Dear reviewer and editor,

Thanks for your important observations, time, and dedication in reviewing this manuscript. We covered all your points as shown below.

Many thanks

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.”

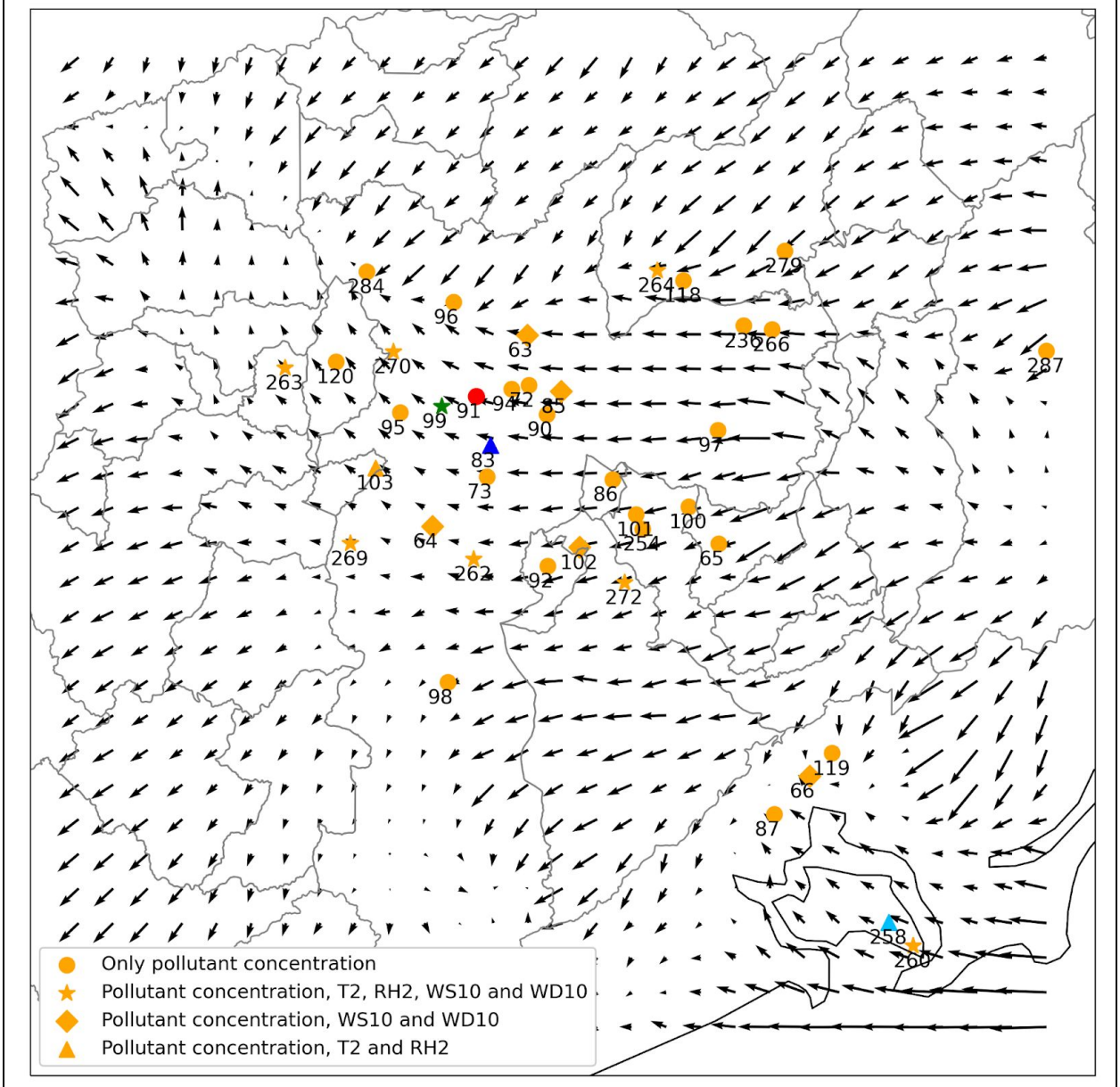

Figure 4. WRF average wind field for the simulation period with CETESB air quality stations (AQS). The green star shows Pinheiros AQS (99), the red circle shows Cerqueira Cesar AQS (91), and the blue triangle shows Ibirapuera AQS (83). Circles represent AQS that only measure pollutant concentrations;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

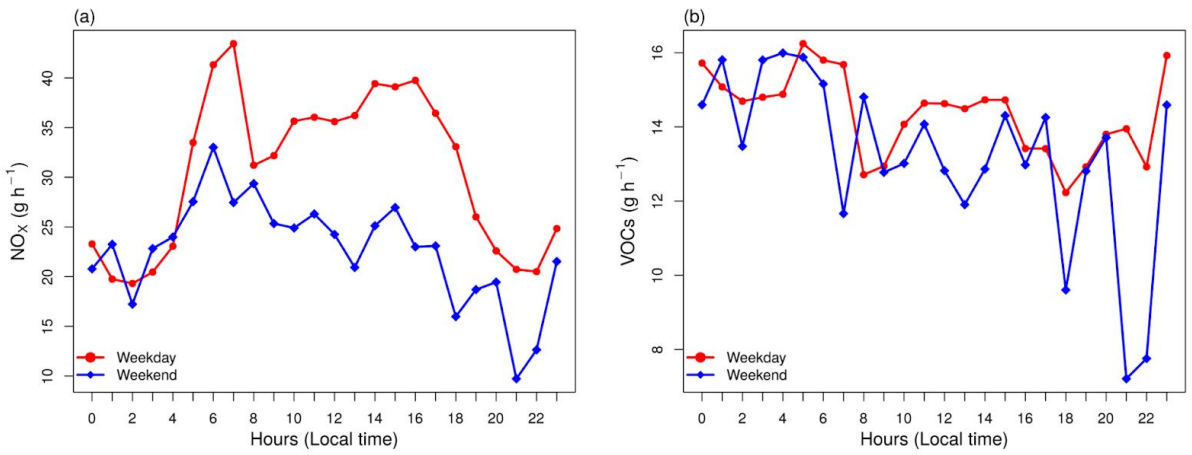

Figure 1. Mean emission from all street links from the Pinheiros neighborhood for (a) NO_X and (b) VOCs for typical weekday and weekend.

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

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

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., 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. 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.

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.

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.