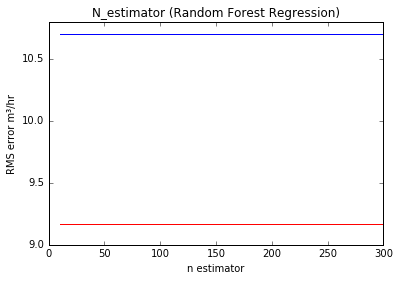
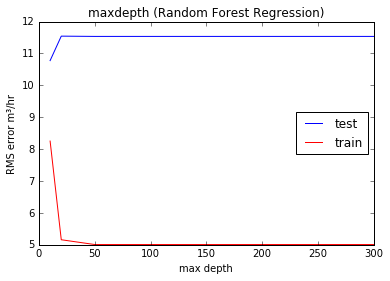
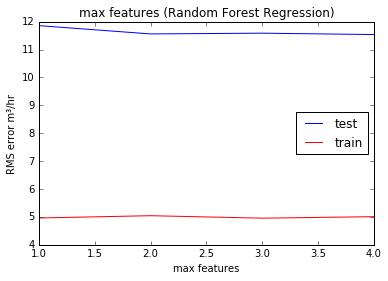
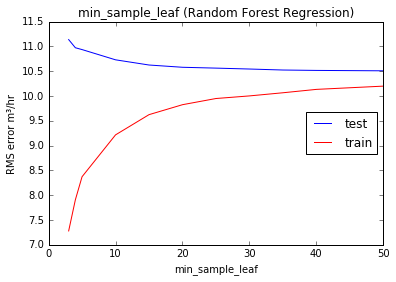
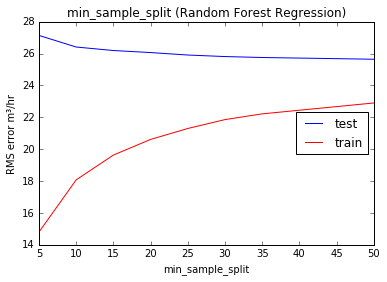
Results and discussion

Different machine learning models were studied and implemented on the same dataset and the results were compared.

The following machine learning models were implemented:

* Random forest

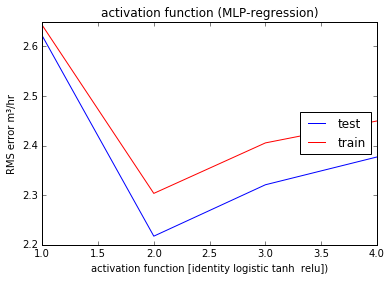
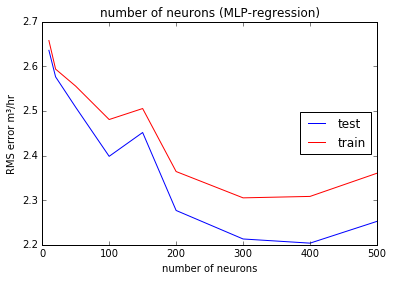
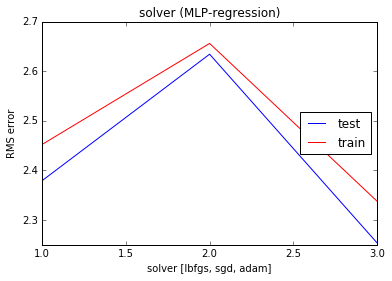
The random forest model was tuned on its parameters (n\_estimators, min\_sample\_leaf, max\_depth, min\_sample\_split, max\_features)



The results show that the model can be tuned and the error ranged from 24 to 5 m3/hr

* Multi-linear perceptron regression

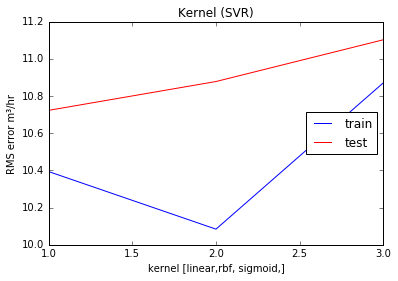
Here the model is tuned on activation function, number of neurons and type of solver



The results show that by parameter tuning we get the rms error ranging from 2.2 to 2.7 m3/hr

* Support vector regression

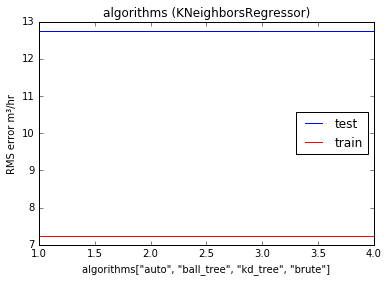
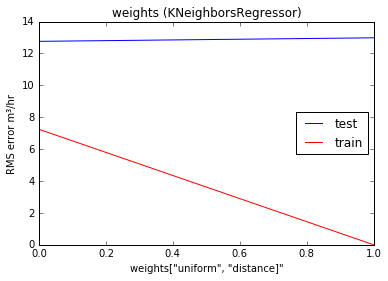
This model is only tuned on its kernel property



The results show that the model can be tuned and the error ranged from 10 to 10.8 m3/hr

* K-nearest neighbor

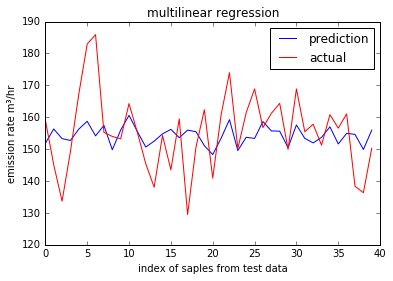
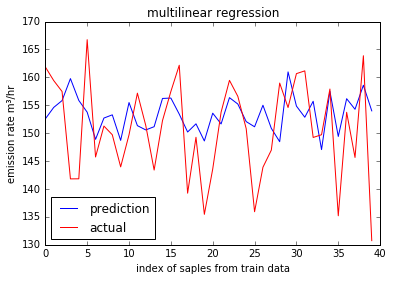
This was tuned on its weight and algorithm parameter, which are used for clustering



The results show that the model can be tuned and the error ranged from 3 to 14 m3/hr

* Multi-linear Regression

This algorithm is an extension of linear regression with added support of higher dimensions in feature space



The results show that the model predicts with an error of 10 m3/hr

to calculate the concentration of the pollutant after emission from the stack, Gaussian air dispersion model was used for simulation, this was implemented using python programing language

