## 1. Comparative evaluation of meteorology

Table S1 summarizes the performance statistics for meteorological parameters over the SPMA, considering the baseline simulation (Case_0) for the 3 km modelling domain. In general, the WRF-Chem simulation captures the daily variations of most of the evaluated parameters reasonably well throughout the study period (see Fig. S1). For temperature and wind speed and direction, the MB and MFB are positive, thereby indicating an overestimation with respect to observations most of the time. Conversely, these same performance statistics, i.e. MB and MFB, are negative for relative humidity, which is coherent since temperature and relative humidity are negatively correlated. In both cases, the smaller MB and MFB indicate bias compensation between under predictions of maximum values and over predictions of minimum values, mostly observed in the second half of the study period when a semi-stationary frontal system was acting close to SPMA. In the case of winds, the WRF-Chem model overall performs better on wind direction compared to wind speed for which less accurate values of R, MFB and MFE are found. Based on the performance statistics and comparing the two sites individually by meteorological parameters, the results for temperature show a better model performance at AF-IAG for most of the statistics; for relative humidity AF-IAG has higher MB and MFB, but slightly better MFE, $RMSE_{UB}$ and R; for wind speed, the model performance is also better at AF-IAG for most of the statistics, except $RMSE_{UB}$ and R, and finally, for wind direction, the results show that the model performs better at INT in terms of MB, MFB and MFE. Large differences in MB and MFB for both wind speed and wind direction is clearly due to overestimations of wind speed at INT and of wind direction at AF-IAG, respectively, as they are poorly reproduced by the model (in full sight from Fig. S1). Since the sites INT and AF-IAG, classified as urban and suburban categories respectively, have different features in terms of land cover, roughness, emissions, etc., then the accuracy of the WRF-Chem model in representing local meteorology (and thus of pollutant concentrations) depends largely on how well these fields are assimilated within the model.

## 2. Comparative evaluation of chemical species

Table S2 presents the performance statistics for gaseous and particulate chemical species over the SPMA, taking also into consideration the baseline simulation for the 3 km modelling domain. Generally, the model captures the temporal variations of all the evaluated PM aerosols, with R greater than 0.5 and $RMSE_{UB}$ less than 12, 8, 3, and 2 μg $m^{-3}$ for $PM_{10}$, $PM_{2.5}$, OC and EC, respectively, but underestimates the observed PM concentrations, with the MB and MFB both negative (note from the Figs. 7, 8 and 14 that the predicted PM concentrations are often lower than their corresponding observed PM concentrations at all sites). Nevertheless, the WRF-Chem performance for both $PM_{10}$ and $PM_{2.5}$ is satisfactory based on the PM model performance criteria proposed by Boylan and Russell (2006), which is defined as "the level of accuracy that is considered to be acceptable for modelling applications" and met when both MFE ≤75% and -60 ≤MFB ≤60%. All of the $PM_{10}$ and $PM_{2.5}$ results are well within the recommended model performance criteria (gray polygon in Fig. S2), indicating that the model is capable of reproducing, with acceptable ranges for bias and error, the observed $PM_{10}$ and $PM_{2.5}$ concentrations. In addition, the MFB and MFE for OC and EC do not meet such criteria showing the difficulty of the model to simulate the mass of these PM compounds. The underestimation of OC, and thus of $PM_{2.5}$ and $PM_{10}$, is related to several factors including underestimation of POA emissions, inaccuracy of SOA formation, inaccurate meteorological predictions, among others. WRF-Chem performs best and worst at INT and NSO, both urban sites located in the southern half and northern SPMA, respectively (see Fig. 1). Compared to $PM_{10}$ performance, the predictions for $PM_{2.5}$ do not show an improved performance with relatively larger MFB and MFE, possibly due to the major complexity of representing the formation of secondary aerosols, which comprise an important fraction of the $PM_{2.5}$. Predicted $PM_{2.5}$ concentrations are identical in both IAG-USP and IPEN-USP as these two sites are sharing the same model grid point for comparison; however, the model performs slightly better at IAG-USP, probably due to its greater proximity to the model grid point that is being used as reference. The worst performance statistics are found for CON. It is worthy to indicate that the sites IPEN-USP and IAG-USP are located in a small green-park (about 7.4 $km^2$) inside the main campus of the University of Sao Paulo in the western SPMA, whereas CON is located in a fully urban area with scarce vegetation in the central region of the city. On the other hand, the model well reproduces the daily variations of $O_3$ (with R ranges from 0.60 for NSO to 0.66 for INT), capturing its decrease during nighttime scavenging periods consistently; however, it is not able to represent adequately some high $O_3$ episodes, clearly underestimating the maximum concentrations, mostly observed in the second half of the study period. MB is negative for most of the sites, ranges from -12.45 for IBI to 10.48 µg $m^{-3}$ for PDP. Finally, the daily cycles of CO and NOx are also reasonably well represented, especially for CO, with maximum concentrations corresponding to the rush hours; however, as found for the other species, the WRF-Chem underestimates the observed concentrations, which is directly related to an underestimation of vehicle emissions as they are, by far, the most important sources of anthropogenic emissions in the SPMA.

## 3. References

Boylan, J. W. and Russell, A. G.: PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models, Atmos. Environ., 40, 4946-4959, 2006.

Tables

Table S1. Performance statistics for WRF-Chem meteorological predictions.

| Meteorology | Site | MB | MFB (%) | MFE (%) | $RMSE_{UB}$ | R |
|---|---|---|---|---|---|---|
| T | AF-IAG | 0.44 | 0.86 | 12.95 | 2.86 | 0.76 |
| | INT | 0.86 | 3.01 | 15.37 | 3.55 | 0.66 |
| RH | AF-IAG | -6.42 | -9.02 | 23.69 | 19.47 | 0.63 |
| | INT | -5.05 | -6.88 | 23.98 | 20.64 | 0.60 |
| WS | AF-IAG | 0.31 | 26.76 | 70.62 | 1.16 | 0.37 |
| | INT | 0.77 | 55.65 | 71.61 | 0.89 | 0.44 |
| WD | AF-IAG | 42.38 | 40.27 | 57.90 | 72.99 | 0.46 |
| | INT | 19.85 | 23.04 | 50.89 | 85.78 | 0.40 |

Table S2. Performance statistics for WRF-Chem chemical predictions.

| Species | Site | MB | MFB (%) | MFE (%) | $RMSE_{UB}$ | R |
|---|---|---|---|---|---|---|
| $PM_{10}$ | NSO | -15.58 | -40.27 | 43.55 | 16.04 | 0.40 |
| | SAN | -16.05 | -40.60 | 40.72 | 11.16 | 0.78 |
| | PDP | -15.60 | -40.96 | 41.05 | 10.55 | 0.75 |
| | MOO | -13.24 | -33.63 | 36.24 | 11.23 | 0.59 |
| | CCE | -8.90 | -27.48 | 31.23 | 8.12 | 0.80 |
| | IAG-USP | -9.13 | -24.33 | 29.15 | 10.83 | 0.76 |
| | IBI | -15.12 | -39.75 | 40.75 | 10.46 | 0.79 |
| | CON | -19.02 | -52.85 | 53.11 | 10.39 | 0.70 |
| | SAM | -14.01 | -41.33 | 42.12 | 9.02 | 0.77 |
| | INT | -13.45 | -40.67 | 41.06 | 8.06 | 0.82 |
| $PM_{2.5}$ | IAG-USP | -7.82 | -40.80 | 41.62 | 6.26 | 0.72 |
| | IPEN-USP | -8.75 | -50.07 | 50.07 | 6.79 | 0.81 |
| | CON | -9.94 | -52.00 | 52.00 | 7.43 | 0.65 |
| OC | IAG-USP | -3.94 | -69.57 | 69.57 | 2.74 | 0.70 |
| EC | IAG-USP | -1.62 | -62.56 | 65.10 | 1.66 | 0.52 |
| $O_3$ | NSO | 4.15 | 41.51 | 82.33 | 28.55 | 0.60 |
| | PDP | 10.48 | 64.02 | 88.12 | 25.49 | 0.62 |
| | MOO | -3.19 | 7.00 | 59.88 | 25.53 | 0.63 |
| | IPEN-USP | -1.79 | 27.22 | 76.53 | 29.60 | 0.63 |
| | IBI | -12.45 | -18.66 | 59.98 | 29.94 | 0.62 |
| | INT | -2.30 | 14.66 | 70.27 | 25.61 | 0.66 |
| NOx | IPEN-USP | 7.06 | 67.49 | 94.52 | 33.40 | 0.43 |
| | IBI | -24.57 | -42.13 | 71.11 | 27.30 | 0.40 |
| CO | IPEN-USP | -0.14 | 4.98 | 83.88 | 0.59 | 0.51 |
| | IBI | -0.39 | -70.04 | 77.97 | 0.54 | 0.57 |

## List of figures

Figure S1. The observed and predicted hourly variations of temperature, relative humidity, and wind speed and direction at two sites in the SPMA for the 3 km modelling domain.

Figure S2. Mean fractional bias (MFB) and mean fractional error (MFE) of different daily-average PM variables: $PM_{10}$, $PM_{2.5,}$ OC and EC. Each point on the scatter plot, displayed with a marker (PM variable) and a color (site), represents the WRF-Chem performance considering the criteria proposed by Boylan and Russell (2006) (gray polygon in the figure).