Informing you on ambient air quality
in the Belgian Regions

Effects of COVID-19 induced lockdown measures on air quality in Belgium

an assessment using a Random Forest model

Vlaanderen is milieu

BRUXELLES ENVIRONNEMENT
LEEFMILIEU BRUSSEL
- IBGE - BIM -

Air Climat
agence wallonne de l'air & du climat

VLAAMSE
MILIEUMAATSCHAPPIJ

Content

List of figures

List of tables

Summary

Estimating the impact of the COVID-19 crisis on air quality is not as simple as it seems. In this report, the effects of the COVID-19 related lockdown on the concentrations of various pollutants are estimated using a "Random Forest" (RF) model (for weeks 12 to 19 in 2020).

With an RF model, the effect of the meteorological conditions can be separated from the effects of changes in the emissions as a result of the corona lockdown which started in mid-March 2020.

This exercise was carried out for different types (traffic, urban, background) of measuring stations in the three Belgian regions[1].

Based on the RF model exercise, it appears that the (soft) lockdown measures had a variable effect on air quality (weeks 12 - 19):

- There is a clear positive impact on the concentrations of typical traffic-related pollutants such as nitrogen oxides (NO_x), nitrogen dioxide (NO_2) and black carbon. Without the COVID-19 crisis, the concentrations of these substances would have been more than 50% (NO_x) and 35-40% (NO_2 and black carbon) higher at the most motor vehicle congested measurement locations. This impact decreases when or where there is less motorised traffic in the vicinity of the measuring stations.
- The corona measures seem to have only a limited impact on the particulate matter concentrations (PM2.5 and PM10). After all, the contribution of the "primary" or directly emitted particulate matter to the total concentration of particulate matter is limited: particulate matter has many sources other than motorised traffic. Industry, households and agriculture are also important sources of particulate matter. These sectors were less affected by the corona crisis.
- There is a negative effect on the ozone concentrations: ground-level ozone (troposphere) is a complex game between ozone formation and depletion. The amount of substances (nitrogen oxides and volatile organic compounds) to form ozone remained high enough despite the reduction in emissions of these substances. However, with less traffic there was less ozone depletion, resulting in an increase in ozone concentrations.

The results from this report were used in the VMM report "het Effect van COVID-19-maatregelen op de luchtkwaliteit in Vlaanderen".

[1] *The air quality monitoring networks are operated by the Vlaamse Milieumaatschappij (https://www.vmm.be) in Flanders, by l´Institut Scientifique de Service Public (https://www.issep.be) and the Agence Wallonne de l'Air et du Climat (AWAC) (http://airclimat.wallonie.be) in Wallonia, and in Brussels by the Brussels Institute for Environmental Management (https://www.ibgebim.be).*

1. Separating the influence of meteorological conditions and changes in emissions

Quantifying the impact of the lockdown measures on air quality on the basis of (only) measurements is not as simple as it seems. After all, the concentrations of harmful substances in the air are not only determined by the emissions, but also by the weather conditions. By comparing the measurements after the start of the lockdown with periods immediately before the lockdown or with periods in previous years, it is not easy to distinguish the effect of changes in emissions and weather conditions. This problem was already outlined in the news report at the beginning of April 2020.

A Chemical Transport Model (CTM) can be used to isolate the individual effect of the corona measures on air quality from the effects of meteorological conditions. CTMs are models that simulate the complex physical and chemical processes in the atmosphere with mathematical algorithms. These models use emissions, meteorological and geographical data as input. If the effect of the corona measures on the emissions can be estimated, a CTM model can be used to determine the impact of only the change in the emission (the meteorological conditions remain the same) on the concentrations in the air. Two model runs must therefore take place: one with “normal” or “Business As Usual” (BAU) emissions where it is assumed that there would have been no corona crisis and one with the emissions as estimated during the lockdown. The difference between the calculated concentrations of the two model runs is the separate effect of the corona measures on air quality.

These kind of model calculations have a number of advantages (you can make a calculation for the entire territory, also in places where no measurements are available) but are complicated. In addition, a correct estimate of the emission changes during the lockdown is necessary. This is not easy: a decrease in motorised traffic volume results in less traffic-related pollution, but it is not easy to estimate exactly how much. It is also not yet clear what the impact is of the corona measures on emissions in other sectors.

2. Random Forest models

Random Forest (RF) models are a relatively new type of machine learning models that are able to determine non-linear correlation between variables from large datasets (https://link.springer.com/article/10.1023/A:1010933404324). For this, an RFmodel is first trained: the model uses decision trees to determine the relationship between a parameter (in this case the concentration of an air pollutant) and variables (for example the wind direction, temperature, ...), that individually have only a limited predictive value in many cases. All the decision trees are then bundled into one so-called "random forest". By combining all the decision trees you obtain an algorithm that makes a robust prediction for a parameter (in this study the concentration of a pollutant in the air) with a set of variables (for example, meteorological parameters) that you used in the training process.

Based on the measured concentrations and the associated set of meteorological (and possibly) other variables from the past, we can “predict” the relationship between the measured concentrations and those weather and other variables using a Random Forest model. Using an RF model we can also try to estimate the individual effect of the lower emissions from motorised traffic during the corona lockdown by comparing the model results with measured concentrations. The big difference with CTM models is that with this approach it is not necessary to use emission data as input.

Various tools are now available to set up a Random Forest model. In this exercise we used the "rmweather" (https://CRAN.R-project.org/package=rmweather) package for the statistical software program "R". The package is specifically designed to study the variation in pollutant concentrations due to varying meteorological conditions using the Random Forest technique.

3. Validation of the Random Forest model

An RF model was set up and trained with the daily mean pollutant concentrations and a series of weather parameters (daily mean, minimum and maximum wind speed, wind direction, daily mean, minimum and maximum temperature, relative humidity, atmospheric mixing layer height and cloudiness). The day in the year (as a seasonal trend), the day of the week (to estimate the effect of the weekday and weekend day) and the elapsed time (as an indicator for the long-term trend) since 1/1/1970 (the so-called "unix timestamp") are used as additional variables to train the model. For the training data from 1/1/2015 to 29/2/2020[2] was used, i.e. the period until just before the introduction of the corona measures.

The meteorological dataset was created based on the measurements in the telemetric weather stations. The list of measuring stations used can be found in Table 1.

Table 1: Weather stations from the monitoring networks of the three regions used for the meteorological dataset.

Station code	Region	Municipality
T2M802	Flemish Region	Antwerp
T4M701	Flemish Region	Ghent
T4N029	Flemish Region	Veurne
T1M003	Brussels-Capital Region	Molenbeek
T5M501	Walloon Region	Charleroi
T3M202	Walloon Region	Liège
T3M205	Walloon Region	Sainte-Ode

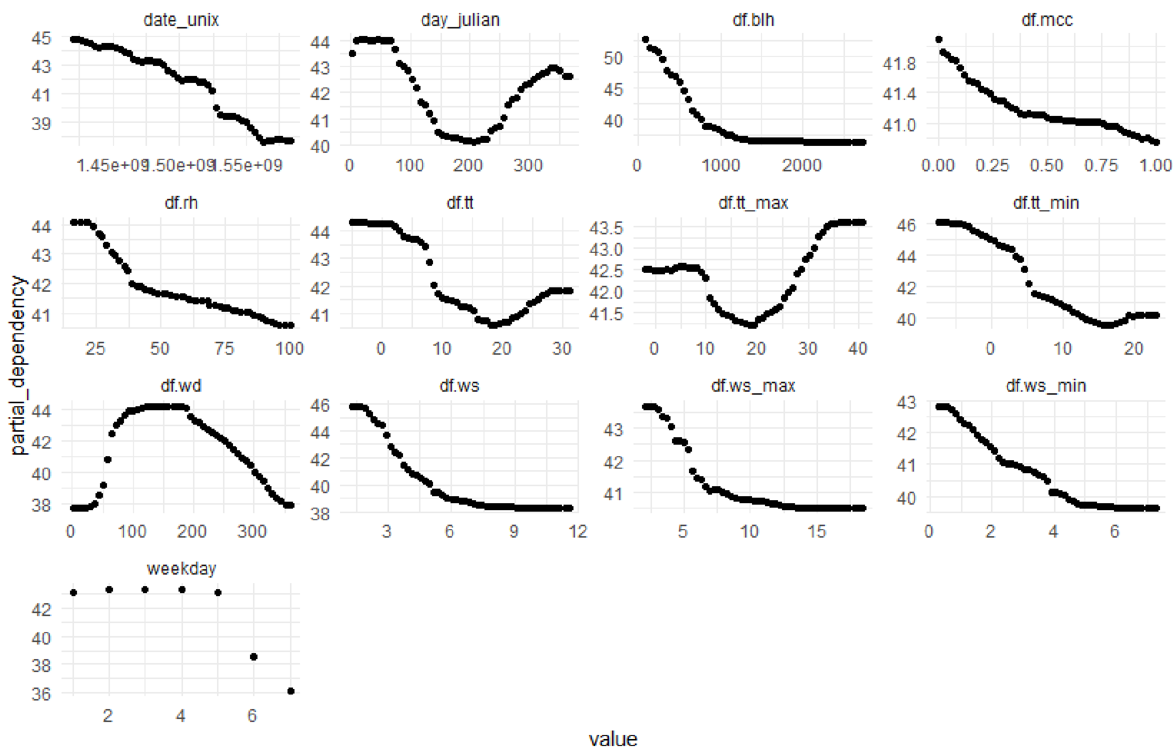
The following parameters measured in these stations were used: wind speed, wind direction, temperature and relative humidity. Data from the European Centre for Medium-Range Weather Forecast (ECMWF) were used for the mixing layer height and the degree of cloudiness.

Figure 1 shows the shared dependencies of the input parameters after training of the RF model for nitrogen dioxide (NO_2) in the traffic measuring station 42R802 in Borgerhout. These figures thus represent the individual effect of a parameter on the modelled concentration. This shows that the

[2] The time series is shorter for a number of measurement locations and pollutants.

model can realistically estimate the relationships that exist between the weather parameters and air pollution (for example, higher NO_2 concentrations at lower wind speed and mixing layer height). The relationships between the time-bound variables and the NO_2 concentrations (such as lower NO_2 in the summer months and during the weekend) are also within expectations.

Figure 1: dependence of the different input parameters after training the Random Forest model.



To assess the performance of the RF algorithm in predicting the concentrations of a pollutant at an air quality measurement station based on the meteorological and time-dependent variables, a random test set was made: 20% of all days in the measurement series between 1/1/2015 and 29/2/2020 were excluded to train the model. The selection of those 20% days was done randomly by the computer program. The remaining 80% of the days were thus used to train the model. When the concentrations calculated by the RF model on the “test days” are compared with the actually measured concentrations, an independent estimation of the performance of the model can be made. Independent because the test days were not included in the training.

Table 2: measurement stations, classification and pollutants for which a Random Forest model was set up

Code	Measuring point	Region	Classification	Weather station	NO_x	NO_2	PM2.5	PM10	BC	O_3
42R801	Antwerp (Borgerhout)	VLA	urban background	T2M802	x	x	x	x	x	x
42R802	Antwerp (Borgerh-straat)	VLA	urban street	T2M802	x	x	x	x	x	
42R803	Antwerp (park spoor N)	VLA	urban background	T2M802	x	x	x	x	x	
42R804	Antwerp (Ring)	VLA	urban street	T2M802	x	x	x	x	x	
42R805	Antwerp (Belgiëlei)	VLA	urban street	T2M802	x	x	x	x	x	
42R817	Antwerp (Wilrijk)	VLA	suburban/inner city	T2M802	x	x	x	x	x	
44R701	Ghent (Baudelo)	VLA	urban background	T4M701	x	x	x	x	x	x
44R702	Ghent (Gustaaf Callier)	VLA	urban street	T4M701	x	x	x	x	x	
44R703	Ghent (Lange Violettestraat)	VLA	urban street	T4M701					x	
44N029	Veurne (Houtem)	VLA	background	T4M029	x	x	x	x	x	x
42N040	Sint-Pieters-Leeuw	VLA	background	T1M001	x	x				x
42N016	Dessel	VLA	background	T2M802[3]	x	x	x	x	x	x
42N046	Lanaken (Gellik)	VLA	background	T2M802**Error! Bookmark not defined.**	x	x				x
41B001	Brussels (Kunst-Wet)	BRU	urban street	T1M001	x	x				
41R001	Molenbeek	BRU	urban street	T1M001	x	x	x	x	x	x
41B004	Brussels (Katelijne)	BRU	inner city	T1M001	x	x				x
41R012	Ukkel	BRU	urban background	T1M001	x	x	x	x	x	x
45R501	Charleroi	WAL	urban background	T5M501	x	x	x	x		
45R502	Charlerloi (Lodelinsart)	WAL	urban background	T5M501	x	x	x	x		x
43R401	Namur	WAL	urban background	T5M501	x	x	x	x		x
43R222	Liège	WAL	urban background	T3M202	x	x	x	x		x
43N060	Havinnes	WAL	background	T5M501	x	x	x	x		x
43N063	Coroy-Le-grand	WAL	background	T5M501	x	x	x	x		x
43N100	Dourbes	WAL	background	T3M202	x	x	x	x		x

[3] For the weather parameters (except mix layer height and cloudiness) in the monitoring stations Dessel (Kempen) and Lanaken (East Limburg), the weather station in Antwerp was used because no weather stations are available in the immediate vicinity of these stations.

An RF model was set up for 6 pollutants and 13 measurement points in Flanders, 4 in Brussels and 7 in Wallonia. Based on these measurement locations, in a next step (see 4), we can obtain a global picture of the impact of the corona lockdown measures on air quality in Belgium (on urban street, urban background and background measurement locations). For more details about the measurement locations and the pollutants per measurement location, see Table 2.

Figure 2 shows the scatter plots for the pollutants (NO_x, NO_2, O_3, PM2.5, PM10 and BC) for the urban background measuring stations Antwerp-Borgerhout (42R801), Molenbeek (41R001) and Namur (43R401). On the x-axis the measured daily mean concentrations are shown, on the y-axis the predicted daily mean concentrations by the random forest, based on the weather conditions and time-dependent variables on those days, are displayed. All of the selected pollutants (6) are measured in these 3 measuring stations. The scatter plots for the other monitoring stations are comparable.

The graphs also contain a number of statistical measures that indicate how the model is performing (for more information about this, see Appendix I: Statistical validation indicator). Table 3 shows the median of these measures for all stations per pollutant. Based on these statistics, we can say that the RF model can reliably estimate the concentrations for NO_x, NO_2, O_3 and BC (high R^2, low MB, and acceptable RMSE). For NO_x, NO_2, O_3 and BC 75 to 80% of the variance can be explained by the RF model. For PM2.5 and PM10 the validation statistics (especially the R^2) show somewhat less agreement, but the results are still acceptable: 50 to 60% of the variance can be explained by the model. It is important to take this into account when the RF model is used to calculate the impact of the corona measures (see 4).

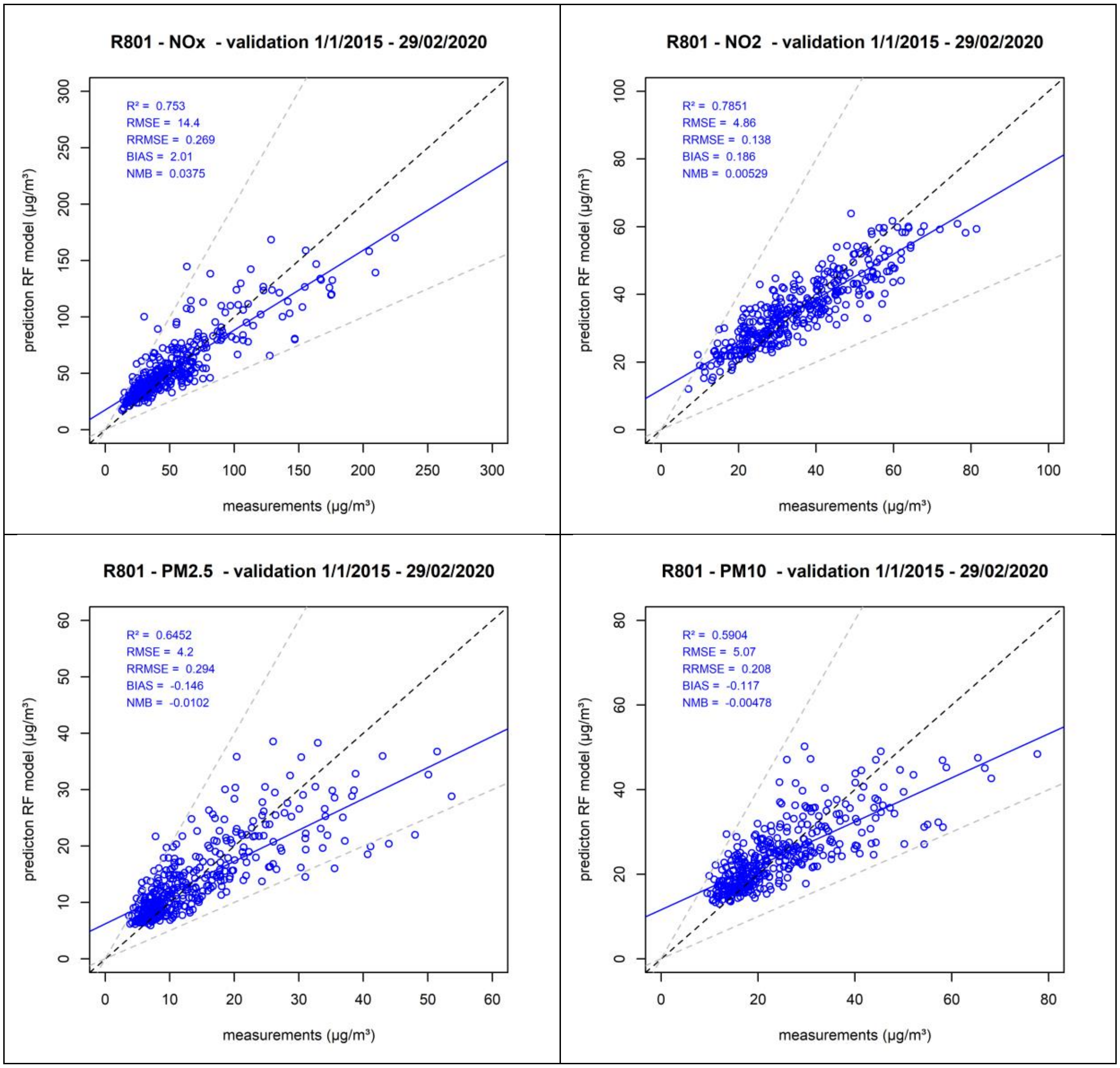
The scatter plots also show that the RF model slightly overestimates the lowest concentrations and slightly underestimates the highest concentrations. This is clearly more pronounced for PM2.5 and PM10.

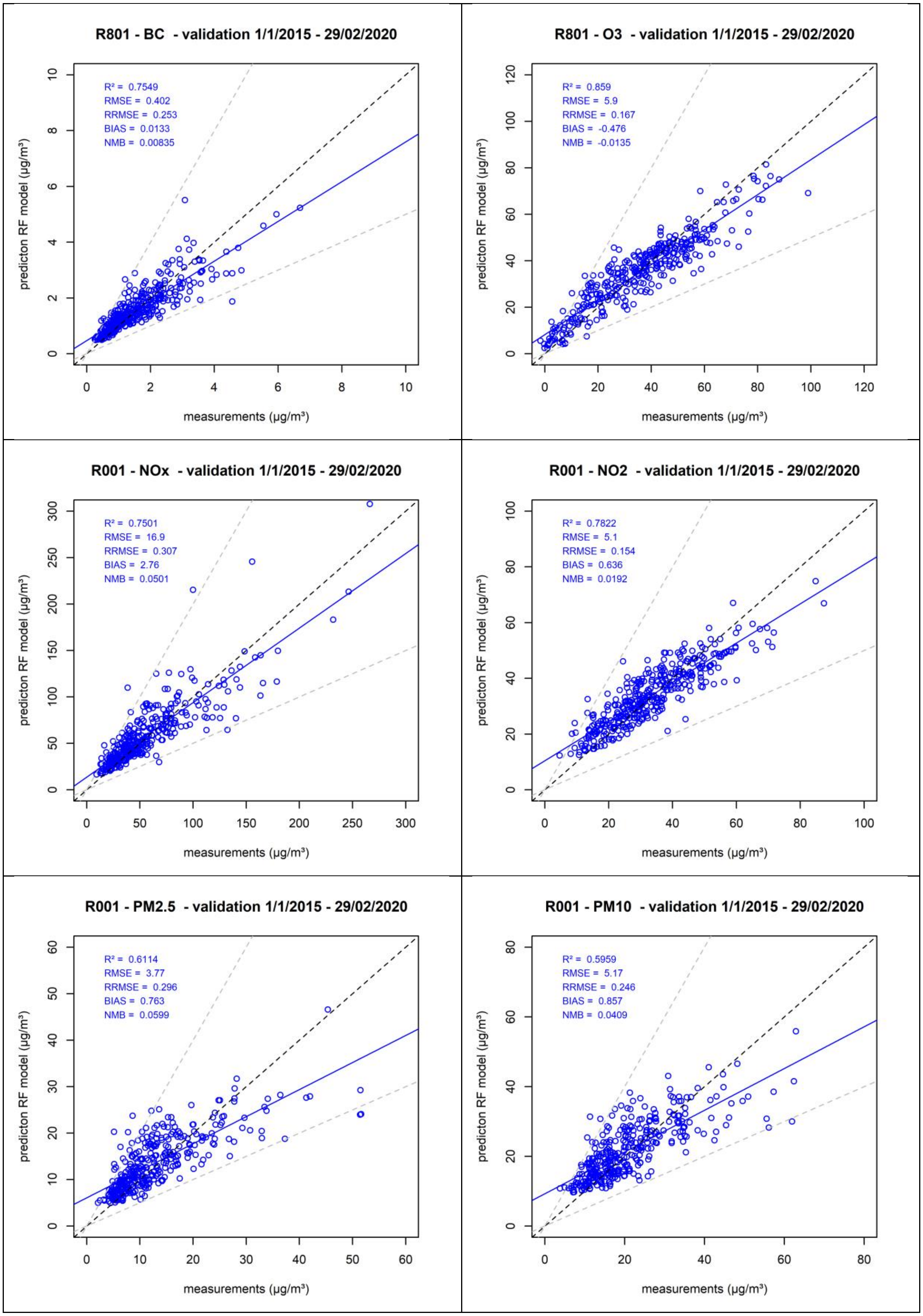
This validation exercise shows that when the temperature, wind direction, wind speed, day of the week, *etc.* are known on a day and when there are no sudden changes in emissions, a reliable estimate can be made of the pollutant concentrations on that day. We can only perform this validation at places where measurements are available and where a sufficiently long time series of pollutant concentrations and weather variables are available.

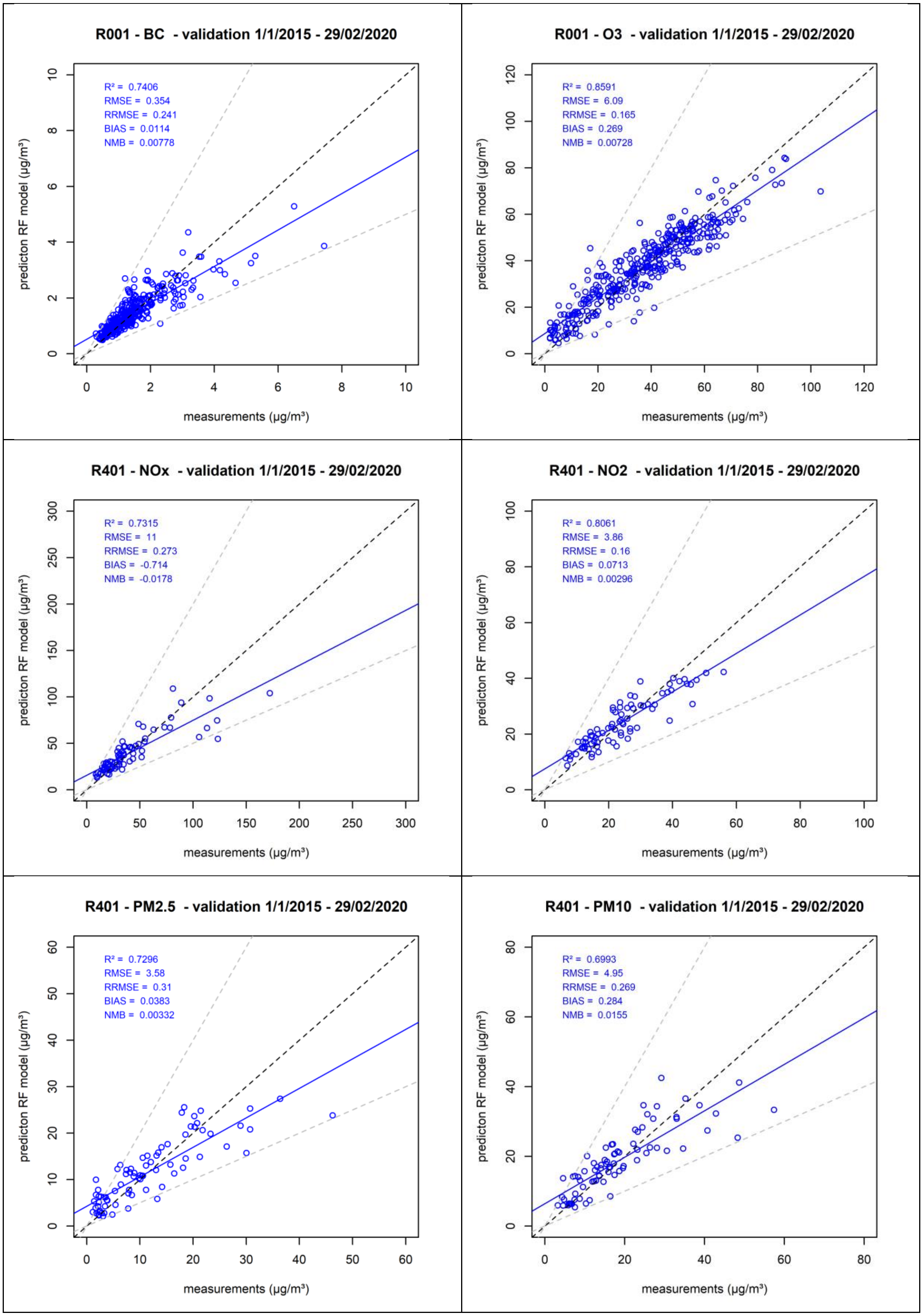
Table 3: Statistical validation parameters (median selection of measuring stations) per pollutant

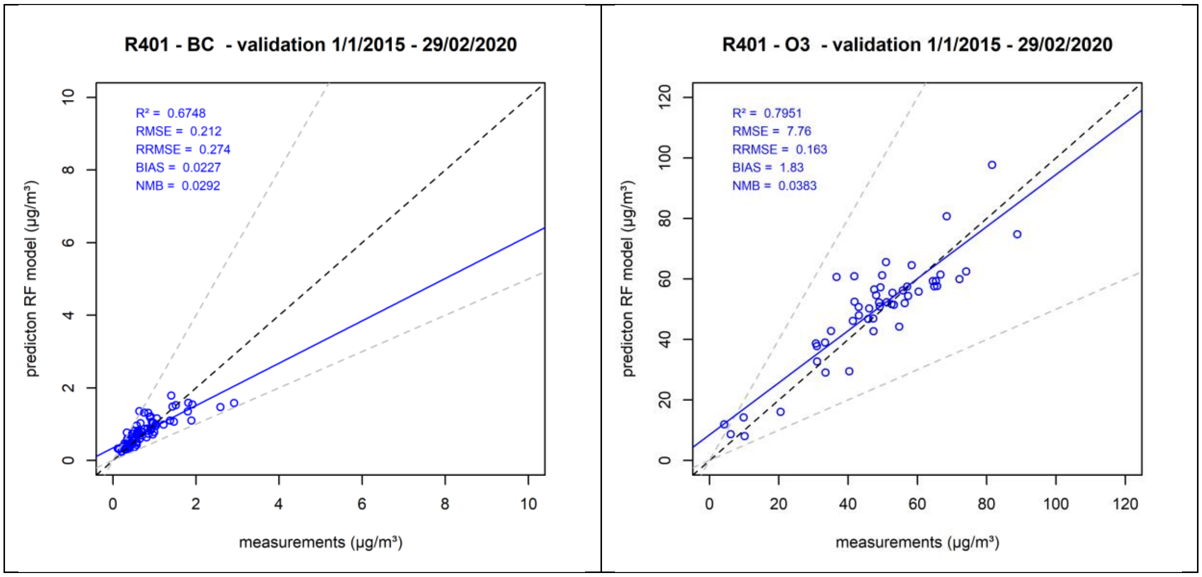
	R^2	RMSE	RRMSE	MB	NMB
NO_x	0.73	11.80	28.8%	0.32	1.7%
NO_2	0.76	4.37	15.5%	0.19	0.5%
PM2.5	0.61	4.33	34.7%	0.34	2.9%
PM10	0.54	5.17	24.4%	0.27	1.8%
BC	0.75	0.36	24.3%	0.03	0.8%
O_3	0.84	6.95	15.6%	0.15	0.1%

Figure 2: validation of the RF model for the urban background measurement stations 42R801 (Antwerp-Borgerhout), 41R001 (Brussels-Molenbeek) and 43R401 (Namur) for 6 pollutants. The measurement results are shown on the x-axis, the RF model results on the y axis. The blue line is the regression line. The black dashed line is the x = y line, the grey dashed lines mark the interval where the measurements and the RF model results differ by a maximum of a factor of 2.




R801 - BC - validation 1/1/2015 - 29/02/2020
R² = 0.7549
RMSE = 0.402
RRMSE = 0.253
BIAS = 0.0133
NMB = 0.00835
predicton RF model (µg/m³)
measurements (µg/m³)
R801 - O3 - validation 1/1/2015 - 29/02/2020
R² = 0.859
RMSE = 5.9
RRMSE = 0.167
BIAS = -0.476
NMB = -0.0135
predicton RF model (µg/m³)
measurements (µg/m³)
R001 - NOx - validation 1/1/2015 - 29/02/2020
R² = 0.7501
RMSE = 16.9
RRMSE = 0.307
BIAS = 2.76
NMB = 0.0501
predicton RF model (µg/m³)
measurements (µg/m³)
R001 - NO2 - validation 1/1/2015 - 29/02/2020
R² = 0.7822
RMSE = 5.1
RRMSE = 0.154
BIAS = 0.636
NMB = 0.0192
predicton RF model (µg/m³)
measurements (µg/m³)
R001 - PM2.5 - validation 1/1/2015 - 29/02/2020
R² = 0.6114
RMSE = 3.77
RRMSE = 0.296
BIAS = 0.763
NMB = 0.0599
predicton RF model (µg/m³)
measurements (µg/m³)
R001 - PM10 - validation 1/1/2015 - 29/02/2020
R² = 0.5959
RMSE = 5.17
RRMSE = 0.246
BIAS = 0.857
NMB = 0.0409
predicton RF model (µg/m³)
measurements (µg/m³)


R001 - BC - validation 1/1/2015 - 29/02/2020
R² = 0.7406
RMSE = 0.354
RRMSE = 0.241
BIAS = 0.0114
NMB = 0.00778
predicton RF model (µg/m³)
measurements (µg/m³)
R001 - O3 - validation 1/1/2015 - 29/02/2020
R² = 0.8591
RMSE = 6.09
RRMSE = 0.165
BIAS = 0.269
NMB = 0.00728
predicton RF model (µg/m³)
measurements (µg/m³)
R401 - NOx - validation 1/1/2015 - 29/02/2020
R² = 0.7315
RMSE = 11
RRMSE = 0.273
BIAS = -0.714
NMB = -0.0178
predicton RF model (µg/m³)
measurements (µg/m³)
R401 - NO2 - validation 1/1/2015 - 29/02/2020
R² = 0.8061
RMSE = 3.86
RRMSE = 0.16
BIAS = 0.0713
NMB = 0.00296
predicton RF model (µg/m³)
measurements (µg/m³)
R401 - PM2.5 - validation 1/1/2015 - 29/02/2020
R² = 0.7296
RMSE = 3.58
RRMSE = 0.31
BIAS = 0.0383
NMB = 0.00332
predicton RF model (µg/m³)
measurements (µg/m³)
R401 - PM10 - validation 1/1/2015 - 29/02/2020
R² = 0.6993
RMSE = 4.95
RRMSE = 0.269
BIAS = 0.284
NMB = 0.0155
predicton RF model (µg/m³)
measurements (µg/m³)



4. Impact of the corona lockdown measures on air quality

Now that it is clear that with the RF model the concentrations can be reliably estimated based on a training set of meteorological and time-bound parameters, we can also use the method to estimate the individual effect of the corona measures. To this end, an RF model was set up per monitoring station and per pollutant with the same historical (daily mean) pollutant concentrations and meteorological measurements as in Chapter 3. The model was trained with data from 1/1/2015 to 29/2/2020 and applied to the period from 1/3/2020 to 10/5/2020. Because the period after the corona lockdown was not included for training the model, the RF model cannot take into account the impact of the corona measures on the pollutant concentrations. The concentrations calculated by the RF model after entering the lockdown are the concentrations as estimated by the model based only on the weather conditions and time-related variables such as long-term trend, seasonal trend and weekly trend in this period. Or in other words, the concentrations according to a Business As Usual scenario ("BAU"), if there would have been no corona crisis and lockdown.

The difference between the concentrations calculated by the RF model and the concentrations actually measured at a measurement location are then a measure of the impact of the corona lockdown measures (including significantly less motorised traffic) on the air quality during the lockdown period.

The RF model was set up for the same pollutants and measuring stations as used in the validation of the model (see 3).

In the period before the lockdown, there were a number of periods (February) with stormy weather and associated high wind speeds that resulted in favourable dilution conditions and little air pollution. The fact that the weather conditions became more adverse after the start of the lockdown is shown in Figure 3. This figure shows the weekly average "ventilation factor" for Antwerp, Brussels, Charleroi and Liège. The ventilation factor is the product of the wind speed and the mixing layer height.

The higher the ventilation factor, the better the dispersion or dilution conditions in the atmosphere. Equal emissions will result in lower concentrations at a high ventilation factor. If the ventilation factor is low, pollution accumulates and the concentrations are higher. This is also clearly visible on Figure 3: during week 4 in 2020 the average ventilation factor was very low. This resulted in very high concentrations (see later). The fact that the meteorological conditions became less favourable for air quality (less dispersion) after entering the lockdown is an element that must be taken into account in what follows.

Figure 3: weekly average ventilation factor (wind speed x mixing layer height) for the measurement locations Antwerp (T2M802), Brussels (T1M003), Charleroi (T5M501) and Liège (T3M202) during week 1 to week 19 in 2020. The vertical blue dotted line indicates when the (soft) lockdown period was introduced in Belgium (19 March 2020).

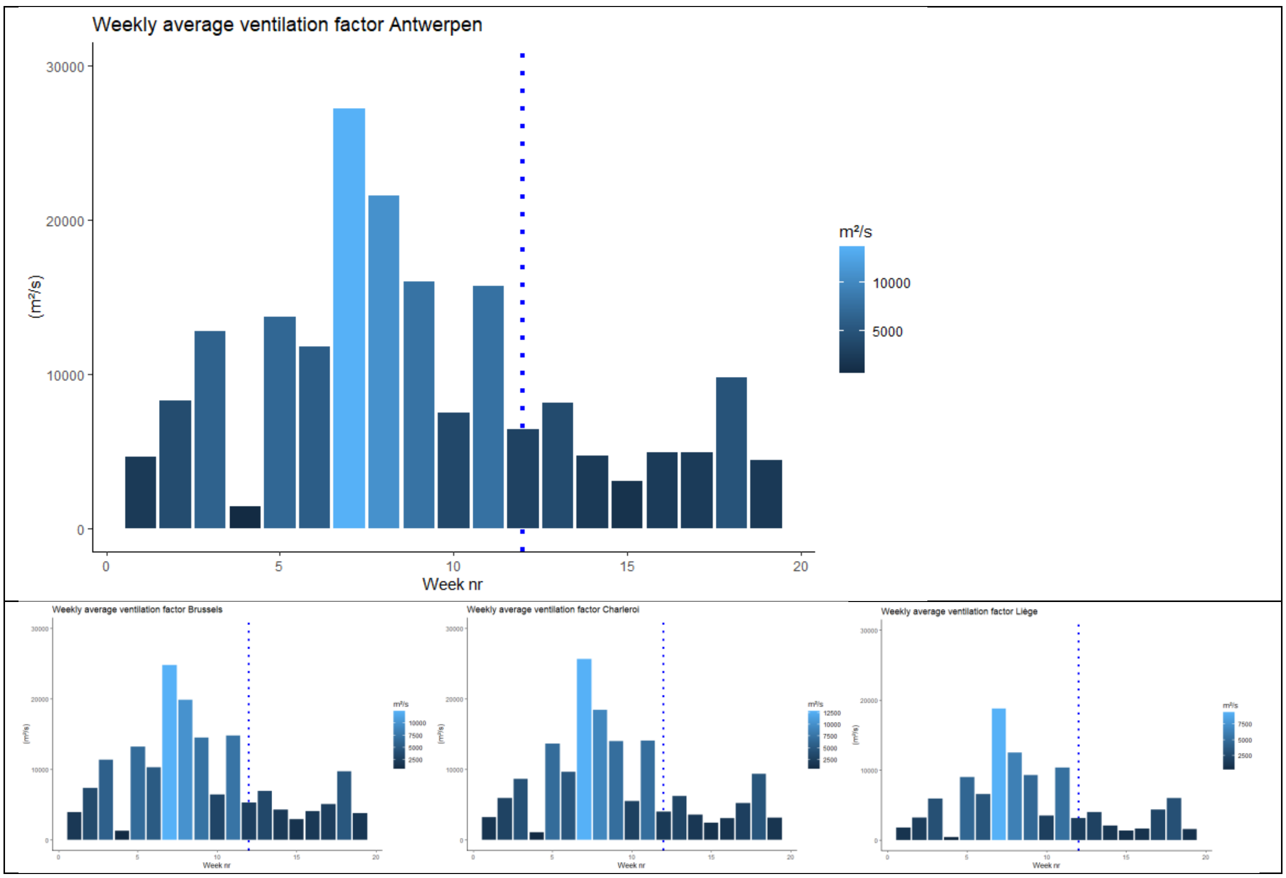
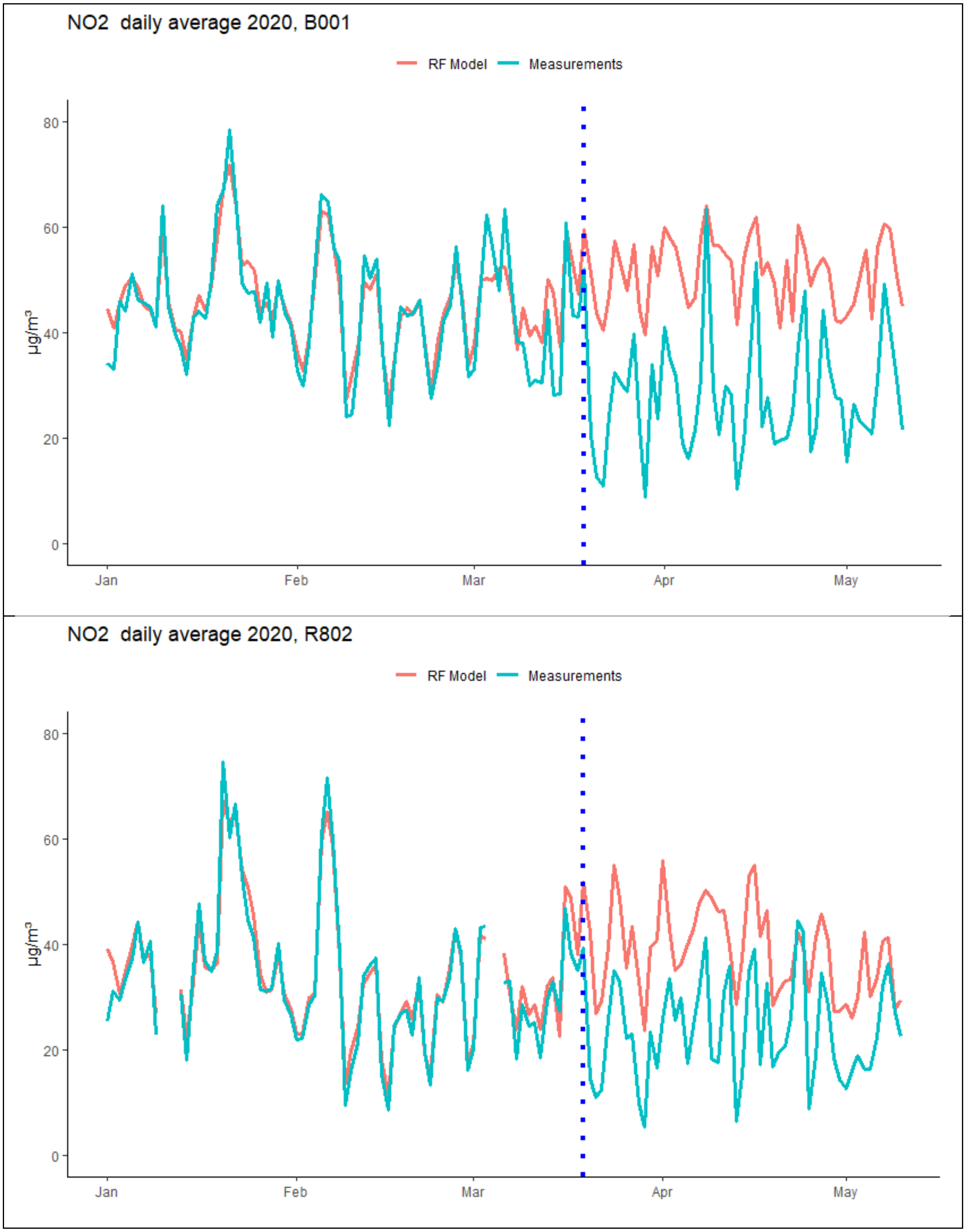


Figure 4 shows (as an example) the measured and modelled daily mean NO_2 concentrations in the period 1/1/2020 - 10/5/2020 for the street stations 41B001 (Brussels, Kunst-Wet) and 42R802 (Antwerp-Borgerhout) and the background stations 44N029 (Veurne close to the coast) and 43N100 (Dourbes in the Ardennes)

Figure 4: measured (blue) and modelled (red) daily mean NO_2 concentrations in the urban street stations in Brussels (41B001) and Antwerp (42R802) and the rural background stations in Veurne (44N029) and Dourbes (43N100). The vertical blue dotted line indicates when the (soft) lockdown period was introduced in Belgium (19 March 2020).



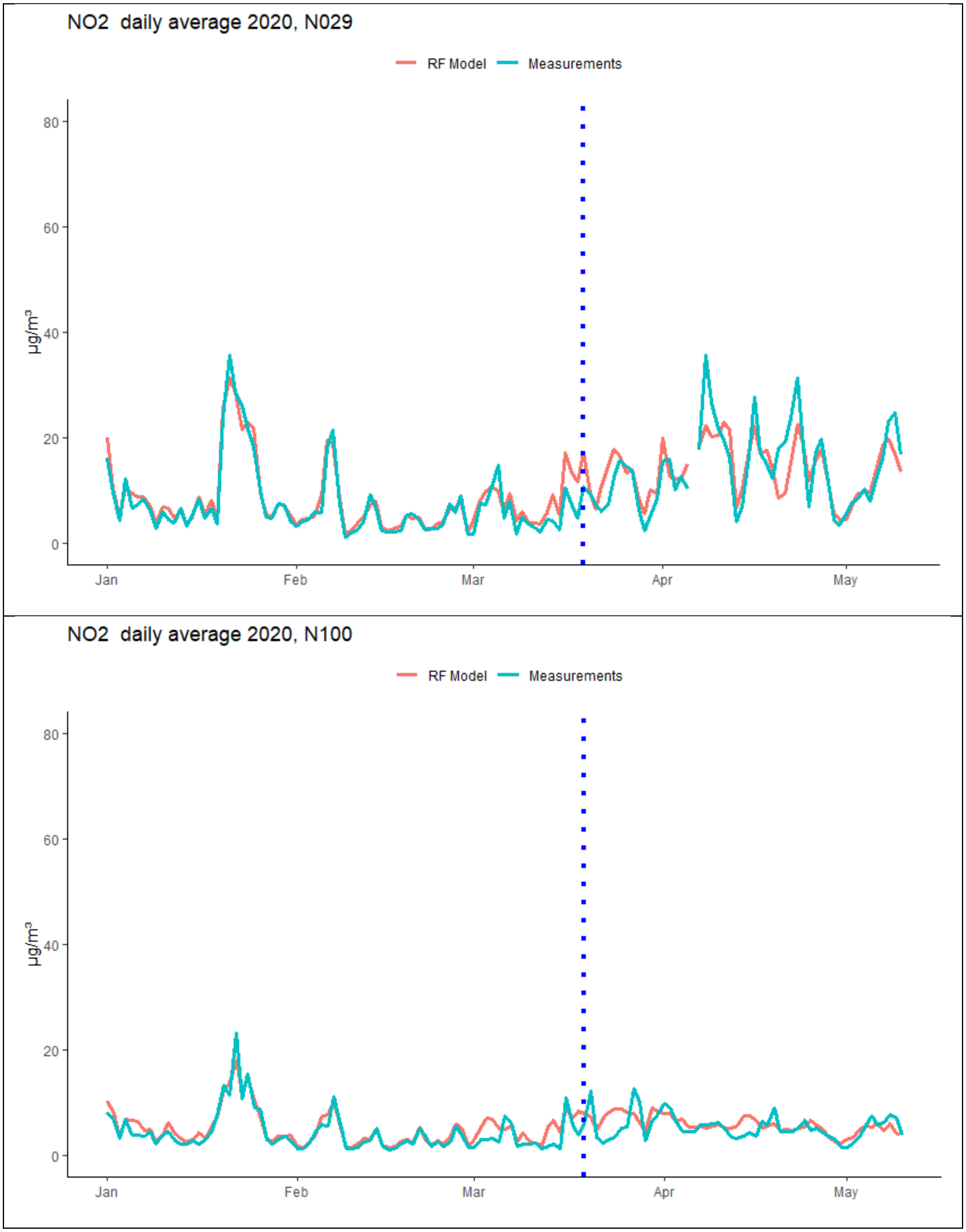


Figure 4 clearly shows that the RF model significantly "overestimates" the NO_2 concentrations in the urban street stations in Antwerp and Brussels on almost every day after the lockdown has started. Based on the weather conditions and the time-bound indicators, the concentrations in a BAU scenario or a scenario “if no lockdown had taken place” would have been significantly higher. The difference between the RF model results (red line) and the levels as measured in practice (blue line) can be considered as the impact of the corona measures on the NO_2 concentrations in the street stations in Antwerp and Brussels. NO_2 is a typical traffic-related pollutant. The reduction in car traffic causes a decrease in NO_2 concentrations at a measurement location that is strongly influenced by emissions from (local) road traffic.

In the background stations in Veurne (coast) and Dourbes (Ardennes) there is less difference between the measured and the calculated NO_2 concentrations. This is an indication that less traffic during the corona lockdown had less impact on the NO_2 concentrations in the monitoring stations that are more remote from motorised traffic sources.

The RF model calculates daily mean concentrations. In the following sections, we use weekly averages to discuss the impact of the corona measures in more detail and per pollutant. For this purpose, the measured and the daily mean concentrations calculated by the model were averaged for week 1 to week 19 in 2020.

4.1 Nitrogen oxides (NO_x)

NO_x (or the molar sum of nitrogen dioxide or NO_2 and nitrogen monoxide or NO) is the form in which nitrogen oxides are emitted by emission sources such as road traffic. When NO_x is emitted by these sources, it is largely nitrogen monoxide (NO). In street stations, motorised traffic in the immediate vicinity is, in particular, the main source of NO_x.

Usually, no separate analysis is made in air reports for NO_x in the open air. For this model exercise, NO_x is used as a separate air pollutant because the amount of NO_x in the air is more directly linked to the emissions (and the amount of) car traffic compared to nitrogen dioxide (NO_2). NO_2 is also directly emitted to a limited extent, but is mainly a secondary component that is formed in the troposphere[4] (among other things from very fast reactions between NO and ozone). The amount of NO_2 in the air, also in street stations, is not only determined by local traffic but also by more distant sources. NO_x is therefore one of the pollutants on which we expect the biggest effect from the lockdown, which resulted in a sharp drop in motorised traffic volume.

Figure 5 shows the weekly mean measured concentration (blue) and the difference between the measured by the RF model and the measured weekly mean NO_x concentration (red) for 4 types of measurement locations and two stations each for the first 19 weeks in 2020. The corona lockdown started in week 12.

It is striking that the measured weekly average NO_x concentration at the urban street monitoring stations is lower from week 12 (after the start of the lockdown) than the weeks before, despite significantly less favourable weather conditions for air quality (see Figure 3 and corresponding explanation above). The red bars, the difference between the expected (BAU) and the measured concentrations, however, indicate that based on this less favourable weather and the time-related

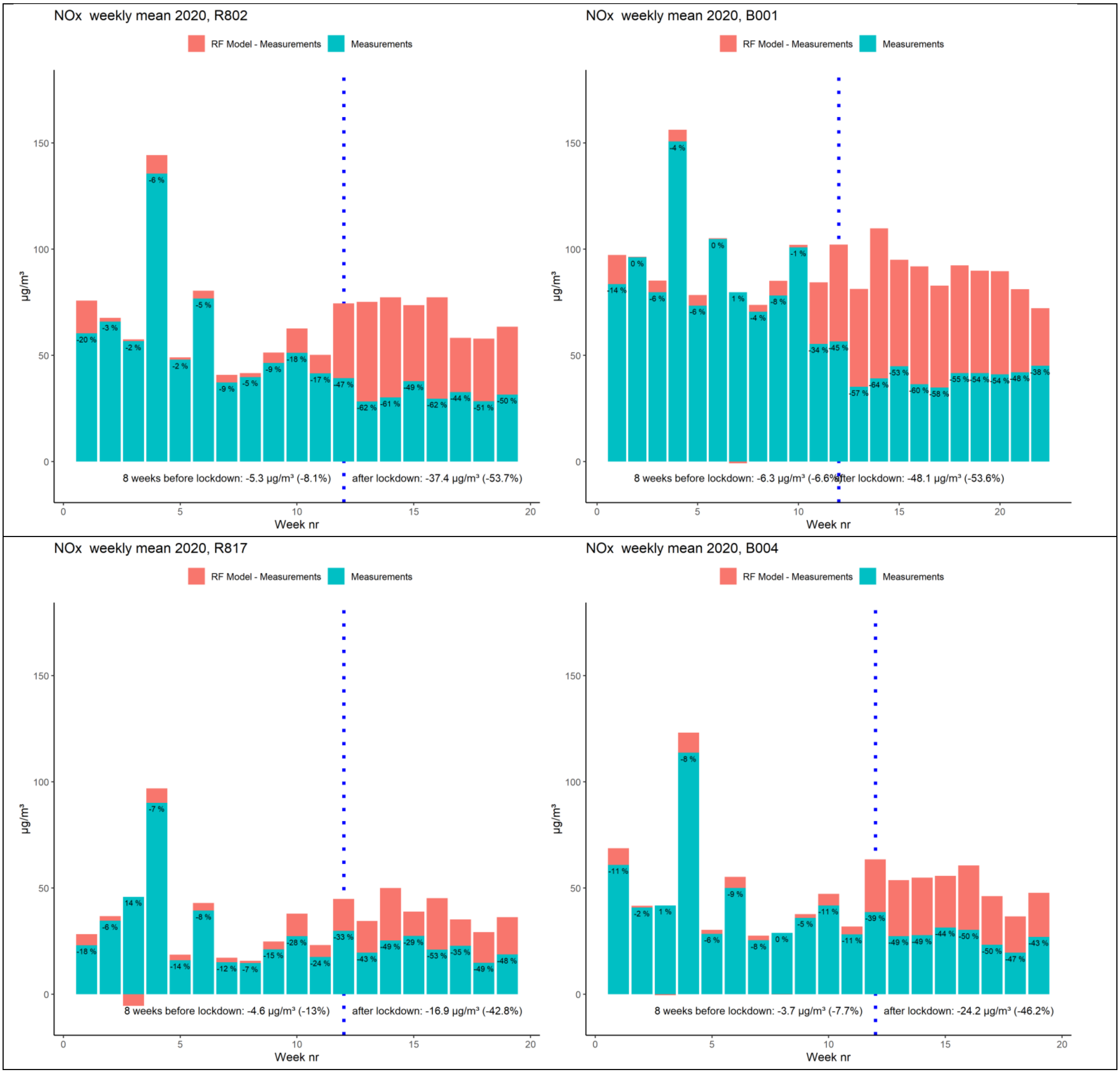
[4] the lowest layer of Earth's atmosphere

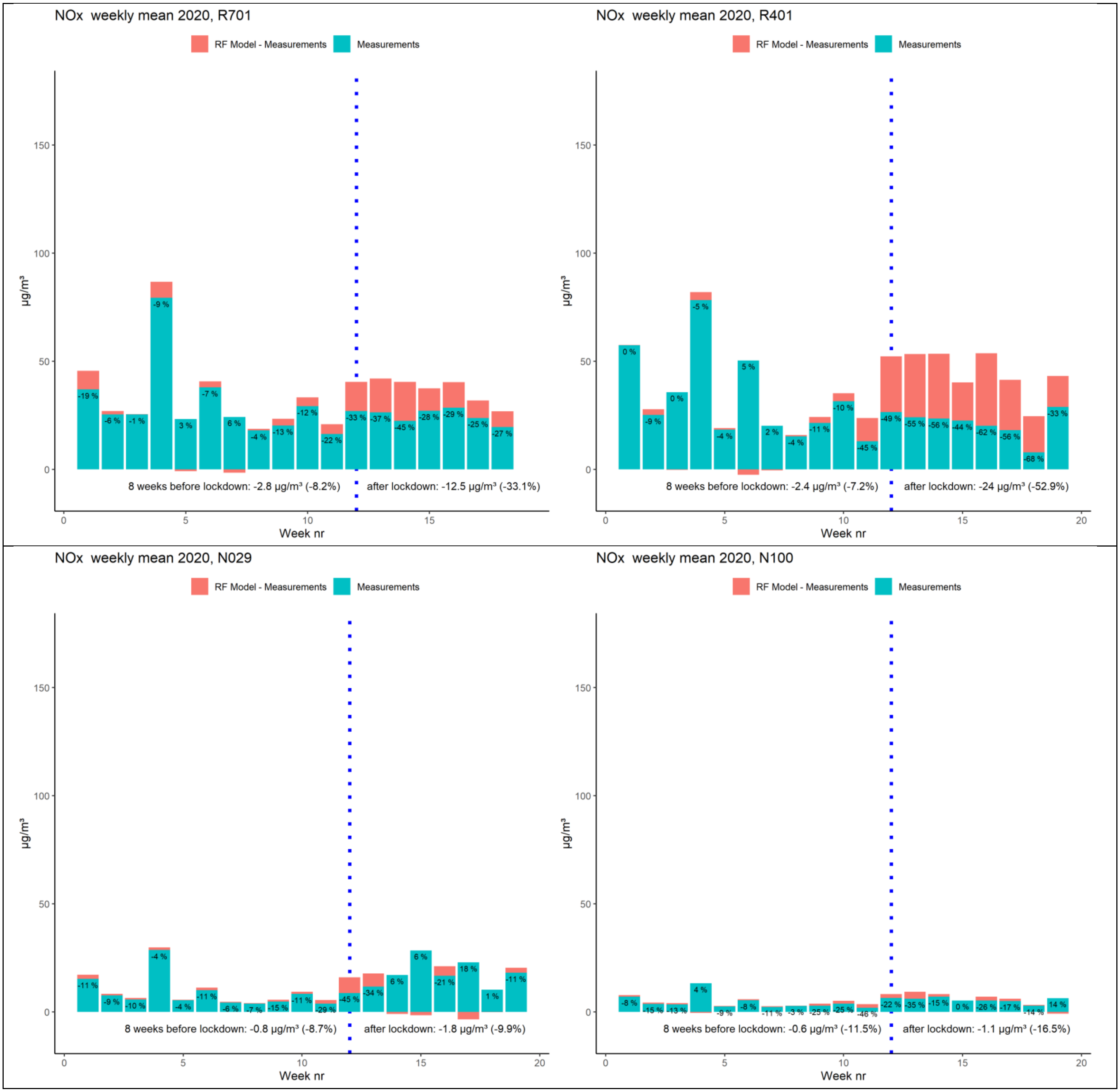
variables, the RF model calculates higher NO_x concentrations than what appears from the measurements. This effect is greatest for the urban street monitoring stations and is slightly less pronounced for the urban background and suburban/inner-city stations and much less for the background stations on the coast and in the Ardennes. The size of the red bar is a measure of the impact of the significantly lower traffic volumes after entering the lockdown, resulting in less NO_x emissions and therefore lower NO_x concentrations.

The (relative) difference is shown by the blue bars per week. The (relative) mean difference for the 8 weeks after the lockdown is indicated at the bottom of the figure. The difference in the 8 weeks before the lockdown is also shown. The fact that the reduction percentage in the weeks before the lockdown is not 0 indicates that the model cannot perfectly predict the concentrations based on meteorological parameters only (see 3).

The decrease (or increase) in the concentrations for the other measuring stations are shown in Table 4 (absolute figures) and Table 5 (relative or percentage figures).

Figure 5: measured weekly mean (blue) and the difference between the expected (RF model) and the measured (red) NO_x concentration for the first 19 weeks in 2020 in 4 types of monitoring stations. From top to bottom: urban street, suburban/inner city, urban background and rural background. The dotted blue line indicates week 12, the start of the corona lockdown measures. The percentages in the blue bars indicate the decrease (or increase) of the weekly mean NO_x concentration compared to the (expected) concentrations calculated by the RF model.




NOx weekly mean 2020, R701
RF Model - Measurements
Measurements
µg/m³
8 weeks before lockdown: -2.8 µg/m³ (-8.2%)
after lockdown: -12.5 µg/m³ (-33.1%)
Week nr
NOx weekly mean 2020, R401
RF Model - Measurements
Measurements
µg/m³
8 weeks before lockdown: -2.4 µg/m³ (-7.2%)
after lockdown: -24 µg/m³ (-52.9%)
Week nr
NOx weekly mean 2020, N029
RF Model - Measurements
Measurements
µg/m³
8 weeks before lockdown: -0.8 µg/m³ (-8.7%)
after lockdown: -1.8 µg/m³ (-9.9%)
Week nr
NOx weekly mean 2020, N100
RF Model - Measurements
Measurements
µg/m³
8 weeks before lockdown: -0.6 µg/m³ (-11.5%)
after lockdown: -1.1 µg/m³ (-16.5%)
Week nr

4.2 Nitrogen dioxide (NO_2)

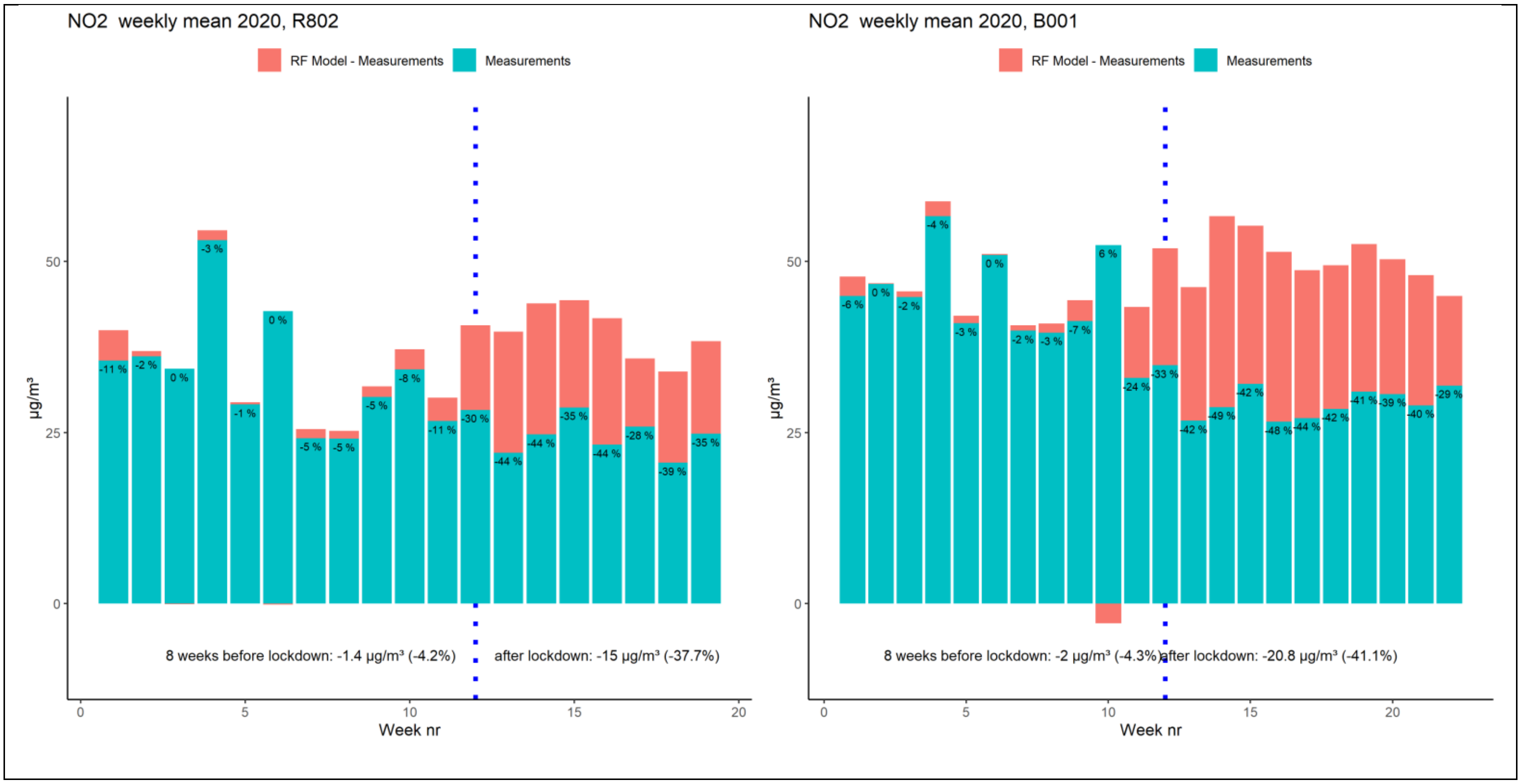
Nitrogen dioxide (NO_2) is a good indicator of traffic-related air pollution in urban environments or places with a lot of traffic, just like NO_x. As mentioned in 4.1 Nitrogen oxides, NO_2 is mainly a secondary (not directly emitted) pollutant and is therefore somewhat less directly related to the traffic emissions of nitrogen oxides. NO_2 is the harmful component of the NO_x mixture and has a direct impact on health. That is why there are European limit values for NO_2 and World Health Organization (WHO) recommended values.

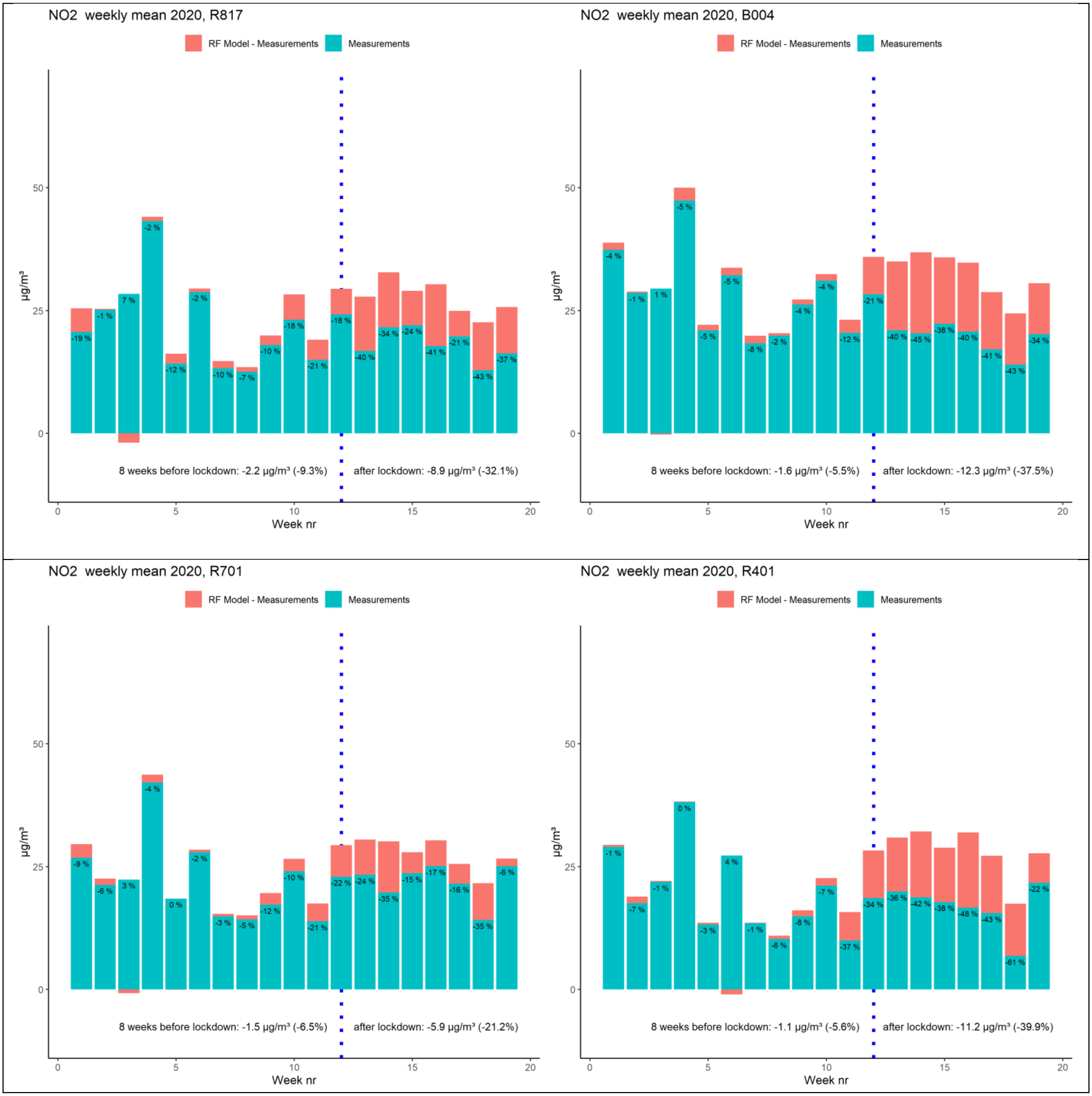
Figure 6 shows the weekly mean NO_2 concentrations (both actually measured and calculated by the RF model) for the same measurement locations as in 4.1 Nitrogen oxides.

The conclusions for NO_2 are similar to those for NO_x: the strongest decrease is in the stations that are most affected by the local emissions of motorised traffic. The decrease in NO_2 concentrations due to the corona lockdown is somewhat less pronounced than for NO_x. More than NO_x, the NO_2 concentrations are also determined by the background concentrations, i.e. by sources further than in the immediate vicinity of the measuring stations. Other sources than motorised traffic (such as industry) that are less affected by the corona measures also contribute to these background concentrations.

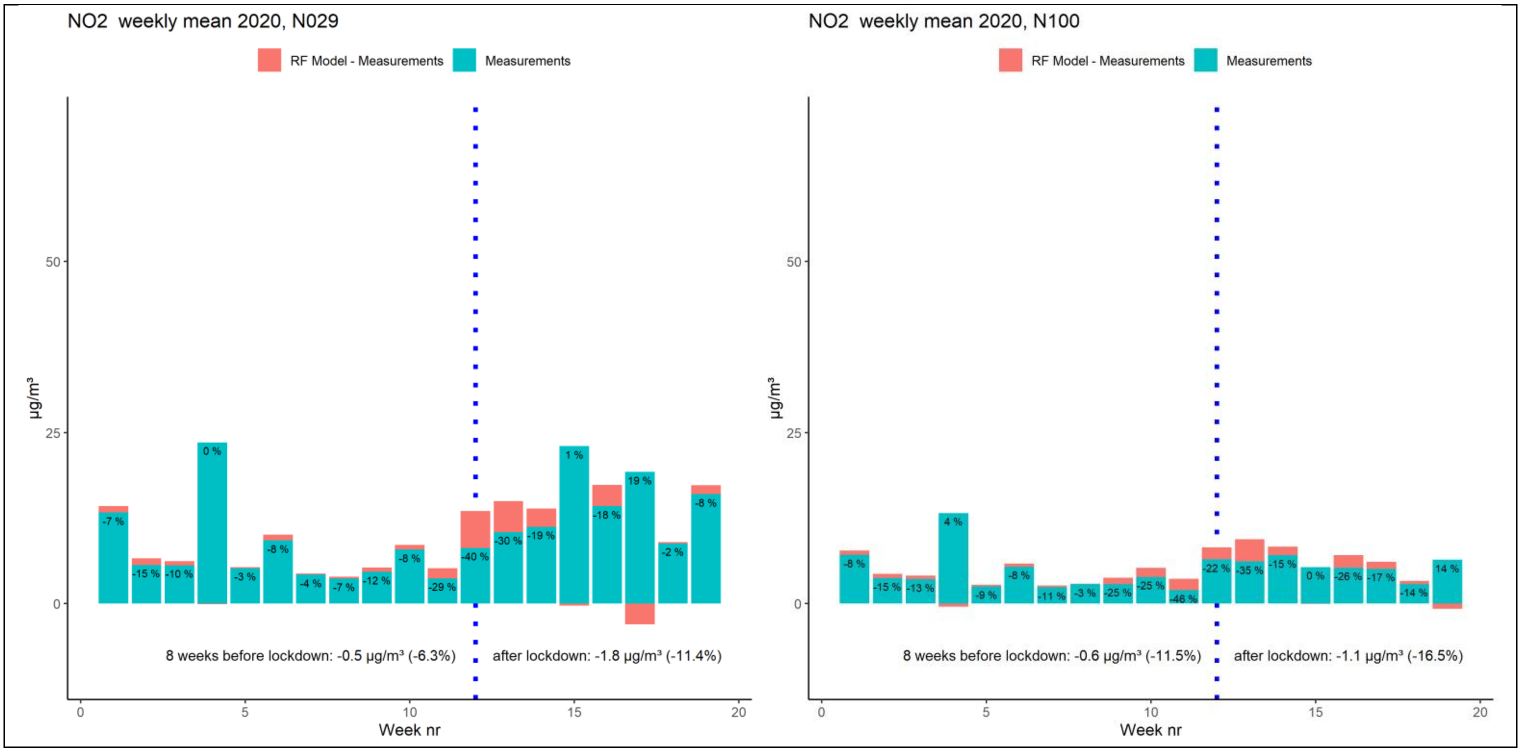
The impact of the corona measures on the NO_2 concentrations for the other stations can be found in Table 4 (absolute difference) and Table 5 (relative difference).

Figure 6: measured weekly mean (blue) and the difference between the expected (RF model) and the measured (red) NO_2 concentration for the first 19 weeks in 2020 in 4 types of monitoring stations. From top to bottom: urban street, suburban/inner city, urban background and rural background. The dotted blue line indicates week 12, the start of the corona lockdown measures. The percentages in the blue bars indicate the decrease (or increase) of the weekly mean NO_2 concentration compared to the (expected) concentrations calculated by the RF model.





NO2 weekly mean 2020, R817
RF Model - Measurements
Measurements
8 weeks before lockdown: -2.2 µg/m³ (-9.3%)
after lockdown: -8.9 µg/m³ (-32.1%)
NO2 weekly mean 2020, B004
8 weeks before lockdown: -1.6 µg/m³ (-5.5%)
after lockdown: -12.3 µg/m³ (-37.5%)
NO2 weekly mean 2020, R701
8 weeks before lockdown: -1.5 µg/m³ (-6.5%)
after lockdown: -5.9 µg/m³ (-21.2%)
NO2 weekly mean 2020, R401
8 weeks before lockdown: -1.1 µg/m³ (-5.6%)
after lockdown: -11.2 µg/m³ (-39.9%)
µg/m³
Week nr



4.3 Particulate matter (PM2.5)

Figure 7 shows the weekly mean PM2.5 concentrations (both actually measured and calculated by the RF model) for the same measurement locations as in 4.1 Nitrogen oxides and 4.2 Nitrogen dioxide (NO_2) (except for urban street station Kunst-Wet-41B001, which is replaced by urban street station Molenbeek-41R001 and inner-city station 41B004 to be replaced by Liège urban background station-43R222).

For PM2.5, the smaller fraction of particulate matter with a diameter smaller than 2.5 micrometre, there is a different development than for NO_x and NO_2: after the start of the corona lockdown, there is an increase in the measured weekly average concentrations. A significant decrease in the traffic volume therefore did not result in a measurable decrease in particulate matter concentrations.

According to the RF model, the concentrations would have been higher had there not been a lockdown, but this is not the case for all weeks.

As stated in chapter 3, the RF model is less capable to predict particulate matter concentrations than the other pollutants. Moreover, the model underestimates the highest concentrations. These high concentrations occur (especially) on days with unfavourable weather conditions. It cannot therefore be ruled out that the model underestimates the concentrations after entering the lockdown due to the more unfavourable weather conditions in this period. The effect of the corona measures on the particulate matter concentrations may therefore be higher than what the RF model estimates.

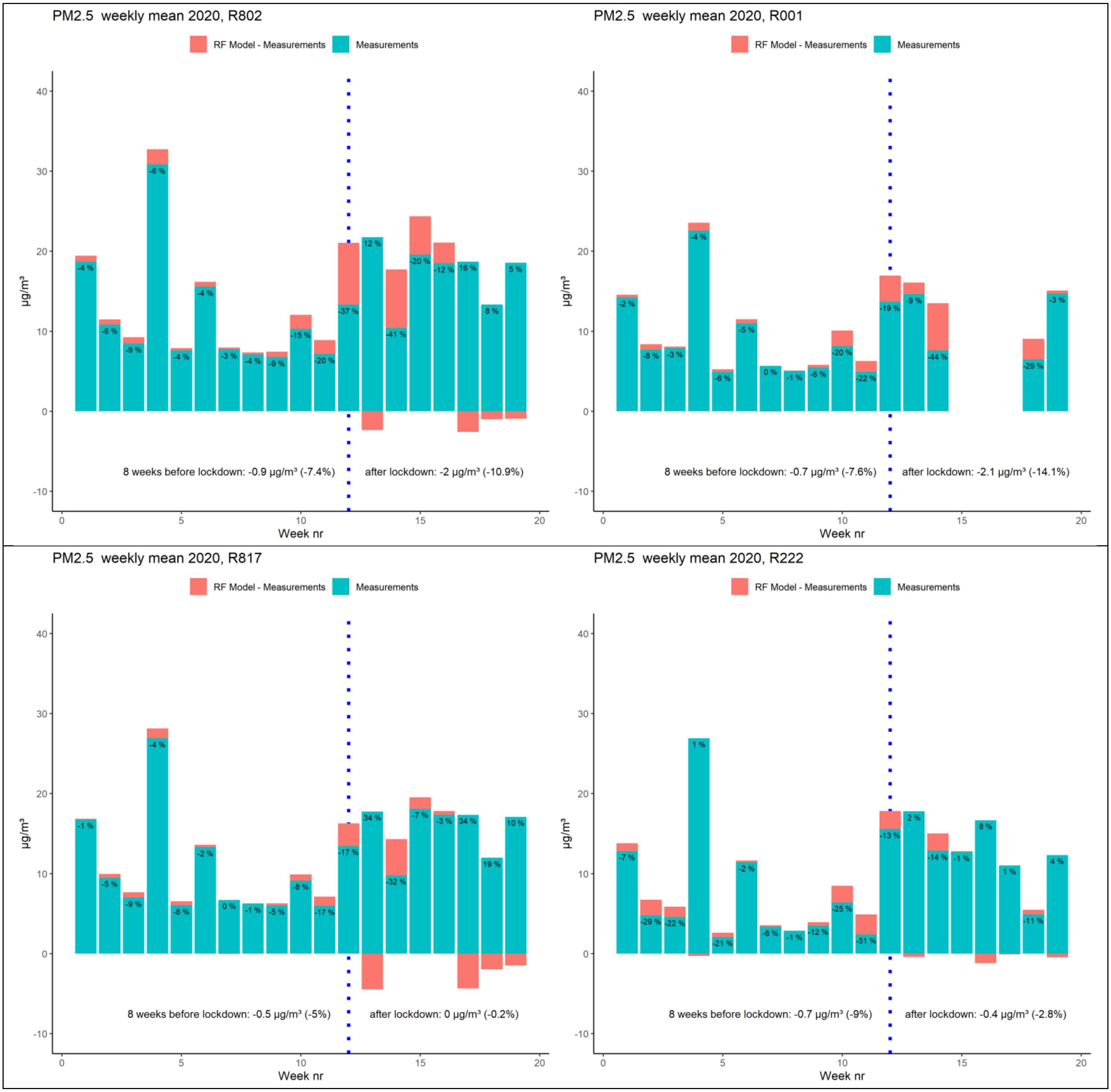
High particulate matter concentrations also occur when polluted air is supplied via continental air currents. The RF model does not yet contain any parameters associated with this. The wind direction at 30 metre, which is included, does not give a full picture of the origin of the air mass. The model will therefore not be able to predict these types of situations properly.

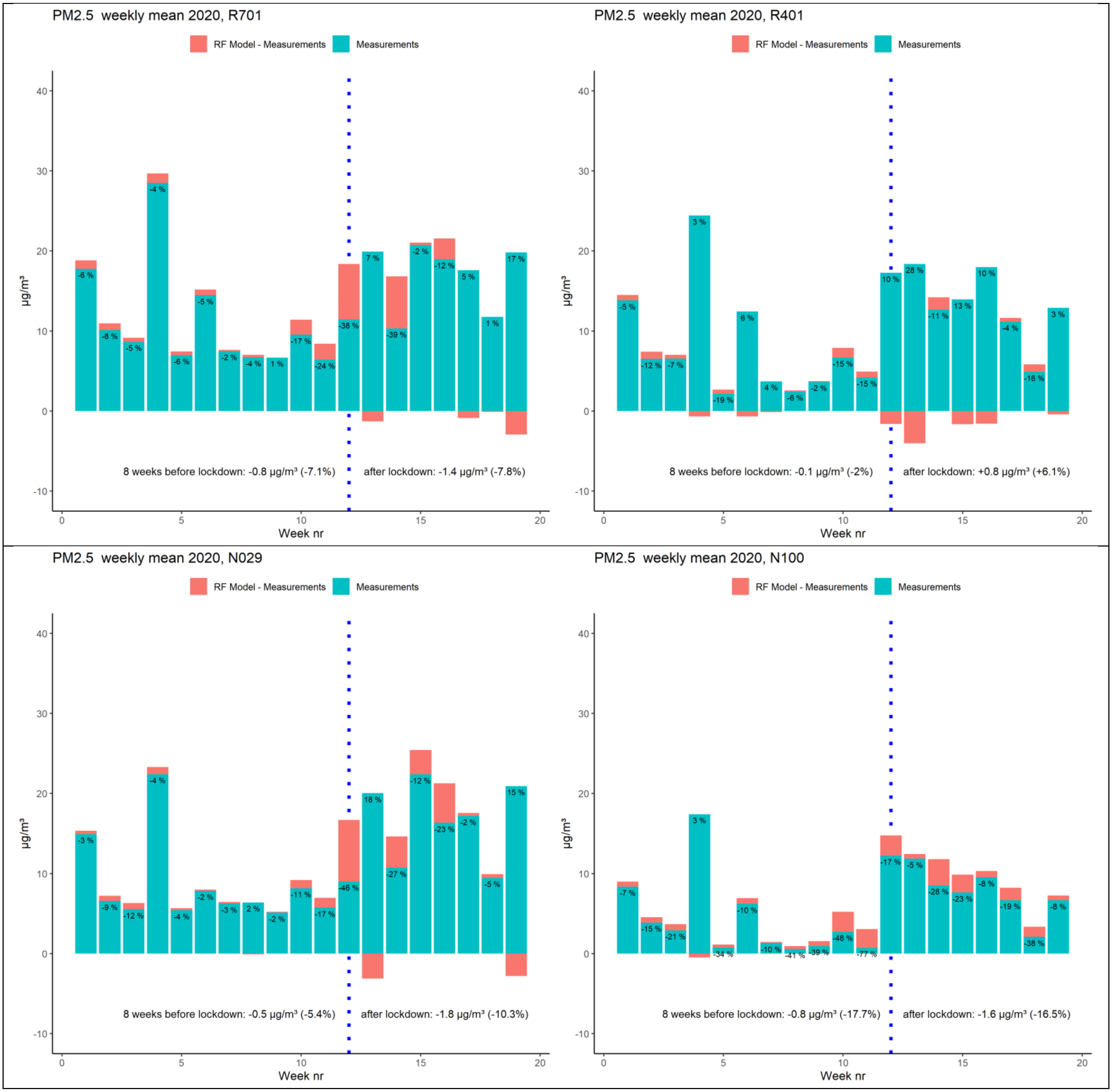
It is clear, however, that the corona measures that cause a significant decrease in the traffic volume and the associated significant decrease in NO_x and NO_2 concentrations, do not cause a comparable decrease in particulate matter concentrations.

This is not entirely unexpected as PM2.5 is not a good traffic-related indicator. After all, the contribution of the “primary” or direct emitted particulate matter to the total mass of PM2.5 is limited. Particulate matter also has many more sources than just motorised road traffic. Industry, households (especially wood combustion) and agriculture are also important sources of particulate matter. In the months of March and April, the ammonia emissions from agriculture (fertilizer application) in unfavourable weather conditions also lead to the formation of secondary ammonium salts (spring smog), which are a component of the “secondary” particulate matter. Secondary particulate matter is particulate matter that does not enter the air directly but is created by physico-chemical processes from gases (ammonia, nitrogen oxides, sulphur dioxide, organic gases). The main component of secondary (inorganic) particulate matter is ammonium nitrate. This is formed when nitrogen oxides (NO_x) react with ammonia. Despite the significant decrease in the amount of NO_x, enough NO_x remained in the air (caused by the remaining traffic and other sources) to form secondary particulate matter with ammonia.

Less motorised traffic therefore only has a limited impact on the total amount of PM2.5 in the air. After the introduction of the corona lockdown, the unfavourable weather conditions resulted in higher PM2.5 concentrations than before.

Figure 7: measured weekly mean (blue) and the difference between the expected (RF model) and the measured (red) particulate matter (PM2.5) concentration for the first 19 weeks in 2020. The blue dotted line indicates week 12, the start of the corona lockdown measures. The percentages in the blue bars indicate the decrease (or increase) of the weekly mean PM2.5 concentration compared to the (expected) concentrations calculated by the RF model.

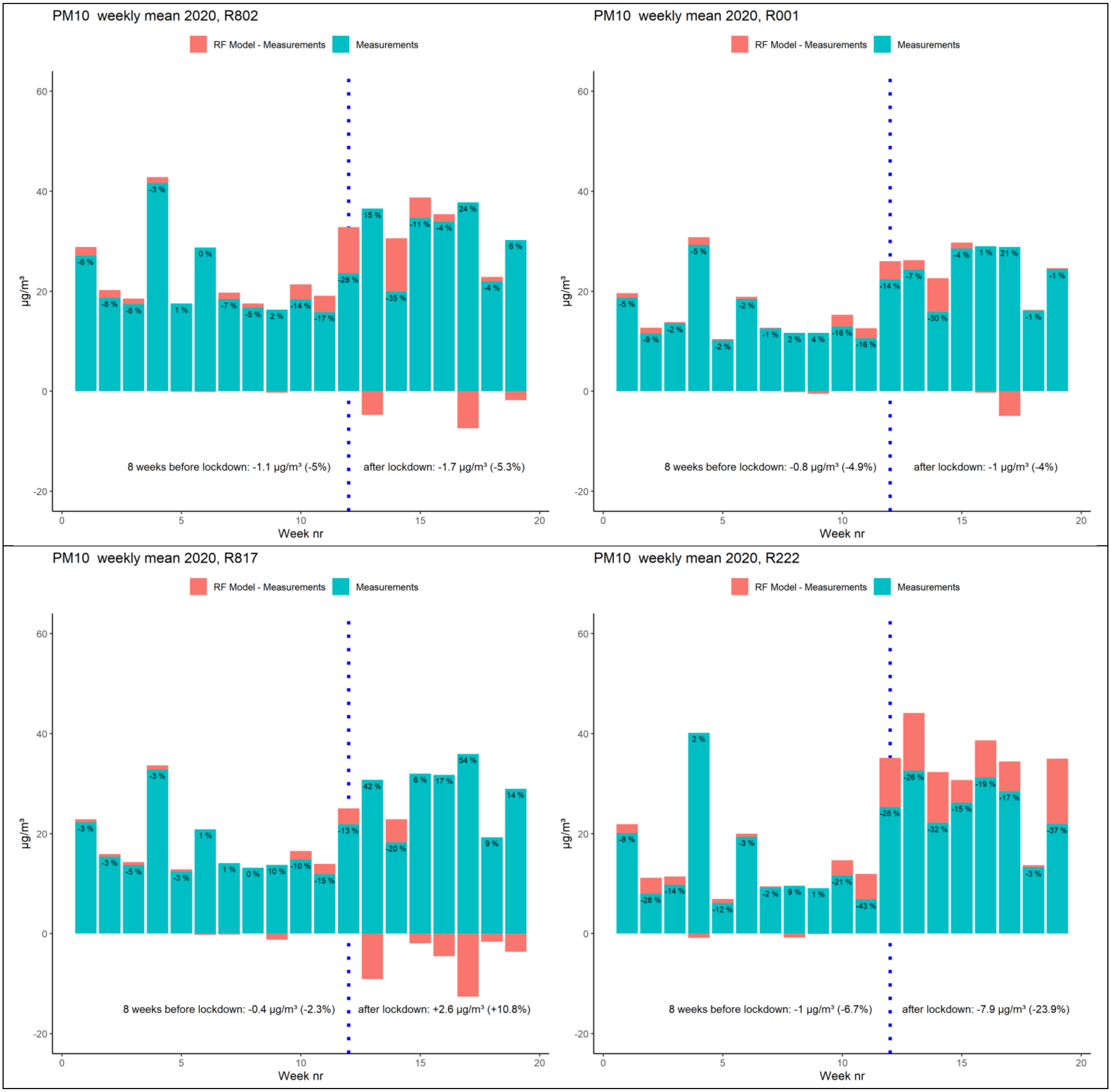


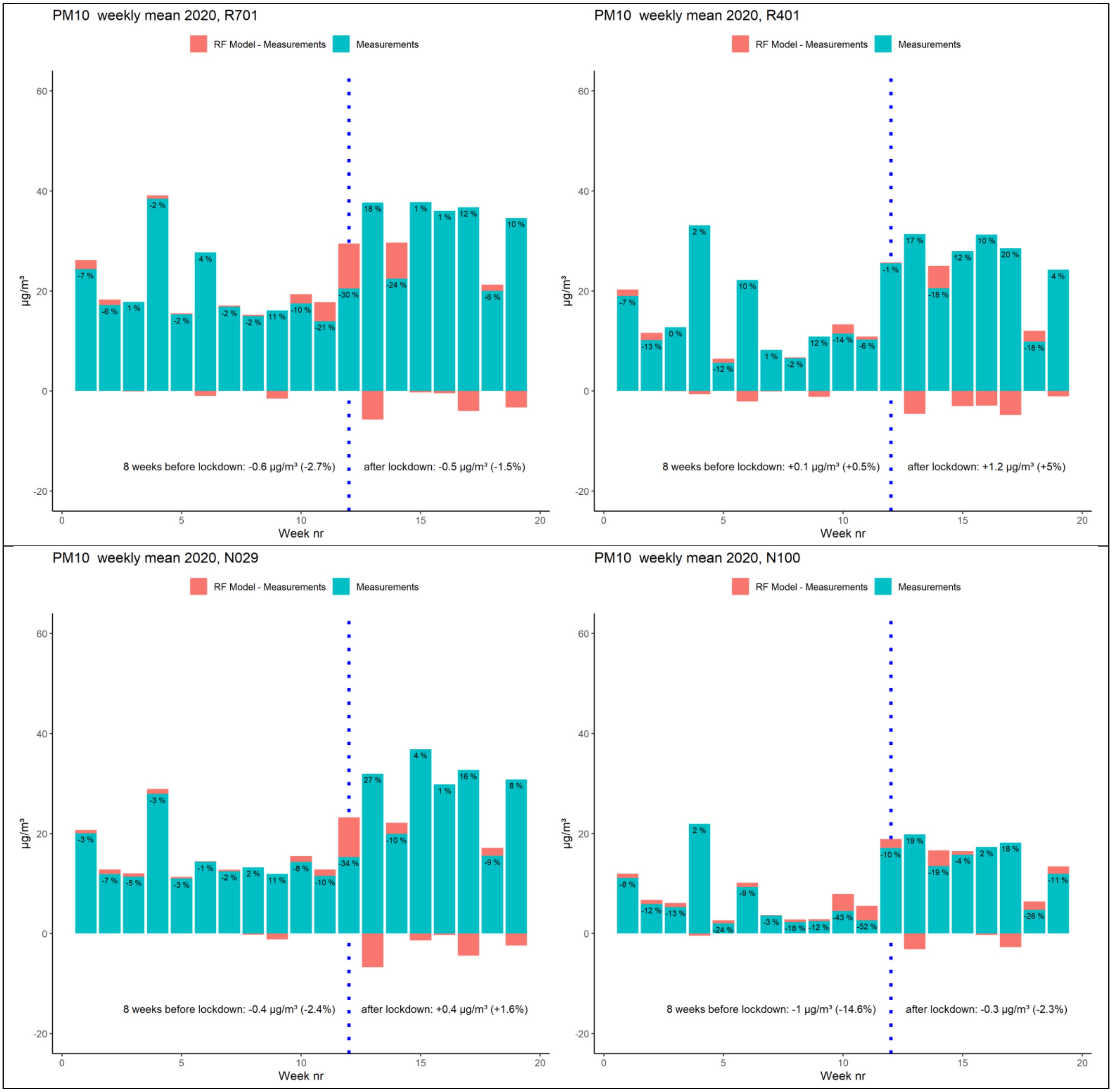

PM2.5 weekly mean 2020, R701
RF Model - Measurements
Measurements
µg/m³
8 weeks before lockdown: -0.8 µg/m³ (-7.1%)
after lockdown: -1.4 µg/m³ (-7.8%)
Week nr
PM2.5 weekly mean 2020, R401
RF Model - Measurements
Measurements
µg/m³
8 weeks before lockdown: -0.1 µg/m³ (-2%)
after lockdown: +0.8 µg/m³ (+6.1%)
Week nr
PM2.5 weekly mean 2020, N029
RF Model - Measurements
Measurements
µg/m³
8 weeks before lockdown: -0.5 µg/m³ (-5.4%)
after lockdown: -1.8 µg/m³ (-10.3%)
Week nr
PM2.5 weekly mean 2020, N100
RF Model - Measurements
Measurements
µg/m³
8 weeks before lockdown: -0.8 µg/m³ (-17.7%)
after lockdown: -1.6 µg/m³ (-16.5%)
Week nr

4.4 Particulate matter (PM10)

Figure 8 shows the weekly mean PM10 concentrations (both actually measured and calculated by the RF model) for the same measurement locations as in 4.3 Particulate matter (PM2.5). As for PM2.5, the PM10 (particulate matter with a diameter of less than 10 micrometres) increases during the lockdown period compared to the previous weeks. The explanation is largely the same: unfavourable weather conditions favouring the formation of spring smog (secondary fine dust). Just as for PM2.5, it is important to take into account the observation (see 3) that the RF model is less capable of predicting the particulate matter concentrations than the other (more traffic-related) pollutants. The results of the model must therefore be interpreted with the necessary caution for particulate matter. On dry days with high wind speeds, the PM10 concentrations can increase significantly. This blowing up of soil dust, which is mainly present in the larger particulate matter fraction, does not occur very often, hence the model cannot properly estimate this phenomenon.

Figure 8: measured weekly mean (blue) and the difference between the expected (RF model) and the measured (red) particulate matter (PM10) concentration for the first 19 weeks in 2020. The blue dotted line indicates week 12, the start of the corona lockdown measures. The percentages in the blue bars indicate the decrease (or increase) of the weekly mean PM10 concentration compared to the (expected) concentrations calculated by the RF model.

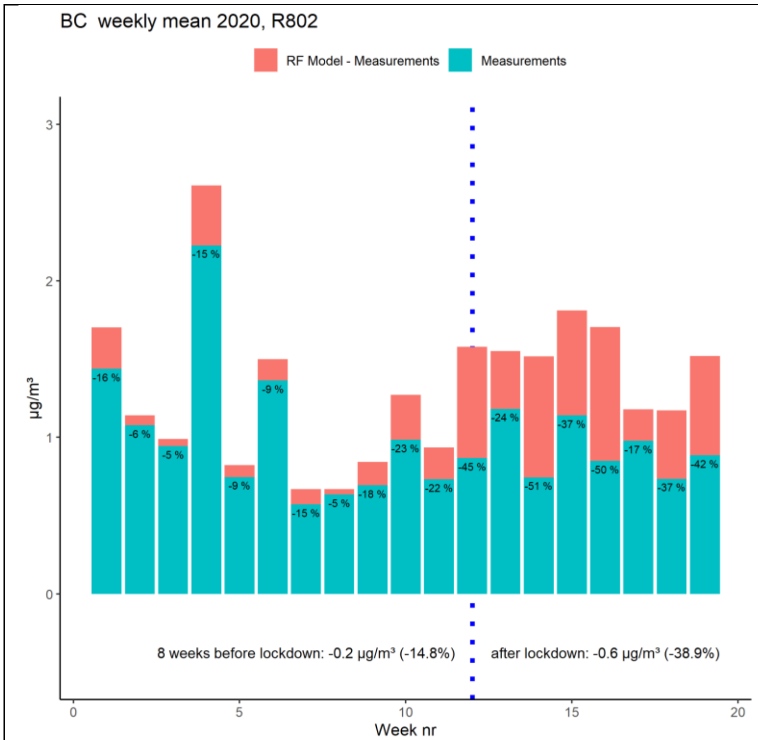


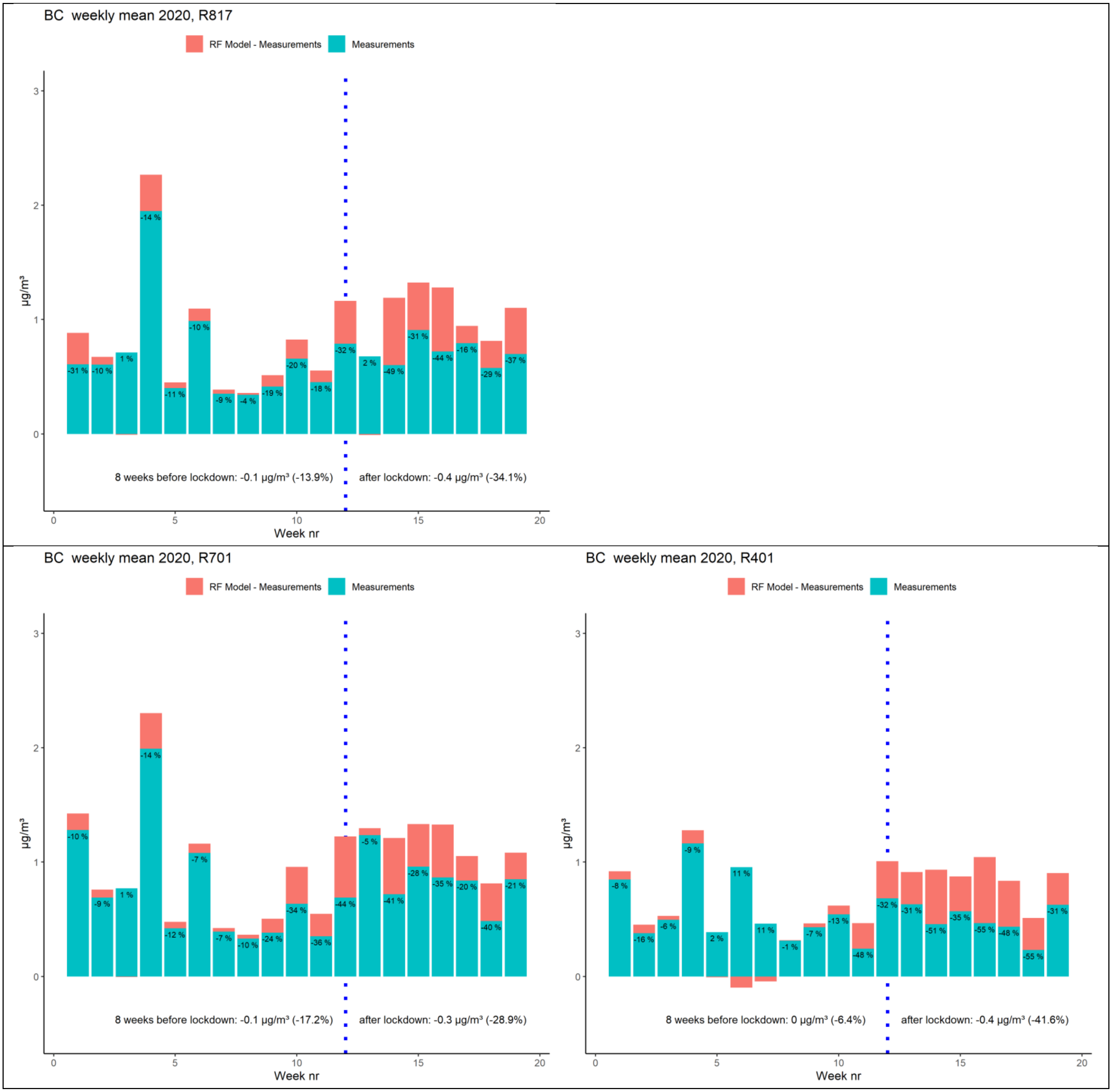

PM10 weekly mean 2020, R701
RF Model - Measurements
Measurements
µg/m³
Week nr
8 weeks before lockdown: -0.6 µg/m³ (-2.7%)
after lockdown: -0.5 µg/m³ (-1.5%)
-7 % -6 % 1 % -2 % -2 % 4 % -2 % -2 % 11 % -10 % -21 % -30 % 18 % -24 % 1 % 1 % 12 % -6 % 10 %
PM10 weekly mean 2020, R401
RF Model - Measurements
Measurements
µg/m³
Week nr
8 weeks before lockdown: +0.1 µg/m³ (+0.5%)
after lockdown: +1.2 µg/m³ (+5%)
-7 % -13 % 0 % 2 % -12 % 10 % 1 % -2 % 12 % -14 % -6 % -1 % 17 % -18 % 12 % 10 % 20 % -18 % 4 %
PM10 weekly mean 2020, N029
RF Model - Measurements
Measurements
µg/m³
Week nr
8 weeks before lockdown: -0.4 µg/m³ (-2.4%)
after lockdown: +0.4 µg/m³ (+1.6%)
-3 % -7 % -5 % -3 % -3 % -1 % -2 % 2 % 11 % -8 % -10 % -34 % 27 % -10 % 4 % 1 % 16 % -9 % 8 %
PM10 weekly mean 2020, N100
RF Model - Measurements
Measurements
µg/m³
Week nr
8 weeks before lockdown: -1 µg/m³ (-14.6%)
after lockdown: -0.3 µg/m³ (-2.3%)
-8 % -12 % -13 % 2 % -24 % -9 % -3 % -18 % -12 % -43 % -52 % -10 % 19 % -19 % -4 % 2 % 18 % -26 % -11 %

4.5. Black Carbon

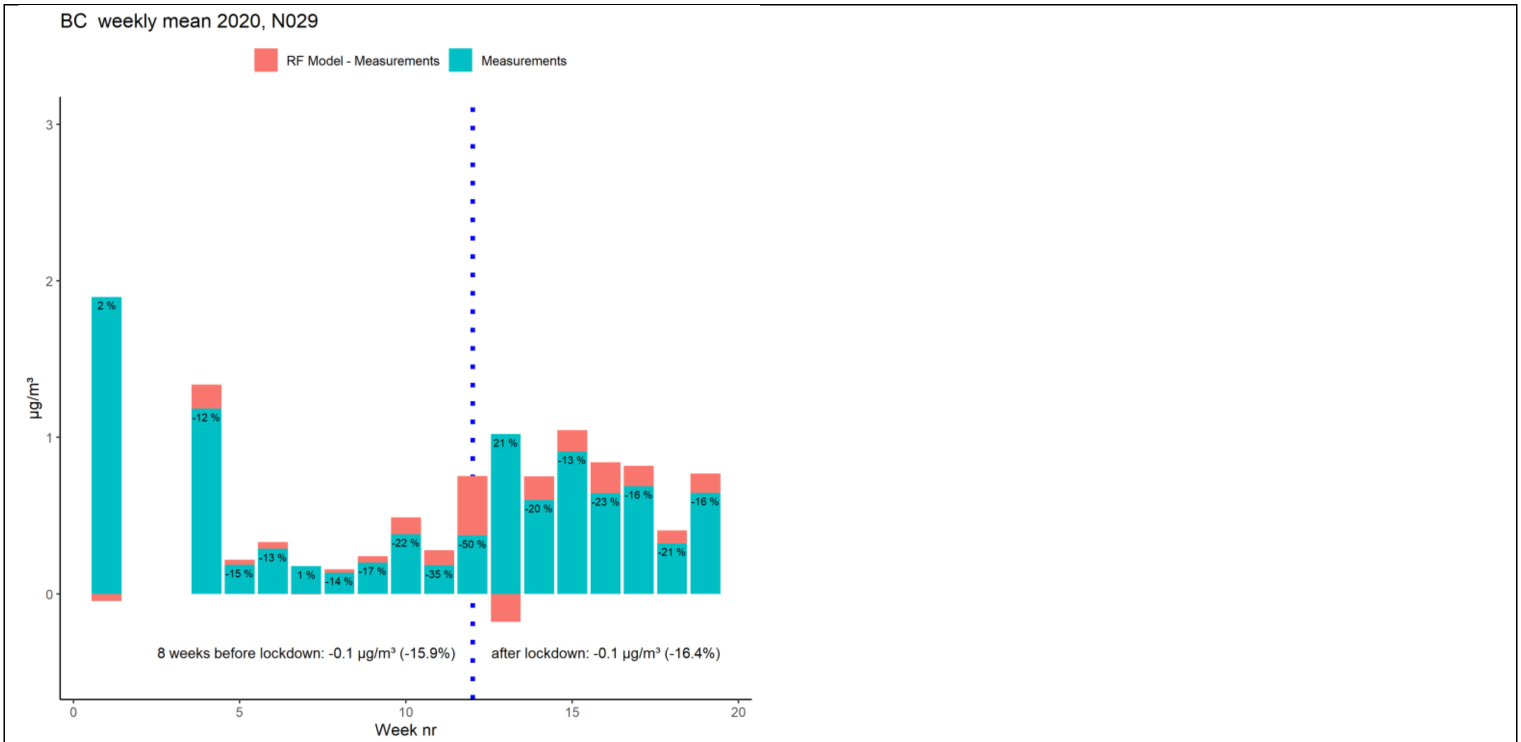
For black carbon (BC), a measure of (diesel) soot, the impact of the corona measures is very similar to that for nitrogen dioxide (NO_2). After all, BC and NO_2 are strongly correlated. Just as for NO_2 (and NO_x), the largest decreases are noted in the traffic-impacted monitoring stations. The influence is lower on urban background and suburban measurement locations. The impact is lowest at background measurement locations further away from traffic sources (see Figure 9).

Figure 9: measured weekly mean (blue) and the difference between the expected (RF model) and the measured (red) BC concentration for the first 19 weeks in 2020 in an urban street station (Anwerpen-Borgerhout-42R802), a suburban station (Antwerp-Wilrijk-42R817), two urban background stations (Gent-44R701 and Namur-43R401) and one background measurement point (Veurne-44N029). The dotted blue line indicates week 12, the start of the corona lockdown measures. The percentages in the blue bars indicate the decrease (or increase) of the weekly mean BC concentration compared to the (expected) concentrations calculated by the RF model.





BC weekly mean 2020, R817
RF Model - Measurements
Measurements
µg/m³
Week nr
8 weeks before lockdown: -0.1 µg/m³ (-13.9%)
after lockdown: -0.4 µg/m³ (-34.1%)
BC weekly mean 2020, R701
RF Model - Measurements
Measurements
µg/m³
Week nr
8 weeks before lockdown: -0.1 µg/m³ (-17.2%)
after lockdown: -0.3 µg/m³ (-28.9%)
BC weekly mean 2020, R401
RF Model - Measurements
Measurements
µg/m³
Week nr
8 weeks before lockdown: 0 µg/m³ (-6.4%)
after lockdown: -0.4 µg/m³ (-41.6%)



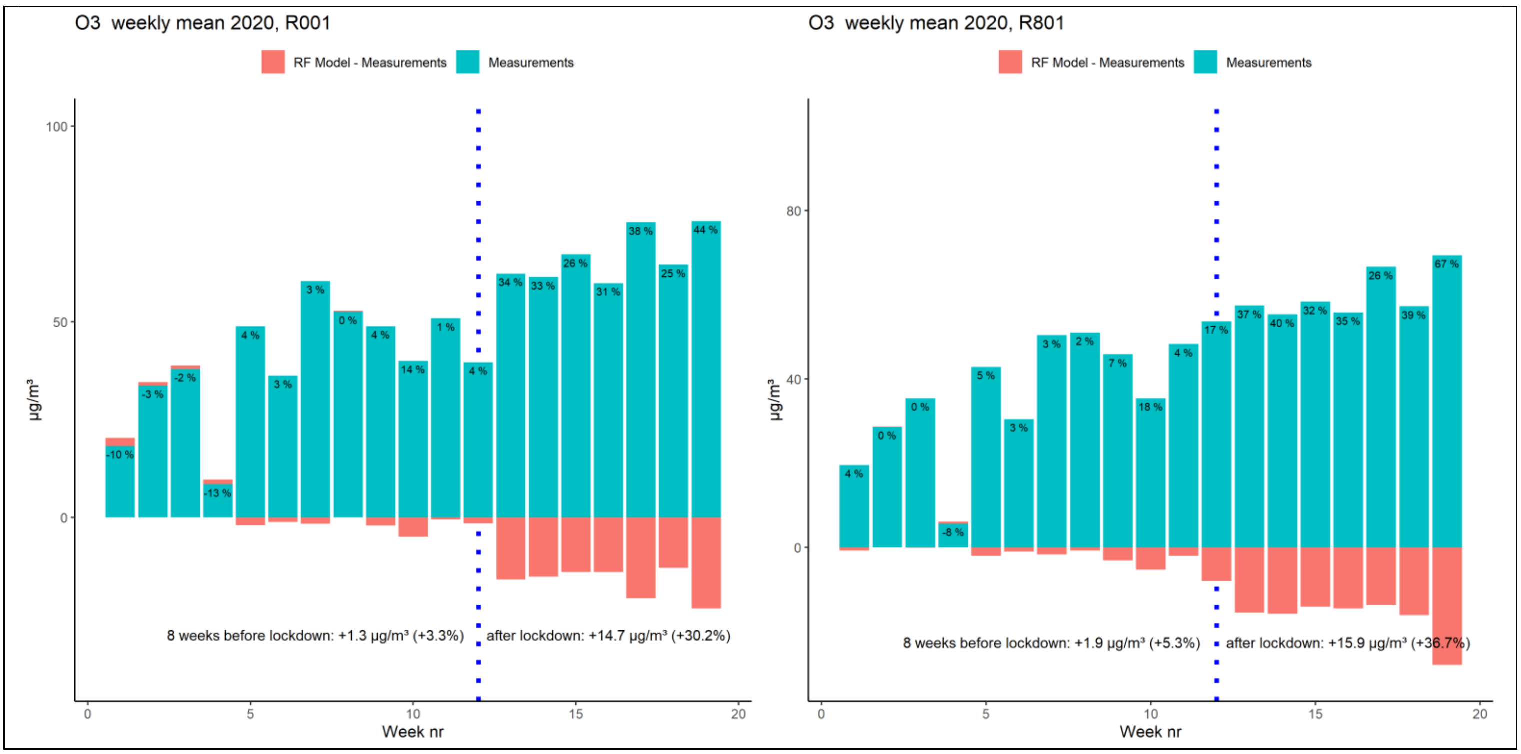
4.6 Ozone (O_3)

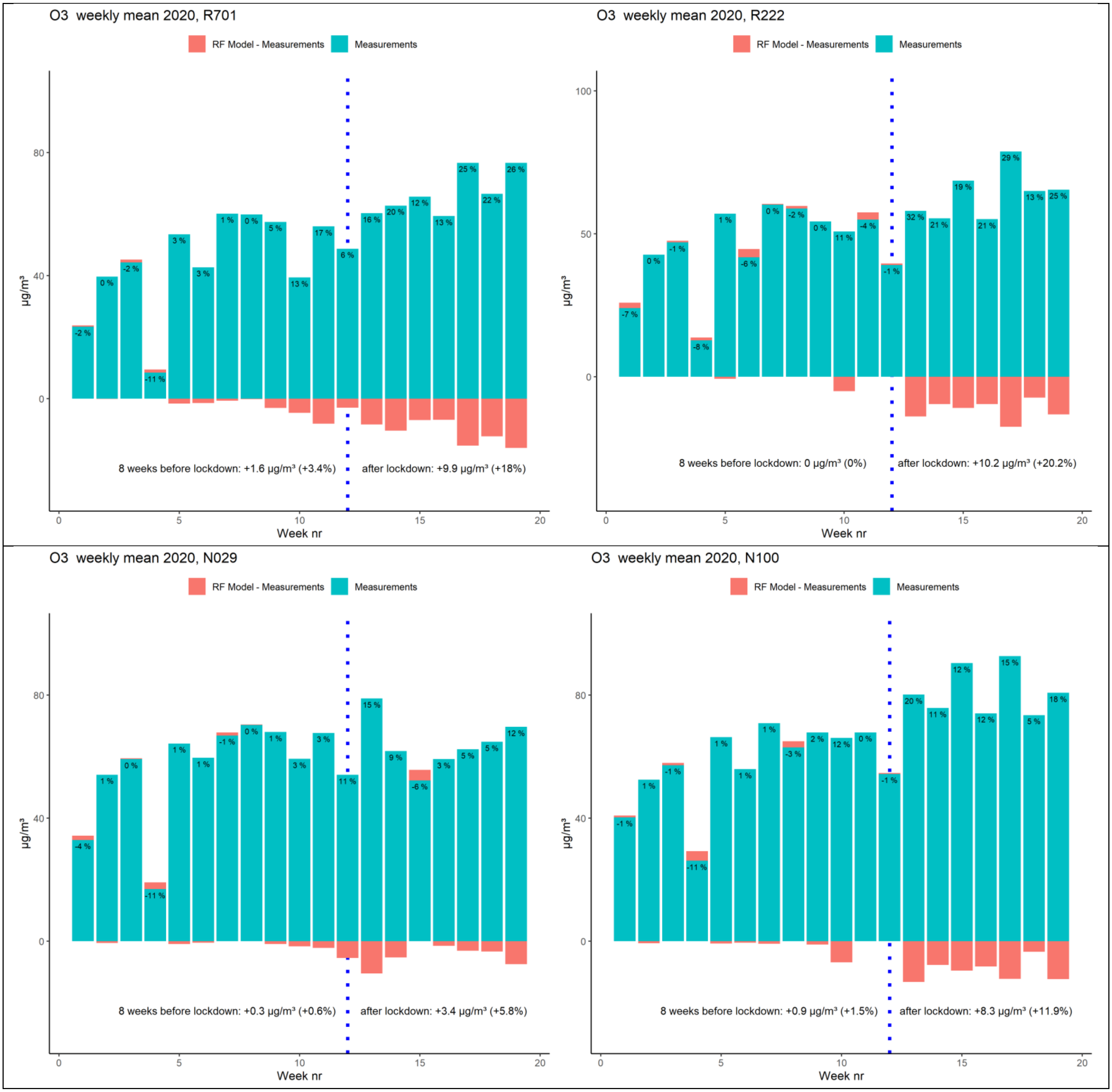
Ozone is a 100% secondary pollutant: ozone is not emitted directly but is created in the atmosphere from reactions between nitrogen oxides and volatile organic components ("ozone precursors"). The highest ozone concentrations occur on warm and sunny days. The chemical reaction that creates ozone in the troposphere (the lower part of the atmosphere) is an equilibrium reaction: ozone is formed but can also be broken down again. At chemical equilibrium, the same amount of ozone is formed as is broken down. The ozone formation reaction is determined by the amount (and ratio of) nitrogen dioxide (NO_2) and volatile organic compounds in the air, the temperature and the intensity of sunlight (UV radiation). Ozone decomposition is mainly determined by the concentration of nitrogen monoxide (NO). In places with a lot of traffic (a lot of NO) ozone depletion usually predominates, so that lower ozone concentrations are generally measured in urban environments and in places with a lot of traffic. This is sometimes referred to as the "ozone paradox".

This ozone paradox is also apparent during the lockdown period (see Figure 10): the ozone concentrations increased compared to what would be expected according to the model. This increase is most pronounced in the traffic stations. It is striking that the RF model estimates that this increase in ozone concentrations will continue over a period of 8 weeks. This would mean that despite significantly less traffic over a long period of time, sufficient "precursors" (NO_2 and volatile organic compounds) would still be present in the air to form ozone, but that the lower nitrogen monoxide (NO) concentrations at urban street measurement locations ensure that the ozone formed is broken down less. Earlier analysis with chemical transport models (CTM) also indicated that the (average) ozone concentrations in Western Europe increase even after drastic reductions in nitrogen oxides.

Another striking finding is that very high ozone concentrations were measured on Saturday 9 May at maximum temperatures as low as 25 °C. With 174 µg/m³ in Ghent, the ozone information threshold of 180 µg/m³ was nearly exceeded. The fact that the ozone concentrations continue to increase, even after a long period with less traffic, indicates that the ozone problem can only be solved by further sustainably and drastically reducing the ozone-forming substances (precursors).

Figure 10: measured weekly average (blue) and the difference between the expected (RF model) and the measured (red) ozone (O_3) concentration for the first 19 weeks in 2020 in the urban street station (Molenbeek-41R001), the urban background stations (Antwerp-Borgerhout- 42R801, Ghent-44R701 and Liège-43R222) and two background measurement sites (Veurne-44N029 and Dourbes-43N100). The blue line indicates week 12, the start of the corona lockdown measures. The percentages in the blue bars indicate the decrease (or increase) of the weekly average O_3 concentration compared to the (expected) concentrations calculated by the RF model.





O3 weekly mean 2020, R701
RF Model - Measurements
Measurements
µg/m³
Week nr
8 weeks before lockdown: +1.6 µg/m³ (+3.4%)
after lockdown: +9.9 µg/m³ (+18%)
O3 weekly mean 2020, R222
RF Model - Measurements
Measurements
µg/m³
Week nr
8 weeks before lockdown: 0 µg/m³ (0%)
after lockdown: +10.2 µg/m³ (+20.2%)
O3 weekly mean 2020, N029
RF Model - Measurements
Measurements
µg/m³
Week nr
8 weeks before lockdown: +0.3 µg/m³ (+0.6%)
after lockdown: +3.4 µg/m³ (+5.8%)
O3 weekly mean 2020, N100
RF Model - Measurements
Measurements
µg/m³
Week nr
8 weeks before lockdown: +0.9 µg/m³ (+1.5%)
after lockdown: +8.3 µg/m³ (+11.9%)

Table 4: Absolute (µg/m³) impact of the corona lockdown measures on the air quality (weeks 12 to 19) at the various measurement locations as estimated by the RF model.

Code	Measuring point	Classification	NO_x	NO_2	PM2.5	PM10	BC	O_3
Flanders								
42R801	Antwerp (Borgerhout)	urban background	-27.6	-12.8	-0.8	-1.6	-0.6	15.9
42R802	Antwerp (Borgerhout straat)	urban street	-37.4	-15	-2	-1.7	-0.6	
42R803	Antwerp (Park Spoor Noord)	urban background	-18.7	-9.5	-1.4	0.5	-0.4	
42R804	Antwerp (Ring)	urban street	-39.4	-14.8	-2.2	-1	-0.8	
42R805	Antwerp (Belgiëlei)	urban street	-33.2	-13.5	-1.2	0	-0.8	
42R817	Antwerp (Wilrijk)	suburban	-16.9	-8.9	0	2.6	-0.4	
44R701	Ghent (Baudelo)	urban background	-12.5	-5.9	-1.4	-0.5	-0.3	9.9
44R702	Ghent (Gustaaf Callier)	urban street	-31.1	-14.4	-0.5	1.1	-0.6	
44R703	Ghent (Lange Violettestraat)	urban street					-0.9	
44N029	Veurne (Houtem)	background	-1.8	-1.8	-1.8	0.4	-0.1	3.4
42N040	Sint-Pieters-Leeuw	background	-8.7	-6.2				7.7
42N016	Dessel	background	-7.4	-5.3	-1.4	1.8	-0.2	11.5
42N046	Lanaken (Gellik)	background	-7.8	-4.3				9.5
Brussels-Capital region								
41B001	Brussels (Kunst-Wet)	urban street	-52.1	-22.1				
41R001	Molenbeek	urban street	-29.7	-13.9	-2.1	-1	-0.3	14.7
41B004	Brussels (Katelijne)	inner city	-24.2	-12.3				21.4
41R012	Ukkel	urban background	-12.2	-9.4	-1	0.9	-0.1	11.7
Wallonia								
45R501	Charleroi	urban background	-25.4	-10.3	0.5	2.5		
45R502	Charlerloi (Lodelinsart)	urban background	-15.1	-7.4	-0.1	1.6		8.9
43R401	Namur	urban background	-24	-11.2	0.8	1.2	-0.4	
43R222	Liège	urban background	-20.2	-9	-0.4	-7.9		10.2
43N060	Havinnes	background	-4.2	-3	-2.5	-3.2		4.7
43N063	Coroy-Le-grand	background	-9.2	-5.8	-2.8	-0.7		12.4
43N100	Dourbes	background	-1.1	-1.1	-1.6	-0.3		8.3

Table 5: Relative (%) impact of the corona lockdown measures on air quality (weeks 12 to 19) at the various measurement locations as estimated by the RF model.

Code	Measuring point	Classification	NO_x	NO_2	PM2.5	PM10	BC	O_3
Flanders								
42R801	Antwerp (Borgerhout)	urban background	-50.5	-36.3	-4.5	-5.1	-40.1	36.7
42R802	Antwerp (Borgerhout straat)	urban street	-53.7	-37.7	-10.9	-5.3	-38.9	
42R803	Antwerp (Park Spoor Noord)	urban background	-42.6	-31.9	-8	1.8	-34.2	
42R804	Antwerp (Ring)	urban street	-50	-35	-12.1	-3.6	-43.3	
42R805	Antwerp (Belgiëlei)	urban street	-49.9	-35.7	-6.9	0.1	-50.4	
42R817	Antwerp (Wilrijk)	suburban/inner street	-42.8	-32.1	-0.2	10.8	-34.1	
44R701	Ghent (Baudelo)	urban background	-33.1	-21.2	-7.8	-1.5	-28.9	18
44R702	Ghent (Gustaaf Callier)	urban street	-56	-41.9	-3.1	4	-41.7	
44R703	Ghent (Lange Violettestraat)	urban street					-55.8	
44N029	Veurne (Houtem)	background	-9.9	-11.4	-10.3	1.6	-16.4	5.8
42N040	Sint-Pieters-Leeuw	background	-42	-38.2				12.9
42N016	Dessel	background	-37.3	-32.7	-8.6	7.1	-21	19.1
42N046	Lanaken (Gellik)	background	-29.6	-22.9				17.3
Brussels-Capital region								
41B001	Brussels (Kunst-Wet)	urban street	-55.8	-42.9				
41R001	Molenbeek	urban street	-46.9	-37.7	-14.1	-4	-19.3	30.2
41B004	Brussels (Katelijne)	suburban/inner street	-46.2	-37.5				49.8
41R012	Ukkel	urban background	-45.4	-50.1	-8.3	5.2	-23.8	19.2
Wallonia								
45R501	Charleroi	urban background	-44	-30.9	3.6	9.9		
45R502	Charlerloi (Lodelinsart)	urban background	-40.9	-30.2	-0.7	6.7		17.4
43R401	Namur	urban background	-52.9	-39.9	6.1	5	-41.6	
43R222	Liège	urban background	-46.6	-33.8	-2.8	-23.9		20.2
43N060	Havinnes	background	-25.5	-22.1	-17.5	-14.5		8.4
43N063	Coroy-Le-grand	background	-39.4	-31.7	-21.3	-3.5		21
43N100	Dourbes	background	-16.5	-16.5	-16.5	-2.3		11.9

Conclusion

The impact of the corona measures on the air quality can be calculated with RF models.

The impact of the corona measures is greatest for the typical traffic-related pollutants (NO_x, NO_2 and BC). In places with abundant road traffic the decreases in NO_x concentrations amount to more than 50%, for NO_2 and 35-40% for BC. These declines are slightly lower at urban background, inner-city and suburban measurement sites depending on the traffic intensity in the vicinity of these stations. In more rural areas, the decreases (depending on the measurement location) of the traffic-related components are clearly less pronounced.

The corona measures seem to have little impact on the concentrations of particulate matter. The particulate matter concentrations increased during the lockdown period compared to the previous 8-week perioddue to the less favourable meteorological conditions in the period with lockdown. This increase would have been slightly higher during a number of weeks, if there had not been a lockdown, but this is not the case for all weeks analysed. When interpreting the results, it should be taken into account that the RF model is less able to predict the concentrations for particulate matter. It is unlikely that the lockdown had a negative impact on particulate matter concentrations.

The ozone concentrations increased after the lockdown. This is most pronounced at traffic-congested measurement locations and the result of lower NO emissions in the period. In the presence of NO, ozone is broken down, a phenomenon known as the ozone paradox. This indicates that the ozone problem can only be solved by further sustainably and drastically reducing the ozone-forming substances (precursors).

References

- Breiman L. 2001. Random forests. Mach. Learn. 45 (1), 5-32. https://doi.org/10.1023.A:1010933404324

- Grange SK, Carslaw D. 2019. Using meteorological normalisation to detect interventions in air quality time series. Sci. Tot. Environ. 653, 578-88. https://doi.org/10.1016/j.scitotenv.2018.10.344.

- https://medium.com/@williamkoehrsen/random-forest-simple-explanation-377895a60d2d

- https://en.wikipedia.org/wiki/Random_forest

- https://CRAN.R-project.org/package=rmweather.

Appendix I: Statistical validation indicators

- Pearson Correlation Coefficient **(R)**

  This indicator provides information about the linear correlation between the measurements and the model values. The more it approaches 1, the more the point couples (model (M) - measurement (O)) are positively correlated or in other words the more they lie on a straight line with a positive slope with gradient 1. If this indicator is zero, this indicates that there is no correlation between the measurements and the model values.

$$R = \sum_{i=1}^{N}\left(M_i - \overline{M}\right)\bullet\left(O_i - \overline{O}\right) \Big/ \sqrt{\sum_{i=1}^{N}\left(M_i - \overline{M}\right)^2} \bullet \sqrt{\sum_{i=1}^{N}\left(O_i - \overline{O}\right)^2}$$

  M_i and O_i are individual model and measurement results, respectively. $\overline{M}$ and $\overline{O}$ are the average model and measurement results. N is the number of data pairs.

  Often in validation exercises the square of the correlation coefficient (R^2) or the determination coefficient is used. The R^2 is used as a percentage for the "explanatory variance".

- Root Mean Square Error (µg/m³) **(RMSE)**

  A measure of the deviation between the modelled and the measured value. The more this indicator approaches zero, the better the overall performance of the model. The RMSE can be strongly influenced by outliers.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(M_i - O_i\right)^2}$$

  In addition to the RMSE, the RRMSE or the "Relative" RMSE is also used.

$$RRMSE = RMSE/\overline{O}$$

- Mean Bias (µg/m³) **(MB)**

  This indicator shows the extent of the general over- or underestimation of the model compared to the measurements.

$$MBias = \frac{1}{N}\sum_{i=1}^{N}\left(M_i - O_i\right)$$

  In addition to the MB, the NMB or the "normalized" MB is also used.

$$NMB = MB/\overline{O}$$

In summary, it can be stated that the more the regression line coincides with the bisector (y = x) in the graph with the modelled concentrations as the y-axis and the measured concentrations as the x-axis, the better the model can predict the measurements. A model performs better the more R (and R^2) goes to 1 and the more the MB and RMSE go to zero.