



Air quality modeling intercomparison and multi-scale ensemble chain for Latin America

Jorge E. Pachón¹, Mariel A. Opazo², Pablo Lichtig³, Nicolas Huneeus², Idir Bouarar⁴, Guy Brasseur⁴, Cathy W. Y. Li⁴, Johannes Flemming⁵, Laurent Menut⁶, Camilo Menares², Laura Gallardo², Michael Gauss⁷, Mikhael Sofiev⁸, Rostilav Kouznetsov⁸, Julia Palamarchuk⁸, Laura Dawidowski³, Nestor Y Rojas⁹, Maria de Fatima Andrade¹⁰, Mario E. Gavidia-Calderón¹⁰, Alejandro H. Delgado Peralta¹⁰, and Daniel Schuch^{10,11}

¹Department of Environmental Engineering, Universidad de La Salle, Bogotá, 111711, Colombia

²Department of Geophysics, Universidad de Chile, Santiago, 8320000, Chile

³Consejo Nacional de Energía Atómica – CNEA, Buenos Aires, C1429BNP, Argentina

⁴Max Planck Institute for Meteorology, Hamburg, 20146, Germany

⁵European Centre for Medium-Range Weather Forecasts– ECMWF, Bonn, 53175, Germany

⁶Laboratoire de Météorologie Dynamique, Palaiseau, 91128, France

⁷Norwegian Meteorological Institute, Oslo, 0313, Norway

⁸Finnish Meteorological Institute, Helsinki, FI-00560, Finland

⁹Department of Chemical and Environmental Engineering, Universidad Nacional de Colombia, Bogotá, 111321, Colombia

¹⁰Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidad de São Paulo, São Paulo, 05508-09B, Brazil

¹¹Civil and Environmental Engineering, Northeastern University, Boston, 02115, USA

Correspondence: Guy Brasseur (guy.brasseur@mpimet.mpg.de)

Abstract. A multi-scale modeling ensemble chain has been assembled as a first step towards an Air Quality forecasting system for Latin America. Two global and three regional models were tested and compared over a shared domain (120W-28W, 60S-30N) to simulate January and July of 2015. Observations from local air quality monitoring networks in Colombia, Chile, Brazil, México, Ecuador and Peru were used for model evaluation. The models generally agreed with observations in large cities such as México City and São Paulo, whereas representing smaller urban areas, such as Bogotá and Santiago, was more challenging. For instance, in Santiago, during wintertime, the simulations showed large discrepancies with observations. No single model had the best performance among pollutants and sites available. Ozone and NO₂ were reproduced better than other pollutants across sites whereas SO₂ was the most difficult. The ensemble, created from the median value of the individual models, was evaluated as well. In some cases, the ensemble showed better results over the individual models and mitigated the extreme over- or underestimation of certain models, demonstrating the potential to establish an analysis and forecast system for Latin America. This study identified certain limitations in the models and global emissions inventories, which should be addressed with the involvement and experience of local researchers.

1 Introduction

Latin America has some of the most populated urban areas in the world, notably, México City and São Paulo have populations exceeding 20 million, while Lima, Bogotá, Rio de Janeiro, and Buenos Aires have more than 10 million inhabitants each



(Nations, 2018). These densely populated regions often experience air pollution events due to large emission sources and due to atmospheric conditions. Other major cities, such as Santiago and Medellin, with a population of ~ 7 and ~ 3.5 million, respectively, are also affected by poor air quality. This urban air pollution not only has long lasting effects on the health of the population but also has a significant negative impact on the environment (Busch et al., 2023; Gouveia et al., 2018; Rodríguez-Villamizar et al., 2018; Romieu et al., 2012).

To better understand the causes of air pollution events in Latin America, it is important to consider the local emission sources. In addition to the usual urban pollution sources (e.g., industrial facilities, residential heating, energy production, and transportation sectors), plumes from biomass burning and long-range dust transport can occasionally reach major cities. In northern South America, increased pollution levels in the dry season have been associated with wildfires (Ballesteros-Gonzalez et al., 2020; Casallas et al., 2023; Mendez-Espinosa et al., 2019) and dust from the Sahara Desert (Mendez-Espinosa et al., 2020). The latter source also affects the Caribbean and central México in early spring (Kramer and Kirtman, 2021; Ramírez-Romero et al., 2021).

Air quality management in Latin America and the Caribbean (LAC) has been traditionally focused on surveillance and building emission inventories (Franco et al., 2019). Modeling activities for LAC are less frequent than North America, Europe, or Asia, mainly due to limited computing resources and scarce information of emission sources. Furthermore, LAC has other challenges: complex terrain where cities are situated in the valleys and canyons of the Andes, varying meteorological conditions due to their proximity to mountains and coastlines, deep convection in the tropics, extensive biomass burning in the Orinoco and Amazonian basins, and the presence of densely populated megacities and urban areas, among others. Despite limitations for applying air quality models in LAC, regional models in the literature have been successfully implemented.

The coupled Aerosol and Tracer Transport model to the Brazilian development of the Regional Atmospheric Modeling System (CCATT-BRAMS) was developed in the region (Longo et al., 2013) to investigate the impact of the Amazonian wildfires on air quality in major Brazilian cities (Pereira et al., 2011; Freitas et al., 2011). The North American Community Multiscale Air Quality Model (CMAQ), coupled with the Weather Research and Forecasting (WRF) meteorological model, has been used in Colombia and Brazil to predict pollutant concentrations and assess reduction strategies (Albuquerque et al., 2019; East et al., 2021; Perez-Peña et al., 2017; Nedbor-Gross et al., 2018; Pachon et al., 2018). The WRF model coupled with Chemistry (WRF-Chem) online has been actively used to study the impact of regional sources on air quality in urban centers across Colombia (Ballesteros-Gonzalez et al., 2020, 2022; Casallas et al., 2024; Mendez-Espinosa et al., 2019; Gonzalez et al., 2018) and São Paulo (Gavidia-Calderon et al., 2024). CHIMERE has been applied in Chile to assess pollutant chemical transformation and dispersion as well as emission reduction strategies (Lapere, 2018; Lapere et al., 2021; Mailler et al., 2017). Additionally, CAMS reanalysis data has been compared against air quality observations, observing well-captured temporal trends for PM_{10} , $PM_{2.5}$ and SO_2 but not for NO_x (Casallas et al., 2024).

This work presents the first model intercomparison and ensemble construction for Latin America, which was assembled under the Prediction of Air Pollutants in Latin America (PAPILA project (<https://papila-h2020.eu/papila>)). The aim of the project was to develop an air quality analysis and forecast system for the region with increasing capabilities in major cities.



50 This work is the first step towards such a system and seeks to examine the differences between the models in diagnostic mode to get an improved forecasting set-up.

2 Methodology

The model intercomparison and construction of the ensemble required relevant activities, such as: the execution of global and regional models in a common domain, harmonization of the model output, ensemble construction, collection of air quality
55 observations, analysis of temporal and spatial variability, and model evaluation.

2.1 Description of the models and modeling set-up

For the model intercomparison, two global models (CAMS and SILAM) and three regional models (CHIMERE, WRF-Chem, EMEP MSC-W) were selected based on the expertise of the research groups working on the PAPILA project (Table 1). WRF-Chem was implemented by two different groups, the Max Planck Institute for Meteorology (MPIM) in Germany and the
60 University of São Paulo (USP) in Brazil, with different set-ups. The different models are briefly described in the following paragraphs.

The Copernicus Atmosphere Monitoring Service (CAMS) provides state-of-the-art global atmospheric composition data based on the IFS (Integrated Forecasting System) model of the European Centre for Medium-Range Weather Forecasts (ECMWF) (Inness et al., 2019). The chemical mechanism of IFS is an extended version of the Carbon Bond 2005 (CB05)
65 and complements the aerosol module (Flemming et al., 2017). The CAMS reanalysis data used for this project is a combination of satellite observations of atmospheric composition and the IFS modeling setup. Anthropogenic emissions from the MACC/CityZen (MACCity) inventory and biomass burning emissions from the Global Fire Assimilation System (GFAS) were used in the simulations (Table 1). The biogenic emissions were simulated off-line by the MEGAN2.1 model (Guenther et al., 2006). CAMS has been extensively evaluated against ozone soundings, aircraft profiles, surface observations, and global satellite
70 retrievals (Flemming et al., 2015).

The system for Integrated modelling of Atmospheric composition (SILAM, <http://silam.fmi.fi>) is a chemical transport model for global-to-local simulations of atmospheric composition and air quality developed at Finish Meteorological Institute (FMI) (Sofiev, 2002; Kouznetsov and Sofiev, 2012; Sofiev et al., 2010, 2006, 2015). For PAPILA, the SILAM simulations were driven by the meteorological IFS model of ECMWF. Anthropogenic emissions were adopted from the CAMS global emission
75 inventory, whereas the biomass burning emissions were generated by the Integrated Monitoring and Modeling System for Wildland fires (IS4FIRES) (Sofiev et al., 2009; Soares and Sofiev, 2014). The biogenic emissions were simulated off-line by the MEGAN2.1 model (Guenther et al., 2006) (Table 1). The model has been extensively evaluated in numerous international retrospective studies (Marecal et al., 2015; Kukkonen et al., 2012; Blechschmidt et al., 2020; Petersen et al., 2019) and real-time operational applications. SILAM is included in the regional European forecasting system provided by CAMS together
80 with CHIMERE, EMEP MSC-W and eight other models (Colette et al., 2020).



Table 1. Description of the models included in the ensemble.

Institution - Model	Model type	Vertical resolution	Projection	IC - BC	Chemical mech - Aerosol model	Meteo- rology	Emissions
FMI - SILAM	Global of- fline v5.7	25 variable layers	lat-lon	Global model	chemistry: gas-phase CBM-4	C-IFS ECMWF	CAMS-REG-AP v3.1 2016, TNO- MACC, IS4FIRES
ECMWF - CAMS	Global	60 levels		Global model	chemistry: CB05 - aerosol: AER bulk	C-IFS	CAMS-REG-AP v3.1 2016. MAC- CITY, MEGAN, GFAS
LMD & UCL - CHIMERE	Regional offline V2017r4.2	8 variable levels up to 500 hPa	Lambert conformal	LMDz- INCA	chemistry: SAPRC-07-A - aerosol: GOCART	C-IFS	EDGAR-HTAP, MEGAN 2.1, min- eral dust, sea-salt.
UCL - EMEP	Regional offline	20 layers up to 100hPa	lat-lon	Global CAMS	aerosol: MARS	C-IFS ECMWF	EDGAR-HTAP Included biogenic emissions, FINN 1.0
MPIM - WRFChem	Regional online V4.1.2		Mercator	GFS - CAM- Chem	chemistry: MOZART - aerosol: GOCART	WRF	CAMS 4.2, MEGAN 2.1, FINN 1.5
USP - WRFChem	Regional online V3.9.1	35 vertical levels	Mercator	GFS 0.25	chemistry: MOZART - aerosol: GOCART	WRF	CAMS, MEGAN, FINN 1.5

Abbreviations: FMI – Finnish Meteorological Institute, ECMWF – European Center for Weather and Modeling Forecast, LMD – Laboratoire de Météorologie Dynamique, MPIM – Max Planck Institute for Meteorology, UCL – University of Chile, USP – University of São Paulo

CHIMERE is a Eulerian chemistry-transport model (CTM) and multi-scale from hemispheric to urban resolutions (Menut et al., 2021; Mailler et al., 2017). The model can be used in offline or online mode and has meteorology forcing from the IFS model by the ECMWF data sets. The biogenic emissions were simulated off-line by the MEGAN2.1 model (Guenther et al., 2006). The model is used in research institutes and in operational centers for forecasting mainly in France and other European countries. In Latin America, CHIMERE has been widely used in Chile to assess pollutant chemical transformation and dispersion as well as emission reduction strategies (Lapere et al., 2021; Mailler et al., 2017; Lapere, 2018). As previously discussed, CHIMERE is also included in the CAMS forecasting ensemble.

The EMEP MSC-W model ('EMEP model' hereafter) is an offline chemical transport model developed at the Norwegian Meteorological Institute (MET Norway). It is used to simulate photo-oxidants as well as organic and inorganic aerosols in scales ranging from local to global scales (Simpson et al., 2012). This model also has meteorological forcing from the IFS model of the ECMWF. Emissions from forest and vegetation fires are taken from the FINN module (Wiedinmyer et al., 2011). The EMEP model has for several decades been the main tool for underpinning air quality policies under the UN ECE convention on long-range trans-boundary air pollution and it is also included in the CAMS regional ensemble. However, it should be noted that the runs for this study were the very first EMEP model simulations ever conducted on a regional scale for LAC and should



95 thus be considered only as a first demonstration of model capabilities. For PAPILA, the EMEP model was run by the modeling team at the University of Chile in Santiago with some support by MET Norway.

The WRF-Chem is the Weather Research and Forecasting (WRF) model coupled with Chemistry, developed at the National Center for Atmospheric Research (NCAR) with the purpose of simulating urban- to regional-scale fields of trace gasses and particulates. The air quality and meteorological components share the same transport and physics scheme, as well as horizontal and vertical grid (Fast et al., 2006; Grell et al., 2005).

CHIMERE, IFS, EMEP, WRF-Chem, LOTOS-EUROS and SILAM models are used in an ensemble mode to configure the MarcoPolo-Panda prediction system in Asia (Brasseur et al., 2019; Petersen et al., 2019) It has been observed that, under specific circumstances, a model ensemble can outperform individual models, demonstrating the potential benefits of this approach. With the desire to replicate the experience in Latin America, the selected models were applied in a common domain, defined by the south-eastern corner at 119°54'W 59°54'S, and the north-eastern corner at 28°6'W 29°54'N. The models were run at a spatial resolution of $\sim 0.2^\circ \times 0.2^\circ$ ($\sim 20 \times 20$ km). Input meteorology and emissions were up to the modeling group (Table 1). The simulation period covers January (southern hemisphere summer) and July (southern hemisphere winter) of 2015.

2.2 Model Evaluation

The models' performance was assessed by comparing the simulated concentrations with the average of the observations for each available city, pollutant, and considered period. For every city and pollutant, the simulated concentration was estimated as the weighted average of the modeled grid cells that intersected with a city's polygon that encompasses the geographical boundaries. The weights were based on the area of the modeled grid cell that overlapped with the city's polygon. The observation's average was constructed by computing the arithmetic mean of all air quality stations available in the network within the city's polygon. This approach was chosen given the objective to assess model performance in cities, rather than for each air quality station. The model evaluation was focused on nitrogen dioxide (NO₂), ozone (O₃), carbon monoxide (CO), Sulphur dioxide (SO₂), and particulate matter less than 2.5 microns (PM_{2.5}).

For each period, pollutant and city, the model evaluation included the following metrics: Model/Observations ratio, mean bias (BIAS), modified normalized bias (MNBIAS), root mean square error (RMSE), fractional gross error (FGE) and correlation coefficient (R). The formulas were replicated from the MarcoPolo-Panda project (Petersen et al., 2019) and are presented in Table A1. These evaluation metrics were computed for all models and the ensemble.

2.3 Air quality monitoring networks in Latin America

Several air quality monitoring networks (AQMN) are available throughout Latin America, especially in major cities. However, worldwide access to the datasets can be difficult due to language barriers and the lack of a centralized platform. A comprehensive list of AQMN in Latin America was assembled for the PAPILA project (<https://papila-h2020.eu/observations>). For the year 2015, we collected air quality data for 12 cities in México, Colombia, Ecuador, Perú, Chile, Brazil, and Uruguay. For all AQMN, a filter ensuring 75% completeness of the air quality database was applied before selecting a site and calculating the city average of the observations, resulting in eight cities with enough data to use for this study. We focus in this study on the



four major cities (from North to South): México City, Bogotá, São Paulo and Santiago. However, data of all available cities were used in the model evaluation. The location of air quality stations in each city is shown in Figure 1.

130 3 Results

Simulated concentrations of all pollutants from all models were compared against observations from every city and for both periods (January and July) in 2015. In this section, we present results from the model evaluation, the spatial and temporal variability of simulate fields and the impact of large versus small urban areas in the model intercomparison.

3.1 Model evaluation

135 The following results are presented for every pollutant: analysis of observations from AQMN, simulated concentrations by the models, comparison of evaluation metrics, discussion of model performance and analysis of model inter-variability.

3.1.1 Nitrogen dioxide - NO₂

Observations

The number of stations per city recording NO₂ is available in Appendix B. In all cities the data availability was 100%. The
140 highest daily average concentration of NO₂ is observed in Santiago during winter at around 40 ppb (Figure 2). This can be attributed to adverse meteorological conditions and emissions from transportation and residential combustion in the surrounding municipalities (Mazzeo et al., 2018; Saide et al., 2016). whereas in the summer NO₂ levels fall to 11 ppb. The second largest values are shown in México City and São Paulo with a daily average NO₂ levels of 24 and 20 ppb respectively, due to the heavy use of fossil fuels in transportation and power generation. The lowest levels of NO₂ are measured in Bogotá with 15 ppb
145 on average.

Model performance

In Bogotá and Santiago, NO₂ is underestimated by the models (Figure 2). In Santiago, the mean of the models is 7.3 ppb in summer and 17.7 ppb in winter, much lower than the mean of the observations. Similarly, in Bogotá the mean of the modeled values is 5.5 ppb. In contrast, in São Paulo and México City, the model fields are above and below the observed concentrations
150 and the average of the modeled values (19.5 ppb and 27 ppb respectively) are in the same order of magnitude of observations.

In São Paulo and México City, the Model/Observations ratios for NO₂ varied from 0.13 to 1.7, indicating that some models overestimate the observations while others show underestimations. These cities exhibit the lowest MNBIAS, FGE and RMSE values (Table A2). The correlation between the models and observations hovers around 0.7, which is larger than the goal benchmark proposed for this pollutant ($r \geq 0.6$) (Zhai et al., 2024). The adequate performance in São Paulo and México City
155 may be attributed to an accurate portrayal of the temporal and spatial variability that is achieved in large urban areas like these (>3500 km²) which encompass at least nine model cells (20 kmx20 km).

In Santiago, Model/Observations ratios range from 0.1 to 0.9. The MNBIAS is consistently negative during both seasons for all models; however, the degree of which the models are underestimating the observations is notably higher in the winter than

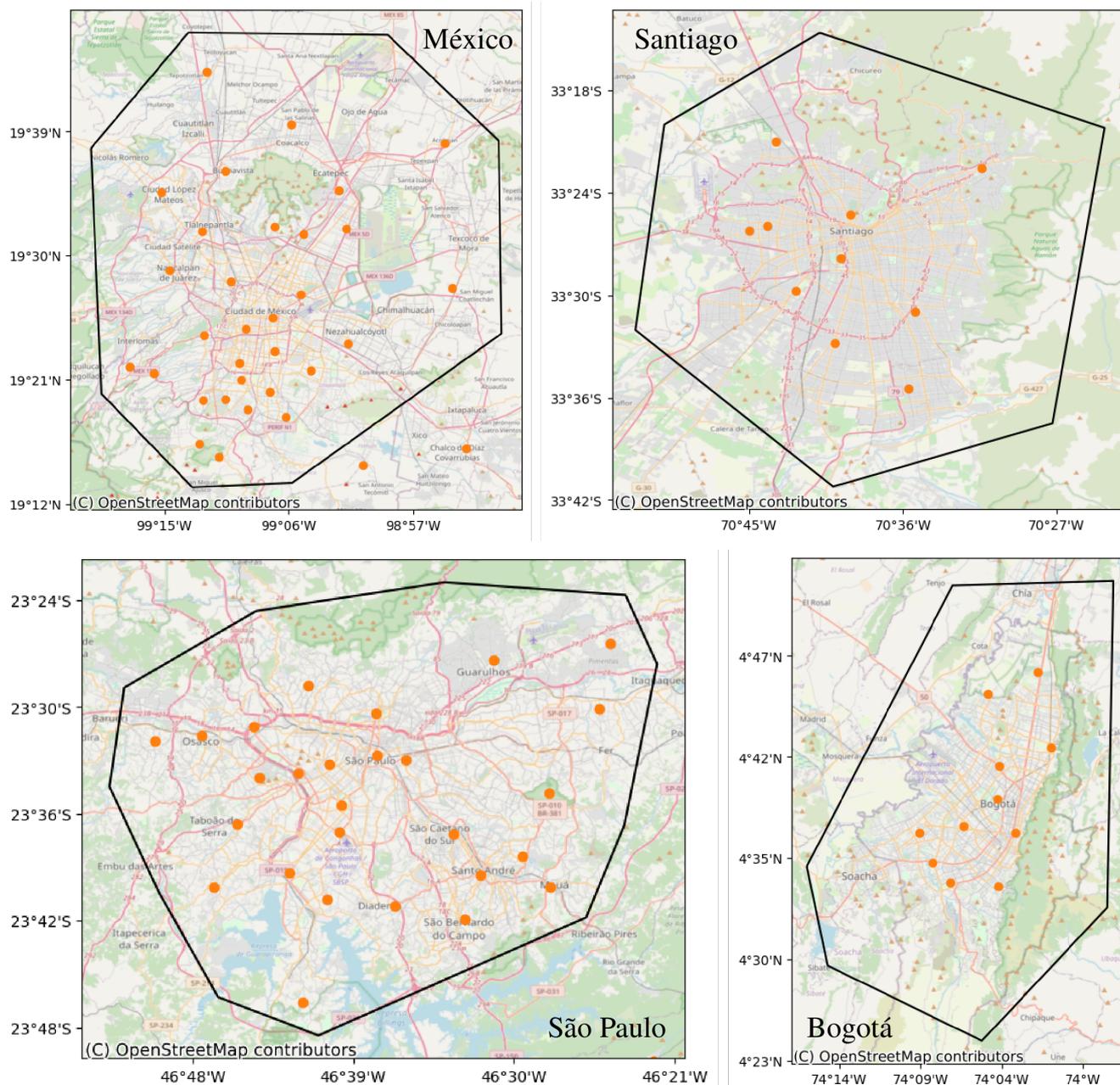


Figure 1. Location of air quality stations in major Latin American cities (Santiago, Bogotá, México City, São Paulo) alongside the city's definition for computing the modeled city average. © OpenStreetMap contributors 2024. Distributed under the Open Data Commons Open Database License (ODbL) v1.0.

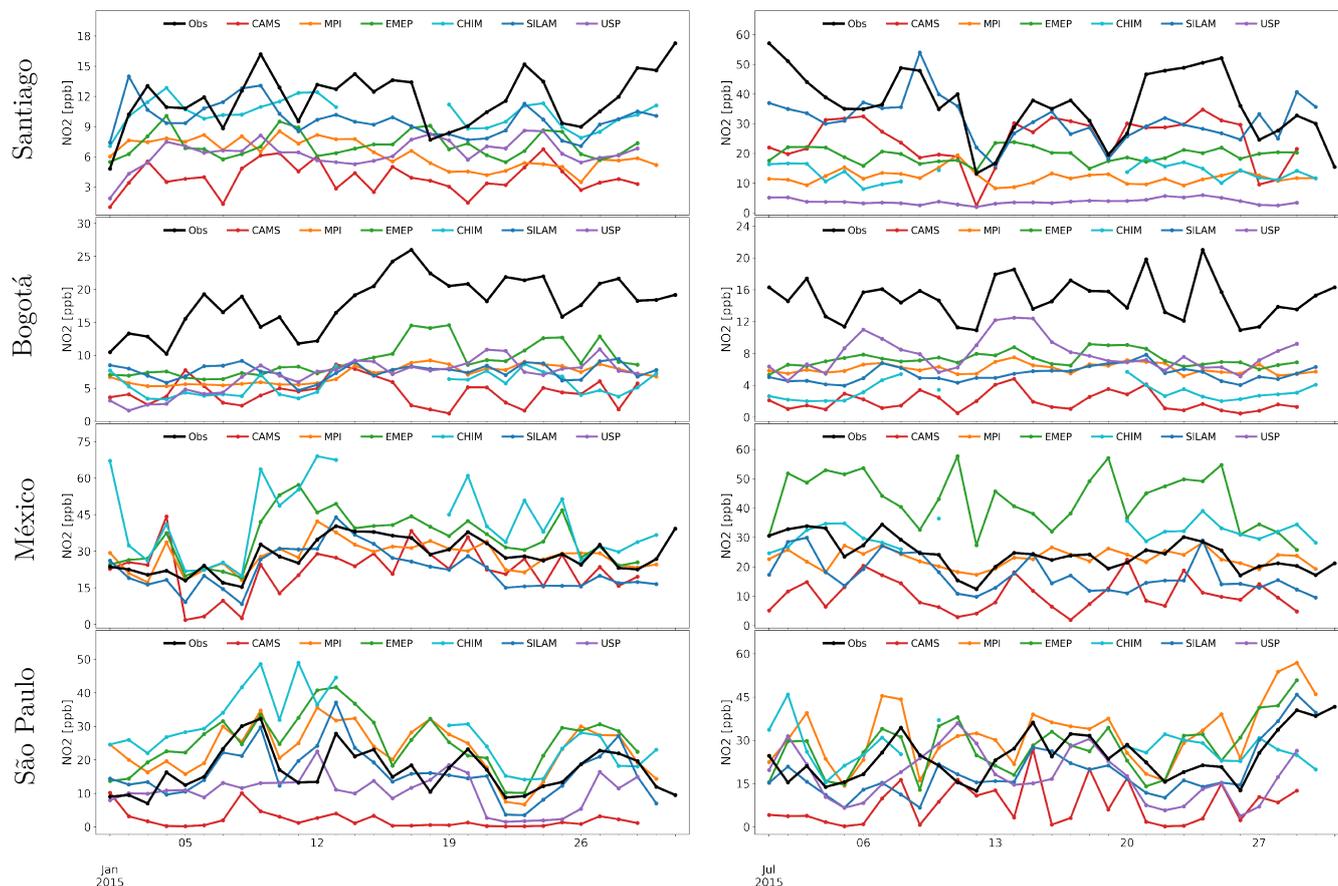


Figure 2. Observed (black) and simulated NO₂ daily mean concentrations in Santiago, (top) Bogotá, México City, São Paulo (bottom) for January (left) and July (right) 2015.

in summer and with a larger error (Table A2). The correlation between the median of the models and observations in Santiago
 160 ranges from 0.5 to 0.7.

In Bogotá the Model/Observations ratios ranged between 0.1 and 0.5, indicating that only 50% of the NO₂ is reproduced
 by the models. The MNBIAS values are large and negative and the FGE varies between 63% and 161% (Table A2). Despite
 these lower scores, the correlation between observations and models are moderate from 0.45 in July to 0.70 in January, meeting
 criteria goals (Zhai et al., 2024) and demonstrate that certain models can successfully replicate the temporal variations but not
 165 the magnitude of the pollutant.

The lower simulated NO₂ levels in Bogotá likely stems from an underestimation of emissions. A study by (Rojas et al.,
 2023) utilized local data to estimate on-road emissions in Colombia and revealed substantial underestimation of NO_x emis-
 sions by global inventories such as EDGAR 6.1, CAMS, and the Community Emissions Data System (CEDS). Their findings
 recommend adjustments to the emission factors used for NO_x, particularly for heavy-duty and passenger vehicles, followed



170 by a recalculation of the resulting emissions. The underestimation of NO₂ can also be noted in other cities such as Medellín, Guadalajara, Lima, and Quito (Figure 8). These cities, along with Bogotá, possess urban areas ranging from 235 to 890 km² and are confined within one or two cells of the models (20km x 20km). It is possible that the average of observations is heavily influenced by local sources, in which case a finer modeling resolution is required to accurately capture the spatial variability of air pollution.

175 *Model intercomparison*

For NO₂, both CAMS and SILAM underestimate the observations in the four cities, with CAMS displaying larger MNBIAS and FGE than SILAM. In general, SILAM reproduces at least 80% of the NO₂ levels, with the exception in Bogotá where only 40% is simulated. The correlation coefficient is better for SILAM (R ~0.6) than for CAMS (R ~0.3). The results from global models suggest that SILAM has a better performance for NO₂ in LAC than CAMS.

180 The results from regional models are very diverse. In general, WRF-MPI, CHIMERE and EMEP have lower values of MNBIAS and FGE for NO₂ in São Paulo and México City (Table A2). In São Paulo, except for WRF-USP, regional models tend to overestimate NO₂ with MNBIAS between 20% and 50%. WRF-USP reproduces about 60% of NO₂ concentrations. In México City, the tendency of regional models is to overestimate the NO₂ levels (MNBIAS: 20 to 50%). In Santiago, CHIMERE achieves the lowest MNBIAS (8%) in January but not in July (-99%). In Bogotá, the MNBIAS in regional models remains
185 consistently negative.

From Figure 2 is visible the model inter-variability as the dispersion between models. In Santiago in winter the range of NO₂ values is 30 ppb, which corresponds to a coefficient of variation (C.V.) of 57% (Table A8), this contrasts with the range in summer of 6.3 ppb (C.V.=33%). Other large dispersion is observed in São Paulo in January (range 26 ppb, C.V. 56%) and México City in July (range 31 ppb, C.V. 47%). On the other hand, lower dispersion is found in January in Bogotá and México
190 City (C.V. 24 and 31% respectively). It's interesting to note the case of Bogotá where all models consistently underestimate NO₂, but the dispersion in the models is the lowest.

3.1.2 Ozone - O₃

Observations

The number of stations per city recording O₃ is available in Appendix B. In January in México City, data availability was 97%.
195 The rest of the cities were 100%. Ozone pollution is particularly significant in México City during the July with an average concentration of 32 ppb. The warm and dry weather creates the ideal conditions for ozone formation, leading to this period being referred to as the “ozone season” (Silva-Quiroz et al., 2019)

The second largest ozone value occurs in São Paulo during January with daily averages of 21 ppb. This is probably due to an abundance of ozone precursors, in particular, volatile organic compounds (VOC) from the use of bio-fuels in the transportation
200 sector (de Fatima Andrade et al., 2017; Gavidia-Calderon et al., 2024). Santiago displays a strong seasonal pattern of ozone concentrations, with summer values of approximately 22 ppb and winter concentrations around 3.3 ppb. In Bogotá, ozone concentrations are the lowest and below 13 ppb.

Model performance

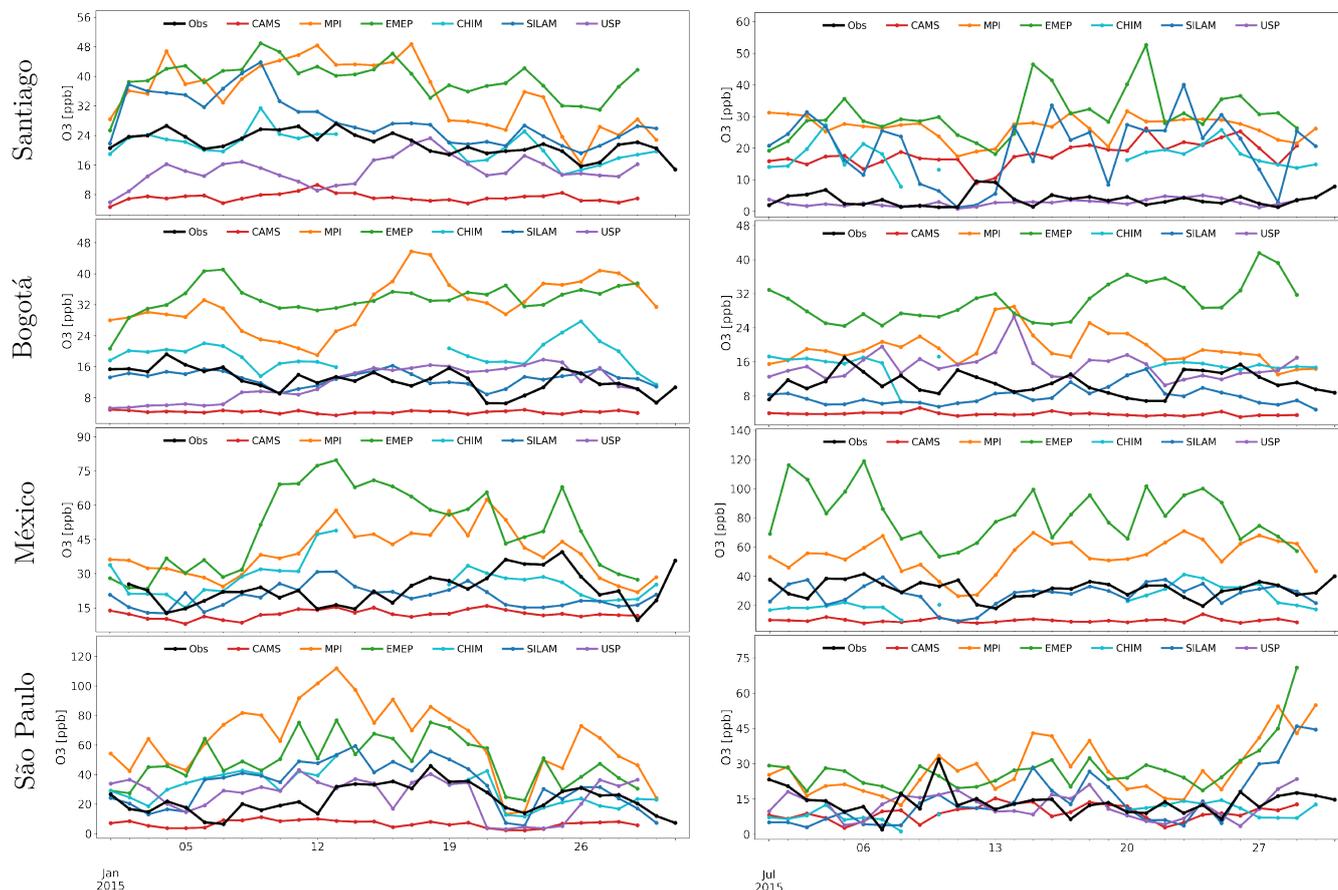


Figure 3. Observed (black) and simulated O_3 daily mean concentrations in Santiago, (top) Bogotá, México City, São Paulo (bottom) for January (left) and July (right) 2015.

In the four cities, simulations of O_3 are mainly overestimated (Figure 3). In the summer in São Paulo and México City, simulations can reach up to 120 ppb, which is significantly above the observations. In Santiago in the winter, the mean of models (~ 20 ppb) is significantly larger than observations, indicating that the models have difficulty reproducing low values of this secondary pollutant. In Bogotá, models estimate an average of 17 ppb with maximum values of 30 ppb.

For O_3 , the Model/Observations ratios vary between 0.3 and 2.8, except in Santiago in July where Model/Observations ratio can be almost 10 (Table A3) and MNBIAS and FGE for most models are larger than 100%. The overestimation of O_3 in Santiago might be related to the underestimation of NO_2 previously described and the inadequate titration of ozone. This situation is also observed in Bogotá where most models overestimate O_3 with MNBIAS between +40% and +93%. In general, correlation coefficients for O_3 are very low ($R < 0.3$), especially in São Paulo and México City, indicating the challenge to adequately reproduce the time variability of this pollutant. Only in Santiago in January, the criteria benchmark for O_3 ($R > 0.5$) is achieved by some models (Emery et al., 2017).



215 *Model intercomparison*

In the case of global models, CAMS generally underestimates O₃ except Santiago during winter, where the Model/Observations ratio reaches about 6.0. Additionally, CAMS tends to have low correlation levels along with large bias and errors. On the other hand, SILAM's O₃ estimates show more variability, ranging between Model/Observations ratios of 0.7 and 1.4 while also maintaining lower bias and errors compared to CAMS. However, just like with CAMS, SILAM significantly overestimates
220 O₃ levels in Santiago during the winter, with a Model/Observations ratio reaching around 7.0. In Bogotá, SILAM accurately simulates ozone levels with Model/Observations ratios ranging from 0.7 to 1.0 and presenting the lowest MNBIAS and RMSE.

In São Paulo, daytime concentrations of ozone are generally overestimated by most models (except for CAMS). The largest overprediction of O₃ (MNBIAS from 30 to 90%) is associated with overestimation of NO₂, especially for MPI, EMEP and
225 CHIMERE models. For the models with NO₂ levels in reasonable agreement with observations (SILAM, USP) the ozone overprediction is lower (MNBIAS <25%). Among the regional models, EMEP and WRF-MPI consistently overestimate O₃ levels in all cities, with relatively high MNBIAS and FGE. In contrast, WRF-USP proves particularly suitable for São Paulo, achieving a Model/Observations ratio of approximately 1.0. CHIMERE also performs well in Santiago, with a Model/Observations ratio of around 0.9, likely owing to local adjustments and parameterizations tailored to these specific cities.

Figure 3 shows a relatively large model intervariability for ozone. The largest ozone dispersion is shown in México City
230 in summertime with a range of 74 ppb and a C.V. of 72% (Table A8). This wide variability is caused by the simulation of the EMEP model (84 ppb) and CAMS (9.8 ppb), that represent the extreme cases of over and underestimation. In a similar manner, in Bogotá, São Paulo and Santiago, the C.V. are approximately 62%, 57% and 50% respectively, explained by the strong underestimation of CAMS and severe overestimation by EMEP and WRF-MPI.

3.1.3 Carbon monoxide - CO

235 *Observations*

The number of stations per city recording CO is available in Appendix B. In January in México City, data availability was 97%. The rest of the cities were 100%. CO levels are generally low in Latin America and daily averages are below 1.5 ppm for all cities. However, during July there are a few instances where values surpass the 1.5 ppm mark in Santiago due to a combination of adverse meteorological conditions and emissions from the transportation sector and residential combustion,
240 commonly employed for heating in neighboring municipalities (Saide et al., 2016; Gallardo et al., 2012).

There is a slight increase of CO in São Paulo in July with respect to January, due to the atmospheric conditions where lower winds and lower boundary layer increased primary pollutant concentration during winter. On the other hand, biomass burning from wildfires which begin in July and peak in August and September for the southern part of the Amazon rainforest (Marlier et al., 2020). Likewise, larger CO concentrations in Bogotá in January are part of the wildfire season in northern South America
245 lasting from the end of December until April (Mendez-Espinosa et al., 2019).

Model performance

México City records the largest simulation of CO with a mean of 1.0 ppm and peak values of 3.0 ppm (Figure 4). Similarly, Santiago during July shows an average of 0.88 ppm and peak values over 3.0 ppm. During the summer in Santiago, CO is

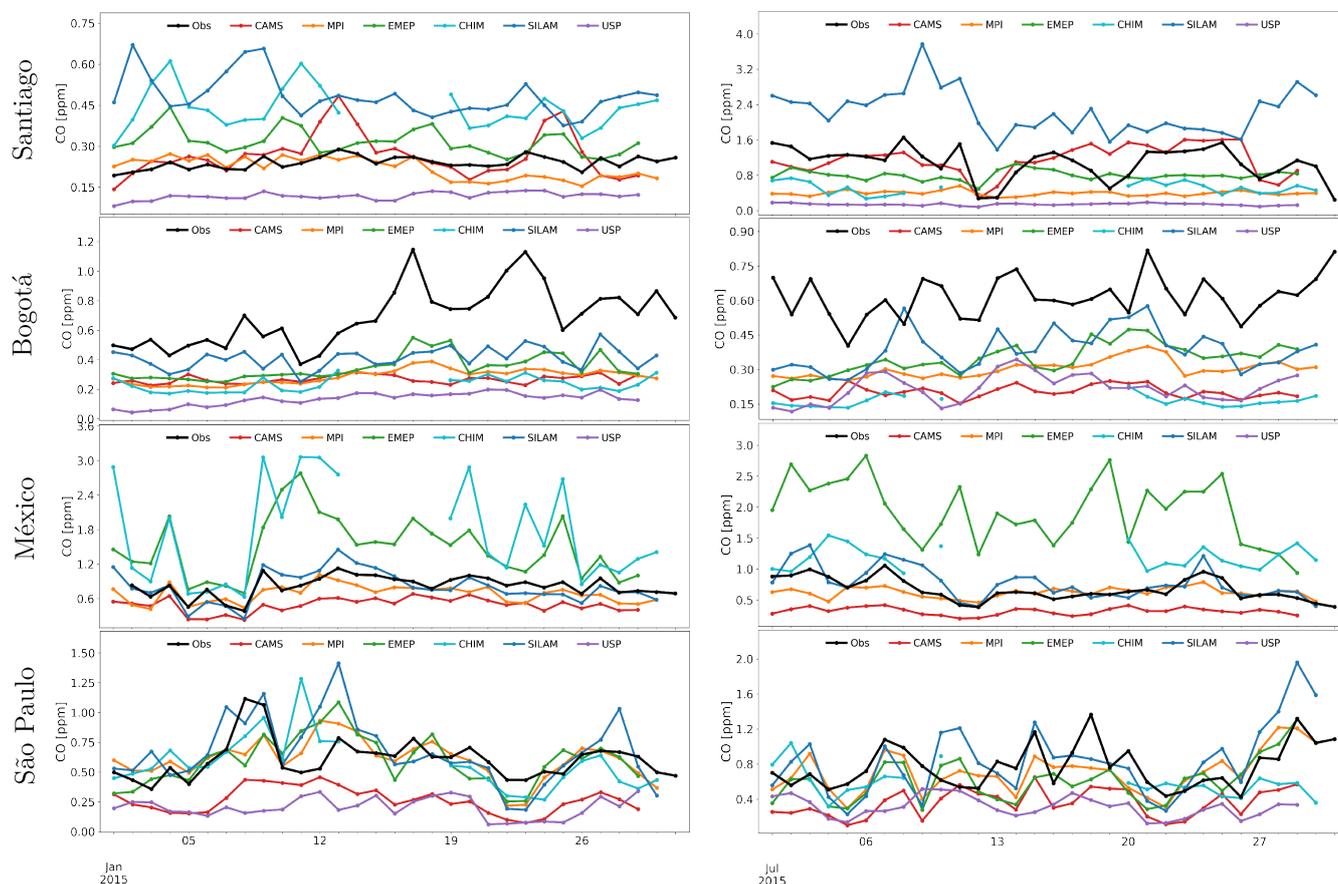


Figure 4. Observed (black) and simulated CO daily mean concentrations in Santiago, (top) Bogotá, México City, São Paulo (bottom) for January (left) and July (right) 2015.

about 0.2 ppm, overestimated by most models. São Paulo displays intermediate values with an average of 0.5ppm, and Bogotá
 250 has the lowest modeled values with an average of 0.27 ppm. In general, simulations are underestimated, particularly in Bogotá
 where only 40% of the concentration is reproducible.

CO simulations in Santiago, São Paulo, and México City, are above and below the observations (Table A4). In Santiago in
 winter, the MNBIAS ranges from -2 to -157% and the FGE vary from 20 to 157%, considerably more pronounced than in
 summer with bias from -67 to 69% and errors from 21 to 69%. This situation could be explained by emissions, synoptic or
 255 the models' simulation of the boundary layer (Mazzeo et al., 2018). In Bogotá, all models consistently underestimate the CO,
 with Model/Observations ratios ranging from 0.2 to 0.6. The correlation between models and observations for CO are within
 goal ($R > 0.4$) and criteria ($R > 0.6$) benchmarks (Zhai et al., 2024) in several cases, demonstrating the model's capability to
 reproduce the time variability of this pollutant, even if the levels are under or overestimated.



The underestimation in Bogotá is similar to that observed for NO₂, which we attributed to a shortfall in emissions. According to the local inventory, CO emissions are predominantly attributed to mobile sources (99%), with motorcycles contributing to 45% of these emissions, automobiles accounting for 36%, and the remainder originating from other vehicles (SDA -Secretaría Distrital de Ambiente, 2018). Notably, it has been identified that motorcycle emissions are underestimated in Colombia (Rojas et al., 2023). The significant rise in the number of motorcycles in the country and their declining condition is not accurately reflected in global emission inventories, such as EDGAR 6.1.

Observed CO mixing ratios are also underestimated in cities such as Medellín, Guadalajara, Quito, and Lima (Figure 8), which might be explained by the coarse resolution of the model not capturing the local characteristics. It is possible that issues with CO emissions in global inventories or excess of OH radicals in the photo-chemistry also contribute to this trend. In São Paulo, five out of six models slightly underestimate CO with a relatively high correlation coefficient. The simulated concentrations for daily values range from 0.1 to 2.0 ppm, similar to that found in other studies (Deroubaix et al., 2024). Nevertheless, concentrations exceeding 1.2 ppm are simulated only for certain days (Jan. 13 and July. 30) and are probably due to wood burning (Figure C1).

Model intercomparison

Global models, particularly CAMS, tend to underestimate CO levels in Bogotá, São Paulo, and México City, with Model/Observations ratios around 0.4. In Santiago, CAMS reproduces CO levels with a Model/Observations ratio of about 1.0, MNBIAS (< ± 10%) and FGE (<20%). The correlation coefficient achieves the criteria benchmark (R > 0.4) proposed by (Zhai et al., 2024). SILAM underestimates CO in Bogotá (Model/Observations ~0.6) and overestimates it in Santiago (Model/Observations ~2.0), while it performs relatively well in São Paulo and México City (Model/Observations ~1.1) and correlation coefficients meeting the goal benchmark (R > 0.6) proposed by (Zhai et al., 2024).

When it comes to regional models, WRF-USP consistently underestimates CO levels with Model/Observations ratios ranging from 0.1 to 0.5 with large bias and errors. WRF-MPI consistently underestimates CO in all cities, with Model/Observations ratios from 0.3 to 0.8 relative to observations, and 1.0 for São Paulo for both periods with correlation coefficients within the goal benchmark (Zhai et al., 2024). EMEP and CHIMERE both largely overestimate observations in México City with values between 4.0 and 6.0ppm, while in São Paulo they closely match observations with ratios around 0.9 and low MNBIAS and FGE. In Santiago, these models tend to overestimate CO in the summer and underestimate it during the winter.

The largest model variability is observed in Santiago during wintertime with a range of 2.1ppm and C.V. of 86% (Table A8). México City shows C.V. of 54% (January) and 63% (July). Bogotá and São Paulo present more consistency between model results with C.V. between 33% and 42%.

3.1.4 Sulfur dioxide - SO₂

Observations

The number of stations per city recording SO₂ is available in Appendix B. In January in México City, data availability was 97%. The rest of the cities were 100%. The largest concentration of SO₂ is observed in México City with values between 3.0 ppb (July) and 4.6 ppb (January) due to the heavy consumption of coal in power generation and cement production, especially



Figure 5. Observed (black) and simulated SO_2 daily mean concentrations in Santiago, (top) Bogotá, México City, São Paulo (bottom) for January (left) and July (right) 2015

in the proximity of the “Tula-Vito-Apasco” industrial area (SEMARNAT and INECC, 2020). On the other hand, SO_2 in Bogotá, Santiago and São Paulo are lower with concentrations ranging from 1.0 to 1.8 ppb.

295 *Model performance*

The largest simulation is shown in México City with an average of 10 ppb SO_2 , followed by São Paulo, with a mean concentration of 6.0 ppb. In Santiago, winter values are around 4.5 ppb and summer values around 3.6 ppb. The lowest modeled values are found in Bogotá with an average of 0.76 ppb.

The models’ simulated SO_2 exhibits significant discrepancies when compared to the observations, with severe overestimation
 300 in Santiago, México City, and São Paulo (Figure 5), with MNBIAS reaching up to 190% and errors up to 200% (Table A5).
 On the contrary, for Bogotá the predicted SO_2 values are in reasonable alignment with the observations (Model/Observations
 ratio around 0.9), except for the WRF-Chem USP simulation, which drastically underestimates SO_2 (MNBIAS -200%) (Table
 A5).



The overestimation of SO₂ levels could stem from issues within global emission inventories. In fact, an overestimation of
305 SO₂ emissions in CAMS was observed for Buenos Aires and Santiago when compared to the PAPILA inventory (Castesana
et al., 2022). These emissions primarily originate from the energy and industrial sectors, where the sulfur content in coal
appears to be significantly contributing to this overestimation.

The good performance in Bogotá might be related to less SO₂ emissions apportioned in the city. In fact, the vast majority
of SO₂ emissions (~90%) in Colombia originate from the industrial and energy production sectors (IDEAM, 2020). However,
310 these facilities are typically located outside major urban areas. Bogotá contributes only 1.5% of the total national SO₂ emissions
(de Ambiente, 2018).

Model intercomparison

With respect to global models, CAMS severely overestimates SO₂ in México City and Santiago with MNBIAS and FGE larger
than 160%. In São Paulo, the bias and errors are lower but still significant (from 90 to 125%). Similarly, SILAM overestimation
315 for these three cities is also large, with MNBIAS and FGE between 86% and 154%. For Bogotá, SILAM demonstrates good
performance in the simulation of SO₂ with MNBIAS between -5% (January) and 4% (July), FGE between 14% and 27% and
correlation coefficients that meet the criteria benchmark ($R > 0.35$) suggested by (Zhai et al., 2024).

The performance of regional models for SO₂ is quite diverse. WRF-USP severely underestimates SO₂ in all cities (MNBIAS
close to -200%). In Santiago, México City and São Paulo the models overestimate SO₂ in a similar fashion than global models.
320 In Bogotá, EMEP shows one of the lowest MNBIAS (from 10 to 17%).

The largest model variability for SO₂ is found in México City where the range of models reach 180 ppb, and the C.V. is
larger than 150% (Table A8). In Santiago in January the C.V. is 130%. São Paulo and Bogotá present intermediate values of
the C.V. between 64% and 88%.

3.1.5 Fine particulate matter - PM_{2.5}

Observations

The number of stations per city recording PM_{2.5} is available in Appendix B. In January in México City, data availability
was 97%. The rest of the cities were 100%. The largest PM_{2.5} concentrations are found in Santiago during the southern
hemispheric winter with values around 60 μgm^{-3} . This can be attributed to adverse meteorological conditions and emissions
from transportation and residential combustion in the surrounding municipalities (Mazzeo et al., 2018; Saide et al., 2016). The
330 second largest values are shown in México City with an average of 21 μgm^{-3} due to local emission sources. In São Paulo,
PM_{2.5} levels are larger in July (18 μgm^{-3}) than January (15 μgm^{-3}), due to the impact of wildfires from the Amazon basin and
sugarcane burning (de Fatima Andrade et al., 2017). In Bogotá, PM_{2.5} concentrations are the lowest in July (13 μgm^{-3}) due
to the influence of the trade winds (Pachon et al., 2018) but with larger values in January (18 μgm^{-3}) due to biomass burning
events and frequent thermal inversions (Ramírez et al., 2018).

Model performance

In Santiago in wintertime, the mean of the models is close to the observation, but with a large standard deviation (62.6 ± 85
 μgm^{-3}). In January, the simulation mean is 65% the observation. In São Paulo, simulated values are approximately double

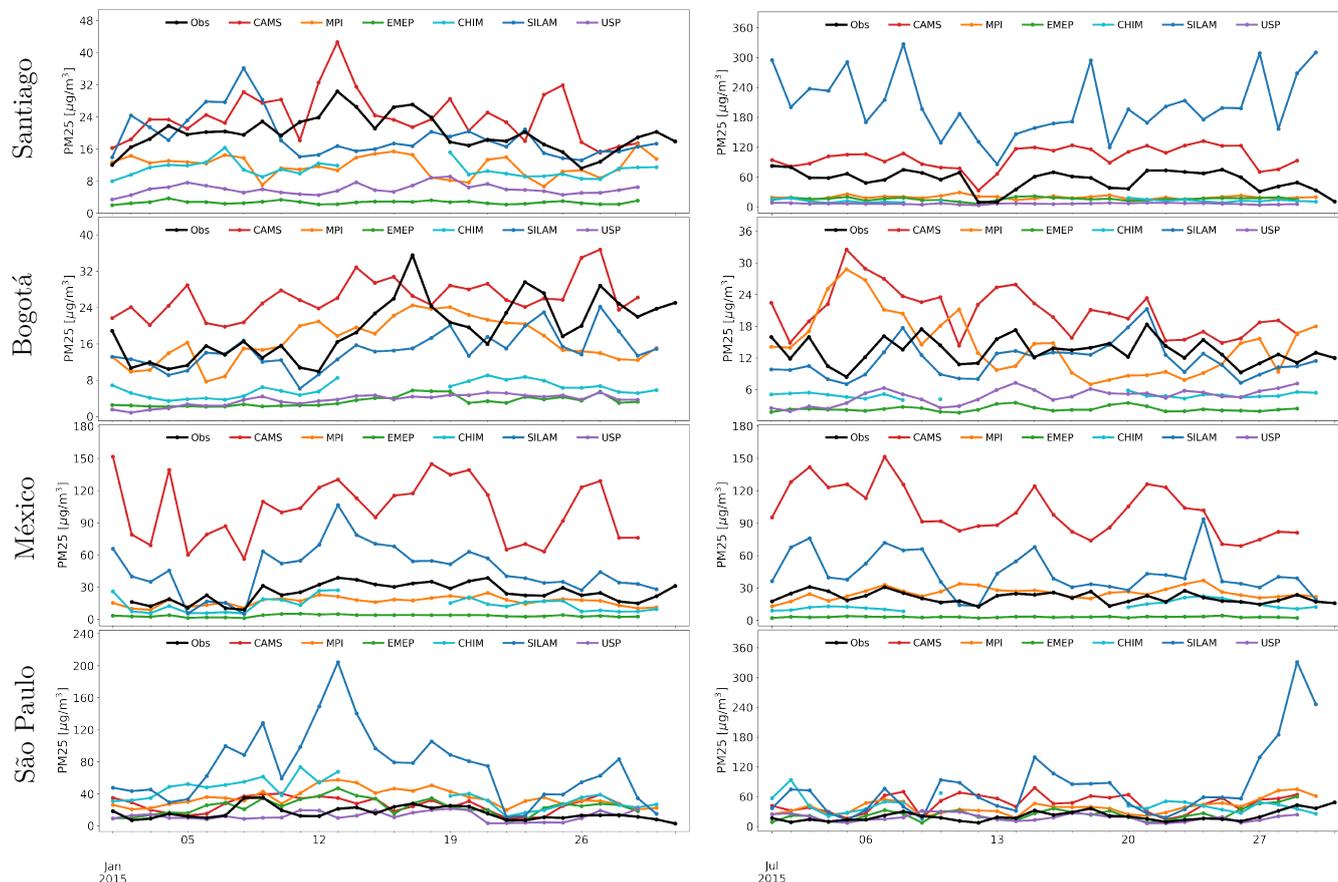


Figure 6. Observed (black) and simulated PM_{2.5} daily mean concentrations in Santiago, (top) Bogotá, México City, São Paulo (bottom) for January (left) and July (right) 2015.

the observation. In México City, simulated values are above and below the observation. In Bogotá, most of the simulations are below the observation (Figure 6).

340 In Santiago, Bogotá, and México City, some models overestimate and others underestimate PM_{2.5} (Table A6). The MNBIAS and FGE are, in general, within the goal or criteria benchmarks suggested by (Boylan and Russell, 2006). In São Paulo, overestimation is observed in all models and may be linked to an excess of fire emissions, as suggested by other studies (Deroubaix et al., 2024). The Model/Observations ratios range from 1.3 to 4.5, and MNBIAS values vary from 25% to 117% except for WRF-USP, whose MNBIAS is -0.8%. The correlation coefficients for PM_{2.5} are in some cases larger than the goal
345 (R > 0.7) or criteria (R > 0.4) benchmarks proposed by (Emery et al., 2017). It's worth noting the case of México City in January and São Paulo in July, where most models achieve the goal metric. In smaller urban areas like Medellin, Lima, and Quito (Figure 8), most models tend to underestimate observations, potentially due to the coarse resolution of the models.



Hourly simulations of $PM_{2.5}$ are useful to understand the discrepancies between model and observations. In Figure C2, we show the hourly data and model outputs. In São Paulo, the highest $PM_{2.5}$ concentrations are simulated by SILAM in January 13 ($> 320 \mu g m^{-3}$) and July 30 ($> 400 \mu g m^{-3}$), which corresponds to days with high simulated CO values as well (Figure C1) and may indicate an overestimation of biomass burning by the IS4FIRES module in SILAM. From Jan 15 to 30 there is also an excess of $PM_{2.5}$ from SILAM.

In México City, the highest $PM_{2.5}$ concentrations are simulated by the CAMS model with about $250 \mu g m^{-3}$ in January and $160 \mu g m^{-3}$ in July (Figure C3) which are severely overestimated. The large $PM_{2.5}$ values are distributed in the whole period rather than specific days and do not correspond with high CO concentrations to suspect the influence of fires. This situation might indicate a local and continuous source of $PM_{2.5}$.

Model intercomparison

Both global models consistently overestimate $PM_{2.5}$ in Santiago, São Paulo and México City, but they behave differently in Bogotá. CAMS generally has a greater overestimation than SILAM throughout the cities, reaching Model/Observations values between 1.5 to 5 while SILAM ranges between 2 to 4 (Table A6). For Bogotá, CAMS displays an overestimation of around 1.5 times the observed values with a poor correlation coefficient, while SILAM slightly underestimates with a Model/Observations ratio of 0.8 but with a correlation coefficient that meets the criteria benchmark suggested by (Emery et al., 2017).

Among the regional models, EMEP typically shows Model/Observations ratios below 0.2, except for São Paulo where it overestimates by 1.5 times with a correlation coefficient that meets the goal benchmark by (Emery et al., 2017) in July. WRF-USP heavily underestimates in Bogotá and Santiago, at 0.3 times the observations, but performs well in São Paulo with the lowest errors. This difference in behavior might be explained by a good adaptation of the model's inputs to the city. The WRF-MPI model meets goal benchmarks for MNBIAS and FGE in Bogotá and México City.

In general, global models achieve more benchmarks (goal or criteria) for $PM_{2.5}$ than regional models.

The largest model intervariability is observed in México City and Santiago during wintertime with a C.V. greater than 100% (Table A8). Santiago in summer and Bogotá present intermediate values (C.V. 65 to 78%), whereas São Paulo shows the least dispersion between models (C.V. 50% to 59%).

3.2 Median Ensemble

In this section, we present the results of the model ensemble based on the median value for every pollutant.

3.2.1 Nitrogen dioxide - NO_2

As it was previously described, NO_2 is underestimated by all models in Santiago and Bogotá. Therefore, the median ensemble also underestimates NO_2 concentration and does not represent any improvement in the evaluation metrics (Table A2). On the contrary, in São Paulo, the ensemble median outperforms individual models for NO_2 . In both summer and winter, the ensemble median presents the lowest RMSE and FGE, with a Model/Observations ratio close to 1.0, a correlation coefficient $R=0.7$, and MNBIAS between -3% (summer) to -12% (winter). The median ensemble also provided adequate statistics in a higher resolution modeling domain in São Paulo (Deroubaix et al., 2024). In México City, the ensemble adequately simulates



NO₂ (Model/Observations ~ 0.9) with lower error and bias than most of the individual models. In January, the correlation coefficient meets the goal benchmark for this pollutant ($R > 0.6$) in all cities, whereas in July the goal benchmark is achieved for São Paulo and the criteria target ($R > 0.5$) for Santiago and México City.

3.2.2 Ozone - O₃

385 In Santiago in January, the median ensemble showed one of the lowest biases (MNBIAS) and errors (FGE, RMSE), surpassed by only one model (Table A3), and achieved the goal benchmark for this pollutant ($R > 0.75$) (Emery et al., 2017). In July, the overestimation of ozone by most models impacts the performance of the ensemble, which also overestimates O₃ concentrations. In Bogotá, the ensemble has the second lowest MBIAS and FGE, both in January and July, and represents an intermediate value between all models. In São Paulo, in wintertime, the ensemble has superior metrics (Model/Observations ratio ~ 1.0 , MNBIAS
390 $\sim -2\%$) compared to any individual model, while in the summer the ensemble overestimates the observations (Model/Observations ratio ~ 1.5) as most models do. In México City, the ensemble median performs better than all individual models with MNBIAS between 7% (summer) and 13% (winter) and FGE less than 30%. Similar to the individual models, for most of the cases, the correlation coefficient for the ensemble does not meet any of the benchmarks (Emery et al., 2017).

3.2.3 Carbon monoxide - CO

395 In the summer in Santiago, the median ensemble outperforms individual models for CO, with MNBIAS of 6% and FGE of 11 less than any other model (Table A4). In winter in Santiago and Bogotá in both periods the ensemble follows the underestimation pattern of all models. In São Paulo, there are models with better performance than the ensemble, but the ensemble results are reasonable with Model/Observations ratio ~ 0.7 , MNBIAS $\sim -30\%$ and $R \sim 0.7$. In México City, the overestimation of CO by the EMEP and CHIMERE models (MNBIAS $> 45\%$) is reduced in the ensemble (MNBIAS -5% in January and of
400 3% in July).

3.2.4 Sulfur dioxide - SO₂

In México City, Santiago and São Paulo, SO₂ is overestimated by all models, except USP. Therefore, the median ensemble also overestimates SO₂ concentration and does not represent any improvement in the evaluation metrics (Table A5). In Bogotá, the ensemble does not display the best metrics, but MNBIAS and FGE are relatively low.

405 3.2.5 Particulate matter - PM_{2.5}

The median ensemble for Bogotá and Santiago does not represent any improvement in the evaluation metrics (Table A6). For São Paulo, all models tend to overestimate PM_{2.5}, so it follows that the ensemble presents the same behavior with Model/Observations ~ 1.9 and MNBIAS 64%. However, the correlation coefficient meets the goal benchmark ($R > 0.7$) in July and the criteria target ($R > 0.4$) in January. In México City the severe overestimation of PM_{2.5} by CAMS (MNBIAS $> 125\%$) and SILAM



410 (MNBIAS>50%) is softened by the construction of the ensemble, resulting in a MNBIAS of -30% and 6% in January and July, respectively. The correlation coefficient meets the goal benchmark ($R>0.7$) in January.

3.3 Spatial seasonal variability of predictions

For all pollutants, models and periods, maps of mean concentrations were constructed to visualize the spatial differences (Appendix D). In order to summarize the results, other spatial plots were also prepared: median ensemble (Figure 7), median
415 absolute deviation (Figure E1), mean standard deviation (Figure E2). In Figure 7, hot pollution spots are clearly visible around major urban areas, in particular, São Paulo in the southeastern coast and México City in the northwestern part of the continent. São Paulo and México City each cover a significant area, of approximately 3600 km², spanning at least nine modeling cells (400 km² each). This extensive coverage offers some spatial representation of the physical and chemical atmospheric processes. Other regions highlighted on the maps include Lima and Santiago on the Pacific coast, Buenos Aires along the southern shore
420 of the Río de la Plata, and cities in the northern part of South America like Quito, Bogotá, Medellín, and Caracas. However, most of these cities are encompassed by three or fewer modeling cells, limiting the potential for significant spatial variation.

The temporal seasonality can also be observed in Figure 7. The left and right panels show results for January and July, corresponding to the southern hemisphere summer and winter respectively. For SO₂, major hot spots appear in México City and surrounding areas, and the Pacific coast in Chile. The SO₂ concentrations are associated with coal use in power generation,
425 cement production and copper smelting that are active in both summer and winter (Huneeus et al., 2006; SEMARNAT and INECC, 2020). Similarly, NO₂ hotspots are in major urban areas due to the major emission source being transportation.

In January, the median ensemble shows high concentrations of PM₁₀ in several areas. In the south of Argentina, the concentrations are primarily due to dust from the Patagonia desertic areas (Gasso and Torres, 2019). In the north of Brazil and the Guianas, increased PM₁₀ levels are most likely associated with fires in the Orinoco basin during the dry season (Hernandez
430 et al., 2019). In a similar manner, PM_{2.5} concentrations across LAC show similar behavior than PM₁₀, with an increase in the northern part of Brazil due to wildfires. Large concentrations of PM_{2.5} in São Paulo in both January and July are probably caused by overestimation of fires, as previously discussed. During the austral summer, São Paulo presents large concentrations of ozone that were simulated by the regional models WRF-Chem, WRF-MPI and EMEP (Figure D3). In January, there is a maximum of CO in the area between north of Argentina, south of Bolivia, Paraguay and south of Brazil, probably related to
435 fires.

In July, concentrations of CO, PM_{2.5} and PM₁₀ are significant in Santiago due to transportation and residential heating emissions under adverse meteorological conditions in the austral winter. PM₁₀ concentrations are large in the Caribbean and central México, primarily due to the transport of Saharan dust into these urban areas (Kramer and Kirtman, 2021; Ramírez-Romero et al., 2021). Similarly, along the Pacific coast between Chile and Peru, increased PM₁₀ is probably explained by
440 anthropogenic emissions of copper smelters in connection with strong eastern wind events (Huneeus et al., 2006). Large concentrations of O₃ are visible in México and in the Andes mountains between northern Chile and central Peru. Ozone is also large in the São Paulo metropolitan area in January.

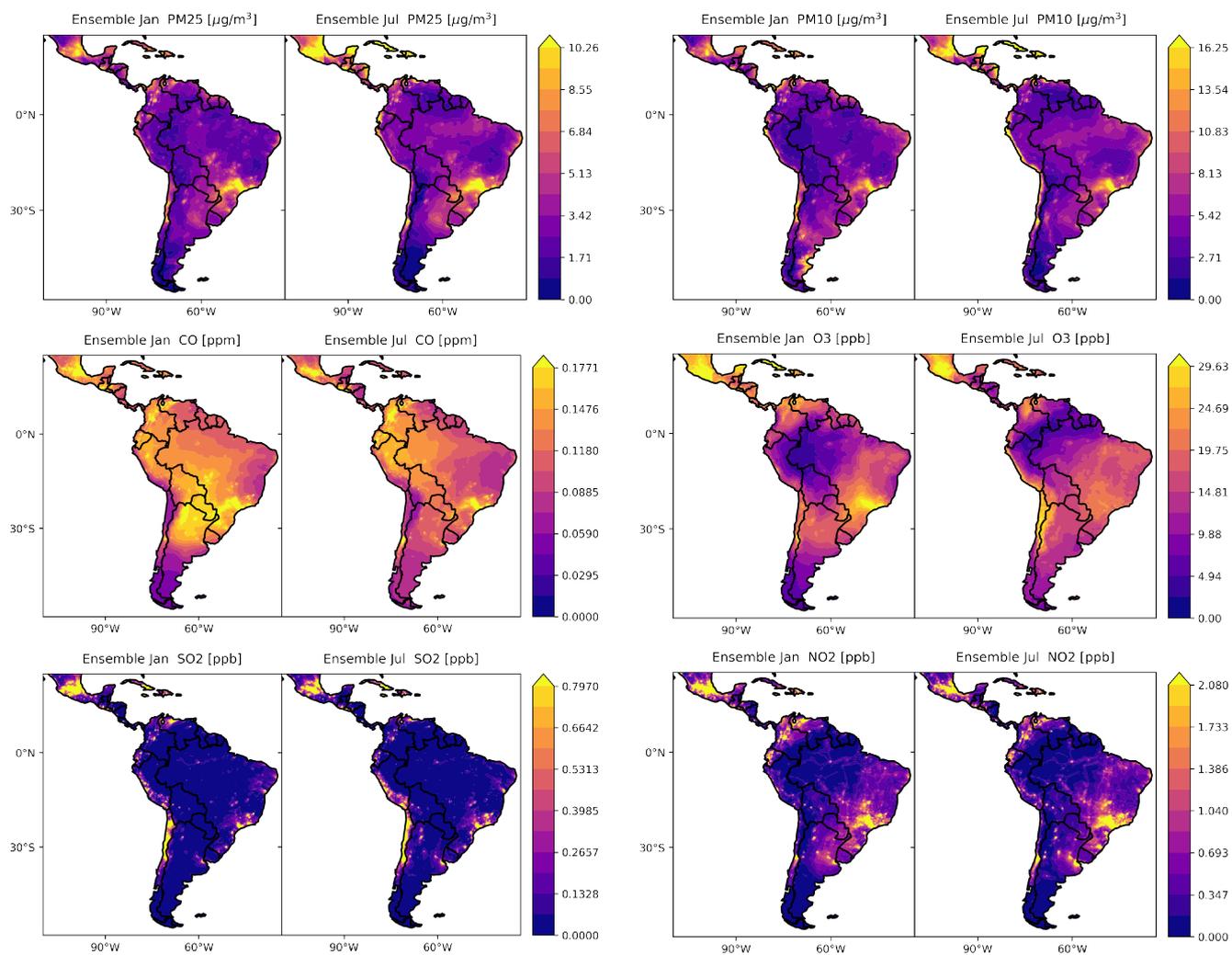


Figure 7. Spatial variability of simulated PM₁₀, PM_{2.5}, O₃, CO in LAC for January and July 2015 (based on the median of the models)

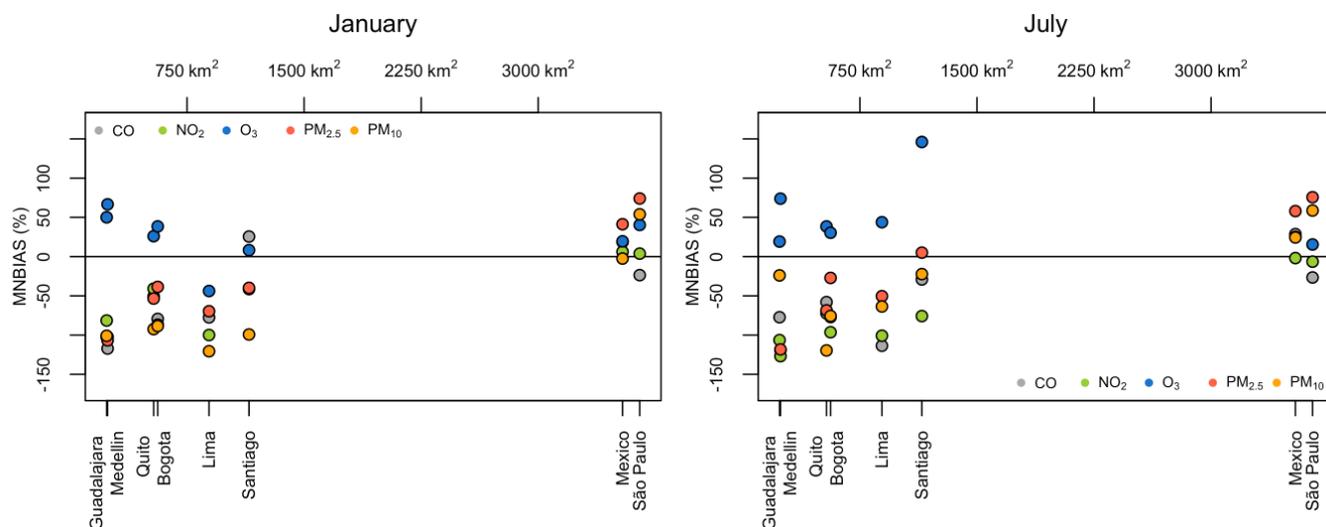


Figure 8. MNBIAS estimated for large and small urban areas.

The median absolute deviation maps (Figure E1) and the standard deviation maps (Figure E2), display spatial differences between model simulations. In particular, for particulate matter (PM₁₀ and PM_{2.5}) notorious dissimilarity is observed in northern Brazil in January, Venezuela in July, and the south of Argentina in both periods. The reason for this disagreement is the simulation of the WRF-MPI model, which contributes with significant PM mass in the mentioned zones, probably due to an overestimation of fires in the northern part of the continent and dust in the southern areas.

3.4 Large versus small urban areas

The coarse resolution used in the modeling systems (0.2°x 0.2°) poses challenges in adequately representing the intricate topography and diverse meteorological conditions of the different cities in LAC. Capturing these physical phenomena can be very difficult and requires a finer scale with much greater computational demand. In the last years, emission inventories for LAC at high spatial and temporal resolution have been constructed (Castesana et al., 2022; Alamos et al., 2022; Puliafito et al., 2015, 2017; Rojas et al., 2023) and it's expected they will complement existing global emission inventories at coarse resolution. We observe that, in large urban areas (> 3500 km²) the models tend in general to have lower and positive MNBIAS compared to medium size (600 < area < 3600 km²) or small (area < 600 km²) cities (Figure 8). For example, for México City and São Paulo, the two largest cities in LAC, the models show the lowest MNBIAS and FGE for CO (-27% to 29%) and NO₂ (-6% to 6%), while in other cities they display larger and negative MNBIAS and FGE (Table A2 and A4). This trend suggests that models typically underestimates CO and NO₂ in medium and small urban areas. The discrepancies in NO₂ have a corresponding impact in the overestimation of O₃. For particulate matter, a similar pattern is observed, with positive MNBIAS for larger urban areas and negative MNBIAS for medium and small cities. Ideally, we would have access to more cities of



various sizes to make this determination with more certainty, unfortunately, local measured data was only available for the cities we considered.

4 Conclusions and future developments

This study performed the first intercomparison and model evaluation in Latin America with interesting and insightful findings for the region. Several challenges were faced and partially overcome. In addition to the intricate topography and diverse meteorological conditions of cities in LAC, some of the individual models were still in an early phase of regional modeling. Limitations in model inputs exist on anthropogenic emissions, spatial and temporal profiles, land use and vegetation types, as well as other data that is relevant for the calculation of biogenic fluxes and wildfires. The latter emission source is crucial for the region, and more relevant under a climate change scenario, for which an adequate parametrization of biomass burning is necessary. The models' boundary condition can be improved, which may be relevant for longer-lived species such as CO and ozone.

Despite the above limitations and the coarse resolution ($0.2^\circ \times 0.2^\circ$) adopted in this work, most models could reproduce air quality observations with the best performance observed for nitrogen dioxide in México City and São Paulo. These enormous urban areas ($> 3500 \text{ km}^2$) outperformed Bogotá and Santiago, cities between 500 and 1000 km^2 . This suggests an accurate portrayal of the temporal and spatial variability in large cities and the need for a finer resolution in smaller cities. During wintertime simulations in Santiago, some pollutants displayed large discrepancies with observations, especially for NO_2 , O_3 and $\text{PM}_{2.5}$. In Bogotá, all models systematically underestimate CO and NO_2 . The discrepancies in NO_2 had a corresponding impact in the overestimation of O_3 . Most models overestimate SO_2 concentrations in all cities, except Bogotá, due to the high sulfur content in solid and liquid fuels attributed to the region.

Global and regional models provided different results. The SILAM model showed a better performance for NO_2 than CAMS. In Bogotá, SILAM presents low bias for ozone concentrations, while CAMS severely underestimated this pollutant. This underestimation was also observed in São Paulo and Santiago. Regional models that have been previously implemented in the cities showed lower bias, such as CHIMERE in Santiago for NO_2 and WRF-Chem in São Paulo for NO_2 and O_3 . Global models show an overestimation of biomass burning emissions, which may explain the overestimation of $\text{PM}_{2.5}$ in São Paulo.

The ensemble median offered a promising avenue for establishing a regional analysis and forecasting system. While certain individual models outperformed the ensemble for specific pollutants and cities, the ensemble provides a useful tool to mitigate the extreme over or underestimation of certain models. In São Paulo, the ensemble median performed better than any model for NO_2 . In México City, the creation of the ensemble softened large overestimation of $\text{PM}_{2.5}$ by global models. In Santiago in the summer, the median ensemble shows one of the lowest biases (MNBIAS) and errors (FGE, RMSE). In México City, the ensemble median for O_3 performed better than all individual models. In the summer in Santiago, the median ensemble outperformed individual models for CO.

This study aimed to assemble a multi-scale ensemble chain as a first step towards an Air Quality forecasting system for Latin America. Before such a prototype can be operative, a thorough analysis of one entire annual cycle with sufficient spin-up time



495 should be conducted. This work only looked at two months (one in summer and one in winter). More AQ observations should also be included for model calibration and evaluation. For 2015, only eight cities in LAC had data that complied with quality and completeness criteria. In recent years, more AQ networks have been implemented and data is more publicly available.

Code and data availability. All model data analyzed in the intercomparison is archived at <https://doi.org/10.5281/zenodo.10934490>. The tool to create the plots, MOSPAT, can be found in GitHub at https://github.com/NeoMOSPAT/NeoMOSPAT_PAPILA.git.

500 *Author contributions.* JE and MO performed the formal analysis of the data; PL, NH, IB, JF, LM, CM, MG, MS, RK, JP, MC, AD, and DS performed the model simulations; JE, MO, PL, NH, and IB prepared the manuscript with contributions from all co-authors; GB and LG provided the financial support for the project leading to this publication; LD, NR, NH, MA coordinated research activities; CL provided technical support; all co-authors reviewed and edited the manuscript.

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

505 *Acknowledgements.* The PAPILA (Prediction of Air Pollutants in Latin America) project was funded by the European Commission under the MSCA action for research and innovation staff exchange, Grant agreement ID: 777544. Support of Academy of Finland HEATCOST and ACCC flagship projects (grants Nbr 334798 and 337552), and H2020 project AQ-Watch (grant Nbr 870301) is acknowledged. Support from the German Climate Computer Center (Deutsches Klimarechenzentrum, DKRZ) is acknowledged. Support from the Center for Climate and Resilience Research (FONDAP/ANID 1523A0002) is acknowledged.



References

- 510 Alamos, N., Huneus, N., Opazo, M., Osses, M., Puja, S., Pantoja, N., van der Gon, H. D., Schueftan, A., Reyes, R., and Calvo, R.: High-resolution inventory of atmospheric emissions from transport, industrial, energy, mining and residential activities in Chile, *Earth Syst. Sci. Data*, 14, 361–379, <https://doi.org/10.5194/essd-14-361-2022>, 2022.
- Albuquerque, T. T., West, J., de F. Andrade, M., Ynoue, R. Y., Andreão, W. L., dos Santos, F. S., Maciel, F. M., Pedruzzi, R., de O. Mateus, V., Martins, J. A., Martins, L. D., Nascimento, E. G. S., and Moreira, D. M.: Analysis of PM_{2.5} concentrations under pollutant emission control strategies in the metropolitan area of São Paulo, Brazil, *Environmental Science and Pollution Research*, 26, 33 216–33 227, <https://doi.org/10.1007/s11356-019-06447-6>, 2019.
- 515 Ballesteros-Gonzalez, K., Sullivan, A. P., and Morales-Betancourt, R.: Estimating the air quality and health impacts of biomass burning in northern South America using a chemical transport model, *Science of The Total Environment*, 739, 139 755, <https://doi.org/https://doi.org/10.1016/j.scitotenv.2020.139755>, 2020.
- 520 Ballesteros-Gonzalez, K., Espitia-Cano, S. O., Rincon-Caro, M. A., Rincon-Riveros, J. M., Perez-Peña, M. P., Sullivan, A., and Betancourt, R. M.: Understanding organic aerosols in Bogota, Colombia: In-situ observations and regional-scale modeling, *Atmospheric Environment*, 284, 119 161, <https://doi.org/https://doi.org/10.1016/j.atmosenv.2022.119161>, 2022.
- Blechschmidt, A.-M., Arteta, J., Coman, A., Curier, L., Eskes, H., Foret, G., Gielen, C., Hendrick, F., Marecal, V., Meleux, F., Parmentier, J., Peters, E., Pinardi, G., Piters, A. J. M., Plu, M., Richter, A., Segers, A., Sofiev, M., a M Valdebenito, Roozendaal, M. V., Vira, J., Vlemmix, T., and Burrows, J. P.: Comparison of tropospheric NO₂ columns from MAX-DOAS retrievals and regional air quality model simulations, *Atmos. Chem. Phys.*, 20, 2795–2823, <https://doi.org/10.5194/acp-20-2795-2020>, 2020.
- 525 Boylan, J. W. and Russell, A. G.: PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models, *Atmospheric Environment*, 40, 4946–4959, <https://doi.org/https://doi.org/10.1016/j.atmosenv.2005.09.087>, 2006.
- Brasseur, G. P., Xie, Y., Petersen, A. K., Bouarar, I., Flemming, J., Gauss, M., Jiang, F., Kouznetsov, R., Kranenburg, R., Mijling, B., Peuch, V.-H., Pommier, M., Segers, A., Sofiev, M., Timmermans, R., van der A, R., Walters, S., Xu, J., and Zhou, G.: Ensemble forecasts of air quality in eastern China – Part 1: Model description and implementation of the MarcoPolo–Panda prediction system, version 1, *Geoscientific Model Development*, 12, 33–67, <https://doi.org/10.5194/gmd-12-33-2019>, 2019.
- 530 Busch, P., Cifuentes, L. A., and Cabrera, C.: Chronic exposure to fine particles (PM_{2.5}) and mortality: Evidence from Chile, *Environmental Epidemiology*, 7, https://journals.lww.com/environepidem/fulltext/2023/08000/chronic_exposure_to_fine_particles__pm2_5__and.3.aspx, 2023.
- 535 Casallas, A., Castillo-Camacho, M. P., Sanchez, E. R., Gonzalez, Y., Celis, N., Felipe, J., Belalcazar, L. C., and Ferro, C.: Surface, satellite ozone variations in Northern South America during low anthropogenic emission conditions: a machine learning approach, *Air Quality, Atmosphere & Health*, 16, 745–764, <https://doi.org/10.1007/s11869-023-01303-6>, 2023.
- Casallas, A., Cabrera, A., Guevara-Luna, M.-A., Tompkins, A., Gonzalez, Y., Aranda, J., Belalcazar, L. C., Mogollon-Sotelo, C., Celis, N., Lopez-Barrera, E., Peña-Rincon, C. A., and Ferro, C.: Air pollution analysis in Northwestern South America: A new Lagrangian framework, *Science of The Total Environment*, 906, 167 350, <https://doi.org/https://doi.org/10.1016/j.scitotenv.2023.167350>, 2024.
- 540 Castesana, P., Resquin, M. D., Huneus, N., Puliafito, E., Darras, S., Gomez, D., Granier, C., Alvarado, M. O., Rojas, N., and Dawidowski, L.: PAPILA dataset: a regional emission inventory of reactive gases for South America based on the combination of local and global information, *Earth System Science Data*, 14, 271–293, <https://doi.org/10.5194/essd-14-271-2022>, 2022.



- 545 Colette, A., Collin, G., and Barre, J.: Update on European Regional Air Quality Forecast in the Copernicus Atmosphere Monitoring Service (CAM5), p. 3487, <https://doi.org/10.5194/egusphere-egu2020-3487>, 2020.
- de Ambiente, S. S. D.: Inventario de emisiones de Bogota, Contaminantes atmosfericos año 2018, 2018.
- de Fatima Andrade, M., 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
550 30 years and future perspectives, *Atmospheric Environment*, 159, 66–82, <https://doi.org/https://doi.org/10.1016/j.atmosenv.2017.03.051>, 2017.
- Deroubaix, A. M., Hoelzemann, J. J., Ynoue, R., Bouarar, I., de Souza Fernandes Duarte, E., Elbern, H., Lichtig, P., Martins, L. D., evora do Rosario, N. M., Brasseur, G. P., rafaela Cruz Alves, de Arruda Moreira, G., phiipp franke, de Fatima Andrade, M., Lange, A. C., Andreao, W. L., rizzieri Pedruzzi, de Almeida Albuquerque, T. T., and Lugon, L.: Intercomparison of air qual-
555 ity models in a megacity: Towards an operational ensemble forecasting system for São Paulo, *JGR Atmospheres*, 129, 1–27, <https://doi.org/https://doi.org/10.1029/2022JD038179>, 2024.
- East, J., Montealegre, J. S., Pachon, J. E., and Garcia-Menendez, F.: Air quality modeling to inform pollution mitigation strategies in a Latin American megacity, *Science of The Total Environment*, 776, 145 894, <https://doi.org/https://doi.org/10.1016/j.scitotenv.2021.145894>, 2021.
- 560 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, *Journal of the Air & Waste Management Association*, 67, 582–598, <https://doi.org/10.1080/10962247.2016.1265027>, 2017.
- Fast, J. D., Jr., W. I. G., Easter, R. C., Zaveri, R. A., Barnard, J. C., Chapman, E. G., Grell, G. A., and Peckham, S. E.: Evolution of ozone, particulates, and aerosol direct radiative forcing in the vicinity of Houston using a fully coupled meteorology-chemistry-aerosol model,
565 *Journal of Geophysical Research: Atmospheres*, 111, <https://doi.org/https://doi.org/10.1029/2005JD006721>, 2006.
- Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A.-M., Diamantakis, M., Engelen, R. J., Gaudel, A., Inness, A., Jones, L., Josse, B., Katragkou, E., Marecal, V., Peuch, V.-H., Richter, A., Schultz, M. G., Stein, O., and Tsikerdekis, A.: Tropospheric chemistry in the Integrated Forecasting System of ECMWF, *Geoscientific Model Development*, 8, 975–1003, <https://doi.org/10.5194/gmd-8-975-2015>, 2015.
- 570 Flemming, J., Benedetti, A., Inness, A., Engelen, R. J., Jones, L., Huijnen, V., Remy, S., Parrington, M., Suttie, M., Bozzo, A., Peuch, V.-H., Akritidis, D., and Katragkou, E.: The CAMS interim Reanalysis of Carbon Monoxide, Ozone and Aerosol for 2003–2015, *Atmospheric Chemistry and Physics*, 17, 1945–1983, <https://doi.org/10.5194/acp-17-1945-2017>, 2017.
- Franco, J. F., Gidhagen, L., Morales, R., and Behrentz, E.: Towards a better understanding of urban air quality management capabilities in Latin America, *Environmental Science & Policy*, 102, 43–53, <https://doi.org/https://doi.org/10.1016/j.envsci.2019.09.011>, 2019.
- 575 Freitas, S. R., Longo, K. M., Alonso, M. F., Pirre, M., Marecal, V., Grell, G., Stockler, R., Mello, R. F., and Gacita, M. S.: PREP-CHEM-SRC – 1.0: a preprocessor of trace gas and aerosol emission fields for regional and global atmospheric chemistry models, *Geosci. Model Dev.*, 4, 419–433, <https://doi.org/10.5194/gmd-4-419-2011>, 2011.
- Gallardo, L., Escribano, J., Dawidowski, L., Rojas, N., de Fatima Andrade, M., and Osses, M.: Evaluation of vehicle emission inventories for carbon monoxide and nitrogen oxides for Bogota, Buenos Aires, Santiago, and São Paulo, *Atmospheric Environment*, 47, 12–19,
580 <https://doi.org/https://doi.org/10.1016/j.atmosenv.2011.11.051>, 2012.
- Gasso, S. and Torres, O.: Temporal Characterization of Dust Activity in the Central Patagonia Desert (Years 1964–2017), *Journal of Geophysical Research: Atmospheres*, 124, 3417–3434, <https://doi.org/https://doi.org/10.1029/2018JD030209>, 2019.



- Gavidia-Calderon, M., Schuch, D., Vara-Vela, A., Inoue, R., Freitas, E. D., de A Albuquerque, T. T., Zhang, Y., de Fatima Andrade, M., and Bell, M. L.: Air quality modeling in the metropolitan area of São Paulo, Brazil: A review, *Atmospheric Environment*, 319, 120301, 585
<https://doi.org/https://doi.org/10.1016/j.atmosenv.2023.120301>, 2024.
- Gonzalez, C. M., Ynoue, R. Y., Vara-Vela, A., Rojas, N. Y., and Aristizabal, B. H.: High-resolution air quality modeling in a medium-sized city in the tropical Andes: Assessment of local and global emissions in understanding ozone and PM10 dynamics, *Atmospheric Pollution Research*, 9, 934–948, <https://doi.org/10.1016/j.apr.2018.03.003>, 2018.
- Gouveia, N., Junger, W. L., Romieu, I., Cifuentes, L. A., de Leon, A. P., Vera, J., Strappa, V., Hurtado-Diaz, M., Miranda-Soberanis, V., 590
Rojas-Bracho, L., Carbajal-Arroyo, L., and Tzintzun-Cervantes, G.: Effects of air pollution on infant and children respiratory mortality in four large Latin-American cities, *Environmental Pollution*, 232, 385–391, <https://doi.org/https://doi.org/10.1016/j.envpol.2017.08.125>, 2018.
- Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled “online” chemistry within the WRF model, *Atmospheric Environment*, 39, 6957–6975, <https://doi.org/https://doi.org/10.1016/j.atmosenv.2005.04.027>, 2005.
- 595 Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), *Atmospheric Chemistry and Physics*, 6, 3181–3210, <https://doi.org/10.5194/acp-6-3181-2006>, 2006.
- Hernandez, A. J., Morales-Rincon, L. A., Wu, D., Mallia, D., Lin, J. C., and Jimenez, R.: Transboundary transport of biomass burning aerosols and photochemical pollution in the Orinoco River Basin, *Atmospheric Environment*, 205, 1–8, 600
<https://doi.org/https://doi.org/10.1016/j.atmosenv.2019.01.051>, 2019.
- Huneus, N., Gallardo, L., and Rutllant, J. A.: Offshore transport episodes of anthropogenic sulfur in northern Chile: Potential impact on the stratocumulus cloud deck, *Geophysical Research Letters*, 33, <https://doi.org/https://doi.org/10.1029/2006GL026921>, 2006.
- IDEAM: Primer inventario indicativo Nacional de emisiones de contaminantes criterio y carbono negro 2010-2014, <https://documentacion.ideam.gov.co/openbiblio/bvirtual/023893/1InventarioBLACK.pdf>, 2020.
- 605 Inness, A., Ades, M., AgustI-Panareda, A., Barre, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition, *Atmos. Chem. Phys.*, 19, 3515–3556, <https://doi.org/10.5194/acp-19-3515-2019>, 2019.
- Kouznetsov, R. and Sofiev, M.: A methodology for evaluation of vertical dispersion and dry deposition of atmospheric aerosols, *Journal of Geophysical Research: Atmospheres*, 117, <https://doi.org/https://doi.org/10.1029/2011JD016366>, 2012.
- 610 Kramer, S. J. and Kirtman, B. P.: Saharan Dust Transport Predictability Utilizing a Subseasonal Experiment (SubX) Model, *Journal of Geophysical Research: Atmospheres*, 126, e2020JD033 802, <https://doi.org/https://doi.org/10.1029/2020JD033802>, 2021.
- Kukkonen, J., Olsson, T., Schultz, D. M., Baklanov, A., Klein, T., Miranda, A. I., Monteiro, A., Hirtl, M., Tarvainen, V., Boy, M., Peuch, V.-H., Poupkou, A., Kioutsioukis, I., Finardi, S., Sofiev, M., Sokhi, R., Lehtinen, K. E. J., Karatzas, K., Jose, R. S., Astitha, M., Kallos, 615
G., Schaap, M., Reimer, E., Jakobs, H., and Eben, K.: A review of operational, regional-scale, chemical weather forecasting models in Europe, *Atmos. Chem. Phys.*, 12, 1–87, <https://doi.org/10.5194/acp-12-1-2012>, 2012.
- Lapere, R.: Modeling the Impacts of Future Emission Scenarios on Air Pollution in Santiago, Chile, pp. 1–44, 2018.
- Lapere, R., Menut, L., Mailler, S., and Huneus, N.: Seasonal variation in atmospheric pollutants transport in central Chile: dynamics and consequences, *Atmos. Chem. Phys.*, 21, 6431–6454, <https://doi.org/10.5194/acp-21-6431-2021>, 2021.



- 620 Longo, K. M., Freitas, S. R., Pirre, M., Marecal, V., Rodrigues, L. F., Panetta, J., Alonso, M. F., Rosario, N. E., Moreira, D. S., Gacita, M. S., Arteta, J., Fonseca, R., Stockler, R., Katsurayama, D. M., Fazenda, A., and Bela, M.: The Chemistry CATT-BRAMS model (CCATT-BRAMS 4.5): a regional atmospheric model system for integrated air quality and weather forecasting and research, *Geosci. Model Dev.*, 6, 1389–1405, <https://doi.org/10.5194/gmd-6-1389-2013>, 2013.
- Mailler, S., Menut, L., Khvorostyanov, D., Valari, M., Couvidat, F., Siour, G., Turquety, S., Briant, R., Tuccella, P., Bessagnet, B., Colette, A., Letinois, L., Markakis, K., and Meleux, F.: CHIMERE-2017: from urban to hemispheric chemistry-transport modeling, *Geosci. Model Dev.*, 10, 2397–2423, <https://doi.org/10.5194/gmd-10-2397-2017>, 2017.
- 625 Marecal, V., Peuch, V.-H., Andersson, C., Andersson, S., Arteta, J., Beekmann, M., Benedictow, A., Bergström, R., Bessagnet, B., Cansado, A., Cheroux, F., Colette, A., Coman, A., Curier, R. L., van der Gon, H. A. C. D., Drouin, A., Elbern, H., Emili, E., Engelen, R. J., Eskes, H. J., Foret, G., Friese, E., Gauss, M., Giannaros, C., Guth, J., Joly, M., Jaumouille, E., Josse, B., Kadygrov, N., Kaiser, J. W., Krajsek, K., Kuenen, J., Kumar, U., Liora, N., Lopez, E., Malherbe, L., Martinez, I., Melas, D., Meleux, F., Menut, L., Moinat, P., Morales, T., Parmentier, J., Piacentini, A., Plu, M., Poupkou, A., Queguiner, S., Robertson, L., Rouil, L., Schaap, M., Segers, A., Sofiev, M., Tarasson, L., Thomas, M., Timmermans, R., a Valdebenito, van Velthoven, P., van Versendaal, R., Vira, J., and Ung, A.: A regional air quality forecasting system over Europe: the MACC-II daily ensemble production, *Geosci. Model Dev.*, 8, 2777–2813, <https://doi.org/10.5194/gmd-8-2777-2015>, 2015.
- 630 Marlier, M. E., Bonilla, E. X., and Mickley, L. J.: How Do Brazilian Fires Affect Air Pollution and Public Health?, *GeoHealth*, 4, e2020GH000331, <https://doi.org/https://doi.org/10.1029/2020GH000331>, 2020.
- Mazzeo, A., Huneus, N., Ordoñez, C., Orfanoz-Cheuquelaf, A., Menut, L., Mailler, S., Valari, M., van der Gon, H. D., Gallardo, L., Muñoz, R., Donoso, R., Galleguillos, M., Osses, M., and Tolvett, S.: Impact of residential combustion and transport emissions on air pollution in Santiago during winter, *Atmospheric Environment*, 190, 195–208, <https://doi.org/https://doi.org/10.1016/j.atmosenv.2018.06.043>, 2018.
- 640 Mendez-Espinosa, J. F., Belalcazar, L. C., and Betancourt, R. M.: Regional air quality impact of northern South America biomass burning emissions, *Atmospheric Environment*, 203, 131–140, <https://doi.org/https://doi.org/10.1016/j.atmosenv.2019.01.042>, 2019.
- Mendez-Espinosa, J. F., Rojas, N. Y., Vargas, J., Pachon, J. E., Belalcazar, L. C., and Ramírez, O.: Air quality variations in Northern South America during the COVID-19 lockdown, *Science of The Total Environment*, p. 141621, <https://doi.org/https://doi.org/10.1016/j.scitotenv.2020.141621>, 2020.
- 645 Menut, L., Bessagnet, B., Briant, R., Cholakian, A., Couvidat, F., Mailler, S., Pennel, R., Siour, G., Tuccella, P., Turquety, S., and Valari, M.: The CHIMERE v2020r1 online chemistry-transport model, *Geoscientific Model Development*, 14, 6781–6811, <https://doi.org/10.5194/gmd-14-6781-2021>, 2021.
- Nations, U.: *The World's Cities in 2018—Data Booklet*, 2018.
- Nedbor-Gross, R., Henderson, B., Perez-Peña, M., and Pachon, J. E.: Air quality modeling in Bogota, Colombia using local emissions and natural mitigation factor adjustment for re-suspended particulate matter, *Atmospheric Pollution Research*, 9, 95–104, <https://doi.org/http://dx.doi.org/10.1016/j.apr.2017.07.004>, 2018.
- 650 Pachon, J., Galvis, B., Lombana, O., Carmona, L., Fajardo, S., Rincon, A., Meneses, S., Chaparro, R., Nedbor-Gross, R., and Henderson, B.: Development and evaluation of a comprehensive atmospheric emission inventory for air quality modeling in the megacity of Bogota, *Atmosphere*, 9, <https://doi.org/10.3390/atmos9020049>, 2018.
- 655 Pereira, G., Shimabukuro, Y. E., Moraes, E. C., Freitas, S. R., Cardozo, F. S., and Longo, K. M.: Monitoring the transport of biomass burning emission in South America, *Atmospheric Pollution Research*, 2, 247–254, <https://doi.org/https://doi.org/10.5094/APR.2011.031>, 2011.



- Perez-Peña, M., Henderson, B., Nedbor-Gross, R., and Pachon, J.: Natural mitigation factor adjustment for re-suspended particulate matter emissions inventory for Bogota, Colombia, *Atmospheric Pollution Research*, 8, <https://doi.org/10.1016/j.apr.2016.07.006>, 2017.
- Petersen, A. K., Brasseur, G. P., Bouarar, I., Flemming, J., Gauss, M., Jiang, F., Kouznetsov, R., Kranenburg, R., Mijling, B., Peuch, V.-H., Pommier, M., Segers, A., Sofiev, M., Timmermans, R., van der A, R., Walters, S., Xie, Y., Xu, J., and Zhou, G.: Ensemble forecasts of air quality in eastern China – Part 2: Evaluation of the MarcoPolo–Panda prediction system, version 1, *Geoscientific Model Development*, 12, 1241–1266, <https://doi.org/10.5194/gmd-12-1241-2019>, 2019.
- 660 Puliafito, S. E., Allende, D., Pinto, S., and Castesana, P.: High resolution inventory of GHG emissions of the road transport sector in Argentina, *Atmospheric Environment*, 101, 303–311, <https://doi.org/https://doi.org/10.1016/j.atmosenv.2014.11.040>, 2015.
- 665 Puliafito, S. E., Allende, D. G., Castesana, P. S., and Ruggeri, M. F.: High-resolution atmospheric emission inventory of the argentine energy sector. Comparison with edgar global emission database, *Heliyon*, 3, e00489, <https://doi.org/https://doi.org/10.1016/j.heliyon.2017.e00489>, 2017.
- Ramírez, O., de la Campa, A. M. S., Amato, F., CatacolI, R. A., Rojas, N. Y., and de la Rosa, J.: Chemical composition and source apportionment of PM₁₀ at an urban background site in a high–altitude Latin American megacity (Bogota, Colombia), *Environmental Pollution*, 233, 142–155, <https://doi.org/https://doi.org/10.1016/j.envpol.2017.10.045>, 2018.
- 670 Ramírez-Romero, C., Jaramillo, A., Cordoba, M. F., Raga, G. B., Miranda, J., Alvarez-Ospina, H., Rosas, D., Amador, T., Kim, J. S., Yakobi-Hancock, J., Baumgardner, D., and Ladino, L. A.: African dust particles over the western Caribbean – Part I: Impact on air quality over the Yucatan Peninsula, *Atmos. Chem. Phys.*, 21, 239–253, <https://doi.org/10.5194/acp-21-239-2021>, 2021.
- Rodríguez-Villamizar, A. L., Rojas-Roa, Y. N., Blanco-Becerra, C. L., Herrera-Galindo, M. V., and Fernandez-Niño, A. J.: Short-Term Effects of Air Pollution on Respiratory and Circulatory Morbidity in Colombia 2011–2014: A Multi-City, Time-Series Analysis, <https://doi.org/10.3390/ijerph15081610>, 2018.
- 675 Rojas, N. Y., Mangones, S. C., Osses, M., Granier, C., Laengle, I. A., J. V. A., and Mendez, J. A.: Road transport exhaust emissions in Colombia. 1990–2020 trends and spatial disaggregation, *Transportation Research Part D: Transport and Environment*, 121, 103–113, <https://doi.org/https://doi.org/10.1016/j.trd.2023.103780>, 2023.
- 680 Romieu, I., Gouveia, N., Cifuentes, L. A., de Leon, A. P., Junger, W., Vera, J., Strappa, V., Hurtado-Díaz, M., Miranda-Soberanis, V., Rojas-Bracho, L., Carbajal-Arroyo, L., Tzintzun-Cervantes, G., and Committee, H. H. R.: Multicity study of air pollution and mortality in Latin America (the ESCALA study)[Abstract], Research report (Health Effects Institute), 2012.
- Saide, P. E., Mena-Carrasco, M., Tolvett, S., Hernandez, P., and Carmichael, G. R.: Air quality forecasting for winter-time PM_{2.5} episodes occurring in multiple cities in central and southern Chile, *Journal of Geophysical Research: Atmospheres*, 121, 558–575, <https://doi.org/https://doi.org/10.1002/2015JD023949>, 2016.
- 685 SEMARNAT and INECC: Calidad del Aire en la Cuenca Atmosferica de Tula, https://www.gob.mx/cms/uploads/attachment/file/571616/Calidad_del_Aire_Cuenca_Atm_de_Tula-FINAL.pdf, 2020.
- Silva-Quiroz, R., Rivera, A. L., Ordoñez, P., Gay-García, C., and Frank, A.: Atmospheric blockages as trigger of environmental contingencies in Mexico City, *Heliyon*, 5, e02099, <https://doi.org/https://doi.org/10.1016/j.heliyon.2019.e02099>, 2019.
- 690 Simpson, D., Benedictow, A., Berge, H., Bergström, R., Emberson, L. D., Fagerli, H., Flechard, C. R., Hayman, G. D., Gauss, M., Jonson, J. E., Jenkin, M. E., Nyiri, A., Richter, C., Semeena, V. S., Tsyro, S., Tuovinen, J.-P., Valdebenito, and Wind, P.: The EMEP MSC-W chemical transport model – technical description, *Atmospheric Chemistry and Physics*, 12, 7825–7865, <https://doi.org/10.5194/acp-12-7825-2012>, 2012.



- 695 Soares, J. and Sofiev, M.: A Global Wildfire Emission and Atmospheric Composition: Refinement of the Integrated System for Wild-Land
Fires IS4FIRES, pp. 253–258, Springer International Publishing, 2014.
- Sofiev, M.: Extended resistance analogy for construction of the vertical diffusion scheme for dispersion models, *Journal of Geophysical
Research: Atmospheres*, 107, ACH 10–1–ACH 10–8, <https://doi.org/https://doi.org/10.1029/2001JD001233>, 2002.
- Sofiev, M., Siljamo, P., Valkama, I., Ilvonen, M., and Kukkonen, J.: A dispersion modelling system SILAM and its evaluation against ETEX
data, *Atmospheric Environment*, 40, 674–685, <https://doi.org/https://doi.org/10.1016/j.atmosenv.2005.09.069>, 2006.
- 700 Sofiev, M., Vankevich, R., Lotjonen, M., Prank, M., Petukhov, V., Ermakova, T., Koskinen, J., and Kukkonen, J.: An operational system for
the assimilation of the satellite information on wild-land fires for the needs of air quality modelling and forecasting, *Atmos. Chem. Phys.*,
9, 6833–6847, <https://doi.org/10.5194/acp-9-6833-2009>, 2009.
- Sofiev, M., Genikhovich, E., Keronen, P., and Vesala, T.: Diagnosing the Surface Layer Parameters for Dispersion Models
within the Meteorological-to-Dispersion Modeling Interface, *Journal of Applied Meteorology and Climatology*, 49, 221–233,
705 <https://doi.org/https://doi.org/10.1175/2009JAMC2210.1>, 2010.
- Sofiev, M., Vira, J., Kouznetsov, R., Prank, M., Soares, J., and Genikhovich, E.: Construction of the SILAM Eulerian atmo-
spheric dispersion model based on the advection algorithm of Michael Galperin, *Geoscientific Model Development*, 8, 3497–3522,
<https://doi.org/10.5194/gmd-8-3497-2015>, 2015.
- Wiedinmyer, C., Akagi, S. K., Yokelson, R. J., Emmons, L. K., Al-Saadi, J. A., Orlando, J. J., and Soja, A. J.: The Fire INventory
710 from NCAR (FINN): a high resolution global model to estimate the emissions from open burning, *Geosci. Model Dev.*, 4, 625–641,
<https://doi.org/10.5194/gmd-4-625-2011>, 2011.
- Zhai, H., Huang, L., Emery, C., Zhang, X., Wang, Y., Yarwood, G., Fu, J. S., and Li, L.: Recommendations on benchmarks for
photochemical air quality model applications in China — NO₂, SO₂, CO and PM₁₀, *Atmospheric Environment*, 319, 120 290,
<https://doi.org/https://doi.org/10.1016/j.atmosenv.2023.120290>, 2024.

<https://doi.org/10.5194/egusphere-2024-815>

Preprint. Discussion started: 25 April 2024

© Author(s) 2024. CC BY 4.0 License.



715 **Appendix A: Evaluation scores**



Table A1. Metrics used for model evaluation.

Metric	Formula for each city, model and month
Model/Observation Ratio	$ratio = \frac{\bar{m}}{\bar{O}} = \frac{\sum_d m_d}{\sum_d O_d}$
Mean Bias	$BIAS = \frac{1}{N} \sum_d (m_d - O_d)$
Mean Normalized Bias	$MNBIAS = \frac{2}{N} \sum_d \frac{m_d - O_d}{m_d + O_d}$
Fractional Error	$FGF = \frac{2}{N} \sum_d \left \frac{m_d - O_d}{m_d + O_d} \right $
Root Mean Square Error	$RMSE = \sqrt{\frac{1}{N} \sum_d (m_d - O_d)^2}$
Correlation Coefficient	$R = \frac{1}{N} \frac{\sum_d (m_d - \bar{m})(O_d - \bar{O})}{\sigma_m \sigma_O}$
Coefficient of Variation	$CV = \frac{\bar{m}}{\sigma_m}$ or $CV = \frac{\bar{O}}{\sigma_O}$
Metric	Formula for each pixel and hour
Median of Models (Ensemble)	$MED = Median(\{m_i\})$ with $i \in N, 1 \leq i \leq 6$
Metric	Formula for each pixel
Median Absolute Deviation	$MAD = Median(\{ m_{i,d} - MED_d \})$ $i \in N, 1 \leq i \leq 6$ and $d \in N, 1 \leq d \leq 31$
Metric	Formula for each city, day and model
Observation	$O_d = \sum_{j,k} g_{A_{j,k}} M_{j,k}$ with $\sum_{j,k} g_{A_{j,k}} = 1$ $j, k \in N$ representing a specific pixel $g_{A_{j,k}}$ proportion of area of the pixel with the area of the polygon of the city



Table A2. NO₂ model evaluation scores (January / July)

		ENSEMBLE		Mean		CAM5		MPI		EMEP		CHIM		SILAM		USP	
NO ₂	City	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul
Model/Observations	Santiago	0.55	0.38	0.61	0.48	0.33	0.67	0.54	0.33	0.61	0.53	0.88	0.37	0.82	0.86	0.54	0.10
	Bogotá	0.37	0.34	0.38	0.37	0.24	0.13	0.39	0.41	0.52	0.49	0.30	0.21	0.42	0.36	0.39	0.53
	México	0.92	0.84	1.03	1.00	0.78	0.44	0.99	0.95	1.25	1.80	1.44	1.29	0.78	0.71		
	São Paulo	0.96	0.88	1.03	0.92	0.13	0.33	1.39	1.39	1.53	1.17	1.65	1.16	0.96	0.78	0.63	0.76
MNBIAS [%]	Santiago	-55.2	-87.8	-44.5	-66.6	-100	-44.0	-58.0	-97.4	-44.4	-57.8	-8.6	-98.1	-16.2	-14.5	-56.3	-161
	Bogotá	-91.2	-97.3	-87.5	-91.9	-120	-155	-85.7	-82.2	-63.3	-67.2	-102	-129	-78.9	-92.6	-90.0	-61.5
	México	-8.0	-17.3	3.2	1.4	-34.3	-84.6	0.5	-3.7	21.5	55.0	37.7	19.7	-27.0	-36.4		
	São Paulo	-4.3	-9.9	3.0	-4.9	-159	-112	31.2	33.2	40.4	18.3	48.9	21.1	-9.1	-25.6	-52.9	-31.2
RMSE [ppb]	Santiago	5.58	25.37	4.85	21.64	7.79	16.35	5.85	27.44	5.06	20.94	2.41	28.14	2.83	11.22	5.56	35.08
	Bogotá	11.78	10.04	11.50	9.67	14.31	13.06	11.32	9.04	9.19	7.98	12.07	11.83	10.90	9.71	11.39	7.55
	México	4.91	5.64	5.38	3.94	10.24	15.27	4.62	5.19	10.52	20.92	18.38	8.81	7.64	8.13		
	São Paulo	4.59	6.54	4.65	5.99	16.16	16.20	8.63	12.66	10.60	9.92	13.41	11.09	4.86	7.92	9.17	10.65
FGE	Santiago	0.55	0.88	0.44	0.67	1.01	0.45	0.59	0.97	0.46	0.61	0.18	0.98	0.23	0.24	0.57	1.61
	Bogotá	0.91	0.97	0.88	0.92	1.20	1.56	0.86	0.82	0.63	0.67	1.02	1.29	0.79	0.93	0.90	0.62
	México	0.16	0.20	0.15	0.13	0.42	0.85	0.13	0.17	0.22	0.55	0.39	0.28	0.30	0.37		
	São Paulo	0.24	0.24	0.23	0.22	1.60	1.14	0.34	0.38	0.42	0.34	0.52	0.36	0.29	0.35	0.61	0.50
R	Santiago	0.52	0.54	0.56	0.65	0.62	0.46	0.25	-0.01	0.11	0.12	0.50	0.36	0.51	0.47	0.39	0.63
	Bogotá	0.72	0.37	0.73	0.38	-0.12	0.52	0.74	0.41	0.66	0.20	0.23	0.11	0.54	0.42	0.62	0.19
	México	0.77	0.71	0.74	0.70	0.55	0.35	0.71	0.42	0.68	0.36	0.65	-0.20	0.78	0.77		
	São Paulo	0.68	0.49	0.66	0.55	0.34	0.43	0.60	0.60	0.56	0.33	0.58	-0.13	0.74	0.62	0.36	0.19

*Median: ensemble based on the median value of the models; CAM5: Copernicus Atmosphere Monitoring Service's (CAM5); MPI: WRF-Chem executed by MPIM; EMEP: European Monitoring and Evaluation Programme; CHIM: CHIMERE transport model; SILAM: System for Integrated modeling of Atmospheric composition; USP: WRF-Chem executed by University of São Paulo.

Table A3. O₃ model evaluation scores (January / July)

		ENSEMBLE		Mean		CAM5		MPI		EMEP		CHIM		SILAM		USP	
O ₃	City	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul
Model/Observations	Santiago	1.13	4.84	1.12	5.07	0.33	4.76	1.62	6.96	1.79	8.04	0.96	4.58	1.30	5.21	0.67	0.72
	Bogotá	1.38	1.36	1.52	1.38	0.34	0.34	2.56	1.74	2.68	2.74	1.51	1.37	1.03	0.72	0.94	1.36
	México	1.20	1.09	1.27	1.28	0.53	0.31	1.69	1.73	2.09	2.58	1.16	0.77	0.86	0.88		
	São Paulo	1.46	1.14	1.52	1.28	0.29	0.67	2.77	2.00	2.13	2.00	1.31	0.67	1.41	1.05	1.14	0.89
MNBIAS [%]	Santiago	10.3	133	10.0	136	-101	132	43.4	151	55.5	155	-5.0	137	23.1	121	-42.6	-19.2
	Bogotá	33.4	31.5	42.3	33.2	-96.0	-95.4	86.0	52.7	89.8	91.8	41.9	30.6	4.9	-32.8	-11.6	29.6
	México	21.8	8.7	26.8	25.0	-56.2	-103	52.9	52.5	68.2	86.8	14.6	-30.6	-11.2	-13.6		
	São Paulo	32.8	12.9	38.3	25.9	-106	-35.2	86.9	64.6	67.7	68.2	32.0	-36.3	22.9	-11.7	-0.8	-12.5
RMSE [ppb]	Santiago	3.73	15.52	3.40	16.24	15.10	15.27	14.85	23.21	17.41	28.14	2.85	14.62	8.05	18.95	9.03	2.56
	Bogotá	5.48	5.26	7.48	5.44	8.95	7.81	20.85	9.47	21.36	19.91	7.14	5.42	2.49	4.91	6.16	5.97
	México	10.57	9.94	12.05	12.96	12.83	22.25	19.56	26.13	31.91	52.61	11.37	13.24	10.17	9.97		
	São Paulo	14.69	7.43	15.70	8.11	19.42	7.52	45.58	17.33	28.80	16.97	15.52	9.11	15.24	11.14	13.28	6.41
FGE	Santiago	0.13	1.35	0.12	1.37	1.01	1.33	0.43	1.52	0.56	1.56	0.11	1.37	0.23	1.34	0.44	0.52
	Bogotá	0.34	0.34	0.43	0.35	0.96	0.95	0.86	0.53	0.90	0.92	0.42	0.38	0.17	0.44	0.41	0.36
	México	0.33	0.25	0.35	0.30	0.57	1.04	0.53	0.55	0.69	0.87	0.31	0.43	0.36	0.28		
	São Paulo	0.44	0.42	0.46	0.42	1.09	0.53	0.89	0.67	0.68	0.70	0.47	0.61	0.45	0.58	0.56	0.42
R	Santiago	0.64	-0.23	0.70	-0.15	0.42	-0.38	0.72	-0.22	0.63	-0.30	0.69	0.31	0.59	0.08	-0.14	-0.06
	Bogotá	0.39	-0.27	-0.01	-0.49	0.03	-0.11	-0.02	-0.27	-0.13	-0.27	0.60	-0.10	0.59	-0.31	-0.46	-0.30
	México	0.07	-0.08	0.09	-0.02	0.20	-0.01	0.34	0.08	0.07	-0.01	0.00	-0.26	-0.28	0.03		
	São Paulo	0.48	0.24	0.47	0.27	0.10	0.00	0.40	0.36	0.45	0.24	0.07	-0.28	0.55	0.23	0.24	0.30

*Median: ensemble based on the median value of the models; CAM5: Copernicus Atmosphere Monitoring Service's (CAM5); MPI: WRF-Chem executed by MPIM; EMEP: European Monitoring and Evaluation Programme; CHIM: CHIMERE transport model; SILAM: System for Integrated modeling of Atmospheric composition; USP: WRF-Chem executed by University of São Paulo.



Table A4. CO model evaluation scores (January / July)

		ENSEMBLE		Mean		CAM5		MPI		EMEP		CHIM		SILAM		USP	
CO	City	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul
Model/Observations	Santiago	1.03	0.54	1.26	0.82	1.11	1.06	0.92	0.36	1.32	0.74	1.83	0.47	2.01	2.07	0.49	0.13
	Bogotá	0.38	0.41	0.41	0.45	0.39	0.33	0.41	0.49	0.50	0.56	0.34	0.26	0.60	0.63	0.19	0.36
	México	0.98	1.06	1.23	1.40	0.60	0.48	0.84	0.94	1.82	2.92	2.09	1.76	0.99	1.15		
	São Paulo	0.76	0.67	0.78	0.72	0.43	0.45	0.99	0.90	0.97	0.76	0.91	0.74	1.08	1.03	0.33	0.40
MNBIAS [%]	Santiago	3.6	-55.2	23.1	-15.4	7.2	4.0	-9.4	-88.7	26.9	-26.8	58.6	-80.1	66.4	69.0	-67.9	-150
	Bogotá	-86.0	-81.5	-80.8	-74.2	-83.3	-99.2	-80.2	-66.7	-63.6	-54.2	-93.6	-114	-47.4	-45.5	-134	-94.4
	México	-5.3	5.4	16.3	33.7	-51.6	-71.1	-18.5	-5.7	53.2	94.4	59.6	46.0	-6.9	10.7		
	São Paulo	-29.5	-37.7	-26.3	-31.8	-83.1	-76.3	-3.1	-10.8	-6.0	-26.5	-10.7	-23.8	1.7	-2.8	-102	-84.1
RMSE [ppm]	Santiago	0.03	0.60	0.07	0.36	0.07	0.34	0.05	0.79	0.09	0.48	0.21	0.74	0.26	1.24	0.12	1.03
	Bogotá	0.46	0.36	0.44	0.34	0.46	0.41	0.44	0.32	0.37	0.27	0.47	0.45	0.33	0.24	0.58	0.40
	México	0.12	0.09	0.28	0.28	0.35	0.39	0.19	0.16	0.80	1.34	1.10	0.50	0.15	0.18		
	São Paulo	0.20	0.32	0.20	0.30	0.38	0.45	0.15	0.24	0.19	0.31	0.21	0.33	0.21	0.32	0.45	0.52
FGE	Santiago	0.10	0.60	0.23	0.32	0.19	0.24	0.19	0.90	0.27	0.43	0.59	0.80	0.66	0.69	0.68	1.50
	Bogotá	0.86	0.82	0.81	0.74	0.83	0.99	0.80	0.67	0.64	0.54	0.94	1.14	0.47	0.46	1.34	0.94
	México	0.12	0.12	0.19	0.34	0.52	0.71	0.22	0.17	0.53	0.94	0.60	0.46	0.17	0.18		
	São Paulo	0.33	0.41	0.31	0.37	0.83	0.77	0.19	0.27	0.25	0.38	0.27	0.36	0.26	0.37	1.02	0.84
R	Santiago	0.28	0.52	0.25	0.53	0.56	0.47	0.06	0.30	0.01	0.07	0.24	0.15	-0.09	0.30	0.49	0.19
	Bogotá	0.75	0.33	0.78	0.37	0.01	0.31	0.78	0.40	0.68	0.26	0.49	0.15	0.59	0.37	0.66	0.14
	México	0.85	0.83	0.74	0.79	0.79	0.63	0.71	0.46	0.58	0.51	0.76	-0.05	0.86	0.78		
	São Paulo	0.59	0.56	0.58	0.55	0.60	0.63	0.56	0.56	0.42	0.41	0.49	0.00	0.63	0.53	0.13	0.20

*Median: ensemble based on the median value of the models; CAM5: Copernicus Atmosphere Monitoring Service's (CAM5); MPI: WRF-Chem executed by MPIM; EMEP: European Monitoring and Evaluation Programme; CHIM: CHIMERE transport model; SILAM: System for Integrated modeling of Atmospheric composition; USP: WRF-Chem executed by University of São Paulo.

Table A5. SO₂ model evaluation scores (January / July)

		ENSEMBLE		Mean		CAM5		MPI		EMEP		CHIM		SILAM		USP	
SO ₂	City	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul
Model/Observations	Santiago	1.81	3.62	3.69	4.72	12.89	9.92	3.93	6.36	0.99	1.88	1.31	3.51	2.52	6.22	0.01	0.01
	Bogotá	0.77	0.68	0.79	0.70	0.72	0.58	1.30	1.23	1.27	1.00	0.41	0.38	0.92	0.88	0.00	0.00
	México	2.92	4.08	11.08	11.14	42.79	38.47	5.07	6.15	1.74	2.49	1.44	1.78	2.46	3.56		
	São Paulo	5.16	4.72	6.66	6.39	2.97	5.18	17.02	16.89	6.37	4.42	4.34	4.34	8.75	6.55	0.01	0.01
MNBIAS [%]	Santiago	56.3	112	113	129	170	162	114	143	-1.8	58.3	24.2	107	84.5	138	-197	-197
	Bogotá	-23.4	-33.0	-22.3	-30.0	-28.0	-47.8	25.8	25.0	21.1	3.9	-80.0	-81.3	-6.8	-7.2	-199	-199
	México	98.6	119	165	166	190	189	134	142	62.2	86.0	44.2	57.1	85.6	108		
	São Paulo	133	130	146	144	94.7	133	176	175	141	123	123	125	156	143	-197	-194
RMSE [ppb]	Santiago	1.28	4.62	3.97	6.53	17.46	15.76	4.56	9.61	0.30	1.72	0.74	4.38	2.32	10.00	1.47	1.78
	Bogotá	0.65	0.68	0.67	0.65	0.69	0.85	0.85	0.60	1.01	0.48	1.11	0.86	0.53	0.58	1.80	1.70
	México	10.16	9.82	48.69	31.49	198.74	116.27	19.96	16.11	4.88	4.76	3.98	2.72	8.02	8.48		
	São Paulo	4.03	4.99	5.49	7.31	1.92	5.80	15.55	21.96	5.28	4.93	3.26	4.96	7.63	7.88	1.02	1.30
FGE	Santiago	0.56	1.12	1.14	1.29	1.71	1.63	1.14	1.44	0.16	0.58	0.31	1.07	0.84	1.38	1.97	1.97
	Bogotá	0.33	0.35	0.35	0.32	0.33	0.50	0.32	0.31	0.37	0.21	0.80	0.81	0.27	0.21	1.99	1.99
	México	1.01	1.19	1.65	1.66	1.91	1.90	1.35	1.43	0.70	0.86	0.64	0.57	0.89	1.09		
	São Paulo	1.33	1.30	1.46	1.45	0.95	1.33	1.77	1.76	1.41	1.23	1.23	1.25	1.57	1.44	1.98	1.95
R	Santiago	-0.22	0.02	-0.19	0.21	-0.06	0.59	-0.32	0.16	-0.07	0.13	-0.45	0.22	-0.13	-0.24	0.12	0.49
	Bogotá	0.21	0.71	0.11	0.71	0.06	0.47	0.04	0.57	-0.02	0.55	0.32	0.39	0.28	0.20	0.11	0.47
	México	0.26	0.02	0.10	-0.04	0.08	-0.13	0.29	-0.14	0.24	0.11	0.19	0.06	0.14	-0.09		
	São Paulo	0.46	0.50	0.55	0.54	0.34	0.53	0.59	0.49	0.35	0.30	0.58	-0.02	0.53	0.47	0.28	0.11

*Median: ensemble based on the median value of the models; CAM5: Copernicus Atmosphere Monitoring Service's (CAM5); MPI: WRF-Chem executed by MPIM; EMEP: European Monitoring and Evaluation Programme; CHIM: CHIMERE transport model; SILAM: System for Integrated modeling of Atmospheric composition; USP: WRF-Chem executed by University of São Paulo.



Table A6. PM_{2.5} model evaluation scores (January / July)

		ENSEMBLE		Mean		CAM5		MPI		EMEP		CHIM		SILAM		USP	
PM _{2.5}	City	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul
Model/Observations	Santiago	0.50	0.38	0.63	1.15	1.21	1.83	0.60	0.36	0.14	0.29	0.55	0.22	0.96	3.79	0.30	0.12
	Bogotá	0.48	0.57	0.62	0.76	1.36	1.55	0.86	1.09	0.17	0.18	0.31	0.37	0.76	0.87	0.19	0.36
	México	0.84	1.22	1.51	1.87	4.16	4.91	0.66	1.20	0.13	0.15	0.52	0.65	1.85	2.06		
	São Paulo	1.78	1.62	2.13	1.94	1.64	2.22	2.16	1.84	1.44	1.10	2.37	1.99	4.41	3.63	0.76	0.83
MNBIAS [%]	Santiago	-65.7	-82.2	-44.1	17.8	18.3	58.7	-49.5	-87.1	-150	-103	-52.5	-128	-4.6	114	-105	-152
	Bogotá	-68.4	-54.6	-40.9	-26.8	35.8	41.4	-11.8	2.2	-139	-139	-96.7	-88.9	-24.1	-16.4	-134	-95.6
	México	-23.5	17.2	41.2	59.7	122	130	-35.1	17.6	-153	-147	-62.7	-44.1	48.1	60.8		
	São Paulo	56.7	53.6	70.8	66.8	44.9	79.0	74.8	64.0	31.4	12.1	84.2	73.9	114	96.7	-32.3	-11.2
RMSE [$\mu\text{g}/\text{m}^2$]	Santiago	10.69	39.84	8.26	18.31	6.68	46.91	9.22	40.96	17.68	44.27	8.72	49.58	6.48	160.50	14.46	53.10
	Bogotá	11.06	6.25	8.67	4.17	9.68	8.69	6.66	7.23	16.64	11.27	12.91	8.49	6.42	3.12	16.41	9.01
	México	7.12	8.71	13.77	19.17	80.13	84.48	10.06	7.99	23.03	18.92	11.56	10.38	25.96	27.52		
	São Paulo	14.68	16.64	20.80	24.48	13.05	30.84	20.43	22.09	11.48	11.19	26.31	29.61	66.31	82.01	9.86	10.01
FGE	Santiago	0.66	0.92	0.44	0.29	0.24	0.59	0.50	0.97	1.50	1.06	0.52	1.29	0.24	1.14	1.06	1.52
	Bogotá	0.68	0.55	0.42	0.30	0.40	0.41	0.30	0.40	1.39	1.40	0.97	0.89	0.26	0.22	1.34	0.96
	México	0.30	0.27	0.41	0.60	1.22	1.31	0.39	0.28	1.53	1.48	0.63	0.46	0.60	0.63		
	São Paulo	0.57	0.56	0.71	0.68	0.48	0.81	0.75	0.66	0.46	0.46	0.84	0.80	1.14	0.99	0.52	0.44
R	Santiago	0.51	0.05	0.58	0.42	0.56	0.52	0.20	0.02	0.24	0.28	0.39	0.07	0.10	0.30	0.32	0.36
	Bogotá	0.72	0.33	0.73	0.17	0.34	0.18	0.41	-0.42	0.80	0.40	0.51	0.12	0.71	0.63	0.60	0.18
	México	0.85	0.32	0.88	0.61	0.69	0.57	0.82	0.05	0.80	0.11	0.75	-0.22	0.85	0.59		
	São Paulo	0.52	0.57	0.54	0.61	0.46	0.47	0.48	0.52	0.34	0.44	0.48	-0.16	0.53	0.62	0.20	0.24

*Median: ensemble based on the median value of the models; CAM5: Copernicus Atmosphere Monitoring Service's (CAM5); MPI: WRF-Chem executed by MPIM; EMEP: European Monitoring and Evaluation Programme; CHIM: CHIMERE transport model; SILAM: System for Integrated modeling of Atmospheric composition; USP: WRF-Chem executed by University of São Paulo.

Table A7. PM₁₀ model evaluation scores (January / July)

		ENSEMBLE		Mean		CAM5		MPI		EMEP		CHIM		SILAM		USP	
PM ₁₀	City	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul
Model/Observations	Santiago	0.25	0.34	0.32	0.88	0.60	1.51	0.29	0.23	0.21	0.37	0.23	0.17	0.47	2.69	0.11	0.07
	Bogotá	0.28	0.31	0.37	0.45	0.76	0.83	0.59	0.83	0.20	0.18	0.14	0.18	0.39	0.44	0.08	0.14
	México	0.52	0.83	1.01	1.25	2.79	3.19	0.38	0.88	0.36	0.44	0.26	0.34	1.08	1.10		
	São Paulo	1.38	1.34	1.71	1.64	1.32	1.87	1.45	1.34	1.72	1.37	1.55	1.55	3.70	3.10	0.43	0.50
MNBIAS [%]	Santiago	-119	-89.6	-101	-7.7	-49.0	42.6	-108	-116	-128	-87.2	-121	-143	-70.8	90.5	-159	-170
	Bogotá	-110	-104	-88.8	-74.2	-22.7	-18.3	-50.2	-27.4	-133	-138	-145	-137	-85.7	-77.8	-170	-149
	México	-68.5	-19.0	0.4	22.6	93.7	103	-85.1	-13.6	-96.3	-76.6	-117	-94.1	-2.3	4.1		
	São Paulo	31.1	32.6	49.6	48.7	22.0	60.9	36.0	30.0	44.8	23.5	44.3	48.9	104	87.6	-82.6	-62.0
RMSE [$\mu\text{g}/\text{m}^2$]	Santiago	43.56	72.44	39.53	33.12	24.59	54.51	41.34	81.91	45.50	70.55	42.96	92.60	32.16	166.82	50.95	95.38
	Bogotá	35.53	24.67	31.77	20.45	16.16	10.52	23.31	18.52	39.47	29.35	39.99	28.69	30.82	20.45	45.32	30.52
	México	26.65	12.02	10.84	15.73	98.60	104.79	34.17	13.64	34.21	27.20	36.93	29.98	18.30	19.14		
	São Paulo	15.21	19.11	23.97	30.66	15.34	38.42	19.56	24.10	26.44	32.33	23.00	32.13	90.06	105.60	20.61	20.91
FGE	Santiago	1.19	0.95	1.01	0.33	0.49	0.46	1.09	1.22	1.29	0.89	1.22	1.43	0.71	0.90	1.60	1.70
	Bogotá	1.10	1.04	0.89	0.74	0.30	0.27	0.54	0.52	1.34	1.38	1.46	1.38	0.86	0.78	1.70	1.49
	México	0.68	0.26	0.19	0.26	0.94	1.03	0.85	0.29	0.96	0.77	1.17	0.94	0.31	0.29		
	São Paulo	0.34	0.42	0.50	0.53	0.34	0.64	0.44	0.47	0.50	0.54	0.47	0.56	1.04	0.90	0.83	0.70
R	Santiago	0.36	-0.12	0.41	0.42	0.41	0.51	0.20	-0.35	0.04	0.13	0.15	0.14	0.07	0.38	0.13	0.45
	Bogotá	0.75	0.40	0.70	0.06	0.39	0.20	0.31	-0.17	0.74	0.14	0.52	0.28	0.66	0.43	0.66	0.02
	México	0.65	0.53	0.72	0.44	0.55	0.20	0.30	0.23	0.60	0.15	0.61	-0.23	0.73	0.42		
	São Paulo	0.48	0.46	0.54	0.54	0.39	0.48	0.10	0.14	0.38	0.37	0.46	-0.05	0.56	0.55	0.20	0.21

ENSEMBLE: ensemble based on the median value of the models; CAM5: Copernicus Atmosphere Monitoring Service's (CAM5); MPI: WRF-Chem executed by MPIM; EMEP: European Monitoring and Evaluation Programme; CHIM: CHIMERE transport model; SILAM: System for Integrated modeling of Atmospheric composition; USP: WRF-Chem executed by University of São Paulo.



Table A8. Coefficient of Variation (CV) per city during January and July

City	NO ₂	O ₃	CO	SO ₂	PM _{2.5}
Santiago	33% 57%	51% 49%	45% 86%	130% 78%	65% 136%
Bogota	24% 44%	62% 63%	35% 33%	64% 67%	78% 73%
São Paulo	56% 46%	57% 57%	42% 38%	88% 88%	59% 50%
Mexico	31% 47%	48% 72%	54% 63%	166% 149%	111% 105%



Appendix B: Air quality observations

Table B1. Stations availability and location for México.

		Obs		ENSEMBLE		CAMS		MPI		EMEP		CHIM		SILAM		USP	
		Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul
CO																	
México	Number Stations	21	24	21	24	21	24	21	24	21	24	21	24	21	24		
	Availability [%]	96.77	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100		
	CV	0.21	0.26	0.28	0.20	0.25	0.19	0.22	0.13	0.35	0.26	0.48	0.15	0.33	0.33		
NO₂																	
México	Number Stations	24	24	24	24	24	24	24	24	24	24	24	24	24	24		
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100		
	CV	0.24	0.22	0.28	0.19	0.44	0.50	0.21	0.13	0.29	0.21	0.37	0.12	0.36	0.35		
O₃																	
México	Number Stations	21	29	21	29	21	29	21	29	21	29	21	29	21	29		
	Availability [%]	96.77	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100		
	CV	0.31	0.19	0.23	0.21	0.15	0.14	0.26	0.21	0.35	0.22	0.29	0.34	0.25	0.28		
PM₁₀																	
México	Number Stations	17	24	17	24	17	24	17	24	17	24	17	24	17	24		
	Availability [%]	96.77	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100		
	CV	0.29	0.20	0.52	0.26	0.28	0.22	0.24	0.27	0.48	0.29	0.51	0.23	0.46	0.40		
PM_{2.5}																	
México	Number Stations	14	16	14	16	14	16	14	16	14	16	14	16	14	16		
	Availability [%]	96.77	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100		
	CV	0.35	0.23	0.50	0.29	0.28	0.22	0.24	0.20	0.33	0.16	0.52	0.29	0.48	0.43		
SO₂																	
México	Number Stations	23	26	23	26	23	26	23	26	23	26	23	26	23	26		
	Availability [%]	96.77	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100		
	CV	0.59	0.29	0.35	0.20	0.25	0.09	0.27	0.13	0.32	0.13	0.37	0.18	0.34	0.27		



Table B2. Stations availability and location for Bogotá

		Obs		ENSEMBLE		CAMS		MPI		EMEP		CHIM		SILAM		USP	
		Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul
CO																	
Bogotá	Number Stations	7	8	7	8	7	8	7	8	7	8	7	8	7	8	7	8
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.28	0.15	0.16	0.15	0.10	0.13	0.17	0.12	0.24	0.18	0.21	0.15	0.17	0.22	0.32	0.27
NO₂																	
Bogotá	Number Stations	8	7	8	7	8	7	8	7	8	7	8	7	8	7	8	7
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.22	0.17	0.19	0.18	0.44	0.59	0.18	0.10	0.27	0.13	0.31	0.34	0.15	0.17	0.35	0.27
O₃																	
Bogotá	Number Stations	10	11	10	11	10	11	10	11	10	11	10	11	10	11	10	11
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.24	0.23	0.10	0.12	0.08	0.11	0.21	0.19	0.11	0.15	0.18	0.15	0.14	0.26	0.35	0.21
PM₁₀																	
Bogotá	Number Stations	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.27	0.22	0.30	0.20	0.15	0.21	0.31	0.52	0.34	0.19	0.25	0.12	0.24	0.24	0.32	0.31
PM_{2.5}																	
Bogotá	Number Stations	9	10	9	10	9	10	9	10	9	10	9	10	9	10	9	10
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.33	0.18	0.30	0.18	0.16	0.22	0.28	0.40	0.33	0.21	0.27	0.10	0.26	0.28	0.32	0.30
SO₂																	
Bogotá	Number Stations	7	6	7	6	7	6	7	6	7	6	7	6	7	6	7	6
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.29	0.34	0.17	0.14	0.10	0.14	0.21	0.16	0.34	0.21	0.24	0.35	0.20	0.15	0.29	0.35



Table B3. Stations availability and location for Santiago

		Obs		ENSEMBLE		CAMS		MPI		EMEP		CHIM		SILAM		USP	
		Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul
CO																	
Santiago	Number Stations	8	9	8	9	8	9	8	9	8	9	8	9	8	9	8	9
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.10	0.34	0.10	0.16	0.30	0.29	0.17	0.14	0.14	0.14	0.17	0.27	0.15	0.22	0.11	0.18
NO₂																	
Santiago	Number Stations	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.23	0.31	0.14	0.14	0.36	0.32	0.22	0.19	0.18	0.13	0.14	0.21	0.17	0.23	0.21	0.25
O₃																	
Santiago	Number Stations	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.14	0.57	0.16	0.22	0.16	0.21	0.25	0.14	0.12	0.24	0.18	0.25	0.23	0.50	0.25	0.39
PM₁₀																	
Santiago	Number Stations	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.20	0.39	0.16	0.20	0.25	0.22	0.26	0.21	0.16	0.26	0.18	0.22	0.23	0.28	0.21	0.19
PM_{2.5}																	
Santiago	Number Stations	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.22	0.37	0.16	0.14	0.25	0.22	0.22	0.18	0.14	0.20	0.18	0.24	0.27	0.30	0.21	0.19
SO₂																	
Santiago	Number Stations	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.14	0.20	0.16	0.12	0.12	0.16	0.27	0.21	0.13	0.23	0.24	0.15	0.18	0.38	0.11	0.16



Table B4. Stations availability and location for São Paulo

		Obs		ENSEMBLE		CAMS		MPI		EMEP		CHIM		SILAM		USP	
		Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul
CO																	
São Paulo	Number Stations	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.26	0.31	0.28	0.29	0.39	0.43	0.25	0.33	0.31	0.38	0.38	0.27	0.40	0.47	0.41	0.36
NO₂																	
São Paulo	Number Stations	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.36	0.32	0.30	0.27	1.18	0.91	0.30	0.33	0.28	0.31	0.34	0.22	0.44	0.47	0.48	0.48
O₃																	
São Paulo	Number Stations	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.40	0.38	0.35	0.41	0.35	0.35	0.36	0.41	0.32	0.35	0.36	0.36	0.46	0.77	0.46	0.44
PM₁₀																	
São Paulo	Number Stations	23	22	23	22	23	22	23	22	23	22	23	22	23	22	23	22
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.42	0.39	0.29	0.31	0.34	0.37	0.28	0.37	0.40	0.64	0.40	0.35	0.53	0.74	0.46	0.41
PM_{2.5}																	
São Paulo	Number Stations	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.51	0.49	0.32	0.33	0.35	0.38	0.29	0.37	0.39	0.51	0.40	0.35	0.59	0.85	0.46	0.41
SO₂																	
São Paulo	Number Stations	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
	Availability [%]	100	100	100	100	96.67	96.67	100	100	96.67	96.67	83.33	66.67	100	100	96.67	96.67
	CV	0.40	0.41	0.29	0.30	0.28	0.38	0.32	0.42	0.34	0.42	0.31	0.43	0.35	0.45	0.33	0.36



Appendix C: Particular hourly simulations

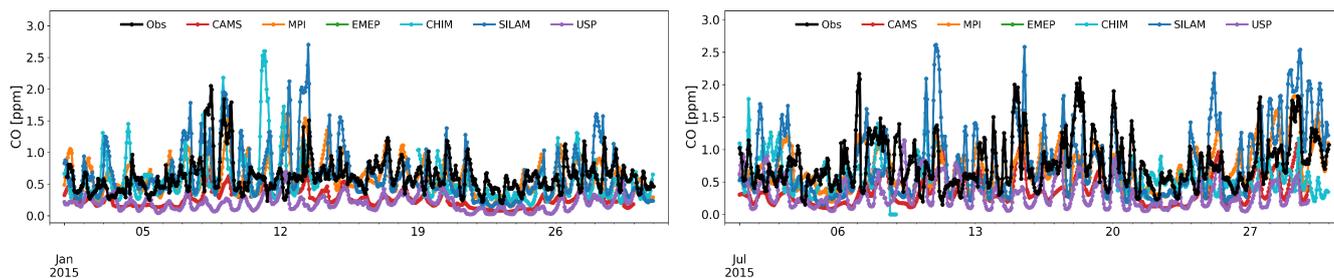


Figure C1. Hourly CO simulations in São Paulo for January and July of 2015

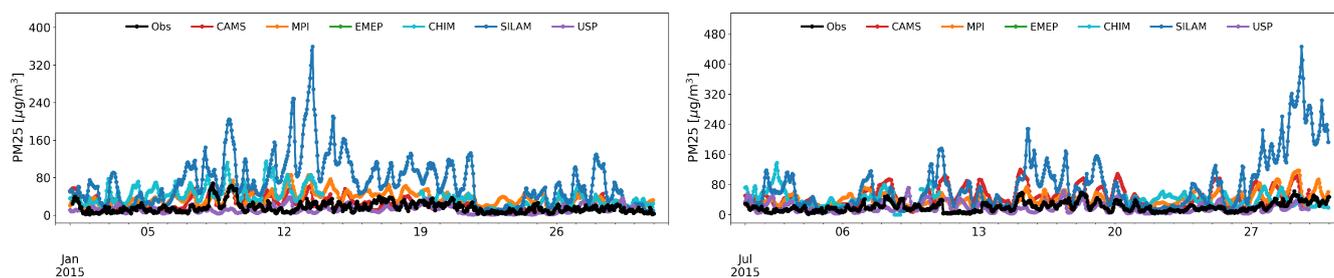


Figure C2. Hourly PM_{2.5} simulations in São Paulo for January and July of 2015

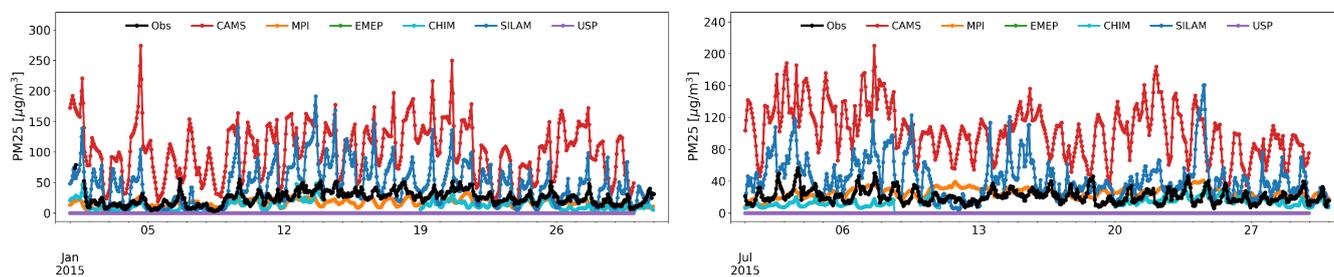


Figure C3. Hourly PM_{2.5} simulations in México City for January and July of 2015



Appendix D: Simulation of all models

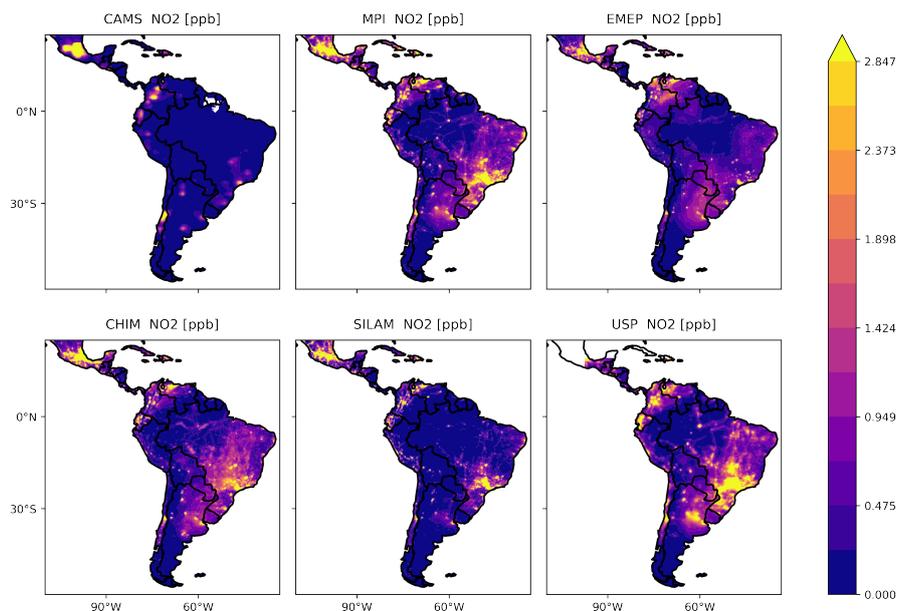


Figure D1. NO₂ simulations of January 2015 for all models

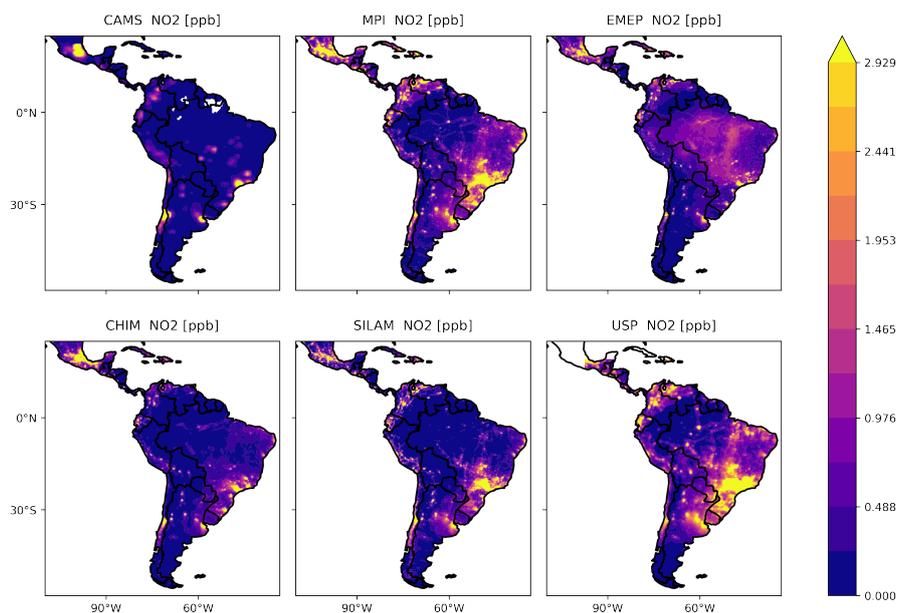


Figure D2. NO₂ simulations of July 2015 for all models

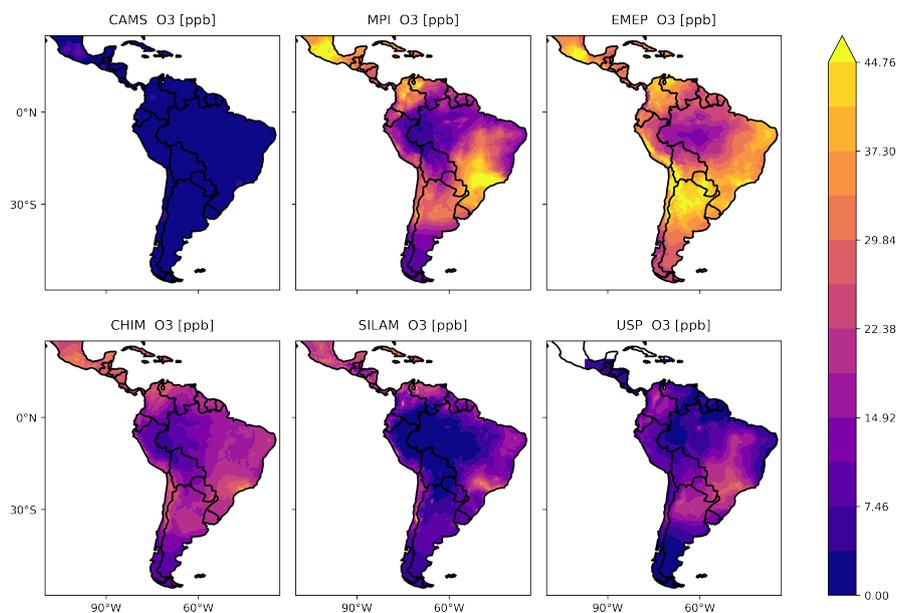


Figure D3. O₃ simulations of January 2015 for all models

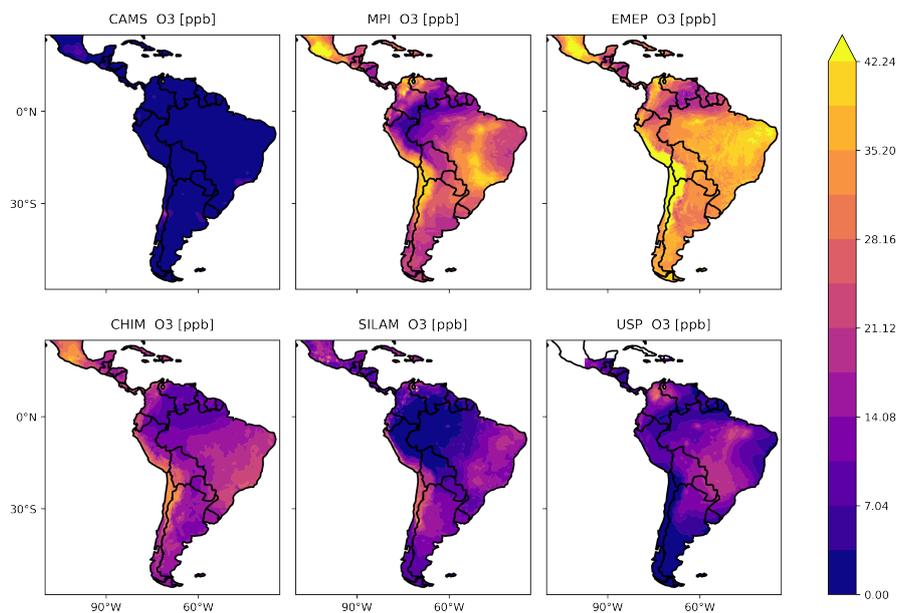


Figure D4. O₃ simulations of July 2015 for all models

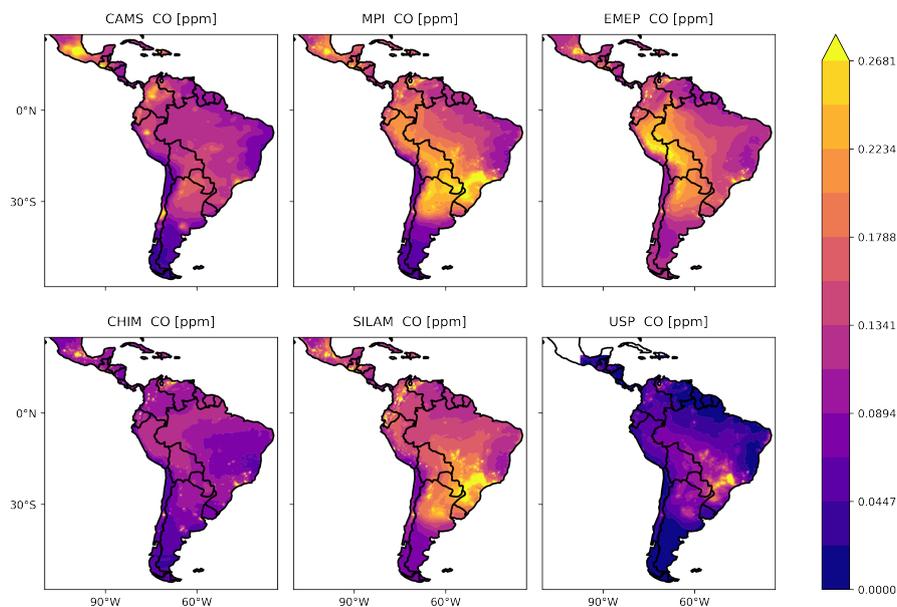


Figure D5. CO simulations of January 2015 for all models

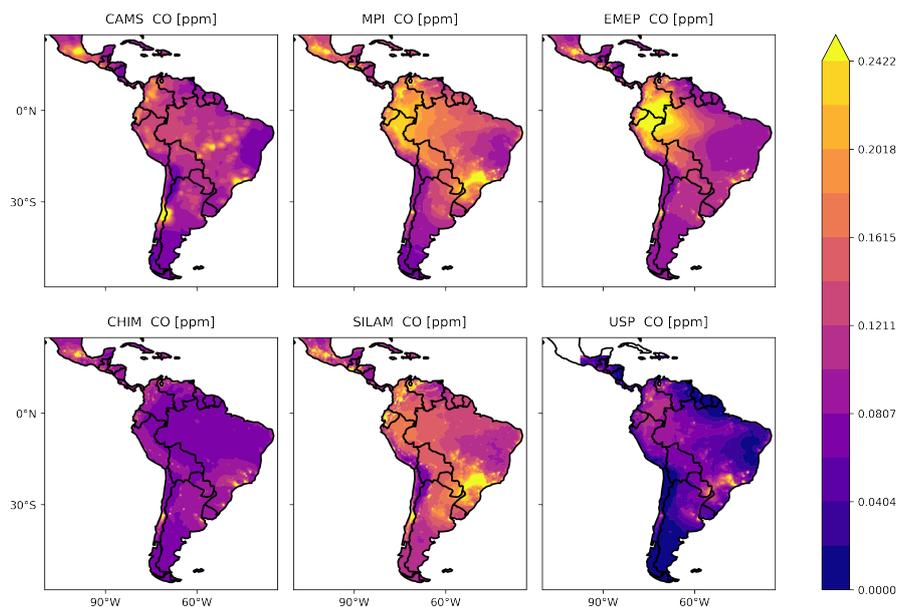


Figure D6. CO simulations of July 2015 for all models

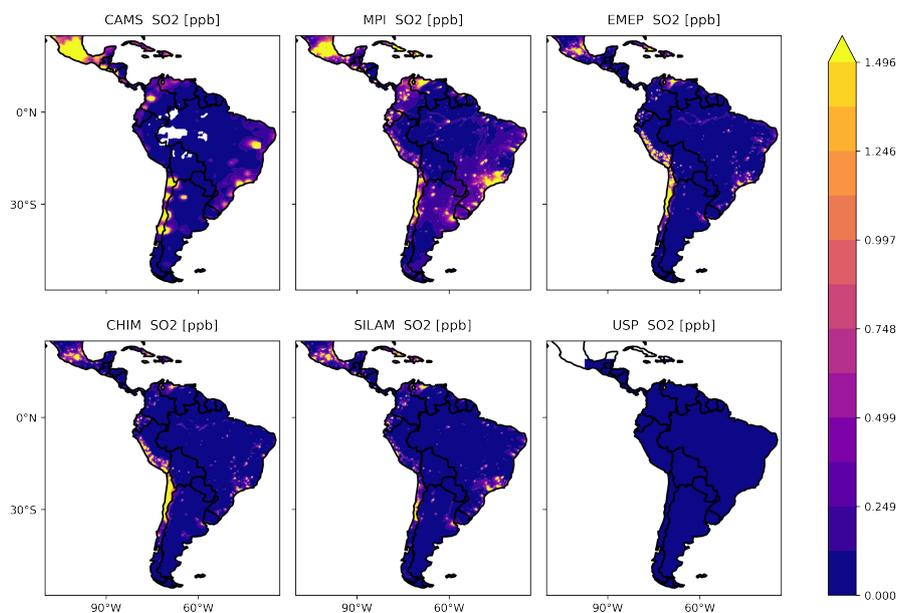


Figure D7. SO₂ simulations of January 2015 for all models

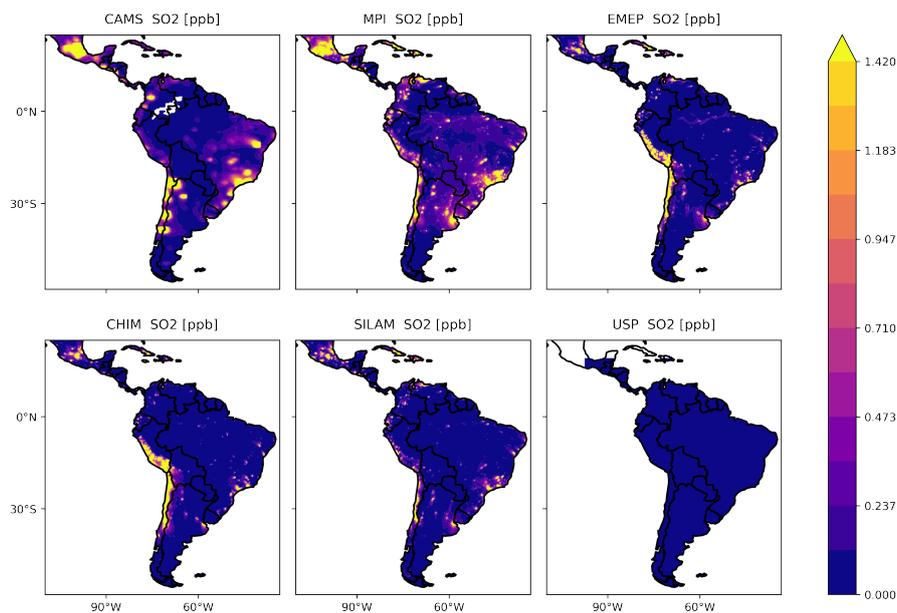


Figure D8. SO₂ simulations of July 2015 for all models

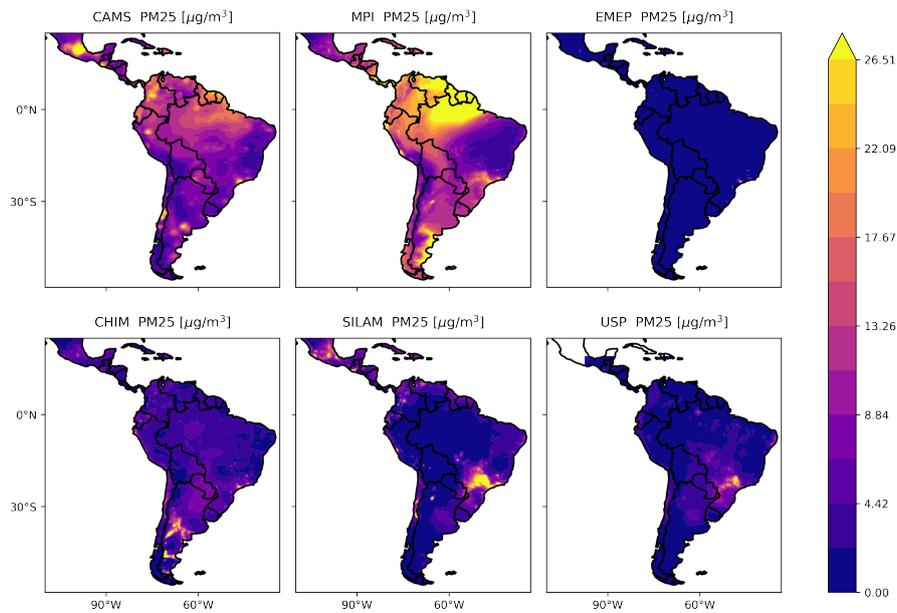


Figure D9. PM_{2.5} simulations of January 2015 for all models

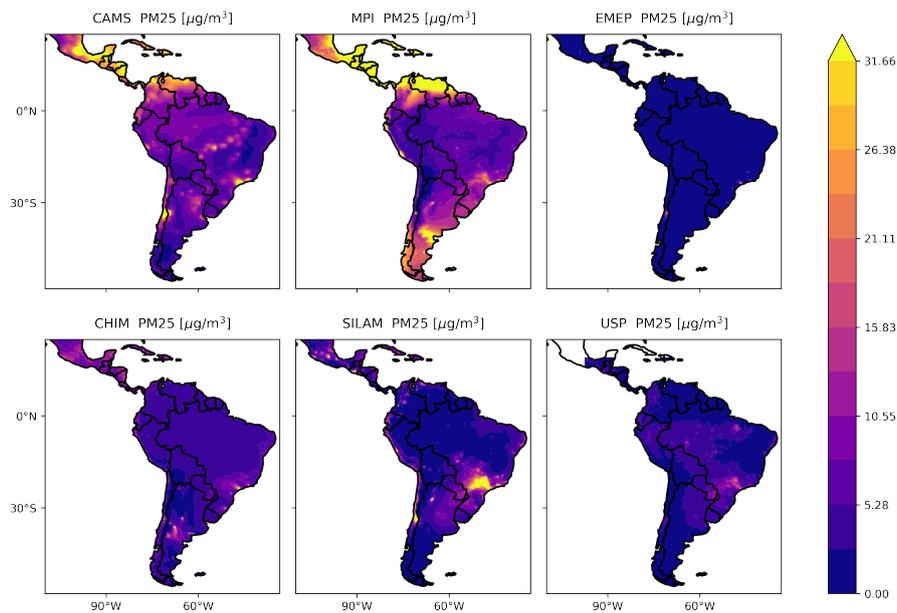


Figure D10. PM_{2.5} simulations of July 2015 for all models



Appendix E: Model deviations

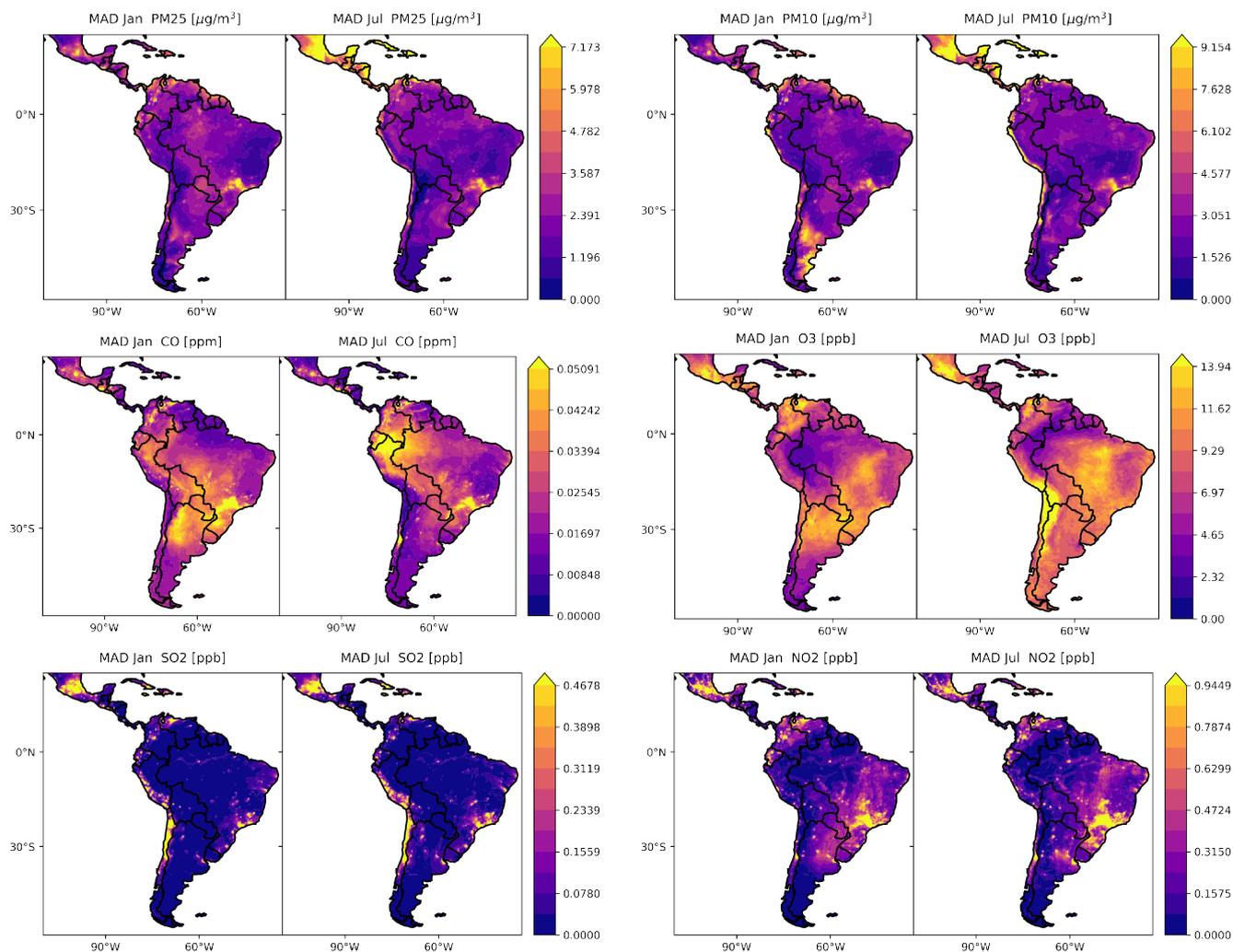


Figure E1. Median absolute deviation of the models with respect to the ensemble for PM₁₀, PM_{2.5}, O₃, CO, SO₂ and NO₂ in LAC for January and July 2015

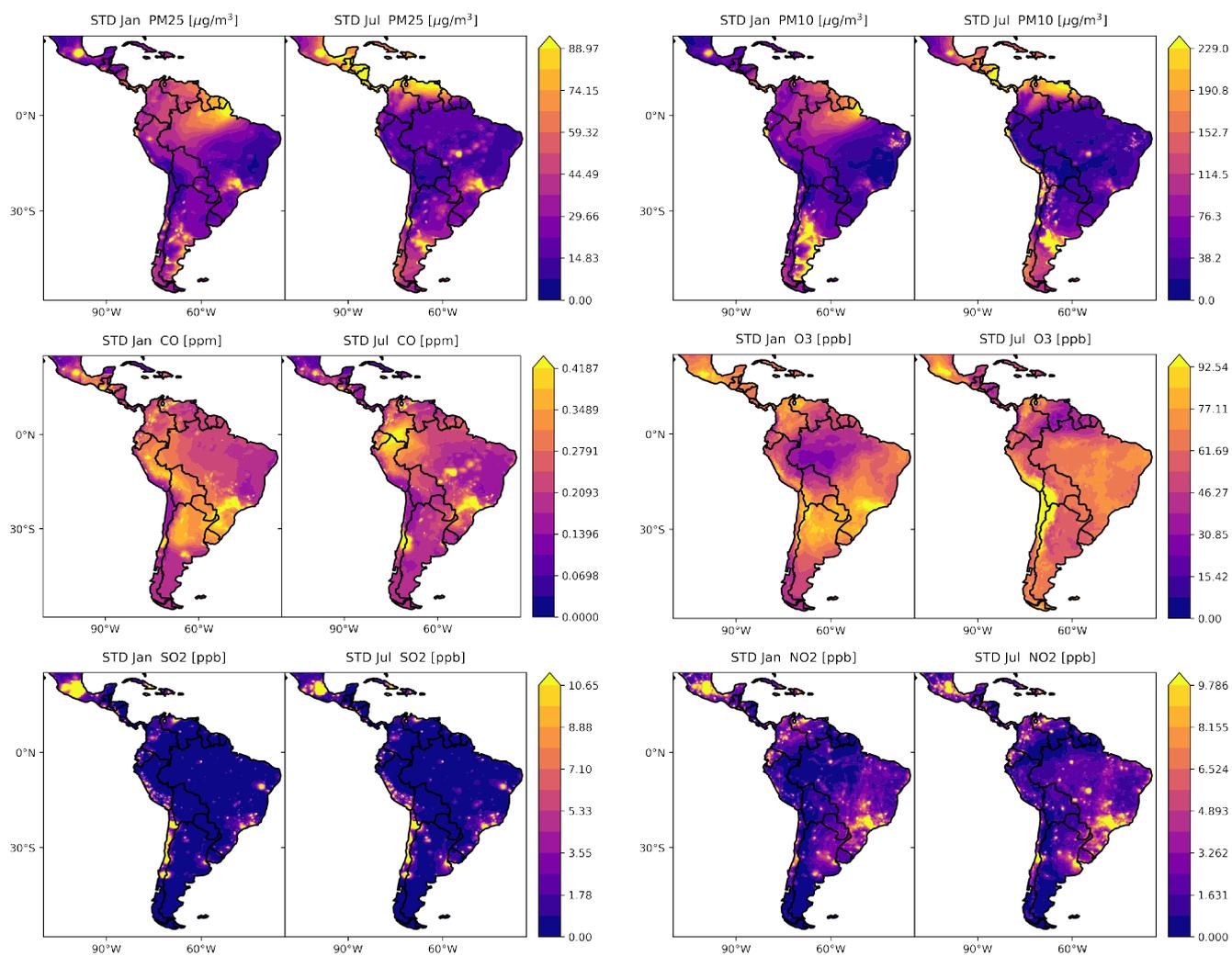


Figure E2. Standard deviation of the models with respect to their mean for PM₁₀, PM_{2.5}, O₃, CO, SO₂ and NO₂ in LAC for January and July 2015