Problem 1

a. random seed = 4

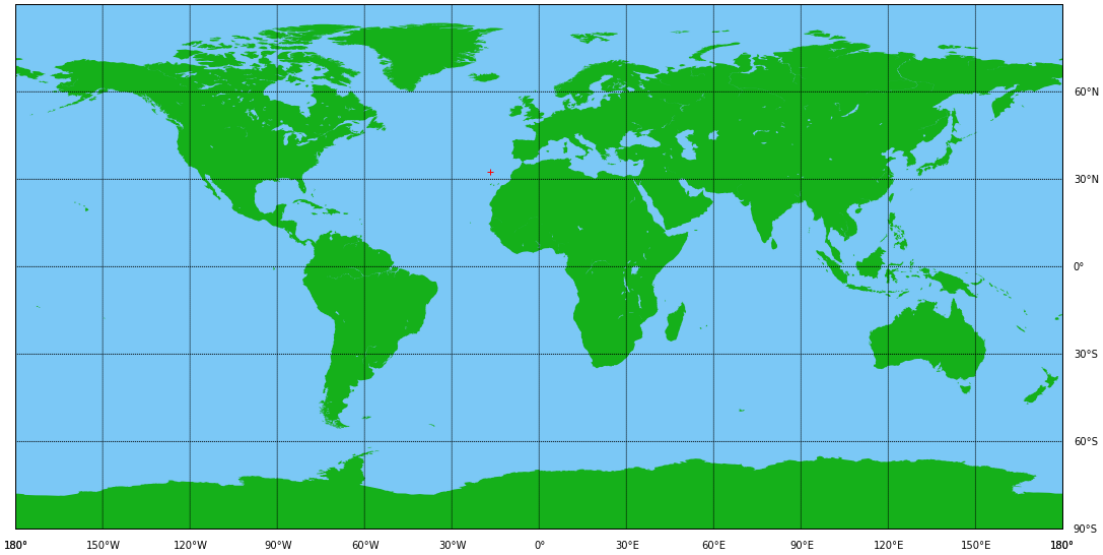

Figure 1.1 The location in the map

b.

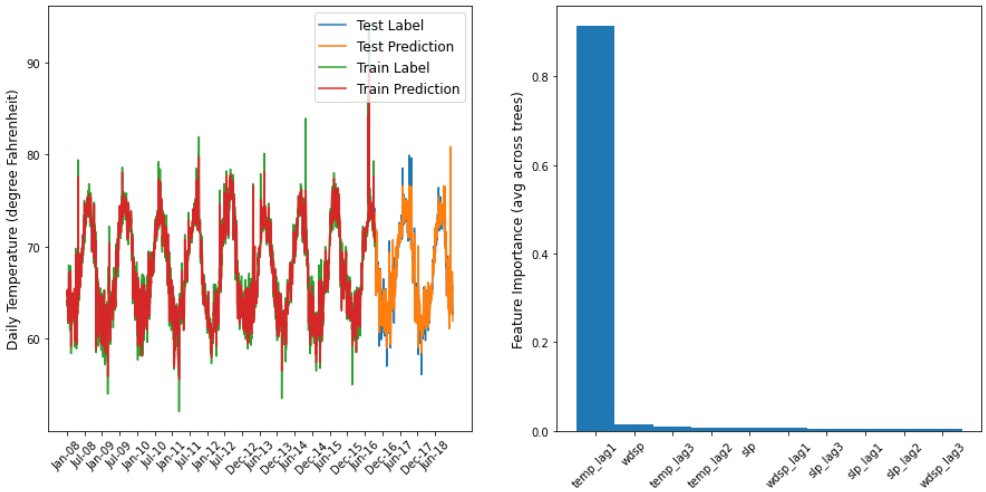

Figure 1.2 Result of temperature prediction

Actually, the temperature is a spatiotemporal sequence, while this model is only time-based. Lacking of one significant element would amplify the error.

Problem 2

a. random seed = 4, predict windspeed (wdsp)

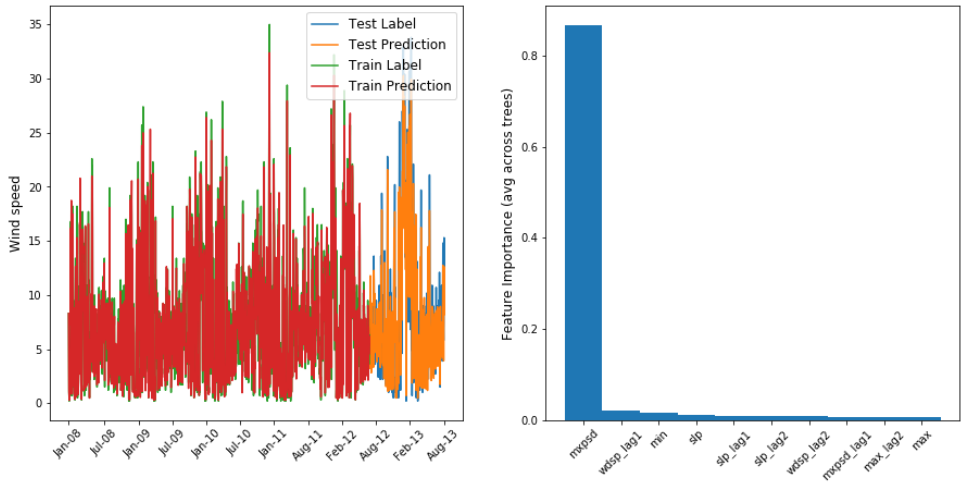

Figure 2.1 Result of windspeed prediction

b. To add spatio factors, I chose the common "Datetime" data from 2 close locations. (2009 groups data)

Name	Lat	Lon
EGVEKINOT	66.350	-179.117
MYS SHMIDTA	68.900	-179.633

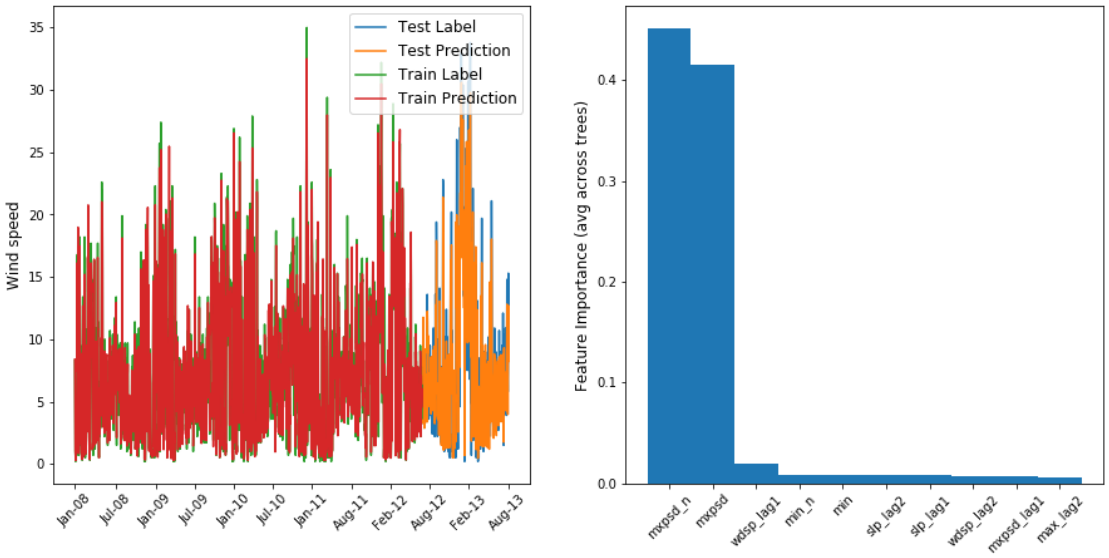

Figure 2.1 Result of spatiotemporal windspeed prediction

Timeseries-only MSE: 4.046597748756221

Spatiotemporal MSE: 4.026684502487564

The MSE is improved, which means spatio factor also influence the prediction. The improvement is slight, because the two area is not close enough.

Problem 3

1. Implement an another ML model for timeseries prediction,

I chose SVM model to replace the RF model. I import the SVM package and fit the train set. Finally predict the validation set and plot them.

Different from RF results, SVM model showed that the recent three days temperature are all highly related with the temperature prediction. While its MSE (which is 2.0492889156009055) is much larger than that of RF (which is 0.8984402600355994).

Genrerally, SVM performs better in linear relations. So we can conclude that the factors about temperature prediction is more likely to be nonlinear.

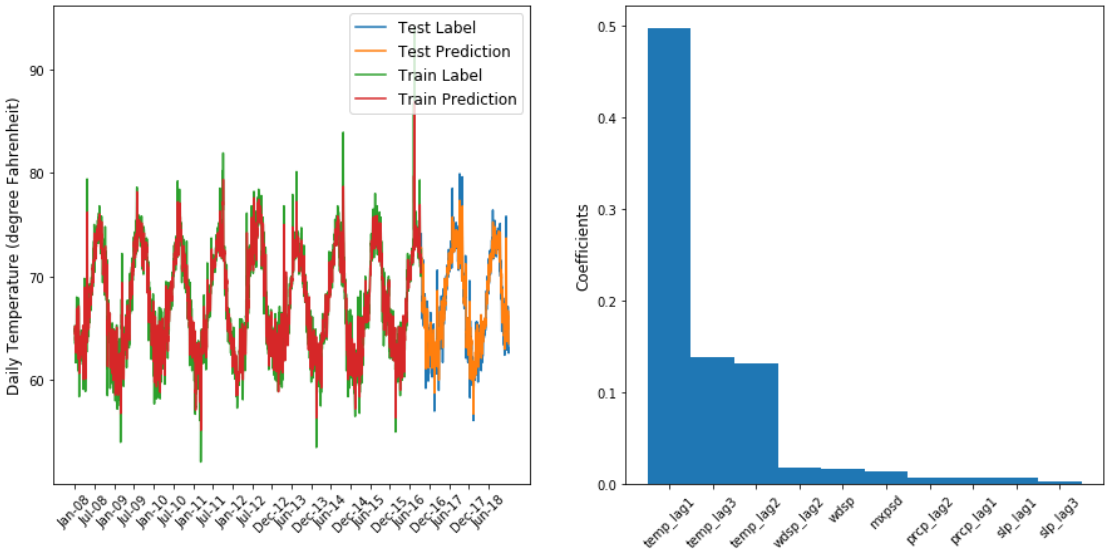

Figure 3.1 Result of temperature prediction by SVM