Supplement of

Air quality simulation with WRF-Chem over south-eastern Brazil, Part I: Model description and evaluation using surface and satellite data

Noelia Rojas Benavente et al.

Correspondence to: Noelia Rojas Benavente (rnoeliab@usp.br)

1 Figures and Tables from the manuscript

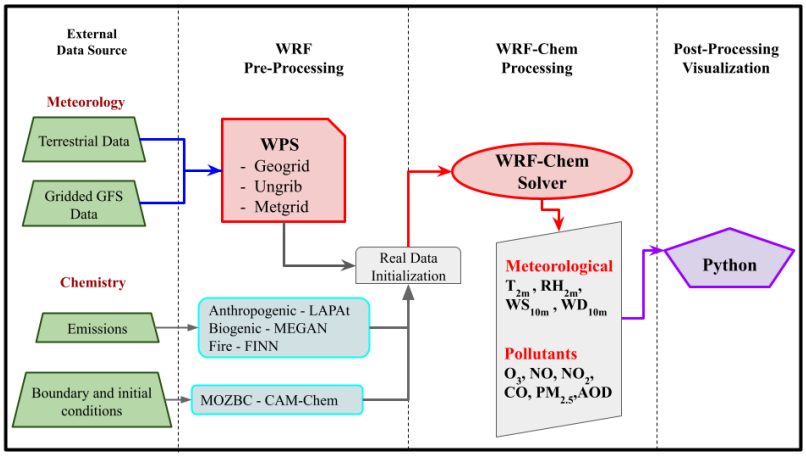

Figure S 1: Flowchart representing the basic steps for running WRF-Chem model. Regional (MEGAN and FINN) and local (LAPAt) emissions are considered to include the chemical and aerosol modules in the WRF modeling.

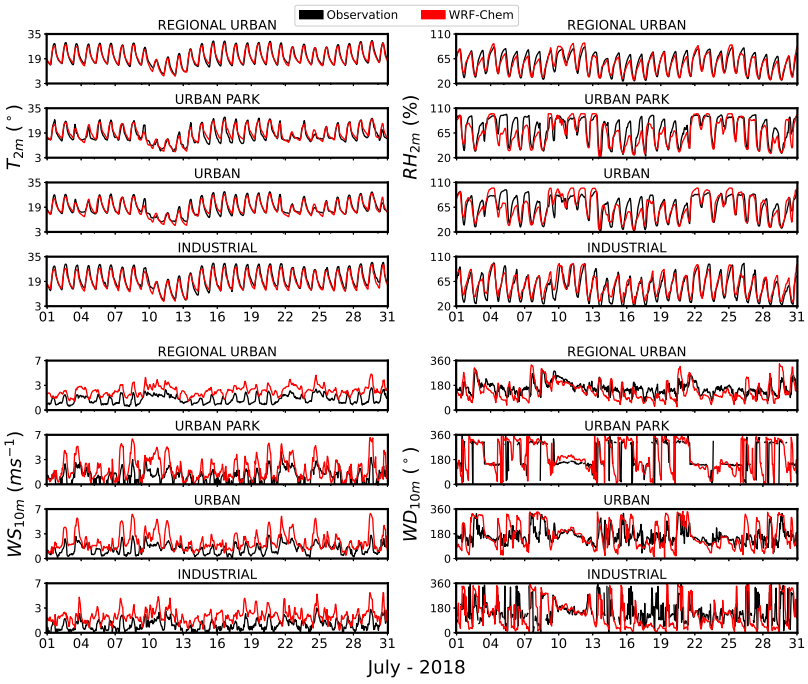

Figure S 2: The observed and simulated T_{2m}, RH_{2m}, WS_{10m} and WD_{10m} variability in the four categories of land use: regional urban, urban park, urban and industrial for July 2018. The measured and simulated results are represented by the black and red line, respectively.

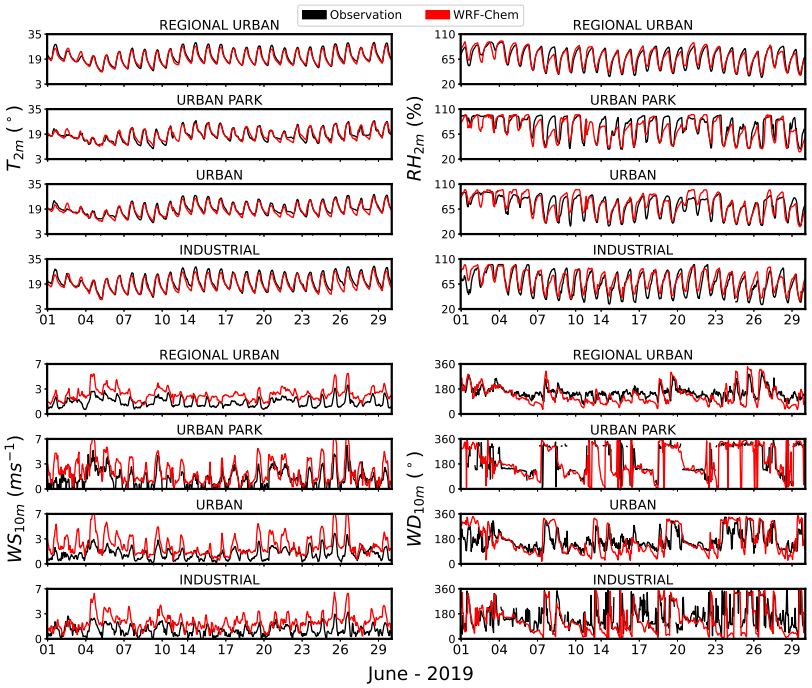

Figure S 3: The observed and simulated T_{2m}, RH_{2m}, WS_{10m} and WD_{10m} variability in the four categories of land use: regional urban, urban park, urban and industrial for June 2019. The measured and simulated results are represented by the black and red line, respectively.

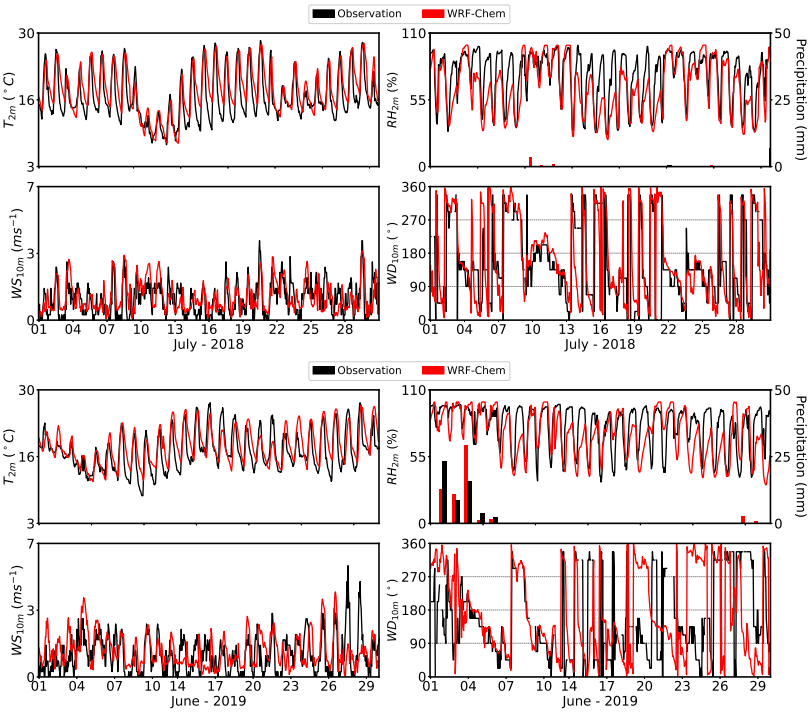

Figure S 4: The observed and simulated T_{2m}, RH_{2m}, WS_{10m} and WD_{10m} variability at IAG-USP station during July 2018 and June 2019. The bars represent the daily accumulated precipitation. The measured and simulated results are represented by the black and red line, respectively.

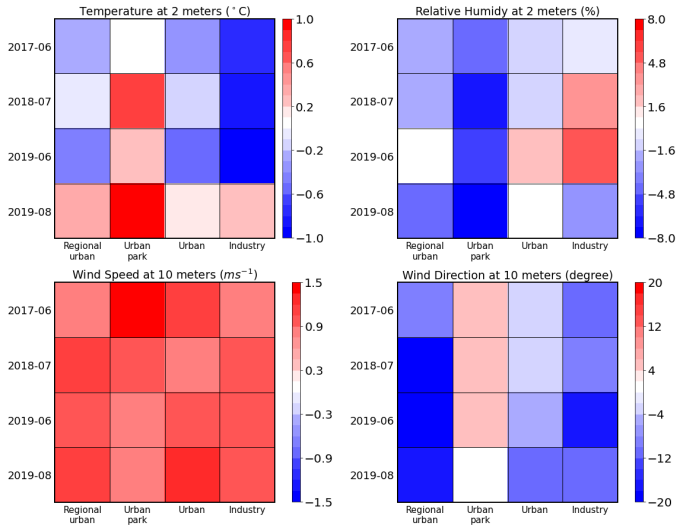

Figure S 5: Statistical histogram of mean bias for T_{2m}, RH_{2m}, WS_{10m} and WD_{10m} in the four categories of land use: regional urban, urban park, urban and industrial for all periods.

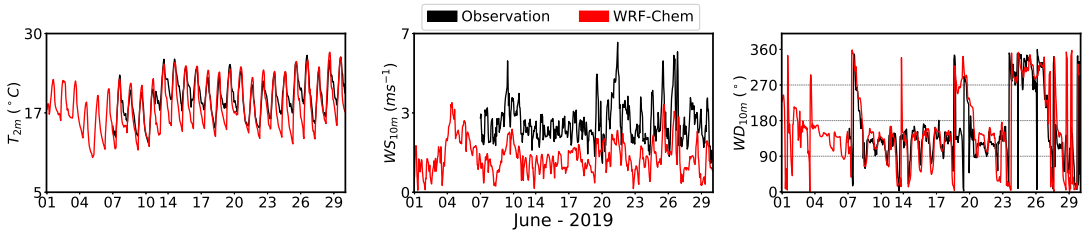

Figure S 6: The observed and simulated T_{2m}, WS_{10m} and WD_{10m} variability at Botucatu station during June 2019. The measured and simulated results are represented by the black and red line, respectively.

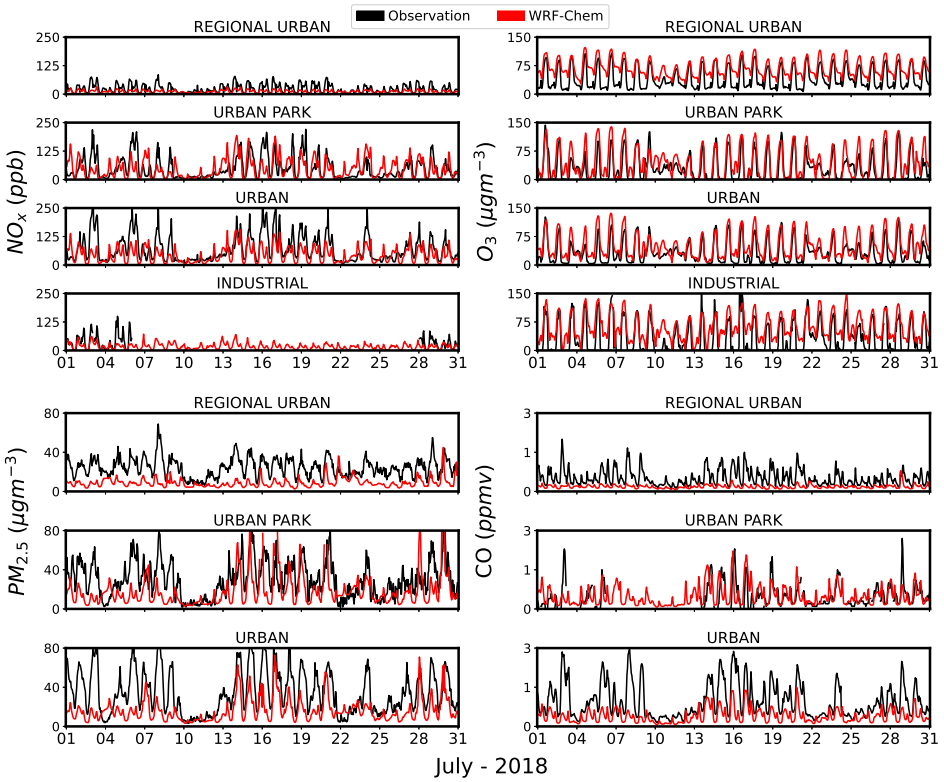

Figure S 7: The observed and simulated NO_x, O_3, $PM_{2.5}$ and CO concentrations variability in the four categories of land use: regional urban, urban park, urban and industrial for July 2018. The measured and simulated results are represented by the black and red line, respectively.

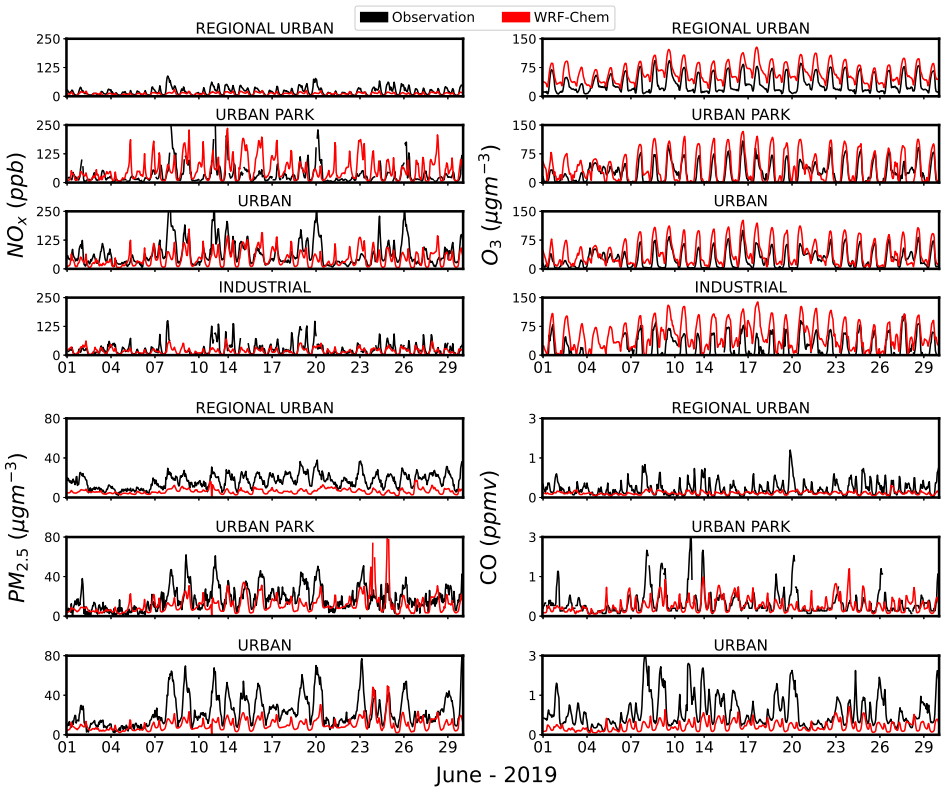

Figure S 8: The observed and simulated NO_x, O_3, $PM_{2.5}$ and CO concentrations variability in the four categories of land use: regional urban, urban park, urban and industrial for June 2019. The measured and simulated results are represented by the black and red line, respectively.

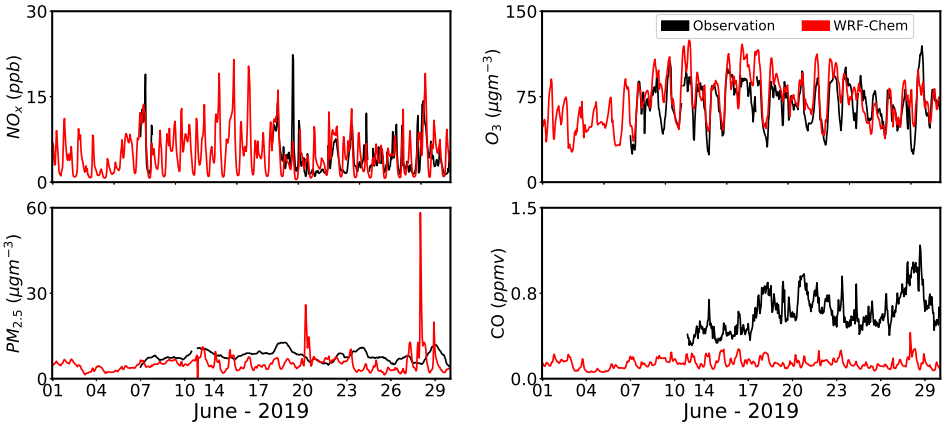

Figure S 9: The observed and simulated NO_x, O_3, $PM_{2.5}$ and CO concentrations variability at Botucatu station during June 2019. The measured and simulated results are represented by the black and red line, respectively.

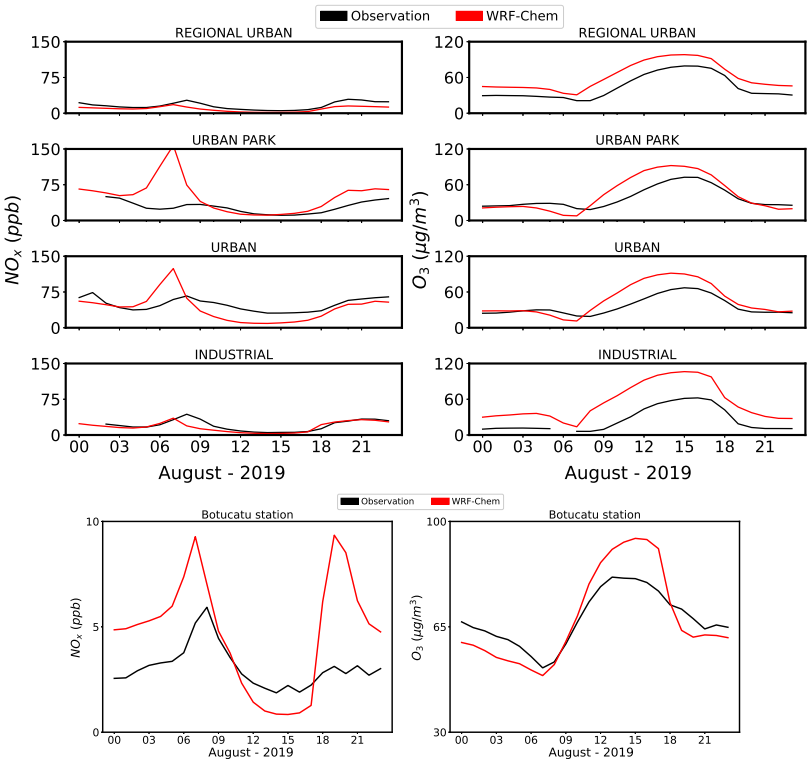

Figure S 10: Simulated and observed hourly mean NO_x and O_3 concentrations in the four categories of land use: regional urban, urban park, urban and industrial over the State of São Paulo and at Botucatu station for August 2019. The measured and simulated results are represented by the black and red line, respectively.

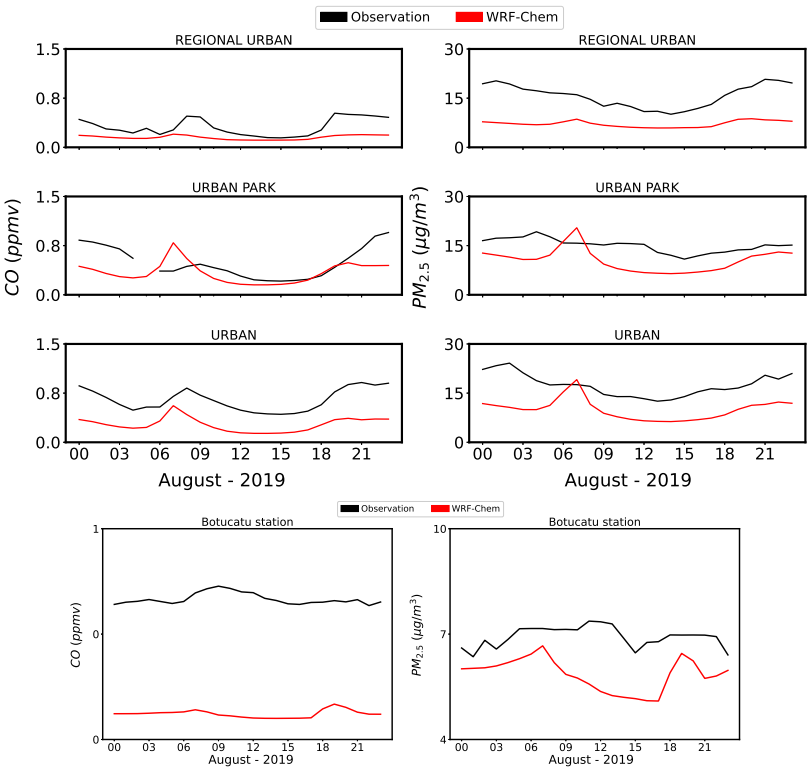

Figure S 11: Simulated and observed hourly mean CO and $PM_{2.5}$ concentrations in the four categories of land use: regional urban, urban park, urban and industrial over the State of São Paulo and at Botucatu station for August 2019. The measured and simulated results are represented by the black and red line, respectively.

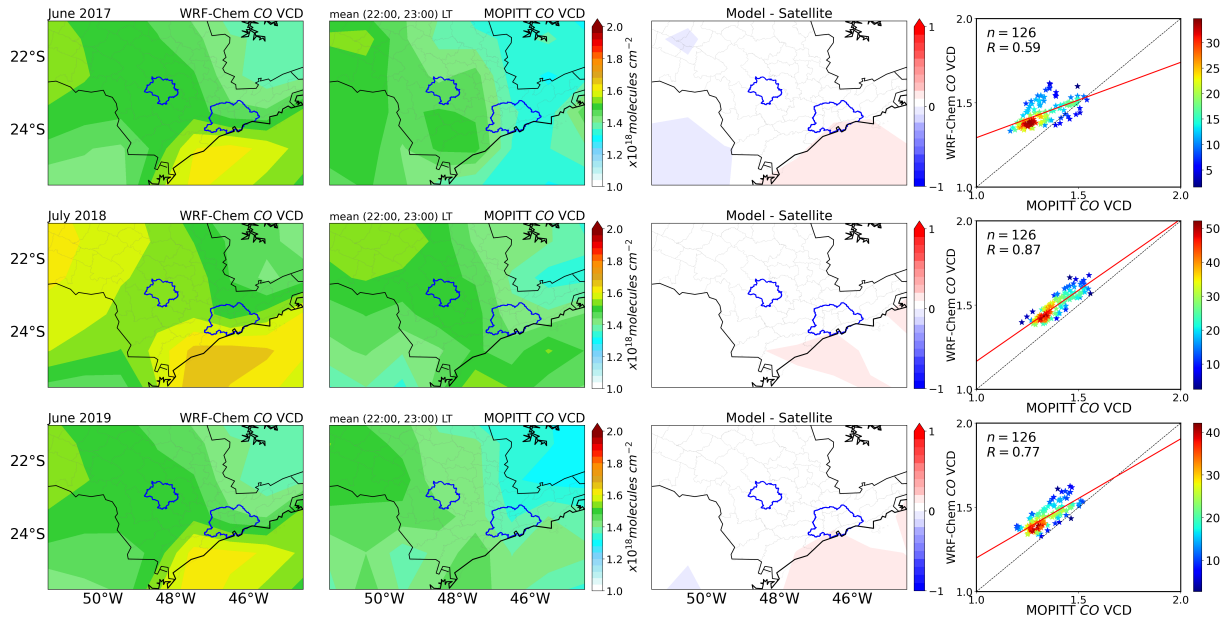

Figure S 12: Simulated (first column) and observed (second column) mean spatial distribution of total CO VCD during 22:00 LT, over southeastern Brazil for June 2017, July 2018 and June 2019. The third and fourth columns represent the relative difference and correlation between both tools. The blue lines, in the first three columns, represent Botucatu city and the MASP region. The gray diamond shape space indicates missing data.

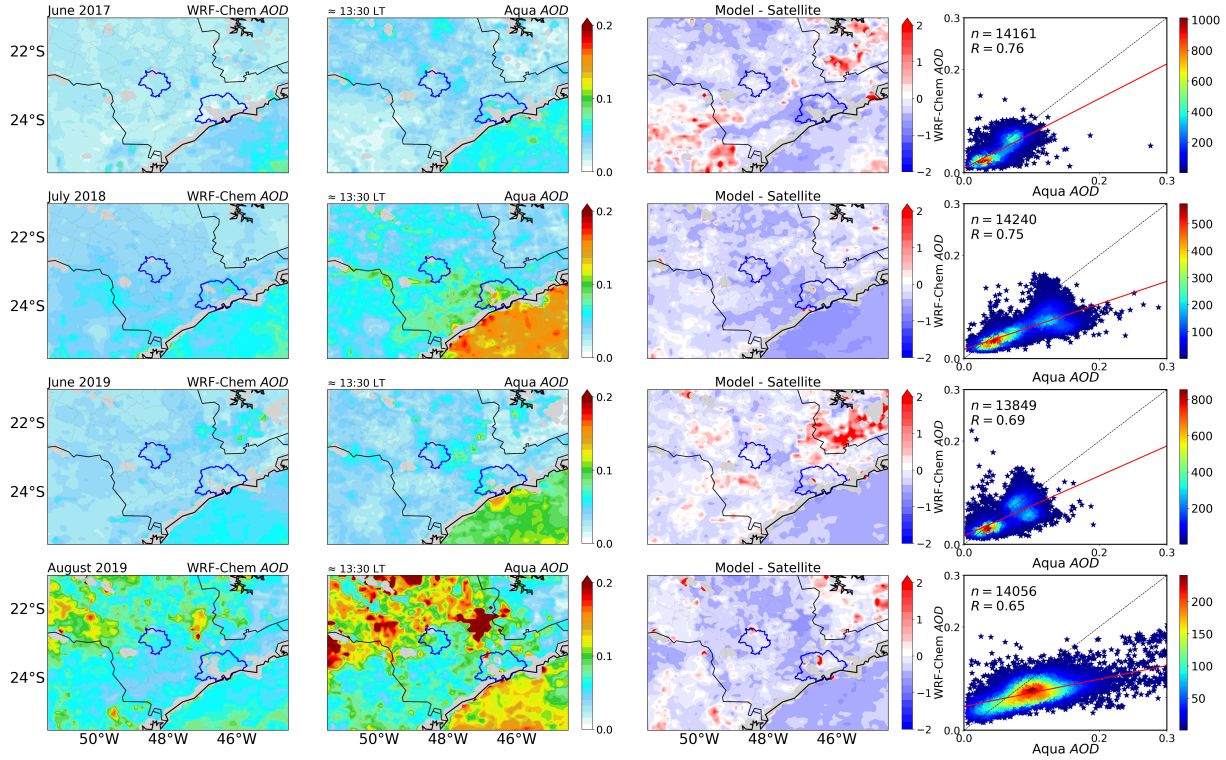

Figure S 13: Simulated (first column) and observed (second column) mean spatial distribution of AOD values at 13:30 LT over southeastern Brazil for June 2017, July 2018 and June and August 2019. The third and fourth columns represent the relative difference and correlation between both tools. The blue lines, in the first three columns, represent Botucatu city and the MASP and the gray space indicates missing data.

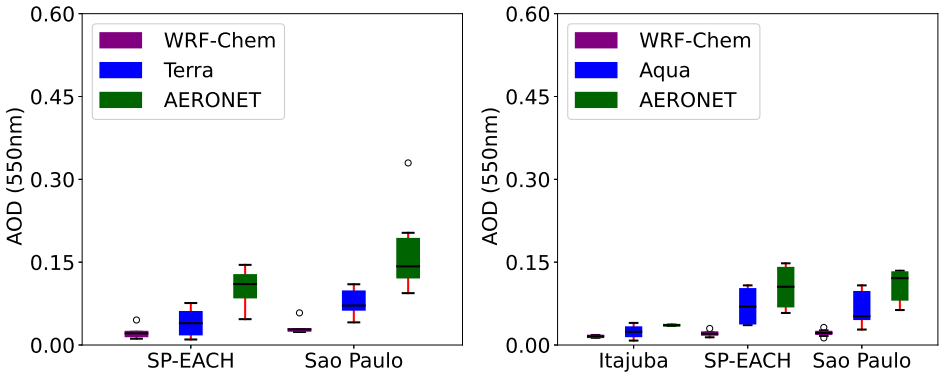

Figure S 14: Box plot of the WRF-Chem, MODIS, and AERONET AOD values at 550 nm over the AERONET stations for June 2017.

Table S 1: Air Quality Stations distributed over São Paulo state.

Stations	Type	Code	Latitude	Longitude	Meteorological Variables				Pollutants			
					T_{2m}	RH_{2m}	WS_{10m}	WD_{10m}	NO_x	O_3	CO	$PM_{2.5}$
Capão Redondo	Urban Park	269	-23.67	-46.78	x	x	x	x	x	x		
Cid.Universitária-USP-Ipen	Urban Park	95	-23.57	-46.74					x	x		x
Ibirapuera	Urban Park	83	-23.59	-46.66					x	x	x	x
Itaquera	Urban Park	97	-23.58	-46.47						x		
Carapicuíba	Urban	263	-23.53	-46.84	x	x	x	x	x	x	x	
Cerqueira César	Urban	91	-23.55	-46.67					x		x	
Congonhas	Urban	73	-23.62	-46.66					x		x	x
Diadema	Urban	92	-23.69	-46.61						x		
Grajaú-Parelheiros	Urban	98	-23.78	-46.70					x	x	x	x
Guarulhos-Paço Municipal	Urban	264	-23.46	-46.52	x	x	x	x	x	x		x
Guarulhos-Pimentas	Urban	279	-23.44	-46.41	x	x	x	x	x	x	x	x
Itaim Paulista	Urban	266	-23.50	-46.42						x		x
Interlagos	Urban	262	-23.68	-46.68	x	x	x	x	x	x		
Marg.Tietê-Pte Remédios	Urban	270	-23.52	-46.74	x	x	x	x	x		x	x
Mauá	Urban	65	-23.67	-46.47					x	x		
Mogi das Cruzes	Urban	287	-23.52	-46.19	x	x	x	x	x	x		
Mooca	Urban	85	-23.55	-46.60	x	x	x	x		x	x	
N.Senhora do Ó	Urban	96	-23.48	-46.69						x		
Osasco	Urban	120	-23.53	-46.79					x		x	
Parque D.Pedro II	Urban	72	-23.54	-46.63	x	x	x	x	x	x	x	x
Pinheiros	Urban	99	-23.56	-46.70	x	x	x	x	x	x	x	x
Santo Amaro	Urban	64	-23.65	-46.71						x	x	
São Caetano do Sul	Urban	86	-23.62	-46.56					x	x	x	
Santana	Urban	63	-23.51	-46.63	x	x	x	x		x		x
S.André Capuava	Urban	100	-23.64	-46.49	x	x	x	x		x		
S.André-Paço Municipal	Urban	254	-23.66	-46.53							x	
S.Bernardo-Centro	Urban	272	-23.70	-46.55	x	x	x	x	x	x	x	x
S.Bernardo-Paulicéia	Urban	102	-23.67	-46.58	x	x	x	x				
Taboão da Serra	Urban	103	-23.61	-46.76					x		x	
Araçatuba	Regional Urban	107	-21.19	-50.44	x	x	x	x		x		
Araraquara	Regional Urban	106	-21.78	-48.19	x	x	x	x	x	x		
Bauru	Regional Urban	108	-22.33	-49.09	x	x	x	x	x	x		
Campinas-Centro	Regional Urban	89	-22.90	-47.06							x	
Campinas-Taquaral	Regional Urban	276	-22.87	-47.06	x	x	x	x	x	x		
Campinas-V.União	Regional Urban	275	-22.95	-47.12	x	x	x	x	x	x		x
Catanduva	Regional Urban	248	-21.14	-48.98	x	x	x	x	x	x		
Guaratinguetá	Regional Urban	289	-22.80	-45.19	x	x	x	x	x	x		
Jaú	Regional Urban	110	-22.30	-48.57	x	x	x	x	x	x		
Jundiaí	Regional Urban	109	-23.19	-46.90	x	x	x	x	x	x		
Limeira	Regional Urban	281	-22.56	-47.41	x	x	x	x	x	x		
Marília	Regional Urban	111	-22.20	-49.96	x	x	x	x	x	x		
Paulínia Sul	Regional Urban	112	-22.79	-47.14					x	x		
Piracicaba	Regional Urban	113	-22.70	-47.65	x	x	x	x	x	x		x
Presidente Prudente	Regional Urban	114	-22.12	-51.41	x	x	x	x	x	x		
Ribeirão Preto	Regional Urban	288	-21.15	-47.83	x	x	x	x	x		x	x
São José Do Rio Preto	Regional Urban	116	-20.78	-49.40	x	x	x	x	x	x		x
Sorocaba	Regional Urban	67	-23.50	-47.48	x	x	x	x	x	x		
S.José Campos	Regional Urban	88	-23.19	-45.87	x	x	x	x	x	x		
S.José Campos-Jd.Satélite	Regional Urban	277	-23.22	-45.89	x	x	x	x	x	x	x	x
S.José Campos-Vista Verde	Regional Urban	278	-23.18	-45.83	x	x	x	x				
Taubaté	Regional Urban	280	-23.03	-45.58	x	x	x	x	x	x	x	x
Tatuí	Regional Urban	256	-23.36	-47.87	x	x	x	x	x	x		
Paulínia	Industrial	117	-22.77	-47.15	x	x	x	x	x	x		
Santa Gertrudes	Industrial	273	-22.46	-47.54	x	x	x	x				

The value "x" represents the existence of meteorological variables and chemical species available for each station.

Table S 2: Estimate of emission from sources of air pollution.

Species/ Contribution (%)	2017					2018					2019				
	CO	HC	NO_x	SO_x	PM	CO	HC	NO_x	SO_x	PM	CO	HC	NO_x	SO_x	PM
Vehicle Emission	96.76	75.30	63.92	16.70	40.	96.52	73.45	62.47	15.55	40.	96.43	72.85	64.90	11.47	40.
Industrial process	3.24	14.91	36.08	83.30	10.	3.48	16.02	37.53	84.45	10.	3.57	16.39	35.10	88.53	10.
Liquid fuel	-	9.80	-	-	-	-	10.53	-	-	-	-	10.77	-	-	-
Particle Resuspension	-	-	-	-	25.	-	-	-	-	25.	-	-	-	-	25.
Secondary Aerosols	-	-	-	-	25.	-	-	-	-	25.	-	-	-	-	25.
Total	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

All contributions are in %.

Table S 3: Meteorological parameter benchmarks (reference: Monk et al., 2019).

Variable	**Terrain Type**	**MB**	**ME**	**RMSE**	**IOA**
T_{2m}	**Simple Terrain**	$\leq \pm 0.5$	≤ 2		≥ 0.8
	Complex Terrain	$\leq \pm 1.0$	≤ 3		
RH_{2m}	**Simple Terrain**	$\leq \pm 10$	≤ 20		≥ 0.6
WS_{10m}	**Simple Terrain**	$\leq \pm 0.5$		≤ 2	≥ 0.6
	Complex Terrain	$\leq \pm 1.5$		≤ 2.5	
WD_{10m}	**Simple Terrain**	$\leq \pm 10$	≤ 30		
	Complex Terrain	$\leq \pm 10$	≤ 55		

The units are T_{2m} in $[^{\circ}C]$, RH_{2m} in [%], WS_{10} in $[ms^{-1}]$ and WD_{10} in [degree].

Table S 4: Benchmarks for photo-chemical model performance statistics (reference: Emery et al., 2017).

	NMB		**NME**		**r**	
Species	**Goal**	**criteria**	**Goal**	**criteria**	**Goal**	**criteria**
1-hr O_3	$\leq \pm 5\%$	$\leq \pm 15\%$	$\leq \pm 15\%$	$\leq \pm 25\%$	≤ 0.75	≤ 0.50
24-hr $PM_{2.5}$	$\leq \pm 10\%$	$\leq \pm 30\%$	$\leq \pm 35\%$	$\leq \pm 50\%$	≤ 0.70	≤ 0.40

The units are O_3 and $PM_{2.5}$ in $[\mu gm^{-3}]$.

Table S 5: Performance statistics of the validation between predicted meteorological parameters and measurements by the AQS.

Type	**Meteorological**	**June - 2017**								**July - 2018**							
of Region	**variable**	**n**	R	MB	NMB	ME	NME	RMSE	IOA	**n**	R	MB	NMB	ME	NME	RMSE	IOA
	T_{2m}	697	0.95	-0.20	-1.07	1.19	6.25	1.49	0.97	721	0.97	-0.10	-0.48	1.06	5.19	1.32	0.98
Regional	RH_{2m}	697	0.89	-2.21	-2.98	5.98	8.08	7.83	0.94	721	0.90	-2.03	-3.48	6.15	10.54	7.63	0.95
Urban	WS_{10m}	697	0.75	0.83	44.57	0.86	46.46	0.97	0.61	721	0.70	1.06	63.58	1.06	63.69	1.16	0.53
	WD_{10m}	697	-	-8.97	-	38.01	-	-	-	721	-	-20.13	-	46.41	-	-	-
	T_{2m}	692	0.91	0.08	0.47	1.56	9.12	2.09	0.94	721	0.94	0.73	4.15	1.47	8.32	1.82	0.96
Urban	RH_{2m}	692	0.84	-4.77	-5.85	8.75	10.73	12.13	0.90	721	0.84	-6.66	-9.16	10.84	14.90	14.07	0.90
Park	WS_{10m}	531	0.67	1.40	86.77	1.61	99.87	2.06	0.51	514	0.67	0.94	59.07	1.17	73.85	1.52	0.60
	WD_{10m}	529	-	5.30	-	35.58	-	-	-	514	-	4.28	-	38.67	-	-	-
	T_{2m}	697	0.91	-0.33	-1.88	1.53	8.70	2.08	0.93	721	0.95	-0.16	-0.85	1.21	6.50	1.52	0.97
Urban	RH_{2m}	697	0.83	-1.21	-1.55	8.35	10.72	10.90	0.90	721	0.83	-1.17	-1.76	9.72	14.62	12.02	0.91
	WS_{10m}	697	0.76	1.06	63.95	1.13	68.20	1.50	0.58	721	0.72	0.86	55.09	0.96	61.84	1.26	0.59
	WD_{10m}	697	-	-2.49	-	51.59	-	-	-	721	-	-3.05	-	55.86	-	-	-
	T_{2m}	697	0.91	-0.74	-3.81	1.75	9.04	2.14	0.94	721	0.96	-0.81	-4.0	1.55	7.61	1.84	0.97
Industry	RH_{2m}	697	0.85	-0.59	-0.82	7.81	10.78	10.38	0.92	721	0.91	3.68	6.68	7.53	13.69	9.49	0.94
	WS_{10m}	659	0.52	0.83	57.27	0.94	65.27	1.13	0.56	669	0.50	1.03	74.96	1.12	81.13	1.32	0.52
	WD_{10m}	658	-	-10.31	-	56.84	-	-	-	669	-	-8.32	-	58.23	-	-	-
		June - 2019								**August - 2019**							
	T_{2m}	649	0.96	-0.44	-2.17	1.06	5.19	1.28	0.98	718	0.96	0.35	1.73	1.12	5.53	1.50	0.98
Regional	RH_{2m}	649	0.91	0.48	0.71	5.56	8.12	7.0	0.95	718	0.92	-4.39	-6.93	6.78	10.69	8.53	0.94
Urban	WS_{10m}	649	0.79	1.02	59.23	1.03	59.67	1.13	0.60	718	0.78	1.08	46.19	1.09	46.78	1.22	0.64
	WD_{10m}	649	-	-21.28	-	41.85	-	-	-	718	-	-16.44	-	30.35	-	-	-
	T_{2m}	629	0.94	0.26	1.45	1.04	5.67	1.35	0.97	713	0.95	0.98	5.87	1.40	8.37	1.79	0.97
Urban	RH_{2m}	629	0.86	-4.85	-6.15	8.24	10.44	11.04	0.91	713	0.89	-7.36	-9.25	9.82	12.34	12.55	0.91
Park	WS_{10m}	483	0.72	0.87	46.67	1.14	61.19	1.48	0.73	631	0.79	0.81	30.78	1.06	39.89	1.34	0.80
	WD_{10m}	481	-	5.62	-	37.47	-	-	-	631	-	1.69	-	25.85	-	-	-
	T_{2m}	649	0.94	-0.52	-2.75	1.13	5.94	1.37	0.96	718	0.95	0.11	0.60	1.20	6.80	1.52	0.98
Urban	RH_{2m}	649	0.85	2.05	2.84	7.89	10.92	9.81	0.92	718	0.87	0.32	0.44	8.60	12.04	10.64	0.93
	WS_{10m}	649	0.77	1.03	64.73	1.10	69.12	1.42	0.61	718	0.80	1.24	60.62	1.28	62.16	1.58	0.58
	WD_{10m}	649	-	-4.86	-	54.96	-	-	-	718	-	-10.53	-	34.39	-	-	-
	T_{2m}	649	0.94	-0.90	-4.46	1.49	7.35	1.84	0.96	718	0.95	0.23	1.13	1.35	6.69	1.78	0.97
Industry	RH_{2m}	648	0.88	5.48	8.26	9.08	13.68	11.12	0.91	718	0.90	-2.44	-3.93	8.10	13.02	10.14	0.94
	WS_{10m}	620	0.62	1.04	75.28	1.13	81.53	1.34	0.56	711	0.68	1.01	54.61	1.11	59.95	1.33	0.65
	WD_{10m}	619	-	-16.59	-	59.55	-	-	-	710	-	-11.14	-	46.06	-	-	-

The units are T_{2m} in $[°C]$, RH_{2m} in $[ms^{-1}]$, WS_{10} in $[ms^{-1}]$ and WD_{10} in [degree].

Table S 6: Performance statistics of the validation between predicted meteorological parameters and measurements by IAG-USP station.

Meteorological	**June - 2017**							
Variable	**n**	R	MB	NMB	ME	NME	RMSE	IOA
T_{2m}	696	0.84	0.41	2.43	2.04	12.13	2.61	0.90
RH_{2m}	696	0.77	-4.13	-5.05	9.75	11.94	13.59	0.86
WS_{10m}	696	0.66	-0.01	-1.06	0.63	45.93	0.81	0.81
				July - 2018				
T_{2m}	720	0.84	1.46	8.58	2.43	14.32	3.03	0.89
RH_{2m}	720	0.79	-8.95	-11.90	12.42	16.51	16.53	0.85
WS_{10m}	720	0.58	0.11	9.87	0.59	53.20	0.74	0.76
				June - 2019				
T_{2m}	648	0.75	0.99	5.60	2.22	12.62	2.84	0.84
RH_{2m}	648	0.57	-5.46	-6.83	13.46	16.84	17.76	0.74
WS_{10m}	648	0.38	0.19	15.95	0.81	69.36	1.02	0.62
				August - 2019				
T_{2m}	718	0.88	1.11	6.78	1.96	11.92	2.58	0.93
RH_{2m}	718	0.81	-4.47	-5.77	18.93	24.47	23.73	0.61
WS_{10m}	718	0.53	0.04	2.18	0.70	42.47	0.88	0.73

The units are T_{2m} in $[^{\circ}C]$, RH_{2m} in $[ms^{-1}]$ and WS_{10} in $[ms^{-1}]$.

2 Building of the anthropogenic emissions (vehicular) and the emissions inventory

2.1 Vehicular emissions

As mentioned in section 2.2.2, vehicular emissions (or ground transport emissions) were considered as the main anthropogenic source to run the WRF-Chem model over southeastern Brazil. The ground transport emissions, specifically on-road vehicles, were derived from the bottom-up transport emission model described by Andrade et al., (2015). This model uses information on number of vehicles, emission factors and intensity of use (mileage) for different vehicle types (motorcycles, light-duty vehicles, and heavy-duty vehicles) and different fuel types (gasohol, ethanol, combination of any proportion of gasohol and ethanol, and diesel) from different agencies, such as DENATRAN, CETESB and the IAG-USP Laboratory of Atmospheric Processes (LAPAt) (see Figure S15). Volatile Organic Compounds (VOCs) speciation and pollutant temporal profiles were derived from in-tunnel field campaigns located in the MASP during 2004 and 2011 (Martins et al., 2006; Nogueira et al.,2021).

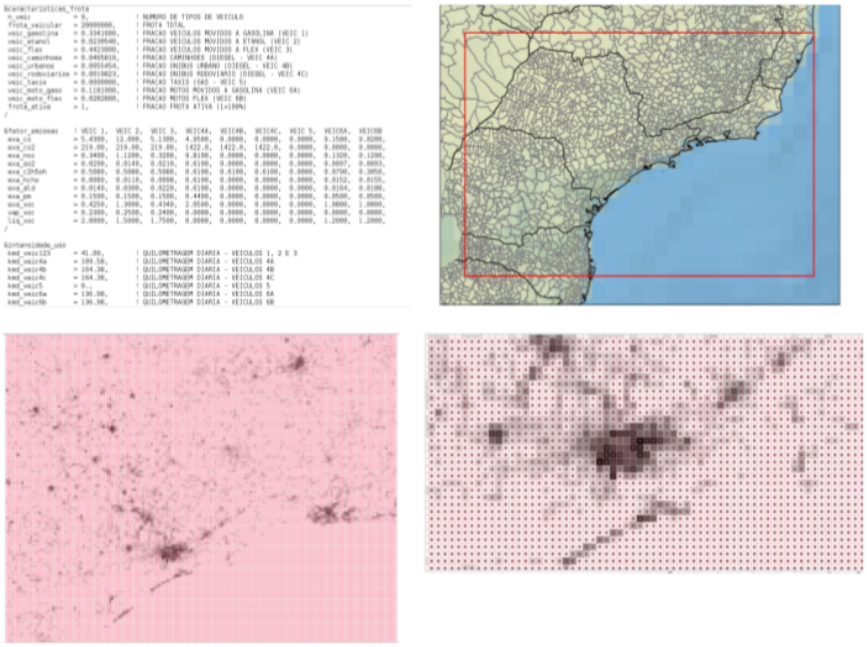

Figure S 15: (a) Information extracted from DENATRAN, CETESB and the IAG-USP Laboratory of Atmospheric Processes (LAPAt), (b) number of municipalities considered in the vehicle emissions inventory, (c) road length (in km) for each 9x9 km grid cell (d) zoom of (c).

For the temporal distribution, the profile for light and heavy vehicles was considered, generating emissions throughout the day. Figure S16 shows the hourly profile of the atmospheric pollutants emissions emitted by light vehicles (CO and Aldehyde) and heavy vehicles (NO and NO_2), considering four different regions (regional urban, urban park, urban and industrial) within the state of São Paulo (these regions are explained in section 2.4). The maximum emissions of light vehicles take place at 10:00 UTC and 19:00 UTC which coincide with the morning and evening rush hours in the MASP. The emissions of heavy vehicles, on the other hand, remain constant over the period between the rush hours.

For the spatial distribution (Figure S17), the number of vehicles in a given grid cell is assumed to be proportional to the road length in the grid cell. Road length in each grid cell was calculated as the sum of various road types found

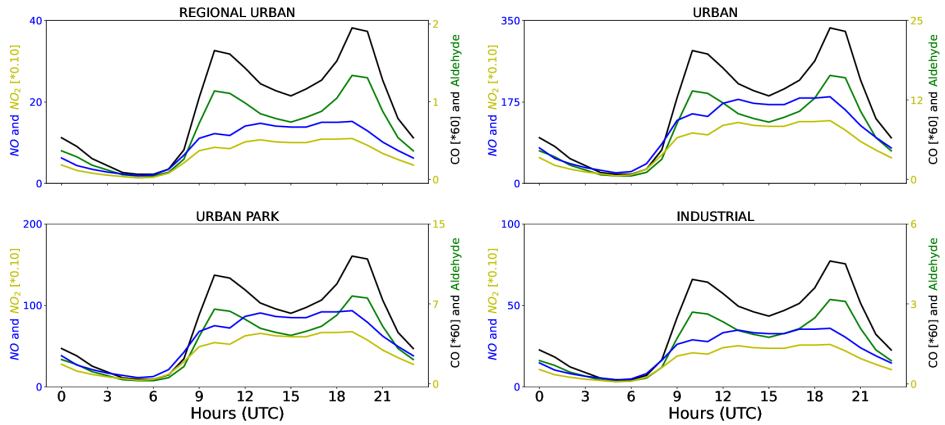

Figure S 16: Temporal distribution of NO, NO_2, CO and Aldehyde emission in $mol\ km^{-2}hr^{-1}$, representing light and heavy vehicular traffic.

within each grid cell, based on OpenStreetMap road length estimates (https://www.openstreetmap.org/). For this study, the total vehicle emissions to distribute over the modeling domain were obtained from the CETESB report for the year 2015 (CETESB, 2015). Figure S17 illustrates the spatial distribution of NO and $PM_{2.5}$ emissions at 20:00 UTC over southeastern Brazil. High emissions of both pollutants can be observed over central areas of the MASP, while low emissions are presented out of it. This contrast of emissions reflects the high flow of vehicles that circulate through MASP (Vara-Vela et al., 2016; Gavidia-Calderón et al., 2018).

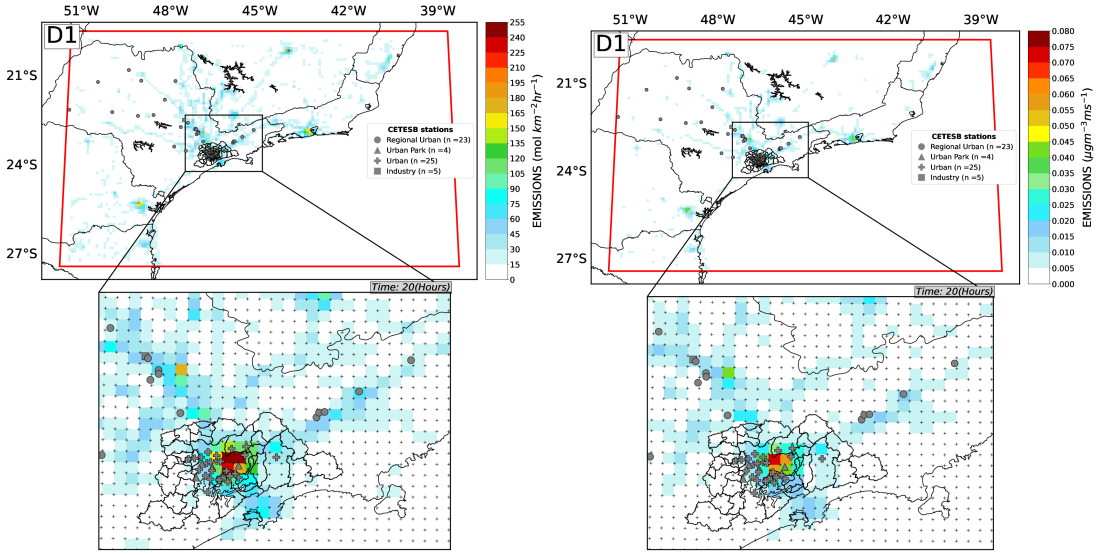

Figure S 17: Spatial distribution of (left side) NO and (right side) $PM_{2.5}$ emission in $mol\ km^{-2}hr^{-1}$ and $\mu gm^{-3}\ ms^{-1}$, respectively, representing the hourly vehicular traffic.

2.2 Biogenic and Fire emissions

Biogenic emission data, calculated by the MEGAN model, estimate the net monthly or annual emission of 20 compound classes generated by urban, rural and agricultural ecosystems, except by biomass burning, and that can influence the atmosphere. This component, breaks down into 150 individual species such as isoprene, monoterpenes, sesquiterpenes, carbon monoxide, nitrogen monoxide, alkanes, alkenes, aldehydes, acids, ketones, and other oxygenated VOCs. MEGAN has a spatial resolution of approximately 1km and it can be used for both regional air quality modeling and global earth system modeling studies (Guenther et al., 2006; Arneth et al., 2011). MEGAN

provides some suitable variables for the WRF-Chem model such as: amount of isoprene (MSEBIO_ISOP), percentage of broad leaf (PFTP_BT), percentage of needle leaf (PFTP_NT), percentage of shrubs (PFTP_SB), percentage of herbaceous biota (PFTP_HB), monthly leaf area index (MLAI), monthly air temperature (MTSA, in K) and monthly downward short wave radiation (MSWDOWN, Wm^{-2}). The upper left panel in Figure S18 represents the spatial distribution of the annual mean emissions of ISOP over southeastern Brazil, with focus on the MASP. The other panels represent the spatial distribution of MLAI for the months of June, July and August. We can seen high ISOP and LAI values on the MASP boundaries and over the state of Parana and Santa Catarina for all periods, although for August the LAI values were higher compared to the other months. The LAI is an important indicator of radiation and precipitation interception, energy conversion, and water balance. It is a reliable parameter for plant growth (Jang et al., 2019).

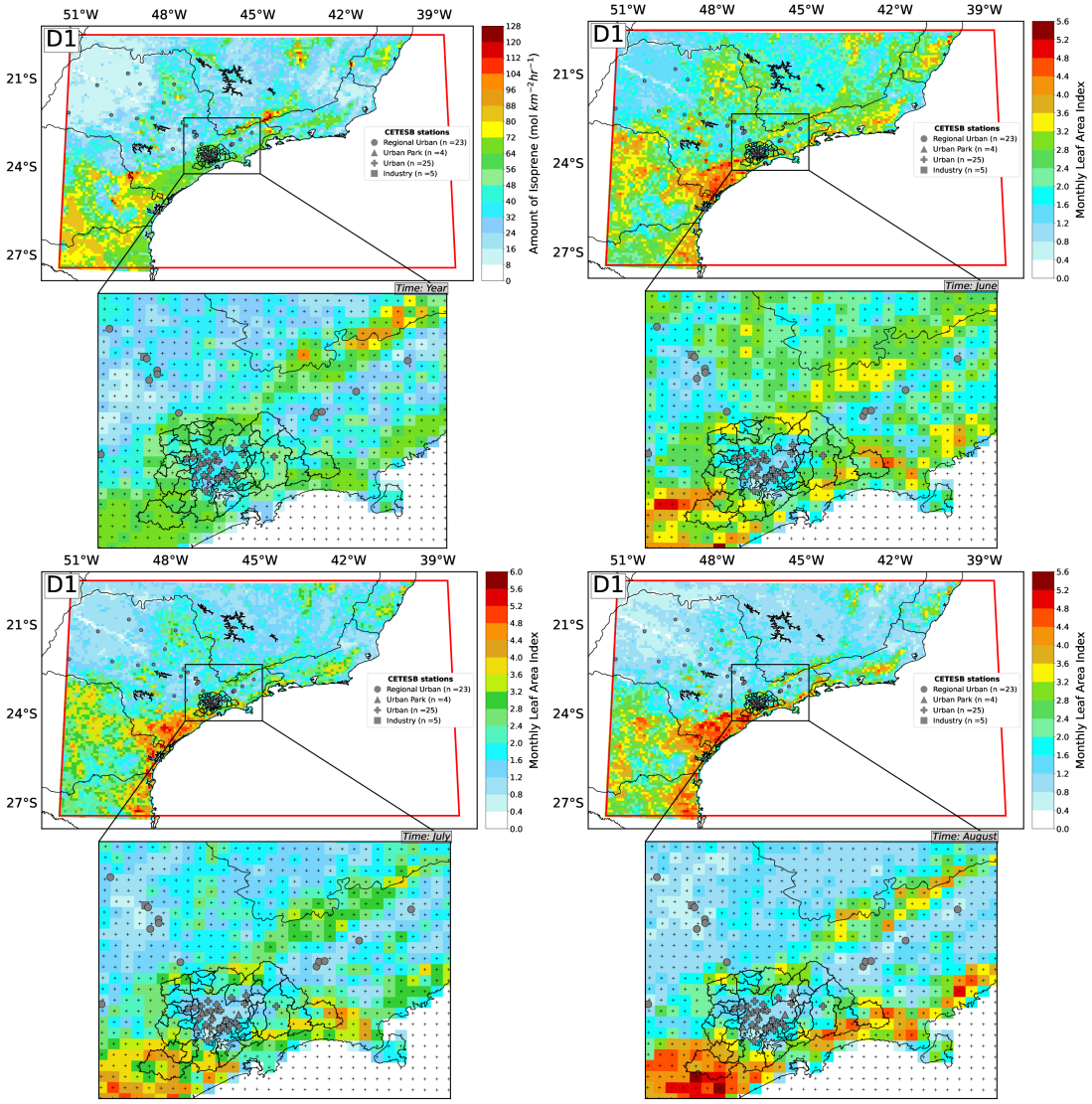

Figure S 18: Spatial distribution of annual mean emissions of isoprene (Top-left), and monthly leaf area index for (Top-right) June, (Bottom-left) July and (Top-left) August derived from the MEGAN model and over the modeling domain.

Biomass burning is a significant air pollution source, with global, regional and local impacts on air quality, public health and climate. The State of São Paulo often experiences high air pollution levels due to emissions of particulate pollutants from local sources and long-range transport of air masses originated from biomass burning. For this work,

we have considered the Fire INventory from NCAR version 1.0 (FINNv1) model to obtain fire emission data over our study domain. This inventory was chosen because it differs from other inventories by providing a unique combination of high temporal and spatial resolution, global coverage, and estimates for a large number of chemical species. FINN model used satellite observations processed by MODIS sensor onboard the Terra and Aqua satellites, to detect active fires and land cover. These data are needed to estimate a high spatial resolution (1km) daily data of global emissions of particulates and trace gases from different types of vegetation fires (wildfires, agricultural burns, and prescribed burns). On the other hand, this inventory does not include biofuels use and burning (Wiedinmyer et al., 2011). The daily data estimated by FINNv1 have been specifically developed to model atmospheric chemistry and air quality in a consistent framework at local to global scales (Freitas et al., 2007; Grell et al., 2011). Emission factors have been updated, particularly, for Non-methane Organic Compounds (NMOC). The list of chemical speciation needed to assign total NMOC emission estimates to grouped species for use in chemical transport models is provided for three widely used chemical mechanisms: SAPRC99, GEOS-CHEM, and MOZART-4. Also, FINN provides estimates of organic compounds, including formaldehyde and methanol (Wiedinmyer et al., 2011).

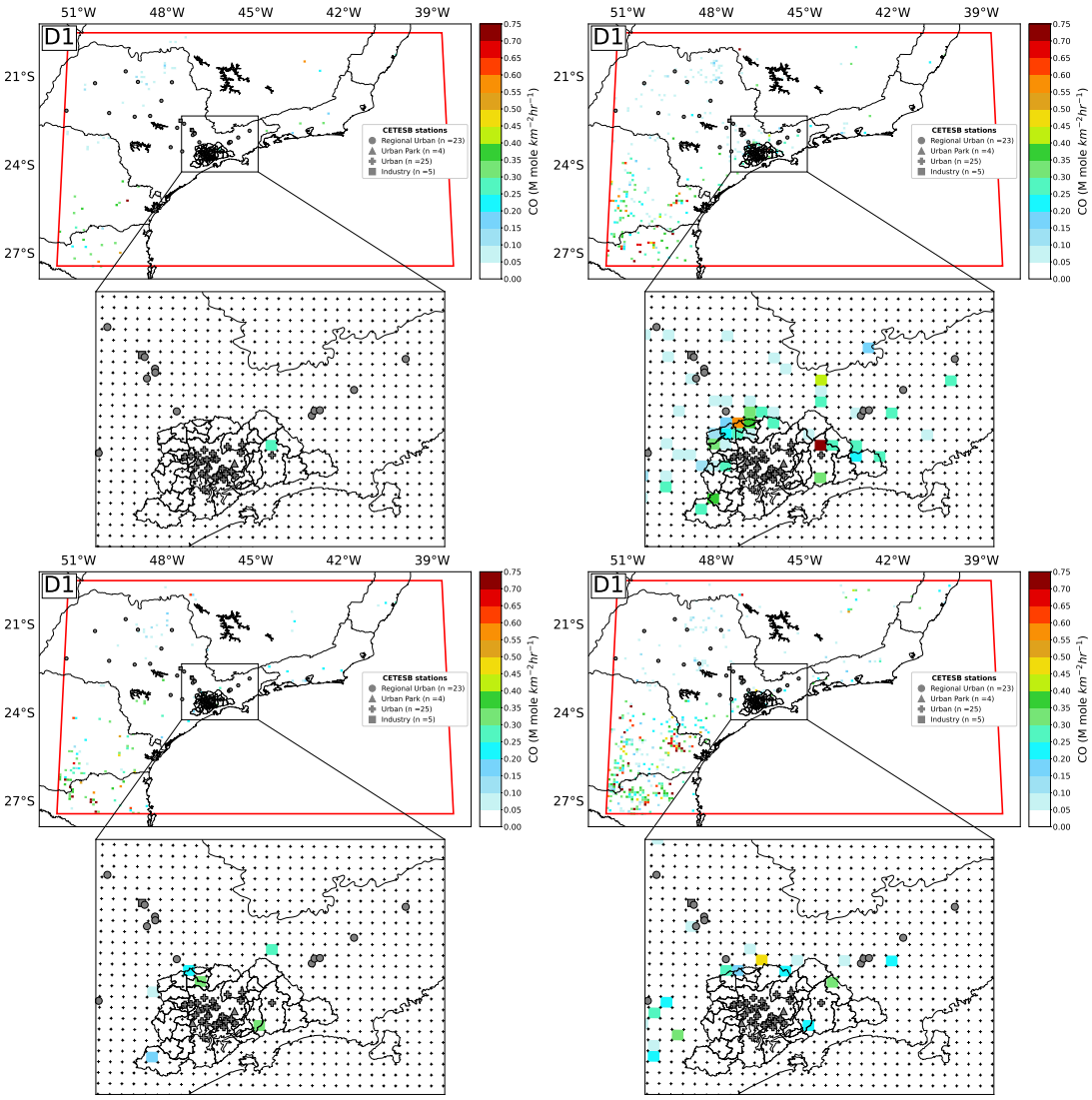

Figure S 19: Spatial distribution of CO emissions from cumulative burning events, estimated by the FINNv1 model, throughout (Top-left) June 2017, (Top-right) July 2018, (Bottom-left) June 2019 and (Bottom-right) August 2019. All of the spatial distributions maps are also showing a zoomed area over the study domain region.

Figures S19 to S21 represent the spatial distribution of CO, $PM_{2.5}$ and NO_2 emissions from monthly cumulative burning events over our study domain and for each period analyzed (June 2017, July 2018, June and August 2019). We can see that during July 2018, there was a higher pollution emission, compared to the other three periods, this will be in the results section.

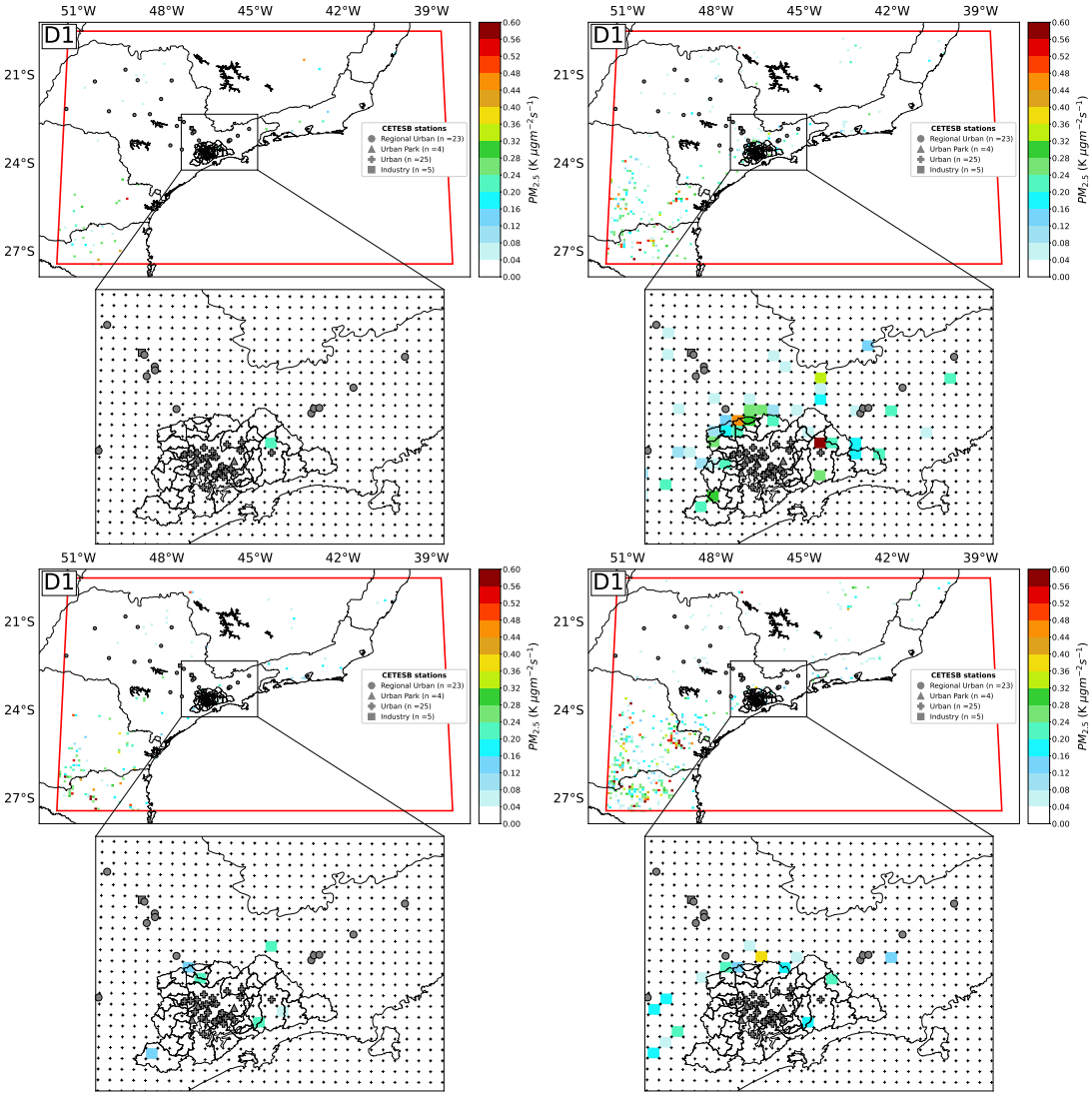

Figure S 20: Spatial distribution of $PM_{2.5}$ emissions from cumulative burning events, estimated by the FINNv1 model, throughout (Top-left) June 2017, (Top-right) July 2018, (Bottom-left) June 2019 and (Bottom-right) August 2019. All of the spatial distributions maps are also showing a zoomed area over the study domain region.

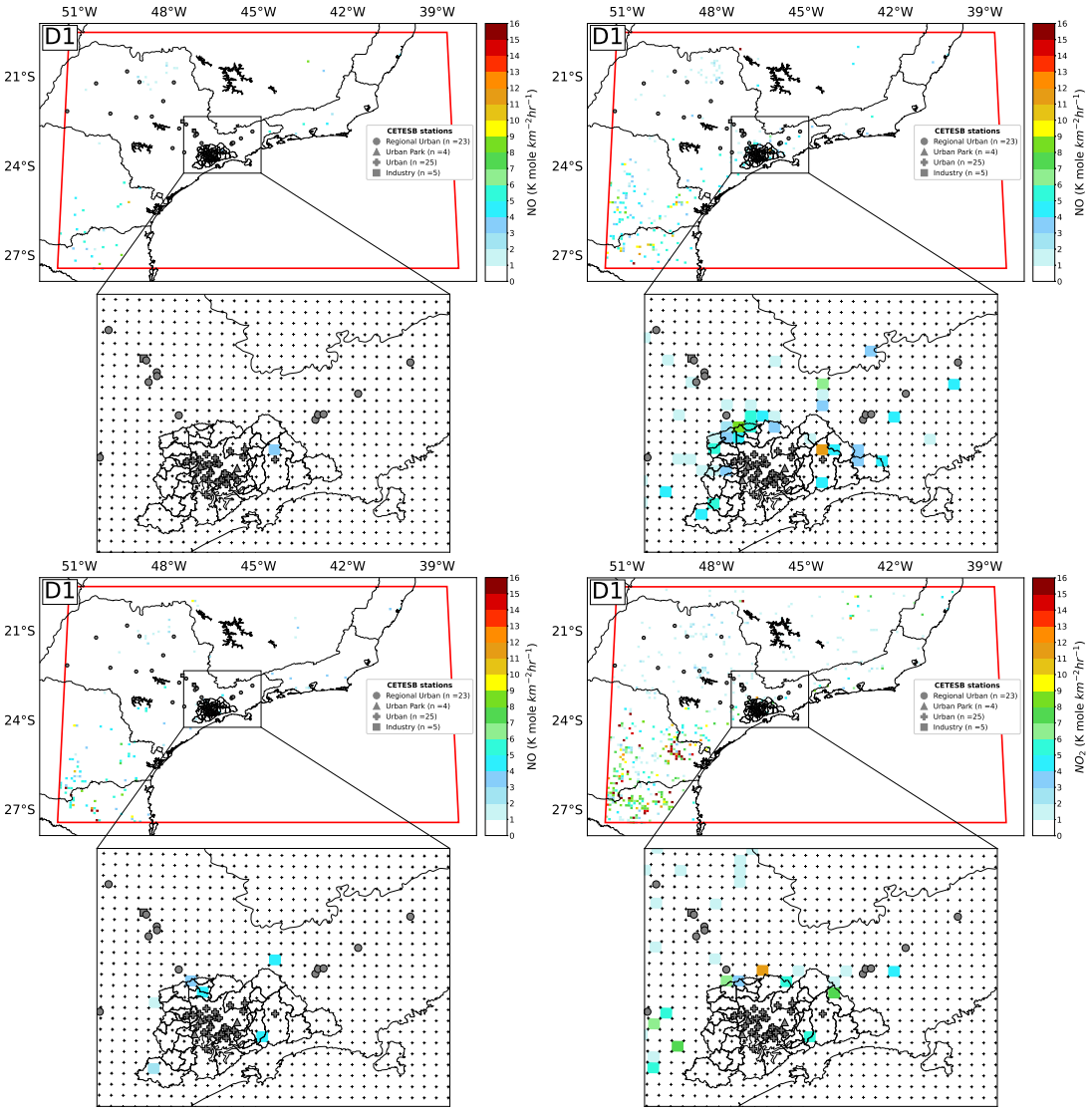

Figure S 21: Spatial distribution of NO_2 emissions from cumulative burning events, estimated by the FINNv1 model, throughout (Top-left) June 2017, (Top-right) July 2018, (Bottom-left) June 2019 and (Bottom-right) August 2019. All of the spatial distributions maps are also showing a zoomed area over the study domain region.

3 Comparison between MODIS products and AERONET stations

In order to find periods with availability of AOD data, a broad 6-year analysis (2014-2019) of both MODIS and AERONET products has been carried out over the southeastern Brazil region. The analysis was performed using a matching methodology, by considering spatio-temporal windows for coincident observations between satellite and ground-truth (Ichoku et al., 2002; Martins et al., 2017). Fig. S22 show the performance of MODIS (for Terra and Aqua satellites, respectively) sensor, in terms of NMB and NME for the period of 6 years. Each panel represents the MODIS performance for a specific spatio-temporal window. The top-left panel in Fig. S22 (top) represents the performance statistics for coincident observations (MODIS, AERONET) that has a distance of less than or equal 1 km and a time span of up to ± 15 minutes (the time refers to satellite passage over the AERONET station). The temporal window changes for each column (t ≤ ± 15, 30 and 60 minutes) and the spatial window changes for each row (d ≤ 1, 3 and 10 km). This same analysis is also considered for the Aqua satellite. The scatterplots from Figure S22 show a difference between the spatial window of 1 x 1 km^2 and the spatial windows of 3 x 3 km^2 and 10 x 10 km^2, while between the latter two the behavior is similar. Regarding the temporal windows, we can be seen that the MODIS performance for these three intervals (± 15, 30, 60 minutes) is almost indistinguishable, using a spatial window of 10 x 10 km^2. Therefore, the most appropriate matchup period for Terra and Aqua satellites is using an interval of 60 minutes with a spatial window of 10 x 10 km^2. The most appropriate matchup is the one with NMB and NME closest to zero. With this spatio-temporal window, a greater number of valid pairs MODIS-AERONET was found, for the y-axis, close and within the second rectangle of each panel (NME < 50 %) and close to zero (NMB ≈ 0) for the x-axis. However, DB-L2 and DT-3K products do not represent well the AOD values for the SP-EACH station (for Terra and Aqua satellites, respectively). Considering the four seasons of the year, Summer (DJF: December, January and February), Autumn (MAM: March, April and May), Winter (JJA: June, July and August) and Spring (SON: September, October and November), and using the matching methodology, we have found different temporal and spatial windows matchups compared to the annual analysis (not shown). Nevertheless, the deficiency of the DB and MAIAC products were also notorious, while the DT-3K and DT-L2 products have performed better in the winter season over most of the AERONET stations over southeast Brazil. Taking into account that the spatio-temporal windows with ± 15, 30 and ± 60 minutes, and 10 km showed similar performance, in this study, the spatio-temporal window of ± 60 minutes and 10 km was chosen for both satellites, considering the largest number of matchups (MODIS - AERONET).

A more detailed analysis to validate MODIS AOD data against AERONET estimates has been carried out, considering the appropriate spatio-temporal windows (d ≤ 10 km and t ≤ ± 60 minutes). Fig. S23 show the temporal variability (left side) and scatter plot (right side) of the AOD values from AERONET at the São Paulo station and retrieved by the MODIS products (MAIAC, DT-3K, DB-L2, DT-DB-L2 and DT-L2) from 2014 to 2019, for the Terra and Aqua satellites. The choice to show the assessment at the São Paulo station was based on the larger amount of data availability compared to other stations. The largest number of matchup average AOD values was found particularly from August to October for each year when forest fires represent a dominant source of aerosols over the central Amazon (Hoelzemann et al., 2009) and can be transported to urban areas in southeastern Brazil, such as the MASP, causing environmental contamination over those regions (Martins et al., 2018; Vara-Vela et al., 2021). Analyzing the Fig. S23, we can observe that the DT (3K/L2) product represents well the AOD values (Terra: R = 0.80/0.78 and Aqua: R = 0.90/0.91) over the study region, while DB product had more deficiency (Terra: R = 0.45 and Aqua: R = 0.78). In relation to MAIAC product, it is not as efficient as the DT product (Terra: R = 0.71 and Aqua: R = 0.77) in representing the AOD values on the São Paulo station, despite having the higher spatial resolution. In addition,

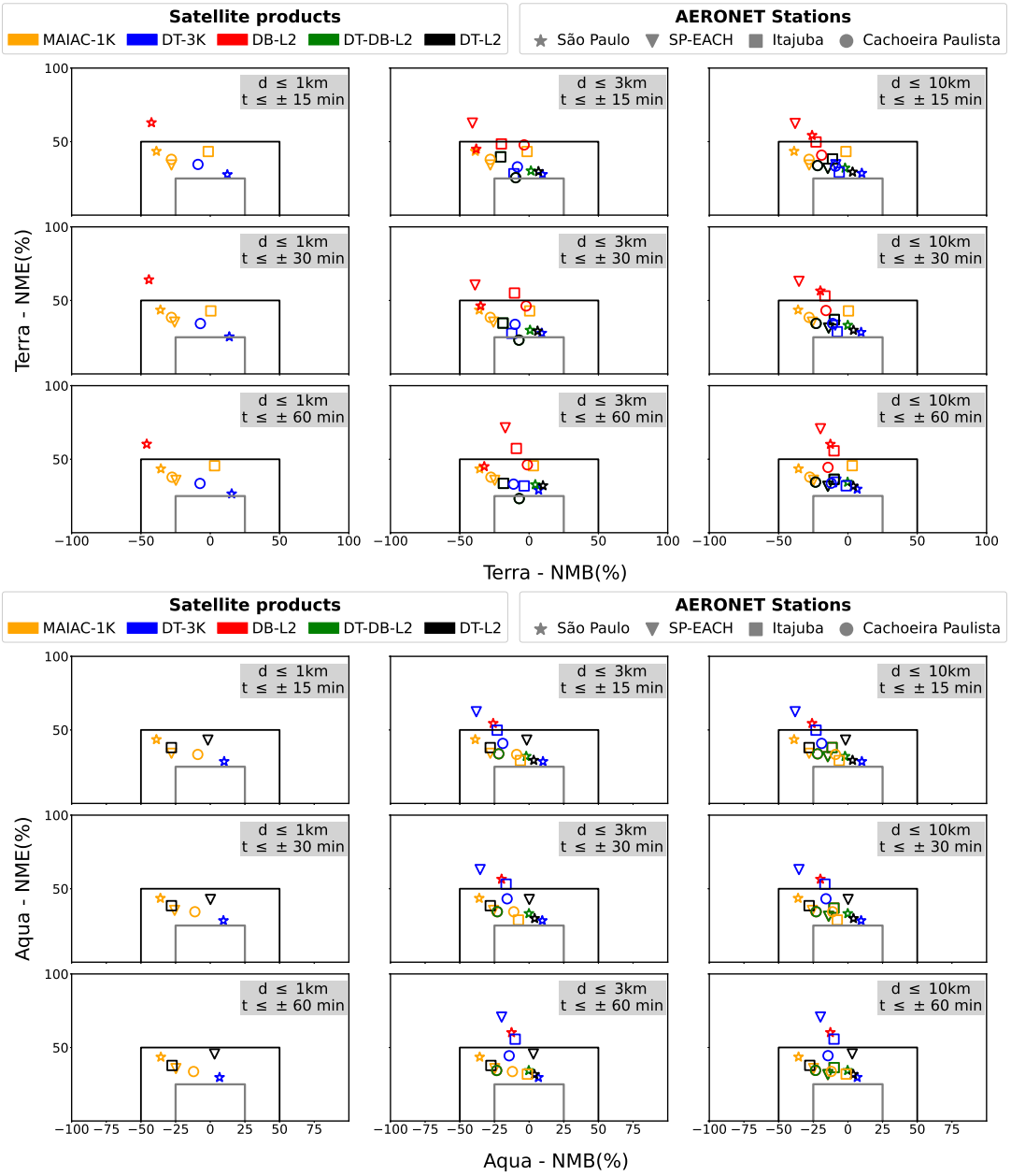

Figure S 22: Scatter plots of NME versus NMB, representing the performance of the MODIS products (dots with different colors) for Terra and Aqua satellite at four AERONET stations (different markers) for the six years of study. Each panel represents the MODIS performance at a specific time window: t $\leq \pm$ 15, 30 and 60 minutes, and a specific spatial window: d $\leq$ 1, 3 and 10 km.

MODIS Aerosol product reaches a good accuracy when more than 66 % of retrievals fall within the expected error (EE) envelope (Remer et al., 2005). Using this condition, 79.7 % (Terra), 89.3 % (Aqua) and 78.2 % (Terra), 89.3 % (Aqua) of retrievals from DT-3K and DT-L2, respectively, fall within the EE envelope, indicating a good accuracy at the São Paulo station. Instead, for the MAIAC and DB products, only 60.8 % (Terra), 71.6 % (Aqua) and 53.9 % (Terra), 57.4 % (Aqua) of retrievals fall within the EE envelope, indicating a bad accuracy for the MAIAC and DB products. A possible justification for MAIAC result is related to the aerosol model used by the MAIAC algorithm which fails over urban areas with land use characteristics different to those found in United States East Coast cities. The MAIAC aerosol model considered over the study area is the same for the United States East Coast, dominated by non-absorbing aerosols, while in the MASP, a more absorbing aerosol model prevails, with lower single scattering albedo values (Damascena et al., 2021). The statistics performances for the other three AERONET stations are

shown in the Table S7. More details are presented in Benavente (2022). This analysis supports the fact that the DT (3K and L2) product is the most suitable MODIS AOD product for studying aerosols over our study area. We have chosen the DT-L2 product because in terms of regriding, it generates low numerical instability and it has a better statistical performance.

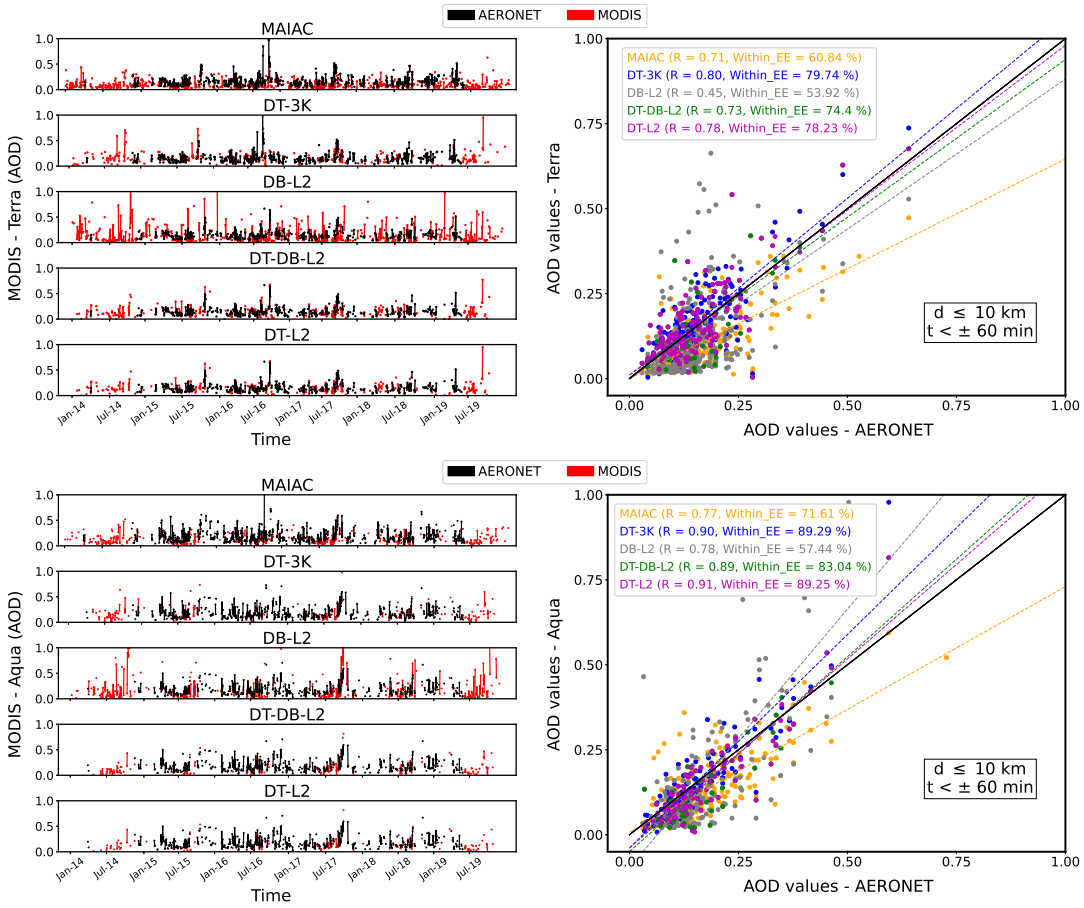

Figure S 23: (Left side) Temporal variability (2014-2019) and (Right side) the correlation of the AOD values at 550nm obtained by the AERONET station (black line) and Aqua satellite (red line), using MAIAC (yellow dots), DT-3K (blue dots), DB-L2 (gray dots), DT-DB-L2 (green dots) and DT-L2 (magenta dots) algorithms at the São Paulo station.

Table S 7: General statistics of the validation between the AOD values measured by the AERONET stations and MODIS products, considering the most appropriate spatio-temporal window (t $\leq \pm$ 60 minutes and d $\leq$ 10 km) from 2014 to 2019 (n = Matchup Number)

Station	**Algorithm/**	**Terra**								**Aqua**							
	Resolution	n	R	MB	NMB	ME	NME	RMSE	EE (%)	n	R	MB	NMB	ME	NME	RMSE	EE (%)
	MAIAC-1km	332	0.71	-0.05	-35.64	0.06	43.52	0.08	60.84	236	0.77	-0.04	-23.35	0.06	35.34	0.07	71.61
	DT-3K	153	0.80	0.01	6.70	0.05	29.68	0.07	79.74	112	0.90	0.0	2.27	0.04	24.47	0.06	89.29
Sao Paulo	DT-L2	147	0.78	0.0	3.90	0.05	31.66	0.07	78.23	93	0.91	-0.02	-11.26	0.04	23.56	0.06	89.25
	DB-L2	293	0.45	-0.02	-12.66	0.09	60.18	0.13	53.92	195	0.78	-0.0	-3.06	0.09	52.14	0.14	57.44
	DT-DB-L2	168	0.73	-0.0	-0.23	0.05	34.38	0.07	74.40	112	0.89	-0.02	-14.22	0.05	28.71	0.07	83.04
	MAIAC-1km	88	0.70	-0.04	-24.78	0.05	35.68	0.07	68.18	62	0.93	-0.03	17.58	0.04	22.94	0.05	88.71
	DT-3K	15	0.49	-0.01	-8.60	0.05	34.47	0.07	80.0	3	-	-	-	-	-	-	-
SP-EACH	DT-L2	33	0.85	-0.03	-14.48	0.06	31.51	0.07	72.73	17	0.94	0.03	18.82	0.04	26.72	0.06	76.47
	DB-L2	98	0.48	-0.03	-19.80	0.12	70.68	0.16	35.71	56	0.60	-0.0	-2.87	0.11	61.16	0.16	42.86
	DT-DB-L2	33	0.85	-0.03	-14.48	0.06	31.51	0.07	72.73	19	0.92	0.03	20.53	0.05	30.72	0.06	73.68
	MAIAC-1km	201	0.74	0.0	3.11	0.04	45.59	0.06	83.58	161	0.57	0.03	44.72	0.05	67.50	0.08	77.64
	DT-3K	61	0.91	-0.0	-1.26	0.04	31.90	0.08	83.61	29	0.97	-0.02	-22.16	0.04	33.48	0.04	93.10
Itajuba	DT-L2	108	0.90	-0.01	-9.96	0.04	36.31	0.04	90.74	37	0.95	-0.04	-27.90	0.05	35.34	0.06	75.68
	DB-L2	171	0.73	-0.01	-10.03	0.05	55.69	0.08	82.46	150	0.81	-0.01	-15.81	0.04	46.38	0.05	82.0
	DT-DB-L2	108	0.90	-0.01	-9.96	0.04	36.31	0.04	90.74	37	0.95	-0.04	-27.90	0.05	35.34	0.06	75.68
	MAIAC-1km	183	0.82	-0.04	-27.53	0.05	37.82	0.07	77.05	157	0.83	-0.03	-26.39	0.05	39.89	0.06	78.98
	DT-3K	197	0.93	-0.02	-13.09	0.05	33.51	0.07	77.21	152	0.91	-0.03	-19.65	0.05	38.01	0.06	77.36
Cachoeira	DT-L2	201	0.78	-0.03	-24.20	0.05	35.34	0.06	78.32	140	0.87	-0.03	-22.40	0.05	39.33	0.07	74.75
Paulista	DB-L2	263	0.56	-0.02	-17.85	0.06	43.33	0.08	75.0	238	0.75	-0.02	-13.60	0.05	39.74	0.07	80.25
	DT-DB-L2	201	0.78	-0.03	-24.20	0.05	35.34	0.06	78.32	140	0.87	-0.03	-22.40	0.05	39.33	0.07	74.75