Identification of Contributing Factors of PM_{10} Concentration at Concrush PTY LTD

Andrew Furst
17701310@students.ltu.edu.au
La Trobe University

Contents

1 Abstract

The effects of ambient pressure, temperature, relative humidity, wind, truck traffic, truck tonnage and machine crushing on PM_{10} concentration were examined through multiple linear regression. The effects of ambient pressure, temperature, relative humidity and truck traffic were significant (p-value < 0.001). Daily Average relative humidity, daily truck traffic, and daily truck tonnage were highly correlated with daily average PM_{10} however daily truck tonnage was not significant due to multicollinearity, when the daily Average relative humidity was less than 80% and the daily truck traffic was over 150 trucks the likelihood of the daily average PM_{10} exceeding 0.05 mg/m^3 increased by 400%. There was no significant evidence that higher wind speeds or winds in certain directions are associated with an increase in PM_{10} concentration. Predictive modelling showed that reducing the number of days with high number of truck traffic whilst still maintaining yearly truck traffic results in a significant reduction in both the yearly average PM_{10} and the daily average PM_{10}. Simulation of increased yearly truck tonnage showed an exponential relationship between increased yearly tonnage and yearly average PM_{10}.

2 Introduction

This study investigates the dependency of PM_{10} levels at Concrush Pty Ltd on weather factors such as wind, temperature, ambient pressure and humidity as well as physical operations at the facility such as truck traffic and crushing operations. PM_{10} refers to particulate matter with an equivalent aerodynamic diameter of 10 micrometres or less. Concrush Pty Ltd is a concrete and building materials recycling company located in Teralba, NSW Australia and must operate in accordance with the National Environment Protection (Ambient Air Quality) Measure 2021 [3] which has a maximum daily average concentration standard of 50 μg/m^3 PM_{10} and a maximum yearly concentration standard of 25 μg/m^3 PM_{10}. The data for the analysis comes from 3 sources: readings from a QAMS Dust Master Pro 7000 [4], truck traffic logs that record the details of trucks that enter and exit the facility and crushing data which contains logs of the material crushing on site. Dust monitoring reports conducted by RCA Australia a company that among other services provides air quality assessment found that for some months of the year PM_{10} exceeds both the daily and annual average criteria. Multiple linear models will be used to identify the relationship between wind speed and wind direction and how it effects PM_{10} concentration. Linear models will be built to identify the statistically significant contributing factors of PM_{10} concentration. The results of these linear models will hopefully be able to guide Concrush towards effective measures for PM_{10} reduction such as windbreakers to block winds from directions that contributed significantly to PM_{10}. Concrush intends to seek approval for increased operations from 250,000 tonnes/year to 350,000 tonnes/year, linear modelling will be used to predict the effect of increasing tonnage/year on PM_{10} concentration.

3 Methodology

3.1 Multiple Linear Regression

The lm function in RStudio fits data to a linear model specified in the function to generate a linear model that maximises R^2. For a dataset with $i = n$ observations where y_i is the dependent variable, x_i are the explanatory variables, β_0 is the constant term, β_j are the β coefficients and ϵ is the error term, the linear model is given by formula 1 below.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_j x_{ij} + \epsilon \quad (1)$$

The summary function in RStudio prints the results and summary statistics for a given model, a brief description of the important summary outputs referred to in this analysis is given below.

R^2: Also known as the coefficient of determination is the proportion of variance for the dependent variable that can be explained by the independent variables. In multiple regression the addition of more independent variables to the model will always result in an increase in R^2.

Adjusted $\mathbf{R^2}$: Adjusted R^2 takes into account the number of independent variables in the model and only increases when new terms improve the model more than would be expected by chance.

Coefficient Estimate: The coefficient estimates are the estimates of the β values which are multiplied by the explanatory variables (x_i) to determine their contribution to the independent variable y_i. The coefficient estimate of β_0 is an exception and is the estimate of the constant term or 'y-intercept' which is the y_i estimate when all x_i are 0.

Coefficient t-value: The coefficient t-value is a measure of how many standard deviations the coefficient estimate is from 0. The magnitude of a coefficients t-value can be used to estimate the importance of the variable.

Coefficient p-value: The p-value for a coefficient tests the null hypothesis that the variable has no correlation with the independent variable, if the p-value is less than 0.05 the null hypothesis is rejected.

3.2 Data Wrangling

RStudio was used for the analysis with the following packages installed: tidyverse, dplyr, corrplot, ggpubr, moments and scales. Microsoft Excel was also used for the purpose of reformatting date and time variables where RStudio was insufficient.

The primary data source provided for the analysis was the dust data which contained the readings from a QAMS Dust Master Pro 7000 with Weather Sensor from Thompson Environmental Systems [4]. The Dust Master is located at the weighbridge near the east side entrance to the facility from Racecourse Rd. The Dust Master records the time, date, dust concentration ($PM_{2.5}$, PM_{10} and PM_{Tot}), weather conditions and Dust Master settings at 5 minute intervals continuously, for the full set of recorded values see table 6. The data files provided contained observations from the 30th of June 2020 to the 28th of February 2021 and were in .txt file format. These data files were imported into RStudio and joined to make a single Dust dataset which contained 69,857 observations. The variable Dew Point contained missing values for 0.005% of the observations in the dataset however as Dew Point was not used in the analysis no missing values were imputed. The dust data contained timestamps given by a date and time of day of recording for each observation, from these variables a single time variables measured in minutes was created. The time for each observation was calculated as the number of minutes that had passed since midnight on April 1st of 2020 (the date of the first truck movement recorded). Due to either their repeated measure as a result of the addition of the weather sensor or their irrelevance to the analysis the following variables were removed from the dataset: SV, LPM, PUMP, M1.WD.SD, M1.T2, S.T, S.RH and M1.AP.

To investigate the hypothesis that there are winds of a certain magnitude and direction that are associated with an increase in PM_{10} levels, the wind speed and wind direction variables were used to impute wind direction 'vectors'. For each observation categorical variables were created to represent different wind direction vectors, each wind direction vector corresponded to a range of values from the Wind direction variable. For each observation each wind direction 'vector' variable was assigned a 0 except for the 'vector' whose direction corresponded to the observations wind direction. Each wind direction vector was then multiplied by the observation's wind speed in order to give magnitude to the variable. Three sets of wind direction 'vectors' were created from the dataset, the first with 4 wind directions, the second with 8 and the third with 16. An observation belonged to a specific wind direction vector if its wind direction was greater than or equal too Angle 1 and less than Angle 2 (see tables 9, 10 and 11), with the exception of North wind where the wind direction needed to be greater than or equal too Angle 1, or less than Angle 2. For example if an observation had a wind speed of 3 m/s and a wind direction of 40° then for 4 wind directions North would be assigned a value of 3 and the others a value of 0, for 8 wind directions North East would be assigned a value of 3 and the others 0 and for 16 wind directions North North East would be assigned a value of 3 and the others 0.

Files containing the the details of truck movements in and out of the facility were imported into RStudio, these files were in .csv format and contained observations from the 1st of April 2020 to the 27th of February 2021 (Site was closed on the 28th of February). These files formed the Truck data which contains the time of entry and exit for all trucks entering the Concrush facility, additionally the weight in tonnes of both the truck and its load as well as a description of the load are recorded (for full set of recorded variables see table 7). The direction for each truck is also recorded, this refers to whether the truck is taking material in or out of the compound and it also refers to when the recording was taken, if the direction was 'IN' then the truck details were recorded when the truck entered the facility through the weighbridge and for 'OUT' the details were recorded when it left the facility. The time in minutes passed since midnight on April 1st of 2020 was calculated for both the Time In and Time Out as well as a new variable called Time on Site which was given by the Time Out - Time In. The original dataset contained 54,198 observations which was reduced to 45,937 observations once repeated entries were removed, the dataset was further reduced to 43,238 observations when observations with 0 Gross, Tare and Net weight were removed.

From the dataset 16,619 (38%) of the observations had a Time on Site of 0, that is that there Time In and Time Out were the same. These observations have a time on site of 0 as a result of their time of entry or exit only being recorded when they've either come 'IN' or 'OUT' of the site, for a truck whose time of entry and exit are equal and has a Direction of 'IN' this means that the time of entry is correct and the time of exit is not, and vice versa for trucks whose time of entry and exit are equal and have a Direction of 'OUT'. To impute these incorrect Time In and Time Out values the dataset was split into two, one dataset containing only entries whose Time on Site is 0 and the other whose Time on Site is greater than 0. The dataset that contained only positive values for Time on Site was then split further in to subsets based on the material type in the observations description. Then for each of these subsets of data a linear model was created using the lm function in RStudio of the form Time on Site (mins) $= \beta_0 + \beta_1 \times$ Net (tonnes) $+ \epsilon$. A linear model of the same form was also created using the entire dataset of non-zero Time on Site observations. For each observation in the full Truck dataset that had a Time on Site of 0, the Time on Site was imputed using the linear model that matched the material type of the observation, where there was no matching material type the model made from the full dataset was used. Using the direction and Time on Site the observations with the same time of entry and exit were corrected, if the truck was entering the site when the observation was recorded the time of exit was corrected by adding to it the time on site, for trucks that were exiting the site the time of entry was corrected by subtracting from it the time on site.

Using the Truck dataset 3 new variables were imputed for the Dust dataset: Trucks In, Trucks Out and Tonnage. The Trucks In is the number of trucks that entered the site in the 5 minutes before the observation Time, the Trucks Out is the number of trucks that exited the site in 5 minutes before the observation Time and the tonnage is the total tonnes (NET) that came in or out of the site in the five minutes before the observation Time.

The final set of data was the Crushing data, which contained timestamped recordings of the crushing operations. 3 sets of data were provided by Concrush: 30 Minute, daily and monthly sensor readings. For the analysis the 30 minute sensor readings were used as they offered more precise readings. 30 minute sensor readings means that whenever the machine begins crushing, every 30 minutes a reading is taken that describes the material type being crushed, the density of the material, amount of material being crushed etc (see table 8 for the full list of variables recorded) with a final reading when the crushing stops which can have a duration of less than 30 minutes. Readings are also taken when the machine is turned on or off and at the start or end of a new measurement. As the crushing operations are of the only concern for the analysis observations containing an amount of 0 were removed from the dataset. The dataset also contained readings that exceeded a rate of 250 tonnes/hour which was identified by Concrush as the upper limit of the machines capability. For each observation that exceeded that 250 tonnes/hours the amount crushed was replaced with the maximum amount that could be crushed within that time period based on the density of the material. The amount crushed was determined to exceed the maximum using the formula below, where density refers to the density of the material being crushed.

$$\text{Maximum Amount Crushed } (m^3) = \frac{250\ (tonnes/h)}{\text{Density } (tonnes/m^3)} \times \frac{\text{Duration } (mins) - \text{Idle Time } (mins)}{60\ (mins/h)}$$

The Crushing dataset contains three timestamp recordings: Server Timestamp, Smartphone Timestamp and Record Timestamp. The former two refer to when the recording was uploaded to either the server or the smartphone, whereas the Record Timestamp refers to the actual time of the recording. From the Record Timestamp the number of minutes that had passed since midnight on April 1st of 2020 was imputed.

Given that data is recorded in 30 minute intervals from when the machine is turned on with a final recording when its turned off a 5 minute average of the amount of material crushed was imputed for each observation in the Crushing dataset using the formula below.

$$\text{5 Minute Average Crushing Amount } (m^3) = \frac{5\ (minutes) \times \text{Amount Crushed } (m^3)}{\text{Measurement Duration } (minutes)}$$

Another variable in the Crushing dataset was imputed, the Start Time for each observation was calculated by subtracting the Duration from the Time. From this dataset the new variable Crushing Amount was imputed for the Dust dataset. Each observation in the Dust dataset was checked against the Crushing dataset, if the Time for an observation in the Dust dataset was between the Start Time and the Time (Time of Recording) for an observation in the Crushing dataset then the Crushing Amount for that observation was the respective 5 Minute Average (m^3) Crushing Amount. Where no observation in the Dust dataset had a matching Start Time and Time from the Crushing dataset a Crushing Amount of 0 was assigned.

Using the Dust dataset that was recorded in 5 minute intervals a 24 Hour Dust dataset was created. The 24 Hour Dust dataset took the daily average of the following variables: PM_{10}, Temperature, Humidity, Ambient Pressure and Wind Speed as well as the daily total of the following variables: Trucks In, Tonnage (in and out) and Crushing Amount. For the wind direction vectors the daily total was also taken, for example the 24 Hour Dust dataset with 4 wind directions contained the variables in table 1 below.

Table 1: 24 Hour Dust Dataset with 4 Wind Directions

Variable	Description	Units/Format
Average PM_{10}	Daily Average PM_{10}	mg/m^3
Traffic Total	Daily Total Trucks In	count
Tonnage Total	Daily Total Tonnage In & Out	tonnes
Average Temperature	Daily Average Temperature	°C
Average Humidity	Daily Average Humidity	%
Average Ambient Pressure	Daily Average Ambient Pressure	mBar
Crushing Amount Total	Daily Total Crushing Amount	m^3
Average Wind Speed	Daily Wind Speed	m/s
North Wind Total	Daily North Wind Total	m/s
East Wind Total	Daily East Wind Total	m/s
South Wind Total	Daily South Wind Total	m/s
West Wind Total	Daily West Wind Total	m/s

3.3 Data Exploration

The correlation coefficients of the non-wind variables in the 24 Hour Dust dataset were produced in RStudio using the cor function with the default method 'pearson' which indicates that the correlation coefficient to be computed is the pearson correlation coefficient. Pearson correlation coefficient (r) is a measure of linear association of two variables [2], variables with a high in magnitude linear association with the dependent variable of a linear model would be expected to be significant whereas variables with a low in magnitude linear association with the dependent variable of a linear model would be expected to be insignificant. Independent variables with a high in magnitude linear association with each other can cause multicollinearity [1]. Given two vectors x and y of length n with respective means m_x and m_y the pearson correlation coefficient r is given by,

$$r = \frac{\sum(x - m_x)(y - m_y)}{\sqrt{\sum(x - m_x)^2(y - m_y)^2}}$$

These correlation coefficients were plotted using the corrplot function from the package of the same name. Scatter plots of all of the non-wind variables against Average PM_{10} were produced using the ggplot function.

3.4 Transformations of the dependent variable

Transformations of the dependent variable Average PM_{10} were performed and 4 linear regression models were computed using the lm function to test the efficacy of transforming the dependent variable. Given that the Average PM_{10} was right skewed (most of the observations had a low Average PM_{10}) a Log, Square Root and Log of the Square Root transformations were tested, with R^2 used to assess them. The four models tested have 4 wind speeds and differ only in the dependent variable (see Equation 2,Equation 3, Equation 4 and Equation 5 below).

$$\begin{aligned}\text{Average PM}_{10} = {} & \beta_0 + \beta_1 \times \text{Traffic Total} + \beta_2 \times \text{Tonnage Total} + \beta_3 \times \text{Average Temperature} \\ & + \beta_4 \times \text{Average Relative Humidity} + \beta_5 \times \text{Average Ambient Pressure} \\ & + \beta_6 \times \text{Crushing Amount Total} + \beta_7 \times \text{North Wind} + \beta_8 \times \text{East Wind} \\ & + \beta_9 \times \text{South Wind} + \beta_{10} \times \text{West Wind} + \epsilon\end{aligned} \tag{2}$$

$$\begin{aligned}\log(\text{Average PM}_{10}) = {} & \beta_0 + \beta_1 \times \text{Traffic Total} + \beta_2 \times \text{Tonnage Total} + \beta_3 \times \text{Average Temperature} \\ & + \beta_4 \times \text{Average Relative Humidity} + \beta_5 \times \text{Average Ambient Pressure} \\ & + \beta_6 \times \text{Crushing Amount Total} + \beta_7 \times \text{North Wind} + \beta_8 \times \text{East Wind} \\ & + \beta_9 \times \text{South Wind} + \beta_{10} \times \text{West Wind} + \epsilon\end{aligned} \tag{3}$$

$$\begin{aligned}\sqrt{\text{Average PM}_{10}} = {} & \beta_0 + \beta_1 \times \text{Traffic Total} + \beta_2 \times \text{Tonnage Total} + \beta_3 \times \text{Average Temperature} \\ & + \beta_4 \times \text{Average Relative Humidity} + \beta_5 \times \text{Average Ambient Pressure} \\ & + \beta_6 \times \text{Crushing Amount Total} + \beta_7 \times \text{North Wind} + \beta_8 \times \text{East Wind} \\ & + \beta_9 \times \text{South Wind} + \beta_{10} \times \text{West Wind} + \epsilon\end{aligned} \tag{4}$$

$$\begin{aligned}\log(\sqrt{\text{Average PM}_{10}}) = {} & \beta_0 + \beta_1 \times \text{Traffic Total} + \beta_2 \times \text{Tonnage Total} + \beta_3 \times \text{Average Temperature} \\ & + \beta_4 \times \text{Average Relative Humidity} + \beta_5 \times \text{Average Ambient Pressure} \\ & + \beta_6 \times \text{Crushing Amount Total} + \beta_7 \times \text{North Wind} + \beta_8 \times \text{East Wind} \\ & + \beta_9 \times \text{South Wind} + \beta_{10} \times \text{West Wind} + \epsilon\end{aligned} \tag{5}$$

3.5 Modeling Wind Speeds

For the wind speed analysis 5 linear regression models were computed to assess and compare the different implementations of wind speed and wind direction variables. Given that the log transformation model performed the best, all 5 of these models used log(Average PM_{10}) as the independent variable. The equation for the first model with no wind variables is given below (see Equation 6), for the other 4 models see equations 13, 14, 15 and 16 in Appendix C.

$$\begin{aligned}\log(\text{Average PM}_{10}) = {} & \beta_0 + \beta_1 \times \text{Traffic Total} + \beta_2 \times \text{Tonnage Total} + \beta_3 \times \text{Average Temperature} \\ & + \beta_4 \times \text{Average Relative Humidity} + \beta_5 \times \text{Average Ambient Pressure} \\ & + \beta_6 \times \text{Crushing Amount Total} + \epsilon\end{aligned} \tag{6}$$

3.6 Modeling Log(Average PM_{10})

Given the results of the wind analysis log(Average PM_{10}) was modelled against all of the variables in the 24 Hour Dust dataset with the exception of the wind speed and wind direction variables (See model 7). To check for multicollinearity, additional models were created by using only one of the highly correlated independent variables at a time (see models 8, 9 and 10).

$$\begin{aligned}\log(\text{Average PM}_{10}) = \beta_0 &+ \beta_1 \times \text{Traffic Total} + \beta_2 \times \text{Tonnage Total} + \beta_3 \times \text{Average Temperature} \\ &+ \beta_4 \times \text{Average Relative Humidity} + \beta_5 \times \text{Average Ambient Pressure} \\ &+ \beta_6 \times \text{Crushing Amount Total}\end{aligned} \tag{7}$$

$$\begin{aligned}\log(\text{Average PM}_{10}) = \beta_0 &+ \beta_1 \times \text{Traffic Total} + \beta_2 \times \text{Average Temperature} \\ &+ \beta_3 \times \text{Average Relative Humidity} + \beta_4 \times \text{Average Ambient Pressure}\end{aligned} \tag{8}$$

$$\begin{aligned}\log(\text{Average PM}_{10}) = \beta_0 &+ \beta_1 \times \text{Traffic Total} + \beta_2 \times \text{Average Temperature} \\ &+ \beta_3 \times \text{Average Relative Humidity} + \beta_4 \times \text{Average Ambient Pressure} \\ &+ \beta_5 \times \text{Crushing Amount Total}\end{aligned} \tag{9}$$

$$\begin{aligned}\log(\text{Average PM}_{10}) = \beta_0 &+ \beta_1 \times \text{Traffic Total} + \beta_2 \times \text{Tonnage Total} + \beta_3 \times \text{Average Temperature} \\ &+ \beta_4 \times \text{Average Relative Humidity} + \beta_5 \times \text{Average Ambient Pressure}\end{aligned} \tag{10}$$

Multicollinearity was assessed by observing changes in variable significance from the results of the initial model with all of the correlated variables included, if any variables were not significant in the presence of the others but were significant on their own there is evidence of multicollinearity. To identify the variables of most importance the absolute value of the coefficients associated t-statistic (t-value) is used, where a higher value in magnitude is associated with a higher variable importance.

3.7 Average PM_{10} Estimates

Using model 8 log(Average PM_{10}) was estimated for the 244 observations in the 24 Hour Dust dataset, transformed to Average PM_{10} and then compared to the actual values of Average PM_{10} in the dataset. Redistribution of the the truck traffic was also tested by sampling each observations Traffic total (only for days where Concrush was open) from a normal distribution centered at 158 which was the average Traffic total for open days for the dataset, a standard deviation of $\frac{70}{3}$ was chosen to keep the distribution highly centered about the mean. The estimated Average PM_{10} levels were then calculated using the new Traffic total values to assess the effect of reducing the number of days with really high Traffic total values without reducing the yearly total traffic.

Given that the average tonnage (NET) per Truck is ≈ 5.588 tonnes, the estimated yearly traffic total for a given yearly total tonnage is given by equation 11 below.

$$\begin{aligned}\text{Yearly Traffic Total} &= \frac{\text{Yearly Tonnage Total (tonnes)}}{\text{Average Tonnage per Truck (tonnes)}} \\ &= \frac{\text{Yearly Tonnage Total (tonnes)}}{5.588 \text{ (tonnes)}}\end{aligned} \tag{11}$$

Equation 11 was used to to predict the Yearly Traffic Total if the yearly tonnage was increased to 300,000, 350,000, 400,000, 450,000 and 500,000 tonnes. Given that the dataset only contains 244 days of data, the Yearly Traffic Total was multiplied by 244/365 and then divided by 198 (number of open days in the dataset) to estimate what the Average Daily Traffic Total for the 198 open days in the datatset would need to be to achieve the given Yearly Tonnage Total (see equation 12).

$$\text{Average Daily Traffic Total} = \text{Yearly Traffic Total} \times \frac{244}{365} \times \frac{1}{198} \tag{12}$$

For each of the 5 Yearly Tonnage Total increases (300,000, 350,000, 400,000, 450,000 and 500,000) a new Traffic Total variable was created by sampling each observations Traffic total (only for days where Concrush was open) from a normal distribution centered at the Average Daily Traffic Total calculated for that Yearly Tonnage Total, a standard deviation of $\frac{70}{3}$ was chosen to keep the distribution consistent with the previous traffic redistribution.

3.8 Visual Analysis

Using the ggplot function a scatter plot was created to assess if there is a link between the Daily Average Humidity and spraying operations, as there is no data regarding spraying operations, a factor variable indicating whether the site was 'Open' or 'Closed' was used instead. For the plot the y-axis was the Average Humidity and the x-axis was the Date with the points on the graph coloured to represent 'Open' or 'closed'.

Using the ggplot function variables in the final linear model were plotted against each other with a factor variable indicating whether the Average PM_{10} for that observation exceeded the daily limit of 0.05 mg/m^3.

4 Results

Initially the linear regression models produced for this analysis used the original dataset containing observations taken at 5 minute intervals however modeling PM_{10} using this data resulted in an $R^2 \approx 0.25$ and modeling $\log(PM_{10})$ improved the model to an $R^2 \approx 0.45$. Given these results and the improvement in the model found when using the 24 Hour dataset there is likely too high a degree of variability in the PM_{10} values when measured at 5 minute intervals to produce an effective model using the 5 minute data, hence the remainder of the analysis was carried out using the 24 Hour Data.

4.1 Data Exploration

The histogram of the Daily Average PM.10 from the 24 hour dataset (see figure 1) shows that while the majority of the 244 days in the dataset have a Daily Average PM_{10} of less than 0.05 mg/m^3 there were 54 days that exceeded that limit.

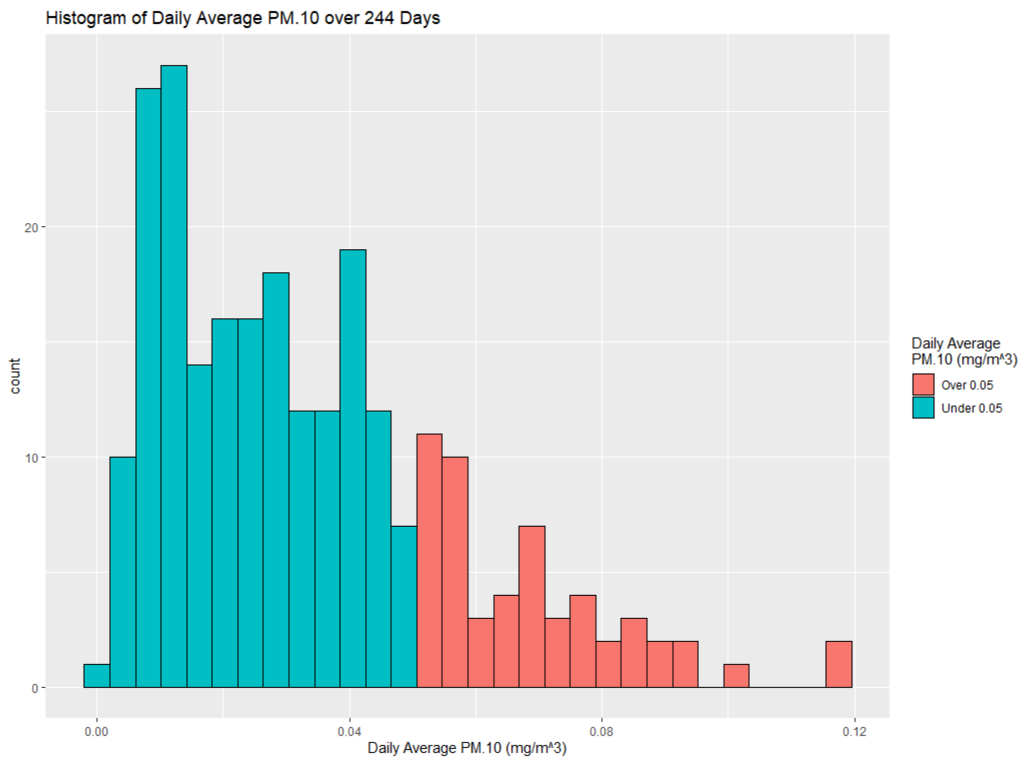


Figure 1: Histogram of Daily Average PM.10 from the 24 Hour Dataset

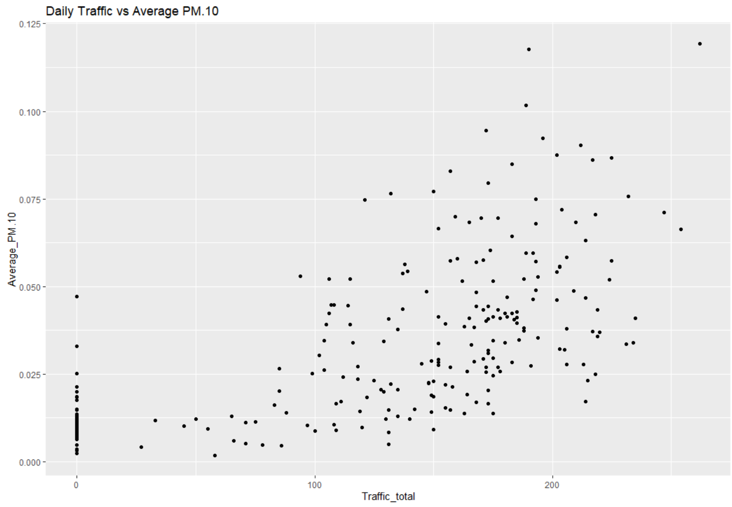
The correlation plot (see figure 2) shows that Traffic total and Tonnage Total have a strong positive correlation with PM_{10} whereas Average Crushing Amount has a moderate positive correlation and both Average Ambient Pressure and Average Wind Speed have a weak positive correlation with Average PM_{10}. Average Humidity has a moderate negative correlation with PM_{10} and Average Temperature has a negative weak correlation with Average PM_{10}. This suggests that Traffic and Tonnage total are the biggest contributing factors of Average PM_{10} in the dataset and conversely Average Humidity contributes the most in reducing Average PM_{10} levels. Traffic total is also highly positively correlated to Tonnage total as well as moderately positively correlated to Crushing Amount total, although this result in not unexpected, it does mean that any linear regression model fitted with more than one of these variables will need to be checked for the effects of multicollinearity.

From the plots of the dependent variables against Average PM_{10} (see figure 3) there appears to be a positive relationship between Traffic total (a) and Average PM_{10} as well as Tonnage total (b) and Average PM_{10}. There also appears to be negative relationship between Average Humidity (d) and Average PM_{10}. There is no obvious relationship between the other dependent variables and Average PM_{10}. These results are as expected based on the correlation plot (see figure 2).

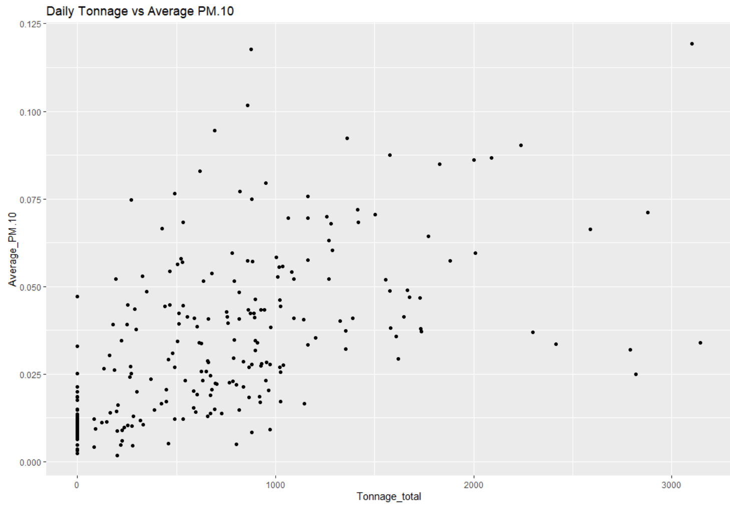
It should be noted that the position of the Dust Master at the weighbridge might be effecting the correlation of the Traffic Total variable with Daily Average PM.10 as all trucks the enter and exit the site do so at the weighbridge. This means that any dust the trucks kick up is more likely to be recorded by the Dust Master than dust produced by crushing (Crushing Amount total) or the drop off and collection of material (Tonnage Total) due to their relative proximity to the Dust Master.

See appendix D for a summary table of the 24 Hour Dataset.

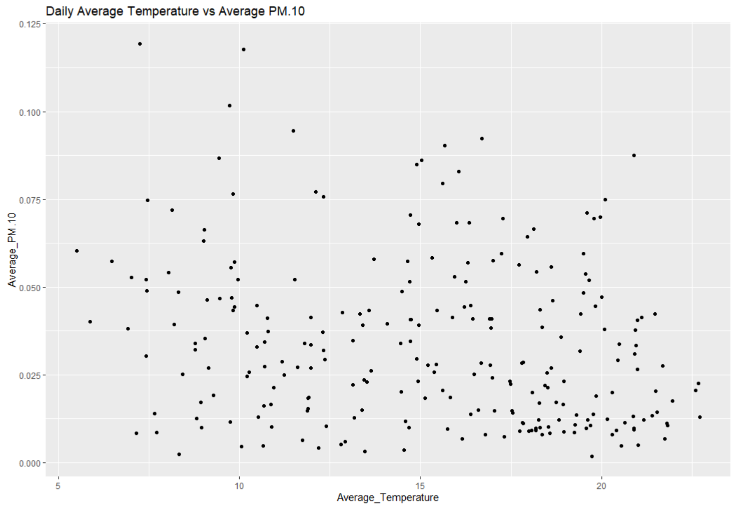
Figure 2: Correlation Plot of 24 Hour Data



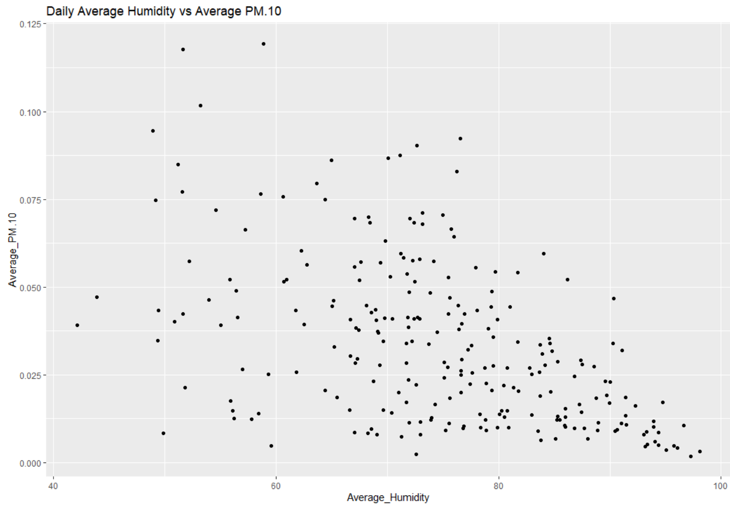
(a) Traffic Total vs Average PM_{10}



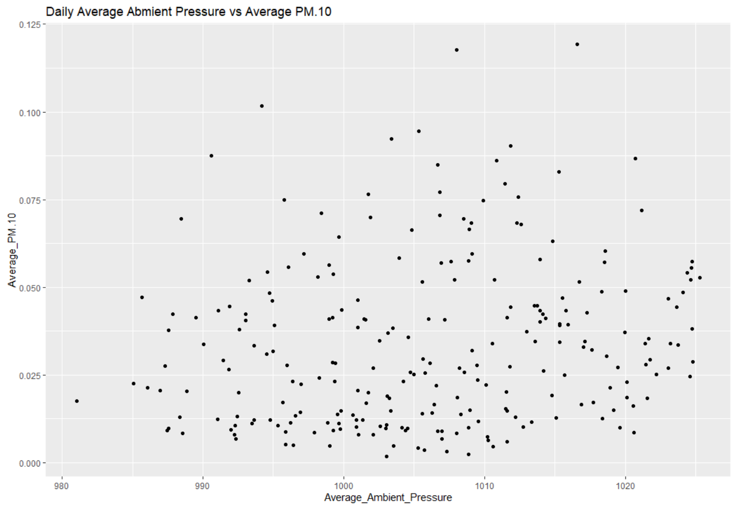
(b) Tonnage Total vs Average PM_{10}



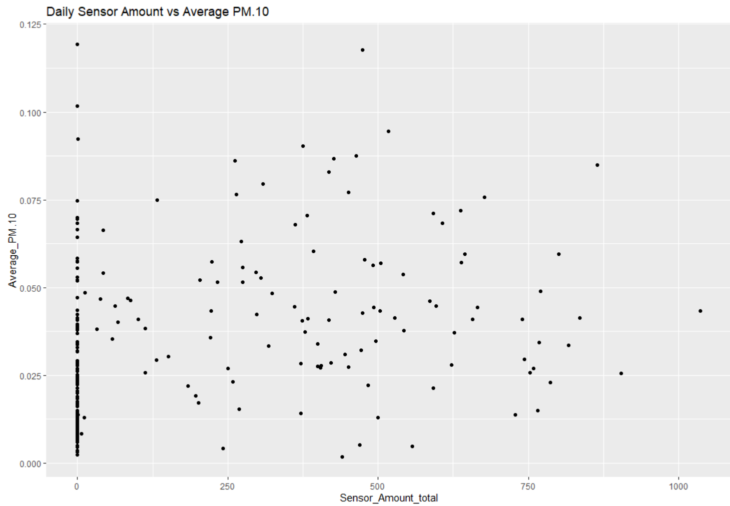
(c) Average Temperature vs Average PM_{10}



(d) Average Humidity vs Average PM_{10}



(e) Average Ambient Pressure vs Average PM_{10}

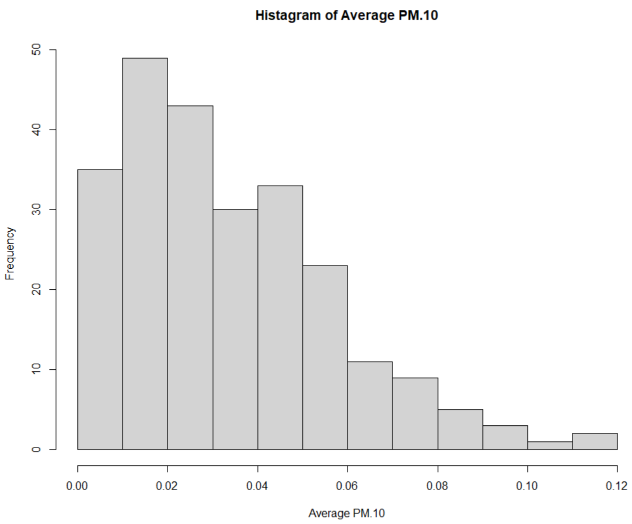


(f) Crushing Amounr Total vs Average PM_{10}

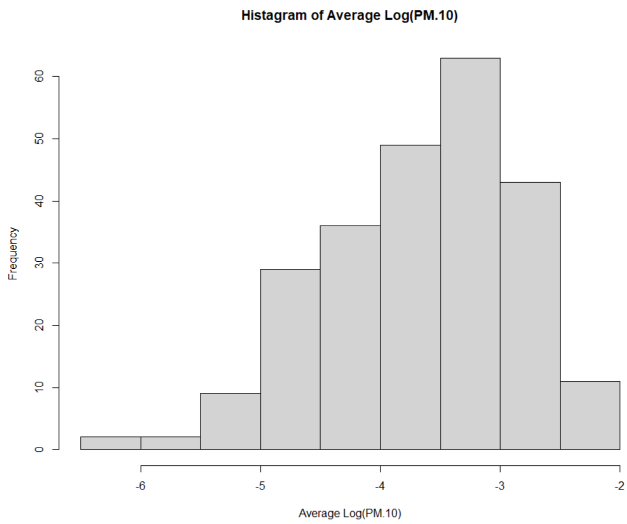
Figure 3: Dependent Variables vs Average PM_{10} for the 24 Hour Dataset

4.2 PM_{10} Transformations

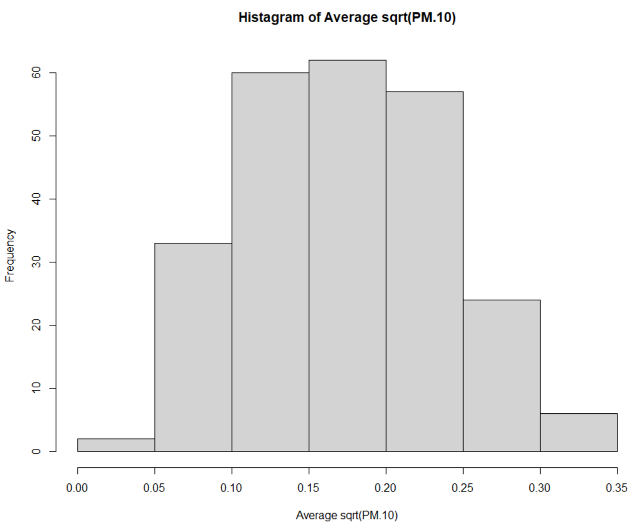
Histograms of Average PM_{10}, log(Average PM_{10}), sqrt(Average PM_{10}) and log(sqrt(Average PM_{10})) are shown in the figure below (4). All three transformations are more normally distributed with both the log(Average PM_{10}) and log(sqrt(Average PM_{10})) being left skewed.



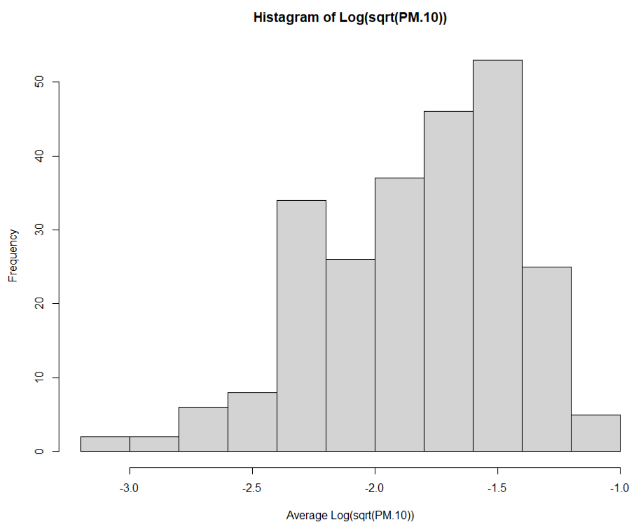
(a) PM_{10}



(b) log(PM_{10})



(c) sqrt(PM_{10})



(d) log(sqrt(PM_{10}))

Figure 4: Histograms of PM_{10} Transformations

From the four linear models created to test the effectiveness of transforming the independent variable the Average PM_{10} model (2) had the lowest R^2 result with $R^2 = 0.613$, both the log(Average PM_{10}) (3) and log(sqrt(Average PM_{10})) (5) models resulting in a significant improvement in R^2 from 0.613 to 0.735, the sqrt(Average PM_{10}) (4) model also showed a significant improvement with an $R^2 = 0.694$.

All four models had 4 common highly significant (p-val < 0.001) effects: Traffic Total, Average Temperature, Average Humidity and Average Ambient Pressure, Tonnage and Crushing Amount were not significant for all four models.

4.3 Wind Analysis

When log(Average PM_{10}) was modeled without any wind effects the model had an $R^2 = 0.71$ and an Adj-$R^2 = 0.703$.

Modeling with only the wind speed effect with no wind direction, the wind speed effect was significant (p-val = 0.000753) and the effect estimate was negative. Hence an increase in the wind speed was associated with a decrease in log(Average PM_{10}). The addition of wind speed to the model improved the model with $R^2 = 0.724$ and Adj-$R^2 = 0.715$.

Modeling with 4 wind directions resulted in a slightly improved model with an $R^2 = 0.735$ and an Adj-$R^2 = 0.723$. East and South wind effect estimates were both negative and significant with p-values < 0.01, while North (p-val = 0.513) and West (p-val = 0.224) effect estimates were both found to be insignificant.

Modeling with 8 wind directions again resulted in a slightly improved model an $R^2 = 0.746$ and an Adj-$R^2 = 0.730$. Only South and Sough East wind direction effects were significant with p-values < 0.01 and both effects were also negative.

Modeling with 16 wind directions resulted in a slightly worse model than with 8 wind directions with an $R^2 = 0.751$ and an Adj-$R^2 = 0.726$. Although R^2 improved the lower Adj-R^2 indicates that this is likely just a result of the increased number of effects in the model. Of the 16 wind direction effects only South wind was found to be significant with p-val=0.0124 and a negative effect estimate.

Table 2: Significant Wind Effect Estimates

Number of Wind Directions	Wind Direction	Angle
4	East	45° to 135°
4	South	135° to 225°
8	South East	112.5° to 157.5°
8	South	157.5° to 202.5°
16	South	168.75° to 191.25°

The analysis of these models shows that there is no significant evidence of an increase in wind speed in any direction being associated with an increase in log(Average PM_{10}) and therefore an increase in Average PM_{10}. The analysis does show that for all three models with wind directions, that a southerly wind and possibly and south easterly or easterly wind is associated with a decrease in log(Average PM_{10}) and hence a decrease in Average PM_{10}. Given the location of the Dust Master Pro 7000 at the weighbridge, a possible reason for these unexpected results could be that these winds are blowing dust out of the work site thus increasing PM_{10} emissions but reducing the PM_{10} levels on site.

4.4 Model Analysis

The multiple linear regression model of log(Average PM_{10}) (8) has an $R^2 = 0.710$ and Adj-$R^2 = 0.703$ which suggest a good fit for the data. The model has 4 significant effects: Traffic total (p-val < 0.001), Average Temperature (p-val < 0.001), Average Humidity (p-val < 0.001) and Average Ambient Pressure (p-val < 0.001). Tonnage total (p-val = 0.560) and Crushing Amount total (p-val = 0.646) were insignificant (see table 3 for the coefficient estimates).

To check for multicollinearity between Traffic total, Tonnage total and Crushing Amount total the model was repeated with only one of the variables included in each model. Of the three models the Traffic total model was significantly better than the other two with an $R^2 = 0.709$ compared to the Tonnage total model ($R^2 = 0.592$) and the Crushing Amount total model ($R^2 = 0.394$). Each of the 3 variable effect estimates were significant (p-val < 0.001) in their respective models. Given that the Traffic total model had a significantly higher R^2 than the other two models it is likely that truck traffic is the primary cause of increased PM_{10} levels and that tonnage and Crushing amount are providing a weaker estimate of the amount of traffic on the site which is why their respective models perform worse, as Tonnage total has a higher correlation to Traffic total than Crushing Amount total its model performs better than the Crushing Amount total model. As a

Table 3: Coefficient Estimates

Coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	$-2.766e+01$	$5.216e+00$	-5.304	$2.60e-07$
Traffic total	$6.562e-03$	$6.871e-04$	9.550	$< 2e-16$
Tonnage total	$4.484e-05$	$7.681e-05$	0.584	0.559875
Average Temperature	$4.800e-02$	$1.217e-02$	3.945	0.000105
Average Humidity	$-3.292e-02$	$2.704e-03$	-12.177	$< 2e-16$
Average Ambient Pressure	$2.472e-02$	$5.121e-03$	4.827	$2.48e-06$
Crushing Amount total	$-5.632e-05$	$1.225e-04$	-0.460	0.646166

result the final model includes only 4 variables: Traffic Total, Average Temperature, Average Humidity and Average Ambient Pressure.

Effect estimates for the final model (4) indicate that Traffic total is the most significant predictor of log(Average PM_{10}) as it has the highest t-value of all 4 effect estimates, followed closely by Average Humidity. As this is a model of log(Average PM_{10}) the effect of these variables on Average PM_{10} is given as a percentage increase or decrease in Average PM_{10}.

For Traffic Total, as $\exp(6.794e-03) = 1.006817132 \approx 1.0068$ we expect to see a 0.68% increase in Average PM_{10} for a 1 unit increase in Traffic total from 0. For an increase of 100 in Traffic total from 0 we would expect a 97.3% increase in Average PM_{10} as $\exp((6.794e-03) \times 100) = 1.972694 \approx 1.973$. Furthermore for an increase of 200 in Traffic total from 0 we would expect a 289% increase in PM_{10} as $\exp((6.794e-03) \times 200) = 3.891521 \approx 3.89$. Given that Traffic total ranges from 0 to 262 in the dataset these estimates highlight the significant effect Truck Traffic has on PM_{10} levels. The effect of increased truck traffic is also not linear as the increase in Average PM_{10} from an increase in 200 to Traffic total is more than double that of the increase in 100 to Traffic total.

These results suggest that either reducing Truck Traffic or reducing the effect of Truck traffic should be the highest priority for Concrush in order to reduce their daily and yearly PM_{10} emissions. Given the there is an exponential relationship between Traffic total and Average PM_{10} reducing the the number of days with high traffic numbers while still maintaining the same total yearly traffic could also be effective at reducing PM_{10} emissions.

Table 4: Coefficient Estimates for Final Model

Coefficients	Estimate	Std. Error	t-value	p-value	exp(Estimate)
(Intercept)	$-2.725e+01$	$5.111e+00$	-5.332	$2.25e-07$	$1.463779e-12$
Traffic total	$6.794e-03$	$3.913e-04$	17.360	$< 2e-16$	1.006817132
Average Temperature	$4.653e-02$	$1.195e-02$	3.895	0.000127	1.04763
Average Humidity	$-3.267e-02$	$2.670e-03$	-12.238	$< 2e-16$	0.9678579
Average Ambient Pressure	$2.431e-02$	$5.016e-03$	4.846	$2.27e-06$	1.024608

From the plots of the final model (see appendix E) there is no evidence of significant non-linearity in the Residuals vs Fitted graph (a) and the residuals are relatively evenly distributed about 0 with no pattern in their distribution, as such there is no evidence of heteroscedasticity. From the Normal Q-Q plot (b) the residuals appear to be approximately normally distributed. From the Residuals vs Leverage plot (d) there appear to be no highly significant data points that have both a high leverage and a high standardized residual.

4.5 Average PM_{10} Estimates

Using the final model (8) and the associated coefficient estimates (4) the Yearly Average PM_{10} was estimated to be 0.0313 mg/m^3 and there were 36 days that exceeded the daily maximum of 0.05 mg/m^3 of Average PM_{10}. These results are lower than the actual Average PM_{10} values in the dataset, where the Yearly Average PM_{10} is 0.0338 mg/m^3 and 54 days exceeded the daily limit. These results show that the model in underestimating Average PM_{10}, given that the difference in the number of days exceeding the 0.05 mg/m^3 is significant compared the the difference in Yearly Average PM_{10} the model is likely underestimating large values of Average PM_{10}. In the histogram below (5) the higher values of Average PM_{10} are mostly actual values whereas the the bars for estimate values are larger for the lower Average PM_{10} values. Scatter plots of the actual vs estimate Average PM_{10} values (12) show that this can be seen especially between September and December (see figure 13) where most of the high Average PM_{10} values are the actual values.

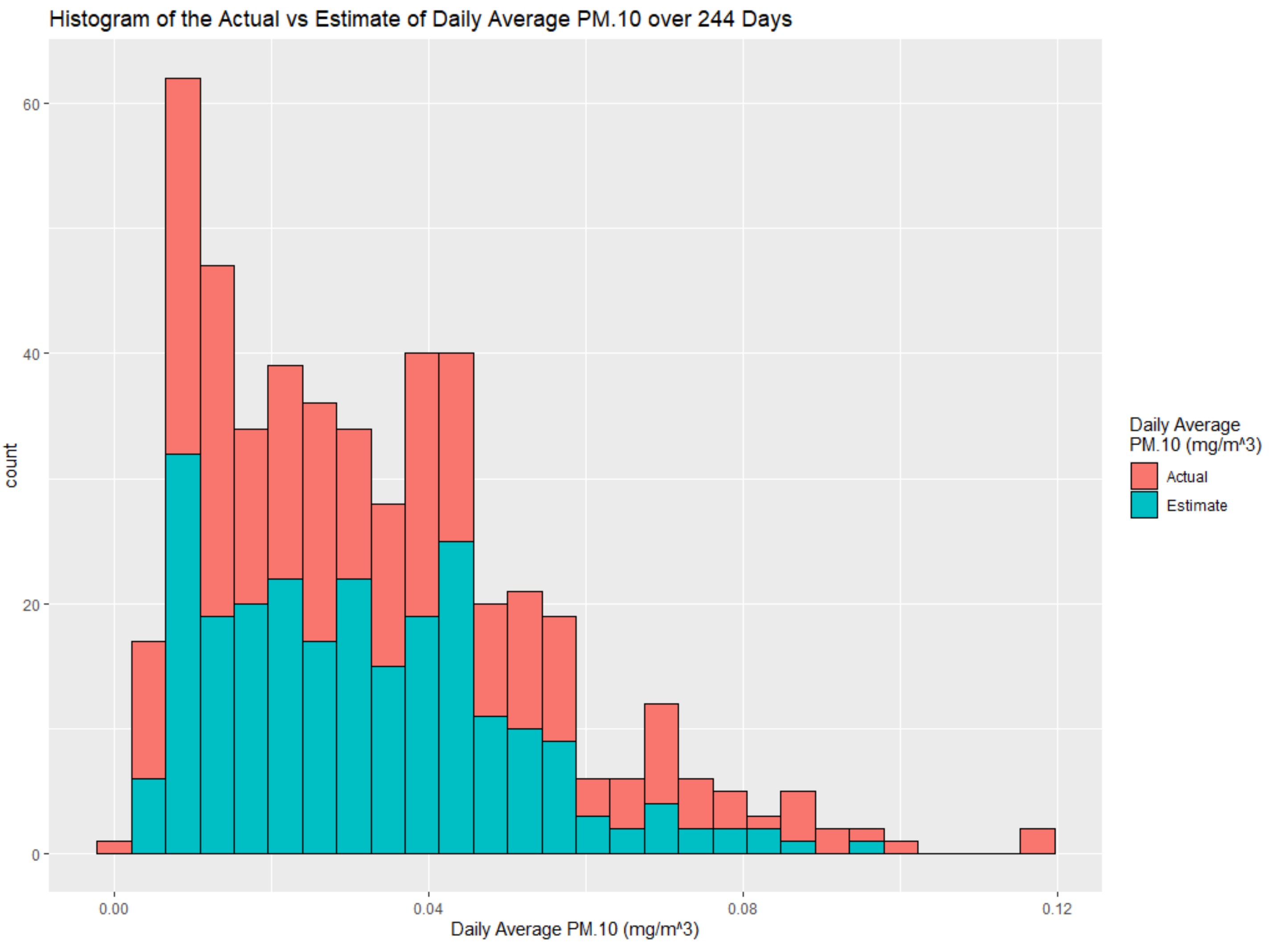


Figure 5: Actual vs Predicted Average PM_{10}

Redistributing the Traffic total variable to reduce the number of days with high Traffic total counts resulted in a reduction in both the estimated Yearly Average PM_{10} (from 0.0313 to 0.0292) and the estimated number of days that exceed the 0.05 mg/m^3 limit (from 36 days to 20 days). See appendix F to see the change in distribution for the Daily Traffic totals while maintaining the same Traffic total over the whole 244 days of the dataset (histogram is of only open days).

The total tonnage in the dataset was 174,803.55 ≈ 174,800 tonnes, hence the total yearly tonnage in the dataset was estimated to be 174,800 × (365/244) ≈ 261,000 tonnes. The estimates for the Yearly Average Tonnage and the number of days over the 0.05 mg/m^3 limit are shown in the table below (5), given that the estimate of Yearly Average PM_{10} for 261,000 tonnes is less than the actual Yearly Average PM_{10} little weight should be given to these findings. Figure 6 shows that there is an exponential relationship between Yearly Tonnage and Estimated Yearly Average PM_{10} that appears relatively linear for yearly tonnages less than 500,000 tonnes.

Table 5: Average PM_{10} Estimates for Increased Yearly Tonnage Total

Yearly Tonnage (tonnes)	Yearly Average PM_{10} (mg/m^3)	Days over 0.05 mg/m^3	Yearly Traffic Total
261,000	0.0292	20	46,707
300,000	0.0339	34	53,686
350,000	0.0411	78	62,634
400,000	0.0499	124	71,582
450,000	0.0608	157	80,529
500,000	0.0742	177	89,477

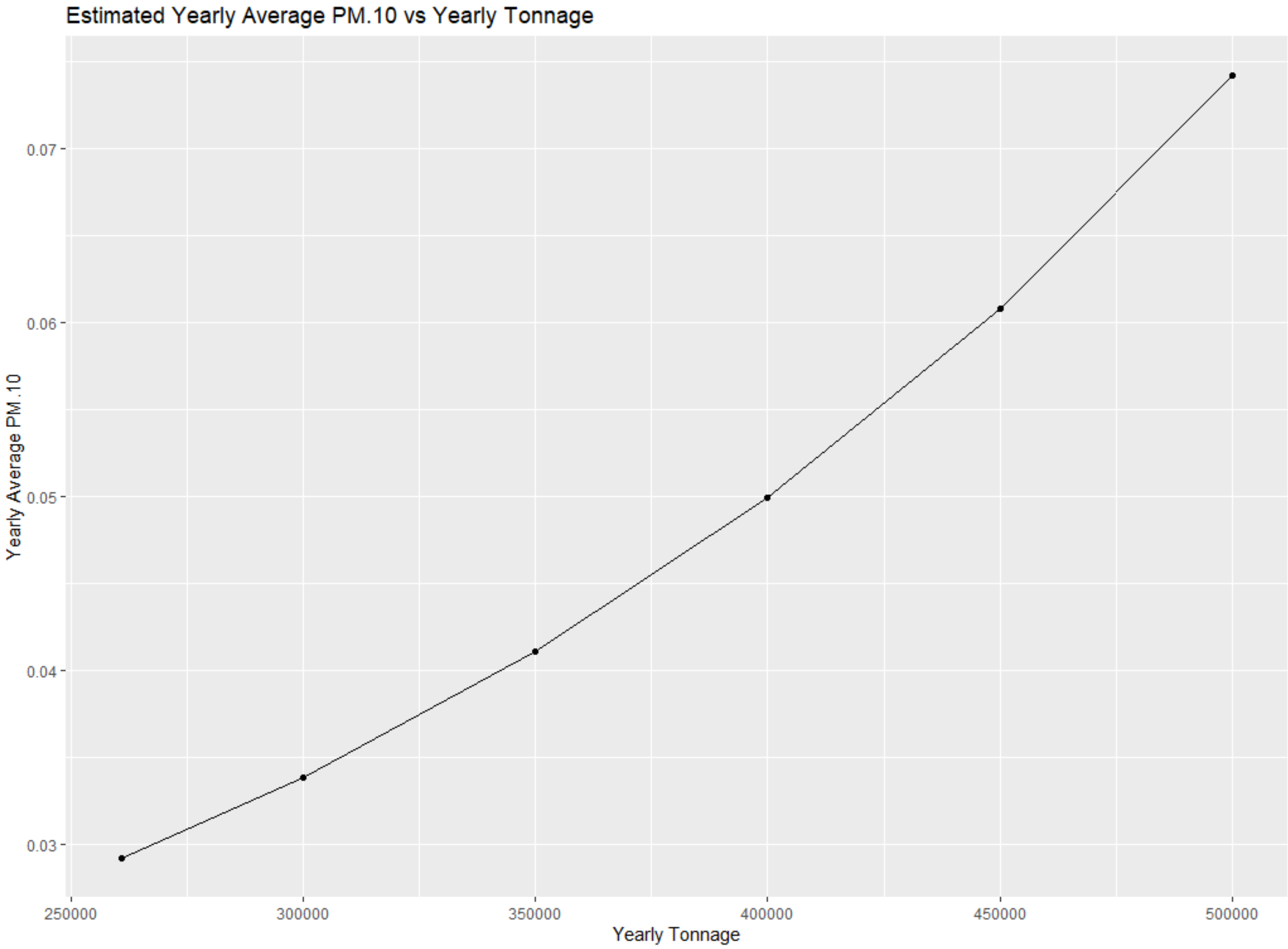


Figure 6: Yearly Tonnage Total vs Estimated Yearly Average PM_{10}

4.6 Further Analysis

4.6.1 Traffic total vs Average Humidity

From figure 7 the majority of days where the daily Average PM_{10} is exceeded are in the bottom right of the graph and are especially concentrated for Average Humidity levels less than 80% and Traffic totals above 150. For days with an Average Humidity of greater than 80% and/or a Traffic total of less than 150, only 7.5% exceeded the 0.05 mg/m^3 daily PM_{10} limit, for only days where the site is open that percentage increased to 11%. The percentage of days that exceed the 0.05 mg/m^3 daily PM_{10} limit with an Average Humidity of less than 80% and a Traffic total of greater than 150 was 44%, which is an increase of 400%.

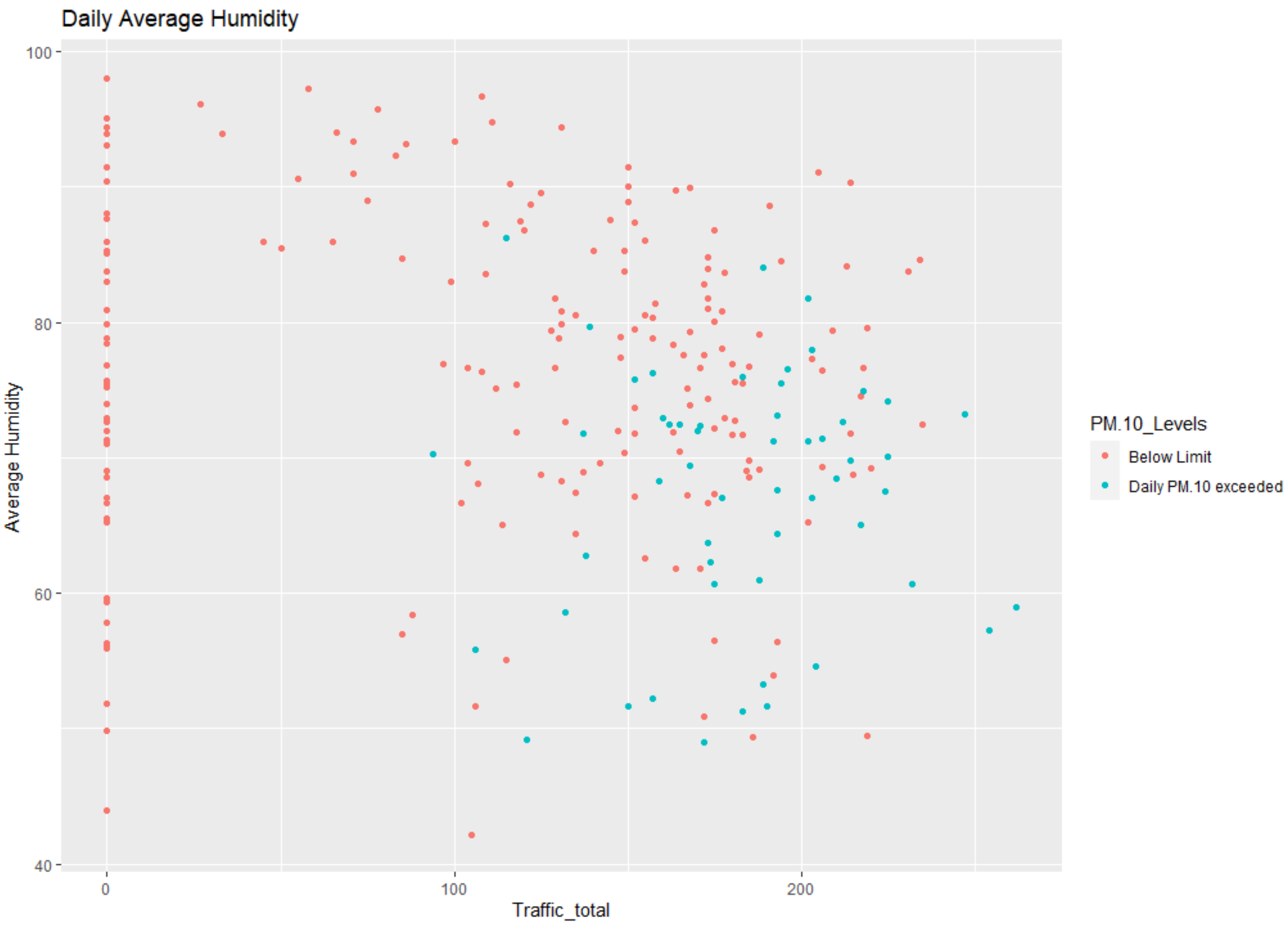


Figure 7: Traffic total vs Average Humidity

4.6.2 Effects of Spraying on Average Humidity

A visual inspection of the daily average humidity over time in the figure below (8) shows that there is no apparent link between the daily average humidity and whether the work site is open or closed, as such it is unlikely that the water spraying at Concrush effects the humidity. That is not to say that the water spraying doesn't reduce PM_{10} emissions only that the negative effect of humidity is independent of water spraying.

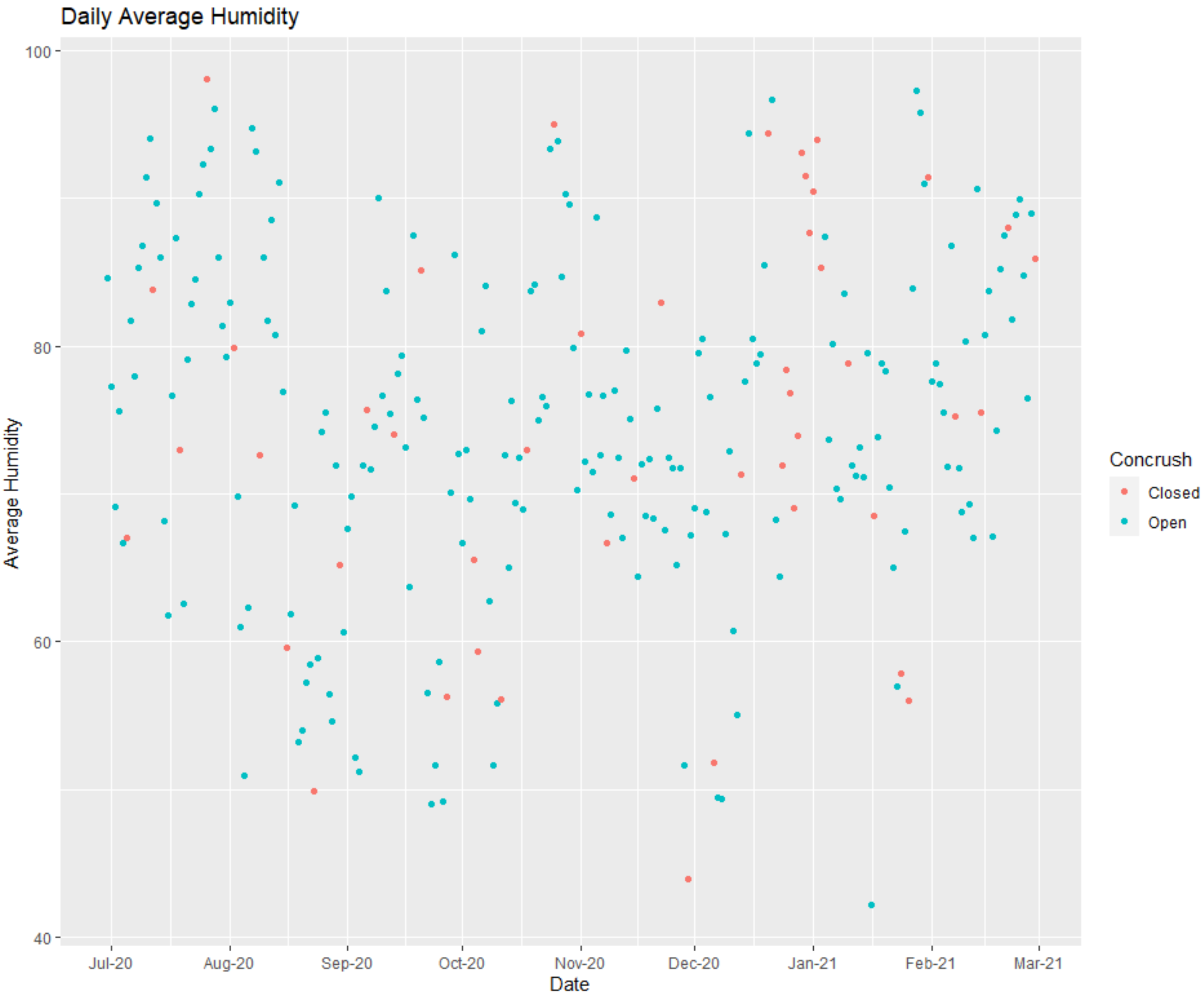


Figure 8: Average Humidity Open vs Closed days

5 Further Work

To help assess the effects of wind on PM_{10} emissions the QAMS Dust Master Pro 7000 (current location indicated by the yellow star in figure 9) could be relocated to the each of the 4 dust depositional bottles (indicated by the Blue and Yellow Pins in figure 9) for a period of a month each. Relocating the Dust Master away from the weighbridge would also help to confirm whether the high significance of the Truck Traffic total variable is effected by the Dust Masters position. Ideally the purchase of multiple Dust Masters would allow for the simultaneous collection of data at different positions in the compound.

Figure 9: Concrush Facility - Dust Master and Depositional Bottle Locations

Given the location of the 4 dust depositional bottles the wind effects analysis could be repeated and the effects of wind passing over the facility before reaching the Dust Master could give a better understanding of the relationship between wind and PM_{10} emissions and might help explain the unexpected findings in the analysis.

While the logarithmic transformation of Average PM_{10} resulted in a significant improvement in the R^2 of the linear models, log(Average PM_{10}) was still left skewed rather than normally distributed. More complex transformations of the dependent variable Average PM_{10} might result in a more accurate model. This might also help the amount at which the model underestimates some of the higher Average PM_{10} observations which will then give more weight to any predicted Yearly Average PM_{10} levels that are estimated for an increased Yearly Tonnage.

6 Conclusion

Multiple linear regression models were successfully fitted to estimate $\log(PM_{10})$ all having an $R^2 > 0.7$. On site operations at Concrush were found to have a significant effect on PM_{10} with Traffic total, Tonnage total and Crushing total all having significant positive effect estimates when estimating $\log(PM_{10})$, whether the effects of Tonnage total and Crushing total are directly associated with $\log(PM_{10})$ or are only significant due to their correlation with Traffic total is unclear. The effects of wind speed and wind direction were found to be either insignificant or associated with a decrease in $\log(PM_{10})$, the position of the Dust Master may have an effect on these results and an analysis of data collected from a Dust Master in different positions may help explain the unexpected results or find different results completely. Linear models of $\log(PM_{10})$ found that the 4 most significant factors for estimating $\log(PM_{10})$ in order of significance were Traffic total, Average Humidity, Average Ambient Pressure and Average Temperature, with Traffic total and Average Humidity being far more important than the other two. Days that exceeded a Traffic Total of 150 trucks and had an Average Humidity of less than 80% were found to be 400% more likely to exceed the daily limit of 0.05 mg/m^3 PM_{10}. To effectively reduce PM_{10} emissions low humidity days should be monitored with a possible limit of traffic set for days with a less than 80% humidity. Continued spraying and the eventual sealing of roads that are used by the trucks should be prioritised and further analysis should be done before the installation of wind barriers.

References

[1] Jim Frost. *Multicollinearity in Regression Analysis: Problems, Detection, and Solutions*. URL: https://statisticsbyjim.com/regression/multicollinearity-in-regression-analysis/#:~:text=Multicollinearity%20causes%20the%20following%20two,variables%20are%20in%20the%20model.&text=Multicollinearity%20reduces%20the%20precision%20of,power%20of%20your%20regression%20model..

[2] "Pearson's Correlation Coefficient". In: *Encyclopedia of Public Health*. Ed. by Wilhelm Kirch. Dordrecht: Springer Netherlands, 2008, pp. 1090–1091. ISBN: 978-1-4020-5614-7. DOI: 10.1007/978-1-4020-5614-7_2569. URL: https://doi.org/10.1007/978-1-4020-5614-7_2569.

[3] *National Environment Protection (Ambient Air Quality) Measure 2021 (Cth) sch II*. URL: https://www.legislation.gov.au/Details/F2021C00475.

[4] Thompson Environmental Systems. *QAMS DUST MASTER PRO 7000 BY TES*. URL: https://www.thomsongroup.com.au/product/tes-dust-master-pro-7000/.

Appendix

A Datasets

Table 6: Dust Master Data

Variable	Description	Units/Format
Date	Date	DD/MM/YY
Time	Time	HH/MM/SS
SV	Supply Voltage	V
PM2.5	PM2.5 concentration	mg/m^3
PM.10	PM10 concentration	mg/m^3
PM.Tot	PM.Total concentration	mg/m^3
LPM	Flowrate	L/min
PUMP	Pump demand	%
S.T	Temperature (onboard sensor)	°C
S.RH	Relative humidity (onboard sensor)	%
S.DP	Dew point (onboard sensor)	°C
S.AP	Ambient Pressure (onboard sensor)	mBar
M1.WS	Wind speed (weather sensor 1)	m/s
M1.WD	Wind direction (weather sensor 1)	°
M1.WD.SD	Wind direction standard deviation (weather sensor 1)	°
M1.WB	Wet bulb temperature (weather sensor 1)	°C
M1.T1	Temperature 1 (weather sensor 1)	°C
M1.T2	Temperature 2 (weather sensor 1)	°C
M1..RH	Relative humidity (weather sensor 1)	%
M1.AP	Ambient Pressure (weather sensor 1)	mBar

Table 7: Truck Data

Variable	Description	Units/Format
Date In	Date	DD/MM/YY
Time In	Time of entry	HH/MM
Time Out	Time of Exit	HH/MM
Vehicle Registration	Vehicle number plate	
Direction	Truck Direction	IN/OUT
Gross	Total Truck Weight	tonnes
Tare	Truck weight	tonnes
Net	Truck load weight	tonnes
Product Description	Truck load description	

Table 8: Crushing Data

Variable	Description	Units/Format
Server Timestamp	Date and Time	DD/MM/YY HH/MM
Smartphone Timestamp	Date and Time	DD/MM/YY HH/MM
Record Timestamp	Date and Time	DD/MM/YY HH/MM
Amount	Volume of Material	m^3
Type	Type of recording	
Indication	Type of recording	
Factor		%
Diameter		mm
Density	Material Density	tonnes/m^3
Idle Time	Machine Idle Time	minutes
Material	Material Size	20mm/40MM
Customer	Machine Owner	Concrush
Sensor	Sensor Number	Sensor 878
Measurement Duration	Measurement Duration	minutes

B Wind Direction Angles

Table 9: 4 Wind Directions

Wind	Angle 1	Angle 2
North	315	45
East	45	135
South	135	225
West	225	315

Table 10: 8 Wind Directions

Wind	Angle 1	Angle 2
North	337.5	22.5
North East	22.5	67.5
East	67.5	112.5
South East	112.5	157.5
South	157.5	202.5
South West	202.5	247.5
West	247.5	292.5
North West	292.5	337.5

Table 11: 16 Wind Directions

Wind	Angle 1	Angle 2
North	348.75	11.25
North North East	11.25	33.75
North East	33.75	56.25
East North East	56.25	78.75
East	78.75	101.25
East South East	101.25	123.75
South East	123.75	146.25
South South East	146.25	168.75
South	168.75	191.25
South South West	191.25	213.75
South West	213.75	236.25
West South West	236.25	258.75
West	258.75	281.25
West North West	281.25	303.75
North West	303.75	326.25
North North West	326.25	348.75

C Wind Speed Models

Model with Average Wind Speed

$$\begin{aligned}\log(\text{Average PM}_{10}) = \beta_0 &+ \beta_1 \times \text{Traffic Total} + \beta_2 \times \text{Tonnage Total} + \beta_3 \times \text{Average Temperature} \\ &+ \beta_4 \times \text{Average Relative Humidity} + \beta_5 \times \text{Average Ambient Pressure} \\ &+ \beta_6 \times \text{Crushing Amount Total} + \beta_7 \times \text{Average Wind Speed} + \epsilon\end{aligned} \tag{13}$$

Model with 4 Wind Directions

$$\begin{aligned}\log(\text{Average PM}_{10}) = \beta_0 &+ \beta_1 \times \text{Traffic Total} + \beta_2 \times \text{Tonnage Total} + \beta_3 \times \text{Average Temperature} \\ &+ \beta_4 \times \text{Average Relative Humidity} + \beta_5 \times \text{Average Ambient Pressure} \\ &+ \beta_6 \times \text{Crushing Amount Total} + \beta_7 \times \text{North Wind} + \beta_8 \times \text{East Wind} \\ &+ \beta_9 \times \text{South Wind} + \beta_{10} \times \text{West Wind} + \epsilon\end{aligned} \tag{14}$$

Model with 8 Wind Directions

$$\begin{aligned}\log(\text{Average PM}_{10}) = \beta_0 &+ \beta_1 \times \text{Traffic Total} + \beta_2 \times \text{Tonnage Total} + \beta_3 \times \text{Average Temperature} \\ &+ \beta_4 \times \text{Average Relative Humidity} + \beta_5 \times \text{Average Ambient Pressure} \\ &+ \beta_6 \times \text{Crushing Amount Total} + \beta_7 \times \text{North Wind} + \beta_8 \times \text{North East Wind} \\ &+ \beta_9 \times \text{East Wind} + \beta_{10} \times \text{South East Wind} + \beta_{11} \times \text{South Wind} \\ &+ \beta_{12} \times \text{South West Wind} + \beta_{13} \times \text{West Wind} + \beta_{14} \times \text{North West Wind} + \epsilon\end{aligned} \tag{15}$$

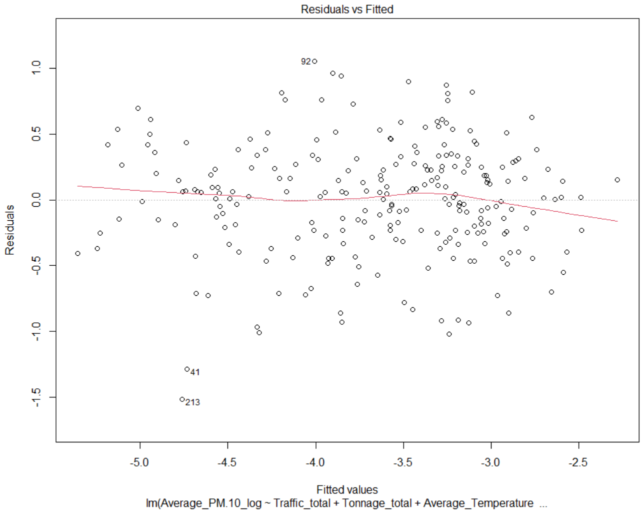
Model with 16 Wind Directions

$$
\begin{aligned}
\log(\text{Average PM}_{10}) = {} & \beta_0 + \beta_1 \times \text{Traffic Total} + \beta_2 \times \text{Tonnage Total} + \beta_3 \times \text{Average Temperature} \\
& + \beta_4 \times \text{Average Relative Humidity} + \beta_5 \times \text{Average Ambient Pressure} \\
& + \beta_6 \times \text{Crushing Amount Total} + \beta_7 \times \text{North Wind} + \beta_8 \times \text{North North East Wind} \\
& + \beta_9 \times \text{North East Wind} + \beta_{10} \times \text{East North East Wind} + \beta_{11} \times \text{East Wind} \\
& + \beta_{12} \times \text{East South East Wind} + \beta_{13} \times \text{South East Wind} + \beta_{14} \times \text{South South East Wind} \\
& + \beta_{15} \times \text{South Wind} + \beta_{16} \times \text{South South West Wind} + \beta_{17} \times \text{South West Wind} \\
& + \beta_{18} \times \text{West South West Wind} + \beta_{19} \times \text{West Wind} + \beta_{20} \times \text{West North West Wind} \\
& + \beta_{21} \times \text{North West Wind} + \beta_{22} \times \text{North North West Wind} + \epsilon
\end{aligned}
\tag{16}
$$

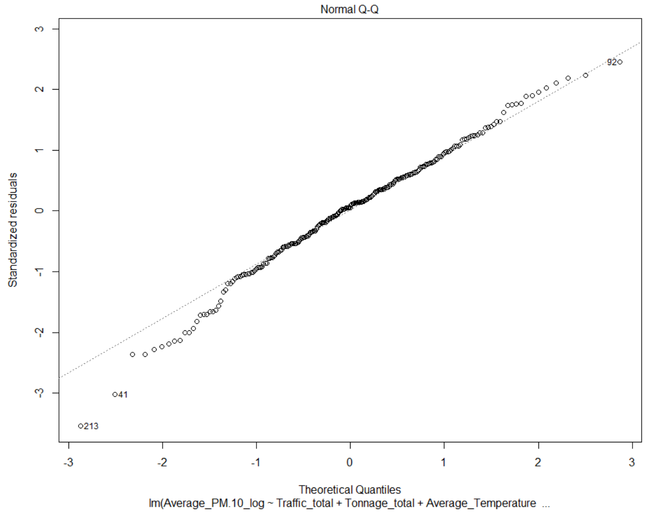
D 24 Hour Data Variable Summary

Table 12: 24 Hour Data Variable Summary

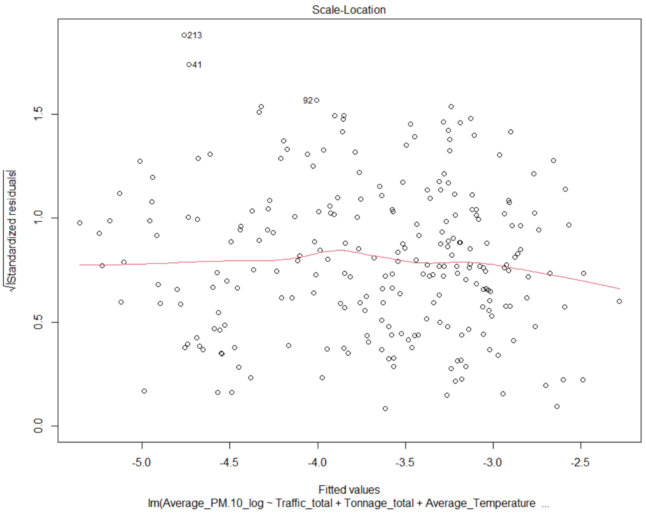
Variable	Units	Mean	Minimum	Maximum
Average PM.10	mg/m^3	0.033753	0.001878	0.119390
Traffic Total	count	128.20	0	262
Tonnage Total	tonnes	716.4	0	3142.9
Average Temperature	°C	15.130	5.509	22.716
Average Humidity	%	74.65	42.18	98.07
Average Ambient Pressure	mBar	1006.0	981.1	1025.3
Crushing Amount Total	m^3	187.1	0	1035.9
Average Wind Speed	m/s	1.3769	0.4065	3.7838



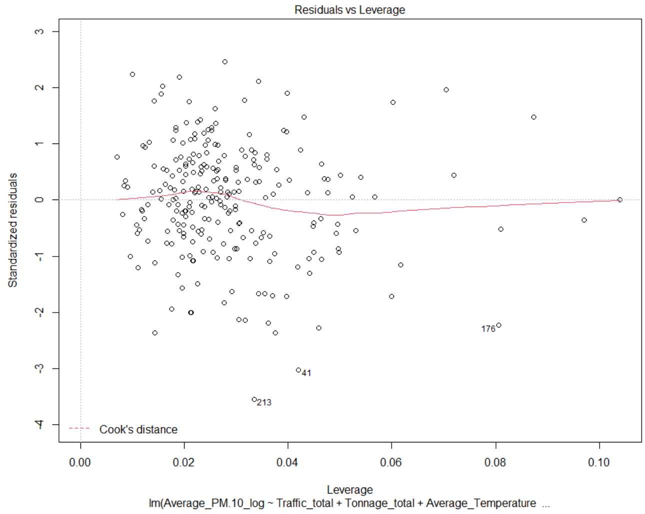
(a) Residuals vs Fitted



(b) Normal Q-Q



(c) Scale-Location



(d) Residuals vs Leverage

Figure 10: Model Plots

E Plots of the Final Model

F Histogram of Actual vs Shifted Daily Traffic Total

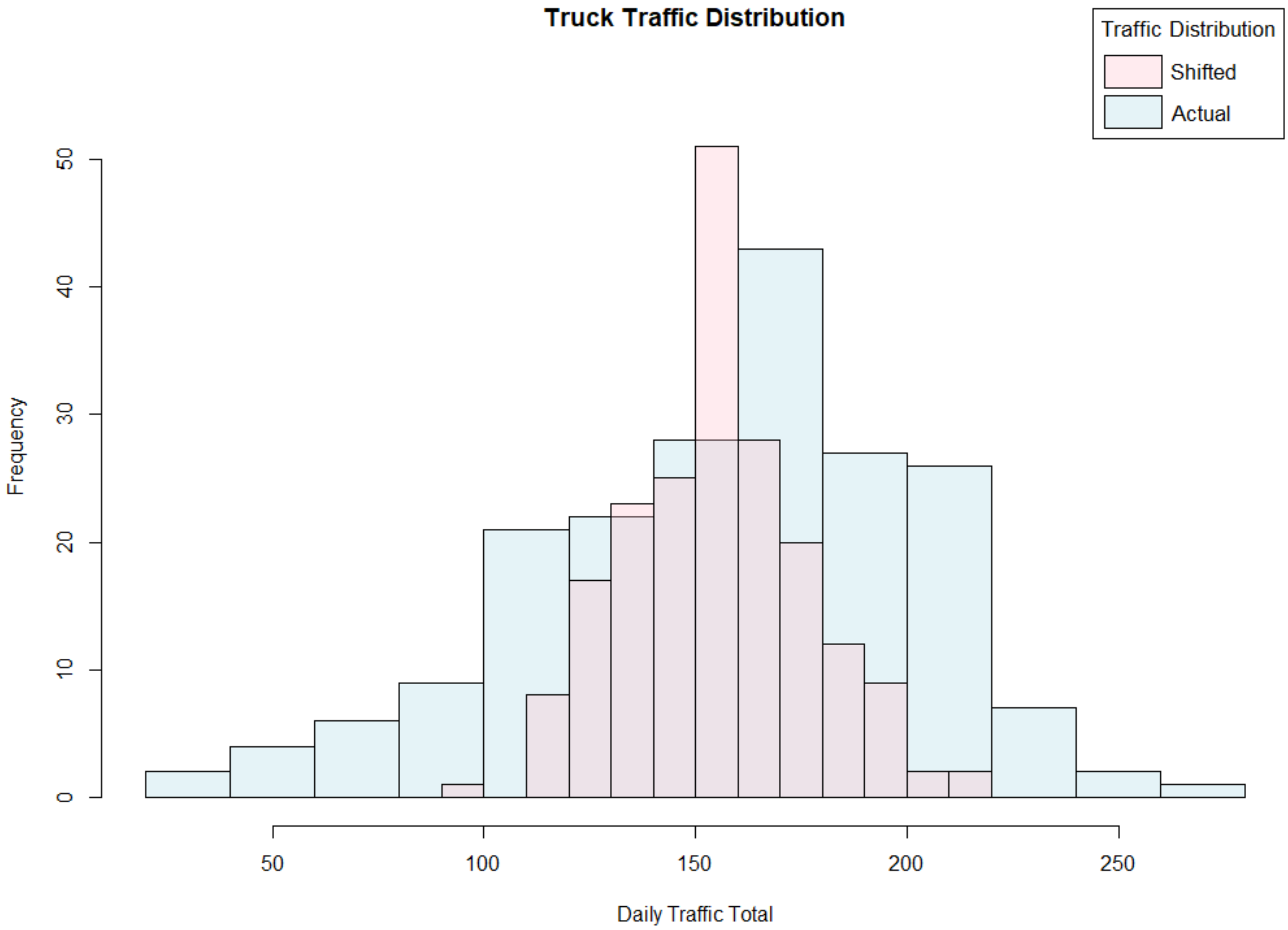


Figure 11: Histogram of Actual vs Shifted Daily Traffic Total

G Actual vs Estimate Daily Average PM.10

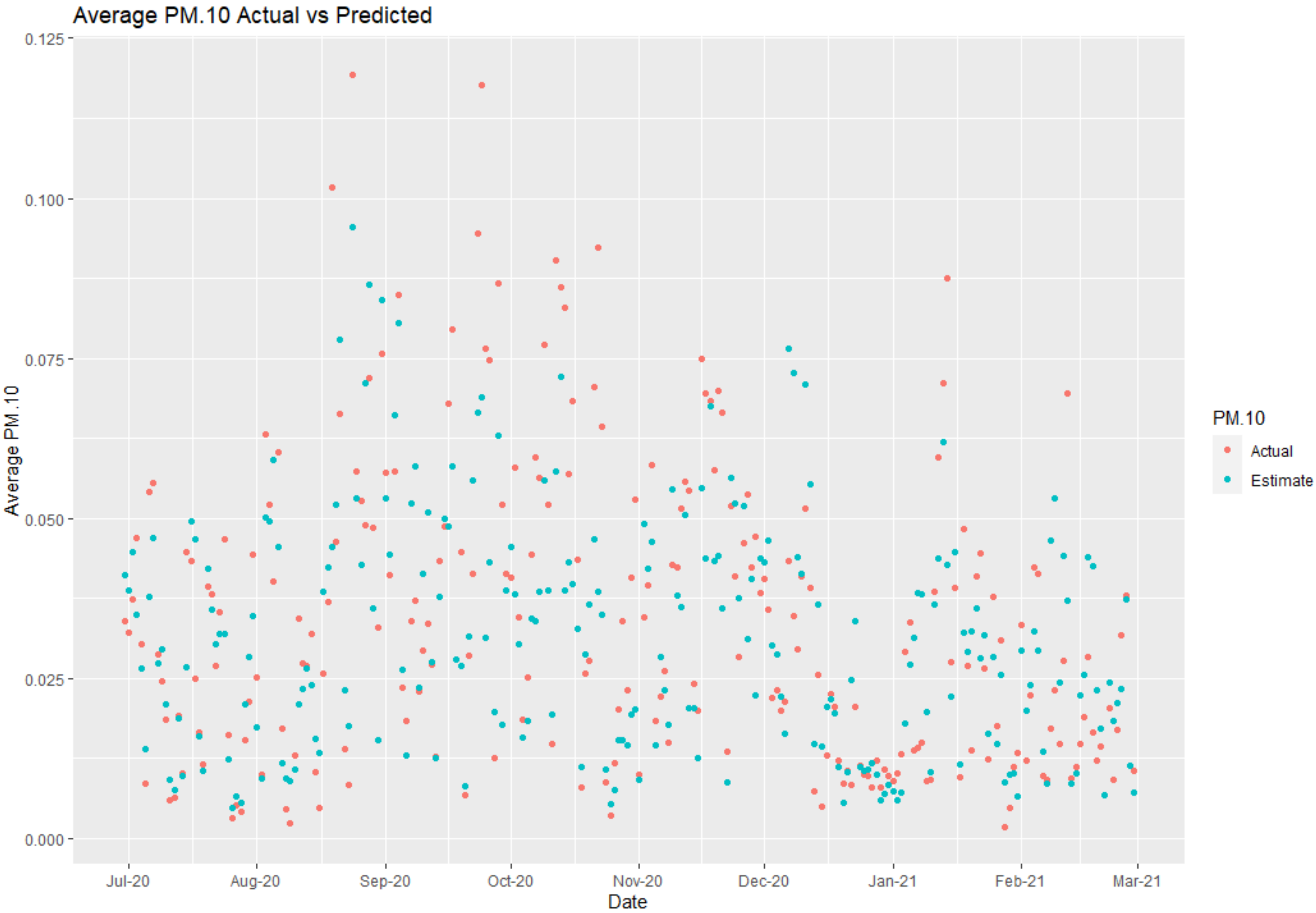


Figure 12: Actual vs Predicted Average PM_{10}

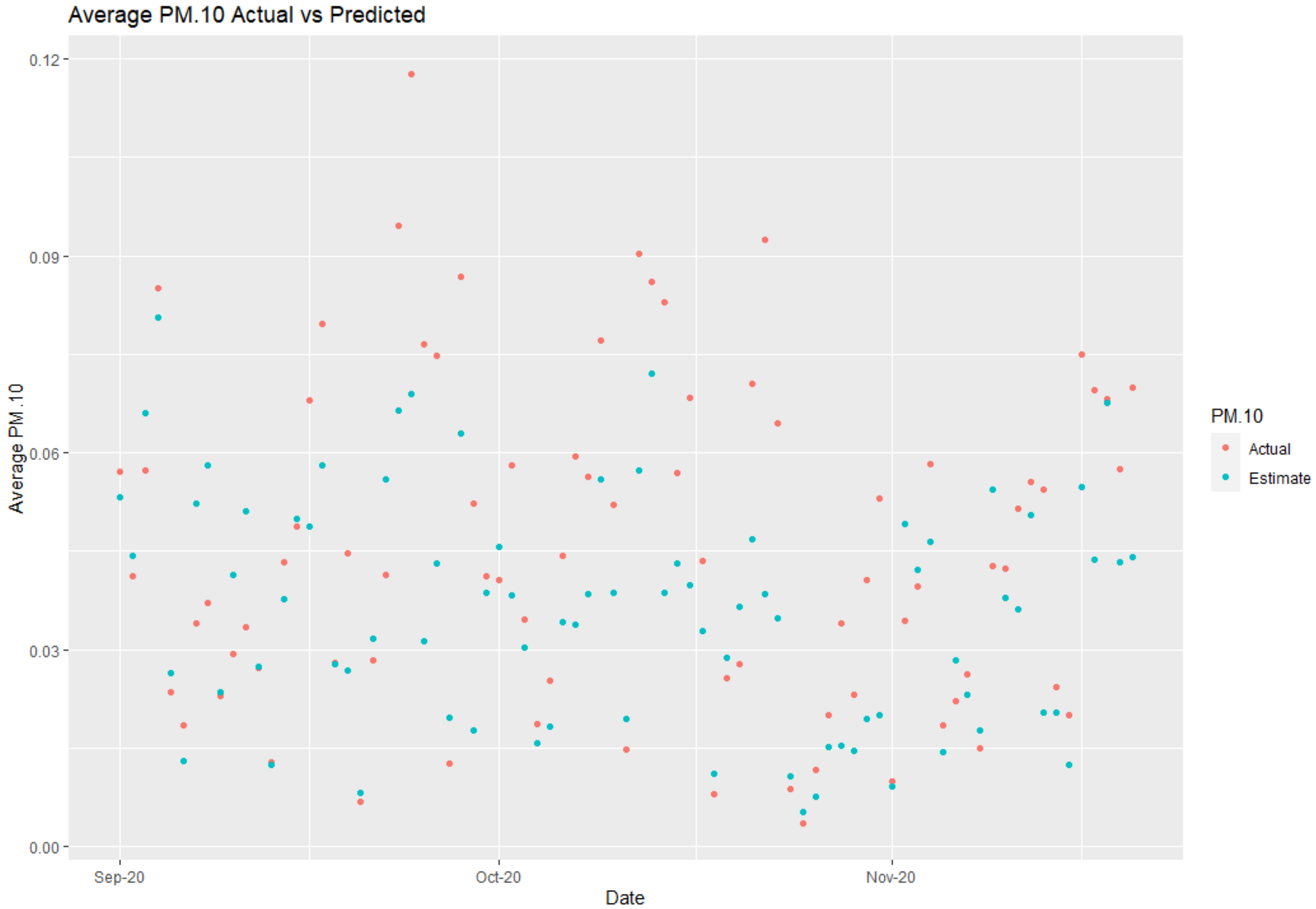


Figure 13: Actual vs Predicted Average PM_{10} for 01/09/2020 to 20/11/2020