Dr. Havala Pye
Topical Editor
GMD

Concerning: Response to reviewers of manuscript gmd-2020-282 entitled ”Simulation of O3 and NOX in Sao Paulo street urban canyons with VEIN (v0.2.2) and MUNICH (v1.0)” by Gavidia-Calderón M. E. et al.

Dear editor and reviewers,

We are thankful to the reviewers for their insightful comments and suggestions that help to improve our manuscript. We covered all your points below. For clarity, the referee’s comments are written in italics followed by our response in blue. In the mark-up version of the manuscript, changes based on Referee 1 comments are displayed in red, while changes based on Referee 2 comments are displayed in blue.

Many thanks,

Mario E. Gavidia-Calderón

Response to reviewers' comments (Manuscript ID: gmd-2020-282)

Referee 1

Comment 1

The paper assumes that the pollutant concentration is mainly contributed by the local sources, not regional sources. In a lot of cases, just the local emission amount may not be accurate. What about regional emission? In particular, O3 typically is a regional source that can be transported from a far way. Without quantifying the ratio between local and regional sources, it is difficult to evaluate the reliability of the model.

Reply: Thank you for this important observation. Previous studies in SPMA identify the vehicular fleet as the main source of air pollution (Andrade et al 2015, 2017). According to Sao Paulo Environmental Agency (CETESB), in 2014 the vehicular fleet was responsible for emitting 97% of CO, 82 % of VOCs, 78 % of NOX, and 40% of particulate matter (PM) emissions in SPMA (CETESB, 2015). To clarify the importance of the local sources, we include the following paragraph in section 2.3.1 Emissions and street link coordinates: "The vehicular fleet is the principal source of air pollution in SPMA (Andrade et al., 2015, 2017). The particularity of this fleet is the extensive use of biofuels (i.e. gasohol, ethanol, and biodiesel). During 2014, vehicular emissions were responsible for emitting 97 % of CO, 82 % of VOCs, 78 % of NOX, and 40 % of particulate matter (CETESB, 2015)."

On the other hand, as we described in section 2.3.4, background concentration in air quality modeling in street canyons accounts for the proportion of air pollutants that aren't emitted in the simulated street-network (Vardoulakis et al., 2003). In our case, we used concentrations of O3, NO2, NO from the Ibirapuera air quality station as background concentration. To explicitly state the air pollutants used as background concentrations, we add the following sentences in section 2.3.4: "In this work, measurements of O3, NO2, and NO in Ibirapuera AQS were used as background concentrations."

Comment 2

2.2 VEIN emission model Line 140-142: "Therefore, if we consider the mean emission factor ratio times the mentioned traffic flow ratio results that the NOX emissions should be approximately 2.37 higher." Is the suggested ratio of 2.37 considering contributions from both light vehicles and heavy vehicles?

Reply: Thank you for your comment. The answer is yes. As we detected less traffic flow by comparing GPS with travel demand models' outputs of light and heavy-duty vehicles, it should be less emissions. That paragraph was reformulated and we recalculated the ratios between real-world and laboratory emission-factors to produce adjustment factors, already implemented in a newer version of the VEIN model (https://atmoschem.github.io/vein/reference/ef_cetesb.html). Specifically, the real-world emissions factors for light-duty vehicles and trucks 1.11 and 1.38 times higher than the emission factors reported by the environmental authority (CETESB, 2015).

We rephrase the paragraph as follows: “The emissions dataset presents two aspects that need to be discussed. The first one is that there are some differences between the traffic flow from travel demand model outputs (TDM) and GPS (Ibarra-Espinosa et al., 2019, 2020). The ratio between traffic flows from TDM and GPS for our study area is 2.22. Regarding the emissions factors used to estimate the emissions, they are based on the average measurement of emissions certification tests (CETESB, 2015), therefore, they may underestimate real-drive emissions (Ropkins et al., 2009). For instance, the real-world emission factors derived from tunnel measurements in São Paulo for NOX were 0.3 g km-1 for light vehicles and 9.2 g km-1for heavy vehicles (Pérez-Martínez et al., 2014), while the respective fleet-weighted CETESB (2015) emission factors are 0.26 g km-1 and 6.68 g km-1, as shown in Fig. S1 in Supplement, resulting in ratios of 1.11 and 1.38. Then, if we consider the mean emission-factor ratio (1.11 + 1.38)/2, times the mentioned traffic flow ratio (2.22) results that the NOX emissions might be approximately 2.73 higher than the estimated using pure CETESB (2015) data. Consequently, we expect that air quality simulations for NOx might be lower than observations.”

Comment 3

Line 145: “We choose Wednesday emission as a typical weekday and Saturday emission for the weekend.” How much difference between typical Saturday and Sunday traffic in SPMA?

Reply: Thank you for your comment. One of the advantages of VEIN is the use of vehicle GPS data that allows a traffic estimation and therefore a better temporal and spatial emission profile. Figure 1 shows the mean emission from all street links from the Pinheiros neighborhood for NOX and VOCs. In the case of NOx emissions, Sunday total emissions are 25 % lower than Saturday total emissions, while in the case of VOC the values are almost the same. According to Ibarra et al. (2020), the difference between NOX emission during the weekday and the weekend is explained by the Buses contribution, which is lower during the weekend, and even lower during Sunday. Figure 1 is added to Supplement. This is an important point to explain NOX and NO overestimation during Sunday for both Pinheiros and Paulita Avenue urban canyons.

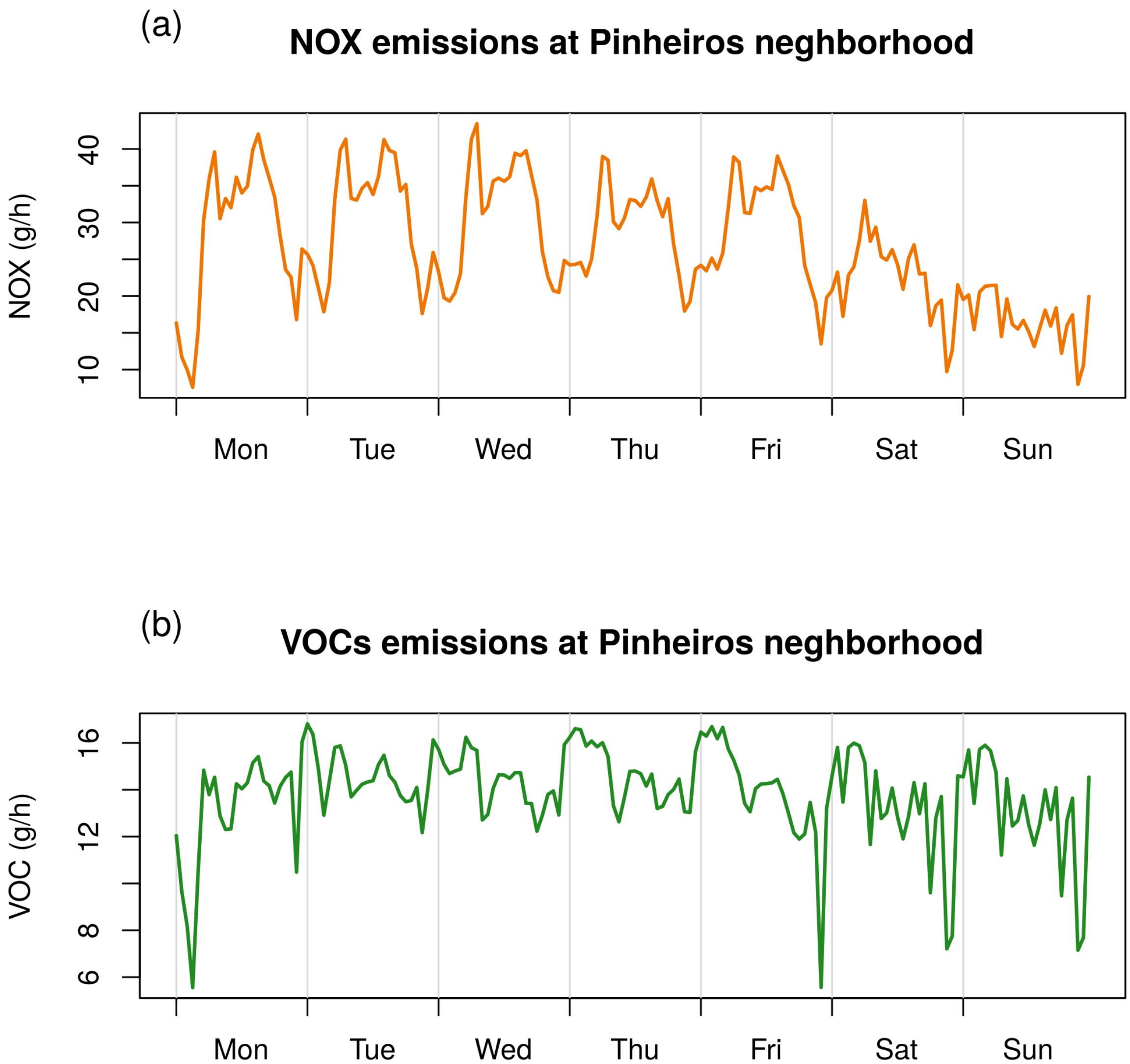


Figure 1. Mean emission from all street links from the Pinheiros neighborhood for (a) NOX and (b) VOCs for a typical week.

Comment 4

2.3.3 Building height and street width Line 176 "Building height is retrieved from the World Urban Database and Access Portal Tools project (WUDAPT) for SPMA (Fig. 3)." How well is WUDAPT describing building height? Especially, LCZ1, "compact high-rise", is having a description of "height of roughness elements >25m". It is also mentioned in line 226 that "Paulista Avenue domain is more uniform, presenting urban canyons with a mean building *height of 45 meters (LCZ1 - Compact high rise).", how is the value of 45 meters obtained? How sensitive is the model to these building height values?*

Reply: Thank you for noticing this. We explain this point by adding the following paragraph in section 2.2.3: "We retrieve the building height from the updated URBPARM.TBL file from WRF-Chem simulations in Pellegati et al. (2019). This file was built with the information described in Stewart et al. (2014), and contains the geomorphological and radiative parameters for each WUDAPT LCZ to be used in the Building Environment Parameterization (BEP) simulation in Pellegati et al. (2019)."

We believe that WUDAPT offers a good reference building height value rather than use a constant building height value. Certainly, this information needs to be improved by comparing with other data sources as Google Earth or by in-situ measurements. We rephrased line 399 in the Discussion and Conclusions section as the following: "On the other hand, now Google Earth allows new features like 3D view, that together with in-situ measurements, can improve WUDAPT building heights estimates."

Furthermore, we also ran a test with different constant building heights (i.e. 30 m, 50 m, 70 m). MUNICH is coherent with previous results where dispersion is restricted in deep urban canyon leading to higher pollutants concentrations (Afiq et al. 2012). As shown in Fig. 2 higher concentrations of NOX are produced inside the urban canyon when we increase the building height, this leads to a decrease of O3, by its reaction with the NOx. As we can see, background concentration and emission rates have a higher impact than the building height in air quality simulation with MUNICH.

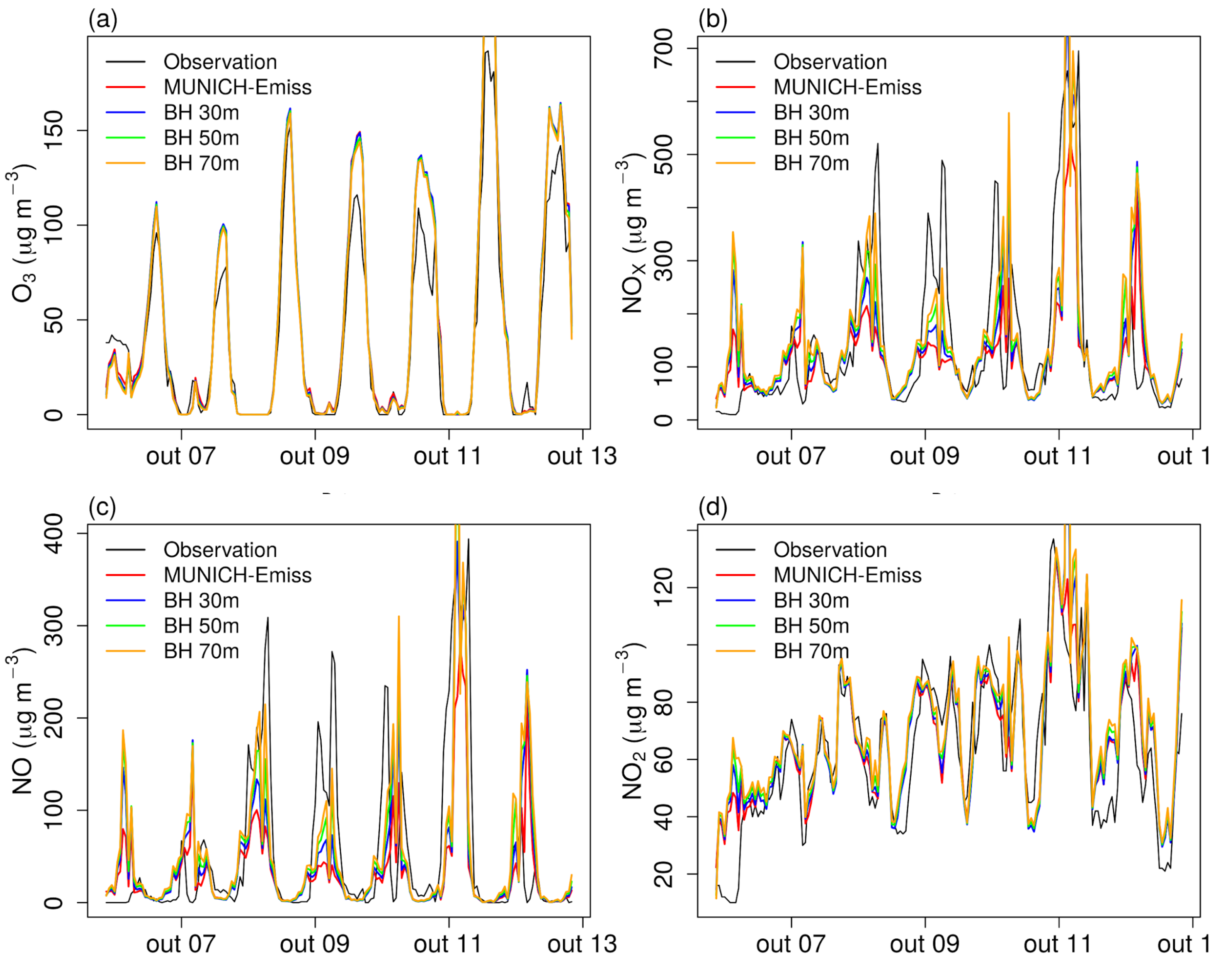


Figure 2. Effect of different building heights on MUNICH air quality simulations.

Comment 5

2.3.4 Background concentration Line 195-198 "With that in mind, by using the mean wind field from WRF simulation for the study period, we select Ibirapuera AQS (83 shown in Fig. 4) measurement as background concentration, which, according to the wind field, advect pollutants to Pinheiros station (99) and Cerqueira Cesar (83) as can be seen in Fig. 4." Is the difference of wind direction from mean during the study period justifying the choice of a single AQS at upwind to provide background concentration. Surely, that single station cannot be upwind for all year round?

Reply: Thanks for bringing this important question. When we analyze the wind fields generated by WRF simulations we can see that there is a different behavior during the daylight (Fig 3.a) and nighttime (Fig 3.b).

During daylight, there is the advection from Ibirapuera AQS to Pinheiros and Cerqueira Cesar AQS, whereas during night time west winds are predominant. As ozone concentrations during the night are low, it is more important to use information from air quality stations that measure the ozone upwind Pinheiros and Cerqueira Cesar AQS during daylight, when ozone concentrations are higher. For that reason, we chose Ibirapuera AQS. Still, as noted in the discussion section, it could be better to use air quality model results as background concentration for MUNICH, not only for a better background concentration estimate but also to address this wind direction implication. Figure 3 is added to Supplement, and we clarify that this assumption is valid during daylight.

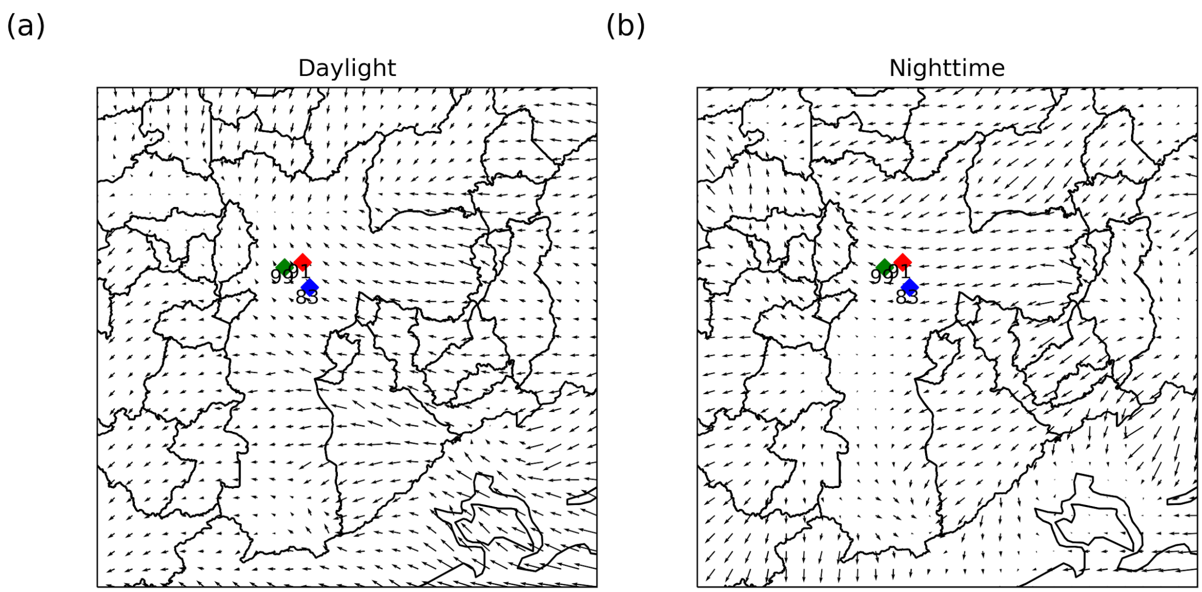


Figure 3. WRF averaged wind fields for daylight and nighttime during the simulation period. The green diamond shows Pinheiros AQS (99), the red diamond shows Cerqueira Cesar AQS (91), and the blue diamond shows Ibirapuera AQS (83)

Comment 6

Figure 4 Minor: in figure Cerqueira Cesar (red diamond) has number 91 instead of 83 as in line 197 and caption. Typo?

Reply: Thank you for noticing this. Yes, Red diamond should have the number 83. It is now corrected in the manuscript.

Comment 7

2.5 Model set up Line 215 “VEIN calculates the emissions for the whole SPMA” Line 219-220 “The red lines are the street links used by VEIN to calculate the emissions, and the yellow *rectangle the urban canyon selected for comparison against observation." I am not quite sure what this means. Are red lines in figure 5(a), (b) all street links in the domain? If there are street links that are not used by VEIN to calculate the emission? If so, how is their emission calculated?*

Reply: Thanks for bringing this up. As detailed in section 2.3.1, VEIN produces emissions for all the street links in SPMA. This information is a simple feature (sf) class object that contains a column with the Municipality/Neighborhood name of each street link. For this work, we subset the street links for Pinheiros neighborhood, and for the neighborhoods that contain the Paulista Avenue urban canyon. Therefore, the red lines in figure 5(a), (b) in the manuscript are a selection of the original VEIN output for SPMA. We clarify this in section 2.5 by adding the following sentence: "VEIN produces emissions for all the street links in SPMA. This information can be filtered by the neighborhood name of the street links. We subset that information for Pinheiros neighborhood (Fig. 5a), and for the neighborhoods that contain the Paulista Avenue urban canyon (Fig. 5b)."

Comment 8

3.2 Emission adjustment Line 263-264 "We ran different scenarios with increased NOX and VOCs emission from VEIN. The best results were produced when doubled the NOX and VOC emissions. This scenario is called MUNICH-Emiss." If there is any reason picking 2x as the adjusted emission? Would it perform better if higher emission, e.g., 2.5x, is used?

Reply: Thank you for your comment. We performed sensitivity tests with different emissions increment scenarios: the original emissions (original VEIN output), doubled emissions, tripled emissions, and quadrupled emissions. We noticed that the increment of emissions improves ozone simulation. Nevertheless, the increment could lead to unreasonable NOX concentrations, as in the case of the quadrupled emission scenario. The tripled emission scenario presented less error in magnitude than the doubled emission scenario, but it presented a lower Pearson correlation coefficient than the doubled emission scenario for NO, NO2, and NOX. To decide the better scenario, we used the index of agreement statistic (IOA). The doubled emission scenario presented higher IOA values for NO, NO2, and NOX. For that reason, we chose the doubled emission scenario as MUNICH-Emiss. We didn't test for 2.5x as the MUNICH-Emiss scenario already provided good results and reached Hanna and Chang (2012) performance criteria.

Comment 9

4 Discussion and conclusions Line 396 "calibrated emissions." What does this mean? Is it the MUNICH-Emiss? Or is it calibrated in some way?

Reply: Yes, in this case, "calibrated emissions" refers to the scenario where emissions are doubled. We have explicitly stated on the manuscript by adding "(i.e. MUNICH-Emiss scenario)" on line 396.

Referee 2

Comment 1

This manuscript demonstrates the street scale air quality modelling system and its evaluation for the city of Sao Paulo. The authors present it as the operational forecast system. However, the forecast system implies that the future atmospheric pollution can be predicted. And "forecast system" seems to be an improbable description of it (Line 85), given that you used real-time air quality observations to force your air pollution forecast. The current system is rather suitable for policymaking and future urban planning or post-accident analysis.

Reply: Thank you for pointing this out. Indeed, the forecast system will be achieved using a photochemical grid model to provide background concentration to MUNICH (like the case of SinG model described in Kim et al. (2018)) or an air quality on-line model that can provide both meteorological information and background concentrations. We briefly mention this point in the Discussion and Conclusions section when we detailed that output from photochemical grid models can improve MUNICH background concentration. Following your observation, we changed "forecast system" to "street-level air quality modeling system". The new paragraph is as follow:

"As the management of secondary pollutants remains a challenge in SPMA, we aim to evaluate MUNICH operational street-network model to simulate O3 and NOx concentration inside urban canyons, coupled with the VEIN emission model, to build a street-level air quality modeling system. This modeling system can be used in air quality and traffic management of Sao Paulo neighborhood, in studies of health effects from traffic emission exposure, in future urban planning, and post-accident analysis."

Comment 2

The meteorological driver (WRF) evaluation was performed in a slightly opaque manner since the authors did not mention neither the location (and number) of meteorological observation sites against which the model was evaluated nor the period of evaluation (perhaps of the same time extent as MUNICH runs). It is also unclear if the WRF output from D03 domain only was evaluated.

Reply: This is an important point. We performed the model evaluation only for our study period, the week from October 6th to 13th, 2014 as described in Table 2. We only evaluated the output from the finest domain (D03) as this is the domain that provided meteorological information to MUNICH. Figure 4 shows the air quality station locations, but not all the stations have meteorological information. Some air quality stations (AQS) only measure pollutant concentrations together with some meteorological parameters. During this period a total of 16 AQS have meteorological data. Only eight AQS measured temperature (T2), relative humidity (RH2), wind speed (WS), and wind direction (WD); five AQS measured only wind speed and direction; and three AQS measured only temperature and relative humidity. We updated Figure

4 to point the AQS with meteorological information. We also clarify these points by the following paragraph in section 2.3.2 WRF simulation:

“Before using the WRF simulation outputs for MUNICH modeling, a model verification is performed. Model verification was carried out for the same period as MUNICH runs and for the finest domain output (D03). We used meteorological information from 16 air quality stations which locations are shown in Figure 4.”

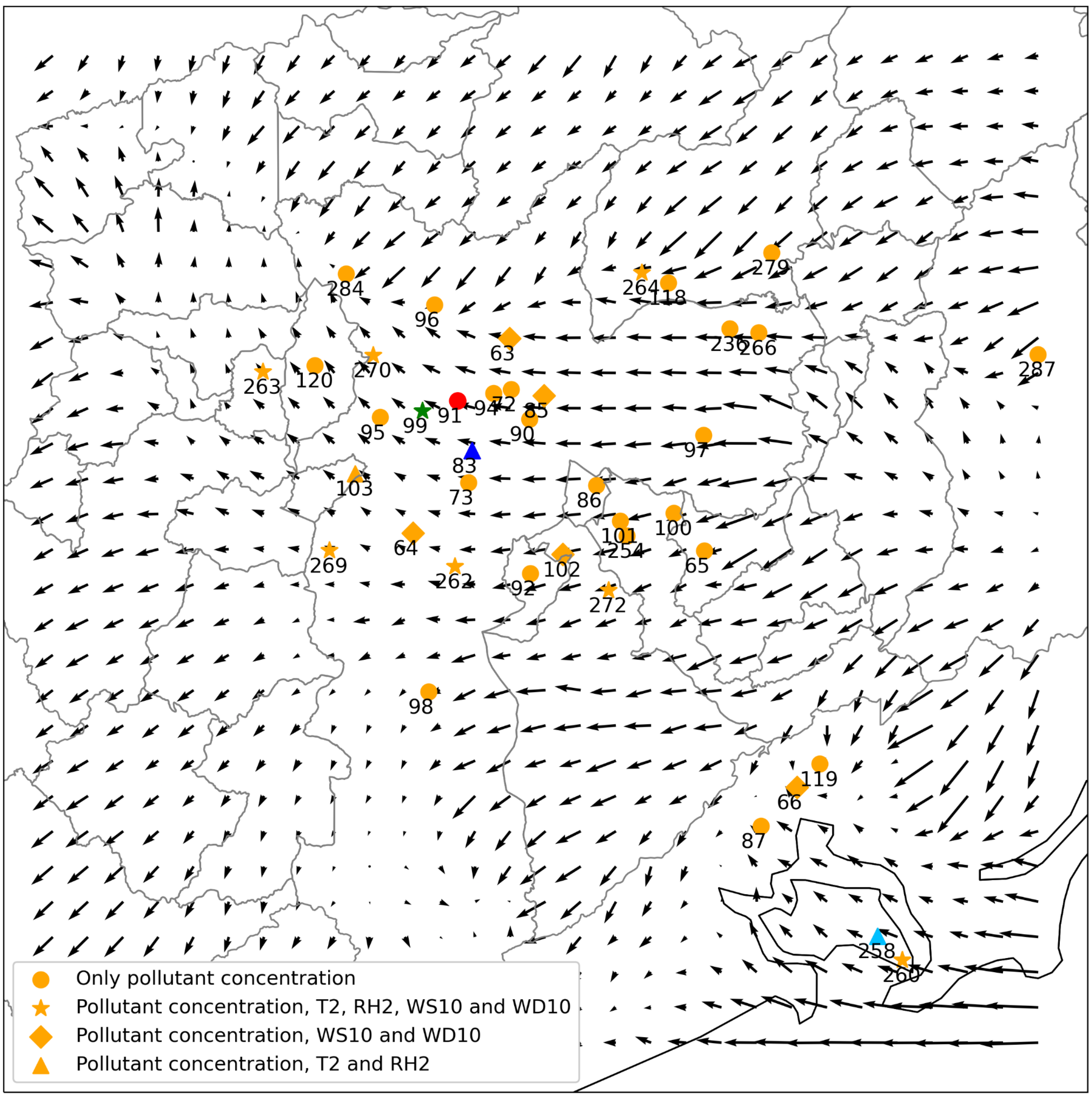


Figure 4. WRF average wind field for the simulation period with CETESB air quality stations (AQS). The green star

shows Pinheiros AQS (99), the red circle shows Cerqueira Cesar AQS (91), and the blue triangle shows Ibirapuera AQS (83). Circles represent AQS that only measure pollutant concentrations; stars represent AQS that also measures T2, RH2, WS10, and WD10; diamonds represents AQS that also measure WS10 and WD10; and triangles represent AQS that also measure T2 and RH2.

Comment 3

Perhaps, the authors could try to pinpoint the cause of large NOx and NO underestimation at Pinheiros AQS during Oct 8-9. Could it be associated with local meteorological conditions (probably unaccounted effect of nearby river, inversion etc.) or very local emissions just during those 2 days?

Reply: Thanks for bringing this up. The underestimation during Oct-8-9 can be explained by a very local emission episode as it did not happen in the Paulista Avenue domain, at least during October 9th where data is available. Still, underestimation of NOX concentration is caused by underestimation of NO concentration which is produced by a lower background concentration and an underestimation of emission factors as discussed in Section 2.3.1 Emissions and street links coordinates. Another factor is that MUNICH uses a single-day emission profile to represent weekdays emission, which can not account for the daily emission variation during the week. Meteorological factors as the overestimation of the wind speed by WRF model enhances dispersion. We add this information in section 3.2. Emission adjustment by rephrasing the paragraph as follows:

“NO_x and NO simulations are still underpredicted, but NO_2 is in the same magnitude as observations. NO_x underprediction is still mainly attributed to the underprediction of NO, especially during October 8th, 9th, and 10th where high observational values of NO were recorded. NO underestimation is explained by the lower NO background concentration, the underestimation of emissions, and the use of a single-day emission profile to represent all weekdays. Wind speed overestimation also affects this underestimation as it enhances dispersion. However, MUNICH can better represent the observed high concentration during Saturday 11th, as MUNICH uses the same emission profile for the weekend and weekdays, this high simulated NO concentration resulted from the influence of meteorology. “

Comment 4

The reasons behind two distinct peaks in NOx and NO observations (not captured by MUNICH) at both AQSs during night time seem to be ambiguous. Did the authors check if those are associated with meteorology? In case they are not related to any issue with meteorology, why did not the authors adjust emissions (one vs. two peaks) to fit the observed concentrations during the nights?

Reply: Thanks for this observation. Errors during nighttime can be caused by wrong representations of meteorology by WRF and by errors in the emission profile. In the case of meteorology, it is common that WRF presents troubles to represent the planetary boundary layer height during nighttime (Hu et al., 2012; McNider & Pour-Biazar, 2020). On the other hand, as shown in the emission profile during weekday and weekend days in Figure 1, NOX emissions do present two emission peaks during 7 hours and 16 hours, and a smaller emission peak around 23 hours, it is probable that this nighttime peak was underestimated. We add the following text in Section 3.3. Application for the Paulista Avenue:

"As in Pinheiros domain, MUNICH did not capture the two peaks of NO and NOX during nighttime. This is caused by WRF limitation in representing planetary boundary layer height during nighttime (Hu et al., 2012; McNider & Pour-Biazar, 2020). Also as shown in Fig. 1a, NOX emission profile during weekday present two peaks during daylight at 7 hours and 16 hours (Local Time), and a smaller emission peak around 23 hours, it is probable that this nighttime peak was underestimated."

Comment 5

Line 125: "street links" is confusing definition of roads, in particular for those who have never dealt with VIEN model. Perhaps, you should define it before using.

Reply: Agreed. Street links are segments of roads split at each vertex. Then, a road is composed of many links. We added this definition in section 2.3.1 Emissions and street links coordinates.

Comment 6

Lines 127-128: Could you please elaborate a bit on how the vehicular composition was obtained from GPS dataset and CETESB (2015) report? The report appears to be in Portuguese language and it might be hard to understand for those who speak/read English only.

Reply: The details about transforming GPS data into vehicular flow are described in Ibarra-Espinosa et al (2019). The details about using these GPS traffic flow to estimate vehicular emissions are described by Ibarra-Espinosa et al (2020). The CETESB report in Portuguese is cited only to cite the source of the emissions factors. CETESB measures and receives emissions laboratory measurements and report the emission factors. The references are below in this reply.

Comment 7

Line 140: The only number which fits the early-mentioned emission factors is 1.46. What is the 0.68 about?

Reply: We detected that real-world heavy trucks emissions factors from tunnel measurements (9.2 g km-1) are higher than laboratory measurements (6.3 g km-1) resulting in a ratio of 9.2/6.68 = 1.38. In the case of light vehicles, tunnel measurements emission factors (0.3 g km-1) are lower than laboratory measurements (0.44 g km-1), resulting in a ratio of 0.3/0.44 = 0.68. Recalling that the traffic is underestimated 2.2 times, the average of ratio emission factors (0.68+1.37)/2 times 2.2, results in approx in 2.3. This was confusing in the text and we apologize for that. But then, we realized that, as the tunnel emission factors are representative of the circulating fleet, we should weigh the CETESB emission factors by the circulating fleet as well. Then, we re-wrote the whole paragraph to improve the clarity as mentioned here:

“The emissions dataset presents two aspects that need to be discussed. The first one is that there are some differences between the traffic flow from travel demand model outputs (TDM) and GPS (Ibarra-Espinosa et al., 2019, 2020). The ratio between traffic flows from TDM and GPS for our study area is 2.22. Regarding the emissions factors used to estimate the emissions, they are based on the average measurement of emissions certification tests (CETESB, 2015), therefore, they may underestimate real-drive emissions (Ropkins et al., 2009). For instance, the real-world emission factors derived from tunnel measurements in São Paulo for NOX were 0.3 g km-1 for light vehicles and 9.2 g km-1for heavy vehicles (Pérez-Martínez et al., 2014), while the respective fleet-weighted CETESB (2015) emission factors are 0.26 g km-1 and 6.68 g km-1, as shown in Fig. S1 in Supplement, resulting in ratios of 1.11 and 1.38. Then, if we consider the mean emission-factor ratio (1.11 + 1.38)/2, times the mentioned traffic flow ratio (2.22) results that the NOX emissions might be approximately 2.73 higher than the estimated using pure CETESB (2015) data. Consequently, we expect that air quality simulations for NOx might be lower than observations.”

Comment 8

Lines 183-185: “The number of lanes is provided by the OpenStreetMap dataset. . .” and “Most OpenStreetMap streets do not include the number of lanes for this region. . .” seem to contradict each other. Both sentences should be reformulated to fit the method you actually used in the manuscript.

Reply: Agreed. The paragraph is rephrased as: “Most OpenStreetMap streets do not include the number of lanes for this region, therefore, they are hole-filled with the average by type of street. Then, street link width is calculated by assuming 3 m of line width and by adding 1.9 m to each side of the street as sidewalk width.”

Comment 9

Lines 196-197: The Ibirapuera AQS (83) does not seem to be the optimal location for background concentration if you look at the mean wind field of upstream region. Perhaps, the mean of observed concentrations from (83) and (94) AQSs would fit better for MUNICH’s forcing. Did the authors consider/try such forcing?

Reply: We chose Ibirapuera because it is located inside a park inside Sao Paulo city. Unfortunately, the air quality station with code 94 (Located at Sao Paulo downtown) does not have measurements of O3, NO, and NO2 for October 2014. So we couldn't consider it as background.

Comment 10

Line 276: phrase "MUNICH uses the same emission profile for the weekend and weekdays" is in contradiction with the section 2.3.1 and Figure 1, where emissions for weekdays and weekends are claimed to be different.

Reply: Agreed. Sentence is rephrased as: " However, MUNICH can better represent the observed high concentration during Saturday 11 th. As MUNICH uses the same emission profile for the weekdays and another emission profile for weekends, this high simulated NO concentration resulted from the influence of meteorology."

Comment 11

Table 4: There are often exceptions, but the fact that the correlation values equal strictly 1 in all 3 cases for ozone is unfortunately hard to believe. Maybe you rounded values or made some error during computations. Adding an extra digit for R values would be a good idea. Since the "Background" concentrations are also observed, it is unclear why authors evaluated and compared them with the street observations and what they tried to achieve by doing that (quality control?).

Reply: Thanks for this important observation. We added two digits for R values in Table 4, R between observations and background concentration was 0.9785, R between observations and MUNICH scenario was 0.9810, and R between observations and MUNICH-Emiss scenario (doubled emission scenario) was 0.9796. We rounded to two digits to R values to save space in Table 4.

We chose to evaluate background concentration against observation to see the difference between observation and background concentration and mainly to assess the influence of the background concentration in MUNICH simulations as previously shown in Wu et al. (2020).

Comment 12

Line 332: "in MUNICH NOx and NO peak happening before observation." Since you have many models and databases interfaced with each other, such mismatch in simulated concentrations could have happened because you did not match timings of datasets and models having them all, for example, in UTC. Are you sure the models and data were perfectly matched?

Reply: We took extremely careful consideration in the input time zone and its transformation to local time for a better visualization of model results. In this sense, all MUNICH input/output (i.e. WRF output, VEIN emissions, and background concentration) are in UTC. Change to local time (America/Sao Paulo) was performed using R functionalities - not manually- to avoid errors.

Response to technical corrections:

1. *Line 95: "before of no precipitation in" probably change to "before dry weather conditions in"*

Reply: Agreed. Sentence changed to "This period is chosen before dry weather conditions in SPMA"

*2. Line 136: please add reference for TDM Lines 146-149: The unit of flux [ug / km / h] is confusing (in Figure 1). Shouldn't it be something like [ug / km*2 / h], typo?*

Reply: Agreed. We added the reference for the TDM (Ibarra-Espinosa et al., 2019, 2020). We chose to plot emissions in ug/km/h because it is the unit that street emissions from VEIN required to be transformed to be read by MUNICH. We updated Figure 1 with emission in g/h which are the units used in VEIN. We also realized that Figure 1 was actually on UTC, now is change to Local Time.

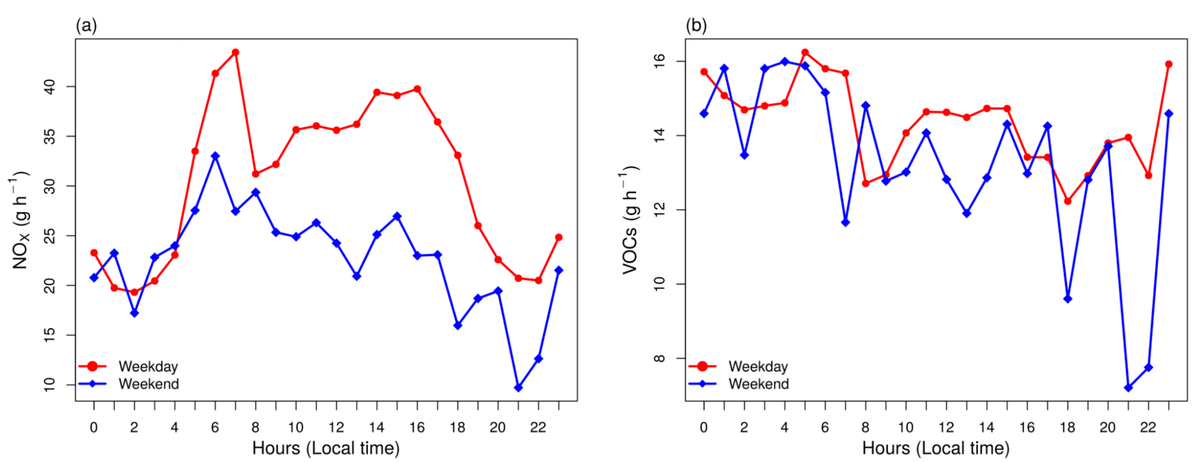


Figure 1. Mean emission from all street links from the Pinheiros neighborhood for (a) NO_X and (b) VOCs for typical weekday and weekend.

3. Line 161/ Figure 2: "WRF simulation domains for domains of. . ." please rephrase

Reply: Agreed. Sentence changed to “WRF simulation domains of 25 km (D01), of 9 km (D02), and of 1 km (D03) spatial resolution”.

4. Line 196: Cerqueira Cesar (83), should not that be 91 (similar typo in Figure 4)?

Reply: Agreed. Corrected to “the red circle shows Cerqueira Cesar AQS (91).”

5. Line 220: “rectangle the urban canyon” change to “rectangle is the urban canyon”

Reply: Agreed and change.

6. Line 229: “adn Paulista Avenue” change to “and Paulista Avenue”

Reply: Agreed and change.

7. Line 309: “We also perform additional” change to “We also performed an additional”

Reply: Agreed and change.

8. Line 319: “COV-limited regime” isn’t it “VOC-limited regime”?

Reply: Agreed and change.

9. Line 320: “with lead to” what does that mean, typo?

Reply: Thank you for noticing this. Sentence corrected to “the increment of NOX emission will lead to a reduction of O3 concentration”

10. Line 331: “but still higher than 0.5” it is imprecise as there are R values of 0.4 and 0.2 in the Table 5.

Reply: Agreed. Rephrased to :

“In this case, R values are lower than those in the Pinheiros case but still higher than 0.4 for NO2 and NOX, confirming that there is a mismatch of simulated concentrations, which is clearer in MUNICH NOX and NO peak happening before observation.”

11. Lines 341, 345: “Note that no O3 observation for Paulista Avenue.” seems grammatically incorrect sentence.

Reply: Agreed. Change to “Note that O3 observations were not available for Paulista Avenue domain.”

12. Line 386: “As the main source of superficial NO” probably you should write “. . . of elevated NO”

Reply: Agreed. That was actually a typo, the corrected sentence is “As the main source of surface NO and NO2 emissions in Sao Paulo are vehicles,”

References

Afiq, W. M. Y., Azwadi, C. S. N. and Saqr, K. M.: Effects of buildings aspect ratio, wind speed and wind direction on flow structure and pollutant dispersion in symmetric street canyons: A review, Int. J. Mech. Mater. Eng., 7(2), 158–165, 2012.

Andrade, M. de F., Kumar, P., de Freitas, E. D., Ynoue, R. Y., Martins, J., Martins, L. D., Nogueira, T., Perez-Martinez, P., de Miranda, R. M., Albuquerque, T., Gonçalves, F. L. T., Oyama, B. and Zhang, Y.: Air quality in the megacity of São Paulo: Evolution over the last 30 years and future perspectives, Atmos. Environ., 159, 66–82, doi:10.1016/j.atmosenv.2017.03.051, 2017.

Andrade, M. de F., Ynoue, R. Y., Freitas, E. D., Todesco, E., Vara Vela, A., Ibarra, S., Martins, L. D., Martins, J. A. and Carvalho, V. S. B.: Air quality forecasting system for Southeastern Brazil, Front. Environ. Sci., 3(February), 1–14, doi:10.3389/fenvs.2015.00009, 2015.

CETESB: Qualidade do ar no estado de São Paulo 2014, São Paulo. [online] Available from: http://www.cetesb.sp.gov.br/ar/qualidade-do-ar/31-publicacoes-e-relatorios, 2015.

Hanna, S. and Chang, J.: Acceptance criteria for urban dispersion model evaluation, Meteorol. Atmos. Phys., 116(3–4), 133–146, doi:10.1007/s00703-011-0177-1, 2012.

Hu, X. M., Doughty, D. C., Sanchez, K. J., Joseph, E. and Fuentes, J. D.: Ozone variability in the atmospheric boundary layer in Maryland and its implications for vertical transport model, Atmos. Environ., 46, 354–364, doi:10.1016/j.atmosenv.2011.09.054, 2012.

Ibarra-Espinosa, S., Ynoue, R. Y., Ropkins, K., Zhang, X. and de Freitas, E. D.: High spatial and temporal resolution vehicular emissions in south-east Brazil with traffic data from real-time GPS and travel demand models, Atmos. Environ., 222(May 2019), 117136, doi:10.1016/j.atmosenv.2019.117136, 2020.

Ibarra-Espinosa, S., Ynoue, R., Giannotti, M., Ropkins, K. and de Freitas, E. D.: Generating traffic flow and speed regional model data using internet GPS vehicle records, MethodsX, 6, 2065–2075, doi:10.1016/j.mex.2019.08.018, 2019.

McNider, R. T. and Pour-Biazar, A.: Meteorological modeling relevant to mesoscale and regional air quality applications: a review, J. Air Waste Manag. Assoc., 70(1), 2–43, doi:10.1080/10962247.2019.1694602, 2020.

Pérez-Martínez, P. J., Miranda, R. M., Nogueira, T., Guardani, M. L., Fornaro, A., Ynoue, R. and Andrade, M. F.: Emission factors of air pollutants from vehicles measured inside road tunnels in São Paulo: case study comparison, Int. J. Environ. Sci. Technol., 11(8), 2155–2168, doi:10.1007/s13762-014-0562-7, 2014.

Ropkins, K., Beebe, J., Li, H., Daham, B., Tate, J., Bell, M. and Andrews, G.: Real-World Vehicle Exhaust Emissions Monitoring: Review and Critical Discussion, Crit. Rev. Environ. Sci. Technol., 39(2), 79–152, doi:10.1080/10643380701413377, 2009.

Stewart, I. D., Oke, T. R. and Krayenhoff, E. S.: Evaluation of the “local climate zone” scheme using temperature observations and model simulations, Int. J. Climatol., 34(4), 1062–1080, doi:10.1002/joc.3746, 2014.

Vardoulakis, S., Fisher, B. E. A., Pericleous, K. and Gonzalez-Flesca, N.: Modelling air quality in street canyons: A review, Atmos. Environ., 37(2), 155–182, doi:10.1016/S1352-2310(02)00857-9, 2003.

Wu, L., Chang, M., Wang, X., Hang, J. and Zhang, J.: Development of a Real-time On-road emission (ROE v1.0)model for street-scale air quality modeling based on dynamic traffic big data, Geosci. Model Dev., (13), 23–40, doi:gmd-13-23-2020, 2020.

Simulation of O_3 and NO_X in Sao Paulo street urban canyons with VEIN (v0.2.2) and MUNICH (v1.0)

Mario E. Gavidia-Calderón[1], Sergio Ibarra-Espinosa[1], Youngseob Kim[2], Yang Zhang[3], and Maria de Fatima Andrade[1]

[1]Institute of Astronomy, Geophysics and Atmospheric Sciences, University of Sao Paulo, Sao Paulo, 05508-090, Brazil
[2]CEREA, Joint Laboratory École des Ponts ParisTech/EDF R&D, Université Paris-Est, 77455 Champs-sur-Marne, France
[3]Department of Civil and Environmental Engi~~e~~neering, Northeastern University, Boston, MA 02115, USA

Correspondence to: Mario E. Gavidia-Calderón (mario.calderon@iag.usp.br)

Abstract. We evaluate the performance of the Model of Urban Network of Intersecting Canyons and Highways (MUNICH) in simulating Ozone (O_3) and Nitrogen Oxides (NO_x) concentrations within the urban street canyons in the Sao Paulo Metropolitan Area (SPMA). The MUNICH simulations are performed inside Pinheiros neighborhood (a residential area) and Paulista Avenue (an economic hub), which are representative urban canyons in the SPMA. Both zones have air quality stations maintained by the Sao Paulo Environmental Agency (CETESB), providing data (both pollutants concentrations and meteorological) for model evaluation. Meteorological inputs for MUNICH are produced by a simulation with the Weather Research and Forecasting model (WRF) over triple-nested domains with the innermost domain centered over the SPMA at a spatial grid resolution of 1 km. Street~~ links~~ coordinates and emission flux rates are retrieved from the Vehicular Emission Inventory (VEIN) emission model, representing the real fleet of the region. The VEIN model has an advantage to spatially represent emissions and present compatibility with MUNICH. Building height is estimated from the World Urban Database and Access Portal Tools (WUDAPT) Local Climate Zone map for SPMA. Background concentrations are obtained from the Ibirapuera air quality station located in an urban park. Finally, volatile organic compounds (VOCs) speciation is approximated using information from Sao Paulo air quality forecast emission file and non-methane hydrocarbons concentration measurements. Results show an overprediction of O_3 concentrations in both study cases. NO_x concentrations are underpredicted in Pinheiros but are better simulated in Paulista Avenue. Compared to O_3, NO_2 is better simulated in both urban zones. The O_3 prediction is highly dependent on the background concentration, which is the main cause for the model O_3 overprediction. The MUNICH simulations satisfy the performance criteria when emissions are calibrated. The results show the great potential of MUNICH to represent the concentrations of pollutants emitted by the fleet close to the streets. The street-scale air pollutant predictions make it possible in the future to evaluate the impacts on public health due to human exposure to primary exhaust gases pollutants emitted by the vehicles.

1 Introduction

Street urban canyons are structures formed by a street and its flanked buildings (Oke et al., 2017). Due to their proximity to emissions from vehicles and their sides function as a compartment that limits pollutant dispersion, the street and the associated urban canyons are considered pollutant hotspots (Zhong et al., 2016). As more people start to live in urban areas (United Nations, 2018), and the ubiquity of urban canyons in the cities, pedestrians, commuters, bikers, and drivers are being exposed to high pollutant concentrations every day (Vardoulakis et al., 2003). Consequently, the study of air pollution inside urban canyons is an important matter when dealing with studies of human health exposure related to traffic emissions.

To estimate the real impact of the pollutants on human health, it is necessary to obtain accurate pollutant concentrations and the lengths of exposure. Most cities are not covered by a high-density network of air quality stations. Even though the measurements provide precise information, it is expensive and also very difficult to cover all of the impacted areas of a city (Zhong et al., 2016). One alternative, that is starting to be contemplated, is the use of numerical modeling to represent the pollutant behavior in urban canyons, which has the advantage of producing pollutant concentration information at high temporal and spatial resolutions.

Computational Fluid Dynamics (CFD) models are considered to be the best modeling approach to understand air pollutant dispersion inside the urban areas. Due to the limitations of high computational resources, these models cannot be applied for long time simulation periods nor for a large area (Fellini et al., 2019; Thouron et al., 2019).

A new type of model, the urban/local scale operational models, overcome these limitations by applying simplifications on urban geometry and parameterizations of the mass transfer processes of air pollutants inside the urban canyons. The Operational Street Pollution Model (OSPM) and the Atmospheric Dispersion Model System (ADMS-urban) are two of the most popular operational models, which have already been tested for different cities around the world (Berkowicz et al., 1997; McHugh et al., 1997). Their main advantage is that they calculate pollutant concentrations when sources and receptors are in the same street urban canyon, but they present a limited treatment for the pollutant transfer between streets and intersections (Carpentieri et al., 2012).

Street-network models are also operational, having the advantage of dealing with the transport of pollutants in city street intersections. The model SIRANE uses parametric relations to solve advection on the streets~~ links~~, the dispersion in the street intersections, and interchange between the streets and the over-roof atmosphere (Soulhac et al., 2011, 2012). Background concentrations at the over-roof atmosphere are estimated using a Gaussian plume model. This estimation method inhibits a comprehensive atmospheric chemistry treatment.

Recently, the Model of Urban Network of intersecting Canyons and Highways (MUNICH) was developed by Kim et al. (2018) using a similar parameterization as SIRANE. MUNICH includes improvements in the treatment of the mean wind profile inside the urban canyon and the turbulent vertical mass transfer at the top of the street. It solves pollutant reactions using a chemical mechanism, so it can also simulate the production of ozone inside the urban canyons. MUNICH has been used to simulate ozone (O_3) and nitrogen oxides (NO_x) by Wu et al. (2020) in Tianhe District of Guangzhou city, and NO_x as part of Street in Grid (SinG) model in Kim et al. (2018), Thouron et al. (2019) and Lugon et al. (2020) in the Paris region.

Significant information is required to run this kind of model. This is explained by Vardoulakis et al. (2003) that, in general, these models need at least information from traffic data, emissions, meteorological data, street geometry, and background concentrations. Recently, the VEIN model, a vehicular emission model, was developed by Ibarra-Espinosa et al. (2018) using information for Sao Paulo. VEIN is suitable to be used in street-network models because it uses the traffic flow, emission factors, and street morphology (i.e., intersection coordinates), to calculate the vehicular emissions. As a matter of fact, due to its architecture, it can be used together with MUNICH.

In Brazil, previous studies of air quality in urban canyons dealt with measurements of black carbon and O_3 inside a street canyon in Londrinas city center (Krecl et al., 2016), and dispersion of NO_x was simulated in Curitiba with the ENVI-met model (Kruger et al., 2011). To our knowledge, this is the first study of modeling O_3 and NO_x inside street urban canyons in Sao Paulo Metropolitan Area (SPMA), the biggest megacity in South America, where it is very often the exceedance of O_3 state air quality standard (Andrade et al., 2017).

As the management of secondary pollutants remains a challenge in SPMA, ~~the biggest megacity in South America,~~ we aim to evaluate MUNICH operational street-network model to simulate O_3 and NO_x concentration inside urban canyons, coupled with the VEIN emission model, to build a ~~forecast~~ street-level air quality modeling system. This ~~forecast~~ modeling system ~~for air pollutant concentrations at street level~~ can be used in air quality and traffic management of Sao Paulo ~~neighborhood~~neighbourhood, ~~and~~ in studies of health effects from traffic emission exposure, in future urban planning, and post-accident analysis.

2. Data and Methods

The experiment consisted of carrying out simulations of O_3, NO_x, NO, and NO_2 concentrations inside the SPMA urban street canyons with the MUNICH model. To evaluate model performance, the model results are compared against the measurements from Sao Paulo Environmental Agency (CETESB) air quality network. We choose Pinheiros urban area to test the model, where there is an air quality station in a mixed residential-commercial area. Once MUNICH and VEIN are calibrated, a study case is prepared by calculating the pollutant concentration inside Paulista Avenue, the economic central area of the city with high canyons. The selected study period covers the week from October 6th to October 13th of 2014. This period is chosen before ~~of no precipitation~~dry weather conditions in SPMA, a period of high O_3 concentrations (Carvalho et al., 2015), the availability of data, and the availability of the emission inventory developed for a typical week in October 2014 (Ibarra-Espinosa et al., 2020).

2.1. MUNICH model

MUNICH is conceptually based on the SIRANE model (Soulhac et al., 2011). It has two main components, the street-canyon component, which deals and solves pollutant concentrations inside the urban-canopy volume, and the intersection component, which calculates the pollutant concentrations inside the intersection volume. MUNICH differs from SIRANE in the treatment of the vertical flux by turbulent diffusion at the roof level (Schulte parameterization, Schulte et al., 2015) and in the mean wind velocity within the street canyon (Lemonsu parameterization, Lemonsu et al., 2004). Currently, MUNICH solves gas-phase pollutants based on the Carbon Bond mechanism version 5 (CB05). Further information is detailed in Kim et al. (2018).

2.2 VEIN emission model

VEIN is an R package (R Core Team, 2019) to estimate vehicular emissions at a street level. VEIN imports functions from the package Spatial Features (Pebesma, 2018), which represent different types of geometries on space and perform geoprocessing tasks, from the data table package (Dowle and Srinivasan, 2020) to perform fast aggregation of databases, and from the units package (Pebesma et al., 2016) to provide binding to the udunits library (https://www.unidata.ucar.edu/software/udunits/). VEIN includes a function to process vehicular flow at each street to generate activity traffic data, different emissions factors, and different sets of emissions calculation and post-processing tools (Ibarra-Espinosa et al., 2018). Specifically, the emissions factors are based on emissions certification tests with dynamometer measurements in laboratories (CETESB, 2015).

2.3 MUNICH input data

Urban canyon models required detailed input information, such as building height and street geometry. Their performance depends on the quality of this information (Vardoulakis et al., 2003). In recent years, new tools have been developed to generate this information. Table 1 summarizes the model input used in this simulation experiment.

Table 1. Summarized MUNICH input data.

Input data	Source
Meteorological input	WRF 3.7.1 simulation centered in SPMA (DX = 1km)
Street links coordinates and with lanes number	VEIN emission model (Ibarra-Espinosa et al., 2018)
Street links emissions	VEIN emission model (Ibarra-Espinosa et al., 2018)
Building height	World Urban Database and Access Portal Tools project (WUDAPT) database for SPMA (http://www.wudapt.org/)
Background concentration	O_3, NO, and NO_2 from the Ibirapuera Air Quality Station (AQS)
VOC speciation	Ethanol, Formaldehyde and acetaldehyde from WRF-Chem emission file from Andrade et al. (2015), other species are based from concentration showed in Dominutti et al. (2016)

2.3.1 Emissions and street links coordinates

The vehicular fleet is the principal source of air pollution in SPMA (Andrade et al., 2015, 2017). The particularity of this fleet is the extensive use of biofuels (i.e. gasohol, ethanol, and biodiesel). During 2014, vehicular emissions were responsible for emitting 97 % of CO, 82 % of VOCs, 78 % of NO_X and 40 % of particulate matter (CETESB, 2015). Vehicular emissions inside SPMA streets were estimated using VEIN emission model (Ibarra-Espinosa et al., 2018).

Street links are segments of roads split at each vertex. Then, a road can be composed of many links. Emission rates inside ~~the~~ these street links in VEIN model are calculated using 104 million GPS vehicles coordinates in southeast Brazil (Ibarra-Espinosa et al., 2019). The GPS dataset is assigned to the OpenStreetMap (2017) dataset and once traffic flow is obtained, the vehicular compositions are generated and assigned with each emission factor reported by CETESB (2015). Emission factors are transformed into speed function, and then the average speed calculated at each street is used to obtain more representative emissions at each hour of a week. In addition, the estimation was calibrated with fuel consumption for the year 2014. Ibarra-Espinosa et al. (2020) described all details regarding the emission estimation, with the emissions dataset in g h^{-1} available at https://github.com/ibarraespinosa/ae1.

The emissions dataset presents two aspects that need to be discussed. The first one is that there are some differences between the traffic flow from travel demand model outputs (TDM) and GPS (Ibarra-Espinosa et al., 2019, 2020). The ratio between traffic flows from TDM and GPS for our study is 2.22. Regarding the emissions factors used to estimate the emissions, they are based on average measurement of emissions certification tests (CETESB, 2015), therefore, they may underestimate real-drive emissions (Ropkins et al., 2019). For instance, the real-world emission factors derived from tunnel measurements in São Paulo for NO_X were 0.3 g km^{-1} for light vehicles and 9.2 g km^{-1} for heavy vehicles (Pérez-Martínez et al., 2014), while the respective fleet-weighted CETESB (2015) emission factors are 0.26 g km^{-1} and 6.68 g km^{-1}, as shown on Fig. S1 in Supplement, resulting in ratios of 1.11 and 1.38. Then, if we consider the mean emission-factor ratio (1.11 + 1.38)/2, times the mentioned traffic flow ratio (2.22) results that the NO_X emissions might be approximately 2.73 higher than the estimated using pure CETESB (2015) data. Consequently, we expect that air quality simulations for NO_X might be lower than observations. ~~The emissions dataset presents two aspects that need to be discussed. The first one is that the regional emissions inventory, which covered southeast Brazil, might not fully represent local reality. For instance, the ratio between travel demand models (TDM) and traffic flow of GPS for the study area is 2.22. Besides, the emission factors are average measurement of emissions certification tests, therefore, they may underestimate real drive emissions (Ropkins et al., 2019). Furthermore, the real-world emission factors derived from tunnel measurements in São Paulo for~~ ~~NO_X~~ ~~were 0.3 g~~ ~~km^{-1}~~ ~~for light vehicles and 9.2 g~~ ~~km^{-1}~~ ~~for heavy vehicles (Pérez-Martínez et al., 2014), while the respective fleet-weighted CETESB (2015) emission factors are 0.44 g~~ ~~km^{-1}~~ ~~and 6.3 g~~ ~~km^{-1}~~~~, as shown on Fig. S1 in Supplement, resulting the ratios of 0.68 and 1.46. Therefore, if we consider the mean emission-factor ratio times the mentioned traffic flow ratio results that the~~ ~~NO_X~~ ~~emissions should be approximately 2.37 higher.~~

Even when VEIN produces hourly emissions for a standard week (Fig. S2 in Supplement), MUNICH only considers a standard day for weekdays and weekends. We choose Wednesday emission as a typical weekday and Saturday emission for the weekend. Figure 1 shows the mean diurnal profile of NO_x and VOCs emission fluxes from street-links in the Pinheiros neighborhood.

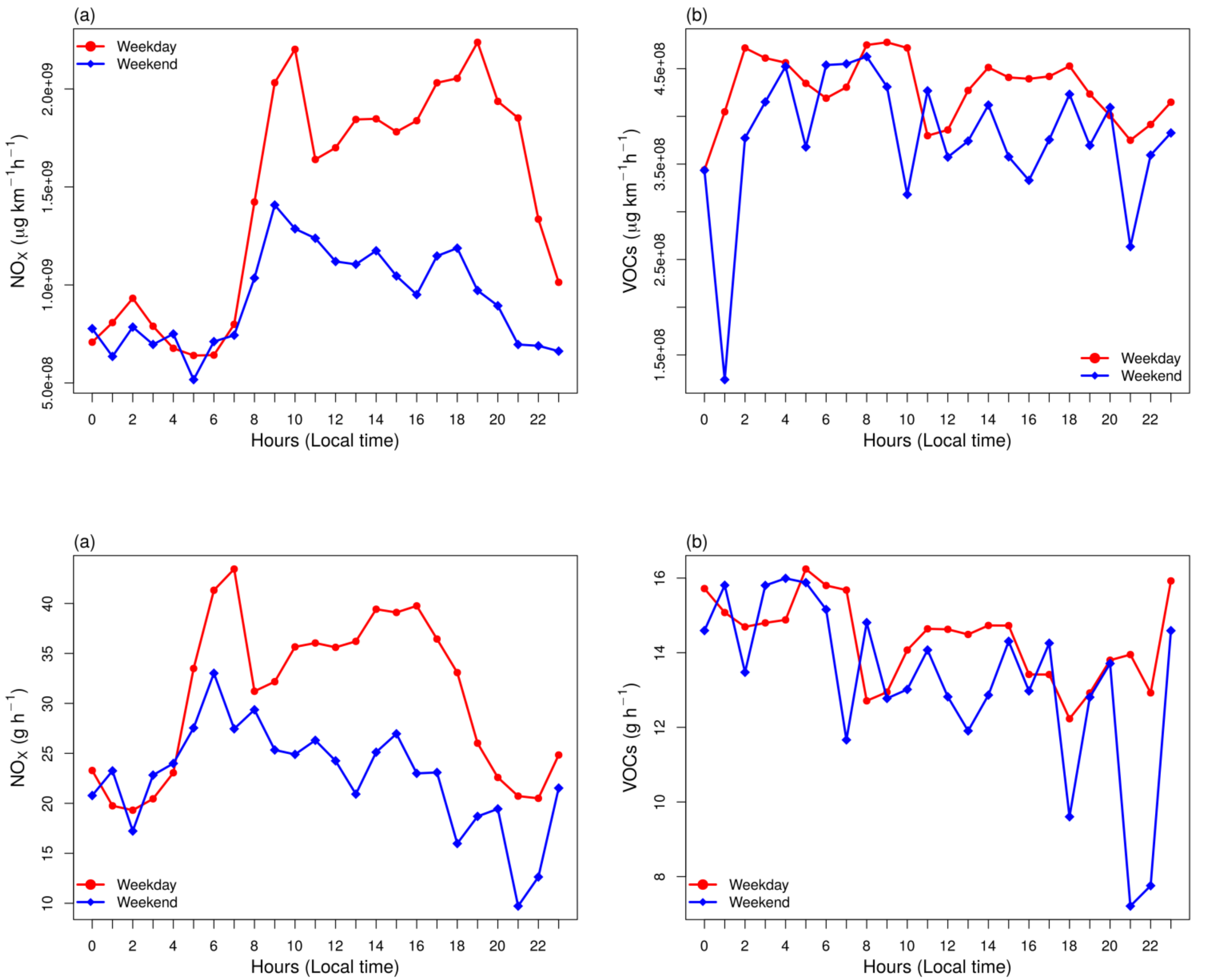


Figure 1. Mean emission from all street links from the Pinheiros neighborhood for (a) NO_X and (b) VOCs for typical weekday and weekend.

2.3.2 WRF simulation

Triple-nested domains are set up centered in SPMA. The mother domain has a spatial resolution of 25 km, the second 5 km, and the finest 1 km. The simulation at 1-km provides MUNICH with meteorological information. Initial and boundary conditions are retrieved from Historical Unidata Internet Data Distribution (IDD) Gridded Model Data (https://rda.ucar.edu/datasets/ds335.0/index.html). Table 2 shows WRF configuration and Fig. 2, the WRF domains.

Table 2. WRF simulation configuration.

Attribute	Configuration
WRF version	3.7.1
Domains spatial resolution	DX= 25 km, 5 km and 1 km
Simulation period	October 3rd to October 13th, 2014 (three first days are spin-up days and not analyzed)
Meteorological IC/BC	Historical Unidata Internet Data Distribution (IDD) Gridded Model Data (DS0335)
Longwave Radiation	RRTMG (Iacono et al., 2008)
Shortwave Radiation	RRTMG (Iacono et al., 2008)
PBL	YSU (Hong et al., 2006)
Surface Layer	Noah (Tewari et al., 2004))
Cumulus cloud	Multi-scale Krain-Fritsch (Zheng et al.,

	2016)
Cloud Microphysics	Morrison double-moment (Morrison et al., 2009)

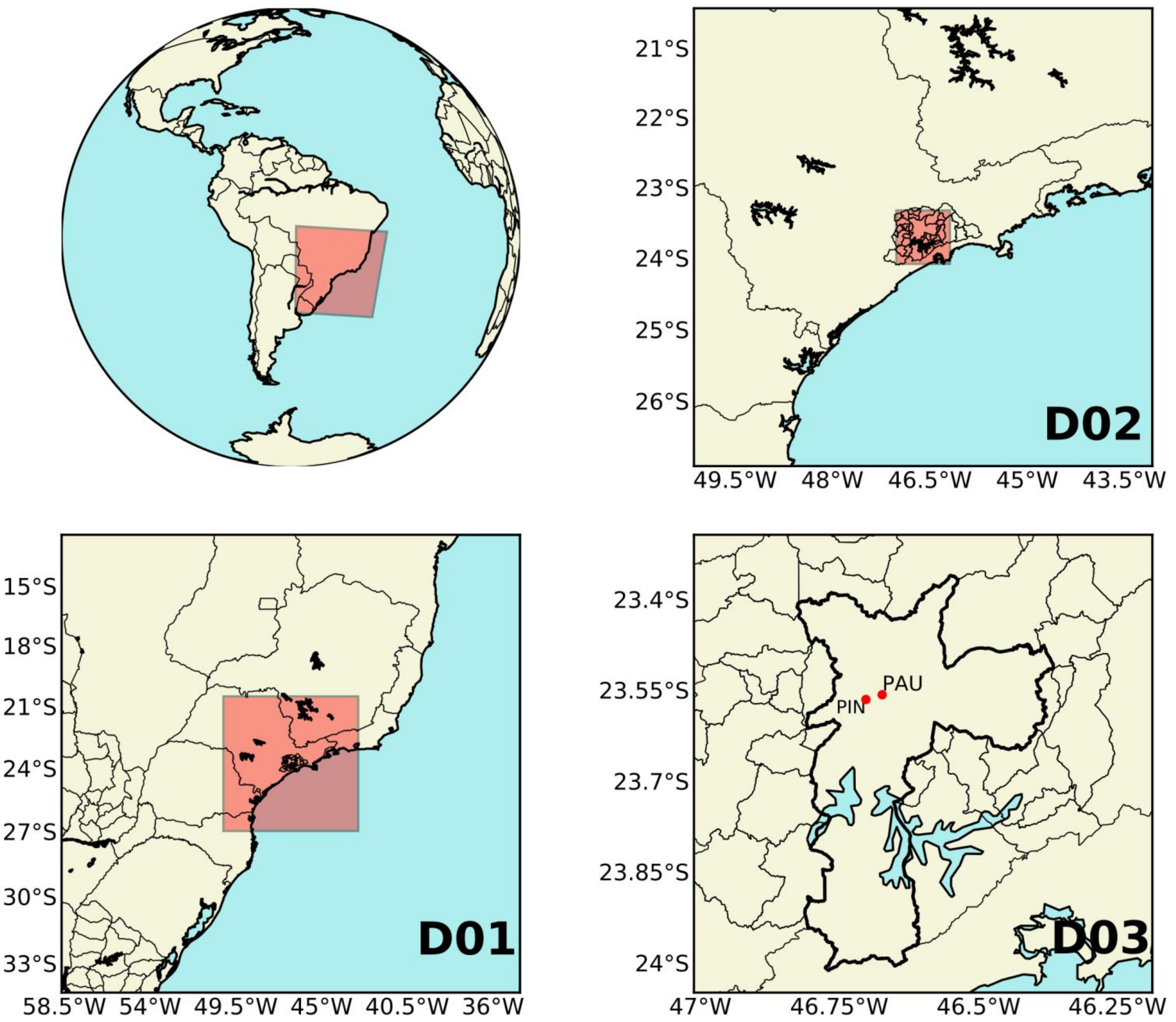


Figure 2. WRF simulation domains ~~for domains~~ of 25 km (D01), of 9 km (D02) , and of 1 km (D03) spatial resolution. D03 provides the meteorological information to MUNICH, Sao Paulo city is outlined in thick black line and the red dots shows MUNICH domains location.

Before using the WRF simulation outputs for MUNICH modeling, a model verification is performed. Model verification was carried out for the same period as MUNICH runs and for the finest domain output (D03). We used meteorological information from 16 air quality stations which locations are shown in Figure 4.

We also use benchmarks suggested by Emery et al. (2001), which were also used in Reboredo et al. (2015) and Pellegati et al. (2019). However, Monk et al. (2019) explained that these benchmarks are suitable for domains in "simple" terrain, they also presented other sets of benchmarks for "complex" terrain, the latter being more suitable for SPMA. The results are detailed in Table 3. The temperature at 2 m (T2) and relative humidity at 2 m (RH) reach the simple terrain benchmarks while wind speed and direction at 10 m (WS10 and WD10, respectively) are very close to them. When compared against complex terrain benchmarks, only the mean bias of WD10 is beyond the benchmark. Finally, T2, RH, and WS10 satisfy the good performance criteria of Keyser and Anthes (1977) and Pielke (2013). More details are shown in Tables S1 and S2 in the Supplement.

Table 3. WRF statistical model verification of simulation quality.

Parameter	Benchmark Simple terrain	Benchmark Complex terrain	Value from the WRF simulation
Temperature at 2m	MB[a] $< \pm 0.5$ K	MB $< \pm 1.0$ K	0.27 K
	MAGE < 2.0 K	MAGE < 3.0 K	1.59 K
	IOA ≥ 0.8		0.83 K
Relative humidity at 2m	MB $< \pm 10.0$ %		-5.02 %
	MAGE < 20 %		9.79 %
	IOA > 0.6		0.74
Wind speed at 10 m	MB $< \pm 0.5$ m.s^{-1}	MB $< \pm 1.5$ m.s^{-1}	0.79 m.s-1
	RMSE ≤ 2 m.s^{-1}	RMSE ≤ 2.5 m.s^{-1}	1.59 m s-1
Wind direction at 10 m	MB $< \pm 10.0$ °	MB $< \pm 10.0$ °	**-16.23** °

[a] MB: Mean bias, MAGE: Mean absolute gross error, IOA: Index of agreement and RMSE: Root mean square error. Results outside the benchmark are highlighted in bold.

MAGE < 30 º	MAGE < 55 º	55 º

2.3.3 Building height and street width

Building height is retrieved from the World Urban Database and Access Portal Tools project (WUDAPT) for SPMA (Fig. 3). WUDAPT classifies urban areas into 17 Local Climate Zones (LCZ). These LCZ are divided into build types, which are LCZ from 1 to 10, and land cover types, which go from A to G. Each of these LCZ presents different thermal, radiative, surface cover, and geometric properties. The building height is the height of roughness elements, which is the geometric average of building heights (Stewart and Oke, 2012). The WUDAPT file for SPMA is a raster with a spatial resolution of 120 m and was previously used in Pellegati et al. (2019).

~~We retrieve the b~~Building height values for each LCZ are extracted from the URBPARM.TBL file from WRF-Chem simulations in Pellegati et al. (2019) and assigned to Sao Paulo WUDAPT raster file. The URBPRAM.TBL~~is~~ file contains the geomorphological and radiative parameters values for each LCZ based on Stewart et al. (2014).~~from WUDAPT data to be used in the Building Environment Parameterization (BEP) urban parameterization simulation test in Pellegati et al. (2019). It is based on Stewart et al. (2014).~~

The number of lanes is provided by the OpenStreetMap dataset, so the street width is calculated by using 3 m of lane width and by adding 1.9 m to each side of the street as sidewalk width. Most OpenStreetMap streets do not include the number of lanes for this region, therefore, they are hole-filled with the average by type of street.

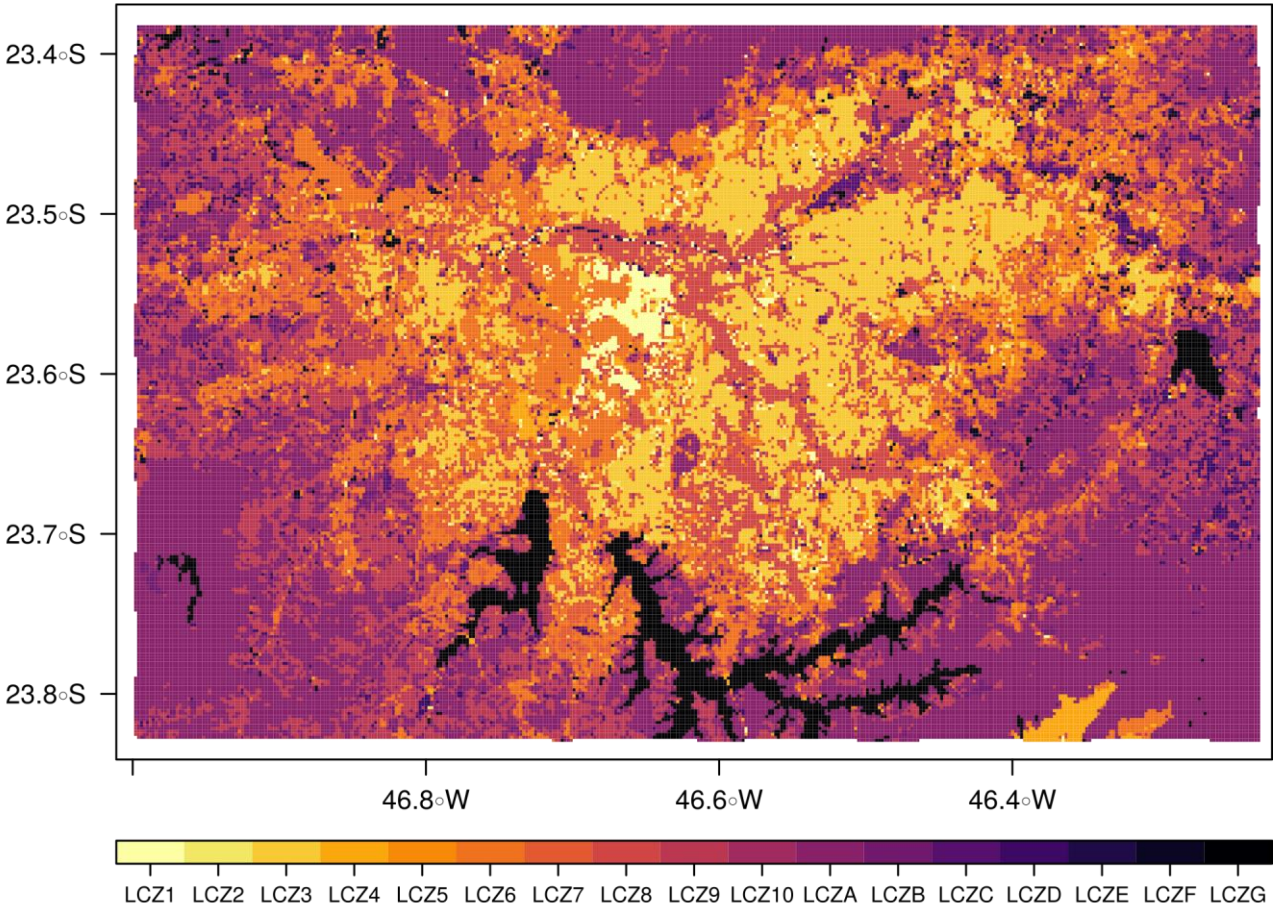
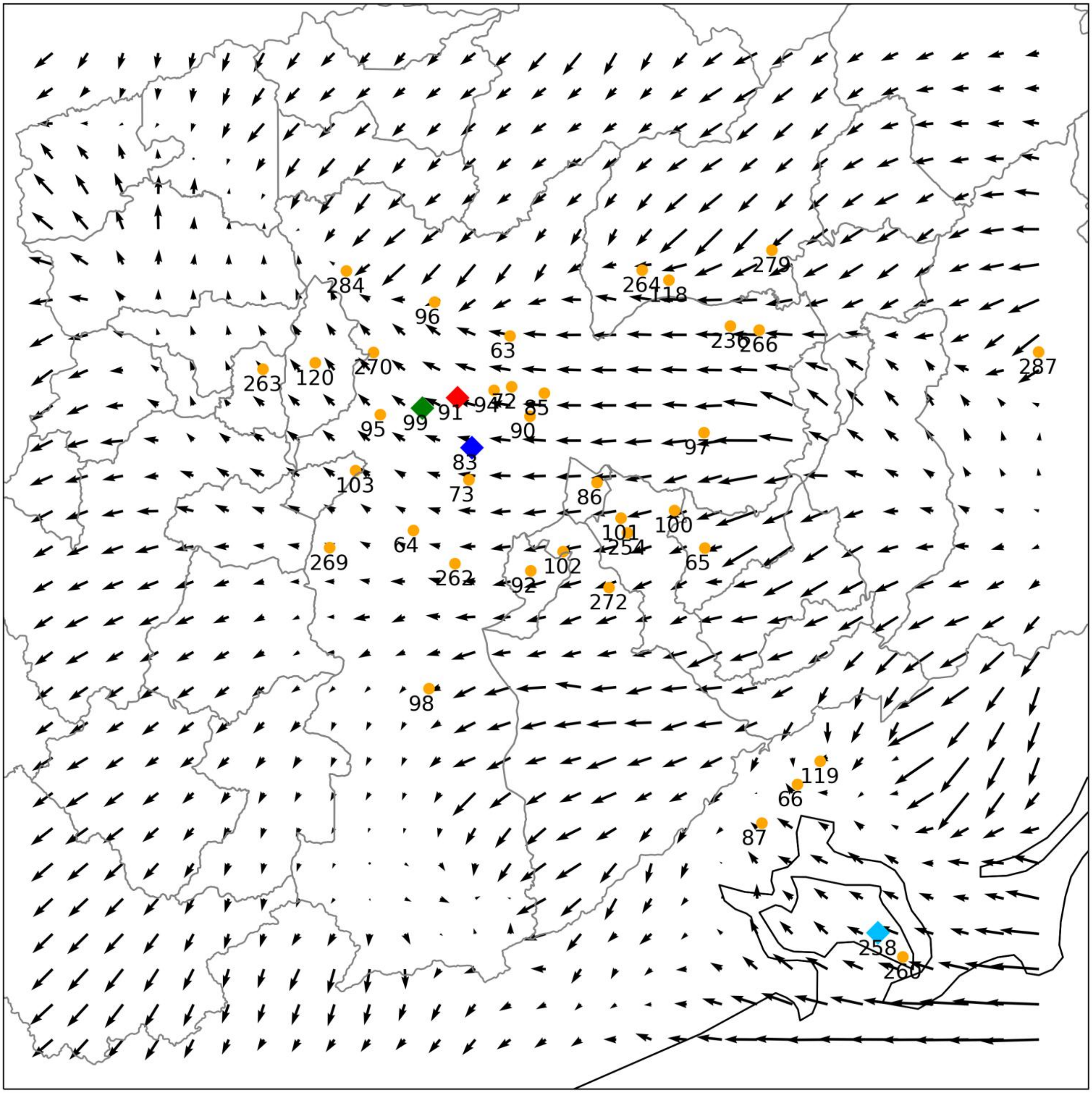


Figure 3. Local Climate Zones for SPMA.

2.3.4 Background concentration

Vardoulakis et al. (2003) explained that the background concentration in street modeling is necessary to include the proportion of air pollutants that are not emitted inside the street~~ links~~. In the SinG model, background concentrations are the concentrations calculated by Polair3D, a mesoscale air quality model (Kim et al., 2018). Wu et al. (2020) chose as the background concentration, measurements from a station located very close to the study zone. Consequently, we consider background concentration the concentration outside the MUNICH domain. With that in mind, by using the mean wind field from WRF simulation for the study period, we select Ibirapuera AQS (83 shown in Fig. 4) measurements as background concentration, which, according to the wind field, advect pollutants to Pinheiros station (99) and Cerqueira Cesar (83) as can be seen in Fig. 4. This assumption is only valid during daylight, when ozone concentrations are higher. As seen in Fig. S3 in Supplement, during nighttime wind presents a westerly direction. Measurements of O_3, NO_2, and NO in Ibirapuera AQS were used as background concentrations.


284
96
264
118
279
63
236
266
287
263
120
270
94
72
85
91
99
95
90
97
83
103
73
86
101
100
254
64
269
65
102
262
92
272
98
119
66
87
258
260

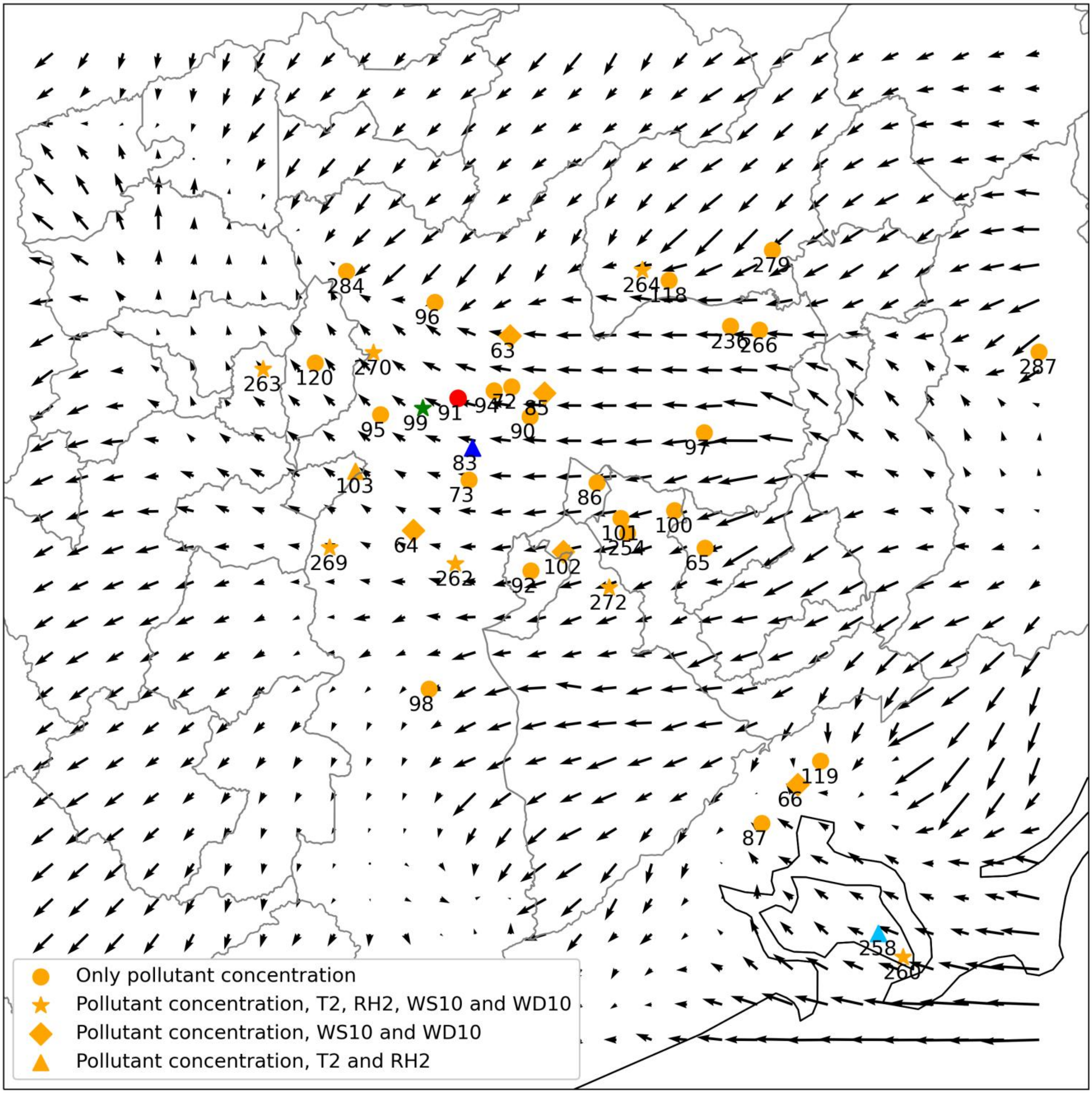


Figure 4. WRF average wind field for the simulation period with CETESB air quality stations (AQS). The green ~~diamond~~ star shows Pinheiros AQS (99), the red ~~diamond~~ circle shows Cerqueira Cesar AQS (~~83~~91), and the blue ~~diamond~~ triangle shows Ibirapuera AQS (83). Circles represent AQS that only measure pollutant concentrations; stars represent AQS that also measures T2, RH2, WS10 and WD10; diamonds represents AQS that also measure WS10 and WD10; and triangles represent AQS that also measure T2 and RH2.

2.4 Measurements and statistical analysis

Meteorological and air pollutant measurements are retrieved from CETESB air quality network. To evaluate WRF simulation in the finest domains, observations from 41 air quality stations (AQS) are used. Background concentration comes from Ibirapuera AQS. Pinheiros AQS is used to evaluate MUNICH performance in the Pinheiros neighborhood, while Cerqueira Cesar is used to evaluate Paulista Avenue. To evaluate model performance we follow the recommendations from Emery et al. (2017). We also use the evaluation statistics from Hanna and Chang (2012): Fractional bias (FB), Normalized mean-square error (NMSE), Fraction of predictions within a factor of two (FAC2), and normalized absolute difference (NAD~~) .~~). The acceptance criteria for urban zones are: |FB| <= 0.67, NMSE <= 6, FAC2 >= 0.3 and NAD <= 0.5. We expand the statistical analysis to the background concentration to see the difference against observation and to assess the influence of background concentration in MUNICH simulations.

2.5 Model set up

We use MUNICH to simulate two urban areas inside SPMA, the first domain is Pinheiros neighborhood and the second one is Paulista Avenue. ~~VEIN calculates the emissions for the whole SPMA, so we retrieve NO_x, NO_2, and VOCs emissions for the streets located in both domains.~~ VEIN produces emissions for all the street links in SPMA. This information can be filtered by the neighborhood name of the street links. We subset that information for Pinheiros neighbourhood (Fig. 5a), and for the neighborhoods that contain the Paulista Avenue urban canyon (Fig. 5b). In MUNICH, NO emissions are estimated from NO_x and NO_2 emissions.

Figure 5 shows MUNICH domain for the Pinheiros neighborhood and Paulista Avenue. The yellow dot represents the location of the air quality stations. The red lines are the street links used by VEIN to calculate the emissions, and the yellow rectangle is the urban canyon selected for comparison against observation.

There are 677 street links for Pinheiros and 535, for Paulista Avenue. Nine points of WRF simulation cover the Pinheiros domains, while twelve WRF points represent Paulista Avenue domains. From WUDAPT we can see that inside Pinheiros there is a variety of buildings with different heights. Pinheiros AQS is located in an urban canyon that has a mean building height of 5 meters (LCZ 6 - Open Low Rise). On the other hand, Paulista Avenue domain is more uniform, presenting urban canyons with a mean building height of 45 meters (LCZ1 - Compact high rise).

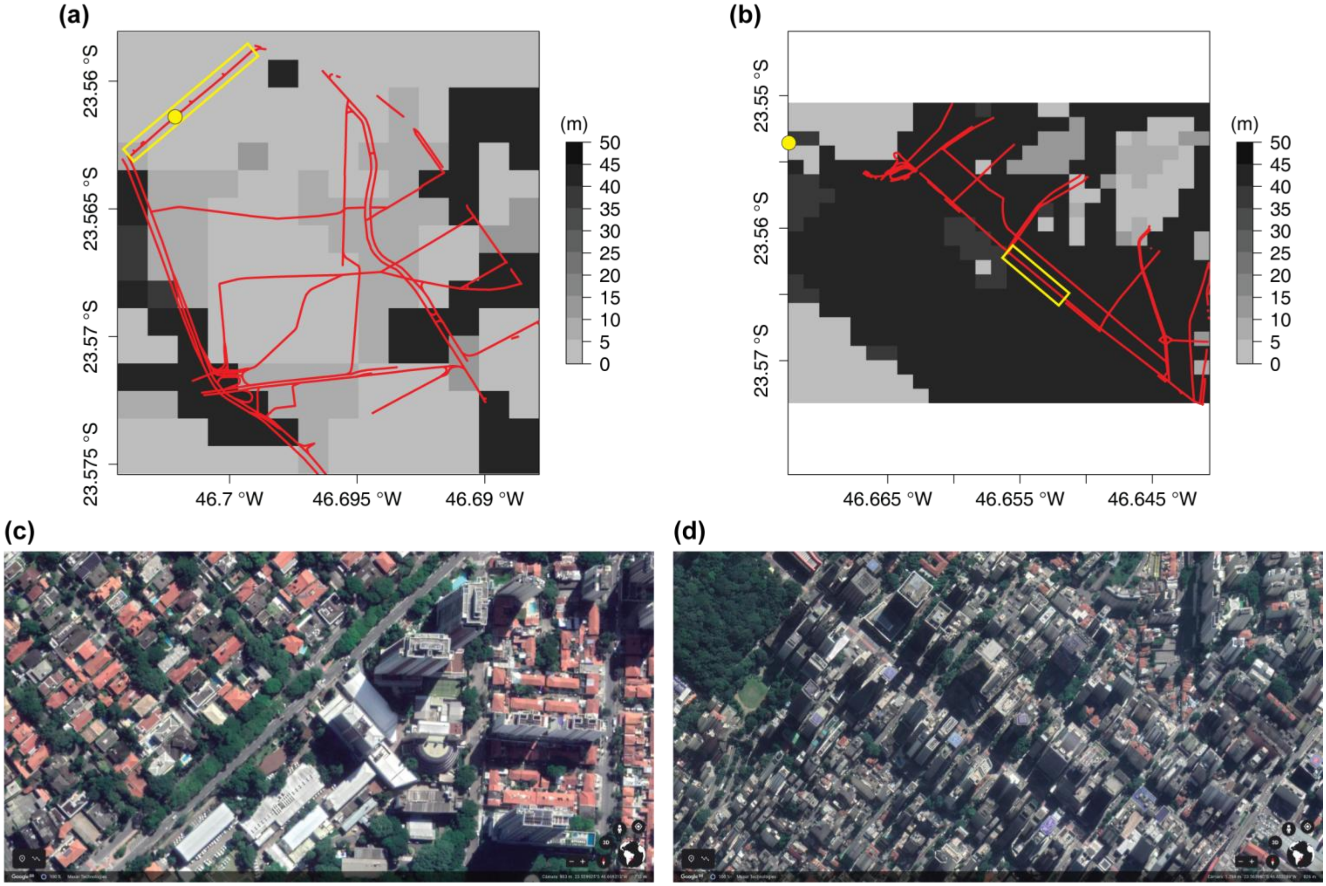


Figure 5. Pinheiros neighborhood (a) a~~d~~nd Paulista Avenue (b) MUNICH domains and building height, the red lines are the streets considered in VEIN, the yellow dot shows Pinheiros AQS and Cerqueira Cesar (AQS). Yellow squares highlight the selected urban canyon for comparison against observation. At the bottom, satellite photos of those urban canyons (Source: © 2019 Google, Image © 2019 Maxar Technologies).

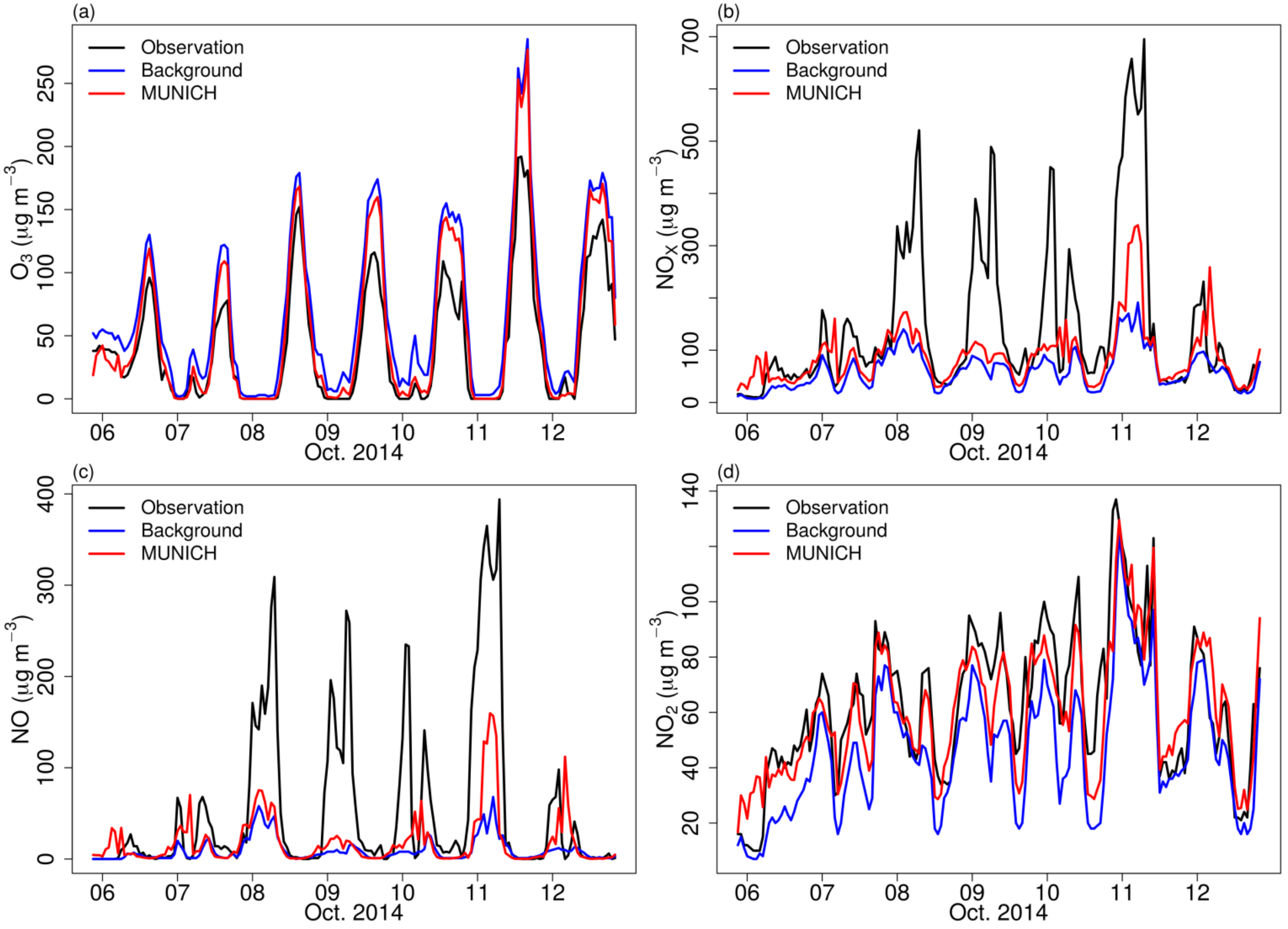
3 Results

Here we present the O_3 and NO_x simulations with MUNICH for a week of October 2014. We first calibrated the input emissions by studying Pinheiros neighborhood, to later simulate NO_x inside Paulista Avenue urban canyon.

3.1 Control case for the Pinheiros neighborhood

Figure 6 shows the results of MUNICH simulation using the original emissions calculated by VEIN for SPMA. MUNICH simulations are very close to background concentrations, which leads to an overprediction of O_3 and underpredicted NO and NO_X concentrations. This is produced by a dependence of MUNICH on background concentration and by emission underestimation. The emission underestimation is caused by emission factors calculated based on average measurements of emissions certification tests, and because emission factors derived from dynamometer, and cycle measurements do not represent real-drive emissions (Ropkins et al., 2019). It’s also probable that the number of vehicles could have been underestimated inside the urban canyon. The underestimation of NO_X is caused by the underestimation of NO concentrations. NO_2 concentration magnitude is well represented by MUNICH.

The diurnal variation of MUNICH simulation, observation, and background concentrations are shown in Fig. 7. MUNICH simulated coherently the temporal variation of O_3 and NO_2 concentration inside the urban canyon. For NO and NO_x, the temporal variation during the day and until midnight is well simulated, while the morning peak at 6 hours is underestimated. After midnight, a higher concentration of NO_X occurs by the increase of heavy-duty vehicles at night that mainly run with diesel. In Pinheiros urban canyons, there is predominantly a flow of light-duty vehicles, even though it is registered high NO_X concentrations that it’s transported from the highway. The mean difference between MUNICH simulation and background concentration for O_3, NO_x, NO, and NO_2 are -13.10 $\mu g\ m^{-3}$, 28.61 $\mu g\ m^{-3}$, 9.25 $\mu g\ m^{-3}$, and 14.43 $\mu g\ m^{-3}$, respectively.



310 **Figure 6. Comparison of MUNICH results against background and observation concentrations of (a) O_3, (b) NO_X, (c) NO, and (d) NO_2 for Pinheiros urban canyon from the control case.**

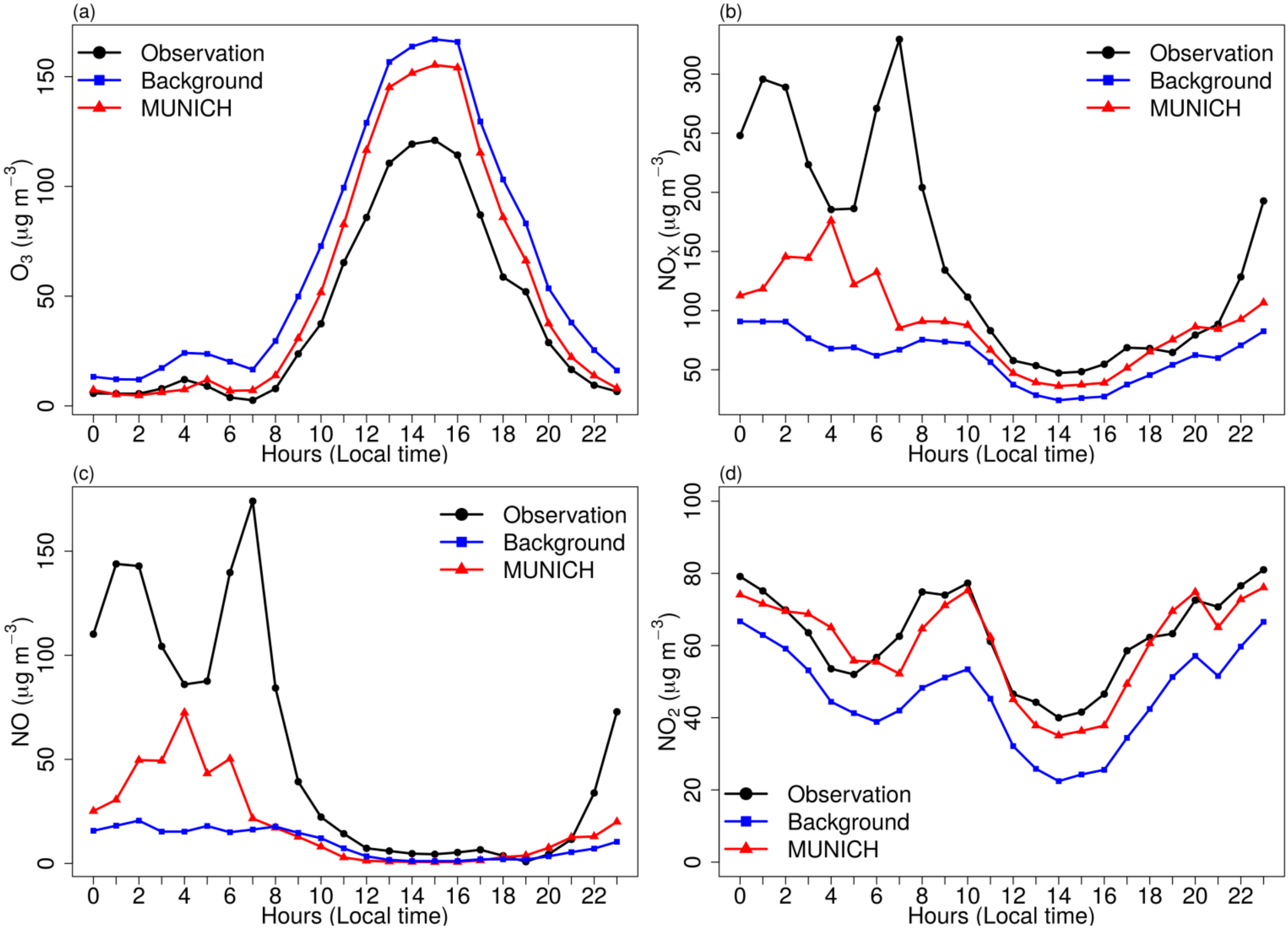


Figure 7. Diurnal profile of MUNICH results, background, and concentrations of (a) O_3, (b) NO_X, (c) NO, and (d) NO_2 for Pinheiros urban canyon from the control case.

3.2 Emission adjustment

We ran different scenarios with increased NO_X and VOCs emission from VEIN. The best results were produced when doubled the NO_X and VOCs emissions, this scenario is called MUNICH-Emiss. With this adjustment, we achieve an overall improvement of MUNICH simulations. Figure 8 shows the new comparison between the model, background concentration, and observations. O_3 is still overpredicted which is caused by the higher value of O_3 background concentration together with a low NO background concentration; nevertheless, the simulated O_3 concentration during night is well represented and daily peaks values are closer to observations.

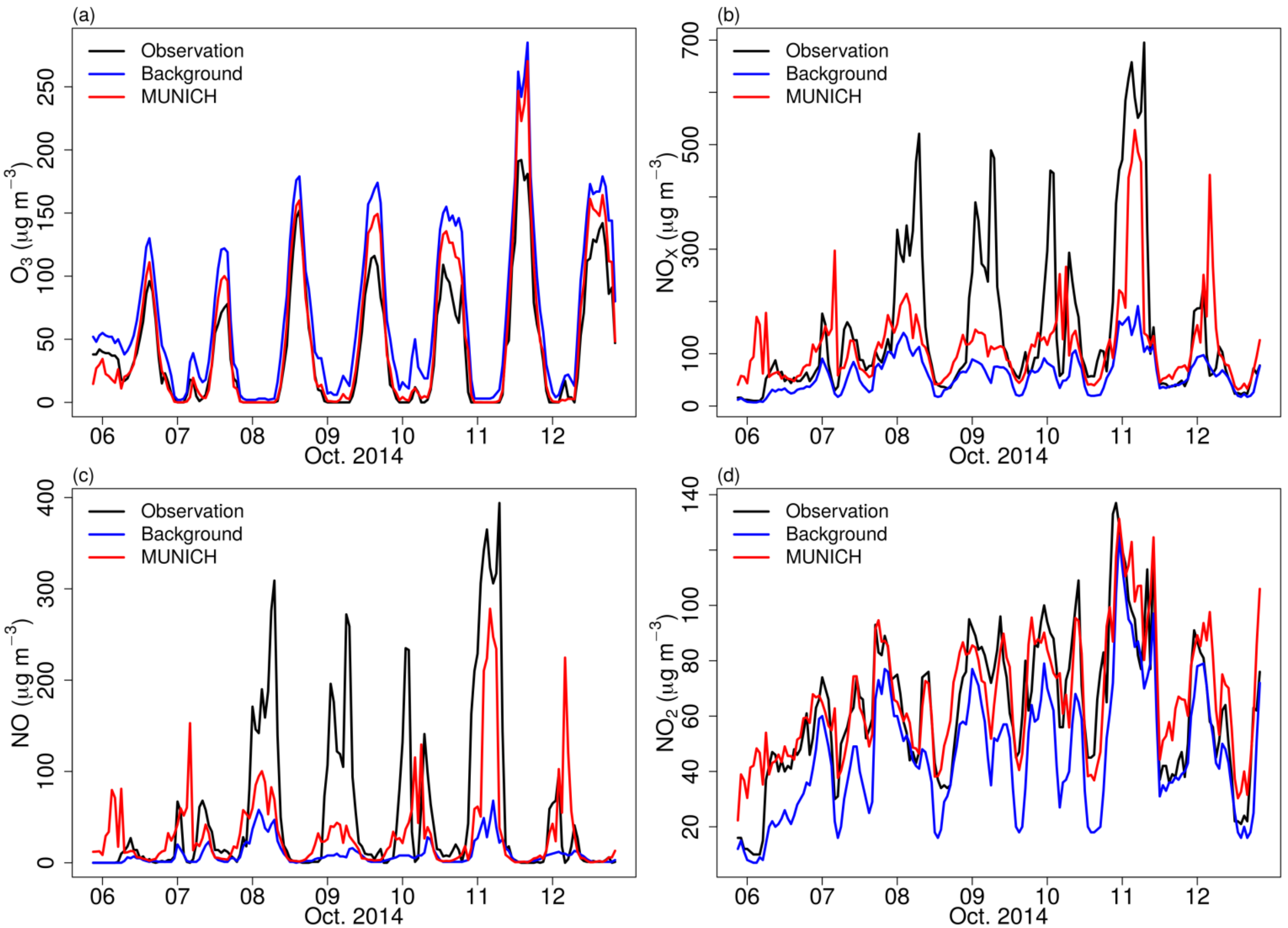


Figure 8. Comparison of MUNICH results against background and observation concentrations of (a) O_3, (b) NO_X, (c) NO, and (d) NO_2 for Pinheiros urban canyon from the MUNICH-Emiss simulation.

NO_x and NO simulations are still underpredicted, but NO_2 is in the same magnitude as observations. NO_x underprediction is still mainly attributed to the underprediction of NO, especially during October 8th, 9th, and 10th where high observational values of NO were recorded. NO underestimation is explained by the lower NO background concentration, the underestimation of emissions, and the use of a single-day emission profile to represent all weekdays. Wind speed overestimation also affects this underestimation as it enhances dispersion. However, MUNICH can better represent the observed high concentration during Saturday 11th, as MUNICH uses the same emission profile for the weekend and weekdays, this high simulated NO concentration resulted from the influence of meteorology.

Figure 9 shows the diurnal profiles for this simulation. The new MUNICH-Emiss profiles are closer to observed concentration profiles, with a better representation of the peak concentrations magnitude of NO_x, NO, and NO_2. The mean difference over the simulation period between simulated and the background concentrations for O_3, NO_x, NO, and NO_2 are -

17.85 μg m^{-3}, -57.26 μg m^{-3}, 23.60 μg m^{-3}, and 21.07 μg m^{-3}, respectively, showing bigger differences than the control case previous scenario and the influence of the reaction with NO emissions.

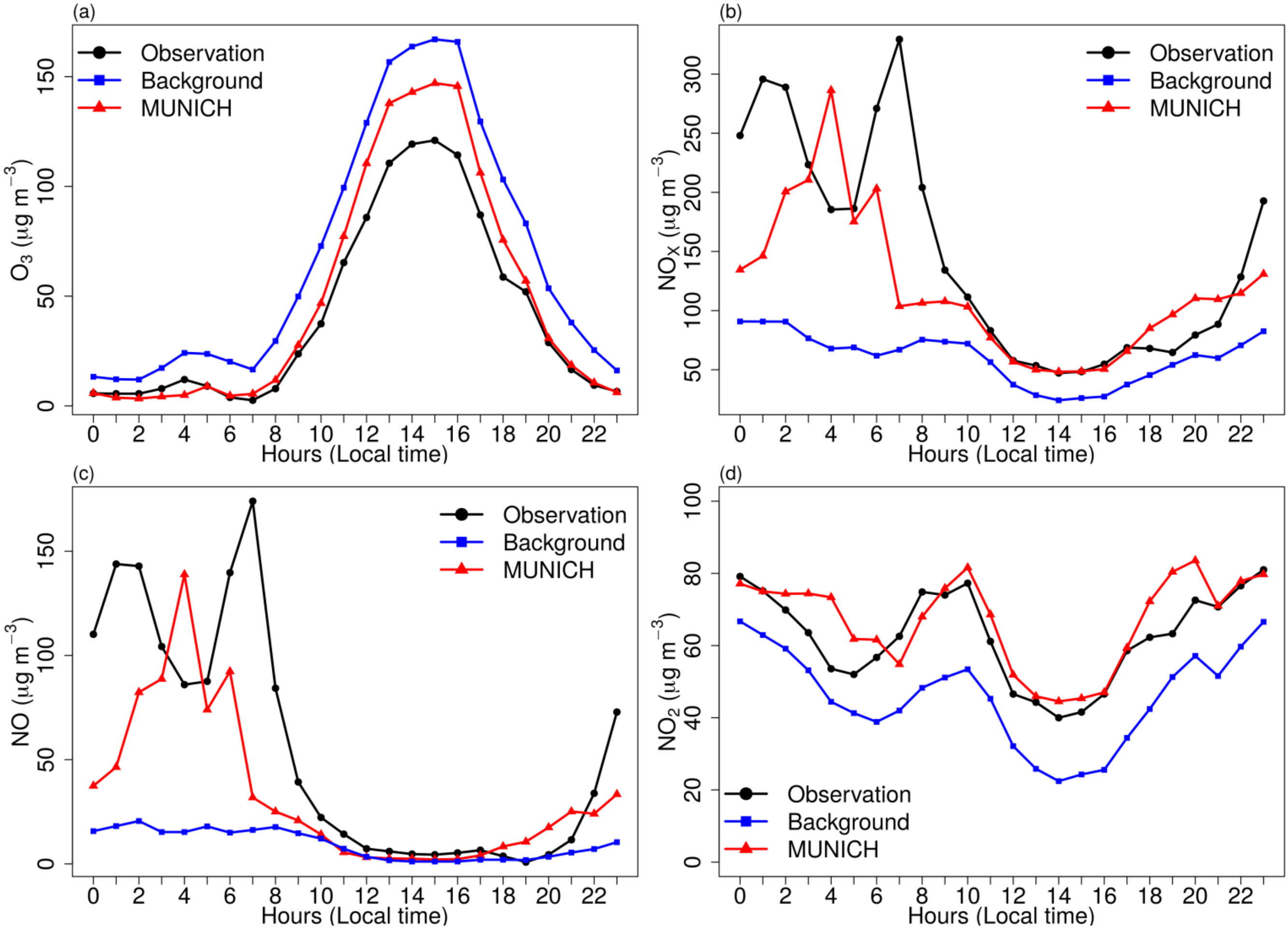


Figure 9. Diurnal profile of MUNICH results, background, and concentration for (a) O_3, (b) NO_X, (c) NO, and (d) NO_2 for Pinheiros urban canyon from the MUNICH-Emiss simulation.

Table 4 summarizes the performance statistics for each scenario and background. The performance statistics from the MUNICH-Emiss case show lower values of MB, NMGE, and RMSE for all pollutants, except NO_2 that presents a slightly increase in these indicators. They also show high values of R (≥ 0.7) for each pollutant in every case, which indicates that the temporal variations of emission and background concentration are in the same phase as the observations. In general, in both MUNICH simulations, NO_2 and O_3 are better simulated. MUNICH-Emiss case performs better and also achieves the recommendations of Hanna and Chang (2012) for O_3, NO_2 NO, and NO_x, whereas MUNICH control case didn't reach these recommendations for NO.

Table 4. Statistical indicators for O_3, NO_X, NO, and NO_2 for comparison between background concentration, MUNICH simulation, and MUNICH-Emiss against observation from Pinheiros AQS.

		$\bar{M}^{b}$	$\bar{O}$	σ_M	σ_O	MB	NMB	NMGE	RMSE	R	|FB|	NMSE	FAC2	NAD
O_3	Background	67.6	41.5	63.2	47.5	26.1	0.6	0.6	32.4	~~1.00~~0.98	**0.5**	**0.4**	**0.5**	**0.2**
	MUNICH	54.5	41.5	62.1	47.5	13.0	0.3	0.3	22.2	~~1.00~~0.98	**0.3**	**0.2**	**0.6**	**0.1**
	MUNICH-Emiss	49.7	41.5	59.5	47.5	8.2	0.2	0.3	18.0	~~1.00~~0.98	**0.2**	**0.2**	**0.6**	**0.1**
NO_X	Background	60.3	146.4	37.3	150.3	-86.0	-0.6	0.6	149.6	~~0.80~~0.79	0.8	**2.5**	**0.5**	**0.4**
	MUNICH	88.9	146.4	57.4	150.3	-57.4	-0.4	0.5	128.5	~~0.70~~0.70	**0.5**	**1.3**	**0.7**	**0.2**
	MUNICH-Emiss	117.6	146.4	85.6	150.3	-28.8	-0.2	0.5	120.0	~~0.60~~0.60	**0.2**	**0.8**	**0.7**	**0.1**
NO	Background	9.5	54.6	12.7	88.9	-45.1	-0.8	0.8	91.5	~~0.80~~0.75	1.4	16.2	**0.3**	0.7
	MUNICH	18.7	54.6	28.7	88.9	-35.9	-0.7	0.8	80.7	~~0.70~~0.70	1.0	6.4	0.1	**0.5**
	MUNICH-Emiss	33.1	54.6	48.5	88.9	-21.5	-0.4	0.8	74.5	~~0.60~~0.60	**0.5**	**3.1**	**0.3**	**0.2**
NO_2	Background	45.8	62.7	23.4	25.9	-16.8	-0.3	0.3	21.2	~~0.90~~0.87	**0.3**	**0.2**	**0.9**	**0.2**
	MUNICH	60.3	62.7	22.8	25.9	-2.4	0.0	0.2	13.3	~~0.90~~0.90	**0.0**	**0.0**	**1.0**	**0.0**
	MUNICH-Emiss	66.9	62.7	22.0	25.9	4.2	0.10	0.2	14.8	~~0.80~~0.80	**0.1**	**0.1**	**0.9**	**0.0**

[b] $\bar{M}$ - Model value mean ($\mu g\ m^{-3}$), $\bar{O}$ - Observation mean ($\mu g\ m^{-3}$), σ_M - model standard deviation ($\mu g\ m^{-3}$), σ_O - observation standard deviation ($\mu g\ m^{-3}$), MB - mean bias ($\mu g\ m^{-3}$), NMB - normalized mean bias, NMGE - normalized mean gross error, RMSE - root mean square error ($\mu g\ m^{-3}$), R - correlation coefficient, FB - fractional mean bias, NMSE - normalized mean-square error, FAC2 - fraction of predictions within a factor of two , and NAD - normalized absolute difference. Values in bold satisfied Hanna and Chang (2012) acceptance criteria.

Figure 10 shows the mean hourly concentration of O_3 and NO_x in the Pinheiros neighborhood, the red diamond points to the location of Pinheiros air quality station. Because the VEIN model can distribute spatially the emissions, there is a variation of concentrations in different street links. For example, the orange diamond shows the location of a traffic light, where traffic jams occur, causing lower O_3 concentrations from higher NO_X emissions.

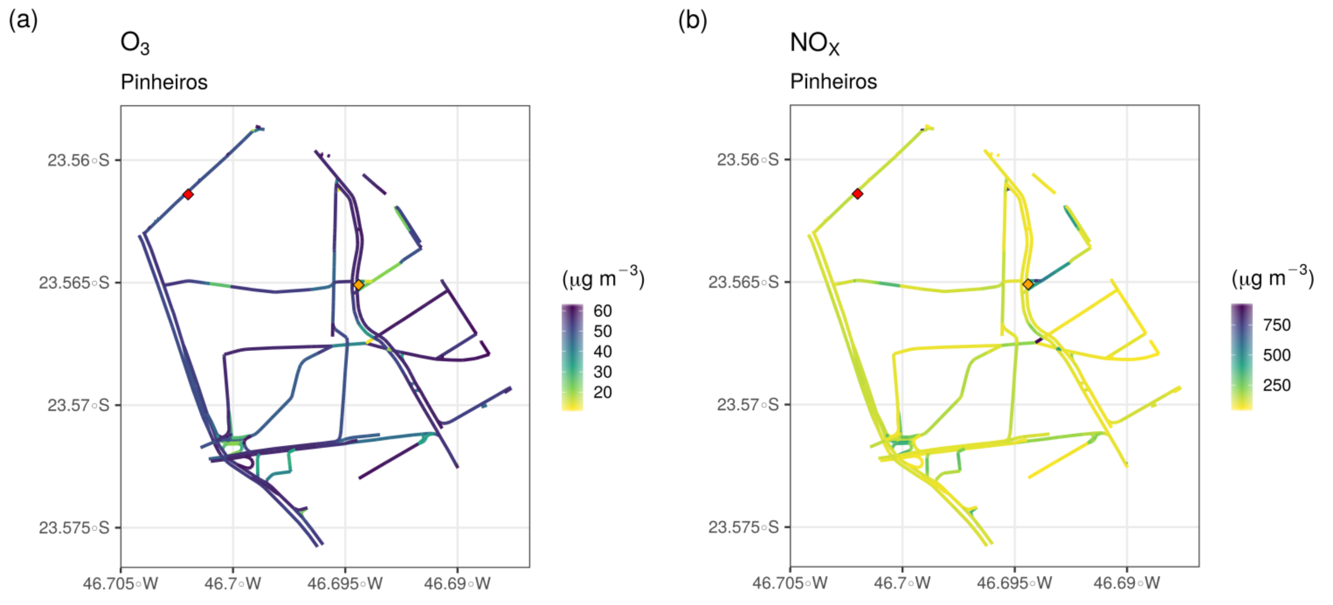


Figure 10. Hourly mean simulated concentration of (a) O_3 and (b) NO_X for Pinheiros neighborhood. Red diamond denotes the location of the Pinheiros AQS and orange diamond denotes traffic light location.

We also perform an additional sensitivity simulation by running MUNICH scenario using the background concentrations from Santos AQS (light blue ~~diamond~~ triangle in Fig. 4). Compared to the Ibirapuera AQS site, measured O_3 and NO_2 concentrations are lower and those of NO concentrations are higher at the Santos AQS. This results in O_3 and NO_2 underprediction and a better simulation of NO concentration magnitude; however, all evaluated pollutants present lower R values and higher NMGE values than MUNICH-Emiss scenario with Ibirapuera AQS as background concentration. Simulated NO_2 and O_3 follow background concentrations, which indicates that the MUNICH simulations have a strong dependence on the background concentration (see Fig. S~~2~~4 and ~~S3~~ Fig. S5 in Supplement).

Lastly, a sensitivity simulation was performed with an only increase of NO_x emission by four and remaining VOCs original emission using Ibirapuera background concentration. This results in a better O_3 representation but unrealistic NO_x, NO, and NO_2 concentration (see Fig. ~~S4~~ S6 and ~~S5~~ Fig. S7 in Supplement). As SPMA has a ~~COV~~VOC-limited regime (Andrade et al., 2017), the increment of NO_X emission ~~with~~ will lead to a reduction of O_3 concentration. Many studies have shown that Sao Paulo atmosphere is VOC-limited (Schuch et al., 2020) due to the high NO_X emission by the heavy-duty that are under old emissions regulations. The new regulations for diesel engine emissions was established recently and are being implemented according to the recycle of the fleet, that is 20 years of use for diesel trucks (CETESB, 2019).

3.3 ~~Aplication~~Application for the Paulista Avenue

The MUNICH simulation is performed with calibrated emissions for a domain that contains a well-defined urban canyon, the Paulista Avenue. The simulation shows a better representation of NO_x, NO, and NO_2 temporal variation and a good representation of concentration magnitude (Fig. 11). Although the MB indicates an overprediction of NO_x, NO, and NO_2 (Table 5), Figure 12 shows that this is caused by an overprediction of these pollutants during night hours, linked to a mismatch of emissions. As in Pinheiros domain, MUNICH did not capture the two peaks of NO and NO_X during nighttime. This is caused by WRF limitation in representing planetary boundary layer height during nighttime (Hu et al., 2012; McNider & Pour-Biazar, 2020). Also, as shown in Fig. 1a, NO_X emission profile during weekday present two peaks during daylight at 7 hours and 16 hours (Local Time), and a smaller emission peak around 23 hours, it is probable that this nighttime peak was underestimated.

Statistics in Table 5 shows an improvement in representing concentration magnitudes of NO_x, NO, and NO_2 with mean simulated concentrations close to observations and very low values of MB, NMB, and RMSE. In this case, R values are lower than those in the Pinheiros case but still higher than 0.4~~5~~ for NO_X and NO_2, confirming that there is a mismatch of simulated concentrations, which is clearer in MUNICH NO_X and NO peak happening before observation. The MUNICH-Emiss simulations achieve Hanna and Chang (2012) performance criteria for NO_x and NO_2. NO_2 is the best simulated species.

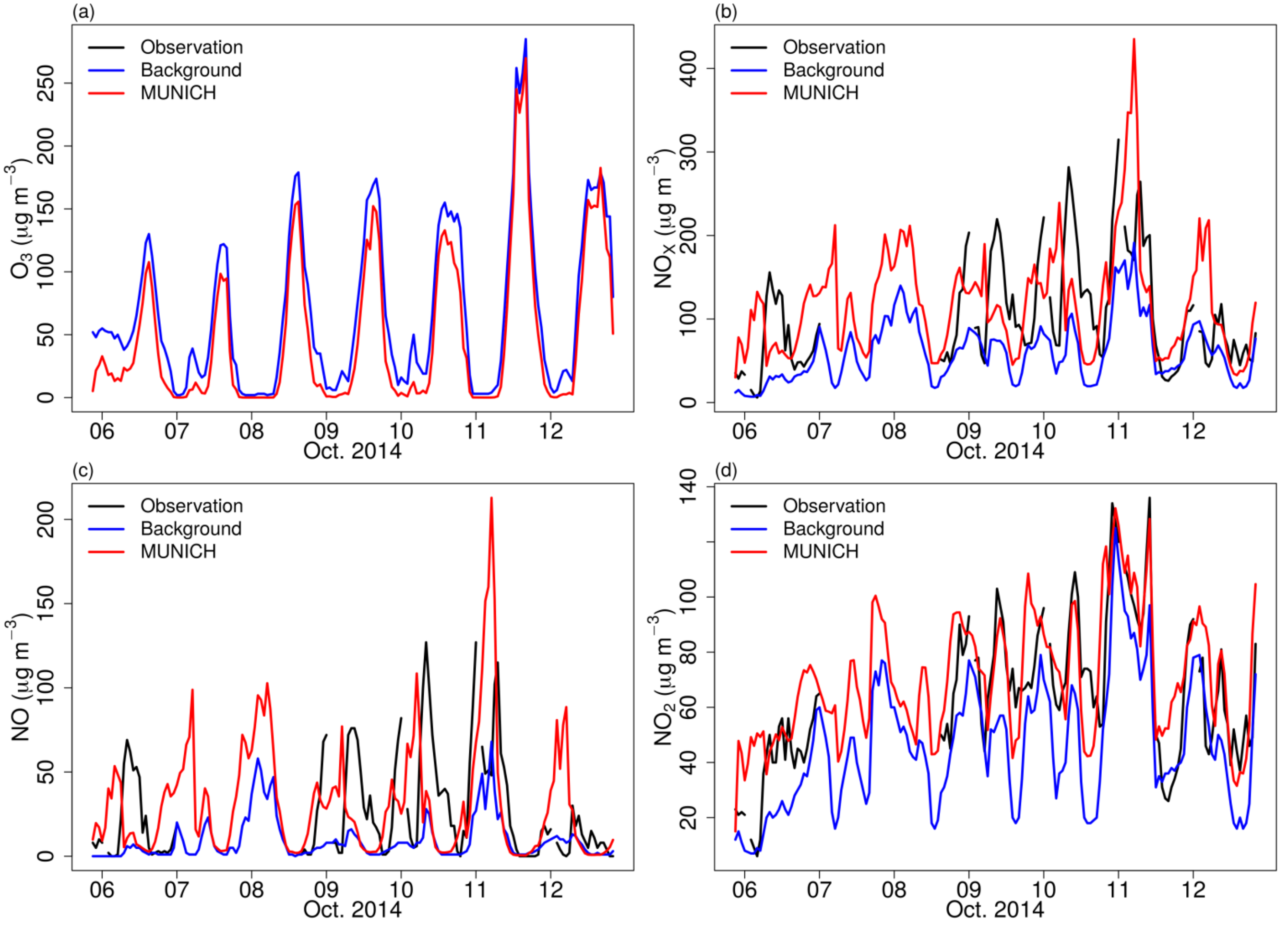


Figure 11. Comparison of MUNICH results against background and observation concentration for (a) O_3, (b) NO_X, (c) NO, and (d) NO_2 for Paulista Avenue urban canyon. Note that O_3 observations were not available ~~O_3 observation~~ for Paulista Avenue domain.

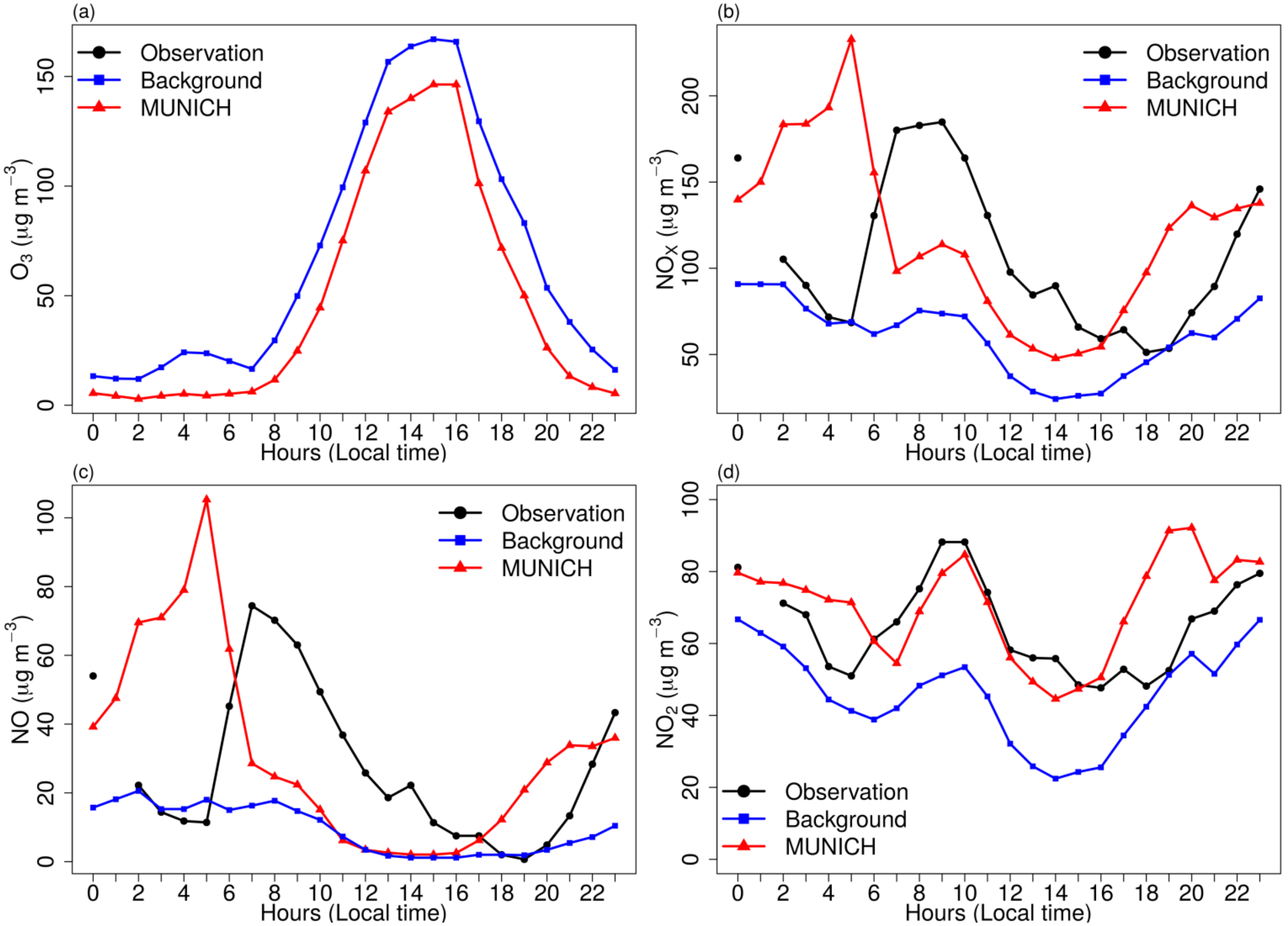


Figure 12. Diurnal profile of MUNICH results, background, and concentration for (a) O_3, (b) NO_X, (c) NO, and (d) NO_2 for Paulista. Note that O_3 observations were not available ~~O_3 observation~~ for Paulista Avenue domain.

Table 5. Statistical indicators for O_3, NO_X, NO, and NO_2 for comparison between background concentration and MUNICH-Emiss against observation from Cerqueira Cesar AQS.

		$\bar{M}$[c]	$\bar{O}$	σ_M	σ_O	MB	NMB	NMGE	RMSE	R	|FB|	NMSE	FAC2	NAD
NO_X	Background	56.8	105.8	36.6	66.8	-49.0	-0.5	0.5	68.9	0.7	**0.6**	0.8	**0.6**	**0.3**
	MUNICH-Emiss	114.8	105.8	68.4	66.8	9.0	0.1	0.6	74.2	0.4	**0.1**	**0.5**	**0.7**	**0.0**
NO	Background	7.3	26.9	10.3	30.7	-19.6	-0.7	0.8	32.5	0.6	1.1	**5.3**	0.2	0.6
	MUNICH-Emiss	28.0	26.9	35.2	30.7	1.1	0.0	1.1	40.8	0.2	**0.0**	2.2	0.2	**0.0**
NO_2	Background	45.5	64.6	24.3	26.5	-19.0	-0.3	0.3	24.2	0.8	**0.3**	**0.2**	**0.8**	**0.2**
	MUNICH-Emiss	71.9	64.6	23.9	26.5	7.4	0.10	0.2	19.1	0.8	**0.1**	**0.1**	**0.9**	**0.1**

[c] $\bar{M}$- Model value mean (μg m^{-3}), $\bar{O}$ - Observation mean (μg m^{-3}), σ_M - model standard deviation (μg m^{-3}), σ_O - observation standard deviation (μg m^{-3}), MB - mean bias (μg m^{-3}), NMB - normalized mean bias, NMGE - normalized mean gross error, RMSE - root mean square error (μg m^{-3}), R - correlation coefficient, FB - fractional mean bias, NMSE - normalized mean-square error, FAC2 - fraction of predictions within a factor of two , and NAD - normalized absolute difference. Values in bold satisfied Hanna and Chang (2012) acceptance criteria.

4 Discussion and conclusions

Simulating air pollutants inside urban street canyons is a challenging task. It is even more difficult in cities as heterogeneous as Sao Paulo, where its urban structure is not always textbook defined. The limited number of air quality stations located inside or near urban canyons, together with the lack of information from detailed emission inventories and urban morphology data, hinder accurate air quality modeling, and consequently the air quality management.

In this paper, we attempt to fill in this gap by using the MUNICH street-network model together with the VEIN vehicular emissions model. The latter provides temporal and spatially detailed emission fluxes inside the main streets and coordinates and width of the streets (i.e., the street network). The urban morphology is completed by extracting the building height from the WUDAPT database for Sao Paulo Metropolitan Area. The advantages of using MUNICH are that, besides solving pollutant dispersion, it also solves photochemistry reactions and it is ~~nature of~~an operational model ~~to~~ that solve pollutant concentration at neighborhood scale considering ~~also~~ street intersections.

Results showed that MUNICH simulations that used adjusted emissions can better represent the temporal variation of O_3, NO_x, NO, and NO_2 concentrations inside urban canyon. Nevertheless, the results are highly dependent on background concentrations and emission fluxes. This background concentration dependence is stronger in secondary pollutants such as O_3, and primary pollutants are more determined by emission fluxes. The reason for the significant contribution of background concentration is that MUNICH is based in SIRANE, and SIRANE also presents a significant contribution from background concentration (Soulhac et al., 2012).

The main cause of O_3 overprediction in our simulation for both tested urban zones is the high value of background O_3 concentration measured in Ibirapuera AQS. In Pinheiros neighboorhood, the underprediction of NO_X concentration is caused by the underprediction of NO concentration in Pinheiros during the second half of the week. This underestimation is caused by the lower NO background concentration together with an emission underestimation. The concentration magnitudes in Paulista Avenue are well represented but there was a mismatching with observed concentration. MUNICH-Emiss scenario ~~with the adjusted emissions~~ fulfills the performance criteria. O_3 concentration simulated in Pinheiros and Paulista Avenue is less than background concentrations, these same results are reported by Wu et al. (2019). As noted in Krecl et al. (2016), this behavior is caused by the high NO_X emissions inside the street urban canyons, which rapidly deplete the formed O_3 and the one from the rooftop (i.e, background concentration).

As the main source of ~~superficial~~ surface NO and NO_2 emissions in São Paulo are vehicles, it is necessary to go deeper into the reasons why the scenario MUNICH-Emiss performs better. The increase of the emissions is necessary because the emissions factors are the average of emission certification tests (CETESB, 2015). It has been shown that emission factors derived from dynamometer and cycle measurements do not represent real-drive emissions (Ropkins et al., 2009). São Paulo does not have an Inspection and Maintenance (I&M) program, therefore, may exist a fraction of the fleet which are high emitters and do not meet the emission standards, more details can be found in Ibarra-Espinosa et al. (2020). Furthermore, the comparison of traffic flow between GPS and TDM data for Pinheiros area showed that TDM traffic flows are 2.22 times higher than GPS. Hence, more representative traffic flows would also improve the emissions compilation. As a conclusion, it is important to develop new and more representative vehicular traffic flow and emission factors for Brazil.

With calibrated emissions (i.e. MUNICH-Emiss scenario), the good performance of MUNICH in representing NO_2 concentrations in both neighborhoods and NO and NO_x in Paulista Avenue urban canyon suggests that VEIN model distributes emissions spatially and temporally efficiently, which proves its potential to be used in other cities. VEIN is being continuously developed and currently offers some utilities to format emissions to the MUNICH model. On the other hand, now Google Earth allows new features as 3D view~~, where information on building height can be retrieved~~that together with in-situ measurements can improve WUDAPT building height estimates. These new features can be used to improve MUNICH input data, and therefore, the model simulation results. Further, a better estimation of background concentrations from photochemical grid models can potentially improve the model performance.

The results obtained show the promising capability of MUNICH to represent the concentrations of pollutants emitted by the fleet close to the streets. As MUNICH uses the CB05 gas-phase mechanism, it can also simulate VOCs inside the urban canyon. Measurements of VOCs inside urban canyons are therefore necessary to validate the model in the future. An accurate prediction of street-scale air pollutant concentrations will enable the future assessment of the impacts on human health due to their exposure to air pollutants emitted by the vehicles.

Appendix A: Statistical indicators

Table A6. Statistical indicator definition.

Statistical indicator	Definition	Reference
Fraction of prediction within a factor of two (FAC2)	$FAC2 = 0.5 \leq \frac{M_i}{O_i} \leq 2.0$	Emery et al (2017)
Mean Bias (MB)	$MB = \frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)$	Emery et al. (2017)
Mean Absolute Gross Error (MAGE)	$MAGE = \frac{1}{N}\sum_{i=1}^{N}\lvert M_i - O_i\rvert$	Emery et al. (2017)
Normalized mean bias (NMB)	$NMB = \frac{\sum_{i=1}^{N}(M_i - O_i)}{\sum_{i=1}^{N} O_i}$	Emery et al. (2017)
Normalized mean error (NME)	$NME = \frac{\sum_{i=1}^{N}\lvert M_i - O_i\rvert}{\sum_{i=1}^{N} O_i}$	Emery et al. (2017)
Root mean square error (RMSE)	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)^2}$	Emery et al. (2017)
Correlation coefficient (R)	$R = \frac{1}{(N-1)}\sum_{i=1}^{N}\left(\frac{M_i - \bar{M}}{\sigma_M}\right)\left(\frac{O_i - \bar{O}}{\sigma_O}\right)$	Emery et al. (2017)
Fractional mean bias (FB)	$FB = 2.0\frac{\overline{O_i - M_i}}{\bar{O} + \bar{M}}$	Hanna and Chang (2012)
Normalized mean-square error (NMSE)	$NMSE = \frac{\overline{(O_i - M_i)^2}}{\bar{O} \times \bar{M}}$	Hanna and Chang (2012)
Normalized absolute difference (NAD)	$NAD = \frac{\overline{\lvert O_i - M_i\rvert}}{\bar{O} + \bar{M}}$	Hanna and Chang (2012)

Data availability. MUNICH input and output data, and scripts to generate the figures and calculations are available on GitHub (https://github.com/quishqa/MUNICH_VEIN_SP) and Zenodo (http://doi.org/10.5281/zenodo.4168056). MUNICH (v1.0) is available on http://cerea.enpc.fr/munich/index.html and Zenodo (http://doi.org/10.5281/zenodo.4168985). VEIN can be installed from CRAN, and it is also available in Zenodo (http://doi.org/10.5281/zenodo.3714187). Additional information and help are available by contacting the authors.

Author contributions. MGC performed the simulations and prepared the manuscript with the support of all co-authors. MGC, MFA and YZ designed the experiment. SIE provided the emissions and street morphology information. YK provided support to set up and run MUNICH. MGC, YZ, MFA, and SIE discussed the results.

Competing interests. The authors declare that they have no conflicts of interest.

Acknowledgements. The authors thank CETESB (São Paulo State Environmental Protection Agency) for providing air pollution and meteorological data, the support from CAPES (Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), and FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo, process 2016/18438-0), and the Wellcome Trust (subaward from Yale University to Northeastern University, subcontract number GR108374).

References

Andrade, M. de F., Kumar, P., de Freitas, E. D., Ynoue, R. Y., Martins, J., Martins, L. D., Nogueira, T., Perez-Martinez, P., de Miranda, R. M., Albuquerque, T., Gonçalves, F. L. T., Oyama, B. and Zhang, Y.: Air quality in the megacity of São Paulo: Evolution over the last 30 years and future perspectives, Atmos. Environ., 159, 66–82, doi:10.1016/j.atmosenv.2017.03.051, 2017.

Andrade, M. de F., Ynoue, R. Y., Freitas, E. D., Todesco, E., Vara Vela, A., Ibarra, S., Martins, L. D., Martins, J. A. and Carvalho, V. S. B.: Air quality forecasting system for Southeastern Brazil, Front. Environ. Sci., 3(February), 1–14, doi:10.3389/fenvs.2015.00009, 2015.

Berkowicz, R., Hertel, O., Larsen, S. E., Sørensen, N. N. and Nielsen, M.: Modelling traffic pollution in streets, Natl. Environ. Res. Institute, Roskilde, Denmark, 10129(10136), 20, doi:10.1287/mnsc.1090.1070, 1997.

Carpentieri, M., Salizzoni, P., Robins, A. and Soulhac, L.: Evaluation of a neighbourhood scale, street network dispersion model through comparison with wind tunnel data, Environ. Model. Softw., 37, 110–124, doi:10.1016/j.envsoft.2012.03.009, 2012.

Carvalho, V. S. B., Freitas, E. D., Martins, L. D., Martins, J. A., Mazzoli, C. R. and Andrade, M. de F.: Air quality status and trends over the Metropolitan Area of São Paulo, Brazil as a result of emission control policies, Environ. Sci. Policy, 47, 68–79, doi:10.1016/j.envsci.2014.11.001, 2015.

CETESB: Emissões veiculares no estado de São Paulo 2014, São Paulo. [online] Available from: https://cetesb.sp.gov.br/veicular/relatorios-e-publicacoes/, 2015.

CETESB: Qualidade do ar no estado de São Paulo 2018, São Paulo. [online] Available from: https://cetesb.sp.gov.br/ar/publicacoes-relatorios/, 2019.

Dominutti, P. A., Nogueira, T., Borbon, A., Andrade, M. de F. and Fornaro, A.: One-year of NMHCs hourly observations in São Paulo megacity: meteorological and traffic emissions effects in a large ethanol burning context, Atmos. Environ., 142, 371–382, doi:10.1016/j.atmosenv.2016.08.008, 2016.

Dowle, M. and Srinivasan, A.: data.table: Extension of “data.frame”. R Package Version 1.12.8, [online] Available from: https://cran.r-project.org/package=data.table, 2019.

Emery, C., Tai, E. and Yarwood, G.: Enhanced meteorological modeling and performance evaluation for two Texas ozone episodes. [online] Available from: https://www.tceq.texas.gov/assets/public/implementation/air/am/contracts/reports/mm/EnhancedMetModelingAndPerformanceEvaluation.pdf, 2001.

Emery, C., Liu, Z., Russell, A. G., Odman, M. T., Yarwood, G. and Kumar, N.: Recommendations on statistics and benchmarks to assess photochemical model performance, J. Air Waste Manag. Assoc., 67(5), 582–598, doi:10.1080/10962247.2016.1265027, 2017.

Fellini, S., Salizzoni, P., Soulhac, L. and Ridolfi, L.: Propagation of toxic substances in the urban atmosphere: A complex network perspective, Atmos. Environ., 198(July 2018), 291–301, doi:10.1016/j.atmosenv.2018.10.062, 2019.

Hanna, S. and Chang, J.: Acceptance criteria for urban dispersion model evaluation, Meteorol. Atmos. Phys., 116(3–4), 133–146, doi:10.1007/s00703-011-0177-1, 2012.

Hong, S.-Y., Noh, Y. and Dudhia, J.: A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes, Mon. Weather Rev., 134(9), 2318–2341, doi:10.1175/MWR3199.1, 2006.

Hu, X. M., Doughty, D. C., Sanchez, K. J., Joseph, E. and Fuentes, J. D.: Ozone variability in the atmospheric boundary layer in Maryland and its implications for vertical transport model, Atmos. Environ., 46, 354–364, doi:10.1016/j.atmosenv.2011.09.054, 2012.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A. and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, J. Geophys. Res. Atmos., 113(13), 2–9, doi:10.1029/2008JD009944, 2008.

Ibarra-Espinosa, S., Ynoue, R. Y., Ropkins, K., Zhang, X. and de Freitas, E. D.: High spatial and temporal resolution vehicular emissions in south-east Brazil with traffic data from real-time GPS and travel demand models, Atmos. Environ., 222(May 2019), 117136, doi:10.1016/j.atmosenv.2019.117136, 2020.

Ibarra-Espinosa, S., Ynoue, R., Giannotti, M., Ropkins, K. and de Freitas, E. D.: Generating traffic flow and speed regional model data using internet GPS vehicle records, MethodsX, 6, 2065–2075, doi:10.1016/j.mex.2019.08.018, 2019.

Ibarra-Espinosa, S., Ynoue, R., O'Sullivan, S., Pebesma, E., Andrade, M. D. F. and Osses, M.: VEIN v0.2.2: an R package for bottom–up vehicular emissions inventories, Geosci. Model Dev., 11(6), 2209–2229, doi:10.5194/gmd-11-2209-2018, 2018.

Keyser, D. and Anthes, R. A.: The Applicability of a Mixed–Layer Model of the Planetary Boundary Layer to Real-Data Forecasting, Mon. Weather Rev., 105(11), 1351–1371, doi:10.1175/1520-0493(1977)105<1351:TAOAMM>2.0.CO;2, 1977.

Kim, Y., Wu, Y., Seigneur, C. and Roustan, Y.: Multi-scale modeling of urban air pollution : development and application of a Street-in-Grid model by coupling MUNICH and, Geosci. Model Dev., (September), 1–24, doi:10.5194/gmd-11-611-2018, 2018.

Krecl, P., Targino, A. C., Wiese, L., Ketzel, M. and de Paula Corrêa, M.: Screening of short-lived climate pollutants in a street canyon in a mid-sized city in Brazil, Atmos. Pollut. Res., 7(6), 1022–1036, doi:10.1016/j.apr.2016.06.004, 2016.

Krüger, E. L., Minella, F. O. and Rasia, F.: Impact of urban geometry on outdoor thermal comfort and air quality from field measurements in Curitiba, Brazil, Build. Environ., 46(3), 621–634, doi:10.1016/j.buildenv.2010.09.006, 2011.

Lemonsu, A., Grimmond, C. S. B. and Masson, V.: Modeling the surface energy balance of the core of an old Mediterranean City: Marseille, J. Appl. Meteorol., 43(2), 312–327, doi:10.1175/1520-0450(2004)043<0312:MTSEBO>2.0.CO;2, 2004.

Lugon, L., Sartelet, K., Kim, Y., Vigneron, J. and Chrétien, O.: Nonstationary modeling of NO2, NO and NO in Paris using the Street-in-Grid model: coupling local and regional scales with a two-way dynamic approach, Atmos. Chem. Phys., 20(13), 7717–7740, doi:10.5194/acp-20-7717-2020, 2020.

McHugh, C. A., Carruthers, D. J. and Edmunds, H. A.: ADMS-Urban: An air quality management system for traffic, domestic and industrial pollution, Int. J. Environ. Pollut., 8(3–6), 666–674, doi:10.1504/IJEP.1997.028218, 1997.

McNider, R. T. and Pour-Biazar, A.: Meteorological modeling relevant to mesoscale and regional air quality applications: a review, J. Air Waste Manag. Assoc., 70(1), 2–43, doi:10.1080/10962247.2019.1694602, 2020.

Monk, K., Guérette, E.-A., Paton-Walsh, C., Silver, J. D., Emmerson, K. M., Utembe, S. R., Zhang, Y., Griffiths, A. D., Chang, L. T.-C., Duc, H. N., Trieu, T., Scorgie, Y. and Cope, M. E.: Evaluation of Regional Air Quality Models over Sydney and Australia: Part 1—Meteorological Model Comparison, Atmosphere (Basel)., 10(7), 374, doi:10.3390/atmos10070374, 2019.

Morrison, H., Thompson, G. and Tatarskii, V.: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes, Mon. Weather Rev., 137(3), 991–1007, doi:10.1175/2008MWR2556.1, 2009.

OpenStreetMap contributors: Planet dump retrieved from https://planet.osm.org , 2017.

Oke, T. R., Mills, G., Christen, A. and Voogt, J. A.: Urban Climates, Cambridge University Press, Cambridge., 2017.

Pebesma, E.: Simple features for R: Standardized support for spatial vector data, R J., 10(1), 439–446, doi:10.32614/rj-2018-009, 2018.

Pebesma, E., Mailund, T. and Hiebert, J.: Measurement units in r, R J., 8(2), 490–498, doi:10.32614/rj-2016-061, 2016.

Pellegatti Franco, D. M., Andrade, M. de F., Ynoue, R. Y. and Ching, J.: Effect of Local Climate Zone (LCZ) classification on ozone chemical transport model simulations in Sao Paulo, Brazil, Urban Clim., 27(December 2018), 293–313, doi:10.1016/j.uclim.2018.12.007, 2019.

Pérez-Martínez, P. J., Miranda, R. M., Nogueira, T., Guardani, M. L., Fornaro, A., Ynoue, R. and Andrade, M. F.: Emission factors of air pollutants from vehicles measured inside road tunnels in São Paulo: case study comparison, Int. J. Environ. Sci. Technol., 11(8), 2155–2168, doi:10.1007/s13762-014-0562-7, 2014.

Pileke, R. A.: Mesoscale Meteorological Modeling, Third., 2013.

R Core Team: A Language and Environment for Statistical Computing, R Found. Stat. Comput., https://www.R-project.org [online] Available from: http://www.r-project.org, 2020.

Reboredo, B., Arasa, R. and Codina, B.: Evaluating Sensitivity to Different Options and Parameterizations of a Coupled Air Quality Modelling System over Bogotá, Colombia. Part I: WRF Model Configuration, Open J. Air Pollut., 04(02), 47–64, doi:10.4236/ojap.2015.42006, 2015.

Ropkins, K., Beebe, J., Li, H., Daham, B., Tate, J., Bell, M. and Andrews, G.: Real-world vehicle exhaust emissions monitoring: review and critical discussion., 2009.

Schuch, D., Andrade, M. D. F., Zhang, Y., Dias de Freitas, E. and Bell, M. L.: Short-Term Responses of Air Quality to Changes in Emissions under the Representative Concentration Pathway 4.5 Scenario over Brazil, Atmosphere (Basel)., 11(8), 799, doi:10.3390/atmos11080799, 2020.

Schulte, N., Tan, S. and Venkatram, A.: The ratio of effective building height to street width governs dispersion of local vehicle emissions, Atmos. Environ., 112, 54–63, doi:10.1016/j.atmosenv.2015.03.061, 2015.

Soulhac, L., Salizzoni, P., Mejean, P., Didier, D. and Rios, I.: The model SIRANE for atmospheric urban pollutant dispersion; PART II, validation of the model on a real case study, Atmos. Environ., 49, 320–337, doi:10.1016/j.atmosenv.2011.11.031, 2012.

Soulhac, L., Salizzoni, P., Cierco, F. X. and Perkins, R.: The model SIRANE for atmospheric urban pollutant dispersion; part I, presentation of the model, Atmos. Environ., 45(39), 7379–7395, doi:10.1016/j.atmosenv.2011.07.008, 2011.

Stewart, I. D. and Oke, T. R.: Local climate zones for urban temperature studies, Bull. Am. Meteorol. Soc., 93(12), 1879–1900, doi:10.1175/BAMS-D-11-00019.1, 2012.

Stewart, I. D., Oke, T. R. and Krayenhoff, E. S.: Evaluation of the "local climate zone" scheme using temperature observations and model simulations, Int. J. Climatol., 34(4), 1062–1080, doi:10.1002/joc.3746, 2014.

Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M. A., Mitchell, K., Ek, M., Gayno, G., Wegiel, J. and Cuenca, R. H.: Implementation and verification of the unified NOAH land surface model in the WRF model, in 20th Conference on weather analysis and forecasting/16th conference on numerical weather prediction, pp. 11–15. [online] Available from: https://www2.mmm.ucar.edu/wrf/users/phys_refs/LAND_SURFACE/noah.pdf, 2004.

Thouron, L., Kim, Y., Carissimo, B., Seigneur, C. and Bruge, B.: Intercomparison of two modeling approaches for traffic air pollution in street canyons, Urban Clim., 27(July 2018), 163–178, doi:10.1016/j.uclim.2018.11.006, 2019.

United Nations: The World 's Cities in 2018. [online] Available from: https://www.un.org/en/events/citiesday/assets/pdf/the_worlds_cities_in_2018_data_booklet.pdf, 2018.

Vardoulakis, S., Fisher, B. E. A., Pericleous, K. and Gonzalez-Flesca, N.: Modelling air quality in street canyons: A review, Atmos. Environ., 37(2), 155–182, doi:10.1016/S1352-2310(02)00857-9, 2003.

Wu, L., Chang, M., Wang, X., Hang, J. and Zhang, J.: Development of a Real-time On-road emission (ROE v1.0)model for street-scale air quality modeling based on dynamic traffic big data, Geosci. Model Dev., (13), 23–40, doi:gmd-13-23-2020, 2020.

Zheng, Y., Alapaty, K., Herwehe, J. A., Del Genio, A. D. and Niyogi, D.: Improving high-resolution weather forecasts using the Weather Research and Forecasting (WRF) model with an updated Kain-Fritsch scheme, Mon. Weather Rev., 144(3), 833–860, doi:10.1175/MWR-D-15-0005.1, 2016.

Zhong, J., Cai, X. M. and Bloss, W. J.: Coupling dynamics and chemistry in the air pollution modelling of street canyons: A review, Environ. Pollut., 214, 690–704, doi:10.1016/j.envpol.2016.04.052, 2016.