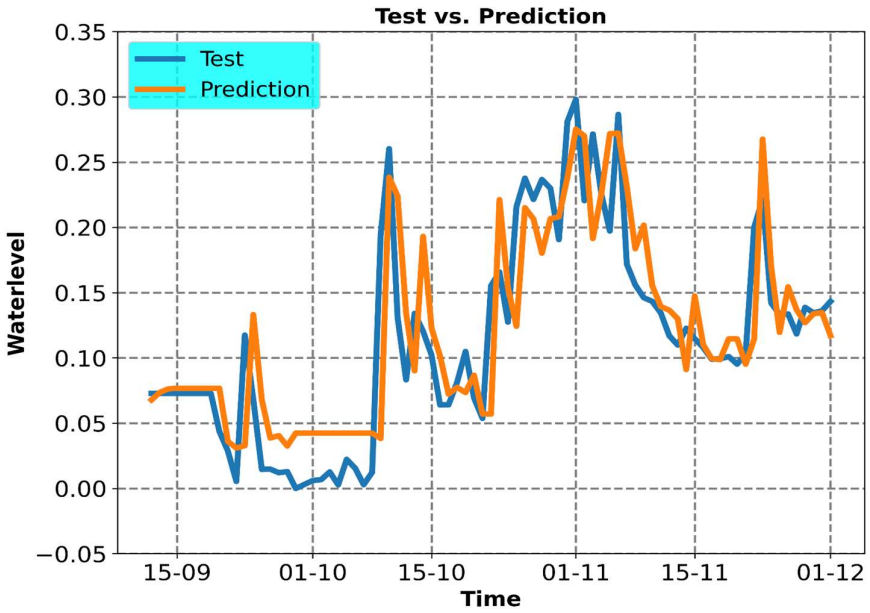

Figure 9: Prediction vs Test at (t+1)

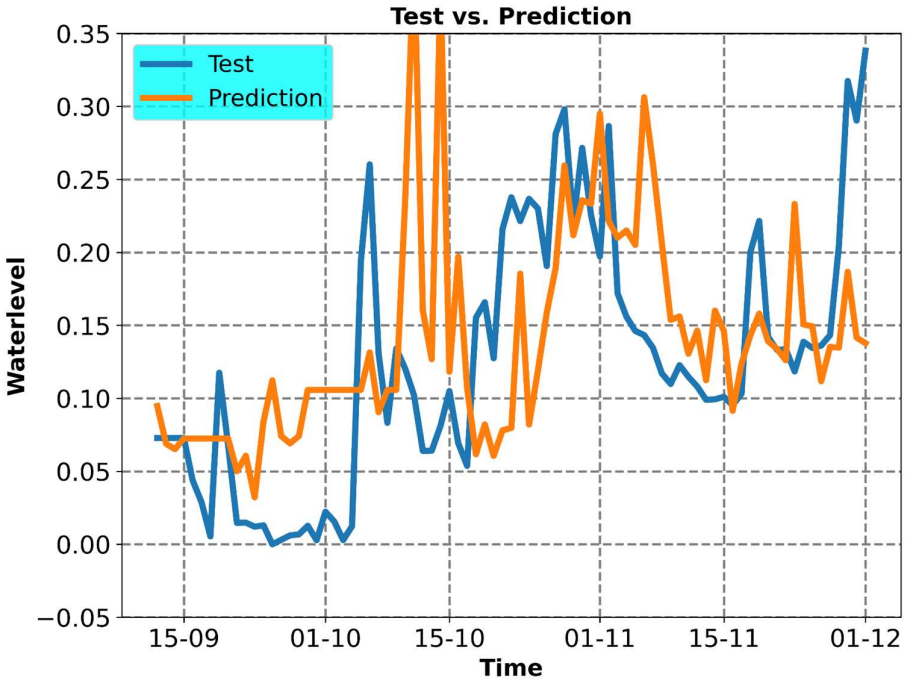

Figure 10: Prediction vs Test at (t+5)