Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
An investigation of the impacts of a successful COVID-19 response and meteorology on air quality in New Zealand

Nick Talbot\textsuperscript{a,c,}\textsuperscript{*}, Akika Takada\textsuperscript{b}, Andrew H. Bingham\textsuperscript{b}, Dan Elder\textsuperscript{b}, Samantha Lay Yee\textsuperscript{a}, Nancy E. Golubiewski\textsuperscript{a}

\textsuperscript{a} Ministry for the Environment, Auckland, New Zealand
\textsuperscript{b} Ministry for the Environment, Wellington, New Zealand
\textsuperscript{c} University of Auckland, School of Environment, Auckland, New Zealand

\textbf{HIGHLIGHTS}

- COVID-19 response produced notable impacts on air quality.
- Decrease in traffic emissions reduced NO\textsubscript{2} notably across New Zealand.
- As the restrictive levels eased, pollution levels returned to near long-term means.
- Particulate concentrations increased during alert level 2, attributed to more home heating emissions from wood burning.
- Machine learning found good R values for NO\textsubscript{2}; however, modelling results were weaker for coastal particulate data.

\textbf{ABSTRACT}

The COVID-19 pandemic brought about national restrictions on people’s movements, in effect commencing a socially engineered transport emission reduction experiment. In New Zealand during the most restrictive alert level (Level 4), roadside concentrations of nitrogen dioxide (NO\textsubscript{2}) were reduced 48–54\% compared to Business-as-usual (BAU) values. NO\textsubscript{2} concentrations rapidly returned to near mean levels as the alert levels decreased and restrictions eased.

PM\textsubscript{10} and PM\textsubscript{2.5} responded differently to NO\textsubscript{2} during the different alert levels. This is due to particulates having multiple sources, many of natural origin and therefore less influenced by human activity. PM\textsubscript{10} and PM\textsubscript{2.5} concentrations were reduced during alert level 4 but to a lesser extent than NO\textsubscript{2} and with more variability across regions. Particulate concentrations increased notably during alert level 2 when many airsheds reported concentrations above the BAU means.

To provide robust BAU reference concentrations, simple 5-year means were calculated along with predictions from machine learning modelling that, in effect, removed the influence of meteorology on observed concentrations. The results of this study show that latter method was found to be more closely aligned to observed values for NO\textsubscript{2} as well as PM\textsubscript{2.5} and PM\textsubscript{10} away from coastal regions.

\textsuperscript{*} Corresponding author. Ministry for the Environment, Auckland, New Zealand.
E-mail address: nick.talbot@mfe.govt.nz (N. Talbot).

https://doi.org/10.1016/j.atmosenv.2021.118322

Received 9 January 2021; Received in revised form 25 February 2021; Accepted 2 March 2021
Available online 11 March 2021

1352-2310/© 2021 Elsevier Ltd. All rights reserved.
1. Introduction

At the end of 2019, a new disease emerged from Wuhan, China initially named SARS-COV2 (now COVID-19). Highly contagious between humans, the virus rapidly spread around the globe with over 200 countries ultimately affected. With no known cure or effective treatment at the initial stages of the outbreak, strict social restrictions were activated in the form of alert levels to help protect vulnerable citizens and prevent health systems becoming inundated (Wilders-Smith and Freedman, 2020; World Health Organisation, 2020).

Globally, research has shown that social intervention restrictions led to notable changes in air pollution levels. For example, significant reductions in nitrogen dioxide (NO\textsubscript{2}) concentrations were reported in Brazil (Nakada and Urban, 2020), Spain (Tobias et al., 2020), Ecuador (Zambrano-Monserrate and Ruano, 2020), the USA (Berman and Ebisu, 2020), and China (Griffith et al., 2020). Concentrations of fine particles (PM\textsubscript{2.5}) also decreased during intervention strategies in China (Wang and Su, 2020), India (Sharma et al., 2020), Kazakhstan (Kerimray et al., 2020) and the USA (Hudda et al., 2020).

Conversely, tropospheric ozone (O\textsubscript{3}) concentrations increased in Europe, China, and South America (Sicard et al., 2020; Siciliano et al., 2020). Particulate results are more variable according to the geographic position (e.g., next to major sources such as oceans or deserts) or the proportion of secondary particulates (Tobias et al., 2020). The influence of natural sources along with human behaviour on particulate concentrations would therefore result in locations experiencing less reduction due to change in activities.

New Zealand is uniquely placed to investigate the impact of behaviour change on air quality given it is so geographically isolated and inherits little transported pollutants from other countries. New Zealand also held one of the strictest and best observed lockdowns of any country, worldwide (Robert, 2020). A four-level response system (https://covid19.govt.nz/alert-system/) was introduced on March 21, 2020, at which point New Zealand entered Level 2. On 23 March, New Zealand moved to Level 3, with limits put on movement, advice to work from home, and schools closed in preparation for a 4-week national strict Level 4 ‘lockdown’ commencing 26 March.

Most New Zealanders stayed at home with only essential services operating through 28 April. Activity slowly resumed as the country eased restrictions: Level 3 (28 April – 13 May) followed by Level 2 (14 May – 8 June) (Table 1). A by-product of New Zealand’s COVID-19 alert levels (Table 1) was nationally regulated reductions of major air pollution emission sources. Both the intensity of the lockdown and the structured easing of restrictions through the alert levels allow for in-depth analysis of observed air quality and the relative importance of key pollution sources. The relationship between changes in the emission sources and the resultant concentrations offers rare insight into each pollutant’s relative importance in the ambient atmosphere.

In a recent study, Patel et al. (2020) reported NO\textsubscript{2} reductions due to COVID-19 related reductions in traffic volume across Auckland, New Zealand’s largest city. Whereas emission inventories and receptor modelling approaches confirm the dominance of traffic sources for nitrogen oxides (NO\textsubscript{x}) (86%) and black carbon (BC) (72%) across the city, observations showed consequent reductions in NO\textsubscript{2} of only 34–57% and in BC, 55–75%. PM\textsubscript{2.5} (also likely to be dominated by traffic emissions) and particulate matter (PM\textsubscript{10}) (seasonal, dominated by sea salt and, to a lesser extent, traffic emissions) were reduced during Level 4 (8–17% for PM\textsubscript{2.5} and 7–20% for PM\textsubscript{10}) (Davy et al., 2017; Patel et al., 2020; Xie et al., 2019).

Patel et al. (2020) investigated three Auckland sites representative of different airsheds—urban peak, urban background, and regional background only during the Level 4 lockdown, not the subsequent levels under which restrictions eased (Table 1). Further investigation is required to understand the trajectory of air quality post lockdown and whether the results in Auckland are representative of a national picture.

The level of reduction in air pollutant concentrations during COVID-

| COVID-19 | Alert System (dates implemented) | Risk Assessment | Key restrictions during level | Traffic volume change (Light vehicles) | Traffic volume change (Heavy Goods Vehicles) |
|---------|----------------------------------|----------------|-------------------------------|----------------------------------------|---------------------------------------------|
| Level 4 (26/3–27/4 2020) | Sustained and intensive community transmission is occurring. | People instructed to stay at home in their bubble. | Reduced ~79% | Reduced ~72% |
| Level 3 (27/4–13/5 2020) | Multiple cases of community transmission occurring. | People instructed to stay in their bubble other than for essential personal movement — including to go to work, school if they must, or for local recreation. | Reduced ~46% | Reduced ~19% |
| Level 2 (14/5–8/6 2020) | Limited community transmission could be occurring. | People can connect with friends and family, and socialise in groups of up to 100, go shopping, or travel domestically, if following public health guidance. | Reduced ~20% | Reduced ~7% |

(continued on next page)
19 restrictions is dependent on a reasonable approximation of a business-as-usual (BAU) value for concentrations during that period. To achieve this, this paper incorporates two methods. Firstly, a simple long-term mean (LTM) for each period was calculated. Secondly, machine learning (ML) was used to account for the impacts of meteorology during each of the alert periods, given that meteorology is a key factor affecting air pollution concentrations (Ebenezer, 2019; Lolli et al., 2020).

The application of machine learning as a prediction tool in the area of atmospheric science is a growing and innovative application of the technology (Feng et al., 2019; Alimissis et al., 2018; Lautenschlager et al., 2020). ML can offer a more robust assessment than the simple LTM to normalise the impacts of meteorological variables (Grange et al., 2018). Meteorological normalisation is a technique that accounts for changes in meteorology in an air quality time series. Controlling for such changes helps support trend analysis because there is more certainty that the observed trends are due to changes in emissions or secondary processes rather than changes in meteorology. This method was successfully used to analyse COVID-19 related national NO2 changes across Spain (Petetin et al., 2020). However, the study did not compare using a 5-year mean against ML algorithm results nor did it consider how meteorological conditions influenced particulate pollution levels. To the authors’ knowledge, ML has not been used for particulate measurements and therefore the response from biogenic sources such as sea salt is unknown. It is due to this uncertainty that LTM is used in conjunction with ML in this paper.

This study addresses these knowledge gaps via a national assessment of air quality, via NO2, PM10, and PM2.5 pollutants, during New Zealand’s first round of COVID-19 restrictions (during alert Levels 4, 3 and 2). Observed changes to air quality under COVID-19 restrictions were compared to two reference values, LTM and ML predictions. The observed changes in pollution concentrations are discussed as well as the model performance considering both location and pollutant.

2. Methodology

2.1. Data collection

We collected data for daily PM10, PM2.5, and NO2 measurements from 36 stations across New Zealand and classified them into five geographical clusters (Table 2). The sites were chosen based on data availability and population, where restrictions were likely to have been more evident and have had greater impacts on air quality. The period considered was from January 1, 2015 to June 8, 2020, although the amount of historical data (2015–2019) varies from station to station. The air quality sites chosen are run according to national standards (Resource Management (National Environmental Standards for Air Quality) Regulations 2004) by local government using instruments that meet regulatory standards for air quality data collection (Ministry for the Environment, 2009).

Where available, meteorological data were retrieved from the same stations recording air pollutant concentrations, including air temperature, relative humidity, wind speed and wind direction. For air quality stations that did not record meteorological parameters, the closest available station was used. All data were aggregated into daily values to align with the air pollutant concentrations.

Both air pollutant concentration and meteorological data were mainly extracted through the Ministry for the Environment’s data pipeline, which directly accesses councils’ databases. Some data were received directly from councils or downloaded through their platform when it could not be accessed directly through the pipeline. Some meteorological data were obtained from the National Institute of Water and Atmospheric Science (NIWA) CLIFLO database (https://cliflo.niwa.co.nz/) when those data were not collected by councils or data were not available due to instrument failures. More details on the location, available air pollutant concentrations, and meteorological data from monitoring stations can be found in Appendix 1.

Daily traffic count for light vehicles and heavy goods vehicles (HGV) were collected by New Zealand Transport Agency (NZTA) at key traffic sites in the main urban centres of Auckland (region NU), Wellington (region NL), and Christchurch (region SL) from December 3, 2018 until June 08, 2020. This information was used as proxy of the amount of reduction on vehicle movements that occurred during each alert level.

To test the relationship between traffic volume and air quality changes, we investigated separate. In this study, we examine data from the full periods for Level 4 (26 March - 27 April 2020), Level 3 (28 April - 13 May), and Level 2 (14 May - 8 June) (Table 1). We compared two methods to estimate the concentrations of each air pollutant that would have been observed without restriction measures: 1) the long-term mean of historical data (2015–2019), and 2) predicted values, incorporating meteorological variability, using the random forest machine learning algorithm.

| Group Name | Regional Cluster | Number of Stations |
|------------|------------------|--------------------|
| North Island | Northland, Auckland, Waikato | 6 3 3 |
| North Central | Hawkes Bay, Gisborne, Taranaki, Bay of Plenty | 5 3 0 |
| North Lower | Horizons, Wellington | 5 3 3 |
| South Island | Nelson, Tasman, Marlborough, West Coast | 5 1 0 |
| South Lower | Canterbury, Otago, Southland | 6 3 2 |

### Table 1 (continued)

| COVID-19 Alert System | Risk Assessment (dates implemented) | Key restrictions during level | Traffic volume change (Light vehicles) | Traffic volume change (Heavy Goods Vehicles) |
|-----------------------|-----------------------------------|-------------------------------|----------------------------------------|---------------------------------------------|
| Businesses can open to the public if following public health guidance including physical distancing and record keeping. | | | | |

| | PM10 | PM2.5 | NO2 |
|---|---|---|---|
| North Island | 6 3 3 |
| North Central | 5 3 0 |
| North Lower | 5 3 3 |
| South Island | 5 1 0 |
| South Lower | 6 3 2 |
We grouped data into historical (LTM) (2015–2019) and observed (OB) (2020) datasets. Historical data were further grouped into average of Julian day (day of a year) to calculate the availability across historical period of each alert level. For example, air quality data on 26th of March in 2015–2019, inclusive, from each site was averaged as a historical 26th of March value and counted as one data point in the historical Level 4 period. This approach was also applied to meteorology data. Given that each alert level lasted for a short period of time (16–33 days), we only considered sites with at least 50% of air quality and meteorological data available during each alert level period and its associated historical period.

Analyses were done using R 3.6.1 (R Core Team, 2020) and the caret (v6.0.86), randomForest (v4.6.14; Liaw and Wiener, 2001), openair (v2.7.4; Carslaw and Ropkins, 2012) and ranger (v0.12.1) packages.

2.2.1. Traffic changes under COVID-19 restrictions

It is understood that mobility was most affected by COVID-19 restrictions, and, as such, transport volume, and therefore emissions, altered. To help understand these changes, traffic data and fleet type data were collected by New Zealand Transport Agency during the COVID-19 alert level periods (Waka Kotahi NZ Transport Agency, 2020). Major road junctions in three New Zealand cities were considered. The data from these sites are extrapolated as proxies for changes in on-road vehicle volume across New Zealand. The similarity of volume changes among the three cities in accordance with alert levels supports this assumption.

2.2.2. Meteorology under COVID-19 restrictions

To understand how ‘typical’ meteorological conditions were during the COVID-19 alert levels for each region, the percentage difference between the two periods has been calculated. The results provided information on the potential changes in air pollutant concentrations under BAU.

2.2.3. Air quality under COVID-19 restrictions

We applied two methods to estimate a reference concentration value for each air pollutant that could have been expected in the absence of COVID-19 restrictions. The first was the average of historical data (2015–2019), named long-term mean (LTM). The second was derived using a random forest machine learning algorithm (ML). We then examined the differences between observed values and BAU estimates using the LTM and ML.

2.2.3.1. Comparing air quality observations under COVID-19 restrictions to long term means. Air quality observations during each alert level were assessed for differences compared to their respective LTM based on 5 years of daily averages (2015–2019), where available. This allows for reasonable comparisons between air quality observations during each COVID-19 alert level and what could be expected under BAU, had pandemic restrictions not been imposed.

2.2.3.2. Comparing air quality observations under COVID-19 restrictions to predicted values taking account of meteorological conditions

2.2.3.2.1. Machine learning using random forest algorithm. Random forest (RF) is a machine learning algorithm used for classification and regression to unfold patterns and relationships and is considered to be one of the most powerful tools in many fields (Chen and Ishwaran, 2012; Herrera et al., 2019). The algorithm builds many regression trees where the target variable takes continuous values to exercise binary recursive partitioning. It is an iterative process that splits the data into partitions the predictor is imputed using these values (using the mean). This approach also centred and scaled the data to ensure that all explanatory variables give equal contribution to the analysis.

For each site, we developed a model for each air pollutant. We used 80% of data as a training set for identifying the best model and 20% of data as a test set for testing the final model. Three hyper parameters were tuned to find the best learners: the number of variables (called mtry), node size (1, 5 and 10) and the number of trees (300, 500 and 800). Ten-fold cross-validation was used to evaluate the RF models, which randomly splits the training data into ten sets of approximately equal size. The model was run ten times for validation, during which each unique group was used as a test set and the remaining as training set. This means each sample has an equal chance of being used as training and test samples, minimising the bias that may occur in a simple, one-off training and test set split. Root Mean Square Error (RMSE), R squared, and Mean Absolute Error (MAE) were calculated in each run to evaluate the model performance. The results from the ten runs were then averaged as the results for the model. After retrieving the results of model performance, the RF algorithm used the whole dataset as a training set to build the final model. Finally, we used the test set to evaluate the model and evaluate the expected accuracy of the models with Normalised Root Mean Square Error (NRMSE), R squared, and MAE.

The potential deviation from the predicted value was quantified using the mean residuals (predicted minus observed concentration of air pollutants) for each site. We computed both average and 7-day running average of the daily residuals, and the associated 5th and 95th percentiles were derived as the uncertainty interval.

2.2.3.2.2. Evaluation of machine learning models. We used MAE, NRMSE and R squared to examine whether the model predictions are valid reference values. We compared the estimated values from the models for the historical period (2015–2019) to its historical observed value (testing dataset) to calculate these metrics.

MAE and RMSE both measure the accuracy of the predictions versus observed values. The main differences between two metrics are: 1) RMSE penalises large errors and increases with the variance of the frequency distribution of error magnitudes and 2) RMSE tends to be much larger as test sample size increases (Chai & Draxler, 2014; Willmott & Matsuura, 2005). Here, we applied NRMSE to remove the scale-dependent nature from RMSE which allows it to compare between accuracies of different datasets. In general, the lower the values, the higher the accuracy. However, zero error is not what machine learning models are aiming for as the model should be flexible and be able to predict in a set of new values instead of memorising the whole training dataset. The other hand, Mean Absolute Error (MAE) measures the proportion of the variance for an air pollutant concentration that is explained by meteorological and time variables that we used. These three metrics together explain whether the models have considered relevant variables that affect air pollutant concentrations from the site and reasonably predicted air pollutant concentrations.

Average variable importance was also examined for the models for each air pollutant to justify the inclusion of the variables and to understand their impacts.
3. Results

3.1. On-road vehicle transport volume changes during COVID-19 restrictions

During the Level 4 lockdown, vehicle flow reduced by approximately 79% for light vehicles and 72% for heavy goods vehicles (HGV) (Waka Kotahi NZ Transport Agency, 2020) across the country (Fig. 2). During Level 3, HGV levels returned to 19% lower than the previous year. However, light vehicle volumes were about 46% lower than the previous year. During Level 2, traffic levels remained lower than 2019 for light vehicles (~20%) and for heavy good vehicles (~7%).

Linear regression (Fig. 3) showed strong relationships between BC concentrations at monitoring stations and traffic counts for both light and heavy vehicles at all three key traffic sites, whereas the relationships in South Lower appeared to be weaker. The South Lower site showed relatively weak relationships between vehicle number and BC concentrations. The lower correlations indicate the likelihood of confounding factors influencing the data, most plausibly local meteorology, with the South Lower collection site close to the Pacific Ocean. Slight deviations of the airflow would bring clean air from the Pacific rather than more polluted city air over the measurement site. Another possibility is that

![Fig. 1. Schematic of the random forest normalisation process employed for this research.](#)

![Fig. 2. Traffic counts for heavy goods vehicles (top row) and light vehicles (bottom row) for regions NU (Auckland), NL (Wellington), and SL (Christchurch). The red lines indicate the days when different alert levels were introduced. (Source: New Zealand Transport Agency).](#)
the South Lower on-ramp vehicle count location was less representative of traffic flow across that region during different alert levels. However, given the similarity in traffic volume changes across regions, this appears unlikely.

Overall, the results here show that extrapolation of traffic volume from one road can be used as a proxy for emitted pollution elsewhere in the same city, albeit with an understanding of meteorology and the pollutant’s characteristics.

### 3.2. Meteorology

Meteorological conditions during COVID-19 alert levels in 2020 differed from past years (2015–2019) to varying degrees, with the most meteorological anomalies occurring during Level 2 (Table 3) (Local microclimatic conditions might mean that some districts reach statistically significant differences not reflected in the mean). Temperature varied only in North Upper during Level 4 and in South Upper during Level 2. Also, during Level 4, relative humidity (RH) was significantly different for North Upper, North Central and South Upper; South Upper also had significantly different wind direction. Level 3 was mostly like the reference period: only RH in North Upper and North Central were statistically significantly different.

On the other hand, meteorological conditions during Level 2 were markedly different than the reference period: RH and wind speed.

#### 3.3. COVID-19 alert level PM$_{10}$ concentrations compared to the long-term mean (LTM)

Concentrations of PM$_{10}$ were reduced during Level 4 across all regions, ranging from 11.5% (South Lower) to 34.1% (South Upper), all lower than predicted by LTM (Table 4). During Level 3, PM$_{10}$ concentrations were closer to the LTM with a regional spread from 6.6% above predicted for the North Central region and 8.1% lower for the North Lower. During Level 2, PM$_{10}$ concentrations were more varied according to region with North Upper reporting observed concentrations 13.8% lower than LTM, whereas the adjoining North Central region showed concentrations almost 30.9% higher than LTM (Table 4).

#### 3.4. COVID-19 alert level PM$_{2.5}$ concentrations compared to the long-term mean (LTM)

During Level 4 all regions had lower PM$_{2.5}$ concentrations than the LTM; the largest reduction was in South Upper with a 22.6% reduction whilst North Upper had just a 2.0% decrease on LTM (Table 5). PM$_{2.5}$

---

**Table 3**

Overview of percentage difference of meteorological conditions compared to the reference period (2015–2019) during each alert level. (AT: air temperature; RH: relative humidity; WS: wind speed).

|                 | Level 4 |       | Level 3 |       | Level 2 |       |
|-----------------|---------|-------|---------|-------|---------|-------|
|                 | AT (%)  | RH (%)| WS(%)   | AT (%)| RH (%)  | WS(%) |
| North Upper (NU)| –4.6    | –6.3  | 7.4     | –3.6  | –7.0    | 27.0  |
| North Central (NC)| –1.7   | –7.6  | –8.6    | –1.2  | –7.0    | –6.1  |
| North Lower (NL)| –3.1    | –1.1  | 0.3     | –2.9  | –2.9    | –11.7 |
| South Upper (SU)| 0.6     | –6.9  | –8.5    | –0.9  | –5.8    | –8.6  |
| South Lower (SL)| –1.6    | –4.1  | 4.4     | 1.9   | –4.6    | 5.1   |
|                 |         |       |         |       |         |       |
|                 |         |       |         |       | 2.7     | –1.2  |
|                 |         |       |         |       | 3.9     | 7.7   |
|                 |         |       |         |       | –1.4    | 6.3   |
|                 |         |       |         |       | –13.2   | 13.4  |
|                 |         |       |         |       | 6.5     | 4.2   |
|                 |         |       |         |       | –15.9   |       |
|                 |         |       |         |       | –32.0   |       |
|                 |         |       |         |       | –17.5   |       |
|                 |         |       |         |       | –44.1   |       |
|                 |         |       |         |       | –19.1   |       |

---

Fig. 3. Linear regression modelling of relationship between BC concentrations set against traffic counts for light vehicles (A) and heavy vehicles (B) at three key traffic sites in New Zealand.
concentrations were variable across regions during Level 3 with the largest reductions in North Central (concentrations 25.7% lower than LTM), whereas North Upper and North Lower had lower concentrations higher than LTM (6.7 and 13.9%, respectively). During Level 2, most regions reported higher concentrations of PM$_{2.5}$ than predicted by LTM. The largest proportional difference was in the SL region, reporting PM$_{2.5}$ concentrations 35.3% above LTM. North Central was the only region to have PM$_{2.5}$ concentrations lower than LTM (12%).

3.5. COVID-19 alert level NO$_2$ concentrations compared to the long-term mean (LTM)

NO$_2$ concentrations markedly reduced during Level 4: all regions reported similar proportional reductions compared to LTM, ranging 48.0%-54.5% below the LTM (Table 6). All regions remained below the LTM during Level 3, although to a lesser degree, with a range from 3.7% (North Lower) to 44.7% (North Upper). The upwards trend in NO$_2$ concentrations continued during Level 2, NO$_2$ concentrations were 11.3% higher than LTM for the North Lower region during Level 2: however, concentrations for North Upper (19.9%) and South Lower (11.9%) were still reduced (Table 6).

3.6. Evaluation of machine learning using random forest algorithm

For PM$_{10}$, MAE ranges from 2.7 to 6.3 (μg/m$^3$) and NRMSE ranges from 0.3 to 0.9 (μg/m$^3$) (Fig. 4). MAE was generally higher in South Lower and lower in North Lower. Large variation of NRMSE is seen in most regional groups, except for South Upper. R squared varied widely from 0.07 to 0.84 with a median of 0.47. R squared values were relatively high at all stations in the South Island, whereas lower R squared were mostly found in the North Island.

ML modelling error results for PM$_{2.5}$ showed MAE ranges from 1.3 to 4.1 (μg/m$^3$) and NRMSE ranges from 0.2 to 0.5 (μg/m$^3$) (Fig. 5). A wide range of R squared values were found, from 0.2 to 0.8 with the median at 0.6. All measures showed large variations in the North Island but less variation from the South Island.

Error measures in NO$_2$ ranged similarly to PM$_{10}$ and PM$_{2.5}$, whereas R squared values were generally higher, with all values over 0.5 (Fig. 6). For NO$_2$, MAE had a range of 1.6–6.7 (μg/m$^3$) across regions and NRMSE were between 0.2 and 0.5 (μg/m$^3$), where variations in each region were found to be the highest in North Upper. NO$_2$ models yielded overall high R-squared from 0.5 to 0.8. The values were higher at stations in the South and lower at stations in the North.

For all air pollutants targeted in this study, the week of the year had been the most important variable for modelling, followed by wind speed (Fig. 7). Wind direction was an important factor as well. Mixed results were shown in air temperature, Unix time and RH, which are all more important for particulates than NO$_2$. Conversely, the day of the week had only been important for NO$_2$.

3.7. Observed COVID-19 air quality concentrations compared to predictions from random forest machine learning modelling

Using the RF machine learning algorithm, a predicted air quality concentration for each alert level was calculated based on observed meteorological conditions (see Section 2.2.3.2). These were then summarised for each region and compared to 2020 measurements and LTM. This comparison helps to support the estimated scale of changes in observed concentrations for different pollutants during the COVID-19 alert levels.

For PM$_{10}$, OB values were notably reduced during Level 4 when compared to estimates from both LTM and ML (Fig. 8), indicative of reduced emissions during this strict lockdown period. The ML predicted much higher concentration for the three North regions than was observed or suggested by LTM. For South Upper and South Lower, ML predictions closely matched the LTM but were, again, higher than observed concentrations. During Level 3, OB concentrations of PM$_{10}$
were notably higher than Level 4, with notable increases across all regions. OB and both LTM and ML methods were closely aligned for the South regions. For North areas, OB was still lower than LTM.

Level 2 OB concentrations closely corresponded to estimates from ML for South Lower and North Central, whilst for North and South Upper regions, OB values were below ML and LTM, which themselves were similar.

For Level 4, PM$_{2.5}$ OB values were lower than reference values produced by LTM or ML. ML modelling estimated higher values than OB or LTM for all Northern regions and were like LTM for Southern regions. During Level 3, PM$_{2.5}$ concentrations increased across all areas. ML estimates were close to OB for North Upper, North Lower and South Upper. During Level 2, OB PM$_{2.5}$ concentrations were higher than either LTM or ML estimates in the South. North Upper and North Lower OB data were close to LTM, whilst OB levels were slightly lower than LTM for the North Central region.

For NO$_2$, OB levels were notably lower than either LTM or ML estimates. Observed NO$_2$ concentrations increased across all regions during Level 3, however, the OB values remained below the LTM and ML values. For the Level 2 period, ML and OB results were similar for the North Upper and North Lower regions with LTM and ML estimates like each other but above the observed concentrations for the South Lower region.

4. Discussion

4.1. Air quality concentrations under COVID-19 restrictions

During the most severe social restrictions of Level 4, NO$_2$ concentrations were most reduced when compared to LTM (between 48.0 and
Fig. 7. Variable importance of Random forest models for each air pollutant concentrations on testing dataset (2015–2019). (week: week of the year; WS: wind speed; AT: air temperature; WD: wind direction; trend: Unix time; RH: relative humidity; wday: day of the week).

Fig. 8. Regional mean PM$_{10}$ concentrations for each COVID-19 alert level. The plots show observed data (OB) (left) compared to the long-term mean (LTM) (middle) and machine learning estimates (ML) by RF (right).

Fig. 9. Regional mean PM$_{2.5}$ concentrations for each COVID-19 alert level. The plots show observed data (OB) (left) compared to the long-term mean (LTM) (middle) and machine learning estimates (ML) by RF (right).
The mixture of natural and anthropogenic particulate sources should reduce PM$_{2.5}$ concentrations in response to human behaviour pattern changes. PM$_{10}$ concentrations were reduced compared to the LTM across New Zealand during Level 4, even though only a small proportion of PM$_{10}$ derives from on-road vehicle combustion. The decreases in observed concentrations may also be explained by less resuspended road dust loading from lower vehicle numbers using the roads.

PM$_{10}$ concentrations are also very heavily influenced by meteorology across New Zealand: higher winds can bring high sea salt concentrations elevating PM$_{10}$ mass across coastal areas (Davy et al., 2017). Conversely, lower wind speeds can help elevate PM$_{10}$ and PM$_{2.5}$ concentrations due to a build-up of wood smoke from home heating sources, and this appears to be the case for the South Upper and South Lower during Level 2, when temperatures lowered, and wind speeds were found to be anomalously lower than LTM. It is likely that during Level 2, household emissions increased due to higher home occupancy during this period when compared to pre-COVID times.

4.2. Impacts of wind speed and wind direction on air quality during COVID-19 restrictions

Wind direction was anomalous during Level 2 across most of New Zealand (Table 3), which likely influenced PM$_{2.5}$ concentrations. A North Upper regional site located in Auckland in the Henderson area provides a good example. The location experienced mostly normal wind conditions during Levels 4 and 3 (Fig. 11). During these alert levels, NO$_2$ concentrations were reduced in line with most roadside measurement stations across New Zealand (Fig. 10). However, PM$_{10}$ concentrations were close to LTM (Fig. 8). During Level 2, an unusual wind direction prevailed emanating from the southeast quadrant instead of the usual south westerlies (Fig. 11).

The result of anomalous wind direction was a reduction in PM$_{10}$ concentrations (Fig. 11C) due to the air flow no longer coming from the Tasman Sea and the associated sea salt loading that dominates Auckland’s air during this season (Davy et al., 2017). Conversely, NO$_2$ concentrations increased during Level 2 compared to the LTM (Fig. 11A and B). The increase in NO$_2$ can be attributed not just to the increase in traffic flow during this period compared to Levels 4 and 3 (Fig. 2) but could also be a product of air being drawn from the south easterly sector, a more urbanised neighbourhood of Auckland with light industry and many busy roads as compared to the prevailing south westerly, which draws air from the Waitakere Forest and, ultimately, the Tasman Sea (Fig. S3).

4.3. Performance of machine learning for NO$_2$, PM$_{10}$ and PM$_{2.5}$ concentrations

In PM$_{10}$ and PM$_{2.5}$, MAE reflected the reasonable amount of prediction errors in the daily average, which showed a good level of prediction from the models (Figs. 4 and 5). The larger errors found in the Southern regions were likely to result from the cooler autumn and early winter period, when particulate concentrations were higher due to residents lighting more fires for home heating purposes, a common practice in New Zealand (Ministry for Environment, 2018). NRMSSE has taken the scale of the air pollutant concentration at each site into account, which is a more comparable model accuracy measure. It appeared that accuracy of PM$_{10}$ models is similar across regional clusters, where outliers were found at most of the regions. With fewer PM$_{2.5}$ sites than PM$_{10}$ across New Zealand, PM$_{2.5}$ results were much more scattered within each region dependent on the key local emission source. High variation in R squared indicated that meteorology and time variables may not be sufficient in explaining much of the concentrations,
especially for Northern stations. Furthermore, many coastal stations, especially in the North Upper region, showed poor (<0.3) R squared values (Fig. 11). This suggested that there might be other variables that are significantly affecting PM$_{10}$ and PM$_{2.5}$ concentrations. An explanation could be due to the dominant natural source of particulates, sea salt, which is hard to predict by meteorology or trend variables (Davy et al., 2017).

The temporal aggregation (daily average) used in this study may have increased the difficulties in differentiating between the impacts of sea salt from the changeable meteorology at coastal stations, where the diurnal pattern of wind speed and direction is often caused by the changes in overland air temperature and sea surface temperature (Tian et al., 2018). The diurnal pattern of sea breezes may have reduced the interpretability of the model to better predict PM$_{10}$ and PM$_{2.5}$ at coastal stations. This would be especially true where competing PM emissions, such as wood smoke from home heating, would possibly peak when wind speeds (and therefore sea salt loadings) are low. On the other hand, overall high R squared values from NO$_2$ models have suggested that

Fig. 11. Anomalous meteorological impacts on air quality measurements: (A) NO$_2$ concentration roses for the Auckland air quality site in Henderson, a suburban peak traffic and residential location (North Upper). (B) NO$_2$ concentration distributions for each of the periods represented by the concentration roses. (C) PM$_{10}$ concentration distributions for each of the periods represented by the roses.

Fig. 12. Machine learning model performance (R squared) for PM$_{10}$, PM$_{2.5}$ and NO$_2$ concentrations across each selected sites in geographical groups on training dataset (2015–2019). R squared values were especially low at coastal stations for both PM$_{10}$ and PM$_{2.5}$, whereas they were high in general for all NO$_2$ stations.
meteorology and time variables account for most of the variation in \( \text{NO}_2 \) concentrations in the historical dataset (Fig. 12). The R squared values described here are commensurate to those of international studies (Petetin et al., 2020). Both metrics saw a larger variation of results in North Upper which may be driven in some part by the higher traffic volume in heavy vehicles. While North Central yielded the lowest MAE, models in South Lower seemed to be more accurate due to lower NRMSE.

Results have indicated that the developed RF algorithm performed better at predicting \( \text{NO}_2 \). While error metrics implied that predictions from all models were acceptable, R squared have hinted that there are more local factors affecting concentration of particulates. More research is needed to reveal other factors that might be helpful in predicting particulates concentrations.

4.4. Assessing the use of long-term mean (LTM) or machine learning (ML) as reference under business-as-usual scenario

The Henderson site is situated in a location with oceanic influences, such as sea salt on its PM\(_{10}\) concentrations. In fact, during summer months sea salt dominates the PM\(_{10}\) fraction at this site (Davy et al., 2018). The stability of the sea salt component and the response of the model might indicate that the model is over-predicting the sea salt component. It has been shown that ML methods had lower R-squared values for coastal influenced sites. To test this, Henderson PM\(_{10}\) data can be compared to an air quality site located inland - Alexandra, a South Island township in the South Lower region that has high wintertime PM\(_{10}\) concentrations dominated by wood smoke from home heating. Observed PM\(_{10}\) concentrations in Henderson were lower than BAU, as indicated by LTM or ML (Fig. 13A and B). The case of ML predictions overestimating PM\(_{10}\) concentrations occurred for several other coastal stations.

By contrast, for the Alexandra South Lower site, ML estimates and LTM are more closely aligned to OB for the alert level periods (Fig. 13C and D). This exemplifies the better model performance on inland stations that have a stronger single PM\(_{10}\) emission source, which in Alexandra’s case is wood smoke from domestic fires.

With significantly lighter winds across many regions during Level 2, coastal regions experienced lower sea salt concentrations, reducing mass loading in the PM\(_{10}\) fraction (Fig. 13A and B), with a marked drop in the observed data when compared to the LTM. By contrast, the same reduction of wind speed for inland areas when cooler conditions prevail would reduce dispersion of wood smoke from home heating, allowing higher concentrations. For air quality monitoring sites that are influenced by both wood smoke and sea salt, such as Henderson, the confounding factors most likely explain, in part, low R squared values of ML models.

To test the ML performance results for \( \text{NO}_2 \), the Henderson site was compared to a North Lower air quality station, Wellington Central (Willis Street). The Wellington Central permanent air quality station is next to a busy road in the city. During Level 4, a sharp drop in \( \text{NO}_2 \) concentrations occurred (Fig. 14A–D), a result of the removal of the direct emission source for the location, namely on-road vehicles. The attribution of \( \text{NO}_2 \) reductions to human behaviour changes in emissions is simpler than that of PM\(_{10}\) given that almost all \( \text{NO}_2 \) derives directly from diesel- and petrol-powered vehicle emissions, with low concentrations from natural sources or imported into New Zealand (Xie et al., 2019). During Level 3, Wellington Central concentrations rapidly returned to values close to the LTM, whilst Henderson remained lower than LTM. In Level 2, observed concentrations were close to LTM and ML estimates. ML is more closely aligned to observed \( \text{NO}_2 \) concentrations than LTM, even at the same Henderson site under the same meteorological conditions (Fig. 14). ML modelling also appears to reflect the nuanced increase in concentrations with the country moving into colder conditions.
Atmospheric Environment 254 (2021) 118322

winter conditions as alert levels decreased. Wintertime in New Zealand generally has the highest NO$_2$ concentrations with increased traffic volume and reduced vertical mixing due to less soil heating of the surface.

5. Summary

This study considers the COVID-19 alert level restrictions that, in effect, created a “natural experiment” on New Zealand’s air quality. As noted elsewhere, the sudden disruption of public movement brought about by lockdowns reduced air pollution significantly. However, the amount of initial reduction varied considerably among different sites and pollutants.

Transport-related NO$_2$ concentrations reduced 48.0% to 54.5% during the most restrictive alert level period, reflecting the immediate reduction in combustion emissions and replicating findings reported in Auckland (Patel et al., 2020) and like reductions measured in other locations with COVID-19 restrictions (e.g., Tobias et al., 2020). Nationally, reductions in NO$_2$ concentrations were similar across the country during Level 4, but some regional differences emerged in NO$_2$ levels as restrictions eased through Levels 3 and 2. These differences are most likely explained by changes in traffic volumes during each alert level along with the proportion of vehicle types (diesel and petrol) using the roads during these periods.

PM$_{10}$ and PM$_{2.5}$ was found to decrease less than NO$_2$ compared to both LTM and ML during Level 4. The scale and regionality of their proportional changes in particulate loading are a result of multiple sources contributing to particulate mass concentrations, including substantial contributions from oceanic sources (sea salt and secondary sulfates) along coastal fringes (Davy et al., 2017). As the alert levels eased, concentrations in both PM$_{10}$ and PM$_{2.5}$ increased; however, the scale of the increase was more notable in South regions where the burning of wood for home heating is more prevalent during this cooler period of the year. While public movements decreased significantly during the COVID-19 alert levels, there was an increase in time spent at home, indoors. This is reflected in increases in PM$_{10}$ and PM$_{2.5}$ concentrations during Level 2, most likely due to increased wood burning for domestic heating. These domestic emissions occurred as the South regions became colder during this period and home occupancy was higher than usual due to travel restrictions.

BAU reference values were developed using LTM and ML modelling. The ML modelling closely represented observed NO$_2$ concentrations across the country once social restrictions were eased. However, during Level 4 both LTM and ML were much higher than observations, a direct result of the reduction of on-road vehicle transport during this time. The closeness of the ML and LTM results during the return of more normal traffic volume during lower alert levels adds confidence to the scale of reduction during Level 4. The ML modelling for PM$_{10}$ and PM$_{2.5}$ results are more mixed, with closer analysis showing that the R squared values of models were lower along coastal fringes, where the predictions were often higher than the observed values. The complexity of the land-ocean interaction on particulate concentrations and, more broadly, the larger number of particulate sources most likely help explain this important finding. Contrary to this, for monitored inland regions, ML predictions more closely matched LTM. The results favour a mixed approach to creating BAU reference values, with ML modelling favoured for gas pollutants and particulates away from coastal influences, while LTM is better for those coastal environments. Using such an approach, future work could include land use factors and direct emissions in modelling computation.

Credit author contribution statement

Nick Talbot: Conceptualization, Methodology, Formal Analysis and Writing – original draft. Akika Takada: Conceptualization, Writing – Original draft, Formal analysis, Software, Visualisation. Nancy E. Golubiewski: Conceptualization, Writing - Review & Editing, Formal Analysis. Funding Acquisition, Project Administration, Supervision.
N. Talbot et al.

Drew Bingham: Writing - Review & Editing. Samantha Lay Yee: Conceptualization, Investigating, Writing - Review & Editing.

Declaration of competing interest
No known competing interests.

Appendix A. Supplementary data
Supplementary data to this article can be found online at https://doi.org/10.1016/j.atmosenv.2021.118322.

References
Alimisis, et al., 2018. Spatial estimation of urban air pollution with the use of artificial neural network models. Atmos. Environ. 191, 205–213. https://doi.org/10.1016/j.atmosenv.2018.07.058.
Batista, G.E., Monard, M.C., 2002. A study of K-nearest neighbour as an imputation method. History 87 (251-260), 48.
Berman, J.D., Ebi, K., 2020. Changes in U.S. air pollution during the COVID-19 pandemic. Science of the Total Environment, 739. https://doi.org/10.1016/j.scitotenv.2020.139864, 139864.
Cardlaw, D.C., Hopkins, K., 2012. Openair — an R package for air quality data analysis. Environ. Model. Software. 27–28, 52–61.
Chai, T., Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. Geosci. Model Dev. (GMD) 7 (3), 1247–1256.
Chen, X., Ishwaran, H., 2012. Random forests for genomic data analysis. In: Genomics, vol. 99. Academic Press, pp. 323–329. https://doi.org/10.1016/j.ygeno.2012.04.005.
Davy, P.K., Trompeter, W.J., Aneel, T., Markwitz, A., 2017. Source Apportionment and Trend Analysis of Air Particulate Matter in the Auckland Region. Issue July.
Ebenezer, G., 2019. The Role of Meteorology in Atmospheric Processes and Air Pollution Studies. December, 0–14.
Feng, et al., 2019. Recurrent Neural Network and random forest for analysis and accurate forecast of atmospheric pollutants: a case study in Hangzhou, China. J. Clean. Prod. 231, 1005–1015. https://doi.org/10.1016/j.jclepro.2019.05.316.
Grange, S.K., Carslaw, D.C., Lewis, A.C., Boleti, E., Hueglin, C., 2018. Random forest meteorological normalisation models for Swiss PM10 trend analysis. Atmos. Chem. Phys. Discuss. 1–28 https://doi.org/10.5194/acp-2017-1092.
Griffith, S.M., Huang, W.S., Lin, C.C., Chen, Y.C., Chang, K.E., Lin, T.H., Wang, S.H., Lin, N.H., 2020. Long-range air pollution transport in East Asia during the first week of the COVID-19 lockdown in China. Sci. Total Environ. 741 https://doi.org/10.1016/j.scitotenv.2020.140214.
Gulia, S., Shiva Nagendra, S.M., Khare, M., Khanna, I., 2015. Urban air quality management: A review. Atmos. Pollut. Res. 6 (2), 286–304. https://doi.org/10.5094/ APR.2015.033.
Herrera, V.M., Khoshgoftaar, T.M., Villanustre, F., Furbh, B., 2019. Random forest implementation and optimization for Big Data analytics on LexisNexis’s high performance computing cluster platform. J. Big Data 6 (1). https://doi.org/10.1186/s40537-019-0232-1.
Hudda, N., Simon, M., Patton, A., Durant, J., 2020. Reductions in traffic-related black carbon and ultrafine particle number concentrations in an urban neighborhood during the COVID-19 pandemic. Sci. Total Environ. 140931. https://doi.org/10.1016/j.scitotenv.2020.140931 xxxx.
Kerimray, A., Baimatova, N., Ibragimova, O.P., Bukenov, B., Kenessov, B., Plotitsyn, P., 2020. COVID-19 pandemic: impacts on the air quality during the partial lockdown in Sáo Paulo state, Brazil. Sci. Total Environ. 730 https://doi.org/10.1016/j.scitotenv.2020.139542.
NIWA, The national climate database (GFiFo). k.d. Retrieved January 1, 2012, from http://cliflo.niwa.co.nz/.
Patel, H., Talbot, N., Salmond, J., Dirks, K., Xie, S., Davy, P., 2020. Implications for air quality management of changes in air quality during lockdown in Auckland ( New Zealand ) in response to the 2020 SARS-CoV-2 epidemic. Sci. Total Environ. 746, 141129. https://doi.org/10.1016/j.scitotenv.2020.141129.
Petetin, H., Bowdalo, D., Soret, A., Guevara, M., Jorba, O., Serradell, K., Pèrez García, Pando, C., 2020. Meteorology-normalized impact of COVID-19 lockdown upon NO2 pollution in Spain. Atmos. Chem. Phys. Discuss. 2, 1–29. https://doi.org/10.5194/acp-2020-446.
R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org.
Robert, A., 2020. Lessons from New Zealand’s COVID-19 outbreak response. Lancet Public Health 5 (11), e569–e570. https://doi.org/10.1016/S2468-2667(20)30237-
Sharma, S., Zhang, M., Ambika Gao, J., Zhang, H., Kota, S.H., 2020. Effect of restricted emissions during COVID-19 on air quality in India. Sci. Total Environ. 728 https://doi.org/10.1016/j.scitotenv.2020.138878.
Sicard, P., De Marco, A., Agathokleous, E., Feng, Z., Xu, X., Paolletti, E., Rodriguez, J.J.D., Calatayud, V., 2020. Amplified ozone pollution in cities during the COVID-19 lockdown. Sci. Total Environ. 735, 139542. https://doi.org/10.1016/j.scitotenv.2020.139542.
Siciliano, B., Carvalho, G., da Silva, C.M., Arbillá, G., 2020. The impact of COVID-19 pandemic on primary pollutant concentrations in the atmosphere of Rio de Janeiro and São Paulo Megacities (Brazil). Bull. Environ. Contam. Toxicol. 105 (1), 2–8. https://doi.org/10.1007/s00128-020-02907-9.
Talbot, N., Leth, R., 2018. The Impacts of Transport Emissions on Air Quality in Auckland’s City Centre (Issue December).
Tian, Y., Liu, J., Han, S., Shi, X., Shi, G., Xu, H., Hu, Y., Zhang, Y., Feng, Y., Russell, A.G., 2018. Spatial, seasonal and diurnal patterns in physicochemical characteristics and sources of PM2.5 in both inland and coastal regions within a megacity in China. J. Hazard Mater. 342, 139–149. https://doi.org/10.1016/j.jhazmat.2017.08.015.
Tobías, A., Carnerero, C., Reche, C., Massaguet, J., Via, M., Minguillon, M.C., Alastuey, A., Querol, X., 2020. Changes in air quality during the lockdown in Barcelona (Spain) one month into the SARS-CoV-2 epidemic. Sci. Total Environ. 726, 138540. https://doi.org/10.1016/j.scitotenv.2020.138540.
Wang, Q., Su, M., 2020. A preliminary assessment of the impact of COVID-19 on environment – a case study of China. Sci. Total Environ. 728 https://doi.org/10.1016/j.scitotenv.2020.139915.
Weekly traffic count information. Weekly traffic count information | Waka Kotahi NZ transport agency. Retrieved April 28, 2020, from https://www.nzta.govt.nz/about-us/coronavirus-disease-covid-19-services-update/weekly-traffic-count-information/.
Wilder-Smith, A., Freedman, D.O., 2020. Isolation, quarantine, social distancing and community containment: pivotal role for old-style public health measures in the novel coronavirus (2019-nCoV) outbreak. Int. J. Travel Med. 27, 15–29. https://doi.org/10.5194/ijtm-2020-002.
Willmott, C.J., Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing model performance. Clim. Res. 30 (3), 79–82.
World Health Organisation, 2020. Coronavirus disease (COVID-19) situation reports. Coronavirus disease (COVID-2019) situation reports. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports.
Xie, S., Sridhar, S., Metcalfe, J., 2019. Auckland Air Emissions Inventory 2016, vol. 4525 environmental releases during COVID-19 on air quality in India. Sci. Total Environ. 728 https://doi.org/10.1016/j.scitotenv.2020.138540.
Zealand ) in response to the 2020 SARS-CoV-2 epidemic. Sci. Total Environ. 746, 141129. https://doi.org/10.1016/j.scitotenv.2020.141129.
Zambrano-Monserrat, M.A., Ruano, M.A., 2020. Has air quality improved in Ecuador during the COVID-19 pandemic? A parametric analysis. Air Quality. Atmos. Health. https://doi.org/10.1007/s11869-020-00866-y.