

# r-aq-analysis

This document explains the application of regression to aq data. The aq-data was collected over couple of months, and a part of it was used to study the applicability of linear regression. Earlier attempt was to use the calibration curve provided by the manufacturer, and then applying temperature compensation, that method had shortcomings, hence explorations into other methods.

## AQ

Summary first:

```
summary(aq.filtered)
```

```
##           X           Series      Time      Value
## Min.      : 2.0000    03:562   Min.      :NA   Min.      :128.7200
## 1st Qu.:144.2500           1st Qu.:NA   1st Qu.:209.6775
## Median :288.5000           Median :NA   Median :225.4250
## Mean     :288.2954           Mean  :NA   Mean     :220.5665
## 3rd Qu.:432.7500           3rd Qu.:NA   3rd Qu.:235.6050
## Max.     :576.0000           Max.   :NA   Max.     :253.6100
##      temp      humidity      aq_resistance
## Min.      :10.55000   Min.      :17.26000   Min.      :3.410000
## 1st Qu.:11.48250   1st Qu.:51.66125   1st Qu.:3.590000
## Median :12.70500   Median :62.61500   Median :3.630000
## Mean     :14.46835   Mean     :58.01204   Mean     :3.618944
## 3rd Qu.:15.57500   3rd Qu.:66.83375   3rd Qu.:3.650000
## Max.     :41.07000   Max.      :69.58000   Max.      :3.690000
##      pred
## Min.      :138.2715
## 1st Qu.:212.6520
## Median :224.7650
## Mean     :220.5665
## 3rd Qu.:234.2539
## Max.     :243.7077
```

Regression with Value and aq resistance

```
##
## Call:
## lm(formula = Value ~ aq_resistance, data = aq.filtered)
##
## Coefficients:
## (Intercept)  aq_resistance
##    -1165.6800      383.0527
```

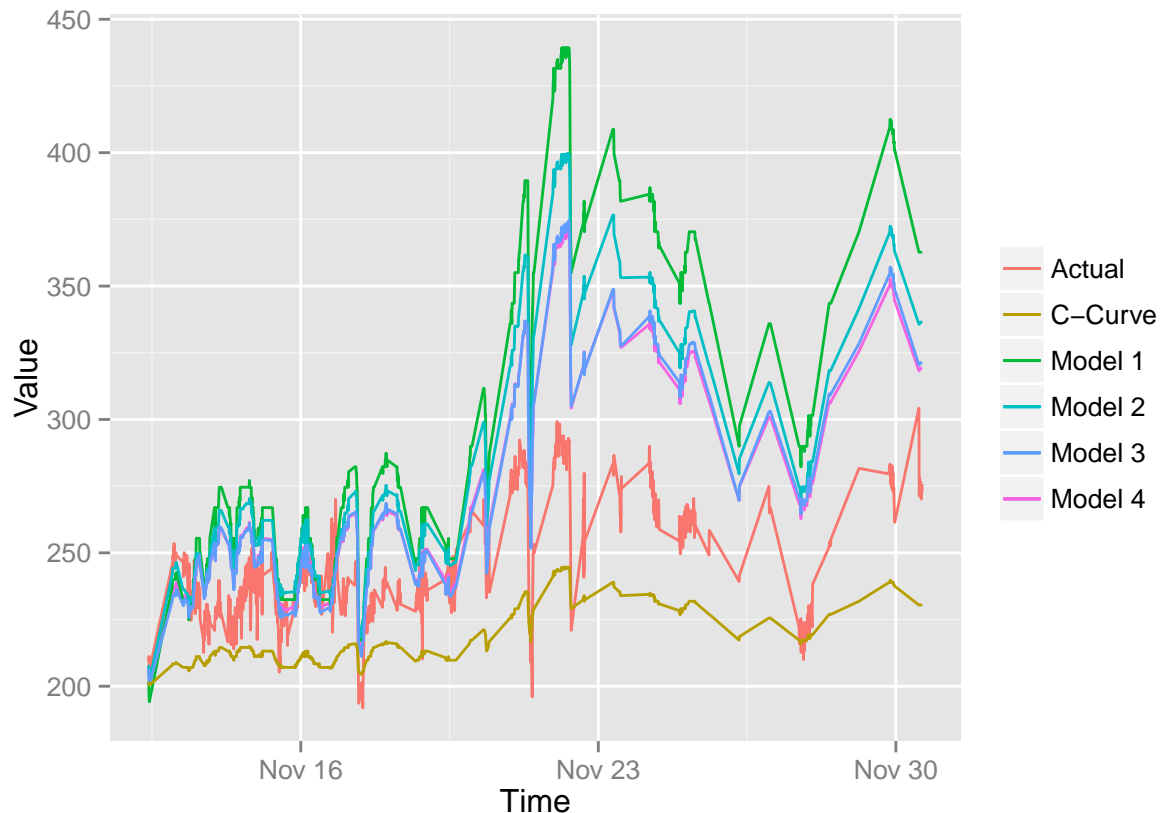
Regression with Value and aq resistance, adding temperature, slightly better  $r^2$ :

```
##
## Call:
## lm(formula = Value ~ aq_resistance + temp, data = aq.filtered)
##
## Coefficients:
## (Intercept)  aq_resistance      temp
##    -804.00231      287.27646     -1.04149
```

Multivariate with 3 , adding humidity and temperature increases the  $r^2$ :

```
##  
## Call:  
## lm(formula = Value ~ aq_resistance + humidity + temp, data = aq.filtered)  
##  
## Coefficients:  
## (Intercept)  aq_resistance      humidity      temp  
## -739.8246423  259.8208715    0.4408594   -0.3774600
```

Regression analysis



As you can see the calibrated curve method has poor response, but the regression models 3 and 4 show almost intandem pattern wrt to aq egg.

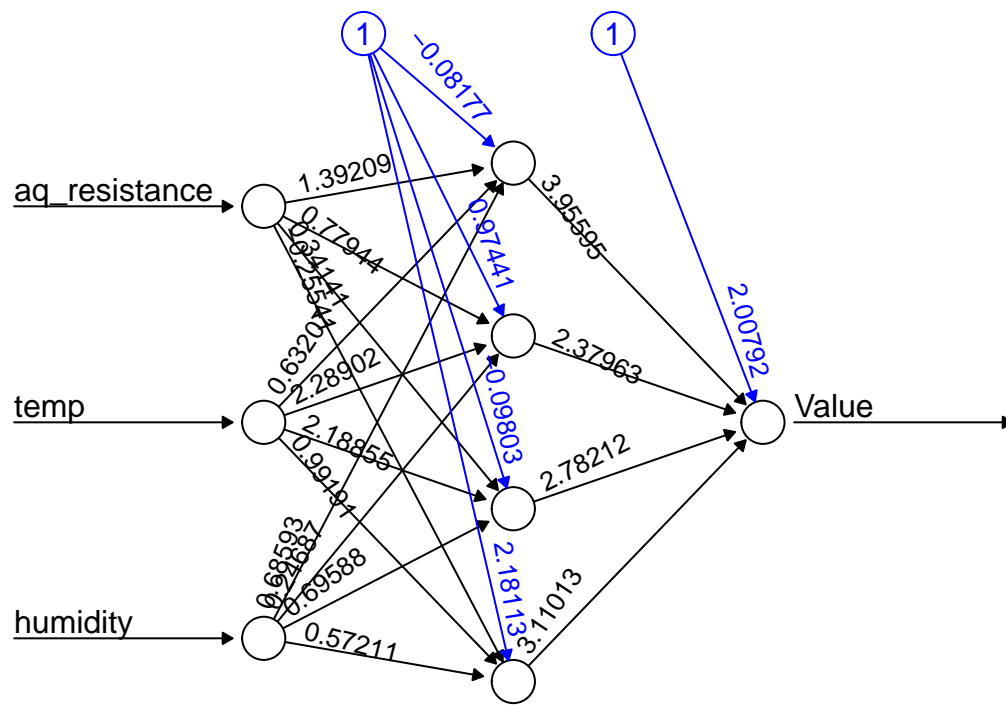
ANN analysis

```
## build the neural network (NN)  
aqnet <- neuralnet(Value ~ aq_resistance + temp+ humidity, aq.filtered, hidden = 4,  
  lifesign = "minimal", linear.output = FALSE, threshold = 0.1)
```

```
## hidden: 4    thresh: 0.1    rep: 1/1    steps:      34    error: 13667701.72    time: 0.01 secs
```

Question here is that, what does the above mean , Steps ~25 and Error in millions

```
## plot the NN  
plot(aqnet, rep = "best")
```



Error: 13667701.721795 Steps: 34