Universidade de São Paulo Instituto de Astronomia, Geofísica e Ciências Atmosféricas Pos-Graduação em Meteorologia

Alejandro Herman Delgado Peralta

Simulation of Tropospheric Ozone Formation in the Metropolitan Area of São Paulo under Climate Change Scenarios

São Paulo

Alejandro Herman Delgado Peralta

Simulation of Tropospheric Ozone Formation in the Metropolitan Area of São Paulo under Climate Change Scenarios

Dissertation submitted to "Instituto de Astronomia, Geofísica e Ciências Atmosféricas" (IAG) of the "Universidade de São Paulo". It is a partial requirement to obtain a Master's degree in Science. Corrected version. The original is available in the Unit.

Concentration area: Meteorology Advisor: Prof.^a Dr.^a Maria de Fátima Andrade

This dissertation work is dedicated to my parents for their love and those who have committed to science. #NoScienceNoFuture #SinCienciaNoHayFuturo

Acknowledgements

This research was made possible by the CAPES (*Coordenação de Aperfeiçonamento de Pessoal de Nível Superior*), who funded the work during the master's studies. The University of São Paulo through the Institute of Astronomy, Geophysics and Atmospheric Sciences, and the laboratory LAPAt, supported this work by providing its facilities as a computer server and the remote connection during the pandemic. I also thank the State Company for the Environment (CETESB) for providing free hourly data from its station network available on its website (QUALAR). I am very grateful to the following people for supporting and encouraging me; without whose help, this work would never have been possible:

- Prof.^(a) Dr.^(a) Maria de Fátima Andrade, who gave me many valuable insights in all stages of this work.
 Professors responsible of each master's courses and during the qualification exam: Adalgiza Fornaro,
 Edmilson Dias de Freitas, Maria Assunção Faus da Silva Dias, Márcia Akemi Yamasoe, Ricardo Hallak,
 and Rita Yuri Ynoue.
- My friends Dr. Angel Vara Vela and Dr. Mario Gavidia Calderón for their valuable suggestions, discussions, support, and teaching related to the WRF-Chem model (setup, create emission files, and postprocessing).
- To my friends Angela Ascarza and Guisselle Castillo for the proofreading of the English writing.
- I also acknowledge Dr. Pieter Tans (NOAA/GML) and Dr. Ralph Keeling (Scripps Institution of Oceanography) for the permission to use the CO₂ trend plots.

Also, I would like to thank my family (especially my wife and my little cat) for their support and encouragement to successfully address this challenge. Finally, thanks to those friendly people to make my life more comfortable during my stay in Brazil and at the University of São Paulo.

Thank God for good memories, muito obrigado, muchas gracias.

Abstract

The impact on ozone formation in the Metropolitan Area of São Paulo (MASP) in 2030 was analyzed through two atmospheric projections based on the Representative Concentration Pathway (RCP) emission scenarios. The modeling system applied in this study was the Weather Research and Forecasting with Chemistry (WRF-Chem) model, maintaining the same emission rates in the MASP and surroundings from road transport, industry, and residential sectors of the base year of 2018. The study of tropospheric ozone (O_3) has a particular interest in air quality and climate change due to its importance as an atmospheric pollutant and a greenhouse gas. It was considered three simulations, corresponding to a control period (Sep-Oct 2018), and two projected scenarios for Sep.-Oct. 2030 based on the RCP emission projections for 4.5 and 8.5 W m⁻² as radiative forcing. We evaluated the model performance for surface ozone and meteorological parameters using available hourly measurements from the Companhia de Tecnologia de Saneamento Ambiental (CETESB, Environmental Protection Agency) and the Instituto de Astronomia, Geofísica e Ciências Atmosféricas (IAG) Climatological Station. Also, emissions representativeness for September and October 2018 were verified through the WRF-Chem model using a configuration based on the MASP for actual conditions. The evaluation showed that surface ozone simulations for stations located in the MASP comply at least with two of three statistical benchmarks. Surface ozone simulations under RCP 4.5 and RCP 8.5 scenarios revealed variations, mainly in the peak ozone concentrations. For September (2018 and 2030), there was an increase in O₃ concentrations under the RCP 8.5 scenario due to the highest temperature. This increment reached $+15.05 \ \mu g \ m^{-3}$ on average in the MASP (urban areas), calculated from the maximum daily rolling mean of 8 hours (MDA8). However, simulations for the RCP 4.5 scenario showed a reduction of surface ozone formation in urban stations (-8.5 μ g m⁻³ on average). In this scenario, there was a slight decrease in temperatures as a monthly average. On the other hand, October (2018 and 2030) simulations presented differences in rainy periods that affected ozone formation. For this month, the RCP 4.5 scenario presented a marked increase of MDA8 surface ozone as a monthly average with $+10.01 \ \mu g \ m^{-3}$ in stations classified as 'Forest preservation.' The RCP 8.5 scenario for October 2030 presented a minor increase (+4.51 $\mu g m^{-3}$ as MDA8 average) for the same stations than the RCP 4.5 scenario. Also, the highest temperature in September 2030 (+2.5 ± 0.12 °C on average) for the RCP 8.5 scenario increased the biogenic emission rates inside the domain what can be a driver for the ozone formation. The relative humidity decreased (-6.76 ± 1.19 % on average) for this specific scenario (RCP 8.5) and month. Regarding accumulated rain, the RCP 4.5 scenario presented higher daily simulated values between September 20-24, 2030. However, simulations of total monthly rain values revealed a decrease for the RCP scenarios, in which the RCP 8.5 presented low values in Sep-Oct 2030. These findings tell us that ozone concentrations increase under future meteorological conditions based on the climate change scenarios (RCP 4.5 and RCP 8.5). The decision-makers with this information can establish policies to mitigate the climate change impact on health.

Resumo

Duas projeções atmosféricas baseadas nos cenários de emissão chamados Representative Concentration Pathways (RCP) foram usadas como condições meteorológicas inicias e de fronteira no modelo Weather Research and Forecasting with Chemistry (WRF-Chem) para avaliar o impacto na formação do ozônio na Região Metropolitana de São Paulo (RMSP) e vizinhança em 2030. O estudo do ozônio (O₃) troposférico apresenta um interesse especial pela importância na qualidade do ar e nas mudanças climáticas por ser um poluente atmosférico e um gás de efeito estufa. Neste estudo foram realizadas três simulações com o modelo WRF-Chem. O primeiro foi um período de controle (setembro e outubro de 2018), e dois cenários projetados para o ano 2030 (setembro e outubro) com base nas projeções de emissão RCP para 4.5 e 8.5 W m⁻² como forçamento radiativo. Foram consideradas as mesmas taxas de emissão de poluentes dos setores de transporte rodoviário, industrial e residencial para o ano base de 2018 e a previsão em 2030. O modelo WRF-Chem teve sua acurácia analisada para as simulações meteorológicas e de qualidade do ar com base em comparações com parâmetros meteorológicos e observações das medições horárias na rede de estações de monitoramento de qualidade do ar da Companhia de Tecnologia de Saneamento Ambiental (CETESB) e da estação meteorológica do Instituto de Astronomia, Geofísica e Ciências Atmosféricas (IAG), localizada em Água Funda. As simulações de ozônio superficial para estações localizadas nas cidades do estado de São Paulo cumprem pelo menos com dois dos três benchmarks estatísticos recomendados para avaliar modelos fotoquímicos. As simulações de ozônio de superfície sob os cenários RCP 4.5 e RCP 8.5 revelaram variações, principalmente nas concentrações de pico de ozônio. Para setembro (2018 e 2030), houve um aumento nas concentrações de O_3 sob o cenário RCP 8.5 devido à temperatura mais alta. O aumento atingiu $+15.05 \ \mu g m^{-3}$ em média na área da RMSP, calculado a partir da média móvel diária máxima de 8 horas (MDA8). No entanto, as simulações para os cenários do RCP 4.5 mostraram uma redução da formação de ozônio superficial (-8.5 $\mu g m^{-3}$ em média nas estações urbanas), em que houve uma ligeira diminuição das temperaturas. Por outro lado, as simulações para outubro (2018 e 2030) apresentaram diferenças nas condições chuvosas que afetam a formação de ozônio. Para este mês, o aumento do ozônio de superfície MDA8 como média mensal foi mais notadamente para o cenário RCP 4.5 com $+10.01 \ \mu g m^{-3} em$ estações classificadas como 'Preservação de florestas'. O cenário RCP 8.5 para outubro de 2030 apresentou um pequeno aumento (+4.51 μ g m⁻³ da MDA8 média) do que o cenário RCP 4.5. A temperatura mais alta em setembro de 2030 (+2.5 ± 0.12 °C como média) para o cenário RCP 8.5 aumentou as taxas de emissão biogênica dentro do domínio, o que pode ser um estímulo para a formação de ozônio. A umidade relativa diminuiu (-6.76 ± 1.19 % como média) para este cenário específico (RCP 8.5). Em relação à chuva acumulada, o cenário RCP 4.5 apresentou maiores valores simulados diários entre 20 e 24 de setembro de 2030. No entanto, as simulações dos valores totais de chuva mensal revelaram uma diminuição para os cenários futuros, em que o RCP 8.5 apresentou valores baixos em setembro-outubro 2030. Os resultados encontrados de aumento das concentrações de ozônio sob condições meteorológicas futuras com base na temperatura mensal, podem ajudar os tomadores de decisão a estabelecer políticas para evitar impactos na saúde e no clima.

Contents

A	Acknowledgements iii					iii
Al	Abstract					\mathbf{v}
Re	Resumo					vii
Co	onter	ts				ix
\mathbf{Li}	st of	Tables				xi
\mathbf{Li}	st of	Figures				xiii
1	Intr	oduction				1
	1.1	Tropospheric ozone formation				2
	1.2	Climate change and global warming				6
		1.2.1 MASP and climate change				8
	1.3	Previous studies for the MASP and motivation				10
	1.4	Objectives				12
2	Met	hodology				13
	2.1	Model description and experiment design				14
		2.1.1 Model configuration				16
		2.1.2 Initial and Lateral-Boundary Conditions (IC/BC)				20
	2.2	Anthropogenic and biogenic emissions				21
		2.2.1 Road transport emissions				21
		2.2.2 Industry and residential emissions				24
		2.2.3 Biogenic emissions				24
	2.3 Surface air quality and meteorological observations				24	
		2.3.1 Air quality data				25
		2.3.2 Meteorological data				25
	2.4	Model performance evaluation				26
		2.4.1 Hypothesis test for Pearson's r				28
3	Res	ults and Discussion				34
	3.1	Evaluation results for current conditions				34

		3.1.1	Anthropogenic emissions	34
		3.1.2	Meteorological model evaluation	37
		3.1.3	Air quality model evaluation	41
	3.2	Future	changes under RCP 4.5 and 8.5 scenarios	47
		3.2.1	Changes in meteorological conditions	49
		3.2.2	Changes in surface ozone	50
4	Con	clusio	15	59
	4.1	Study	limitations and suggestions for future works	61
		4.1.1	Limitations	61
		4.1.2	Suggestions for future works	61
Bi	bliog	graphy		61
\mathbf{A}	WR	F-Che	m files	69
Α	WR A.1	C F-Che Namel	m files ist used to run the WRF-Chem model	69 69
A	WR A.1	C F-Che Namel A.1.1	m files ist used to run the WRF-Chem model	69 69 69
Α	WR A.1	EF-Che Namel A.1.1 A.1.2	m files ist used to run the WRF-Chem model	69696970
Α	WR A.1 A.2	XF-Che Namel A.1.1 A.1.2 Anthre	m files ist used to run the WRF-Chem model NCEP FNL and RCP namelist.wps NCEP FNL and RCP namelist.input opogenic Emissions Calculation	 69 69 69 70 73
A	WR A.1 A.2	F-Che Namel A.1.1 A.1.2 Anthro A.2.1	m files ist used to run the WRF-Chem model NCEP FNL and RCP namelist.wps NCEP FNL and RCP namelist.input NCEP FNL and RCP namelist.input pogenic Emissions Calculation EDGAR-HTAP processing for industry and residential sectors	 69 69 70 73 73
A	WR A.1 A.2	EF-Che Namel A.1.1 A.1.2 Anthro A.2.1 A.2.2	m files ist used to run the WRF-Chem model NCEP FNL and RCP namelist.wps NCEP FNL and RCP namelist.input opogenic Emissions Calculation EDGAR-HTAP processing for industry and residential sectors LAPAt emission preprocessor for road transport	 69 69 70 73 73 74
A	WR A.1 A.2	EF-Che Namel A.1.1 A.1.2 Anthro A.2.1 A.2.2 A.2.3	m files ist used to run the WRF-Chem model NCEP FNL and RCP namelist.wps NCEP FNL and RCP namelist.input opogenic Emissions Calculation EDGAR-HTAP processing for industry and residential sectors LAPAt emission preprocessor for road transport Biogenic emissions	 69 69 70 73 73 74 81
A B	WR A.1 A.2	EF-Che Namel A.1.1 A.1.2 Anthro A.2.1 A.2.2 A.2.3 Qualit	m files ist used to run the WRF-Chem model NCEP FNL and RCP namelist.wps NCEP FNL and RCP namelist.input opogenic Emissions Calculation EDGAR-HTAP processing for industry and residential sectors LAPAt emission preprocessor for road transport Biogenic emissions y and Meteorological Information	 69 69 70 73 73 74 81 83
A B C	WR A.1 A.2 Air Stat	EF-Che Namel A.1.1 A.1.2 Anthro A.2.1 A.2.2 A.2.3 Qualit	m files ist used to run the WRF-Chem model NCEP FNL and RCP namelist.wps NCEP FNL and RCP namelist.input opogenic Emissions Calculation EDGAR-HTAP processing for industry and residential sectors LAPAt emission preprocessor for road transport Biogenic emissions y and Meteorological Information metrics	 69 69 70 73 73 74 81 83 93

List of Tables

2.1	WRF-Chem model configuration	17
2.2	Vehicular fleet by type and fuel for São Paulo State (CETESB, $2019a$)	21
2.2	Vehicular fleet by type and fuel for São Paulo State (CETESB, $2019a$)	22
2.3	Emissions factors (g/km) based on Pérez-Martínez et al. (2014) and Andrade et al. (2015, 2019)	23
2.4	Intensity of use by vehicle type and fuel	23
2.5	CETESB stations considered to monthly analysis based on five years (2014-2018) \ldots	25
2.6	Statistical benchmarks for surface ozone (1-hr or MDA8) suggested by Emery et al. (2017)	27
2.7	Statistical benchmarks for meteorology parameters suggested by Monk et al. (2019)	28
3.1	Emission rates (kt/year) for the second modeling domain based on activity data for the September	
	2018 period	37
3.2	Statistical results for meteorological parameters for Sep-Oct 2018 (IAG/USP station)	38
3.3	Statistical results for meteorological parameters for Sep-Oct 2018 (all stations) $\ldots \ldots \ldots$	38
3.4	Summary of compliance of statistical benchmarks	41
3.5	Global statistical results for air quality parameters	42
3.6	Statistical results for surface ozone by station type	43
3.7	t-test values for correlation coefficient (r) for Sep. 2018	43
3.8	Summary of compliance of the ozone statistical benchmarks by station types with a cutoff value	
	for NMB and NME	44
3.9	Statistical results for toluene in Sep-Oct 2018 period	47
4.1	MDA8 ozone monthly average in $\mu g m^{-3}$ by station type $\dots \dots \dots$	60
B.1	Air quality and weather stations network in the MASP	85
B.2	Air quality and weather stations network in surrounding areas the MASP in the São Paulo State	86
D.1	Statistical results for meteorological parameters for Sep-Oct 2018 by type	96

List of Figures

1.1	Tropospheric ozone formation in urban areas in the presence of VOC and NO_x		
1.2	Carbon dioxide concentrations in the atmosphere monitored at NOAA's Mauna Loa station,		
	Hawaii observatory from 1958 to the present (Tans and Keeling, 2021). With permission from		
	www.esrl.noaa.gov.	7	
1.3	Radiative forcing by emissions and drivers relative to 1750 up to 2011 (ilustration adapted from		
	IPCC, 2013)	8	
1.4	Comparisons between two climate global simulations and observation data from 1850 up to 2010		
	(ilustration adapted from IPCC, 2013)	9	
1.5	Effective radiative forcing for RCP scenarios (IPCC, 2013)	10	
2.1	WRF-Chem simulation domain	18	
2.2	WRF-Chem simulation for September and 2030	18	
2.3	Flowchart for running the WRF-Chem model, applied to this study	19	
2.4	Temporal distribution of emissions for light-duty (LDV, including motorcycles) and heavy-duty		
	(HDV) vehicles, according to Andrade et al. (2015).	23	
2.5	Air quality and meteorological stations network	30	
2.6	Station classification type based on location in the São Paulo State. Maps Data: Google $©2020,$		
	Maxar Technologies.	31	
2.7	Hourly mean concentrations for surface ozone (O_3) , nitrogen monoxide (NO) , nitrogen dioxide		
	(NO_2) , and carbon monoxide (CO) by station type in the MASP	32	
2.8	Hourly mean concentration by month (Sep-Oct, 2018), based on measurements in Pinheiros and		
	S. André-Capuava stations.	33	
3.1	NO emissions for second modeling domain in SE Brazil	35	
3.2	Emission rates by species (part a) and temporal distribution (part b) for all days for NO emission		
	rates and by sector	36	
3.3	Comparison between observed and modeled values for meteorological parameters for the IAG/USP $$		
	station during the period Sep-Oct 2018	39	
3.4	Wind rose plot for Sep-Oct 2018 period at the IAG/USP station, for modeled and observed values.		
	Wind speed simulations less than 0.5 m/s were replaced by 0 m/s to compare with observations.	40	
3.5	Comparison between observed and modeled values for total daily rain rate during Sep-Oct 2018		
	for the IAG/USP station	40	
3.6	Statistical results for surface ozone during the period of September 2018	45	

3.7	Comparison between modeled and simulated values for surface ozone during September and	
	October 2018 period considering all stations classified by type	46
3.8	Comparison between modeled and simulated values based on MDA8 surface ozone during (a)	
	September and (b) October 2018 period for station types	47
3.9	Average hourly concentration of model simulations during the course of a day (Sep-Oct 2018) of	
	some important pollutants	48
3.10	Average hourly concentrations comparison between observed (Obs.) and modeled (Mod.) values	
	during the course of a day	49
3.11	Monthly mean temperature by scenario and station: (a) September, (b) October	50
3.12	Monthly mean relative humidity by scenario and station: (a) September, (b) October. \ldots .	51
3.13	Daily and monthly total rain by scenarios for the IAG/USP weather station	52
3.14	Surface ozone concentrations for Sep-Oct 2018 compared with RCP 4.5 and 8.5 for Sep. 2030	53
3.15	Maximum daily 8-hr rolling mean for surface ozone concentrations for Sep 2018 and 2030 (RCP $$	
	4.5 and 8.5)	54
3.16	Maximum daily rolling mean for surface ozone concentrations for Oct 2018 and 2030 (RCP 4.5	
	and 8.5)	55
3.17	MDA8 ozone mean spatial changes between Sep 2030 (RCP 4.5 and 8.5) and Sep. 2018 $\hfill \ldots$.	56
3.18	MDA8 ozone mean spatial changes between Oct. 2030 (RCP 4.5 and 8.5) and Oct. 2018	57
3.19	Mean differences of surface maximum daily 8-h average (MDA8) by month between both scenarios	
	(RCP 4.5 and 8.5 for 2030) and the current scenario (2018). \ldots \ldots \ldots \ldots \ldots \ldots	58
A.1	VOC fractions	73
A.2	Spatial distribution of MEGAN version 2 in the 15 km parent modeling domain for September	82
B.1	Air quality and meteorological hourly data, downloaded from CETESB (QUALAR)	84
B.2	Hydrocarbon measurements in some CETESB stations as Toluene and Benzene.	87
B.3	Time series of maximum daily rolling 8-hr mean for ozone and daily mean for temperature	88
B.4	Hourly mean concentrations for NO and NO_2 comparison for September 2018 for station types	
	in the MASP.	89
B.5	Surface ozone variation by month and station type in São Paulo State $\ldots \ldots \ldots \ldots \ldots$	90
B.6	Hourly time series of meteorological parameters registered at IAG/USP weather station	91
B.7	Total daily rain and cloud cover at IAG/USP weather station	92
B.8	Total daily rain and cloud cover in September 2018 based on data collected at IAG/USP weather	
	station	92
D.1	Model results for 2-m relative humidity (Sep-Oct 2018) compared with observations from CETESB $$	
	and IAG stations by types.	97
D.2	Model results for 2-m temperature (Sep-Oct 2018) compared with observations from CETESB $$	
	and IAG stations by types.	98

D.3 Model results for wind speed (Sep-Oct 2018) compared with observations from CETESB $% \left({{\left[{{{\rm{D}}_{\rm{s}}} \right]}_{\rm{spec}}} \right)$			
	IAG stations by station types		
D.4	Model results for NO $_x$ (Sep-Oct 2018) compared with hourly time series from CETESB measure-		
	ments		
D.5	Model results for CO (Sep-Oct 2018) compared with hourly time series from CETESB measure-		
	ments by station types		
D.6	Model results for Toluene compared with CETESB measurements		
D.7	Daily mean of temperature by scenario and station type $\ldots \ldots \ldots$		
D.8	Daily mean of relative humidity by scenario and station type $\ldots \ldots \ldots$		
D.9	Daily mean of wind speed by scenario and station		
D.10	Daily total rain for scenarios and station type		

Chapter 1

Introduction

Bad men need nothing more to compass their ends, than that good men should look on and do nothing.

> John Stuart Mill, inaugural address at St Andrews University, 1867.

The state of São Paulo is located in Brazil's Southeastern region and it is the most populous and developed (measured by the *Índice de Desarrollo Humano* or IDH) region of the country. It is an important state, accounting for 32% of Brazil's Gross Domestic Product (GDP). It has a high population of approximately 46.29 million (estimated for 2020) with a density of 166.25 inhabitants/km², and 29 million vehicles quantified for 2018 (Brazilian Institute of Geography and Statistics, IBGE, 2020*a*). The main urban area in this state is the Metropolitan Area of São Paulo (MASP), which contains 39 municipalities with a population of 21.9 million inhabitants (IBGE, 2020*b*), located around 770 m above sea level and around 45 km from the coast. This megacity presents air pollution episodes associated with the exceedances of air quality standards for particulate matter (less than 10 μ m (PM₁₀) and 2.5 μ m (PM_{2.5})) and ozone concentrations, which the vehicular emissions are the primary driver due to burning different types of fuels (CETESB, 2019*a*). Also, short-lived climate pollutants (SLCP) such as tropospheric ozone and black carbon contribute significantly to climate change (von Schneidemesser et al., 2015).

Air quality is very susceptible to weather conditions, even if emissions are constant and the wind direction does not change (Visscher, 2014). Therefore, as climate change affects the weather conditions, it can impact air pollutant concentrations over time, worsening them, as is observed with surface ozone concentrations in the MASP (CETESB, 2019b). According to IPCC (2013), the potential effect of climate change on surface ozone in polluted regions suggests a 'climate penalty' in which air emissions control will have to consider the temperature rise projection to achieve a specific target of ozone formation reduction.

According to World Health Organization (WHO) and recently studies (Nuvolone et al., 2018), people exposed to higher surface ozone concentrations, above WHO air quality guideline value (100 μ g m⁻³ daily maximum rolling 8-hour mean), can have adverse health effects associated with respiratory diseases: trigger asthma, reduce lung function and cause lung diseases (WHO, 2006). Short-term exposure to higher ozone concentrations is associated with lung function impairment. It is identified as cough and pain on deep inspiration with significant variability for each individual depending on gender, age, and pre-existing pulmonary diseases (Nuvolone et al., 2018). Tropospheric ozone is relevant because it is also a greenhouse gas, and there is robust evidence that it has a detrimental impact on vegetation physiology (IPCC, 2013; von Schneidemesser et al., 2015).

As part of the Metroclima project (FAPESP 2016/18438-0), this study examines variations in the tropospheric (surface) ozone formation as a response to future changes in meteorological conditions in the MASP, without taking into account the future emission rate. The objective of the Metroclima project is to examine the role of the São Paulo megacity emissions as drivers for regional air quality degradation and climate change. The project integrates multi-platform measurements and modeling tools to describe the atmospheric behavior (greenhouse gases and SLCP) and the effects of climate change on the air quality in the MASP. Hence, based on the Metroclima project, this study analyzes two scenarios for the year 2030, where we take meteorological conditions into account based on the Radiative Concentration Pathways (RCP) concept for 4.5 and 8.5 W/m² as radiative forcing, named as RCP 4.5 and RCP 8.5, respectively (van Vuuren et al., 2011).

This dissertation presents general background information in **Chapter 1**, such as an overview of tropospheric ozone formation, climate change, previous studies related to this research for the MASP, motivations and the objectives of this study. **Chapter 2** presents a brief description of the methods related to the air quality model (setup, emissions inventory), meteorological data sets used for initial and lateral boundary conditions, static data (e.g., land use, land cover, surface terrain data), and the model performance evaluation. **Chapter 3** presents results about anthropogenic emissions, the model performance evaluation for air pollutants and weather parameters, and changes about contributions of meteorological projections under RCP scenarios to surface ozone in the MASP and surround it. Additionally, **Chapter 3** presents explanations about ozone formation changes related to emission and meteorological factors. Finally, in **Chapter 4**, this dissertation shows conclusions that respond to the primary and specific objectives. We also suggest future works based on the study limitations related to the period of analysis and the emission inventory projection based on local policy decisions about mitigation controls.

This study can support decision-makers because the ozone formation in the MASP can increase in the future, considering 2020 as one of the three warmest years on record globally according to the World Meteorological Organization (Nullis, 2020). Consequently, in September 2020, the ozone formation reached higher concentrations, according to the CETESB report¹.

1.1 Tropospheric ozone formation

Tropospheric ozone is a secondary pollutant formed by reactions of nitrogen oxides ($NO_x=NO+NO_2$), volatile organic compounds (VOC), carbon monoxide (CO), and methane (CH₄) in the presence of solar radiation (von Schneidemesser et al., 2015).

Ozone is present everywhere. Rates of ozone formation are higher in polluted regions than in the remote² troposphere (i.e., oceans, according to Wolfe et al. 2019). It depends on two major classes of precursors: VOCs

¹Retrieved in this link.

 $^{^{2}}$ The definition of "remote" is somewhat ambiguous and should not be confused with "pristine". According to Wolfe et al. (2019), remote troposphere means non land areas far away from forest and urban areas with significant air emission sources.

(from anthropogenic and biogenic sources) and NO_x . Ozone lifetimes vary in the troposphere depending on altitude, latitude, and season. For example, during summer, the higher water vapor concentration and solar radiation reduce its lifetime (Seinfeld and Pandis, 2016).

Tropospheric ozone photolysis is the principal source of hydroxyl radical (OH) in the presence of water vapor (Brasseur et al., 1999). This OH is often referred to as the 'atmosphere detergent' because it defines the oxidizing capacity of the troposphere. It reacts with most trace species (such as all organic compounds, CO, CH₄, and nitrogen and sulfur species) relevant to climate and air quality to produce carbon dioxide (CO₂) and water (H₂O). Furthermore, the OH is the most important reactive species in the ozone formation because there is a competition between VOCs and NO_x for the OH (Seinfeld and Pandis, 2016).

The basic photochemical cycle of NO_x and O₃ occurs when the solar radiation ($\lambda < 420$ nm) dissociates NO₂ where atomic and excited oxygen (O³P) is released as we can see in the following reaction (Seinfeld and Pandis, 2016):

$$NO_2 + hv \rightarrow NO + O(^3P)$$
 $\lambda < 424 \ nm;$ (1.1)

then, this $(O^{3}P)$ reacts with molecular oxygen (O_{2}) to produce ozone

$$O(^{3}P) + O_{2} + M \to O_{3} + M$$
; (1.2)

finally, nitrogen monoxide reacts with ozone to produce nitrogen dioxide and oxygen

$$NO + O_3 \to NO_2 + O_2. \tag{1.3}$$

Where M could be O_2 or N_2 that absorbs the excess energy. These reactions are known as the photostationary state that controls the ozone mixing ratio. However, as Wallace and Hobbs (2006) and Seinfeld and Pandis (2016) mention, there are other reactions with net ozone production rather than the photostationary state in the remote troposphere as well as regional and urban areas. So, additional species (i.e., CO, CH₄, VOC) lead to the atmosphere's net ozone production. The ozone formation is almost always initiated by reactions between a primary³ hydrocarbon (RH), other organic or CO with the hydroxyl radical (Sillman, 2014; Seinfeld and Pandis, 2016). The reaction of the RH with OH radical removes hydrogen to produce a RO₂ radical, as shown in reaction (1.4). The CO oxidation exhibits many of the key reactions to analyze and understand the troposphere's chemistry. In those reactions, it is relevant to consider the limits of low and high NO_x concentrations. The equivalent reaction for CO forms HO₂, a radical with many chemical similarities to the various RO₂ radicals (Sillman, 2014):

$$RH + OH \xrightarrow{O_2} RO_2 + H_2O \tag{1.4}$$

$$CO + OH \xrightarrow{O_2} HO_2 + CO_2$$
 (1.5)

³For instance, methane (CH₄) and propane (C₃H₈) (Sillman, 2014).

Other reactions with NO produce the conversion to NO_2 ,

$$RO_2 + NO \xrightarrow{O_2} R'CHO + HO_2 + NO_2$$
 (1.6)

The R'CHO represents intermediate organic species or secondary VOC, typically including aldehydes and ketones.

$$HO_2 + NO \rightarrow OH + NO_2$$
 (1.7)

Then, photolysis of NO₂ results in the formation of atomic oxygen (O), which reacts with atmospheric O_2 to form ozone via reactions (1.1) and (1.2). Also, by a reaction to itself (HO₂) and with nitrogen dioxide, highly soluble products are removed by wet deposition,

$$2HO_2 \to H_2O_2 + O_2 \tag{1.8}$$

$$OH + NO_2 + M \to HNO_3 + M. \tag{1.9}$$

Ozone formation is highly dependent on sunlight. Several authors and studies mention that the highest concentrations of ozone occur during the spring and summer periods due to high surface solar irradiation and temperature (von Schneidemesser et al., 2015; Carvalho et al., 2015). Urban areas (metropolitan and surrounding) present different chemical regimes when ozone is formed, referred to as NO_x-saturated (VOC-sensitive) or NO_x-sensitive (VOC-saturated). These regimes are closely associated with their sources (produced by photolysis) and sinks of the odd hydrogen radicals (H, OH, HO₂, in general HO_x) (von Schneidemesser et al., 2015).

The VOC/NO_x ratio is essential to understand how the ozone precursors are relevant to its formation or the increase/decrease behavior. In areas with VOC/NO₂ ratio less than 5.5:1 predominates the OH-NO₂ reaction, retarding the further production of O₃; on the other hand, when the ratio exceeds 5.5:1, OH reacts with VOCs, accelerating O₃ production (Seinfeld and Pandis, 2016). Other authors mention ratios between 8-12 between VOC and NO₂ (for instant, ratio of 11 in the MASP, according to Orlando et al., 2010). This analysis can be done using of the ozone isopleth plot⁴ that shows the formation of ozone according to the VOC/NO_x ratio. Usually, many urban areas have a higher concentration of NO_x, called VOC-sensitive; on the other hand, there is a higher concentration of VOC in a rural area, called NO_x-sensitive (von Schneidemesser et al., 2015).

In urban areas, such as megacities (more than 10 million people according to World Urbanization Prospects⁵ for 2018), different gaseous pollutants are released into the atmosphere, mainly due to vehicle emissions and industrial activities. In South America, we have megacities with air quality problems such as Buenos Aires (Argentina), São Paulo (Brasil), Rio de Janeiro (Brasil), Lima (Peru), and Bogotá (Colombia). Figure 1.1 illustrates complex photochemical reactions, when for one OH radical produced from one O_3 , then two O_3 are provided from a complex mechanism of reactions. According to von Schneidemesser et al. (2015), many urban areas have higher NO_x concentrations, which regime tends to be NO_x -saturated or VOC-sensitive. Also, in

 $^{{}^{4}}$ It is a helpful diagram to make the right decision about which pollutant emissions must be reduced (Seinfeld and Pandis, 2016). 5 Retrieved in this link.



Figure 1.1 Tropospheric ozone formation in urban areas in the presence of VOC and NO_x . Note. Adapted from Jacob (1999). R is an organic group.

urban areas is essential to consider a phenomenon called "NO_x titration," in which, NO is the ozone sink via reaction (1.3).

According to CETESB (2019*a*) and Andrade et al. (2017), vehicles in the MASP are responsible for the emissions of the majority of air pollutants. Vehicles contribute to a high percentage of air pollutant emissions (CETESB, 2019*a*): 97% of CO, 75% of HC, 64% of NO_x, 17% of sulfur oxides (SO_x), and 40% of particulate matter (PM). There are different fuel types for the road transport sector in Brazil. Heavy-duty vehicles use diesel-fueled which is a significant source of NO_x emissions. Flex-fuel vehicles can burn both gasoline C (around 75% gasoline mixed with a range from 18% to 27% anhydrous ethanol, also called gasohol) or hydrous ethanol (7.5% of maximum water content) (CETESB, 2019*a*). Although ethanol in vehicles may lead to some reductions in CO and VOC emissions, it produces aldehydes during combustion mainly to form acetaldehyde in the exhaust emissions (Gaffney and Marley, 2009). Primary acetaldehyde leads to O₃ formation, H₂O₂, formic acid, CO, peroxyacetyl nitrate (PAN), acetic acid, and peracetic acid (Gaffney and Marley, 2009).

In Brazil, unlike other countries, high concentrations of acetaldehyde have been found in the atmosphere (Nogueira et al., 2014), which in the MASP presents a NO_x-saturated or VOC-sensitive condition (Sánchez-Ccoyllo et al., 2006; Alvim et al., 2018). In this context, VOC and CO have reactions with hydroxyl radicals that generate peroxy radicals. These (RO₂) compete with O₃ when they react with NO to produce NO₂, as shown in Figure 1.1. Alvim et al. (2018) used the ozone isoplet Package for Research (OZIPR) trajectory model to determine the significant ozone precursors as VOC in São Paulo. They found that the ten most abundant VOC during the 2011-2012 period were ethanol, acetaldehyde, formaldehyde, acetone, propane, ethane, ethene, butane, 1-ethyl-4-methyl benzene, and 1,2,4-trimethylbenzene. Also, Alvim et al. (2018) presented an ozone isopleth by season, in which they illustrate one for spring based on the September 2011 period. They found in the MASP VOC/NO_x ratios less than 4, representing polluted urban areas with high concentrations ratio of NO_x

to VOCs; therefore, the VOC/NO_x ratio is low, and ozone formation will depend on the VOC concentrations. According to the same authors, the aldehydes (acetaldehyde and formaldehyde) were responsible for 74% of the ozone formation, followed by aromatics (14.5%). Therefore, simulation results in the MASP done by Alvim et al. (2018) showed that the most effective alternative for limiting the ozone formation is to reduce the VOC emissions through aldehydes from ethanol burning. Those findings are in agreement with a recent study (Dominutti et al., 2020), in which their results suggest a strong influence of vehicular emissions in the VOC levels in the MASP, mainly associated with the large consumption of ethanol.

1.2 Climate change and global warming

Climate is an average meteorological condition over a long time, at least 30 years, according to the World Meteorological Organization (WMO) (IPCC, 2013). So, climate change is a variation of the normal meteorological conditions that persist over time. Hence, weather and climate are not the same phenomenon. Weather conditions always change every day and it is interesting for scientists to forecast extreme episodes using numerical weather models.

Climate change and global warming are different concepts but both are related. Global warming depends on the greenhouse gases (GHG), however this last term does not mean something "bad". The greenhouse effect is important to maintain the life on Earth where the GHG trap the heat, absorbing a fraction of the infrared (IR) waves emitted from the warm surface and is re-radiated by these GHG back to the Earth's surface (Farmer and Cook, 2013). By definition, GHG are chemical species that:

"absorb and emit radiation at specific wavelengths within the spectrum of terrestrial radiation emitted by the Earth's surface, the atmosphere itself, and by clouds" (IPCC, 2013).

Water vapor (H₂O), carbon dioxide (CO₂), nitrous oxide (N₂O), methane (CH₄), and tropospheric ozone (O₃) are the primary GHG (IPCC, 2013; von Schneidemesser et al., 2015).

 CO_2 is a very important GHG due to its capacity to warm the lower atmosphere efficiently, and it is known as the Earth's thermostat. Life has evolved when CO_2 levels are approximately 280 ppm (Farmer and Cook, 2013), considering only natural forcing. Also, life has changed the composition of the Earth's atmosphere, governing the dynamics of CO_2 on the planet today (Kasting, 1993).

Since the Industrial Revolution, burning fossil fuels are the primary source of CO_2 . We can note in Figure 1.2 that the CO_2 levels have increased. In 1960, it was 316 ppm with the rate of increase less than 1.0 ppm per year, and in 2020, it reached 417 ppm with the rate of increase of 2.4 ppm per year (Tans and Keeling, 2021; Letcher, 2021). Present-day atmosphere CO_2 levels exceed the natural equilibrium of absorption (oceans, biota, and land) with accumulated effect due to CO_2 has a prolonged life-time in the atmosphere because it is very unreactive (Letcher, 2021). Unfortunately, this rising of CO_2 levels does not stop despite the warnings by scientists and institutions like the IPCC (2013).

We have experimented an increase of global mean temperature every year due to IR absorption from terrestrial radiation by rising CO_2 concentrations, re-radiated back toward the Earth's surface. There can be more evaporation of water (e.g., oceans), increasing the vapor content in the troposphere, which more IR absorption



Figure 1.2 Carbon dioxide concentrations in the atmosphere monitored at NOAA's Mauna Loa station, Hawaii observatory from 1958 to the present (Tans and Keeling, 2021). With permission from www.esrl.noaa.gov.

(Letcher, 2021). This warming effect by CO_2 is related to the climate sensitivity⁶ expressed by the feedback effect, whereby an increasing temperature causes an increasing concentration of water vapor in the atmosphere from oceans, which causes more increasing temperature with negative effects (permanent glacial melting, disappearing Arctic ice cap, change in cloud patterns, ocean acidification) (Farmer and Cook, 2013; Letcher, 2021). Hence, this increase of global mean temperature is called global warming; moreover, Tuckett (2021) mentions that this name has become *global heating*.

The interaction between gases and aerosols is very complex, some have negative and positive radiative forcing $(RF)^7$. However, there are estimates and associated uncertainties about those interactions related to radiative forcing, published by the IPCC (2013). Figure 1.3 illustrates the relationship between emitted compounds (i.e., gases and aerosols) and resulting atmospheric drivers considering the RF relative to the pre-industrial revolution (1750) up to 2011. This chart shows that the cloud adjustments' level of confidence due to aerosols is low, which means that there are many uncertainties related to the cloud formation and the RF changes. We again note that the tropospheric ozone is responsible for the positive RF more than the negative RF by decreasing the stratospheric ozone concentration.

Considering all interactions, climate scientists demonstrated that human activities impacts the global climate due to the increase of GHG. Figure 1.4 shows model results where we can see comparisons between two climate global simulations (i.e., only natural forcing and both natural and anthropogenic forcing) and observation data (IPCC, 2013). We note again the model simulations that include the anthropogenic forcing are closer to the observations. This behavior is a warning about how humankind will face the adverse effects in different environmental components (extended droughts, increasing wildfires, increasing urban air pollution, insect infestations, intensifies storms, changing rainfall, and agricultural patterns) (Farmer and Cook, 2013).

To understand future impacts due to atmospheric GHG and other pollutants, the IPCC (2013) adopted four scenarios, using the concept of representative concentration pathways (RCP). According to the target level

⁶IPCC (2013, Glossary Annex III): "The effective climate sensitivity (units: $^{\circ}$ C) is an estimate of the global mean surface temperature response to doubled carbon dioxide concentration that is evaluated from model output or observations for evolving non-equilibrium conditions. It is a measure of the strengths of the climate feedbacks at a particular time and may vary with forcing history and climate state, and therefore may differ from equilibrium climate sensitivity."

⁷According to IPCC (2013), "radiative forcing is the change in the net, downward minus upward, radiative flux (expressed in W m⁻²) at the tropopause or top of atmosphere due to a change in an external driver of climate change, such as, for example, a change in the concentration of carbon dioxide or the output of the Sun". The IPCC report refers to RF as the change relative to the year 1750 as the global and annual average value.



Figure 1.3 Radiative forcing by emissions and drivers relative to 1750 up to 2011 (ilustration adapted from IPCC, 2013) Notes. Very height (VH), height (H), medium (M), low (L).

for 2100, the RCP emission scenario depends on the radiative forcing caused by GHG and other agents such as changes in land use and aerosol concentrations (van Vuuren et al., 2011). Figure 1.5 shows the effective radiative forcing trajectories, known as RCP 2.6, RCP 4.5, RCP 6.0, and RCP 8.5. These scenarios include "time series of emissions and concentrations of GHG and aerosols and chemically active gases, as well as land use/land cover" as mentioned in (Most et al., 2000; cited in IPCC, 2013).

In this dissertation, two scenarios (RCP 4.5 and RCP 8.5) were analyzed for 2030; in that year, those projections are close to others (RCP 2.6 and RCP 6.0), as shown in Figure 1.5. After that year, uncertainties about land cover changes and anthropogenic emissions would increase significantly. The RCP 4.5 is a stabilization scenario and assumes all nations will comply with emission mitigation through changes in the energy system, including shifts to electricity from lower emissions energy technologies and carbon capture and geologic storage technology (Thomson et al., 2011). The RCP 8.5 is the very high baseline emission scenario, representing the range of non-climate policy known as "business as usual," combined with the growing population and high demands of fossil fuel and food (Riahi et al., 2011).

1.2.1 MASP and climate change

According to several authors (Andrade et al., 2017; de Lima and Magaña Rueda, 2018), the MASP has a climate with mild temperatures, with a defined dry season (June to August) and humid summers (December to February). There are rainfalls over the year (mainly during the summer season), influenced by weather systems such as cold fronts, the South Atlantic Convergence Zone, squall lines, sea breezes, and the urban effect



Figure 1.4 Comparisons between two climate global simulations and observation data from 1850 up to 2010 (ilustration adapted from IPCC, 2013)

(Andrade et al., 2017; de Lima and Magaña Rueda, 2018). Three relevant factors dominate the air circulation (Oliveira et al., 2003): (i) sea breeze, (ii) mountain-valley circulation, and (iii) urban effects, such as roughness, building-barrier, and urban heat island (UHI) effects. The sea breeze circulation influences the temperature differences within the MASP, producing a strong convergence zone during its lifetime (diurnal variation), which moves from southeast to northwest across the city (de Lima and Magaña Rueda, 2018). The same authors also mentioned the UHI phenomenon's contribution to precipitation patterns changes due to convective air circulation when the synoptic-scale winds are weak.

The MASP has suffered from climate change events since the year 1930 (Marengo et al., 2020), with marked effects since 1960s (de Lima and Magaña Rueda, 2018) due to the most remarkable growth of the urban spot and accelerated population increase due to the plentiful supply of jobs in the MASP during those years. The main effects in the MASP due to climate change from 1960 until 2019 are: (i) increases in air temperature, extreme precipitation events, and atmospheric stability; (ii) reduction in the number of days with light precipitation, and relative humidity (mainly during the nights) (Marengo et al., 2013; de Lima and Magaña Rueda, 2018; Nobre et al., 2019; Marengo et al., 2020). For the future, climate projection for Brazil showed trends of rising temperatures (Nobre et al., 2019). Locally, the MASP may suffer the increase in the intensity and frequency of heavy precipitation and negative trends of light rain with the possibility of prolonged dry periods of continuous days (Marengo et al., 2013).



Figure 1.5 Effective radiative forcing for RCP scenarios (IPCC, 2013)

Notes:

WMGHG is Well-Mixed Green House Gases in dash-dot. Long dashes with squares are ozone; short dashes with diamonds are aerosol. RCPs 2.6, 4.5, and 6.0 net forcings at 2100 are approximate values using aerosol projected for RCP8.5, according to IPCC (2013).

1.3 Previous studies for the MASP and motivation

Several modelling and observational studies about air quality in the MASP are available since 2006, mainly for fine particles and tropospheric ozone due to their impacts on human health and the environment. Sánchez-Ccoyllo, Ynoue, Martins and Andrade (2006) analyzed the tropospheric ozone formation in the MASP through the California Institute of Technology (CIT) model to propose emissions reduction to improve the air quality, based on VOC/NO_x ratio. They found that the urban area in the MASP has a VOC-sensitive condition, and the surrounding areas were slightly NO_x-sensitive. Based on this information, they proposed that a reduction of VOC anthropogenic emissions could control ozone formation.

Vara-Vela (2013) studied the impact of ozone formation in the MASP due to changes in its precursors' emission factors. He mentioned the importance of the ozone precursors, being necessary to study with more relevance, the VOC. This pollutant group encompasses several reactive compounds (e.g., isoprene, aromatics, aldehydes) that enhance ozone formation in urban areas. Also, they concluded that the grid domain with 3 km of spatial resolution represented better the temporal ozone formation in the simulations with the WRF-Chem model. Finally, regarding model parameterizations, they found that physical and chemical configurations in the WRF-Chem model were coherent to represent the ozone formation and its transport.

Regarding the influence of climate change over ozone formation, Mazzoli da Rocha (2013) used the database from the global climatic model CCSM3 into the WRF-Chem model. She analyzed two future years (2020 and 2050) compared with a case-control (the year 2011). She found minor differences in the ozone formation for the coming years when emissions do not change with time. However, she mentioned that if the air pollutant emissions increase, ozone formation will have a considerable impact on the MASP.

Several studies, for instance, Carvalho et al. (2015), Andrade et al. (2015), and Andrade et al. (2017) have been analyzing the air quality conditions related to emission sources, atmospheric chemistry, air quality modeling, and the evolution of pollutant concentrations due to regulations on air quality and emission sources. Carvalho et al. (2015) highlighted that high ozone and particle concentrations are mostly associated with vehicular emissions, in which high ozone concentrations seasonal behavior occurs during the spring, probably due to lower cloud cover. They also presented recommendations for improving the air quality based on better public transport systems and economic incentives for clean-air technology.

Andrade et al. (2015) applied a novel bottom-up approach for road transport emission inventory based on the real conditions in the MASP, verified through the WRF-Chem model. Simulations compared with measurements showed good quality for ozone; however, the approach requires improving the NO_x and fine particles simulation results. They mentioned uncertainties of emission inventories and the boundary conditions as aspects that have to be evaluated. In that study, the operational model configurations for the gas-phase chemistry mechanism was the carbon-bond mechanism, version Z (CBM-Z; Zaveri and Peters, 1999). This gas-phase mechanism included ethanol and other oxygenated compounds to be represented explicitly, considering that the fuel consumption in the MASP has a high percentage of ethanol.

Andrade et al. (2017) emphasized the greatest challenge of controlling secondary pollutants such as ozone and fine particles. They mentioned the ethanol biofuel impacts on the MASP as a relevant contributor of ozone formation in two ways:

- acetaldehyde emissions due to incomplete combustion,
- and the direct evaporative emission of ethanol.

Their review recommended improving the emission inventories for their use in the air quality modeling, including the stationary sources (use of wood and charcoal for cooking in restaurants, industrial process) and evaporative emissions (including gas stations). They suggested an effective way of improving air quality to scrap old vehicles (those 10-15 years of age) in the MASP according to recommended practices in most megacities. Finally, they recommended studies based on the impact of climate change for different scenarios on air quality and human health.

Gavidia-Calderón et al. (2018) studied how the chemical boundary conditions (CBC) impact the WRF-Chem model related to the tropospheric ozone formation in the MASP. He found the CBC did not considerably impact the spring period due to the increase of local sources and photochemical reactions during that time. By following Warner (2011), he used a grid-resolution (9 km x 9 km) with sufficient domain area to avoid propagation errors from meteorological boundary conditions related to the MASP, located at the center of the modeling domain.

Climate change projections and their assessment based on RCP scenarios for Brazil regions were developed until the year 2100 by Chou et al. (2014), Cunningham et al. (2017), Marengo et al. (2018), and Nobre et al. (2019). As mentioned by Nobre et al. (2019), Brazil was vulnerable to extreme climate events in the past (e.g., the year 2014). For the future, climate projection for Brazil showed trends of rising temperatures (Nobre et al., 2019). Consequently, as the term "climate penalty" effect suggests (mentioned in IPCC 2013), tropospheric ozone concentrations could increase for future years. However, few studies (Mazzoli da Rocha, 2013; Schuch et al., 2020) analyzed the weather projections under climate change scenarios and their effects on the ozone formation in the MASP.

In recent work, Schuch et al. (2020) analyzed the sensitivity of the ozone concentration and fine particles (less than 2.5 μ m) to change in emissions under the RCP 4.5 scenario over Brazil to determine the signal and spatial patterns, using short-period simulations. Despite uncertainties about future changes in emissions and land use, they found a decrease in O₃ concentrations, located at the São Paulo and Rio de Janeiro metropolitan areas. Schuch et al. (2020) mentioned a possible explanation that the difference is caused by the increase of NO_x emissions in different VOC/NO_x regimes. However, in this study, the "climate penalty" effect due to temperature increase was not discussed related to its role in the tropospheric ozone formation, as suggested by the IPCC (2013).

Therefore, the primary motivation for carrying out this dissertation is to complement further analysis based on the climate penalty effect. Mainly associated with weather projections under two RCP scenarios that depict the intermediate and worst-case scenario, maintaining the same air pollutant emission rates for the evaluated period (2018) and the projected period (2030). This dissertation's findings can be important for decision-makers and public policymakers about the ozone formation in urban and regional areas in the São Paulo state. Also, this dissertation shows future works related to improving limitations found in this work, such as how the urban environment and land use will evolve in the following decades. Furthermore, how can we add these land-use changes in the model configuration?

1.4 Objectives

This dissertation aims to study the impact of future climate change scenarios on tropospheric ozone formation in the Metropolitan Area of São Paulo (MASP) in 2030 using of the WRF-Chem model. Thus, specific objectives are:

- Prepare and update the emission files for the year 2018, representing the modeling domain areas centered in the São Paulo state.
- Setup the WRF-Chem model options for physical and chemistry modules to analyze ozone formation based on the MASP emission sources' features.
- Evaluation of the WRF-Chem model results, based on the case-control study (September and October 2018).
- Obtain surface ozone concentrations from the WRF-Chem model based on the RCP scenarios and compared them with the model results representative of the case-control study.
- Study the current conditions that affect the tropospheric ozone formation through the WRF-Chem model compared with observations.
- Study how changes in future meteorological conditions in 2030 impact the surface ozone formation in the MASP and around it.

Chapter 2

Methodology

It is not enough to have a good mind, the main thing is to use it well.

René Descartes, Discourse on the Method

This chapter describes the relevant information to run the WRF-Chem model and to evaluate the simulation results through recommended statistical benchmarks applied for photochemical models. The model configuration, the meteorological initial and boundary conditions (IC/BC), and the emission approaches are also presented. The description of how the anthropogenic and biogenic emissions files were built is included, considering assumptions and their limitations for future projections. Additionally, measured air data inside the model domain area and its geographical information is described. Finally, this section describes statistical benchmarks used for the model performance evaluation for September to October 2018. Statistical results show us the model performance for current conditions and give us an insight about the model precision to simulate the ozone formation.

Impacts of climate change over surface ozone formation preferentially require analyzing multiyear ensemble simulations. However, a coupled atmosphere-chemistry model demands high computing time to do this task. Due to this computational limitation, five years of data (2014-2018) was analyzed to review historical ozone measurements in the São Paulo state. After that, as an alternative way, we chose only two months when high surface ozone episodes frequently occurred as a worst-case scenario. September and October represent these two months in the spring season when higher ozone concentrations are measured due to low cloud cover and high frequency of sunny days, also consistent with findings of Carvalho et al. (2015).

2018 was chosen as a period that represents current conditions. Based on the air quality report published by CETESB (2019*b*), that year presented a high percentage of good air quality conditions than previous years (2014-2017) due to few days of meteorology conditions that enhanced ozone formation. Only September and December 2018 had a high number of frequent violations of São Paulo air quality standard for ozone (140 μ g m⁻³, 8-h rolling mean). June, August, and October did not have exceedances of the air quality standards for ozone. Nonetheless, the same report mentioned that there isn't a tendency for many years due to photochemical reactions depending on many factors, which have a non-linear relation.

Two scenarios were analyzed using the meteorological datasets as meteorological IC/BC based on the RCP 4.5

(stabilization scenario) and the RCP 8.5 (worst-case scenario) for two month (2030) as future projections. Only the current simulation for short periods of the year 2018 (Sep. and Oct.) was evaluated through statistic parameters for meteorology and surface ozone concentrations based on recommended statistical benchmarks for photochemical models. To isolate the effects of changes in future weather conditions, these three sets of simulations (current and future) share the same anthropogenic emission files and geographical data (i.e., land use/land cover, topography height). This means that any variation on future simulations is caused by the weather conditions' effects, affecting the biogenic emissions inside the WRF-Chen due to they depend on the temperature (Guenther et al., 2006).

However, the same chemical boundary conditions, adding the same anthropogenic emissions, and geographical data may cause uncertainty in the model results for future scenarios. Likewise, according to van Vuuren et al. (2011), land use influences the climate system due to interactions with the local atmosphere (i.e., albedo and surface roughness) and biogenic emissions. For that reason, only the year 2030 was analyzed instead of include the year 2050 or even more future years, considering land use and anthropogenic emission could change in the future, and the uncertainty may increase significantly for 2050.

2.1 Model description and experiment design

The Weather Research and Forecasting with chemistry model (WRF-Chem) (Grell et al., 2005) is an open-source community model, developed by many different groups. Several research institutes collaborated for the development of the meteorology model Weather Research and Forecasting (WRF) in 1990 through a partnership of the National Center for Atmospheric Research (NCAR), the National Oceanic and Atmospheric Administration (NOAA) -represented by NCEP and ESRL-, the United States Air Force, the Naval Research Laboratory, the University of Oklahoma, and the Federal Aviation Administration (Skamarock et al., 2019). The WRF model is an Eulerian non-hydrostatic designed for use in atmospheric research and operational forecasting. Many options of physical parameterizations let us represent subgrid processes that the model cannot explicitly calculate. Details about WRF meteorological module can be found in Skamarock et al. (2019).

The model integrates the meteorological and chemical modules. The mass coordinate version called Advanced Research WRF (ARW) is the only dynamical core coupled to the chemical module. To summarize, ARW's equations are cast in flux form from conserved variables; non-conserved variables such as pressure and temperature are diagnosed from the conserved prognostic variables (Grell et al., 2005; Skamarock et al., 2019). As mentioned by Skamarock et al. (2019), the main features of the ARW system (version 4) are:

- Equations: Fully-compressible. Conserves dry air mass and scalar mass.
- *Prognostic variables*: Velocity components, and perturbation variables. Optionally, turbulent kinetic energy and any number of scalars.
- Vertical coordinate: Terrain-following. Top of the model is a constant pressure surface.
- Horizontal grid: Arakawa C-grid staggering.
- Time integration

- Spatial discretization: 2nd- to 6th-order advection options in horizontal and vertical.
- Turbulent mixing and model filters
- Initial conditions: Three dimensional for real-data.
- Lateral, top and bottom boundary conditions
- Earth's rotation: Full Coriolis terms included.
- *Mapping to sphere*: polar stereographic, Lambert conformal, Mercator, and latitude-longitude. Curvature terms included.
- Nesting: One-way, two-way, and moving nests.
- Nudging: Grid, spectral, and observation nudging capabilities.
- Global grid: Global simulation capability.
- Tropical channel.

The physic parameterizations of the model are summarized as follow Skamarock et al. (2019):

- *Microphysics*: They are related to the development of hydrometeors such as water vapor, cloud (ice crystals), and precipitation processes (rain drops) based on how the size distributions of particle types are represented. They are extremely important in climate modeling and their performance can depend on season and the meteorological process that prevail in specific geographic regions (Warner, 2011).
- *Cumulus parameterizations*: Deep and shallow convection, adjustment, mass-flux, and scale aware schemes available.
- *Surface physics*: Multi-layer land surface models ranging from a simple thermal model to full vegetation and soil moisture models, including snow cover and sea ice. Urban parameterizations are available.
- Planetary boundary layer physics: Turbulent kinetic energy prediction or non-local K schemes.
- Atmospheric radiation physics: Longwave and shortwave schemes with multiple spectral bands and a simple shortwave scheme suitable for climate and weather applications. Cloud effects and surface fluxes are included.

The physical and chemical atmosphere processes have interactions that affect meteorological conditions (e.g., clouds and precipitation) and air pollutants concentrations (i.e., ozone, nitrogen oxides, carbon monoxide, aerosols). For instance, the chemistry can also affect the meteorological conditions through its effect on the radiation budget and aerosols' interaction with cloud condensation nuclei (Grell et al., 2005).

The WRF-Chem model considers those interactions; for that reason is an "online" model. The chemistry module is integrated simultaneously with the meteorology module, allowing both components to have the same transport scheme (mass and scalar preserving), the same grid, the same physics options (Grell et al., 2005). However, this coupled can demand high computational time to resolve physical and chemical processes. As mentioned by Grell et al. (2005) about the first version of WRF-Chem,

"the chemistry package consists of dry deposition, biogenic emission, the chemical mechanism from RADM2, a complex photolysis scheme, and a state of the art aerosol module (MADE/SORGAN aerosol parameterization)."

The details of the chemical aspects are covered in Zaveri and Peters (1999), Grell et al. (2005), Fast et al. (2006), and Gustafson Jr. et al. (2007). So, the selection of chemical mechanisms and parameterizations will depend on the study's aims (i.e., secondary aerosol or ozone formation, or both).

2.1.1 Model configuration

The WRF-Chem model (version 4.1.3) simulated three scenarios. A control period (September and October 2018) was denoted as "current" and two simulations for September and October 2030 for RCP 4.5 and RCP8.5 were named as "future."

Model configurations with two nested domains were used where the MASP is at the center of the domains (Figure 2.1), a 15-km resolution parent, and a 3-km resolution in a one-way concurrent run domain. This design was chosen because the spatial resolution of meteorology lateral boundary conditions (LBC) for future conditions has one degree, and it is necessary to reduce errors from the dynamical downscaling procedure. According to Warner (2011), there is evidence of errors in regional downscaling modeling due to meteorological LBC. Thus, the second domain is located far away from the boundary of the parent modeling domain borders.

As mentioned by Vara-Vela (2013), model configurations for 3-km resolution represented better the ozone formation than 1-km resolution. Therefore, only the 3-km resolution domain was analyzed in this study. Table 2.1 shows the WRF-Chem configuration for these two domains. The gas-phase chemical mechanism set up in the model is the CBM-Z (Zaveri and Peters, 1999) without Dimethylsulfide (DMS). This mechanism was used to simulate tropospheric ozone formation according to findings in Andrade et al. (2015); Gavidia-Calderón et al. (2018) for the MASP conditions. This chemical mechanism was chosen due to the use of ethanol as fuel consumption in the MASP.

For running the WRF-Chem model, we started with two days of first spin-up with five days of data analyzed. We continued the simulation with one day of spin-up, reinitializing meteorology with previously calculated chemistry. This means that IC is updated every five days of simulation, preserving the previous calculated chemical species concentrations (i.e., chemical IC is not used). This kind of run helps us to reduce meteorological IC error in our long time simulations.

For instance, as shown in Figure 2.2, two days (August 30-31) of spin-up were considered in the WRF-Chem model as a suggestion by Warner (2011). The term "spin-up" means generating a model solution of features not entered by the initial conditions such as inertia-gravity waves to reach approximate hydrostatic and geostrophic balance (Warner, 2011). For initialization (red cells), an additional 24 hours before the five days (blue cells) is needed to avoid the adverse effects due to newly meteorology initial conditions. This last consideration of the temporal modeling setup was based on suggestions by Ritter (2013), who used 12 hours before the 5-day steps.

Description	Option	WRF codes
Domain configuration ^(a)		
Points to the west-east direction	90, 151	e_we
Points to the south-north direction	60, 121	e_sn
Vertical levels ^(b)	32, 27	num_metgrid_levels
Geographical dataset	30 seg, 30 seg	
Grid spacing of first domain	15 km	dx = 15000, dy = 15000
Grid spacing of second domain	3 km	dx = 3000, dy = 3000
Map projection	Mercator	<pre>map_proj = 'mercator'</pre>
Center latitude	-23.57	ref_lat=-23.57
Center longitude	-46.61	ref_lon=-46.61
Physical parameterizations		
Long-wave radiation	RRTM	ra_lw_physics=1, 1
Short-wave radiation	RRMG	ra_sw_physics=4, 4
Boundary layer	BouLac	<pre>bl_pbl_physics=8, 8</pre>
Surface layer	Revised MM5 scheme	<pre>sf_sfclay_physics = 1, 1</pre>
Land-surface	Noah	<pre>sf_surface_physics=2, 2</pre>
Cumulus cloud	Grell 3D	cu_physics=5, 5
Cloud microphysics	Morrison double-moment scheme	mp_physics=10, 10
Urban surface	Urban canopy model	sf_urban_physics=1
Chemical options		
Chemical lateral	Idealized profile	chem_in_opt=0
	re-initialization	chem_in_opt=1
Gas-phase mechanism	CBMZ without DMS	chem_opt=6, 6
Photolysis scheme	Fast-J	phot_opt=2
Emissions	Two 12 h files	io_style_emissions=1
	$\operatorname{CBMZ}/\operatorname{MOSAIC}$ anthropogenic	emiss_opt=4, 4
	emissions	
	RADM2 speciation	<pre>emiss_inpt_opt=102, 102</pre>
	MEGAN2	bio_emiss_opt=3
		ne_area = 70
	No GOCART dust emissions	dust_opt=0

Table 2.1 WRF-Chem model configuration

^(a) This configuration uses information from the WRF-Chem Model User's Guide (Wang et al., 2019) and WRF-Chem version 3.9.1.1 User's Guide (NOAA et al., 2018).

^(b) NCEP-FNL (ds083.2) and NCAR CESM (ds316.1) have 32 and 27 metgrid levels, respectively.

The "namelist.wps" controls the WRF Preprocessor System (WPS) configuration while the "namelist.input" controls WRF configuration (Figure 2.3). Both namelists are shown in Appendix A.1. The WPS module prepares geography and meteorology data based on the modeling domain area. Geography data has limitations because it can not represent current features for years 2018 and 2030. However, the WPS used the high resolution available (30s) from MODIS and USGS data, downloaded from the WRF webpage. In the WRF module, the namelist.input requires chemical options that are located in &chem section, like the chemical mechanism, the biogenic emission scheme, or the chemical LBC. Different from a meteorological run, WRF-Chem requires emission files in NetCDF format, called 'wrfchemi'. The real.exe builds initial (wrfinput_<domain>) and lateral-boundary conditions (wrfbdy d01). Finally, wrf.exe uses IC/BC and the emission files to simulate



Figure 2.1 WRF-Chem simulation domain

Notes:

The MASP is at the center of the domain area (23.5°S and 46.6°W) in the São Paulo state. The height ground terrain was based on geo_em.d01.nc and geo_em.d02.nc, generated both by WPS Preprocessor.





Note:

Green cells are two days of spin-up; blue cells are data output analyzed for this study; red cells are one day of spin-up for meteorology initialization.


Figure 2.3 Flowchart for running the WRF-Chem model, applied to this study

Note:

Geogrid defines the model domain and creates static files of terrestrial data. Ungrib: decodes GRIB-formatted data only for global data for 2018. Metgrid: interpolates meteorological data to the model domain. Emission files are generated from LAPAt preprocessor emissions through the Bottom-up method for road transport (Sep. and Oct. 2018), adding industry and residential emissions from EDGAR-HTAP (2010) using ANTHRO_EMISS. LBC: Lateral-Boundary Conditions; IC: Initial Conditions. WRF is ran based on Real Data ARW System. hourly meteorological variables and pollutant concentrations in the defined spatial resolution. After running the first seven days (Figure 2.2), re-initialization uses wrf_chem_input_d01 and wrf_chem_input_d02 linked to the first hour of the previous day, using Linux codes, for instance:

ln -sf ../wrfout_reinit/wrfout_d01_2018-09-05_00:00:00 wrf_chem_input_d01
ln -sf ../wrfout_reinit/wrfout_d02_2018-09-05_00:00:00 wrf_chem_input_d02
Also, it is necessary edit options in the namelist.input:

io_form_auxinput12 = 2
auxinput12_inname = 'wrf_chem_input'
chem_in_opt = 1, 1

2.1.2 Initial and Lateral-Boundary Conditions (IC/BC)

Meteorology IC/BC

For the 'current' scenario (Sep-Oct 2018), the WRF-Chem model was initialized from the NCEP Final Analysis (available online at ds083.2), which has a temporal resolution of six hours and a horizontal resolution of 1×1 degree. According to NCEP et al. (2000), parameters include surface pressure, sea level pressure, geopotential height, temperature, sea surface temperature, soil values, ice cover, relative humidity, zonal (u) and meridional (v) winds, vertical motion, vorticity and ozone.

For the 'future' scenario (Sep-Oct 2030), the model was initialized from the NCAR's Community Earth System Model (CESM1) outputs with 1×1 degree resolution every six hours (available online at ds316.1). CESM1 participated in phase 5 of the Coupled Model Intercomparison Experiment (CMIP5) (Monaghan et al., 2014), which supported the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR-5). Two datasets for future scenarios based on RCP 4.5 and RCP 8.5 were used. These contain all the variables needed for the meteorological IC/BC for simulations with WRF-Chem, provided in the Intermediate File Format specific to WRF. This means that running "*ungrib*"¹ in WPS is not required. The variables have been biascorrected using the European Centre for Medium-Range Weather Forecast (ECMWF) Interim Reanalysis (ERA-Interim). The datasets have 26 pressure levels and the meteorological variables are air temperature, boundary-layer winds, geopotential height, humidity, sea ice concentration, sea level pressure, sea surface temperature, skin temperature, snow water equivalent, soil moisture/water content, soil temperature, surface pressure, surface winds, upper-air temperature, upper-level winds.

Chemical IC/BC

According to Gavidia-Calderón et al. (2018), dynamic chemical BC affects more ozone simulation with WRF-Chem in the autumn season (May) than in the spring season when photochemical activity is a higher. The same authors show in their results that the impact of dynamic chemical conditions could be neglected when the photochemical activity is higher (e.g., spring season). The study concludes that the dynamic chemical BC

¹Extract meteorological fields from GRIB files.

does not significantly impact near-surface ozone concentration due to ozone formation much depending on the vehicular emission. For that reason, this study considered chemical BC based on the default configuration.

According to Grell et al. (2005) and WRF-Chem User's Guide, default chemical IC/BC are based on northern hemispheric, mid-latitude, clean environment conditions with data based upon results from a NOAA-Aeronomy Laboratory Regional Oxidation Model (NALROM). For the initial simulation period with two days of spinup, the model was initialized with chem_in_opt = 0 in the 'namelist.input'. After that, the re-initialization only applied to meteorological IC/BC and considered the first hour of the previous day based on the WRF-Chem output, preserving the previous calculated chemical species concentrations, as shown in Figure 2.2. This procedure was used for 'current' and 'future' simulations.

2.2 Anthropogenic and biogenic emissions

Anthropogenic emissions used as input in the model represent three main sectors (i) road transport, (ii) industry, and (iii) residential. For the representation of the transport sector emission it was used the LAPAt pre-processor emissions model (Andrade et al., 2015) based on a bottom-up approach for emissions inventory. The bottom-up approach provides detailed emission rates with spatial and temporal variation, using local detailed emission factors and activity data (Wang et al., 2009).

For the industry and residential sectors, anthro_emiss tool (Kumar, 2020) was used to assimilate EDGAR-HTAP emissions to our two modeling domains. Biogenic emissions were calculated online using the Model of Emissions of Gases and Aerosols from Nature (MEGAN), through bio_emiss tool. Both tools (anthro_emiss and bio_emiss) are available on the NCAR UCAR web page. Appendix A.2 presents details about the data source used to generate emission files required by the WRF-Chem model.

2.2.1 Road transport emissions

LAPAt pre-processor emissions model, cited in Andrade et al. (2015), created the WRF-Chem emission files for road transport based on a bottom-up approach.

ID	Category	Fuel	$\mathbf{Veh}/\mathbf{year}$
VEIC 1	Passenger Car (PC)	Gasoline C	2686720
VEIC 2	Passenger Car (PC)	Ethanol	203893
VEIC 3	Passenger Car (PC)	Flex-fuel	7402653
VEIC 1	Light Commercial Vehicles (LCV)	Gasoline C	503522
VEIC 2	Light Commercial Vehicles (LCV)	Ethanol Hid.	19460
VEIC 3	Light Commercial Vehicles (LCV)	Flex-fuel	959325
VEIC 4A	Light Commercial Vehicles (LCV)	Diesel	427788
VEIC 4A	Light Truck (LT)	Diesel	32662
VEIC 4A	Semi Light Truck (SLT)	Diesel	107579
VEIC 4A	Medium Truck (MT)	Diesel	61148
VEIC 4A	Semi Heavy Truck (SHT)	Diesel	112008
VEIC 4A	Heavy Truck (HT)	Diesel	124368
		Continue 1	

Table 2.2 Vehicular fleet by type and fuel for São Paulo State (CETESB, 2019a)

Continued on next page

ID	Category	Fuel	$\mathbf{Veh}/\mathbf{year}$
VEIC 4B	Small Urban Bus (SUB)	Diesel	62351
VEIC 4B	Urban Bus (UB)	Diesel	15069
VEIC $4C$	Urban Bus Articulated (UBA)	Diesel	28243
VEIC 6A	Motorcycle (MC)	Gasoline C	1823493
VEIC 6B	Motorcycle (MC)	Flex-fuel	696079

Table 2.2 Vehicular fleet by type and fuel for São Paulo State (CETESB, 2019a)

After running test simulations using different emission rates, we corrected only the emission rates for the road transport sector through a correction factor to achieve simulations closer to performance benchmarks suggested by Emery et al. (2017). Road transport emission approximation through the LAPAt pre-processor model required the following information:

- *Road length*: Two files for each modeling domain (grid3km_d02.txt and grid15km_d01.txt) represent the sum of motorway, trunk, primary, secondary, and tertiary types of roads inside each grid cell.
- Fraction fleet by type and vehicle number: Table 2.2 shows fraction fleet by type was obtained from the emission report published by CETESB (2019*a*). It is mainly based of fuel type and light or heavy duty vehicles. The vehicle number for September and October 2018 was obtained from DENATRAN webpage. There is a specific information about monthly vehicle numbers by type, for each municipality and state. Two estimations of total vehicle fleet were obtained for the first (56 410 975 vehicles in Sep. 2018 and 56 604 008 vehicles in Oct. 2018) and second (24 259 504 vehicles in Sep. 2018 and 24 332 719 vehicles in Oct.) domain.
- Emission factors by vehicle and fuel types were used as shown in Table 2.3 in units of g km⁻¹. Different vehicle types (motorcycles, light-duty vehicles, and heavy-duty vehicles) and fuel types (gasohol, ethanol, ethanol-blended gasohol, and diesel) release different pollutant emission rates (CO, NO_x, particles, SO_x, VOC).
- Use intensity mean: Kilometer traveled by vehicle type (km day⁻¹), shown in Table 2.4, was calculated based on CETESB information.
- Emission correction factor: To obtain the best simulation for the current scenario (September 2018), 11 experiments were developed running 5 days with no precipitation (September 6 12 and 24 28, 2018). The experiment 10 was chosen because they reached two of three (correlation coefficient and normalized mean bias) statistical benchmarks suggested by Emery et al. (2017) for surface ozone. The experiment 10 only required a correction factor for road transport emission (fc_nox = 0.8) indicated in the LAPAt pre-processor model.
- *Temporal distribution*: Normalized average vehicles counts by light-duty (LDV) and heavy-duty (HDV) were done from the tunnel experiments as is mentioned in Andrade et al. (2015), shown in Figure 2.4.
- **Relative fractions of VOC emissions**: This information was shown in Andrade et al. (2015) related to VOC fractions for exhaust, evaporative, and liquid emissions from gasohol, ethanol, and diesel. The VOC

Pollutants ^a	Light-dut	ty vehicles	b	Heavy-duty	y vehicles ^c	Taxis ^d Motorcycles ^e			
	VEIC 1	VEIC 2	VEIC 3	VEIC 4A	VEIC 4B	VEIC 4C	VEIC 5	VEIC 6A	VEIC 6B
exa CO	4.4000	12.000	4.4000	1.5000	1.5000	1.5000	0.0000	9.1500	9.0200
$exa CO_2$	219.00	219.00	219.00	1422.0	1422.0	1422.0	0.0000	0.0000	0.0000
exa NO_x	0.2000	1.1200	0.2000	6.9000	6.9000	6.9000	0.0000	0.1320	0.1290
$exa SO_2$	0.0290	0.0140	0.0210	0.6100	0.6100	0.6100	0.0000	0.0097	0.0093
exa C_2H_5OH	0.5080	0.2500	0.5080	0.6100	0.6100	0.6100	0.0000	0.0790	0.3050
exa HCHO	0.0089	0.0110	0.0098	0.6100	0.6100	0.6100	0.0000	0.0152	0.0155
exa Aldehyde	0.0140	0.0300	0.0220	0.6100	0.6100	0.6100	0.0000	0.0164	0.0188
exa PM	0.0200	0.0200	0.0200	0.2770	0.2770	0.2770	0.0000	0.0500	0.0500
exa VOC	0.4250	1.3000	0.4340	2.0500	2.0500	2.0500	0.0000	1.0800	1.0800
vap VOC	0.2300	0.2500	0.2400	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
liq VOC	2.0000	1.5000	1.7500	0.0000	0.0000	0.0000	0.0000	1.2000	1.2000

Table 2.3 Emissions factors (g/km) based on Pérez-Martínez et al. (2014) and Andrade et al. (2015, 2019)

^a Exhaust (exa), evaporative (eva), and liquid (liq).
 ^b VEIC 1 (gasoline), VEIC 2 (ethanol), VEIC 3 (flex-fuel).
 ^c Diesel as fuel: VEIC 4A (semi light, medium, and semi heavy trucks), VEIC 4B (small and urban bus), VEIC 4C (urban bus articulated).

 $^{\rm d}\,$ Tax is use natural gas as fuel.

^e VEIC 6A (motorcycle, gasoline), VEIC 6B (motorcycle, flex-fuel).

Vehicle type	Fuel	Use (km/day)
	Gasohol	
Light-duty vehicles (VEIC 123)	Ethanol	39.39
	Flex-fuel	
Truck (VEIC 4A)	Diesel	91.02
City buses (VEIC 4B)	Diesel	86.70
Intercity busses (VEIC 4C)	Diesel	106.18
Taxis (VEIC 5)	Natural gas	0
Motorcycle (VEIC 6A)	Gasohol	32.16
Motorcycle (VEIC 6B)	Flex-fuel	35.78

Table 2.4 Intensity of use by vehicle type and fuel

Based on available data provided in emission report of CETESB (2019a).



Figure 2.4 Temporal distribution of emissions for light-duty (LDV, including motorcycles) and heavy-duty (HDV) vehicles, according to Andrade et al. (2015).

speciation was made considering measurements performed in tunnels and with dynamometers (Pérez-Martínez et al., 2014; Nogueira et al., 2014).

2.2.2 Industry and residential emissions

Global EDGAR-HTAP (Janssens-Maenhout et al., 2015; Crippa et al., 2020) emission inventory has data for the year 2010 and by sector (i.e., industry and residential). This dataset provides monthly grid maps with a spatial resolution of $0.1^{\circ} \ge 0.1^{\circ}$ (latitude by longitude). The utility *anthro_emiss*, Fortran based preprocessor, creates WRF-Chem anthropogenic emission files from global inventories on a local coordinates. This utility provides industry and residential emissions files, respectively:

- wrfchemi_{00z,12z}_<domain>_ind
- wrfchemi_{00z,12z}_<domain>_res.

EDGAR-HTAP doesn't have ground transport emissions with highly temporal and spatial resolution. Instead, LAPAt preprocessor emission model outputs were used.

2.2.3 Biogenic emissions

MEGAN version 2 (Guenther et al., 2006) through bio_emiss preprocessor is an essential tool to estimate natural emissions based on driving variables such as ambient temperature, solar radiation, leaf area index, and plant functional type. This tool was setup for September and October conditions. After running MEGAN v2 model, two biogenic emission files were created, called as wrfbiochemi_d01 and wrfbiochemi_d02. The Appendix A.2.3 shows useful information to run the MEGAN v2 model, applied for this study.

2.3 Surface air quality and meteorological observations

The São Paulo state Environmental Agency (CETESB) is responsible for the automatic air quality stations network (approximately 61 stations), according to CETESB (2019b). The air quality stations inside and outside the MASP are located at different altitudes so the modeling domain can be considered as complex terrain.

As shown in Figure 2.5, stations were classified by their surroundings. The station types inside the MASP were classified as "Forest Preservation", "Urban", and "Urban Park", stations outside the MASP were called as "Regional Urban", and "Industry", as shown in Table B.2. Each station was visualized in Google Earth Pro for classification. For instance, Figure 2.6 shows satellite images for each station type: Forest preservation (Pico do Jaraguá), Urban (São Caetano do Sul), Urban park (Capão Redondo), Regional urban (Jaú), and Industry (Santa Gertrudes). The classification used in this work is different from CETESB classification based only on urban and industrial characteristics.

For this study, we used two downloaded datasets from CETESB: (i) five years (2014-2018) of hourly concentrations for ten stations (Table 2.5), and (ii) two months (September and October 2018) of hourly concentrations for all stations inside the second modeling domain. Appendix B shows utilities used to download hourly of meteorological and air quality data.

Type	Station
Forest preservation	Pico do Jaraguá
Industry	Paulínia
Regional urban	Campinas-Taquaral
Regional urban	Sorocaba
Urban	Interlagos
Urban	Carapicuíba
Urban	Parque D.Pedro II
Urban	Pinheiros
Urban park	Ibirapuera
Urban park	Itaquera

Table 2.5 CETESB stations considered to monthly analysis based on five years (2014-2018)

2.3.1 Air quality data

The period from 2014 to 2018 was analyzed, as shown in Figures B.1 and B.3. September and October presented higher hourly monthly mean surface ozone concentrations in many station types (e.g., Industry, Regional urban, Urban), as shown in Figure B.5. September presented more frequently higher surfer ozone concentrations than the other months and, for that reason, it can be considered the worst-case scenario. The maximum ozone concentration was reached at 15 hours (local time) in many station types (e.g., Forest preservation, Industry, Regional urban, and Urban). After 15 hours, during the evening, mean ozone concentrations decrease more than those recorded for December and January, as shown in Figure B.5.

Regarding the interactions between surface ozone with other pollutants (CO, NO_x), Figure 2.7 shows charts by station type as hourly mean concentration based on data collected from September and October, 2018. The surface ozone reached its maximum levels between 14:00-15:00 hours (local time). CO and NO_x (NO and NO₂) highest concentrations occurred at early morning (7:00-10:00 h) and evening (18:00-20:00 h) hours, positively correlated with the temporal distribution of emissions for light-duty and heavy-duty vehicles (Figure 2.4). Reduction of these pollutant concentrations between 11:00-19:00 hours is related to photochemical reactions that enhance ozone formation. Nocturnal ozone maximum concentration occurred around three and six hours (local time). This behavior was mentioned by Carvalho et al. (2015), and can be associated with atmospheric transport from vertical levels, as it is mentioned in Mazzoli da Rocha (2013) and Andrade et al. (2017).

As mentioned in CETESB (2019b), few stations monitor some hydrocarbons such as Benzene and Toluene (Figure B.2). The hourly analysis shows a reduction of these VOC during the photochemical activity as shown in Figure 2.8.

2.3.2 Meteorological data

CETESB stations register hourly values for meteorological parameters such as surface temperature, relative humidity, wind speed and direction. However, these stations do not register hourly precipitation. Complementary, hourly data registered from the IAG climatological station in Água Funda was used for this study to compare with WRF-Chem model results. This data was requested on the IAG/USP station website:

http://www.estacao.iag.usp.br/sol_dados.php.

The data was received as Excel files (hourly in rows and day in columns) and ordered as time series in rows and

26

meteorological parameters in columns. Figure B.6 in Appendix shows hourly time variation for September and October 2018 for eight meteorological parameters. Figure B.7 in Appendix shows total daily rain and mean cloud cover; which in September (107 mm) was less than October (152 mm). According to Carvalho et al. (2015), the weather condition in spring enhances ozone formation and is related to low cloud cover. September presented less values of cloud cover favoring the ozone formation due to photochemical activity.

2.4 Model performance evaluation

As recommended by Seinfeld and Pandis (2016), three different performance index can be used to analyze the urban ozone models in the ability to reproduce the peak ozone concentrations, described below:

- Analysis of predictions and observations paired in space and time.
- Comparison of predicted and observed maximum concentrations.
- Comparisons paired in space but not in time.

Other recommendations are to use charts, plots, and statistical metrics for evaluating model results, letting us understand the model performance and its behavior, as suggested by Emery et al. (2017). There are recommendations about statistics and benchmarks to assess photochemical model performance. Emery et al. (2017) analyzed statistical benchmarks applied to North American. For ozone, they recommended calculating statistics over temporal scales of one week (an episode), not longer than one month, and compared with recommended goals² and criteria³ benchmarks shown in Table 2.6. Although these statistical benchmarks were based on North American modeling studies, Emery et al. (2017) recommendations are relevant to any such applications of state-of-science photochemical models. Therefore, they can be used for Brazil. In case one or two did not comply with the recommended criteria for Normalized Mean Bias (NMB), Normalized Mean Error (NME), or Pearson correlation coefficient (r), the modeling application should not be considered "failure" (Emery et al., 2017). The same authors recommend to restrict periods when observations exceed a minimum threshold value ("cutoff"), particularly for ozone. As conclusion in the paper published by Emery et al. (2017), they recommend:

"applying a cutoff value of 40 ppb (observations) when calculating only NMB and NME for 1 hour ozone, for several physical and regulatory reasons beyond the need for statistical stability, but not for correlation as it is best characterized over the entire concentration distribution. The choice of 40 ppb is not absolute and should consider the chemical climatology of the region being modeled."

On the other hand, even when the primary aim of this study is to understand ozone formation, it is also essential to consider meteorological conditions. Why do we need to evaluate meteorological results? Emery et al. (2017) and Monk et al. (2019) suggest that it is essential due to its crucial role in the air quality simulations (e.g., ozone) since weather conditions are also a significant air quality driver.

²"The more restrictive goals around the 33rd percentiles indicate statistical values that about one-third of top performing past applications in many U.S. modeling studies have met, and should be viewed as the best a model can be expected to achieve" (defined in Emery et al., 2017).

 $^{^{3}}$ "The less restrictive criteria around the 67th percentile indicate statistical values that about two-thirds of past applications have met, and should viewed as establishing historical context that a majority of models have achieved" (defined in Emery et al., 2017).

Statistical metric	Goal	Criteria
NMB	$<\pm5~\%$	$<\pm15~\%$
NME	${<}15~\%$	< 25~%
r	> 0.75~%	> 0.5~%

Table 2.6 Statistical benchmarks for surface ozone (1-hr or MDA8) suggested by Emery et al. (2017)

Notes. Normalized Mean Bias (NMB), Normalized Mean Error (NME), correlation coefficient (r).

For this study, model results for September and October 2018 were compared with hourly measured data at each station. This evaluation let us to compare the analysis of modeled and observed values in space and time, suggested by Seinfeld and Pandis (2016). Previously, units concentrations from model results (excluding CO) were converted from mixing ratio ppm(v) to $\mu g m^{-3}$, using the following equation (Seinfeld and Pandis, 2016):

Concentration
$$[\mu g/m^3] = \frac{p \ M_i}{8.314 \ T} \times \zeta_i \ [ppm]$$
 (2.1)

Where p is the atmospheric pressure in Pa (N m⁻²), M_i is a molecular mass (g mol⁻¹), T is in Kelvin, and ζ_i is the mixing ratio in ppm, considering a molecular gas constant (R) equals to 8.1314 J K⁻¹ mol⁻¹ or Pa m⁻³ K⁻¹ mol⁻¹. Meteorological simulations of temperature and atmospheric pressure were used to convert concentration from ppm to μ g m⁻³.

The analysis of surface ozone was performed in two forms: 1-hr time series and the maximum daily 8-hr rolling mean (MDA8). Thus, the comparison of predicted and observed maximum concentrations is made based on MDA8 values for ozone concentrations through time plots by station types. Other pollutants as nitrogen oxides (NO, NO₂), CO, and toluene were also analyzed in time and space. We analyzed toluene only for some stations where it was measured during September and October 2018 period. It is important to remark that the WRF-Chem model did not include benzene in the results. Therefore, it is excluded from the statistical analysis and the comparison with observations.

Regarding meteorological evaluation, São Paulo state can be considered as a complex terrain due to areas with high altitudes (e.g., Pico do Jaraguá) and high buildings in the MASP. For this reason, meteorology model results were evaluated using specific statistical benchmarks suggested by Monk et al. (2019) for complex terrain (Table 2.7).

More details, such as utilities, can be found in Appendix C. Statistic equations are described as follows based on the description by Emery et al. (2017):

Parameter	Statistical metric	Complex terrain
Temperature [K]	MAGE	≤ 3
	MB	$\leq \pm 1$
	IOA	≥ 0.8
Relative humidity	MAGE	< 20~%
	MB	$<\pm10~\%$
	IOA	≥ 0.6
Wind speed $[ms^{-1}]$	RMSE	≤ 2.5
	MB	$\leq \pm 1.5$
	IOA	≥ 0.6
Wind direction [degree]	MAGE	≤ 55
	MB	$\leq \pm 10$

Table 2.7 Statistical benchmarks for meteorology parameters suggested by Monk et al. (2019)

Notes. Mean Absolute Gross Error (MAGE), Mean Bian (MB), index of agreement (IOA), Root Mean Square Error (RMSE).

$$MAGE = \frac{1}{N} \sum |P_j - O_j| \tag{2.2}$$

$$MB = \frac{1}{N} \sum (P_j - O_j) \tag{2.3}$$

$$IOA = 1 - \frac{\sum (P_j - O_j)^2}{\sum (|P_j + \overline{O}| + |O_j + \overline{O}|)^2}$$
(2.4)

$$NMB = \frac{\sum (P_j - O_j)}{\sum O_j} \times 100 \tag{2.5}$$

$$NME = \frac{\sum |P_j - O_j|}{\sum O_j} \times 100 \tag{2.6}$$

$$r = \frac{\sum [(P_j - \overline{P}) \times (O_j - \overline{O})]}{\sqrt{\sum (P_j - \overline{P})^2 \times \sum (O_j - \overline{O})^2}}$$
(2.7)

2.4.1 Hypothesis test for Pearson's r

The correlation coefficient (r) gives us information about how two datasets have a linear relationship or not. A positive perfect correlation (+1) corresponds to all the pairs lying on a straight line with a positive slope when a scatter diagram is used. It is desired to achieve in air quality modeling (a perfect model!). However, a few paired values could generate errors for the correlation coefficient, which could not be robust or significant. Significance and confidence interval can be calculated for correlation coefficient based on t-test (two-tailed). Based on Navidi et al. (2019) t-test statistical evaluation methodology, we defined Null (H_o) and Alternative hypothesis (H_A), considering an alpha (α) equals to 0.05 for degree freedom (df) equals to n-2, where n is the number of paired values:

- $H_o: \rho = 0$ (samplings of x (obs) and y (mod) are not correlated)
- $H_A: \rho \neq 0$

Where ρ denotes the population correlation between x and y. Then, assuming that H_o is true, a statistical test is computed and used to asses the strength of the evidence against H_o. Also, we compute the P-value of the test that is the probability and is also called the **observed significance level**. The P-value measures the

plausibility of H_o . So, if the P-value is sufficiently small, we may be willing to abandon the assumption that H_o is true and believe H_A instead (rejecting the null hypothesis). As a practical rule, if the t critical value (calculated from $\alpha=0.05$ and n-2 degree freedom) is less than the t-statistic, we will reject the null hypothesis and accept the H_A . Therefore, we can state "there is enough evidence to conclude that there is a significant linear relationship between x (obs) and y (mod) because the correlation coefficient is significantly different from zero." The formula for the t-statistic value calculation is (Navidi et al., 2019):

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$
(2.8)

Where r is the sample correlation of the n points. The t-statistic value has the same sign as the correlation coefficient.



 $Figure \ 2.5$ Air quality and meteorological stations network

Note:

The MASP map is in gray color. All stations except IAG belong to CETESB.





Urban



Regional urban





Figure 2.6 Station classification type based on location in the São Paulo State. Maps Data: Google ©2020, Maxar Technologies.

Urban park



Figure 2.7 Hourly mean concentrations for surface ozone (O_3) , nitrogen monoxide (NO), nitrogen dioxide (NO_2) , and carbon monoxide (CO) by station type in the MASP.

Note. Shaded area corresponds to the standard deviation.



Figure 2.8 Hourly mean concentration by month (Sep-Oct, 2018), based on measurements in Pinheiros and S. André-Capuava stations.

Note: Toluene (Tol). Benzene (Ben). Shaded area corresponds to ozone concentration standard deviation.

Chapter 3

Results and Discussion

Assure me that I yet may change these shadows you have shown me.

Charles Dickens, A Christmas Carol

This chapter presents the emission inventory results for September and October of 2018 and WRF-Chem model results for the second modeling domain ($\delta X = 3$ km of spatial resolution) for all scenarios. Two scenarios (RCP 4.5 and RCP 8.5 for 2030) are future predictions oddly reminiscent of Dickens.

In Section 3.1, this study evaluates the WRF-Chem model results for the current conditions (September and October 2018 period) based on the meteorological IC/BC, emission for anthropogenic (road transport, industry, and residential), and biogenic sources.

Statistical evaluation for surface ozone simulations covers detailed analyses for the global data, station types, and station locations. Further analyses show variations for some ozone precursors (NO_x, CO, and toluene) by hour per day as an average for the all analyzed period (i.e., Pinheiros station).

In Section 3.2, future meteorological conditions and surface ozone concentrations from the WRF-Chem model are presented. These results are based on the same emission rates used to evaluate the model for current conditions, with the change for meteorological IC/BC provided from the CESM1 datasets: RCP 4.5 and RCP 8.5 scenarios; representing a set of possible humankind future pathways based on GHG emissions, air pollutants concentrations, and land use/land cover changes.

3.1 Evaluation results for current conditions

3.1.1 Anthropogenic emissions

This section shows the spatial and temporal anthropogenic emissions distribution. The results are based on emissions inventory approaches from the bottom-up methodology for the road transport sector (Andrade et al., 2015) and from the EDGAR-HTAP datasets (Janssens-Maenhout et al., 2015) interpolation in space and time using the *anthro* emiss emission pre-processor (Kumar, 2020) for the industrial and residential sectors.

We tested 11 experiments, running the model for seven¹ non-rainy days based on different emission files 1 Two first days are spin-up, not included in the analysis.

(the base emission without correction factor and some pollutants with a correction factor, such as nitrogen oxides). Considering Emery et al. (2017) benchmarks, experiment 10 with a correction factor of 0.8 for NO_x road transport emission files represented better the ozone formation in the MASP for September and October 2018. However, these tests showed the difficulty of improving other pollutants' simulations (NO_x , CO, and toluene) compared with observations. The emission files are the primary error source for simulations and require to be corrected as many works have been reported for São Paulo and other cities outside of Brazil (Russell and Dennis, 2000; Holnicki and Nahorski, 2015; Andrade et al., 2017; Ibarra-Espinosa et al., 2020). Vehicular emissions have limitations on the spatial and temporal distribution (e.g., roads only for heavy-duty vehicles), even when considering a bottom-up approach, as it was reported by Andrade et al. (2015).

Figure 3.1 shows the spatial distribution of NO emissions for the second modeling domain. Other species (e.g., NO₂, CO, SO₂) have the same emission spatial distribution. Based on CBM-Z chemical mechanism, 21 species of emission rates were considered into the WRF-Chem emission files (i.e., wrfchemi_{00z, 12z}_{d01, d02}). Figure 3.2 shows emission rates and the temporal distribution of Industry, Residential, and Road Transport sectors. Only road transport emissions present a temporal distribution based on information available in Andrade et al. (2015) (Figure 2.4). Table 3.1 shows emission rates in units of kilo-tonnes per year (kt/year) accumulated in the second domain, based on the calculated emissions for September 2018.



Figure 3.1 NO emissions for second modeling domain in SE Brazil

Note:

Only Total emission files were used as input to the WRF-Chem model. Anthropogenic emissions were summed using a Python script. LT= Local Time. X and Y axis labels are in local coordinates.



Figure 3.2 Emission rates by species (part a) and temporal distribution (part b) for all days for NO emission rates and by sector

	ID	Road	Industrial	Residential
Species				
Carbon monoxide	СО	1412.29	510.65	191.80
Nitrogen oxide	NO	244.23	134.21	7.59
Nitrogen dioxide	NO_2	27.14	22.87	1.29
Nitrogen oxides	NO_x	271.37	157.08	8.88
Sulfur dioxide	SO_2	30.74	133.09	12.92
Ammonia	NH_3	4.71	21.41	0.13
Isoprene	ISO	0.52	0.00	0.00
Ethane	ETH	8.31	6.53	0.73
Propane	HC3	52.41	39.19	4.39
Alkanes (k _{OH} $0.5 - 1$)	HC5	79.96	45.73	5.12
Alkanes (k_{OH} 1 - 2)	HC8	126.25	45.73	5.12
Xylenes	XYL	55.87	19.60	2.19
Alkenes (internal)	OL2	41.90	52.26	5.85
Alkenes (terminal)	OLT	75.85	45.73	5.12
Alkenes (primary)	OLI	48.41	26.13	2.92
Toluene	TOL	80.50	26.13	2.92
Formaldehyde	HCHO	28.17	32.66	3.66
Aldehydes	ALD	38.17	26.13	2.92
Ketones	KET	0.45	0.00	0.00
Methanol	CH_3OH	0.56	0.00	0.00
Ethanol	C_2H_5OH	396.17	287.42	32.17

Table 3.1 Emission rates (kt/year) for the second modeling domain based on activity data for the September 2018 period

3.1.2 Meteorological model evaluation

WRF-Chem model results were compared with measurements from CETESB stations network and the IAG/USP climatological station located in Água Funda. The meteorological parameters extracted from the model for the statistical evaluation were 2-m temperature [°C], 2-m relative humidity [%], accumulated rain (processed from rainc and rainnc parameters) [mm], wind speed [ms⁻¹] and direction [degree] at 10 m above the ground.

Model simulations at IAG/USP location were analyzed. This station complies with the World Meteorological Organization standards to locate a weather station (WMO, 2018). Furthermore, this station recorded the hourly rainfall rate, a parameter not recorded by the CETESB stations network.

IAG/USP climatological station

Figure 3.3 shows a comparison between model simulations and observed values for each meteorological parameter, such as temperature, relative humidity, rain rate, wind speed, and wind direction. Table 3.2 shows statistical model evaluation results and t-test values for correlation coefficient for the IAG/USP station.

Temperature simulations comply with MAGE and IOA statistical benchmarks suggested by Monk et al. (2019) for complex terrain. However, MB result is not less than 1 °C. There is an overestimation of +1.56 °C on average compared with observations.

Relative humidity simulations comply with the three statistical benchmarks (MAGE and MB units are in %, apply for the MASP). Thus, model results presented high accuracy to simulate this meteorological parameter.

Wind speed model results comply with the RMSE statistical benchmark, but not for the others (MB and IOA). The model overestimated in ± 1.69 m/s the average wind speed observations. Wind direction model

Statistic ^(a) 2-n	n Temp. (°C)	2-m RH (%)	Rain rate (mm)	W. Speed (m s ⁻¹)	W. Dir. (°)
n	1461.00	1461.00	1461.00	1461.00	1461.00
MB	1.56	-9.00	0.13	1.69	5.02
MAGE	1.97	11.15	0.42	1.84	32.57
RMSE	2.66	14.66	1.72	2.22	-
IOA	0.89	0.78	0.23	0.43	-
r	0.85	0.71	0.11	0.37	-
Mm	20.32	74.30	0.30	3.37	-
Om	18.76	83.30	0.18	1.69	-
Msd	3.96	15.68	1.20	1.50	-
Osd	3.89	14.46	1.35	0.91	-
t-stat	61.63	38.51	4.23	15.21	-
t-crit	1.96	1.96	1.96	1.96	-

Table 3.2 Statistical results for meteorological parameters for Sep-Oct 2018 (IAG/USP station)

^(a) MB = Mean bias, MAGE = Mean Absolute Gross Error, RMSE = Root Mean Square Error, Mm = Mean of modeled values, Om = Mean of observed values, Msd = Standard deviation of modeled values, and Osd = Standard deviation of observed values. Units depend on the meteorological parameter. Correlation coefficient (r) is in dimensionless units. Statistical parameters are t-test statistical (t-stat) and t critical (t-crit).

results comply with statistical benchmarks for MAGE (value less than 55) and MB (value less than 10). Thus, the model has a good capability to simulate wind directions at IAG/USP station. Figure 3.4 shows the two wind rose plots, one with the modeled values and the other for the observation values. The frequency of calm winds is greater in observations than in the model results.

Statistical results for rain simulations presented a low performance, so the model was not accurate to represent rain observations recorded at the IAG/USP station. However, the hypothesis t-test suggests a significant linear relationship between modeled and observed values based on t-critical being less than the t-statistic value (Table 3.2). Despite a low correlation, the rainfall simulation values have significant statistical indexes with the observations. Figure 3.5 shows comparisons between modeled and observed values for total daily rain with reasonable accuracy for September 14 and October 24, 2018.

CETESB and IAG/USP grouped by station types

Statistical results considered all stations with non missing values for temperature, relative humidity, wind speed, and wind direction. Table 3.3 shows global statistical results for each meteorological parameter.

	n	MB	MAGE	RMSE	IOA	r	Mm	Om	Msd	Osd
2-m Temp. ($^{\circ}C$)	38727	1.30	1.94	2.54	0.92	0.89	21.99	20.69	4.36	4.74
2-m RH (%)	37298	-7.32	11.18	14.46	0.84	0.76	68.87	76.19	17.71	18.49
W. Speed (m s^{-1})	45412	1.38	1.71	2.13	0.52	0.39	3.36	1.98	1.70	1.12
W. Dir. (°)	43189	-21.66	50.58	-	-	-	-	-	-	-

Table 3.3 Statistical results for meteorological parameters for Sep-Oct 2018 (all stations)

Temperature simulations comply with two statistical benchmark (MAGE and IOA). The model overestimated temperature results with +1.30 °C based on modeled (Mm) and observed (Om) mean values, also shown as MB result in Table 3.3.

In general terms, the model has a good performance for relative humidity because simulations comply with



Figure 3.3 Comparison between observed and modeled values for meteorological parameters for the IAG/USP station during the period Sep-Oct 2018

all the benchmarks (MAGE < 20%, -10 < MB < +10%, and IOA ≥ 0.6). However, simulations underestimated measurements as we can see in the negative MB value.

Model simulations for wind speed comply with two (RMSE and MB) of three statistical benchmarks. The IOA value does not reach the statistical benchmark (≥ 0.6). Regarding wind direction, only one statistical value (MAGE) complies with the benchmark (≤ 55). The MB (as absolute value) for wind direction is greater than the statistical benchmark (≤ 10).

Further specific analysis by station type is shown in Appendix Table D.1, and summarized in Table 3.4, which shows model performance results for each meteorological parameter and station type. The WRF-Chem model has a good performance for the relative humidity simulations. Figure D.1 in Appendix shows time series as a daily mean for relative humidity, where model results and observations are compared by station type. Temperature model results comply with two of three statistical benchmarks. Figure D.2 in Appendix shows temperature values as a daily mean for model results compared with observations by each station type. Only model results for "Industry" station type presented good performance.

Finally, the wind was difficult to simulate and did not present good results compared with observed values.



Figure 3.4 Wind rose plot for Sep-Oct 2018 period at the IAG/USP station, for modeled and observed values. Wind speed simulations less than 0.5 m/s were replaced by 0 m/s to compare with observations.



Figure 3.5 Comparison between observed and modeled values for total daily rain rate during Sep-Oct 2018 for the IAG/USP station

Statistical results suggest low capability of the model to simulate wind speed in urban and forest areas (i.e., Pico do Jaraguá station). However, this is not particularly surprising given that CETESB stations aim to measure air quality parameters, and many stations could not comply with WMO recommendations to install a weather station (WMO, 2018). Furthermore, there are errors in the model due to low resolution to represent the topography, such as hills in the higher terrain (i.e., Pico do Jaraguá), and by the low resolution (one degree)

	Urban	U. park	R. urban	Ind.	F. pre.	
2-m Temp. (°C) (3 benchmarks)	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark \checkmark \checkmark$	$\checkmark\checkmark$	11
2-m RH (3 benchmarks)	$\checkmark \checkmark \checkmark$	$\checkmark\checkmark$	$\checkmark \checkmark \checkmark$	$\checkmark \checkmark \checkmark$	$\checkmark \checkmark \checkmark$	14
W. Speed (m s^{-1}) (3 benchmarks)	$\checkmark\checkmark$	\checkmark	\checkmark	\checkmark		6
W. Dir. (°) (2 benchmarks)	\checkmark	$\checkmark\checkmark$		\checkmark	\checkmark	5
	8	7	7	8	6	

Table 3.4 Summary of compliance of statistical benchmarks

of the meteorological global data used in the WRF-Chem as IC/BC. Figure D.3 in Appendix shows wind speed values as a daily mean for model results and observations; modeling values belonging to "Forest preservation sites" highly overestimated observed values more than other station types.

3.1.3 Air quality model evaluation

This section shows the model output for current conditions (Sep. and Oct. 2018), their evaluation through statistical analysis, and the comparison between model simulations and observations from CETESB air quality stations network. The parameters extracted from the model for the statistical evaluation were nitrogen monoxide (NO in $\mu g m^{-3}$), nitrogen dioxide (NO₂ in $\mu g m^{-3}$), carbon monoxide (CO in ppm), toluene ($\mu g m^{-3}$), and surface ozone (O₃ in $\mu g m^{-3}$). Ozone 8-hr rolling mean was also calculated from hourly time series for model simulated and observed values. Some stations located close to a coastal zone were not included as part of statistical evaluation due to low performance of WRF-Chem model, which are:

- Santos
- Santos-Ponta da Praia
- Cubatão-Centro
- Cubatão-Vale do Mogi
- Cubatão-V. Parisi

Furthermore, hourly model simulations for September 14-15 (2018) were not considered in the statistical analysis due to cloud and rainfall conditions did not represent appropriately by the WRF-Chem model. Figure B.8 in Appendix shows mean cloud cover and total rain by day. Thus, 57 stations with hourly observations were compared with WRF-Chem model results. Also, hourly simulations were compared against to observations by station type (i.e., Forest Preservation, Urban, Urban Park, Regional Urban, and Industry). As highlighted in Figure 2.5, the station types inside the MASP correspond to Forest Preservation, Urban, and Urban Park.

Global statistical results in Table 3.5 show correlation coefficients (r) for ozone above 0.50 that comply with the criteria level as statistical benchmark suggested by Emery et al. (2017). Average values of NMB and NME for ozone comply with the goal ($\leq \pm 5$ % for NMB) and criteria (≤ 25 % for NME) benchmarks. Based on MB results, positive values are indicators that simulations have overestimated the observations. Otherwise, negative values indicate underestimations.

Primary pollutants (NO_x and CO) have low values of correlation coefficients. On average, simulations overestimated the observations for NO₂ concentrations due to positive values for MB (+4.09 μ g m⁻³ for Sep.

Month	Sep. 2018					Oct. 2018				
Parameter	O_3	NO	NO_2	NO_x	CO	O_3	NO	NO_2	NO_x	CO
$\rm Statistic^{(a)}$										
n	23921.00	21693.00	21693.00	21693.00	9837.00	26034.00	24099.00	24099.00	24099.00	12152.00
MB	9.32	-0.55	4.09	3.54	-0.26	13.15	-4.10	1.39	-2.71	-0.31
MAGE	24.01	10.50	18.33	27.75	0.29	22.51	8.73	15.44	22.99	0.33
RMSE	30.07	23.78	25.77	43.63	0.40	29.43	21.89	21.79	37.44	0.44
$\rm NMB^{(b)}$	19.24	-6.54	15.56	10.20	-52.28	31.65	-47.84	5.92	-8.46	-61.76
NME ^(b)	49.57	124.67	69.75	79.99	59.26	54.17	101.81	65.74	71.70	64.64
NMB ^(c)	2.24	-	-	-	-	1.93	-	-	-	-
NME ^(c)	21.66	-	-	-	-	20.80	-	-	-	-
IOA	0.80	0.46	0.62	0.58	0.45	0.76	0.31	0.65	0.56	0.44
r	0.67	0.25	0.39	0.34	0.19	0.64	0.16	0.42	0.34	0.17
Mm	57.75	7.87	30.36	38.23	0.24	54.71	4.47	24.87	29.35	0.19
Om	48.43	8.42	26.27	34.70	0.49	41.56	8.58	23.48	32.06	0.51
Msd	37.75	17.50	25.47	38.68	0.12	33.27	8.57	20.81	26.71	0.08
Osd	31.36	20.97	19.99	36.90	0.31	27.49	21.14	19.47	36.83	0.32

Table 3.5 Global statistical results for air quality parameters

^(a) MB, MAGE, RMSE, Mm, Om, Msd, and Osd values are in μ g m⁻³. Correlation coefficient is in dimensionless units.

^(b) No cutoff was applied.

^(c) A cutoff value of 80 μ g m⁻³ was applied to calculate NMB and NME only for 1-hr ozone, suggested by Emery et al. (2017). Units are in percentage. Values in bold blue comply with the Goal benchmark and in only blue with the Criteria benchmark for ozone.

2018) and NMB (15.56 % for Sep. 2018). However, simulations for September 2018 period underestimated NO and CO concentrations which NMB values are -6.54 % and -52.3 %, respectively. These underestimations were higher for October 2018 period with -47.84 % and -61.76 %, respectively. Probably, a possible explanation is related to NO and CO emissions that could be underestimated, mainly for the October 2018 period. The spatial distribution of NO_x emission is subjected to errors, considering that heavy-duty vehicles in the MASP use specific roads where trucks are concentrated on motorways around the city (Ibarra-Espinosa et al., 2020).

Surface ozone

Ozone model results were also evaluated for each station type, as shown in Table 3.6. Three statistical values (NMB, NME, and r) for many station types complied with at least the criteria level of benchmarks suggested by Emery et al. (2017). If we consider a cutoff value of 80 μ g m⁻³, NMB values comply with the criteria benchmark. Some NMB values ($< \pm 5\%$) for station types presented a good performance because they comply with the goal benchmark, such as "Urban Park" (September and October), "Regional urban" (September and October), and "Forest Preservation" (only for October).

Only for stations classified as "Industry" (e.g., Santa Gertrudes), the correlation coefficient value for September 2018 complies with the goal level as a statistical benchmark (r > 0.75). Based on the hypothesis t-test evaluation, there is a significant linear relationship between observed and simulated values because the correlation coefficient is significantly different from zero, as shown in Table 3.7 ($\alpha = 0.05$ at confidence interval of 95 %). We can see that t-statistic values are greater than t critical values, which mean P-values are sufficiently small to reject the null hypothesis and accept the alternative hypothesis: "there is enough evidence to conclude that there is a significant linear relationship between observations and simulations".

Based on the score shown in Table 3.8, both months presented a better performance of simulations for ozone if a cutoff value is considered. Stations belonging to the MASP comply at least with two of three statistical benchmarks. Table 3.8 shows that surface ozone simulations presented better values for station types as 'Forest preservation' (only for October), 'Urban park', 'Industry' (only for September), and 'Regional urban' (both

Month type Statistic ^{(a}	Sep. 201 F. pre.	18 Urb	U. park	Ind	R. urb.	Oct. 201 F. pre.	18 Urb	U. park	Ind	R. urb.
n	638.00	11348.00	2211.00	639.00	9085.00	711.00	12418.00	2670.00	708.00	9527.00
MB	0.45	3.98	-1.74	17.03	18.76	4.35	10.25	4.48	19.88	19.52
MAGE	24.42	23.25	23.36	27.26	24.85	22.15	21.25	21.04	28.69	24.13
RMSE	30.61	29.52	29.24	32.53	30.73	28.40	28.45	27.69	35.09	30.73
$\rm NMB^{(b)}$	0.91	8.85	-3.62	33.65	35.65	10.47	27.56	11.35	45.84	40.93
$\rm NME^{(b)}$	48.83	51.76	48.58	53.85	47.21	53.34	57.15	53.26	66.16	50.59
NMB ^(c)	7.19	7.39	-0.58	-7.26	-0.28	-4.48	8.97	4.75	-8.53	-1.89
NME ^(c)	30.59	24.75	22.67	17.46	19.22	24.36	23.26	23.02	20.63	18.57
IOA	0.81	0.80	0.82	0.85	0.76	0.74	0.77	0.80	0.76	0.69
r	0.72	0.68	0.69	0.80	0.69	0.60	0.66	0.67	0.69	0.59
Mm	50.46	48.90	46.33	67.64	71.39	45.88	47.43	43.98	63.25	67.23
Om	50.01	44.93	48.08	50.61	52.63	41.53	37.18	39.50	43.37	47.70
Msd	43.84	39.91	40.11	36.72	28.55	34.73	34.91	36.53	30.57	25.17
Osd	28.94	28.93	31.59	46.12	32.54	24.67	25.88	27.87	39.46	27.36

Table 3.6 Statistical results for surface ozone by station type

^(a) MB, MAGE, RMSE, Mm, Om, Msd, and Osd values are in $\mu g m^{-3}$. Correlation coefficient is in dimensionless units. ^(b) No cutoff was applied, not recommended by Emery et al. (2017).

^(c) A cutoff value of 80 μ g m⁻³ was applied to calculate NMB and NME only for 1-hr ozone, suggested by Emery et al. (2017). Units are in percentage. Units are in percentage. Values in bold blue comply with the Goal benchmark and in only blue with the Criteria benchmark for ozone. Values in red don't comply with the statistical benchmarks.

Table 3.7 t-test values for correlation coefficient (r) for Sep. 2018

	F. pre.	Urb	U. park	Ind	R. urb.
t-statistic t critical	$26.21 \\ 1.96$	$98.94 \\ 1.96$	$44.87 \\ 1.96$	$33.70 \\ 1.96$	$91.00 \\ 1.96$

months). The station type "Forest preservation" presented less performance of the simulations for September. Despite cloudy and rain conditions, simulations for October 2018 presented good performance for this station type. However, the MB of ozone for October presented higher values than ozone simulations for September, which they mean ozone simulations are overestimated. Similarly, with the correlation coefficient, simulations for October presented low performance compared with September. As was mentioned at the beginning of this chapter, the WRF-Chem model, due to its configuration, could not adequately simulate heavy rainy days and consequently their influence on the ozone formation. In that sense, a possible explanation for this low performance based on MB and correlation coefficient values is attributed to rainy conditions in October. For instance, the cold front system reached the MASP for six days during October 5-10, 2018, according to the *Centro de Hidrografia da Marinha* (CHM, 2020). A possible explanation for the low performance attributed to "Urban" stations could be related to the model resolution, in which spatial emission rates distribution (3 km \times 3 km) do not represent local emission for stations closer to main roads.

If we can see details, Figure 3.6 shows statistical evaluations for each station. Correlation (r) values comply with the criteria level in all stations, and few of them comply with the statistical goal level. However, in some locations (eight stations), the model results do not comply with the criteria level for NMB. The statistical evaluation also calculated the maximum model overestimation value (NMB equals 41.40% in September 2018), located in Sorocaba station, belonging to the "Regional urban" station type. The model simulations only in twelve stations on September 2018 underestimated ozone concentrations between -19.17% (Araraquara, as an "Regional urban" station type) and -2.02% (Piracicaba, as an "Regional urban" station type) as NMB values.

Finally, model simulations of ozone for many stations (25) on September 2018 complied with the criteria level for the NME statistical benchmark suggested by Emery et al. (2017). Only three 'Regional urban' stations complied with the goal level, which are Limeira, Guaratinguetá, and Taubaté. IOA results are also shown in Figure 3.6 with very high values above 0.6. According to Willmott (1984), "the index of agreement varies between 0 and 1 where a value of 1 expresses perfect agreement between O (observations) and P (predictions) and 0 describes complete disagreement." IOA values greater than 0.8 suggest that the model is highly accurate for those stations that comply with that criteria (e.g., Paulínia, Campinas Taquaral, Piracicaba, Americana, Jundiaí, Limeira, Carapicuíba, Santana, Pico do Jaraguá, Cid. Universitária USP IPEN, Ibirapuera, São Caetano do Sul, Diadema, Guarulhos-Paço Municipal, Parque D. Pedro II, Interlagos, and Itaquera).

Figure 3.7 shows us the comparison in time between observed (black dots) and modeled (green line) surface

Month	Statistic	F. pres.	Urban	U. park	Ind	R. urb.	
Sep. 2018	NMB	\checkmark	\checkmark	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	7
	NME		\checkmark	\checkmark	\checkmark	\checkmark	4
	r	\checkmark	\checkmark	\checkmark	$\checkmark\checkmark$	\checkmark	6
	Total	2	3	4	4	4	17
Oct. 2018	NMB	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	8
	NME	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	5
	r	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	5
	Total	4	3	4	3	4	18

Table 3.8 Summary of compliance of the ozone statistical benchmarks by station types with a cutoff value for NMB and NME

Note. $\checkmark \checkmark =$ goal and criteria levels compliance suggested by Emery et al. (2017).



Figure 3.6 Statistical results for surface ozone during the period of September 2018

ozone values for September and October 2018. September 14-15 were not considered due to the high inaccuracy of the WRF-Chem model to simulate cloud. Ozone concentrations as daily maximum 8-h rolling mean (MDA8) were also calculated for observed and simulated values, shown in Figure 3.8. The are many stations to plot model simulations compared with observations. For that reason, comparison values are classified by station type. Both figures show peak values for model simulations and observations. This consideration is essential for model validation where model capabilities reproduce peak ozone concentrations (Seinfeld and Pandis, 2016). Differences could be associated with weather conditions as cloudy days of September (4-5, 16-18, 26-30) of 2018, shown in Appendix (Figure B.8).

Other pollutants related to ozone formation

Charts by station type for remaining pollutants (NO_x, CO, and Toluene) are shown in Appendix D. Pinheiros and S.André-Capuava stations measured toluene and the comparison against WRF-Chem simulation are presented in Figure D.6. Figure 3.9 shows the primary (CO, NO_x) and secondary (O₃) pollutants' diurnal profile. Maximum ozone concentrations were reached between 12:00-16:00 hours, with peak concentration at 13:00 hours on average. However, this does not correspond with observations when peak concentration is reached between 14:00-15:00 hours, as shown in Figure 2.7. One of the ozone precursors (NO_x) builds up during the morning rush hour and ending afternoon time, associated with traffic rush hours in the MASP. Photochemical activity reduces NO₂ concentrations and enhances ozone formation. This behavior depends on the VOC/NO_x regime. In the MASP, a VOC-sensitive (NO_x-saturated) predominates. Furthermore, flex-fuel vehicles can burn hydrous ethanol and gasohol. They contribute with aldehydes (Nogueira et al., 2014) which at high concentrations can lead to an increase in troposphere reactivity as a driver in the ozone formation.

Model simulations for Pinheiros station as an hourly mean concentration of the simulated period are compared with observations, shown in Figure 3.10. These results extend our knowledge about the influence of hydrocarbons in photochemical activity. Toluene modeled values appear to be over-predicted in night-time



Figure 3.7 Comparison between modeled and simulated values for surface ozone during September and October 2018 period considering all stations classified by type.

hours and underestimated during daylight hours. This hourly variation during daylight time is an indicator of remain hydrocarbons and their contributions to ozone formation. The statistic evaluation for toluene shows no significant linear correlation for the S. André Capuava station (Table 3.9). Pinheiros station and others (Paulínia, SJC, and SJC-VV) have significant linear correlations.

There are several possible explanations for this result. Temporal distribution of road transport emission may likely have contributed to inaccuracies because it represents the year 2014 (Andrade et al., 2015). Another source of inaccuracy is related to vehicle fleet averaged by month, which could not represent daily variations along the months, such as the weekend effect. Further data collection, such as vehicle fleet by day and hour, would be needed to determine how temporal distribution and road transport emissions affect toluene and ozone simulations. Thus, the Vein model (Ibarra-Espinosa et al., 2018) could improve these results based on high spatial and temporal resolution of the vehicle fleet from real-time GPS.



Figure 3.8 Comparison between modeled and simulated values based on MDA8 surface ozone during (a) September and (b) October 2018 period for station types.

Note. Daily maximum 8 h rolling mean (MDA8) of surface ozone concentrations.

Station	Paulínia	Pinheiros	S.André-C.	SJC	SJC-VV
Statistic					
n	1153	1441	1384	1455	1457
MB	0.70	3.65	3.14	2.31	-0.10
MAGE	3.85	6.68	5.03	3.25	3.34
RMSE	6.52	9.36	7.02	4.89	5.10
NMB	17.36	60.26	74.84	196.9	-2.47
NME	96.04	110.11	119.77	276.35	81.50
IOA	0.39	0.42	0.31	0.33	0.51
r	0.19	0.15	0.03	0.13	0.26
Mm	4.70	9.72	7.34	3.49	4.00
Om	4.01	6.06	4.20	1.17	4.10
Msd	3.47	6.98	5.28	3.08	3.53
Osd	6.16	6.22	3.55	3.45	4.73
t-statistic	6.57	5.76	1.12	5.00	10.27
t critical	1.96	1.96	1.96	1.96	1.96
Significant	True	True	False	True	True

Table 3.9 Statistical results for toluene in Sep-Oct 2018 period

San José-Campos (SJC). S.José Campos-Vista Verde (SJC-VV).

3.2 Future changes under RCP 4.5 and 8.5 scenarios

Model results were obtained for Sep-Oct 2030 period, considering two RCP emission scenarios as meteorology IC/BC from the CESM Bias-Corrected dataset. We used the same emissions files in September-October 2018,



Figure 3.9 Average hourly concentration of model simulations during the course of a day (Sep-Oct 2018) of some important pollutants Note.

Shaded area is the standard deviation.

therefore any variation on O_3 concentration is caused by the projections of the meteorological conditions.

The RCP 4.5 is a stabilization scenario representing an emission mitigation through changes in the energy system, including shifts to electricity from lower emissions energy technologies, carbon capture and geologic storage technology (Thomson et al., 2011). The worst-case scenario analyzed is the RCP 8.5. It represents a very high baseline emission scenario, based on a non-climate policy known as "business as usual", combined with growing high population and high demands for fossil fuel and food (Riahi et al., 2011). These two scenarios were compared with simulations values from "current" conditions (Sep-Oct 2018).



Figure 3.10 Average hourly concentrations comparison between observed (Obs.) and modeled (Mod.) values during the course of a day.

3.2.1 Changes in meteorological conditions

Model simulations show different temperature variations for September and October. As shown in Appendix (Figure D.7), the RCP 8.5 scenario presented higher temperature values than RCP 4.5 and current (Sep. 2018) scenarios. Daily mean temperatures, based on the RCP 8.5, reach 32°C at the end of September and in the first week of October 2030. However, temperature simulations for October 2030 under two RCP scenarios presented similar variation ranges as in October 2018.

For monthly average, we found higher temperatures in September 2030 based on the RCP 8.5 scenario as shown in Figure 3.11. The rising temperature affects the biogenic emissions inside the WRF-Chem because they are dependent on the temperature. The most intriguing result arises from the comparison of monthly mean temperatures between the current 2018 and the RCP 4.5 scenarios. It is interesting to note that temperatures for Sep. 2018 are very similar to the RCP 4.5 scenario for September 2030 (getting ahead 12 years in the future!). Moreover, October's temperature simulations were similar, where monthly mean values for October 2018 are slightly above the RCP 4.5 scenario but below the RCP 8.5. These results suggest that the RCP scenarios could underestimate the future rising temperatures for the São Paulo state, which could be higher. Daily mean relative humidity variations are showed in Figure D.8 for current and future scenarios. These results are inversely proportional to the temperature values as shown in Figure 3.12. Low values of relative humidity could occur for the RCP 8.5 scenario, mainly for September 2030. In October, the current scenario presented high relative humidity values than the RCP 4.5 and 8.5 scenarios.

Wind speed and direction did not show marked differences between current and future RCP scenarios (Figure D.9). Regarding accumulated rain (Figure D.10), the RCP 4.5 scenario presented higher daily modeled values between September 20-24 (2030). However, total monthly rain values reveal a decrease for the RCP scenarios (Figure 3.13), in which the RCP 8.5 could present low values of monthly rain in the future (2030).



Figure 3.11 Monthly mean temperature by scenario and station: (a) September, (b) October.

3.2.2 Changes in surface ozone

The changes of surface O_3 concentrations from the 2018 to 2030 were analyzed by month due to different weather conditions. MDA8 for ozone values were calculated for each scenario and station. After that, mean differences



Figure 3.12 Monthly mean relative humidity by scenario and station: (a) September, (b) October.

of MDA8 between the future and the current scenario were obtained for each station and for all the second modeling domain area.

September

Figure 3.14 shows the comparison among simulated surface ozone concentrations for the three scenarios: Current (Sep. 2018), RCP 4.5 (Sep. 2030), and RCP 8.5 (Sep. 2030). From the chart, we can observe that there are peak concentrations in some days (Sep. 8-11, 22-24) of the current scenario greater than in future conditions. The industry type station's model simulations presented peak concentrations associated mainly with the RCP 4.5 scenario greater than the RCP 8.5 scenario. For the lasts days of September, the RCP 4.5 scenario also presented peak values in the MASP stations. There are many days with peak concentrations associated with the RCP 8.5 scenario, mainly between Sep. 25-29 in the MASP.

On average, the most remarkable result to emerge from the model simulations is that the RCP 8.5 scenario



Figure 3.13 Daily and monthly total rain by scenarios for the IAG/USP weather station

presented higher peak concentrations in the MASP, represented by station types as Urban, Urban park, and Forest preservation. Simulations for the RCP 4.5 show decreases of surface O_3 concentrations in the MASP, whereas the RCP 8.5 simulations presented increases, mainly in urban areas. This result confirms previous findings in the time series analysis in which the RCP 8.5 scenario presented many days with higher peak concentrations due to the increase in temperature. Regarding changes of surface ozone for the RCP 4.5, decreasing concentrations are also confirmed in previous findings by Schuch et al. (2020), considering results for 2030 under two emission scenarios (i.e., Current Legislation, Mitigation, and Maximum Feasible Reduction),



Figure 3.14 Surface ozone concentrations for Sep-Oct 2018 compared with RCP 4.5 and 8.5 for Sep. 2030.

with anthropogenic emissions variation depending on the chosen scenario.

Further analyses carried out for the MDA8 for ozone concentrations confirmed the initial findings regarding ozone increases in urban areas inside the MASP. Figure 3.15 reveals many days with higher MDA8 ozone concentrations in the MASP, based on the RCP 8.5 scenario as meteorology IC/BC. However, there are few days with higher concentrations belonging to the current scenario (Sep. 2018) and to the RCP 4.5 at the beginning and the last days of September. Based on MDA8 ozone results, Figure 3.17 shows spatial changes of surface ozone concentration as monthly mean difference between Sep. 2030 (RCP 4.5 and 8.5) and Sep. 2018.



Figure 3.15 Maximum daily 8-hr rolling mean for surface ozone concentrations for Sep 2018 and 2030 (RCP 4.5 and 8.5)

October

Model simulations for October are different with higher peak concentrations when compared with the September simulation results shown in the previous section. Figure 3.14 shows a clear trend in the increase of surface ozone concentrations for both RCP scenarios in 2030 compared with October 2018. Although the RCP 8.5 is the worst-case scenario, we found much higher simulated surface ozone values for the RCP 4.5 than the RCP 8.5, markedly in days between Oct. 20 to 25. The most remarkable finding from the data analysis is that in "Urban" stations occurred a higher peak concentration with almost 320 μ g m⁻³ corresponded to the RCP 4.5 scenario. Some days for the RCP 8.5 presented higher concentrations than the other scenarios (current and the RCP 4.5).

Model results of MDA8 ozone concentrations, shown in Figure 3.16, revealed some days with much higher
concentrations for the RCP 8.5 and 4.5 scenarios. However, the RCP 8.5 scenario presented for some days (20-25 Oct.) low values of MDA8 ozone concentrations than the current and the RCP 4.5 scenarios. In the last days of October, MDA8 ozone concentrations for the RCP 8.5 present lower values than the October 2018 period. These last days presented rainy conditions for the RCP 8.5 scenario.



Figure 3.16 Maximum daily rolling mean for surface ozone concentrations for Oct 2018 and 2030 (RCP 4.5 and 8.5)

These results thus need to be interpreted with caution. A possible explanation is that cloudy and low temperatures influenced differently for each scenario. Changes in MDA8 surface ozone concentrations were identified in Figure 3.18. From the chart, we can note that the RCP 4.5 presented more increases in the MDA8 ozone concentrations than the RCP 8.5. Some locations for the RCP 8.5 presented decreases in ozone concentrations, especially when they are far away from the urban center of the MASP.

On average, the two RCP scenarios showed increases in the MDA8 of ozone concentration in the urban area of the MASP, as we can see in Figure 3.19. Ozone simulations for September 2030 based on the RCP 4.5 scenario only showed decreases in many zones of the second modeling domain area. In contrast, it is very remarkable to observe the MDA8 increases on average for September 2030 based on the RCP 8.5 scenario. On the other hand, simulations for October only showed increases for both scenarios, especially in some areas in the north-west (RCP 4.5 and RCP 8.5) and south-east (RCP 8.5) of the MASP. On average, a decrease in ozone concentrations is observed in the ocean area under the RCP4.5 scenario. Also, MDA8 decreases are shown to the northeast and southwest of the MASP under the RCP 8.5 scenario.



Figure 3.17 MDA8 ozone mean spatial changes between Sep 2030 (RCP 4.5 and 8.5) and Sep. 2018



Figure 3.18 MDA8 ozone mean spatial changes between Oct. 2030 (RCP 4.5 and 8.5) and Oct. 2018



Figure 3.19 Mean differences of surface maximum daily 8-h average (MDA8) by month between both scenarios (RCP 4.5 and 8.5 for 2030) and the current scenario (2018).

Chapter 4

Conclusions

If not now, when? If not you, who?

Paraphrased from Hillel the Elder

It is a challenging to control tropospheric ozone's adverse effects on human health and climate in a regional and global context. The use of chemical transport models (CTM) demonstrated to be a powerful tool to study the ozone and secondary pollutants formation. The WRF-Chem model is a CTM that allows us to understand the driving factors that led to the ozone formation in urban areas and how this pollutant concentration can change at different weather conditions when maintaining anthropogenic emission rates.

Here, we examine the impact of future changes in meteorological conditions for short periods (two months), maintaining the same emission rates of the precursors for the surface ozone formation. This study enhances our understanding of the impact of future meteorological conditions on ozone formation in urban and regional areas based on RCP as climate change scenarios used by the IPCC (2013).

The emission approach considered anthropogenic (road transport, industry, and residential) and biogenic sources for September and October 2018 period (named in this study as 'current' conditions). The RCP 4.5 and RCP 8.5 were considered as future scenarios during September and October 2030. Model evaluation results for ozone formation during the base case scenario suggest that part of the emission approach represents the current conditions. NO_x and VOC emissions are the main parameters that have to be represented for ozone formation. However, NO_x and CO did not present good agreement with observations, whereby CO emissions were underestimated in the model simulations by comparison with observations. Despite these findings, surface ozone simulations for September 2018 comply with two of three statistical benchmarks for the stations inside the MASP, satisfying the model evaluation reasonably well. Model performance evaluation for October 2018 was satisfactory compared with September because it complied with all benchmarks at the criteria level suggested by Emery et al. (2017). However, this modeling period had some limitations due to low values of r compared with ozone formation simulation for September 2018 during rainy conditions. It can be related to the model configurations (i.e., microphysics and cumulus parameterizations). Other limitations are related to wind directions and errors due to the spatial distribution of emission sources.

Surface ozone concentrations based on RCP 4.5 and RCP 8.5 scenarios for future months (Sep-Oct 2030) presented differences compared with current conditions results, mainly in the peak ozone concentrations. Sim-

ulations, mainly for September 2030, suggest that surface ozone can change depending on the RCP scenario. Despite model limitations, this work brings us an insight into the future changes in weather conditions that could affect surface ozone concentrations. Humankind can follow the worst-case scenario (the RCP 8.5) with negative impacts in urban areas by increasing ozone concentration. It is interesting to note that temperatures for September 2030 from the RCP 4.5 scenario are very close to that from September 2018 as the monthly mean. An implication of this comparison is the possibility that the RCP scenarios as meteorology IC/BC could underestimate future weather conditions for the São Paulo state, which means that even a worse scenario can occur. These future meteorological changes also reveal a decrease in monthly accumulated rain, worsening when the RCP increases the radiative forcing. In October, it was observed different rainy periods for each scenario that affected the ozone formation.

In conclusion, this work, with the application of the WRF-Chem model, found negative impacts in the surface ozone formation over the MASP in September due to changes only in meteorological conditions under the RCP 8.5 scenario, maintaining the emission rates and land use in 2030 similar to 2018.

Month	Location	Туре	2018	2030 (RCP 4.5)	2030 (RCP 8.5)
Sep.	Outside	Industry Regional urban	100.45 97.04	94.35 (-6.1) 90.66 (-6.4) 91.67 (-6.4)	98.55 (-1.9) 99.93 (+2.9) 100.12 (-10.5)
	MASP	Forest preservation Urban Urban park	92.67 90.47 90.24	$84.77 (-7.9) \\81.82 (-8.7) \\81.99 (-8.3)$	$\begin{array}{c} 106.13 \ (+13.5) \\ 105.48 \ (+15.0) \\ 105.37 \ (+15.1) \end{array}$
Oct.	Outside	Industry Regional urban	$95.04 \\ 91.42$	104.81 (+9.8) 98.15 (+6.7)	98.64 (+3.6) 91.00 (-0.4)
	MASP	Forest preservation Urban Urban park	$79.69 \\ 80.04 \\ 80.38$	$\begin{array}{c} 89.70 \ (+10.0) \\ 87.73 \ (+7.7) \\ 87.36 \ (+7.0) \end{array}$	$\begin{array}{c} 84.20 \end{array} \left(\begin{array}{c} +4.5 \\ 80.26 \end{array} \right) \\ 80.98 \end{array} \left(\begin{array}{c} +0.2 \\ +0.6 \end{array} \right)$

Table 4.1 MDA8 ozone monthly average in $\mu g m^{-3}$ by station type

^(a) MDA8 ozone variations are shown in parentheses.

^(b) Increases marked in red are greater than +2.749 based on MB result for temperature (shown in Table 3.2) and the relationship on average of the increase between temperature and ozone for September (RCP 8.5 - current).

The application of the meteorological IC/BC conditions from the RCP 4.5 scenario for September 2030 decreases the daily maximum rolling 8-hr mean surface ozone concentrations in the MASP compared with September 2018 (Table 4.1). The rise of temperature ($\pm 2.50 \pm 0.12$ °C on average) for the RCP 8.5 scenario was the main driver of ozone formation. The MDA8 ozone as monthly mean in the station types showed reductions for the RCP 4.5 scenario and increases and one reduction ("Industry") for the RCP 8.5 scenario.

However, due to more model uncertainties in weather representation for October attributed to the mean bias, the impact in surface ozone concentrations could not be better represented, mainly due to different rainy periods. However, there are impacts for the RCP 4.5 in the ozone formation higher than the RCP 8.5 (Table 4.1). Simulations presented differences between years in rainy periods that affected ozone formation. Changes in temperature on monthly average for both scenarios (0.88 ± 0.14 °C for the RCP 4.5 and $+2.05 \pm 0.15$ °C for the RCP 8.5) are positives compared with September (-0.16 ± 0.27 °C for the RCP 4.5 and $+2.5 \pm 0.12$ °C for the RCP 8.5). It could also be a period for future research to analyze the impact of ozone formation under changes

in rainy conditions.

4.1 Study limitations and suggestions for future works

4.1.1 Limitations

This study has limitations related to emission inventory, mainly associated with spatial and temporal distribution of emission sources. Despite using the fleet for 2018 (emissions factor), the temporal distribution for road transport emission may have contributed to inaccuracies because it represents the year 2014, described in Andrade et al. (2015).

The present study has only examined future changes in meteorological conditions for short-periods. Consequently, the static data used as the land cover does not represent the urban expansion for 2018 and the next decade (2030).

4.1.2 Suggestions for future works

Future work will concentrate on two tasks, as shown as follows:

- The first task depends on high computational resources to simulate extended periods (i.e., 30 years) representing a climate change scenario. However, it is needed to consider the following:
 - Land cover and anthropogenic emission changes are relevant factors and these represent future years based on environmental and economic policies.
 - A critical issue to resolve for future studies is improving the cumulus parameterizations in the WRF-Chem model to get better results during rainy days.
 - Despite significant values, low Pearson correlation results for primary pollutants (NO_x, CO, and toluene) suggest the following direction for future research to improve or evaluate other emission estimation approaches. The VEIN model (Ibarra-Espinosa et al., 2018) could improve spatial emission distribution, increasing the model accuracy and precision.
- The second task is to evaluate different socioeconomic climate change scenarios based on local policy decisions for the São Paulo state. Some initiatives to reduce emission contributions from the road transport sector are:
 - As is suggested in Andrade et al. (2017), as practice in other megacities, could be to scrap old vehicles (those 10-15 years of age) in the São Paulo State.
 - Ethanol as a vehicle fuel type contributes to ozone formation, so another scenario is to reduce this emission source. One of the probably environmental policies could be improving public transportation, applying cleaner fuel as electric buses.
 - Related to vehicle use intensity, less exhaust emission from vehicles should mean less traffic flow.
 Thus, an implication of this is to improve public transportation reducing congested traffic, mainly in central urban areas.

Bibliography

- Alvim, D. S., Gatti, L. V., Corrêa, S. M., Chiquetto, J. B., Santos, G. M., Rossatti, C. d. S., Pretto, A., Rozante, J. R., Figueroa, S. N., Pendharkar, J. and Nobre, P. (2018), 'Determining VOCs reactivity for ozone forming potential in the megacity of São Paulo', Aerosol and Air Quality Research 18(9), 2460–2474. URL: http://www.aaqr.org/doi/10.4209/aaqr.2017.10.0361
- Andrade, M. d. F., Kumar, P., de Freitas, E. D., Ynoue, R. Y., Martins, J., Martins, L. D., Nogueira, T., Perez-Martinez, P., de Miranda, R. M., Albuquerque, T., Gonçalves, F. L. T., Oyama, B. and Zhang, Y. (2017), 'Air quality in the megacity of São Paulo: Evolution over the last 30 years and future perspectives', Atmospheric Environment 159, 66–82.
 URL: http://dx.doi.org/10.1016/j.atmosenv.2017.03.051
- Andrade, M. D. F., Ynoue, R. Y., Freitas, E. D., Todesco, E., Vara Vela, A., Ibarra, S., Martins, L. D., Martins, J. A. and Carvalho, V. S. B. (2015), 'Air quality forecasting system for Southeastern Brazil', Frontiers in Environmental Science 3(February), 6975.
 URL: http://journal.frontiersin.org/Article/10.3389/fenvs.2015.00009/abstract
- Andrade, M. F., Oliveira, G., Miranda, R. M. and Nogueira, T. (2019), 'Evolution of vehicular emission factors, from 2001 to 2018, in sao paulo, brazil, evaluated by tunnel measurements'.
 URL: https://agu.confex.com/agu/fm19/meetingapp.cgi/Paper/581795
- Brasseur, G. P., Orlando, J. J. and Tyndall, G. S. (1999), Atmospheric Chemistry and Global Change, Oxford University Press, New York.
- Carvalho, V. S. B., Freitas, E. D., Martins, L. D., Martins, J. A., Mazzoli, C. R. and Andrade, M. d. F. (2015),
 'Air quality status and trends over the Metropolitan Area of São Paulo, Brazil as a result of emission control policies', *Environmental Science & Policy* 47, 68–79.
 URL: http://www.sciencedirect.com/science/article/pii/S1462901114002081
- CETESB (2019a), Emissões Veiculares no Estado de São Paulo 2018, Série Relatórios, São Paulo.
- CETESB (2019b), Qualidade do ar no estado de São Paulo 2018, 21 edn, Série Relatórios, São Paulo. URL: https://cetesb.sp.gov.br/ar/publicacoes-relatorios/
- CHM (2020), 'Cartas sinóticas', https://www.marinha.mil.br/chm/dados-do-smm-cartas-sinoticas/ cartas-sinoticas. Accessed: 2020-09-30.

Chou, S. C., Lyra, A., Mourão, C., Dereczynski, C., Pilotto, I., Gomes, J., Bustamante, J., Tavares, P., Silva, A., Rodrigues, D., Campos, D., Chagas, D., Sueiro, G., Siqueira, G. and Marengo, J. (2014), 'Assessment of Climate Change over South America under RCP 4.5 and 8.5 Downscaling Scenarios', American Journal of Climate Change 03(05), 512–527.

URL: https://file.scirp.org/Html/7-2360202_52887.htm

- Crippa, M., Solazzo, E., Huang, G., Guizzardi, D., Koffi, E., Muntean, M., Schieberle, C., Friedrich, R. and Janssens-Maenhout, G. (2020), 'High resolution temporal profiles in the emissions database for global atmospheric research', *Scientific Data* 7(1). URL: https://www.nature.com/articles/s41597-020-0462-2
- Cunningham, C., Cunha, A. P., Brito, S., Marengo, J. and Coutinho, M. (2017), Climate Change and Drought in Brazil, *in* 'Reduction of vulnerability to disasters: from knowledge to action', Rima, São Paulo, chapter 14, p. 16.
- de Lima, G. N. and Magaña Rueda, V. O. (2018), 'The urban growth of the metropolitan area of sao paulo and its impact on the climate', Weather and Climate Extremes 21, 17 – 26.
 URL: http://www.sciencedirect.com/science/article/pii/S2212094718300082
- Dominutti, P., Nogueira, T., Fornaro, A. and Borbon, A. (2020), 'One decade of vocs measurements in são paulo megacity: Composition, variability, and emission evaluation in a biofuel usage context', Science of The Total Environment 738, 139790.
 UDL http://www.common.com/www.common.com/www.common.com/www
 - **URL:** https://www.sciencedirect.com/science/article/pii/S0048969720333106
- Emery, C., Liu, Z., Russell, A. G., Odman, M. T., Yarwood, G. and Kumar, N. (2017), 'Recommendations on statistics and benchmarks to assess photochemical model performance', *Journal of the Air & Waste Management Association* 67(5), 582–598. PMID: 27960634. URL: https://doi.org/10.1080/10962247.2016.1265027
- Farmer, G. T. and Cook, J. (2013), Climate Change Science: A Modern Synthesis, Springer Netherlands, Dordrecht. URL: http://link.springer.com/10.1007/978-94-007-5757-8
- Fast, J. D., Gustafson Jr., W. I., Easter, R. C., Zaveri, R. A., Barnard, J. C., Chapman, E. G., Grell, G. A. and Peckham, S. E. (2006), 'Evolution of ozone, particulates, and aerosol direct radiative forcing in the vicinity of houston using a fully coupled meteorology-chemistry-aerosol model', *Journal of Geophysical Research:*

URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JD006721

Atmospheres **111**(D21).

Gaffney, J. S. and Marley, N. A. (2009), 'The impacts of combustion emissions on air quality and climate – from coal to biofuels and beyond', Atmospheric Environment 43(1), 23–36. Atmospheric Environment - Fifty Years of Endeavour.

URL: https://www.sciencedirect.com/science/article/pii/S1352231008009175

- Gavidia-Calderón, M., Vara-Vela, A., Crespo, N. and Andrade, M. (2018), 'Impact of time-dependent chemical boundary conditions on tropospheric ozone simulation with WRF-Chem: An experiment over the Metropolitan Area of São Paulo', Atmospheric Environment 195, 112–124. URL: https://linkinghub.elsevier.com/retrieve/pii/S1352231018306216
- Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C. and Eder, B. (2005),
 'Fully coupled 'online' chemistry within the WRF model', Atmospheric Environment 39(37), 6957–6975.
 URL: https://linkinghub.elsevier.com/retrieve/pii/S1352231005003560
- Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I. and Geron, C. (2006), 'Estimates of global terrestrial isoprene emissions using megan (model of emissions of gases and aerosols from nature)', Atmospheric Chemistry and Physics 6(11), 3181–3210. URL: https://acp.copernicus.org/articles/6/3181/2006/
- Gustafson Jr., W. I., Chapman, E. G., Ghan, S. J., Easter, R. C. and Fast, J. D. (2007), 'Impact on modeled cloud characteristics due to simplified treatment of uniform cloud condensation nuclei during neages 2004', *Geophysical Research Letters* 34(19).

URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007GL030021

- Holnicki, P. and Nahorski, Z. (2015), 'Emission data uncertainty in urban air quality modeling—case study', Environmental Modeling & Assessment 20(6), 583–597.
 URL: https://link.springer.com/article/10.1007/s10666-015-9445-7
- Ibarra-Espinosa, S., Ynoue, R., O'Sullivan, S., Pebesma, E., Andrade, M. d. F. and Osses, M. (2018), 'VEIN v0.2.2: an R package for bottom-up vehicular emissions inventories', *Geoscientific Model Development* 11(6), 2209–2229.

URL: https://www.geosci-model-dev.net/11/2209/2018/

Ibarra-Espinosa, S., Ynoue, R. Y., Ropkins, K., Zhang, X. and de Freitas, E. D. (2020), 'High spatial and temporal resolution vehicular emissions in south-east Brazil with traffic data from real-time GPS and travel demand models', Atmospheric Environment 222, 117136.

URL: https://linkinghub.elsevier.com/retrieve/pii/S1352231019307757

- IBGE (2020a), Cidades e Estados. accessed October, 2020. URL: https://www.ibge.gov.br/cidades-e-estados
- IBGE (2020b), Estimativas da população. accessed October, 2020.URL: https://www.ibge.gov.br/estatisticas/sociais
- IPCC (2013), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, United Kingdom, New York.
- Jacob, D. J. (1999), Introduction to Atmospheric Chemistry, first edit edn, Princeton University Press, Harvard University.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Dentener, F., Muntean, M., Pouliot, G., Keating, T., Zhang, Q., Kurokawa, J., Wankmüller, R., van der Gon, H. D., Kuenen, J. J. P., Klimont, Z., Frost, G., Darras, S., Koffi, B. and Li, M. (2015), 'HTAP_v2.2: a mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution', *Atmospheric Chemistry and Physics* 15(19), 11411– 11432.

URL: https://doi.org/10.5194/acp-15-11411-2015

- Kasting, J. (1993), 'Earth's early atmosphere', Science 259(5097), 920–926.
- Kumar, R. (2020), 'Wrf-chem tools for the community', https://www2.acom.ucar.edu/wrf-chem/ wrf-chem-tools-community. Accessed: 2020-09-30.
- Letcher, T. M. (2021), 1 global warming—a complex situation, in T. M. Letcher, ed., 'Climate Change (Third Edition)', third edition edn, Elsevier, pp. 3–17.
 URL: https://www.sciencedirect.com/science/article/pii/B9780128215753000013
- Marengo, J. A., Ambrizzi, T., Alves, L. M., Barreto, N. J. C., Reboita, M. S. and Ramos, A. M. (2020), 'Changing trends in rainfall extremes in the metropolitan area of são paulo: Causes and impacts', Frontiers in Climate 2.
 UPL: https://doi.org/10.2280%2Efclim.2020.00002

URL: https://doi.org/10.3389%2Ffclim.2020.00003

- Marengo, J. A., Souza, C. M., Thonicke, K., Burton, C., Halladay, K., Betts, R. A., Alves, L. M. and Soares, W. R. (2018), 'Changes in Climate and Land Use Over the Amazon Region: Current and Future Variability and Trends', Frontiers in Earth Science 6(December), 1–21. URL: https://www.frontiersin.org/article/10.3389/feart.2018.00228/full
- Marengo, J., Valverde, M. and Obregon, G. (2013), 'Observed and projected changes in rainfall extremes in the metropolitan area of são paulo', *Climate Research* 57(1), 61–72.
 URL: https://doi.org/10.3354%2Fcr01160
- Mazzoli da Rocha, C. R. (2013), Estudo Numérico da influência das Mudançãs Climáticas e das Emissões Urbanas no Ozônio Troposférico da Região Metropolitana de São Paulo, Tese de doutorado, Universidade de São Paulo.
- Monaghan, A. J., Steinhoff, D. F., Bruyere, C. L. and Yates, D. (2014), 'Ncar cesm global bias-corrected cmip5 output to support wrf/mpas research'. URL: https://doi.org/10.5065/D6DJ5CN4
- Monk, K., Guérette, E.-A., Paton-Walsh, C., Silver, J. D., Emmerson, K. M., Utembe, S. R., Zhang, Y., Griffiths, A. D., Chang, L. T.-C., Duc, H. N., Trieu, T., Scorgie, Y. and Cope, M. E. (2019), 'Evaluation of regional air quality models over sydney and australia: Part 1—meteorological model comparison', *Atmosphere* 10(7).

URL: https://www.mdpi.com/2073-4433/10/7/374

Navidi, W., Navidi, W. and of Mines, C. S. (2019), *Statistics for Engineers and Scientists*, McGraw-Hill Education.

URL: https://books.google.com.pe/books?id=-EcdxAEACAAJ

- NCEP, NWS, NOAA and of Commerce, U. D. (2000), 'NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999'. URL: https://doi.org/10.5065/D6M043C6
- NOAA, ESRL and GSD (2018), 'WRF-Chem Version 3.9.1.1 User's Guide'. URL: https://ruc.noaa.gov/wrf/wrf-chem/user-support.htm
- Nobre, C. A., Marengo, J. A. and Soares, W. R., eds (2019), Climate Change Risks in Brazil, Springer International Publishing, Cham. URL: http://link.springer.com/10.1007/978-3-319-92881-4
- Nogueira, T., Dominutti, P. A., de Carvalho, L. R. F., Fornaro, A. and de Fatima Andrade, M. (2014), 'Formaldehyde and acetaldehyde measurements in urban atmosphere impacted by the use of ethanol biofuel: Metropolitan area of sao paulo (masp), 2012–2013', Fuel 134, 505–513. URL: https://www.sciencedirect.com/science/article/pii/S0016236114005535
- Nullis, C. (2020), '2020 on track to be one of three warmest years on record', WMO website. World Meteorological Organization.

URL: WMO website

- Nuvolone, D., Petri, D. and Voller, F. (2018), 'The effects of ozone on human health', Environmental Science and Pollution Research 25(9), 8074–8088. URL: http://link.springer.com/10.1007/s11356-017-9239-3
- Oliveira, A. P., Bornstein, R. D. and Soares, J. (2003), 'Annual and diurnal wind patterns in the city of são paulo', Water, Air, & Soil Pollution: Focus 3(5), 3–15. URL: https://doi.org/10.1023/A:1026090103764
- Orlando, J. P., Alvim, D. S., Yamazaki, A., Corrêa, S. M. and Gatti, L. V. (2010), 'Ozone precursors for the São Paulo Metropolitan Area', Science of the Total Environment 408(7), 1612–1620. URL: http://dx.doi.org/10.1016/j.scitotenv.2009.11.060
- Pérez-Martínez, P. J., Miranda, R. M., Nogueira, T., Guardani, M. L., Fornaro, A., Ynoue, R. and Andrade, M. F. (2014), 'Emission factors of air pollutants from vehicles measured inside road tunnels in são paulo: case study comparison', *International Journal of Environmental Science and Technology* 11(8), 2155–2168. URL: https://doi.org/10.1007/s13762-014-0562-7
- Reboredo, B., Arasa, R. and Codina, B. (2015), 'Evaluating Sensitivity to Different Options and Parameterizations of a Coupled Air Quality Modelling System over Bogotá, Colombia. Part I: WRF Model Configuration', Open Journal of Air Pollution 4(2), 47–64.

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N. and Rafaj, P. (2011), 'RCP 8.5 - A scenario of comparatively high greenhouse gas emissions', *Climatic Change* 109(1-2), 33–57.

URL: http://link.springer.com/10.1007/s10584-011-0149-y

- Ritter, M. (2013), An air pollution modeling system for Switzerland using WRF-Chem: development, simulation, evaluation, PhD thesis, Faculty of Science. Doctoral Thesis.
- Russell, A. and Dennis, R. (2000), 'Narsto critical review of photochemical models and modeling', Atmospheric Environment 34(12), 2283–2324.

URL: https://www.sciencedirect.com/science/article/pii/S1352231099004689

Sánchez-Ccoyllo, O. R., Ynoue, R. Y., Martins, L. D. and Andrade, M. d. F. (2006), 'Impacts of ozone precursor limitation and meteorological variables on ozone concentration in São Paulo, Brazil', Atmospheric Environment 40(SUPPL. 2), 552–562.

URL: https://linkinghub.elsevier.com/retrieve/pii/S1352231006006030

Schuch, D., Andrade, M. d. F., Zhang, Y., Dias de Freitas, E. and Bell, M. L. (2020), 'Short-term responses of air quality to changes in emissions under the representative concentration pathway 4.5 scenario over brazil', *Atmosphere* 11(8).

URL: https://www.mdpi.com/2073-4433/11/8/799

- Seinfeld, J. H. and Pandis, S. N. (2016), Atmospheric Chemistry and Physics: From Air Pollution to Climate Change, 3 edn, John Wiley & Sons, Inc.
- Sillman, S. (2014), 11.11 tropospheric ozone and photochemical smog, in H. D. Holland and K. K. Turekian, eds, 'Treatise on Geochemistry (Second Edition)', second edition edn, Elsevier, Oxford, pp. 415–437. URL: https://www.sciencedirect.com/science/article/pii/B9780080959757009116
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. and yu Huang, X. (2019), A description of the advanced research wrf model version 4, Technical report, NCAR UCAR.
- Tans, P. and Keeling, R. (2021), 'Trends in atmospheric carbon dioxide'. URL: https://www.esrl.noaa.gov/gmd/ccgg/trends/mlo.html
- Thomson, A. M., Calvin, K. V., Smith, S. J., Kyle, G. P., Volke, A., Patel, P., Delgado-Arias, S., Bond-Lamberty, B., Wise, M. A., Clarke, L. E. and Edmonds, J. A. (2011), 'RCP4.5: a pathway for stabilization of radiative forcing by 2100', *Climatic Change* 109(1-2), 77–94. URL: http://link.springer.com/10.1007/s10584-011-0151-4
- Tuckett, R. (2021), Chapter 2 greenhouse gases and the emerging climate emergency, in T. M. Letcher, ed., 'Climate Change (Third Edition)', third edition edn, Elsevier, pp. 19–45.

URL: https://www.sciencedirect.com/science/article/pii/B9780128215753000025

- van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J. and Rose, S. K. (2011), 'The representative concentration pathways: an overview', *Climatic Change* 109(1-2), 5–31.
 URL: http://link.springer.com/10.1007/s10584-011-0148-z
- Vara-Vela, A. (2013), Avaliação do impacto da mudança dos fatores de emissão veicular na formação de ozônio troposférico na Região Metropolitana de São Paulo (RMSP), Master dissertation, University of São Paulo.
- Visscher, A. D. (2014), Air Dispersion Modeling, John Wiley & Sons, Inc, Hoboken, NJ. URL: http://doi.wiley.com/10.1002/9781118723098
- von Schneidemesser, E., Monks, P. S., Allan, J. D., Bruhwiler, L., Forster, P., Fowler, D., Lauer, A., Morgan, W. T., Paasonen, P., Righi, M., Sindelarova, K. and Sutton, M. A. (2015), 'Chemistry and the Linkages between Air Quality and Climate Change', *Chemical Reviews* 115(10), 3856–3897.
 URL: http://pubs.acs.org/doi/10.1021/acs.chemrev.5b00089
- Wallace, J. M. and Hobbs, P. V. (2006), Atmospheric Science An Introductory Survey, elsevier edn, University of Washington, London.
- Wang, H., Fu, L., Lin, X., Zhou, Y. and Chen, J. (2009), 'A bottom-up methodology to estimate vehicle emissions for the beijing urban area', Science of The Total Environment 407(6), 1947–1953. URL: https://www.sciencedirect.com/science/article/pii/S0048969708011303
- Wang, W., Bruyere, C., Michael, D., Duhia, J., Dave, G., Kavulich, M., Werner, K., Chen, M., Hui-Chuan, L., Michalakes, J. and Rizvi, S. (2019), Weather Research & Forecasting Model, User's Guide, january 20 edn, National Center for Atmospheric Research.
- Warner, T. T. (2011), Numerical Weather and Climate Prediction, first edn, Cambridge University Press, New York.
- WHO (2006), 'Air quality guidelines global update 2005: particulate matter, ozone, nitrogen dioxide and sulfur dioxide'.
- Willmott, C. J. (1984), On the Evaluation of Model Performance in Physical Geography, Springer Netherlands, Dordrecht, pp. 443–460.
- WMO (2018), 'Guide to meteorological instruments and methods of observation'.
- Wolfe, G., Nicely, J., St.+Clair, J., Hanisco, T., Liao, J., Oman, L., Brune, W., Miller, D., Thames, A., González+Abad, G., Ryerson, T., Thompson, C., Peischl, J., McKain, K., Sweeney, C., Wennberg, P., Kim, M., Crounse, J., Hall, S., Ullmann, K., Diskin, G., Bui, P., Chang, C. and Dean-Day, J. (2019), 'Mapping hydroxyl variability throughout the global remote troposphere via synthesis of airborne and satellite formaldehyde observations', *PNAS; Proceedings of the National Academy of Sciences* **116**(23), 11171–11180.
- Zaveri, R. A. and Peters, L. K. (1999), 'A new lumped structure photochemical mechanism for large-scale applications', Journal of Geophysical Research: Atmospheres 104(D23), 30387–30415.
 URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/1999JD900876

Appendix A

WRF-Chem files

A.1 Namelist used to run the WRF-Chem model

A.1.1 NCEP FNL and RCP namelist.wps

WPS stage in WRF-Chem modeling requires a namelist.wps to prepare information related to geographical and meteorological data for a limited area for initial and lateral-boundary conditions. For current (Sep-Oct 2018) simulation, WPS processing required the follow namelist:

```
&share
 wrf_core = 'ARW',
 \max dom = 2.
                                                     ! Number of domains
 start_date = '2018-08-30_00:00:00','2018-08-30_00:00:00','2018-09-04_00:00:00'
end_date = '2018-10-31_00:00:00','2018-10-31_00:00:00','2018-09-14_00:00:00'
 interval_seconds = 21600
 io_form_geogrid = 2,
&geogrid
 parent_id
                           1.
                                   1. 2.
                                                    ! field information from grid resolution
 parent_grid_ratio = 1, 1, 1, 2,
parent_grid_ratio = 1, 5, 3,
i_parent_start = 1, 30, 43,
j_parent_start = 1, 20, 42,
                                                    ! ratio e.g. 1:5 to get high resolution
                                                    !
                       = 90, 151, 81,
                                                   ! (60-30)*5+1 = 2nd grid
 e_we
                       = 60, 121, 86,
                                                    ! (42-18)*5+1 = 2nd grid
 e_sn
                    = '30s', '30s', '30s',
 geog_data_res
 dx = 15000
dy = 15000
                                                       ! resolution in meters
                                                      ! resolution in meters
 map_proj = 'mercator',
ref_lat = -23.5700,
ref_lon = -46.6100,
truelat1 = -30.0000,
                                                     ! center of MASP
                                                     ! center of MASP
! truelat2 = -24.0000,
! stand_lon = -45.0000,
 geog_data_path = '/scr2/alejandro/WRF/DATA/GEOG'
&ungrib
 out_format = 'WPS',
 prefix = 'FILE',
&metgrid
 fg_name = 'FILE'
                          ! change this according to LBC name
 io_form_metgrid = 2,
```

For future simulation (Sep-Oct 2030), the namelist used is shown below. The only difference with the previous namelist is the line "fg_name" source name. For instance, CCSM4_CMIP5_MOAR_BC_RCP45 represents global projections for the RCP 4.5 emissions scenario.

```
&share
 wrf_core = 'ARW',
 max_dom = 2,
                                                    ! Number of domains
 mal_dom = 2,
start_date = '2030-08-30_00:00:00', '2030-08-30_00:00:00', '2018-09-04_00:00:00'
end_date = '2030-10-31_00:00:00', '2030-10-31_00:00:00', '2018-09-14_00:00:00'
 interval_seconds = 21600
 io_form_geogrid = 2,
&geogrid
 parent_id
                          1, 1, 2,
                                                    ! field information from grid resolution
                       =
 parent_grid_ratio = 1, 1, 1, 2,
parent_grid_ratio = 1, 5, 3,
i_parent_start = 1, 30, 43,
j_parent_start = 1, 20, 42,
                                                    ! ratio e.g. 1:5 to get high resolution
                                                    !
                       = 90, 151, 81,
                                                   ! (60-30)*5+1 = 2nd grid
 e_we
                                                   ! (42-18)*5+1 = 2nd grid
 e_sn
                      = 60, 121, 86,
                       = '30s', '30s', '30s',
 geog_data_res
 dx = 15000
                                                      ! resolution in meters
 dy = 15000
                                                     ! resolution in meters
 map_proj = 'mercator',
             = -23.5700,
                                                    ! center of RMSP
 ref_lat
 ref_lon = -46.6100,
truelat1 = -30.0000,
                                                    ! center of RMSP
! truelat2 = -24.0000,
! stand_lon = -45.0000,
 geog_data_path = '/scr2/alejandro/WRF/DATA/GEOG'
&ungrib
 out_format = 'WPS',
 prefix = 'FILE',
&metgrid
 fg_name = 'CCSM4_CMIP5_MOAR_BC_RCP45'
                                                  ! change this according to LBC name
 io_form_metgrid = 2,
```

A.1.2 NCEP FNL and RCP namelist.input

The first simulation period with two days of spin-up, the namelist.input changed in:

```
&time_control
io_form_auxinput12 = 0
&chem
chem_in_opt = 0, 0,
```

After that, the namelist.input for the next simulation period requires chemical initial conditions to continue the simulation, where wrf_chem_input must be linked with the wrfout first hours results of the period simulation. For instance, the follow namelist.input started in September 5 00:00 h, so the wrfout results of the first simulation must be correspond to that hour and linked to wrf_chem_input.

<pre>%time_control</pre>					
run_days	=	0,			
run_hours	=	0,			
run_minutes	=	0,			
run_seconds	=	0,			
start_year	=	2018,	2018,	2004,	2004
start_month	=	09,	09,	09,	09,
start_day	=	05,	05,	05,	05,
start_hour	=	00,	00,	00,	00,
start_minute	=	00,	00,	00,	00,
start_second	=	00,	00,	00,	00,
end_year	=	2018,	2018,	2004,	2004,
end_month	=	09,	09,	09,	09,
end_day	=	11,	11,	11,	11,
end_hour	=	00,	00,	06,	06,
end_minute	=	00,	00,	00,	00,
end_second	=	00,	00,	00,	00,
interval_seconds	=	21600			
input_from_file	=	.true	.,.tru	e.,.tru	1e.,.true.,
history_interval	=	60, 6	50, (30, 30),
frames_per_outfile	=	1, 1,	1000,	1000,	
restart	=	.fals	э.,		
restart_interval	=	7200,			
io_form_history	=	2			
io_form_restart	=	2			

APPENDIX A. WRF-CHEM FILES

<pre>io_form_input io_form_boundary io_form_auxinput4 io_form_auxinput2 io_form_auxinput6 io_form_auxinput5 io_form_auxinput12 auxinput5_interval_m ! auxinput6_interval_h auxinput1_inname auxinput1_inname auxinput12_inname debug_level force_use_old_data /</pre>	<pre>= 2 = 2 = 2 = 2 = 2 = 0 = 2 = 60, 60, 60, 60 = 24, 24, = "met_em.d<domain>.<date>" = 'wrfbiochemi_d<domain>' = 'wrf_chem_input' = 0 = .true.,</domain></date></domain></pre>	<pre>! biogenic ! biogenic ! biogenic ! re-initialization ! suggested by mgavidia</pre>
&domains		
<pre>time_step time_step_fract_num time_step_fract_den max_dom s_we e_we s_sn e_sn s_vert e_vert num_metgrid_levels num_metgrid_soil_levels dx dy grid_id parent_start j_parent_start j_parent_start parent_grid_ratio parent_time_step_ratio feedback smooth_option p_top_requested zap_close_levels interp_type t_extrap_type force_sfc_in_vinterp use_levels_below_ground use_surface lagrange_order sfcp_to_sfcp /</pre>	<pre>= 40, = 0, = 1, = 2, = 1, 1, 1, 1, 1, = 90, 151, 118, 100, = 1, 1, 1, 1, 1, = 60, 121, 118, 100, = 1, 1, 1, 1, 1, = 35, 35, 35, 35, = 32 = 4 = 15000, 3000, 3000, 1000, = 1, 2, 3, 4, = 1, 1, 2, 3, = 1, 20, 34, 33, = 1, 20, 34, 33, = 1, 5, 3, 3, = 1, 5, 3, 3, = 1, 5, 3, 3, = 0, = 0 = 5000 = 1 = 2 = 0 = .true. = 1 = .true.;</pre>	! 3600/(dx/1000*6) < integer
<pre>&physics mp_physics ra_lw_physics ra_sw_physics radt sf_sfclay_physics sf_surface_physics bl_pbl_physics ! topo_wind bldt cu_physics cudt isfflx ifsnow icloud surface_input_source num_soil_layers sf_urban_physics mp_zero_out mp_zero_out_thresh maxiens maxens2 maxens3 ensdim cu_rad_feedback / &fdda / &dynamics</pre>	<pre>= 10, 10, 2, 2, ! Morris = 1, 1, 1, 0, 0, = 1, 1, 1, 1, 1, ! RRTM = 4, 4, 2, 2, ! RRTMG = 15, 3, 15, 15, = 1, 1, 1, 1, 1, ! = 2, 2, 2, 2, 2, ! Noah 1 = 8, 8, 1, 1, ! boula = 2, 2, 2, 2, 2, ! to r = 0, 0, 0, 0, 0, = 5, 5, 5, 5, 0, ! Multi- = 0, 0, 0, 0, 0, = 1, = 1, = 1, = 1, = 4, = 1, ! Urban = 2, = 1.e=8, = 1, = 3, = 3, = 3, = 16, = 144, = .true.,</pre>	on double-moment scheme = 10 = 1 shortwave Land Surface Model = 2 c = 8 educe winds intensity scale Kain-Fritsch scheme = 11, GRELL 3D = 5 canopy model 3-category UCM = 1
<pre>«aynamics rk_ord w_damping</pre>	= 3, = 1,	

diff_opt	= 1,			
km_opt	= 4,			
damp opt	= 290.			
zdamp	= 5000.,	5000.,	5000.,	5000.,
dampcoef	= 0.01,	0.01,	0.01	, 0.01,
diff_6th_opt	= 0,			
diff_6th_factor	= 0.12,		<u>^</u>	•
Khdlf kwdif	= 0,	0,	0,	0,
non hydrostatic	= 0,	true	true.	, true
moist_adv_opt	= 2,	2,	2,	2,
scalar_adv_opt	= 2,	2,	2,	2,
chem_adv_opt	= 2,	2,	2,	2,
tke_adv_opt	= 2,	2,	2,	2,
time_step_sound	= 4,	4, 5	4, E	4, F
n_mom_adv_order	- 5, = 3.	3.	3.	3.
h_sca_adv_order	= 5,	5,	5,	5,
v_sca_adv_order	= 3,	З,	з,	3,
hybrid_opt	= 0,			
use_theta_m	= 0,			
/				
&bdy_control	-			
spec_bdy_width	= 5,			
spec_zone	= 1, = 4			
specified	- 4, = .trne.	.false	false	false
periodic_x	= .false.	,.false.	,.false	.,.false.,
- symmetric_xs	= .false.	,.false.	,.false	.,.false.,
symmetric_xe	= .false.	,.false.	,.false	.,.false.,
open_xs	= .false.	,.false.	,.false	.,.false.,
open_xe	= .false.	,.false.	, false.	.,.false.,
symmetric vs	= .false.	, false.	, false	.,.false.,
symmetric_ye	= .false.	,.false.	,.false	.,.false.,
open_ys	= .false.	,.false.	,.false	.,.false.,
open_ye	= .false.	,.false.	,.false	.,.false.,
nested	= .false.	, .true.	, .true	.,.true.,
/				
/ &grib2 /				
/ &grib2 / &chem				
/ &grib2 / &chem kemit	= 1,		!	Number of vertical levels
/ &grib2 / &chem kemit chem_opt	= 1, = 6,	6,	!	Number of vertical levels CBMZ chemical mechanism without DMS = 6
/ &grib2 / &chem kemit chem_opt bioemdt	= 1, = 6, = 15,	6, 15,	! ! !	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes
/ &grib2 / &chem kemit chem_opt bioemdt photdt chemdt	= 1, = 6, = 15, = 15, = 2	6, 15, 15, 2	!	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes
<pre>/ &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions</pre>	= 1, = 6, = 15, = 15, = 2, = 1,	6, 15, 15, 2,	! ! !	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes
<pre>/ &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt</pre>	= 1, = 6, = 15, = 15, = 2, = 1, = 102,	6, 15, 15, 2, 102,	:	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt</pre>	= 1, = 6, = 15, = 15, = 2, = 1, = 102, = 4,	6, 15, 15, 2, 102, 4,	! ! !	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt </pre>	= 1, = 6, = 15, = 15, = 2, = 1, = 102, = 4, = 1,	6, 15, 15, 2, 102, 4, 1,	! ! !	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gras_dryden_opt </pre>	= 1, = 6, = 15, = 15, = 2, = 1, = 102, = 4, = 1, = 2, = 1	6, 15, 15, 2, 102, 4, 1, 2,	! ! !	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt aer drvdep opt </pre>	= 1, = 6, = 15, = 15, = 2, = 1, = 102, = 4, = 1, = 2, = 1, = 1,	6, 15, 15, 2, 102, 4, 1, 2, 1,	! ! !	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt aer_drydep_opt bio_emiss_opt</pre>	= 1, = 6, = 15, = 15, = 2, = 1, = 102, = 4, = 1, = 2, = 1, = 1, = 3,	6, 15, 15, 2, 102, 4, 1, 2, 1, 1,	!	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt aer_drydep_opt bio_emiss_opt ne_area</pre>	= 1, = 6, = 15, = 15, = 2, = 1, = 102, = 4, = 1, = 2, = 1, = 3, = 70,	6, 15, 15, 2, 102, 4, 1, 2, 1, 1,	!	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt aer_drydep_opt bio_emiss_opt ne_area dust_opt</pre>	= 1, = 6, = 15, = 2, = 1, = 102, = 4, = 1, = 2, = 1, = 1, = 3, = 70, = 0,	6, 15, 15, 2, 102, 4, 1, 2, 1, 1,	1 1 1	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt aer_drydep_opt bio_emiss_opt ne_area dust_opt dmsemis_opt </pre>	= 1, = 6, = 15, = 2, = 1, = 102, = 4, = 1, = 2, = 1, = 1, = 3, = 70, = 0, = 0,	6, 15, 15, 2, 102, 4, 1, 2, 1, 1,	1 1 1	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt aer_drydep_opt bio_emiss_opt ne_area dust_opt dmsemis_opt seas_opt gas_opt gas_opt</pre>	= 1, = 6, = 15, = 2, = 1, = 102, = 4, = 1, = 2, = 1, = 1, = 2, = 1, = 0, = 0, = 0, = 1	6, 15, 15, 2, 102, 4, 1, 2, 1, 1,	8 9 8	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt bio_emiss_opt ne_area dust_opt dmsemis_opt gas_bc_opt gas_bc_opt gas_bc_opt </pre>	= 1, = 6, = 15, = 15, = 1, = 1, = 102, = 4, = 1, = 2, = 1, = 1, = 3, = 70, = 0, = 0, = 1, = 1,	6, 15, 15, 2, 102, 4, 1, 2, 1, 1, 1,	! ! !	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt aer_drydep_opt bio_emiss_opt ne_area dust_opt dmsemis_opt seas_opt gas_bc_opt gas_ic_opt aer_bc_opt </pre>	= 1, = 6, = 15, = 15, = 1, = 1, = 1, = 1, = 2, = 1, = 1, = 0, = 0, = 0, = 1, = 1, = 1,	6, 15, 15, 2, 102, 4, 1, 1, 1, 1,	! ! !	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt aer_drydep_opt bio_emiss_opt ne_area dust_opt dmsemis_opt seas_opt gas_bc_opt gas_ic_opt aer_ic_opt aer_ic_opt </pre>	= 1, = 6, = 15, = 15, = 1, = 1, = 1, = 1, = 1, = 1, = 1, = 0, = 0, = 0, = 1, = 1, = 1, = 1, = 1, = 1, = 1, = 1	6, 15, 15, 2, 102, 4, 1, 2, 1, 1, 1, 1,	! ! !	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt aer_drydep_opt bio_emiss_opt ne_area dust_opt dmsemis_opt seas_opt gas_bc_opt gas_ic_opt aer_ic_opt gaschem_onoff</pre>	= 1, = 6, = 15, = 2, = 1, = 102, = 4, = 1, = 1, = 1, = 3, = 70, = 0, = 0, = 1, = 1, = 1, = 1, = 1, = 1, = 1, = 1	6, 15, 15, 2, 102, 4, 1, 2, 1, 1, 1, 1, 1, 1,	! ! !	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt bio_emiss_opt ne_area dust_opt dmsemis_opt seas_opt gas_bc_opt gas_ic_opt aer_bc_opt aer_bc_opt gasche_opt aer_ic_opt gasche_opt gasche_opt aer_ic_</pre>	= 1, = 6, = 15, = 12, = 1, = 1, = 1, = 3, = 70, = 0, = 1, = 1, = 1, = 1, = 1, = 1, = 3, = 70, = 0, = 1, = 0, = 0, = 0, = 0, = 0, = 0, = 0, = 1, = 1, = 0, = 0, = 0, = 1, = 1, = 1, = 1, = 1, = 0, = 0, = 1, = 1,	6, 15, 15, 2, 102, 4, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1,	! ! !	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt aer_drydep_opt bio_emiss_opt ne_area dust_opt dmsemis_opt gas_bc_opt gas_ic_opt aer_ic_opt aer_ic_opt gaschem_onoff eldebm_emiff</pre>	= 1, = 6, = 15, = 12, = 1, = 1, = 1, = 1, = 70, = 0, = 0, = 1, = 1, = 1, = 1, = 1, = 0, = 0, = 0, = 1, = 0, = 0, = 0, = 1, = 1, = 1, = 0, = 0, = 1, = 0, =	6, 15, 15, 2, 102, 4, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 0,	1 1 1 1	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt aer_drydep_opt dmsemis_opt me_area dust_opt dmsemis_opt gas_bc_opt gas_bc_opt gas_ic_opt aer_bc_opt gascav_onoff vetscav_onoff cldchem_onoff </pre>	= 1, = 6, = 15, = 2, = 1, = 102, = 4, = 1, = 2, = 1, = 2, = 1, = 3, = 70, = 0, = 1, = 0, = 0, = 0, = 1, = 1,	6, 15, 15, 2, 102, 4, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1	1 1 1	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt aer_drydep_opt bio_emiss_opt me_area dust_opt dmsemis_opt gas_bc_opt gas_ic_opt aer_bc_opt gasclopt aer_bc_opt gaschem_onoff vetscav_onoff cldchem_onoff chem conv tr</pre>	= 1, = 6, = 15, = 15, = 2, = 1, = 102, = 4, = 1, = 2, = 1, = 2, = 1, = 2, = 1, = 0, = 0, = 1, = 1, = 1, = 1, = 1, = 102, = 1, = 1, = 102, = 2, = 1, = 102, = 2, = 1, = 102, = 1, = 1, = 102, = 2, = 1, = 1, = 1, = 1, = 1, = 1, = 1, = 1	6, 15, 15, 2, 102, 4, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,		Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt bio_emiss_opt ne_area dust_opt dmsemis_opt gas_bc_opt gas_ic_opt aer_bc_opt aer_ic_opt gaschem_onoff cldchem_onoff chem_conv_tr biomass_burn_opt</pre>	= 1, = 6, = 15, = 15, = 2, = 1, = 102, = 4, = 1, = 2, = 1, = 2, = 1, = 2, = 1, = 0, = 0, = 1, = 1, = 1, = 1, = 1, = 102, = 4, = 1, = 102, = 4, = 1, = 102, = 4, = 1, = 102, = 4, = 1, = 102, = 1, = 1, = 102, = 1, = 1, = 1, = 1, = 1, = 1, = 1, = 1	6, 15, 15, 2, 102, 4, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,		Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt aer_drydep_opt bio_emiss_opt ne_area dust_opt dmsemis_opt gas_bc_opt gas_ic_opt aer_ic_opt aer_ic_opt aer_ic_opt chem_onoff chem_conv_tr biomass_burn_opt plumerisefire_frq</pre>	= 1, = 6, = 15, = 12, = 1, = 102, = 4, = 1, = 2, = 1, = 2, = 1, = 2, = 1, = 0, = 0, = 0, = 1, = 1, = 1, = 1, = 1, = 102, = 4, = 1, = 102, = 4, = 1, = 102, = 4, = 1, = 102, = 4, = 1, = 1, = 1, = 1, = 1, = 1, = 1, = 1	6, 15, 15, 2, 102, 4, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 30,		Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt bio_emiss_opt ne_area dust_opt dmsemis_opt gas_bc_opt gas_bc_opt aer_ic_opt aer_ic_opt aer_ic_opt aer_ic_opt chem_onoff vetscav_onoff chem_conv_tr biomass_burn_opt plumerisefire_frq aer_ra_feedback</pre>	= 1, = 6, = 15, = 1, = 1, = 1, = 2, = 1, = 2, = 1, = 2, = 1, = 2, = 1, = 2, = 1, = 0, = 0, = 1, = 1, = 1, = 2, = 1, = 2, = 1, = 2, = 1, = 0, = 1, = 1, = 0, = 1, = 1, = 0, = 1, = 1, = 0, = 1, = 1, = 1, = 0, = 1, = 1, = 1, = 0, = 1, = 1, = 1, = 1, = 0, = 1, = 0, = 0, = 0, = 0, = 0, = 0, = 0, = 0, = 0, = 1, = 1, = 1, = 0, = 0	6, 15, 15, 2, 102, 4, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 30, 0,		Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt bio_emiss_opt ne_area dust_opt dmsemis_opt gas_bc_opt gas_bc_opt gas_ic_opt aer_ic_opt aer_ic_opt aer_ic_opt chem_onoff vetscav_onoff chem_conv_tr biomass_burn_opt plumerisefire_frq aer_ra_feedback have_bcs_chem /</pre>	<pre>= 1, = 6, = 15, = 15, = 2, = 1, = 102, = 4, = 1, = 2, = 1, = 1, = 2, = 1, = 0, = 0, = 0, = 1, = 1, = 1, = 1, = 1, = 1, = 1, = 2, = 1, = 2, = 1, = 1, = 102, = 4, = 1, = 102, = 4, = 1, = 1, = 1, = 1, = 1, = 2, = 1, = 1, = 1, = 2, = 1, = 1, = 1, = 2, = 1, = 1, = 2, = 1, = 1, = 1, = 2, = 1, = 0, = 0, = 0, = 0, = 0, = 1, = 1, = 1, = 0, = 0, = 0, = 0, = 1, = 1, = 1, = 0, = 0, = 0, = 1, = 1, = 1, = 1, = 0, = 0, = 0, = 1, = 1, = 1, = 1, = 0, = 0, = 0, = 1, = 1, = 1, = 1, = 1, = 1, = 1, = 1</pre>	6, 15, 15, 2, 102, 4, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 30, 0, 1, 1,31se	.,	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>/ &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt aer_drydep_opt bio_emiss_opt ne_area dust_opt dmsemis_opt gas_bc_opt gas_ic_opt aer_ic_opt aer_ic_opt gaschem_onoff chem_convtr biomass_burn_opt plumerisefire_frq aer_ra_feedback have_bcs_chem /</pre>	= 1, = 6, = 15, = 2, = 1, = 102, = 4, = 1, = 2, = 1, = 1, = 2, = 1, = 1, = 0, = 0, = 1, = 1, = 1, = 1, = 1, = 1, = 1, = 1	6, 15, 15, 2, 102, 4, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,	! ! !	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt aer_drydep_opt bio_emiss_opt ne_area dust_opt dmsemis_opt gas_bc_opt gas_bc_opt aer_ic_opt aer_ic_opt gaschem_onoff chem_convtr biomass_burn_opt plumerisefire_frq aer_ra_feedback have_bcs_chem / &namelist_quilt</pre>	= 1, = 6, = 15, = 2, = 1, = 102, = 4, = 1, = 2, = 1, = 1, = 2, = 1, = 1, = 0, = 0, = 1, = 1, = 1, = 1, = 1, = 1, = 1, = 1	6, 15, 15, 2, 102, 4, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,	.,	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt aer_drydep_opt bio_emiss_opt ne_area dust_opt dmsemis_opt gas_bc_opt gas_ic_opt aer_ic_opt aer_ic_opt aer_ic_opt chem_onoff cldchem_onoff cldchem_onoff chem_convtr biomass_burn_opt plumerisefire_frq aer_ra_feedback have_bcs_chem / &namelist_quilt nio_tasks_per_group = 0, </pre>	= 1, = 6, = 15, = 2, = 1, = 102, = 4, = 1, = 2, = 1, = 1, = 2, = 1, = 1, = 0, = 0, = 1, = 1, = 1, = 1, = 1, = 1, = 1, = 1	6, 15, 15, 2, 102, 4, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,	! ! !	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species
<pre>% &grib2 / &chem kemit chem_opt bioemdt photdt chemdt io_style_emissions emiss_inpt_opt emiss_opt chem_in_opt phot_opt gas_drydep_opt bio_emiss_opt ne_area dust_opt gas_bc_opt gas_ic_opt gas_copt gaschem_onoff aerchem_onoff cldchem_onoff cldchem_onoff chem_conv_tr biomass_burn_opt plumerisefire_frq aer_iftedback have_bcs_chem / &namelist_quilt nio_tasks_per_group = 0, nio_groups = 1, /</pre>	= 1, = 6, = 15, = 15, = 2, = 1, = 102, = 4, = 1, = 2, = 1, = 1, = 2, = 1, = 1, = 0, = 0, = 0, = 1, = 1, = 1, = 1, = 1, = 1, = 1, = 1	6, 15, 15, 2, 102, 4, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,	! ! !	Number of vertical levels CBMZ chemical mechanism without DMS = 6 biogenic in minutes ! MEGAN 2 ! MEGAN 2 No of chemical species



A.2 Anthropogenic Emissions Calculation

A.2.1 EDGAR-HTAP processing for industry and residential sectors

EDGAR-HTAP has monthly emissions datasets (yearly and monthly files) for different sectors (e.g., industry, residential), available in its website. According to it, datasets contain 0.1×0.1 deg gridmaps of CH₄, CO, SO₂, NO_x, NMVOC, NH₃, PM₁₀, PM_{2.5}, BC and OC for the years 2008 and 2010. These datasets "use nationally



Figure A.1 VOC fractions

reported emissions combined with regional scientific inventories in the format of sector-specific gridmaps. The gridmaps are complemented with EDGARv4.3 data for those regions where data are absent." However, these datasets don't contain specific emissions about NMVOC speciation (e.g., ethane, toluene, aldehydes, formaldehydes, etc.) needed for the CBM-Z mechanism in the WRF-Chem for photochemical reactions (i.e., ozone formation). Before running the anthro_emiss for each sector, the NMVOC emission files were converted to different VOC species based on relative fractions of VOC emissions from different processes and fuels (Figure 6 in Andrade et al. 2015). Procedures to obtain speciation emission files from NMVOC are as follow:

- voc_frac.py uses information from LAPAt emission preprocessor for road transport (wrfchemi_00z_d01_veic) to generate a voc_frac_round.csv (Figure A.1).
- htap2AE_year.py uses the voc_frac_round.csv file and NMVOC emission files from EDGAR-HTAP to generate VOC speciation emission files.

WRF-Chem emission files (wrfchemi_<hh>z_<domain>) for industry and residential are generated through running ANTHRO EMISS in the Linux terminal:

./anthro_emiss < cbmz.inp > cbmz.out

For instance, the cbmz.inp namelist contents the follow information for the industry sector:

&CONTROL
anthro_dir = '/scr2/alejandro/WRF/DATA/util/EDGAR-HTAP/industry'
wrf_dir = '/scr2/alejandro/WRF/SEP18/WRF'
domains = 2
src_file_prefix = 'edgar_HTAP_'
src_file_suffix = '_emi_INDUSTRY_2010.0.1x0.1_AE.nc'

Python scripts (voc_frac.py and htap2AE_year.py) can get on this GitHub.

A.2.2 LAPAt emission preprocessor for road transport

LAPAt preprocessor emission model used in Andrade et al. (2015) contents a namelist (namelist_fc.emi) and a NCL script (wrfchemi_cbmz_fc.ncl) developed by researchers of IAG as a utility for generating ready emission files for WRF-Chem. The NCL script (wrfchemi_cbmz_fc.ncl) calculates road transport emission based through the Bottom-Up methodology based on data files called in the namelist_fc.emi. For instance, the follow namelist_fc.emi applies to the parent domain with 15 km × 15 km for September 2018:

```
&arquivos_necessarios
path_wrfinput = ./wrfinput_d01.nc
path_map_vias = ./grid15km_d01.txt
             = ./voc_split_cbmz.txt
path voc
&caracteristicas_grade
                          ! about grid roads, NOT of the model domain area
                          ! NUMERO DE PONTOS EM X (WEST-EAST)
             = 103,
nx
             = 63,
                          ! NUMERO DE PONTOS EM Y (SOUTH-NORTH)
ny
             = 15,
                           ! ESPACAMENTO DE GRADE EM X (KM)
dx
              = 15,
                           ! ESPACAMENTO DE GRADE EM Y (KM)
dy
1
&caracteristicas_frota
                          ! setembro 2018
            = 9.
                          ! NUMERO DE TIPOS DE VEICULO
n veic
veic_gasolina = 0.2089720, ! FRACAO VEICULOS MOVIDOS A GASOLINA (VEIC 1)
             = 0.0146304,
                          ! FRACAO VEICULOS MOVIDOS A ETANOL (VEIC 2)
veic_etanol
veic_flex
             = 0.5477388,
                         ! FRACAO VEICULOS MOVIDOS A FLEX (VEIC 3)
veic_caminhoes = 0.0566967, ! FRACAO CAMINHOES (DIESEL - VEIC 4A)
veic urbanos
             = 0.0050713, ! FRACAO ONIBUS URBANO (DIESEL - VEIC 4B)
= 0.0000000, ! FRACAO TAXIS (GAS - VEIC 5)
veic_taxis
veic_moto_flex = 0.0455956,
                          ! FRACAO MOTOS FLEX (VEIC 6B)
                          ! FRACAO FROTA ATIVA VEICULOS LEVES (1=100%)
frota_ativa
              = 1,
1
```

[!] Fatores de emissao baseados em Perez-Martinez et al. (2014) e Andrade et al. (2015, 2018) &fator_emissao ! VEIC 1, VEIC 2, VEIC 3, VEIC4A, VEIC4B, VEIC4C, VEIC 5, VEIC6A, VEIC6B

exa_co	= 4.4000,	12.000,	4.4000,	1.5000,	1.5000,	1.5000,	0.0000,	9.1500,	9.0200,
exa_co2	= 219.00,	219.00,	219.00,	1422.0,	1422.0,	1422.0,	0.0000,	0.0000,	0.0000,
exa_nox	= 0.2000,	1.1200,	0.2000,	6.9000,	6.9000,	6.9000,	0.0000,	0.1320,	0.1290,
exa_so2	= 0.0290,	0.0140,	0.0210,	0.6100,	0.6100,	0.6100,	0.0000,	0.0097,	0.0093,
exa_c2h5oh	= 0.5080,	0.2500,	0.5080,	0.6100,	0.6100,	0.6100,	0.0000,	0.0790,	0.3050,
exa_hcho	= 0.0089,	0.0110,	0.0098,	0.6100,	0.6100,	0.6100,	0.0000,	0.0152,	0.0155,
exa_ald	= 0.0140,	0.0300,	0.0220,	0.6100,	0.6100,	0.6100,	0.0000,	0.0164,	0.0188,
exa_pm	= 0.0200,	0.0200,	0.0200,	0.2770,	0.2770,	0.2770,	0.0000,	0.0500,	0.0500,
exa_voc	= 0.4250,	1.3000,	0.4340,	2.0500,	2.0500,	2.0500,	0.0000,	1.0800,	1.0800,
vap_voc	= 0.2300,	0.2500,	0.2400,	0.0000,	0.0000,	0.0000,	0.0000,	0.0000,	0.0000,
liq_voc	= 2.0000,	1.5000,	1.7500,	0.0000,	0.0000,	0.0000,	0.0000,	1.2000,	1.2000,
/									
&fator_correcao		! Mod	lified by	Alejandro	o Delgado				
fc_co	= 1,	! Emi	ssion con	rection f	factor				
fc_co2	= 1,								
fc_no	= 0.8,								
fc_no2	= 0.8,								
fc_so2	= 1,								
fc_c2h5oh	= 1,								
fc_hcho	= 1,								
fc_ald	= 1,								
fc_pm	= 1,								
fc_voc	= 1,								
&intensidade_uso		! y	ear 2018	calculate	ed using a	a Python s	cript		
kmd_veic123	= 39.39,	! (UILOMETRA	AGEM DIARI	IA - VEICU	JLOS 1, 2	Е З		
kmd_veic4a	= 91.02,	! (UILOMETRA	AGEM DIARI	IA - VEICU	JLOS 4A			
kmd_veic4b	= 86.70,	! (UILOMETRA	AGEM DIARI	IA - VEICU	JLOS 4B			
kmd_veic4c	= 106.18,	! (UILOMETRA	AGEM DIARI	IA - VEICU	JLOS 4C			
kmd_veic5	= 0.,	! (UILOMETRA	AGEM DIARI	IA - VEICU	JLOS 5			
kmd_veic6a	= 32.16,	! (UILOMETRA	AGEM DIARI	IA - VEICU	JLOS 6A			
kmd_veic6b	= 35.78,	! (UILOMETRA	AGEM DIARI	IA - VEICU	JLOS 6B			
/									
&perfil_diario	! HURA UTC	used in	Andrade e	et al. (20)15)			0 074	
nrspit_co	= 0.019, 0	.012, 0.0	108, 0.004	i, 0.003,	0.003, 0.	006, 0.01	7, 0.047,	0.074, 0	0.072, 0.064,
h1+	0.055, 0	0.052, 0.0	10 0.048	3, 0.052,	0.057, 0.	015 0.08	0.085,	0.057, 0	0.035, 0.024,
hrspit_no	= 0.019, 0	.015, 0.0), 0.000,	0.009, 0.	015, 0.03	0, 0.040,	0.053, 0	0.051, 0.061,
/	0.004, 0		01, 0.000), 0.000,	0.005, 0.	.005, 0.00	0, 0.050,	0.044, (.035, 0.027,
/									
&fracionamento as	rossois								
f fina total	= 0.670	I FRACA	0 FINA TO	TAT.					
f fina so4	- 0.070	I EDACA	O ETNA DE	2 904					

f_fina_so4	= 0.070,	! FRACAO FINA DE SO4
f_fina_no3	= 0.016,	! FRACAO FINA DE NO3
f_fina_org	= 0.420,	! FRACAO FINA DE ORG
f_fina_ec	= 0.190,	! FRACAO FINA DE EC
f_fina_pm25	= 0.304,	! FRACAO FINA DE PM25
f_grossa_total	= 0.330,	! FRACAO GROSSA TOTAL
f_grossa_so4	= 0.000,	! FRACAO GROSSA DE SO4
f_grossa_no3	= 0.000,	! FRACAO GROSSA DE NO3
f_grossa_org	= 0.000,	! FRACAO GROSSA DE ORG
f_grossa_ec	= 0.000,	! FRACAO GROSSA DE EC
i_so4	= 0.136,	! MODA AITKEN DE SO4
j_so4	= 0.864,	! MODA ACUMULACAO DE SO4
i_no3	= 0.230,	! MODA AITKEN DE NO3
j_no3	= 0.770,	! MODA ACUMULACAO DE NO3

i_org	= 0.190,	! MODA AITKEN DE ORG
j_org	= 0.810,	! MODA ACUMULACAO DE ORG
i_ec	= 0.940,	! MODA AITKEN DE EC
j_ec	= 0.060,	! MODA ACUMULACAO DE EC
i_pm25	= 0.250,	! MODA AITKEN DE PM25
j_pm25	= 0.750,	! MODA ACUMULACAO DE PM25
/		

LAPAt pre-processor emissions model created WRF-Chem ready emissions files for two modeling domains, considering the following structure: wrfchemi_{00z, 12z}_{d01, d02}_veic, using the follow code in the Linux terminal of the Master IAG server (svan2):

ncl wrfchemi_cbmz_fc.ncl

The total number of vehicle approach for each modeling domain area was based on the spatial data sets of Brazil (available in geobr Python package) and information about vehicle types by each municipality downloaded from DENATRAN. A Python script '01_Vehicles.py' was used to calculate the total number of vehicles and fraction by type and fuel consumption, available in this GitHub.

The primary data files are grid15km_d01.txt and grid03km_d02.txt, which are the results from QGIS and Python processing, shown below (next page). These datasets represent the total road length as the sum for each grid cell based on the motorway, trunk, primary, secondary, and tertiary types.

Create a Road Length Grid for the Modeling Domain using **QGIS 3.14**

September 4, 2020

- 1. Merge all shapefile roads downloaded from http://download.geofabrik.de/ according to the extent of the modelling domain
 - a. Vector/Data Management Tools.../Merge Vector Layers...
 - i. Save as all roads.shp (Figure 1).
 - b. Select in Layer Properties/Settings the coordinate reference as EPSG:4326 WGS 84 (Figure 1 and rectangle in red).



Figure 1 Layer Properties/Settings and selection of the coordinate system.

2. Create a vector grid of 15 km and 3 km of spatial resolution:

- a. Vector/Research Tools.../Create Grid...b. Grid type: Rectangle (Polygon)
- Grid extent for modelling domain: C.
 - i. First modeling domain (d01): -53.53, -39.69, -27.70, -19.30. The first two values are coordinates in longitude; last two values are in latitude.
 - ii. Second modeling domain (d02): -49.01, -44.36, -25.05, -21.64. The MASP is at the center of the grid.
- d. Horizontal spacing / Vertical spacing:
 - i. Domain d01: 15 km ~ 0.135135 degrees.
 - ii. Domain d02: 3 km ~ 0.027027 degrees.
- e. Horizontal overlay / Vertical overlay: 0
- Save as 1 Grid15km d01, 1 Grid3km d02 as shapefile, where d01 and d02 are <domain>, and 15km and 3km <res>. Take a few seconds to run this process.

3. Extract vectors based on 1 Grid3km <domain>

- a. Vector/Geoprocessing Tools/Intersection...
- b. Input Laver: all roads
- Overlay layer: 1 Grid<res> < <domain> C.
- Intersection: Saved as 2 GridDomain<res> <domain>.shp. Take many minutes to run d. this process for high resolution (3 km).

4. In 2 GridDomain<res> <domain>:

- a. Open Attribute Table
- Remove columns as "ref", "oneway", "bridge", "tunnel", "maxspeed" b.
- Open Calculator Field (Figure 2): С
 - i. Output field name: kmlen
 - ii. Output field: Decimal number (real), length: 10, precision: 5.
 - iii. Expression: \$length*111. The value 111 is in km, and it is equivalent to 1 degree for São Paulo state. The result is the length of each road in km. So, it is important to consider five decimals in the values.
 - iv. Save.
 - v. Don't close yet the window, proceed with the next step.
- 5. In 2 GridDomain<res> <domain>, create with Calculator Field "vehroads" that represents the main roads where the vehicle flow is probably and frequently (Major

Roads and Highway links). According to https://www.geofabrik.de/data/geofabrik-osm-gisstandard-0.7.pdf, we have roads classified as Major Roads, Minor Roads, Highway links (sliproads/ramps), Very small roads, Paths unsuitable for cars, Ferries, and Unknown. Not all of these types of ways have vehicle flows.

- a. Output field name: vehroadsb. Output field: decimal number (real), length 10, precision 5.
- c. In "Expression" (Figure 2):

case			
when	fclass	is	'motorway' then kmlen
when	fclass	is	'motorway_link' then kmlen
when	fclass	is	'trunk' then kmlen
when	fclass	is	'trunk_link' then kmlen
when	fclass	is	'primary' then kmlen
when	fclass	is	'primary_link' then kmlen
when	fclass	is	'secondary' then kmlen
when	fclass	is	'secondary_link' then kmlen
when	fclass	is	'tertiary' then kmlen
when	fclass	is	'tertiary_link' then kmlen
else	0		
end			

d. Close Table Attributes and save.

(

ight Open fie	eld calculator (Ctrl+I) r
2_GridDoma	ain15km — Field Calculator
 Only update 0 selected features Create a new field Create virtual field Output field name vehroads Output field type Decimal number (real) 	Update existing field
Output field length 10 + Precision 5	
Expression Function Editor	
<pre>case when fclass is 'motorway' then kmlen when fclass is 'inotorway' then kmlen when fclass is 'trunk' then kmlen when fclass is 'trunk' then kmlen when fclass is 'primary' then kmlen when fclass is 'primary' then kmlen when fclass is 'secondary' is 'secondary_link' then kmlen when fclass is 'tertiary' = + - / + ^ () \n' Feature 29011719 Preview: 28686.81878</pre>	Show Helprow_numberAggregatesArraysColorConditionalsConditionalsConversionsDate and TimeFields and P.Fuzzy MatcGeometryMap LayersMathOperatorsRastersRecord andStringVariablesRecent (fiel
You are editing information on this OK, edit mode will automatically be Help	layer but the layer is currently not in edit mode. If you click e turned on. Cancel OK

Figure 2 Creation of 'vehroads' using Field Calculator.

6. In Vector General/Join Attributes by Location (summary). See Figure 3.



Figure 3 Join Attributes by Location (summary). In this example, a resolution of 15 km was processed; similarly, this procedure applies for 3 or 9 km grid cell of spatial resolution.

- a. Input layer: 1_Grid<res>_<domain>
- b. Join layer: 2_GridDomain<res>_<domain>
- c. Geometric predicate: intersects
- d. Summaries to calculate: sum
- e. Joined layer: 3_roadstypegrid_<domain>.shp. This process could take many hours, ensure to select only the 'vehroads' and 'id' columns to join, with this we can avoid delays in the process. If you do not select only 'vehroads' and 'id', you can wait days to complete the process for the modelling domain. After you can see the results as shown in Figure 4.
- f. Open Attribute Table
 - i. Remove columns: ID_sum, ... sum, except vehroads_s
 - ii. Update vehroads_s: case when vehroads_s is NULL then 0 else vehroads_s end
 - iii. Save and exit of edition mode.



Figure 4 Road length (in km) for each 3 x 3 km grid cell for 'vehroads' field. This figure was created using QGIS 3.14, selecting Properties/Graduated/Mode: Natural Breaks (Jenks) in the layer 3_roadstypegrid_d02.shp.

7. In Vector/Geometry Tools/Centroids...

- a. 3_roadstypegrid_<domain>
- b. Save as 4_final<res>grid_<domain>

8. Open Attribute Table of 4_final<res>grid_<domain>:

- a. Remove columns except for vehroads_s and id
- b. Open Field Calculator and create two fields:
 - i. Ylat (decimal, length 10, and precision 2). In expression, write \$y.
 - ii. Xlon (decimal, length 10, and precision 2). In expression, write \$x.
 - iii. Save
- 9. Right-click in 4_final3kmgrid_d02 and select Export, then save as csv file a. Name: grid<res>_<domain>.csv
- 10. Calculate nx and ny based on the number of latitude and longitude values.
 - a. In order to do this, run the Python codes in Jupyter Notebook or using the following codes as a script in Python 3:

```
import pandas as pd
import numpy as np
path = '../2_Emissions_inventory/roads/'
df3 = pd.read_csv(path+'grid03km_d02.csv')
df15 = pd.read_csv(path+'grid15km.csv')
print("nx: ", df3.Xlon.nunique())
print("ny: ", df3.Ylat.nunique())
print("nx: ", df15.Xlon.nunique())
print("ny: ", df15.Ylat.nunique())
def order(df=df3):
      df.loc[:,'id']=range(len(df.id))
df = df[['id','Xlon','Ylat','vehroads_s']]
df = df.round({"id":0,"Xlon":5, "Ylat":5, "vehroads_s":3})
df = df.sort_values(by=['Ylat','Xlon'],ascending=[False,True])
df.loc[:,'id']=np.arange(0.0,len(df.id))
      return df
df3 = order(); df15 = order(df=df15)
def vias(fname = 'grid03km_d02.txt',file = df3):
      np.savetxt(path+fname, file, delimiter='
                 fmt= ['%12.0f', '%5.5f', '%14.5f', '%15.2f'])
vias()
vias(fname='grid15km final.txt', file=df15)
print("Successfully")
```

A.2.3 Biogenic emissions

The megan_bio_emiss utility reads transformed MEGAN biogenic input files and create wrfbiochemi_d<nn> files needed to run the WRF-Chem model. All the executables can be created with the make_util script. MEGAN version 2 can be run using the follow linecode in the Linux terminal:

./megan_bio_emiss < megan_bio_emiss.inp</pre>

The megan_bio_emiss.inp file for this study contents the follow information and cover September and October:

```
&control
domains = 2,
start_lai_mnth = 7,
end_lai_mnth = 11,
wrf_dir = '/scr2/alejandro/WRF/Y2018/WRF'
megan_dir = '/scr2/alejandro/WRF/DATA/util/MEGAN/data'
/
```

Two biogenic emission files were created (wrfbiochemi_d01 and wrfbiochemi_d02). For instance, Figure A.2 shows information for the parent domain (wrfbiochemi_d01).



Figure A.2 Spatial distribution of MEGAN version 2 in the 15 km parent modeling domain for September.

Appendix B

Air Quality and Meteorological Information

There are several air quality and meteorological stations managed by CETESB for the São Paulo state. The parameters were downloaded from QUALAR or the next url: https://qualar.cetesb.sp.gov.br/ and correspond to:

- Meteorological parameters
 - Temperature at 2 m above ground $[^{o}C]$
 - Relative humidity at 2 m above ground [%]
 - Solar radiation $[W/m^2]$
 - Wind speed at 10 m above ground [m/s]
 - Wind direction at 10 m above ground [degrees]
- Air quality parameters
 - Surface ozone concentration $[\mu g m^{-3}]$
 - Nitrogen monoxide concentration $[\mu g m^{-3}]$
 - Nitrogen dioxide concentration $[\mu g m^{-3}]$
 - Carbon monoxide concentration [ppm]

For September and October, air quality and meteorological parameters were downloaded from CETESB, considering all stations with hourly data available for five years (2014-2018). For other months, only ten stations were downloaded in order to recognize and justify which month recorded high surface ozone concentrations. Figure B.1 shows monthly distribution of hourly data by station type, and covers five years as time series. These stations represent different land use types and they were named in this study as "Industry", "Regional urban", "Urban park", "Urban", and "Forest preservation". Not all station types are inside the MASP (i.e., Regional urban and Industry). Hourly time series of weather and air quality parameters were automatically downloaded

using a Python scripts, named downloaded_CETESB.py and qualar_py.py (developed by M. Gavidia). The repository of these scripts used in this study are available in this GitHub link.



Figure B.1 Air quality and meteorological hourly data, downloaded from CETESB (QUALAR)

Daily maximum rolling 8-hour mean (MDA8) for ozone concentration were compared with air temperature daily mean considering five years and by station type as shown in Figure B.3. Temperature has a positive correlation with surface ozone when the weather is not cloudy as occurred during highest rainfall in the MASP from November to March (de Lima and Magaña Rueda, 2018). When solar radiation is not reduced by cloud cover, months between September and December presented high hourly mean ozone concentration as shown in Figure B.5. September and October stand out as the months with the highest concentrations of ozone.

Table B.1 Air quality and weather stations network in the MASP

Name	Latitude	Longitude	Type	Abb
IAG	-23.651200	-46.622400	Forest preservation	IAG
Pico do Jaraguá	-23.456269	-46.766098	Forest preservation	PdJr
Santo Amaro	-23.654977	-46.709998	Urban	SAm
S.André-Capuava	-23.639804	-46.491637	Urban	SACp
S.André-Centro	-23.645616	-46.536335	Urban	SACt
S.André-Paço Municipal	-23.656994	-46.530919	Urban	SAPM
S.Bernardo-Centro	-23.698671	-46.546232	Urban	SBC
S.Bernardo-Paulicéia	-23.671354	-46.584668	Urban	SBP
Grajaú-Parelheiros	-23.776266	-46.696961	Urban	GrP
Marg.Tietê-Pte Remédios	-23.518706	-46.743320	Urban	MTPR
Guarulhos-Pimentas	-23.440117	-46.409949	Urban	GrlP
Cambuci	-23.567708	-46.612273	Urban	Cmb
S.Miguel Paulista	-23.498526	-46.444803	Urban	SMPa
Itaim Paulista	-23.501547	-46.420737	Urban	ItPa
Taboão da Serra	-23.609324	-46.758294	Urban	TdS
Interlagos	-23.680508	-46.675043	Urban	Intr
Guarulhos	-23.463209	-46.496214	Urban	Grlh
São Caetano do Sul	-23.618443	-46.556354	Urban	SCdS
Santana	-23.505993	-46.628960	Urban	Stna
Carapicuíba	-23.531395	-46.835780	Urban	Crp
Diadema	-23.685876	-46.611622	Urban	Dia
Mauá	-23.668549	-46.466000	Urban	Maua
Mogi das Cruzes	-23.518172	-46.186861	Urban	MdC
Mooca	-23.549734	-46.600417	Urban	Mca
N.Senhora do Ó	-23.480099	-46.692052	Urban	NSdO
Parque D.Pedro II	-23.544846	-46.627676	Urban	PDPII
Osasco	-23.526721	-46.792078	Urban	Ossc
Cerqueira César	-23.553543	-46.672705	Urban	CeC
Pinheiros	-23.561460	-46.702017	Urban	Pinh
Centro	-23.547806	-46.642415	Urban	Ctro
Guarulhos-Paço Municipal	-23.455534	-46.518533	Urban	GPM
Congonhas	-23.616320	-46.663466	Urban	Cng
Capão Redondo	-23.668356	-46.780043	Urban park	CRe
Cid.Universitária-USP-Ipen	-23.566342	-46.737414	Urban park	USP
Itaquera	-23.580015	-46.466651	Urban park	Itqr
Ibirapuera	-23.591842	-46.660688	Urban park	Ibir

Notes:

IAG station belongs to USP.

Abb = abbreviation.

Table	B.2 Ai	quality an	d weather	• stations	network	in	surrounding	areas	the	MA	SP	in t	he	São	Paulo	$^{\rm s}$	tate
-------	--------	------------	-----------	------------	---------	----	-------------	-------	-----	----	----	------	----	-----	-------	------------	------

Name	Latitude	Longitude	Type	Abb
Santa Gertrudes	-22.459955	-47.536298	Industry	StGrt
Paulínia	-22.772321	-47.154843	Industry	Pln
Jundiaí	-23.192004	-46.897097	Regional urban	Jnd
Limeira	-22.563604	-47.414314	Regional urban	Lmr
Taubaté	-23.032351	-45.575805	Regional urban	Tbt
Paulínia Sul	-22.786806	-47.136559	Regional urban	PlnS
Piracicaba	-22.701222	-47.649653	Regional urban	Prcb
Pirassununga-EM	-22.007713	-47.427564	Regional urban	PrEM
Presidente Prudente	-22.119937	-51.408777	Regional urban	\Pr Pr
Ribeirão Preto	-21.153942	-47.828481	Regional urban	RbPr
São José do Rio Preto	-20.784689	-49.398278	Regional urban	SJRP
S.José Campos	-23.187887	-45.871198	Regional urban	SJCp
S.José Campos-Jd.Satelite	-23.223645	-45.890800	Regional urban	SJCJ
S.José Campos-Vista Verde	-23.183697	-45.830897	Regional urban	SJCV
Sorocaba	-23.502427	-47.479030	Regional urban	Srcb
Tatuí	-23.360752	-47.870799	Regional urban	Tt
Jaú	-22.298620	-48.567457	Regional urban	Jau
Jacareí	-23.294199	-45.968234	Regional urban	Jcr
Americana	-22.724253	-47.339549	Regional urban	Ame
Catanduva	-21.141943	-48.983075	Regional urban	Cnt
Araçatuba	-21.186841	-50.439317	Regional urban	Ara
Araraquara	-21.782522	-48.185832	Regional urban	Arrq
Bauru	-22.326608	-49.092759	Regional urban	Bau
Campinas-Centro	-22.902525	-47.057211	Regional urban	CpC
Campinas-Taquaral	-22.874619	-47.058973	Regional urban	CpT
Campinas-V.União	-22.946728	-47.119281	Regional urban	CpV
Marília	-22.199809	-49.959970	Regional urban	Mrl
Guaratinguetá	-22.801917	-45.191122	Regional urban	Grtg



 $Figure \ B.2$ Hydrocarbon measurements in some CETESB stations as Toluene and Benzene.



Figure B.3 Time series of maximum daily rolling 8-hr mean for ozone and daily mean for temperature

Note:

Station types as Forest preservation (Pico do Jaraguá), Urban (Interlagos, Carapicuíba, Parque D.Pedro II, Pinheiros), and Urban park (Ibirapuera, Itaquera) are inside the MASP. Others as Industry (Paulínia) and Regional urban (Campinas-Taquaral, Sorocaba) are outside the MASP, and they belong to São Paulo state.



Figure B.4 Hourly mean concentrations for NO and NO₂ comparison for September 2018 for station types in the MASP.



Figure B.5 Surface ozone variation by month and station type in São Paulo State

Note:

Station types as Forest preservation (Pico do Jaraguá), Urban (Interlagos, Carapicuíba, Parque D.Pedro II, Pinheiros), and Urban park (Ibirapuera, Itaquera) are inside the MASP. Others as Industry (Paulínia) and Regional urban (Campinas-Taquaral, Sorocaba) are outside the MASP, and they belong to São Paulo state.


Figure B.6 Hourly time series of meteorological parameters registered at IAG/USP weather station. Note. Wind speed (ws), wind direction (wd), 2-m temperature (tc), relative humidity (rh), atmospheric pressure (press), rain rate (rr), sunlight duration (sun), cloud cover (cc).



Figure B.7 Total daily rain and cloud cover at IAG/USP weather station.



 $Figure \ B.8$ Total daily rain and cloud cover in September 2018 based on data collected at IAG/USP weather station Note:

Gray highlight represents excluding days for the paired statistical evaluation between modeled and observed values.

Appendix C

Statistical metrics

For this study, the statistical evaluation was developed using a Python script called mod_stats. This script was based on the original script model_stats developed by M. Gavidia's GitHub. Functions that are included in the script mod_stats (available in this GitHub) are :

- aq_stats evaluates air quality paired values
- met_stats evaluates meteorological paired values
- wind_dir_diff calculates difference between wind directions based on its periodic property based on Reboredo et al. (2015).
- wind_dir_mb calculates wind direction mean bias based on Reboredo et al. (2015).
- wind_dir_mage calculates wind direction mean absolute error based on Reboredo et al. (2015).
- r_pearson_sign calculates Pearson's r significance based on t-test with a two-tail (non-directional). The p-value is calculated using scipy.stats Python module with n-2 degrees of freedom.
- r_pearson_confidence_interval calculates Pearson's r confidence intervals using two-tail t-test.

These functions (aq_stats, met_stats) calculate statistical parameters described below:

- *n* is the number of paired values considered (modeled and observed values)
- MB is the Mean Bias
- MAGE is the Mean Absolute Gross Error
- *RMSE* is the Root Mean Square Error, which the ideal value is 0.
- *NMB* is the Normalize Mean Bias and reports mean paired modeled and observation differences normalized by the mean observation. Positive value corresponds to overprediction; negative value corresponds to underprediction.
- *NME* is the Normalized Mean Error and reports mean paired modeled and observation differences normalized by the mean observation as a positive value.

- *IOA* is the Index Of Agreement, which the perfect value is 1.
- r is the correlation coefficient based on Numpy module in Python
- Mm is the mean of modeled values
- Om is the mean of observed values
- Msd is the standard deviation of modeled values
- Osd is the standard deviation of observed values

Appendix D

Model results and Post-processing

This appendix shows results for the primary pollutants (NO_x, CO, and toluene), and meteorological parameters. WRF-Chem model outputs were processed through Python scripts (wrf_extract.py, wrf_rain.py, mod_stats.py). These scripts extracted pollutants and meteorological parameters for each station location, and for statistical analysis. Full codes are available on this GitHub.

Table D.1 Statistical results for meteorological parameters for Sep-Oct 2018 by type

	n	MB	MAGE	RMSE	IOA	r	Mm	Om	Msd	Osd	type
\mathbf{tc}	13040	1.33	1.85	2.44	0.91	0.88	20.98	19.65	4.05	4.17	Urban
\mathbf{rh}	13031	-7.01	10.58	13.77	0.81	0.73	71.24	78.25	16.26	15.80	Urban
ws	18748	1.36	1.65	2.02	0.49	0.34	3.35	1.98	1.49	1.04	Urban
wd	17869	-21.15	44.19	-	-	-	-	-	-	-	Urban
tc	1416	1.48	1.87	2.48	0.91	0.88	20.42	18.94	4.00	4.09	Urban park
$^{\mathrm{rh}}$	1416	-10.03	12.24	15.34	0.79	0.74	74.06	84.08	15.93	16.20	Urban park
ws	1429	1.58	1.72	2.08	0.52	0.47	3.39	1.82	1.48	1.05	Urban park
wd	1264	-10.24	32.67	-	-	-	-	-	-	-	Urban park
tc	18465	1.26	1.98	2.59	0.92	0.88	22.91	21.65	4.32	4.80	Regional urban
$^{\mathrm{rh}}$	17045	-7.44	11.46	14.72	0.84	0.77	65.96	73.40	17.92	19.09	Regional urban
ws	19427	1.20	1.64	2.08	0.58	0.46	3.31	2.11	1.86	1.24	Regional urban
wd	18641	-25.68	58.98	-	-	-	-	-	-	-	Regional urban
tc	2890	0.77	1.76	2.34	0.94	0.90	23.60	22.83	4.24	5.05	Industry
$^{\mathrm{rh}}$	2890	-5.16	11.26	14.85	0.87	0.80	65.01	70.17	19.74	22.96	Industry
ws	2888	1.71	1.99	2.49	0.43	0.34	3.29	1.58	1.88	0.99	Industry
wd	2538	-13.24	50.95	-	-	-	-	-	-	-	Industry
tc	2916	1.90	2.26	2.88	0.88	0.86	19.87	17.97	4.09	4.15	Forest preservation
$^{\mathrm{rh}}$	2916	-8.87	11.65	15.10	0.80	0.72	76.61	85.48	16.68	16.14	Forest preservation
ws	2920	2.18	2.27	2.69	0.31	0.33	3.74	1.57	1.67	0.75	Forest preservation
wd	2877	-11.27	43.36	-	-	-	-	-	-	-	Forest preservation



Figure D.1 Model results for 2-m relative humidity (Sep-Oct 2018) compared with observations from CETESB and IAG stations by types.

Note. Time series as daily mean. Shaded area is the standard deviation.



 $Figure \ D.2$ Model results for 2-m temperature (Sep-Oct 2018) compared with observations from CETESB and IAG stations by types.

Note. Time series as daily mean. Shaded area is the standard deviation.



 $Figure \ D.3$ Model results for wind speed (Sep-Oct 2018) compared with observations from CETESB and IAG stations by station types.

Note. Time series as daily mean. Shaded area is the standard deviation.



Figure D.4 Model results for NO_x (Sep-Oct 2018) compared with hourly time series from CETESB measurements. Note. Representative station types for Forest preservation (Pico do Jaraguá), Urban (Carapicuíba), Urban park (Ibirapuera), Industry (Paulínia), and Regional urban (Campinas-Taquaral).



Figure D.5 Model results for CO (Sep-Oct 2018) compared with hourly time series from CETESB measurements by station types.



Figure D.6 Model results for Toluene compared with CETESB measurements.



 $\label{eq:Figure D.7 Daily mean of temperature by scenario and station type$ The shaded area is the standard deviation.



 $\label{eq:Figure D.8} \mbox{ Daily mean of relative humidity by scenario and station type}$ The shaded area is the standard deviation.



Figure D.9 Daily mean of wind speed by scenario and station



Figure D.10 Daily total rain for scenarios and station type