Assessment of meteorological settings on air quality modeling system—a proposal for UN-SDG and regulatory studies in non-homogeneous regions in Brazil

Mauricio Soares da Silva1 · Luiz Cláudio Gomes Pimentel2 · Fernando Pereira Duda1 · Leonardo Aragão3 · Corbiniano Silva4 · Ian Cunha D’Amato Viana Dragaud4 · Pedro Caffaro Vicentini5

Abstract
Air quality models are essential tools to meet the United Nations Sustainable Development Goals (UN-SDG) because they are effective in guiding public policies for the management of air pollutant emissions and their impacts on the environment and human health. Despite its importance, Brazil still lacks a guide for choosing and setting air quality models for regulatory purposes. Based on this, the current research aims to assess the combined WRF/CALMET/CALPUFF models for representing SO2 dispersion over non-homogeneous regions as a regulatory model for policies in Brazilian Metropolitan Regions to satisfy the UN-SDG. The combined system was applied to the Rio de Janeiro Metropolitan Area (RJMA), which is known for its physiographic complexity. In the first step, the WRF model was evaluated against surface-observed data. The local circulation was underestimated, while the prevailing observational winds were well represented. In the second step, it was verified that all CALMET three meteorological configurations performed better for the most frequent wind speed classes so that the largest SO2 concentrations errors occurred during light winds. Among the meteorological settings in WRF/CALMET/CALPUFF, the joined use of observed and modeled meteorological data yielded the best results for the dispersion of pollutants. This result emphasizes the relevance of meteorological data composition in complex regions with unsatisfactory monitoring given the inherent limitations of prognostic models and the excessive extrapolation of observed data that can generate distortions of reality. This research concludes with the proposal of the WRF/CALMET/CALPUFF air quality regulatory system as a supporting tool for policies in the Brazilian Metropolitan Regions in the framework of the UN-SDG, particularly in non-homogeneous regions where steady-state Gaussian models are not applicable.

Keywords Air quality modeling · Regulatory purpose · SGDs · WRF · CALMET · CALPUFF · Metropolitan Regions

Introduction
Cities are important agents of change to build actions that identify virtues and vulnerabilities, as evidenced in the 17 Sustainable Development Goals (SDG) of the 2030 Agenda of the United Nations (UN) (UN 2015), contributing to the urban environment, improving from urban climate to the global climate, life quality of the population worldwide, and tackling key challenges established in the global pact signed by 193 UN member states, including Brazil (UN 2015, 2018; Acuti et al. 2020; Baklanov et al. 2016, 2018).

Air pollution represents the most significant environmental risk to health and affects all regions, scenarios, socioeconomic, and age groups. Every year, about 7 million deaths worldwide are attributed to ambient air pollution, and 91% of the world’s population lives in places exceeding WHO
air quality guidelines (WHO 2022). In 2019, 99% of the world’s population lives in places exceeding WHO air quality guidelines (WHO 2022). Countries can reduce the burden of diseases caused by stroke, heart disease, lung cancer, and both chronic and acute respiratory diseases, including asthma, by lowering air pollution levels (WHO 2022). The approaches seeking out an improving air quality are directly connected to access to clean energy services, climate mitigation goals, waste management, and other aspects of socioeconomic development (UN 2015, 2018; Rao et al. 2017; Rafaj et al. 2018; Manisalidis et al. 2020). Additionally, cleaner transportation, energy-efficient homes, power generation, industry, and better municipal waste management policies and investments would all help to reduce major sources of outdoor air pollution (WHO 2022).

Modeling tools inform the formulation of strategies for environmental analysis, land use, and land cover (LULC) effects on urban regions and megacities (Baklanov et al. 2016), to guide public policies in the air pollutant emissions control and their impacts on the air quality and population health in the context of the UN SDGs (e.g., USEPA 2004; Wang and Hao 2012; Requia et al. 2016; Ajtai et al. 2020; Chen et al. 2019; Higgins et al. 2019). The use of air quality (AQ) models is a complementary strategy to environmental monitoring and provides information to estimate the adverse impacts caused by gaseous and particulate emissions from industrial, vehicular, and other sources on health and ecosystems (Març et al. 2015; Thunis et al. 2016; Yang and Wang 2017; Gulia et al. 2020). AQ models are the only diagnostic and/or prognostic method that quantifies the deterministic relationship between emissions sources and concentrations/depositions at the receptor, making it possible to define and evaluate the effectiveness of mitigation strategies (Daly and Zannetti 2007; Nguyen 2014; Yadav et al. 2020). Such computational platforms with georeferencing techniques enable efficient implementation of environment and health management programs, based on air pollution episodes, spatial detailing of the pollutant’s concentration, and delimitation of the impact area, assist in the optimized design of the environmental monitoring network (EEA 2011).

Despite all the potentialities of the AQ models, these tools have uncertainties associated with sources known with meteorological fields, physical–chemical parameterizations, and emission inventory (Holnicki and Nahorski 2015; Haupt et al. 2019a, b), which should be minimized through sensitivity and performance studies. Regarding the meteorological fields, the use of diagnostic and prognostic computational modeling has constituted an alternative to complement or replace observational weather information to the AQ models (Lalas and Ratto 1996; Chandrasekar et al. 2003; Jackson et al. 2006; Abdul-Wahab et al. 2011a; De Visscher 2013; Lee et al. 2014; Wu et al. 2018; Rzeszutek 2019; Haupt et al. 2019a).

It is worth highlighting those diagnostic models such as California Meteorological Model–CALMET (Scire et al. 2000a). Quick Urban and Industrial Complex (QUIC) (Pardyjak and Brown 2003), Stationary Wind Flow and Turbulence (SWIFT) (Geai 1987), and MCSCIPUF (Sykes et al. 1984) do not represent a full solution to the atmospheric governing equations and cannot forecast atmospheric conditions since they are not elaborated through the time integration of the conservation equations (Lalas and Ratto 1996). Prognostic atmospheric models or Numerical Weather Prediction (NWP) models as the Fifth-generation Pennsylvania State University–NCAR Mesoscale Model–MM5 (Dudhia 1993), Advanced Regional Prediction System (ARPS) (Chow et al. 2006), and the Weather Research and Forecasting (WRF) model (Skamarock et al. 2019) are based on the solution of mass, momentum, and heat conservation equations. Although significant improvements have been made in the NWP models, some deficiencies remain, such as high computational costs, the difficulty in representing microscale effects, ample use of physical parameterizations to represent the atmospheric process, and difficulties in representing non-homogeneous boundary layer (e.g., Chow et al. 2006; Paiva et al. 2014; Dragaud et al. 2019; Rzeszutek 2019).

Klausmann et al. (2003) combined the Eta/CALMET/CALPUFF models to represent the dispersion and deposition of radioactive pollutants in situations with little meteorological information. Pfender et al. (2006) operated the MM5/CALMET/CALPUFF models to simulate the transport and deposition of phytopathogenic fungi spores. Aiming at estimating the impacts of air quality on the city of Chongqing, China. Yang et al. (2008) utilized the MM5/CALMET/CALPUFF in the transport of sulfur dioxide (SO2) to supplant the lack of measured meteorological data. Yim et al. (2010) using the MM5/CALMET/CALPUFF models were capable of reproducing regional monsoon patterns in Hong Kong making it possible to observe the variability of SO2 concentrations during the summer and winter seasons. Abdul-Wahab et al. (2011b) investigated the SO2 dispersion in the coastal region of Oman with WRF/CALMET/CALPUFF. Cui et al. (2011) evaluated the MM5/CALMET/CALPUFF models for a tracer experiment with synoptic conditions favorable for light winds in a complex terrain setting. Besides the detailed studies, other more recent investigations available in the scientific literature seek to demonstrate the skill of the atmospheric combined models to represent the wind field in high resolution in the pollutant dispersion model (Hernández-Garcés et al. 2021; Hao and Xie 2018; Rzeszutek 2019).

The AQ models are still not explored by the Brazilian legislation for the formulation of air quality control strategies at urban, regional, or national levels. Despite the diversity of AQ models available in the international literature in recent years (Holmes and Morawska 2006; Leelössy et al. 2016). Quick Urban and Industrial Complex (QUIC) (Pardyjak and Brown 2003), Stationary Wind Flow and Turbulence (SWIFT) (Geai 1987), and MCSCIPUF (Sykes et al. 1984) do not represent a full solution to the atmospheric governing equations and cannot forecast atmospheric conditions since they are not elaborated through the time integration of the conservation equations (Lalas and Ratto 1996). Prognostic atmospheric models or Numerical Weather Prediction (NWP) models as the Fifth-generation Pennsylvania State University–NCAR Mesoscale Model–MM5 (Dudhia 1993), Advanced Regional Prediction System (ARPS) (Chow et al. 2006), and the Weather Research and Forecasting (WRF) model (Skamarock et al. 2019) are based on the solution of mass, momentum, and heat conservation equations. Although significant improvements have been made in the NWP models, some deficiencies remain, such as high computational costs, the difficulty in representing microscale effects, ample use of physical parameterizations to represent the atmospheric process, and difficulties in representing non-homogeneous boundary layer (e.g., Chow et al. 2006; Paiva et al. 2014; Dragaud et al. 2019; Rzeszutek 2019).

Klausmann et al. (2003) combined the Eta/CALMET/CALPUFF models to represent the dispersion and deposition of radioactive pollutants in situations with little meteorological information. Pfender et al. (2006) operated the MM5/CALMET/CALPUFF models to simulate the transport and deposition of phytopathogenic fungi spores. Aiming at estimating the impacts of air quality on the city of Chongqing, China. Yang et al. (2008) utilized the MM5/CALMET/CALPUFF in the transport of sulfur dioxide (SO2) to supplant the lack of measured meteorological data. Yim et al. (2010) using the MM5/CALMET/CALPUFF models were capable of reproducing regional monsoon patterns in Hong Kong making it possible to observe the variability of SO2 concentrations during the summer and winter seasons. Abdul-Wahab et al. (2011b) investigated the SO2 dispersion in the coastal region of Oman with WRF/CALMET/CALPUFF. Cui et al. (2011) evaluated the MM5/CALMET/CALPUFF models for a tracer experiment with synoptic conditions favorable for light winds in a complex terrain setting. Besides the detailed studies, other more recent investigations available in the scientific literature seek to demonstrate the skill of the atmospheric combined models to represent the wind field in high resolution in the pollutant dispersion model (Hernández-Garcés et al. 2021; Hao and Xie 2018; Rzeszutek 2019).

The AQ models are still not explored by the Brazilian legislation for the formulation of air quality control strategies at urban, regional, or national levels. Despite the diversity of AQ models available in the international literature in recent years (Holmes and Morawska 2006; Leelössy et al. 2016).
of superseding this absence for use in regulatory models as a manner to recommend the use of various regions of Brazil (WMO 2020, 2021a, b) has led to the establishment of local observation network and depicts violations of air quality that vary in space-time (e.g., fumigation, breeze, mountain circulations), besides not demand high computational costs and many inputs such as the WRF-Chem (Grell et al. 2005) and CMAQ (Byun and Schere 2006) models that are still not feasible for regulatory purposes. Appropriate AQ models for the licensing of the ports, oil exploration activities, highways, and airports still must be established in the state and/or federal legislation.

The current scenario is worrying due to the disordered expansion of urban-industrial areas in the country and the consequent increase in the concentration levels of atmospheric pollutants such as particulate matter (PM) and ozone (O₃) (Soares da Silva and Pimentel 2017; Perlmutt and Cronmar 2019; Bodor et al. 2020; Bálâ et al. 2021; Da Silveira et al. 2021; Gao et al. 2021; Liu et al. 2022). The Rio de Janeiro Metropolitan Area (RJMA) is the second-largest among the 74 metropolitan areas in Brazil (IBGE 2020) and presents the second-highest concentration of industries and vehicles in the country (INEA 2015). These aspects in conjunction with the heterogeneity of LULC, anthropogenic source location, wind regime spatial and temporal variability, and complex topography play an important role in the atmospheric pollutant dispersion modeling in the non-homogeneous RJMA (Pimentel et al. 2014a, b; Paiva et al. 2014; Soares da Silva et al. 2014; Silveira et al. 2021). Also valuable is the use of AQ models to design and support the local observation network and depict violations of air quality standards, as well as to establish and guide public policies for urban-industrial planning and RJMA environmental protection areas (Silveira et al. 2021).

The lack of observed meteorological information in various regions of Brazil (WMO 2020, 2021a, b) has led certain environmental agencies to recommend the use of results from prognostic atmospheric models as a manner of superseding this absence for use in regulatory models of air quality. Among the many available models, the WRF (Skamarock et al. 2019) has been frequently recommended for this purpose. However, as discussed in Rzeszutke (2019), meteorological and air quality modeling in complex terrains is one of the main challenges nowadays. This argument is valid for the RJMA, as verified by previous studies on the simulation of pollutants dispersion and atmospheric flow over the surface in this region (Pimentel et al. 2014b; Paiva et al. 2014; Giannaros et al. 2017; Dragaud et al. 2019). Care must be taken to obtain predictions with a universal configuration for the NWP models, as proposed by State Environmental Institute (INEA), and the selection of input data of adequate quality for different regions.

This study aims to evaluate meteorological settings of the combined WRF/CALMET/CALPUFF models to represent the dispersion of SO₂ in hourly and 24-h time scales over non-homogeneous regions as a proposition to a regulatory model for land use policies in Brazilian Metropolitan Regions in the context of UN Sustainable Development Goals–2030 Agenda.

Study area

The RJMA comprises 22 cities, with a territory of 7530.386 km², a population of about 13 million, and a density of 1743.81 inhab·km⁻², and is currently the second largest among the 74 metropolitan areas in Brazil (IBGE 2020). It is a highly industrialized region, whose current process has been leveraged by the metal-steel and logistical-oil sectors, integrated with other industries and their associated activities, and increased by logistical investments (articulated by highways, railways, and ports) and industrial, productive partnerships, financial and tax with the installed companies (Kjerfve et al. 1997; Moraes et al. 2017; da Silveira et al. 2021). The RJMA is characterized as a pole of attraction for investments and expansion of already consolidated sectors and the newer ones, where the Rio de Janeiro Petrochemical Complex (COMPERJ) and the enterprises located in the Sepetiba Bay (Porto de Itaguaí, Termium, Gerdau, Usiminas, Petrobrás, LLX, among others) stand out (Da Silva and Zurita 2019; Kjerfve et al. 2021; da Silveira et al. 2021).

According to the criteria established by the State Environmental Institute (INEA 2010, 2015), the leading industrial activities in the Rio de Janeiro state are of high or medium potential for pollution, with RJMA being the one with the highest levels of air quality degradation. It occurs mainly due to the high concentration of emission sources since it is the second-largest concentration of vehicles, industries, and polluting sources in the country, whose consequences generate severe problems of air pollution. In the report INEA (2010), it was highlighted that this region showed the greatest compromise in air quality in the Rio de Janeiro State, being 77% of all inventoried emitted pollutants from
vehicular sources, and 23% coming mostly from industrial sectors like petrochemical, naval, chemical, food, and energy transformation.

Its geo-bio-physiographic characteristics enhance problems associated with air quality and are challenging elements for atmospheric modeling studies. It highlights the Serra do Mar Mountain range to the north-northwest, the Atlantic Ocean to the south, the presence of water bodies such as lakes, the Sepetiba Bay to the southwest, and the Guanabara Bay around the center of the coastline. The Guanabara Bay measures 28 km width from approximately west to east, 30 km from south to north, has a narrow entrance of 1.6 km wide, and has a water surface area of 328 km$^2$ (Kjerfve et al. 1997). In comparison, Sepetiba Bay has maximum south to north shore-normal width of 15 km, 40 km from west to east, and a total water surface area of 305 km$^2$ (Kjerfve et al. 2021). The biggest urban area in the region is delimited by the water bodies (the Atlantic Ocean and the bays) and three massifs (Fig. 1). These elements produce a complex and heterogeneous land surface and airflow in terms of the distribution and dispersion of pollutants; the tropical climate, which favors photochemical processes, generating other pollutants; as well as the critical and intense process of land use, which highlights urban and industrial areas, and the largest urban forest on the planet (Tijuca Forest), among other uses (Figs. 1 and 2).

The South Atlantic Subtropical Anticyclone (SASA) is on the synoptic-scale the system that predominantly influences the region, favoring clear and slightly cloudy skies (Dereczynski and Menezes 2017) and inducing the near-surface winds to blow from the first quadrant ($0^\circ$–$90^\circ$) (Dereczynski and Menezes 2017; Oliveira-Júnior et al. 2017; Dragaud

![Study area, with emphasis on: Serra dos Órgãos mountain range to the north; Atlantic Ocean to the south; Gericinó-Mendanha, Pedra Branca and Tijuca massifs, and Guanabara and Sepetiba bays](image-url)
The influence of SASA is interrupted by transient meteorological systems, such as extratropical cyclones, frontal systems (FS), mesoscale convective systems (MCS), upper tropospheric subtropical cyclonic vortices, and the South Atlantic Convergence Zone (SACZ), which promote increased cloudiness and rainfall (Oliveira Júnior et al. 2014; Dereczynski and Menezes 2017; Oliveira-Júnior et al. 2017; Brito et al. 2017; Bonnet et al. 2018; Nielsen et al. 2019; Correia Filho et al. 2021). Under the influence of an atmospheric frontal system, winds come from the southern quadrant (Stech and Lorenzzetti 1992; Dourado and Oliveira 2001) and inhibit the local circulations.

Local thermally driven circulations can overcome the SASA winds, and near the shore, the influence of the land, sea, and bay breeze systems make the near-surface wind regime perpendicular to the coastline (Pimentel et al. 2014b; Aragão et al. 2017; Oliveira-Júnior et al. 2017; Kjerfve et al. 2021; Correia Filho et al. 2021). The sea breeze comes from the Atlantic Ocean (Paiva et al. 2014) and couples with the bay breeze circulation over the Guanabara Bay (Oliveira-Júnior et al. 2017) and Sepetiba Bay (Aragão et al. 2017). As the near-surface flow approaches the mountains, it can become channeled and influenced by the valley-mountain circulation (Oliveira et al. 2013; Pimentel et al. 2014b; Aragão et al. 2017; Sobral et al. 2018; Kjerfve et al. 2021; Correia Filho et al. 2021). The stages of the bay breeze at Sepetiba Bay were investigated by Aragão et al. (2017). The preparatory phase starts when the land surface temperature becomes higher than the sea surface temperature, the convective boundary layer begins to build, and the wind direction is normal aligned, from the bay to land close to the shoreline. In the development phase, the greatest bay-land temperature and pressure gradients occur when the near-surface wind directly blows from the bay to land and
the maximum wind speeds occur. During the mature phase, there is a decrease in air temperature and an increase in wind speed, and the vertical circulation cell between land and bay is established. Near sunset, the final phase occurs, and there is the transition to the land breeze circulation, and the bay-land temperature gradient drops close to zero. The Sepetiba Bay joins the Ilha Grande Bay, forming a coupled bay system (Kjerfve et al. 2021). At the bottom of Ilha Grande Bay, de Freitas Ramos Jacinto et al. (2021) analyzed the thermally and dynamically driven local atmospheric circulations. Over this region, the authors characterized the local wind regime as a result of the combined action of multiscale weather systems, such as the valley-mountain, land-sea, and bay breeze circulations; synoptic-scale systems; and the influence of the relief (de Freitas Ramos Jacinto et al. 2021).

Pollution and air quality are identified in connection with the SDGs (UN 2018) and have been incorporated into the RJMA municipalities. The SDGs directly related to air pollution (WHO 2016) include health (SDG 3, target 3.9), cities (SDG 11, target 11.6), and energy (SDG 7, target 7.A). Air quality is implicit in other development goals: food production and safety (SDG 2, target 2.4); economic growth and employment (SDG 8, target 8.4); reduced inequality (SDG 10); urgent actions to mitigate climate change (SDG 13, target 13.2); conservation and sustainable use of oceans, seas, and marine resources (SDG 14, target 14.1); sustainably manage of forests, combat desertification, halt and reverse land degradation, face the biodiversity loss (SDG 15, targets 15.1, 15.4, and 15.5).

The methodology is based on the implementation of the Sustainable Cities Development Index–Brazil (IDSC-BR) (https://idsc-br.sdgindex.org/methodology), produced by the Sustainable Development Solutions Network (SDSN) (Lafortune et al. 2018). The scores highlight the biggest challenges and evaluate the performance of services and public policies implemented by municipal management. The evaluation of the evolution of the municipalities in an indicator is measured in four intervals, where the green category (> 60 to 100) indicates that the municipality has reached the SDG, yellow (50–60) indicates that there are challenges, orange (40–50) indicates significant challenges, and red (< 40) indicates major challenges. Based on the adopted methodology, only 2 of the 21 RJMA cities evaluated reached the SDGs associated with pollution and air quality, with the others in different situations of vulnerability, where 12 present challenges, 6—significant challenges, and 1—major challenges (Fig. 2).

Materials and methods

Period

The study period was from January 1 to May 31, 2011, defined by the simultaneous availability of the required datasets (meteorological, air quality, and emissions), which comprises the main meteorological systems acting over the RJMA: SASA, FS, SACZ, MCS, and breeze systems (valley-mountain, land-sea, and bay).

Observed data

Data recorded in three types of monitoring stations were used: meteorological, pluviometric, and air quality (Fig. 2). Meteorological data was recorded in five World Meteorological Organization standard stations located in the airports of RJMA. From the five stations—Galeão Airport (SBGL), Campos dos Afonsos Airport (SBAF), Santos Dumont Airport (SBRJ), Jacarepaguá Airport (SBJ), and Santa Cruz Airport (SBSC)—only the SBGL station registers both surface and upper-air observations. At surface level, the records are hourly, and at the upper air two soundings are launched daily at 9 a.m. and 9 p.m. local time. Three surface stations (SBGL, SBAF, and SBSC) operate on a full-time basis, that is, 24 h a day, while the SBRJ and SBJ stations operate from 6 a.m. to 11 p.m. and 6 a.m. to 10 p.m. local time, respectively. The wind data used in this study presents measurements above 0.5 m.s⁻¹ (≈ 1 knot). Wind speed values below this threshold are considered calm winds. To elaborate on the diagrams, the wind classes defined in the Beaufort scale were considered (WMO 2018).

In addition to traditional meteorological stations, 41 pluviometric stations with hourly accumulated rainfall data are considered to represent the wet removal mechanism. 32 stations come from the Municipality of Rio de Janeiro network and 9 from the State of Rio de Janeiro network (Fig. 2). The AQ monitoring stations present in the RJMA are carried out by INEA (State Institute of the Environment), SMAC (Secretariat of Environment of the Municipality of Rio de Janeiro), in addition to several private companies. This study considers hourly averages of SO₂ concentrations from five AQ monitoring stations: Copacabana, São Cristóvão, Tijuca, Jardim Primavera, and São Bento (Fig. 2). The first three are managed by SMAC and the last two by private companies. Although there are more AQ stations for the study period, only the stations with 75% or more recorded data were considered.

WRF model and uncertainty thresholds for the evaluation of wind speed

Due to the advances in solution techniques, improvements to the representation of physical processes, and an increase in computational power, prognostic models currently represent state-of-the-art approaches to simulating atmospheric phenomena in various time-space scales (Pielke 2002; Goger et al. 2016; Yano et al. 2018; Haupt et al. 2019a, b, c).
The WRF is an NWP model developed for operational purposes and atmospheric research (Skamarock et al. 2019). Initial and boundary conditions used in the simulations were produced with the Global Forecasting System (GFS) model, which generates analysis with a horizontal spatial resolution of 0.5° (~55 km at RJMA latitudes) and a temporal resolution of 6 h. Four nested domains centered on the Rio de Janeiro state were adjusted with one-way interaction (Fig. 3). The largest domain, D1, possesses 24 km of horizontal grid resolution with 130 ×100 grid cells; the second domain (D2) has an 8-km grid resolution with 148 × 106 grid cells; the third (D3) has a 2.666-km grid resolution with 172 × 112 grid cells; and the last one, D4, has a 0.888-km grid resolution with 181 × 112 grid cells, covering the whole RJMA. All domains have 45 vertical levels, from the ground level up to 50 hPa. The physics parameterization schemes used are the Rapid Radiative Transfer Model for General Circulation Models (RRTMG) (Iacono et al. 2008), Revised MM5 (Jiménez et al. 2012), Noah – MP Land Surface Model (Niu et al. 2011), Yonsei University (YSU) (Hong et al. 2006), WRF single-moment three-class (WSM3) (Hong et al. 2004), and Kain–Fritsch Scheme (Kain 2004). The horizontal diffusion was computed in physical space since it is more realistic and suggested for regions with complex terrain (Chen 2020).

Several studies have been developed to evaluate the accuracy of atmospheric models through the comparison with observational data using statistical indexes, such as bias and RMSE (root-mean-square error) comparing the simulated results with the observed data in sites with in situ monitoring (e.g., Hanna and Yang 2001; Chow et al. 2006; Zhang et al. 2011; Paiva et al. 2014; and Yver et al. 2013). However, there are different uncertainty levels not only to the setups applied on models (e.g., schemes, grids, initial, and boundary conditions) but also due to the geographical complexity that influences the wind regimes. Table 1 and 2 are summarized the uncertainty thresholds for the evaluation of wind speed.
**CALPUFF modeling system**

The California puff model (CALPUFF) modeling system consists of two main processing modules and a post-processing one: the meteorological module CALMET, the dispersion module CALPUFF, and the CALPOST. The latter is responsible for the post-processing of the results generated by the CALMET and CALPUFF modules (Scire et al. 2000a, b). CALMET is a diagnostic meteorological model that produces a 3D-gridded field of wind and temperature, estimates micrometeorological parameters (for input in the CALPUFF dispersion model), and assures mass conservation, being called a mass–consistent model (Scire et al. 2000a). The relative simplicity and operability of diagnostic models, combined with the skill to run much faster than prognostic models, make them attractive for many practical purposes. It should also be considered that these models do not require much input data available to operate efficiently in the generation of a 3D-gridded meteorological field, which satisfies some physical constraints such as kinematic terrain effects, slope flows, and blocking effects (Lalas and Ratto 1996; Scire et al. 2000a; Haupt et al. 2019a, b, c).

The domain utilized in CALMET was adjusted to cover the largest part of the RJMA: 160 × 140 horizontal grid cells with 0.5 km of resolution (Figs. 1 and 2) and 12 vertical layers from the ground level up to the free atmosphere (4000 meters). Topography data from the Shuttle Radar Topographic Mission project (Farr et al. 2007) were used with 90 m of spatial resolution, as well as soil use and occupation data from the Glob Cover project (Arino et al. 2012) with a 300-m resolution. The required parameters TERRAD, R1, RMAX1, R2, and RMAX2 were set with the values 10 km, 1 km, 2 km, 10 km, and 20 km, respectively.

To define the best meteorological setting in the WRF/CALMET/CALPUFF, three numerical experiments were developed (Fig. 4):

- **CONFIG 1**: Meteorological data measured at five surface meteorological stations, one upper air meteorological station, and 41 precipitation stations.
- **CONFIG 2**: Meteorological data from WRF model and 41 precipitation stations.
- **CONFIG 3**: Combination of the data used in CONFIG 1 and CONFIG 2.

Since the evaluation period is long and includes rainy conditions, all simulations considered the wet removal process with rainfall data measured in 41 stations, and the SO₂

---

**Table 1** Summarized statistical indexes from other studies for wind speed and wind direction

| Studies                   | Index  | Wind Speed (m s⁻¹) | Wind direction (°) | Ideal |
|---------------------------|--------|--------------------|--------------------|-------|
|                           |        | Best               | Worst              |       |
|                           |        |_best_             |_worst_             |_ideal_|
| Hanna and Yang (2001)     | RMSE   | 1.6               | 2.5               | 51    | 76    | 0     |
|                           | bias   | −0.1              | 1.5               | −2    | −14   | 0     |
| Zhang et al. (2011)       | RMSE   | 2.6               | 3.7               | 172.6 | 193.4 | 0     |
|                           | bias   | 1.4               | 2.5               | −20.9 | −80.9 | 0     |
| Paiva et al. (2014)       | RMSE   | 1.12              | 4.26              | 59.73 | 105.11| 0     |
|                           | bias   | 0.03              | 3.18              | 2.56  | −36.47| 0     |
| Chow et al. (2006)        | RMSE   | 0.53              | 1.56              | 53.14 | 93.21 | 0     |
|                           | bias   | 0.04              | −1.01             | −1.50 | −35.79| 0     |
| Yver et al. (2013)        | RMSE   | 1.00              | 3.00              | 40.00 | 140.00| 0     |
|                           | bias   | −     | −     | −     | −     | 0     |

**Table 2** Uncertainty grades for wind speed, according to Clifford (2011)

| Grade    | RMSE (m.s⁻¹) | MAE (m.s⁻¹) |
|----------|--------------|-------------|
| Poor     | > 3.0        | > 3.0       |
| Acceptable | < 3.0        | < 3.0       |
| Good     | < 2.5        | < 2.5       |
| Excellent | < 2.0        | < 2.0       |
standard CALPUFF coefficient (Scire et al. 2000b), that is, scavenging coefficient of 3.0E-05 s⁻¹. To avoid errors associated with the WRF precipitation field and to allow a more focused assessment of the wind, only measured data were employed, including CONFIG 2.

Before being used in the CALMET model, the WRF information was pre-processed by the CALWRF, a code also developed and distributed by the developers of the CALPUFF modeling system.

The CALPUFF model (Scire et al. 2000b) is a multi-layer, multi-species, non-steady-state Lagrangian modeling system of the atmospheric dispersion of puffs with Gaussian distribution. Its main characteristics are the capacity to simulate the effects of time and space varying meteorological conditions, calm wind conditions, first-order chemical transformations, and removal processes during transportation. The method based on the similarity theory was used to calculate the dispersion coefficients, and the method considered for the representation of the plume was the Slug (non-circular puff elongated in the direction of the wind when near its source).

**Sulfur dioxide total emissions at RJMA**

The primary pollutant SO₂ was employed in the present study, where the inventory of fixed and mobile emissions used in the data entry process of WRF/CALMET/CALPUFF is the same one utilized by Vicentini et al. (2011) and Vicentini et al. (2011). In the latter study, all sulfur dioxide fixed and mobile sources of the RJMA were considered according to the INEA pollutant emission inventory prepared in 2009, based on the year 2008 (Vicentini 2011; Vicentini et al. 2011). It is noteworthy that fixed sources are responsible for about 88% of SO₂ emissions in RJMA (Vicentini 2011; INEA 2010).

The inventory of fixed sources for the RJMA updated in 2009 considered the details of the inventory of the ninety largest companies out of the 425 considered in the inventory of 2003 (Pires 2005), maintaining the other companies’ emission factors. All of these were treated as point sources, totaling 1917 anthropogenic fixed sources.

The inventory of mobile sources was also updated and expanded in 2009, adding to the 260 sources initially registered (2003) another 290 relating to secondary roads. The flow on the inventoried roads was measured using counts of light and heavy vehicle traffic at certain times of the day based on the manual of the National Department of Transport Infrastructure (DNIT 2006) and data from the Department of Transport of the Rio de Janeiro (DETRAN-RJ). The RJMA’s hourly variation of vehicular traffic was represented considering 24 emission factors. These factors make no distinction between weekend days’ and business days’ traffic.

This inventory was formatted to be assimilated into CALPUFF, resulting in 499 area sources representing the main traffic routes in the RJMA, and 492 point sources representing the main industrial sources cataloged in the region. Figure 5 shows the areas with the highest emissions from fixed sources divided by 5 x 5 km areas. However, it should be noted that these emissions were configured in the CAQMS as point sources. Although the emission factors for fixed sources may vary over time, this variation was not considered since it would require information that is not available. Furthermore, the main emission activities in the region operate continuously under constant load most of the time.

**Evaluation of combined WRF/CALMET/CALPUFF system**

Even with the advances, NWP models still do not satisfactorily resolve local flow forced by terrain characteristics such as the topography and land use, for example. Otherwise, diagnostic models can better represent microscale phenomena but do not provide predicted meteorological fields (Lalas and Ratto 1996). As an attempt to attain the best performance of the AQ models, both approaches can be combined, exploring the complementarity of their skills, comprising the so-called combined air quality modeling system (CAQMS).

The evaluation of the CAQMS skill to simulate the wind field and SO₂ hourly and 24-h averages concentration was done through two steps. The first one focused on the evaluation of the WRF skill to simulate the wind field over the non-homogeneous RJMA. In this step, WRF hourly data of wind speed and direction were extracted using the nearest neighbor technique for the five meteorological surface stations. Thus, simulated and observed data were evaluated via frequency distribution diagrams (wind roses) and statistical indexes. The literature on the performance of prognostic meteorological models proposes a wide variety of statistical methods for model evaluation. However, the indexes root mean square error (RMSE), mean absolute error (MAE), and bias are commonly utilized in these evaluations to quantify the uncertainty of the models (Hanna and Yang 2001; Chow et al. 2006; Zhang et al. 2011; Paiva et al. 2014).

The second step evaluated the different meteorological settings strategies used on the CAQMS—CONFIG 1, CONFIG 2, and CONFIG 3—modified the SO₂ hourly and 24-h average concentrations results. This analysis considers the pairing in time and space of the simulated results and the AQ monitoring stations data using a set of statistics indexes widely employed on AQM assessments: bias, normalized mean square error (NMSE), Pearson correlation coefficient (R), factor of exceedance (FOEX), the fraction of predictions within a factor of two of observations (FAC2), the fraction of predictions within a factor of five of observations (FAC5). In particular, FOEX was used to bring information about
the percentage of samples overestimated/underestimated by the model without considering the magnitude of this error. The FOEX ranges between +50 and −50% and the ideal value are 0%. While a FOEX equal to +50% means that all values are overpredicted, a FOEX equal to −50% means that all values are under-predicted. Since FA2 (FA5) represents the percentage of estimates between 1/2 (1/5) ≤ observed/predicted ≤ 2 (5) (Mosca et al. 1998).

Additionally, a final assessment was presented focused on verifying whether the performance of the CAQMS is compatible with what is expected of a regulatory model. For this type of assessment, some assumptions different from the previous ones were considered to equalize with the assessments commonly carried out for regulatory models of air quality. In this sense, only the pairing in space between the simulated and observed data was considered, since the main objective of this type of analysis is not to verify the ability of a model to predict the concentration in time and space, but to predict the magnitude of concentrations, more specifically, the highest concentrations (USEPA 1992). In addition to the less rigorous pairing, the set of statistical indices most frequently used for this type of analysis was also considered (Chang and Hanna 2004), which are the fractional bias (FB), the geometric mean bias (MG), the NMSE, the geometric variance (VG), the Spearman correlation coefficient (Rsprm), and FAC2.

Results and discussion

Analysis of the WRF skill to simulate the wind field

Figure 6A–J present the frequency distribution diagrams of the wind data extracted from the simulations with the WRF model and from the data measured in the respective surface meteorological stations.

At Galeão station, the most frequently observed wind directions were Southeast (30%) and East (16%) (Fig. 6A), associated with the afternoon-occurring sea breeze and bay breeze systems (Pimentel et al. 2014a, b; Oliveira-Júnior et al. 2017). The results obtained with the WRF (Fig. 6B) indicate that the model represents the prevailing...
wind directions at a frequency like those observed at the Galeão station (Fig. 6A). This result is above expectations since the representation of the wind at the SBGL station is difficult because it is located on an island at the west boarder of Guanabara Bay, close to the region where the bay has its maximum width (Fig. 2). Over this region, the breeze acts in different directions in a short distance, following the irregular shoreline. Regarding the wind speed, the observed winds in all directions are more intense than the simulated (Fig. 6A–B), in general. This difference in intensity occurs in the order of one class, that is, approximately 1.5 m.s⁻¹, which is excellent considering the scales of uncertainties proposed by Clifford (2011) (Table 2). For calm winds, the model results represent a percentage (4.86%) close to that observed (6.96%). Wind occurrences from the north quadrant (315°-45°) (Fig. 6A–B)
are associated with the land breeze system, which is less intense than the bay and sea breeze, and with the SASA influence (Oliveira-Júnior et al. 2017).

At the Campo dos Afonsos station, there are mainly observed winds from the south direction (Fig. 6C). The southern winds are related to the sea breeze, which comes from the Atlantic Ocean, enters the continent, and is channeled by the Tijuca and Pedra Branca massifs (Pimentel et al. 2014a, b), which reaches heights above 1000 m. The occurrence of the east and west wind direction is probably related to the bay breeze propagation from the Guanabara and Sepetiba Bays, respectively. The winds from the north quadrant (315°–45°) (Fig. 6C) are associated with the land breeze system (Pimentel et al. 2014b) and with the SASA influence (Correia Filho et al. 2021). According to Paiva et al. (2014), this station is also influenced by the valley-mountain breeze system. The incidence of light winds with around 14% of calm winds (Fig. 6C) usually occur during the night, early morning, and morning periods, and is caused by the influence of a topographic barrier (Pimentel et al. 2014b), can also be highlighted. The results obtained with the WRF (Fig. 6D) demonstrate that the model was able to reproduce wind patterns like those registered at the SBAF station (Fig. 6C). It is possible to recognize the ability of the model in the reproduction of the variability of wind directions, indicating its sensibility in simulating the coupling of processes in several scales: micro, mesoscale, and synoptic, which result in the wind regime in this region. It is also noteworthy that the frequency of the prevailing winds, the winds from the south, are like those recorded at the station, around 20% (Fig. 6C–D). Likewise, the simulated frequency of the calm winds indicates the optimal accuracy of the model (Fig. 6C–D). Regarding the wind speed, subjectively, it can be said that the results are consistent with the observed data indicating excellent performance according to criteria given by Clifford (2011), and thus being better than the results modeled for the SBGL station.

The Santos Dumont (SBRJ) station (Fig. 6E), situated at the entrance of the Guanabara Bay, experiences a strong influence from the land and sea breezes regime, where the directions South (sea breeze) and North (land breeze) are predominant (Pimentel et al. 2014b; Oliveira-Júnior et al. 2017). However, since the observational data available for this study were registered only during the period of 6 a.m. to 11 p.m. local time, it was not possible to observe the action of the land breeze (North winds) in the same proportion as the sea breeze (Fig. 6E). In the simulated results which have been extracted for the same periods as the observed data, the predominant winds are southerly winds as well as the observed data, but 10% less often (Fig. 6E–F). This difference between the observed and modeled frequency is balanced by the southeast winds, which are modeled at a frequency of approximately 10% higher than those observed (Fig. 6E–F). From this, it is assumed that the model can represent the sea breeze phenomenon at the same frequency as that observed. It is noteworthy that the second most frequent phenomenon at the station was also modeled by WRF, that is, the land breeze associated with the north winds (Fig. 6F). The intensity of the most frequent winds varied between 3.6 and 5.7 m.s⁻¹ for both observed and modeled results (Fig. 6E–F). As well as the results obtained for SBGL and SBAF, the performance of WRF to model periods of calm winds on the SBRJ site is excellent, with only about a 1% difference between those registered at the station. Although both SBRJ and SBGL are located inside Guanabara Bay, the first is in a narrower region of the Bay and is closer to the ocean and mountains in the surrounding area. These physiographical characteristics make the wind pattern differ at these sites. SBRJ has more influence on the sea breeze and less influence on the bay breeze (Oliveira-Júnior et al. 2017) than SBGL, and the wind is channeled by the presence of the bay (Pimentel et al. 2014b) and the surrounding mountains.

For Jacarepaguá station, which is located to the South of the Campo dos Afonsos station and at 2 km away from the shoreline, meteorological observations were realized only between 6 a.m. and 10 p.m. local time. In this station, it is possible to observe a higher occurrence frequency of the south to north winds (Fig. 6G). According to Pimentel et al. (2014b), the southern winds occur mainly during the afternoon due to the sea breeze, and the northern winds occur predominantly during the morning due to the land breeze. However, this site is close to the ocean and more exposed to the influence of synoptic-scale disturbances. The FS and migratory high-pressure systems induce southwesterly winds (Stech and Lorenzetti 1992; Cavalcanti and Kousky 2009), while the SASA induces winds from the east quadrant (45°–135°) (Dereczynski and Menezes 2017; Oliveira-Júnior et al. 2017; Dragaud et al. 2019; Kjerfve et al. 2021; Correia Filho et al. 2021). The occurrence of low-speed winds and the high percent of calm winds (10.15%) in the period of the study should be highlighted (Fig. 6G). Although it does not present the highest frequency specifically for the southern winds, it appears that the model represents the phenomenon of the sea breeze at a frequency like that observed, if assuming the southeast winds as one of the possible consequent directions of this phenomenon (Fig. 6G–H). It should also be noted that the northern winds associated with the land breeze are represented by the model at a frequency like the observation (Fig. 6G–H). Note that the WRF model simulates (Fig. 6H) more intense winds than those observed, as well as verified for the SBGL station, and this average difference in intensity is about 1.5 m.s⁻¹. These are excellent results according to the criteria of Clifford (2011). However, it is noted that the model does not present good accuracy in representing calm winds for this site, i.e., 3.9% modeled against 10.15% observed, the worst result among all sites.
The data observed at the Santa Cruz (SBSC) station, located on the west portion of the RJMA, near the east border of Sepetiba bay, is presented in Fig. 6f. It can be noted that southwest and northeast winds are predominant, besides the 8% of registered calm winds at this site (Fig. 6f). These wind directions are due to the action of the land-sea and bay breeze circulations due to the presence of the Sepetiba bay and the Atlantic Ocean (Pimentel et al. 2014b; Aragão et al. 2017). Regarding the simulations, it is highlighted that the results demonstrate that the model could represent the phenomenon of the bay breeze system associated with southwest winds. However, the present simulated results did not represent the wind direction associated with land breezes (northeastern winds) as well as the bay breeze (southwest winds) (Fig. 6f). The adequate simulation of the bay breeze represents a gain compared to the results from Paiva et al. (2014), who mentioned that the model fails to detect when the wind turns completely from the southwest direction, even using numerical grids with higher horizontal resolution (300 m). On the other hand, the high frequency of the north winds simulated by the WRF (Fig. 6f) may be a model response to represent the land breeze phenomenon since the meridional component of the wind has an important weight in the composition of the northeast winds. Other possibilities raised here are an excessive weight given by the model to the flow of synoptic-scale that produces winds from the north in this region, or even a flow induced by mountain breezes to the north of this region (Aragão et al. 2017). Regarding the wind speed, both observed and modeled results indicate that there are two most frequent classes, 2.1 to 3.6 m.s⁻¹ and 3.6 to 5.7 m.s⁻¹ (Fig. 6f). It is also verified that the model represents only about 50% of the cases of calms observed in the SBSC station.

Although the WRF model represented the near-surface wind regime with characteristics of the breeze systems and the most frequent wind directions as the observations (Fig. 6), in most of the evaluated sites, the percentage of the north and east components of the wind were overestimated. Over the coastal region, the near-surface winds resulted from the combined contribution of the synoptic forcing, predominantly the SASA, and the thermally induced mesoscale forcing, the land-sea breeze (Dragaud et al. 2019). As the SASA induces the near-surface winds to blow from the first quadrant (0°–90°), probably the simulations are representing a greater influence of this system on the local circulation. These discrepancies in the representation of the local winds could be related to the synoptic-scale influence through the atmospheric boundary conditions (Giannaros et al. 2017), underestimation of the local land-sea thermal gradient through the ocean (Dragaud et al. 2019) and land boundary conditions, and a miss-representation of the terrain effects on the wind (Paiva et al. 2014).

In Table 3, the statistical indexes obtained for wind direction and speed are presented. As a most notable result, the bias index indicates a systematic trend of the model for all evaluated sites: a counterclockwise deviation of −4° to −12° about the wind direction observations. This trend of deviation in the wind direction agrees with the differences shown in the wind roses, where the model results always tend to deviate counterclockwise from observations, as demonstrated in the results for SBRJ, SBJR, and SBSC stations (Fig. 6), that is, the three stations most influenced by the land-sea breeze. This may be evidence that the model has difficulties in representing mesoscale systems such as breezes, which in turn, maybe due to the excess weight given to the synoptic forces, or even the poor representation of the land-sea temperature gradient, which was remarked by Dragaud et al., (2019) for the coastal region of Rio de Janeiro state. Considering the bias and RMSE indexes the wind speed results are better in general than those demonstrated in previous literature by Hanna and Yang (2001), Zhang et al. (2011), Paiva et al. (2014), and Yver et al. (2013), even considering the FDFA technique applied in Hanna and Yang (2001) and the RJMA ultra-high-resolution surface databases used by Paiva et al. (2014). The values for bias evidence a slight overestimation of the simulated winds. According to Jiménez and Dudhia (2012), this pattern of stronger simulated winds is frequently located over plains and valleys in complex terrain regions. Based on the criteria established by Clifford (2011) (Table 2), it can be verified in Table 3 that during the 5 months where the WRF model skill was evaluated, its simulated results were considered excellent. The only exception was the RMSE index for SBSC (Table 3), considered a good result according to Clifford (2011). Since the results obtained through the WRF model demonstrate satisfactory performance and even

| Stations | Wind speed (m.s⁻¹) | Wind direction (°) |
|----------|--------------------|--------------------|
|          | Bias   | MAE   | RMSE  | Bias   | MAE   | RMSE  |
| SBGL     | −1.00  | 1.46  | 1.85  | −10.27 | 56.52 | 73.77 |
| SBRJ     | 0.25   | 1.36  | 1.80  | −4.32  | 52.01 | 71.65 |
| SBAF     | 0.07   | 1.15  | 1.51  | −7.42  | 67.40 | 85.08 |
| SBJR     | 0.96   | 1.34  | 1.69  | −12.57 | 53.74 | 73.93 |
| SBSC     | 0.32   | 1.61  | 2.10  | −12.47 | 60.79 | 80.09 |
superior to studies with ultra-high resolution and FDDA, it is assumed that the uncertainties from the atmospheric model are acceptable for air quality modeling.

**Evaluation of the SO₂ hourly average concentrations results by using CONFIG 1, CONFIG 2, and CONFIG 3**

In Table 4, the statistical results for the simulated concentrations with CONFIG 1, CONFIG 2, and CONFIG 3, at each AQ monitoring station, are shown. The analysis related to the NMSE index demonstrates that the CONFIG 3 performs better in all AQ stations and the CONFIG 1 presents the worst results except for the São Cristóvão station. It should also be noted that the NMSE values obtained for CONFIG 3 are of the same order of magnitude or lower than those obtained by Cui et al. (2011), a study developed for the region of the Gan Jiang River, in China, whose values varied between 3 and 6. Systematically, all values for the bias index are positive, with CONFIG 2 results being the worst of the three settings strategies. Although the bias indicates that the simulations overestimated the concentrations observed, the FOEX index demonstrates that the number of overestimated events (hours) is not always more frequent than the underestimated events. In this case, the results obtained for the Jardim Primavera station stand out, in which the underestimated events for CONFIG 1 and CONFIG 3 occur in 69% and 65% of the cases. Still discussing FOEX, the results indicate that CONFIG 3 presents the best balance between the number of underestimated and overestimated events.

Regarding the FA2 and FA5 indexes, in general, the results presented for CONFIG 3 are also better, while the results for CONFIG 2 are the worst. Despite important recent advances in atmospheric models, it is noted that the isolated use of simulated results in AQ models may not be the best strategy. The values of FAC2 and FAC5 attained through CONFIG 3 suggest a better representation of the simulated pollutant dispersion than those obtained by Cui et al. (2011). Its maximum values were 27% and 61% for FAC2 and FAC5, respectively, which reinforces the model’s inclination to better results with this setting strategy. The standard deviation analysis indicates that, for all air quality monitoring stations, the magnitude of the simulated index is higher than that observed with the air quality monitoring data, with greater discrepancy on the Jardim Primavera and São Bento stations.

The results for the R index (Table 4) and the daily cycle of SO₂ (Fig. 7) demonstrate that the simulations are outdated in relation to that observed for all air quality monitoring stations, and there is a significant overestimation of the SO₂ concentration for the stations of São Cristóvão, São Bento, and Jardim Primavera. It is not surprising the R index values close to 0, since the present evaluation is audacious in the sense of trying to model a highly urban environment with infinite casualties, unlike well-controlled experiments. However, the decision to present the correlation result is to encourage future studies and more investments in the data used in non-steady-state dispersion models. To be successful in modeling the variability of pollutant concentrations in the atmosphere, two main factors must be well represented: (1) emissions and (2) advective and turbulent transport.

| AQ Stations     | Std. dev. (μg.m⁻³) | Runs       | Std. dev. (μg.m⁻³) | NMSE | Bias (μg.m⁻³) | FOEX (%) | FAC2 (%) | FAC5 (%) | R     |
|-----------------|---------------------|------------|---------------------|------|--------------|----------|----------|----------|-------|
|                 |                     | CONFIG 1   | 4.63                | 2.33 | 1.61         | 1.00     | 39.0     | 77.0     | 0.16  |
|                 |                     | CONFIG 2   | 4.72                | 2.02 | 1.86         | 3.00     | 38.0     | 76.0     | 0.17  |
|                 |                     | CONFIG 3   | 4.02                | 1.79 | 1.25         | 0.00     | 41.0     | 80.0     | 0.20  |
| Copacabana      | 2.52                |            |                     |      |              |          |          |          |       |
| São Cristóvão   | 6.46                | CONFIG 1   | 8.68                | 3.19 | 1.31         | −5.00    | 40.0     | 75.0     | 0.10  |
|                 |                     | CONFIG 2   | 19.92               | 5.49 | 13.85        | 16.00    | 33.0     | 66.0     | 0.05  |
|                 |                     | CONFIG 3   | 9.08                | 3.15 | 3.20         | 5.00     | 43.0     | 79.0     | 0.08  |
| Tijuca          | 3.54                | CONFIG 1   | 5.07                | 2.35 | 0.68         | −9.00    | 28.0     | 62.0     | 0.18  |
|                 |                     | CONFIG 2   | 5.78                | 2.09 | 2.40         | 5.00     | 33.0     | 68.0     | 0.18  |
|                 |                     | CONFIG 3   | 5.11                | 1.84 | 1.77         | 4.00     | 35.0     | 70.0     | 0.20  |
| Jardim Primavera| 3.01                | CONFIG 1   | 10.54               | 8.70 | 2.68         | −19.00   | 33.0     | 69.0     | 0.05  |
|                 |                     | CONFIG 2   | 16.69               | 8.18 | 10.31        | 6.00     | 32.0     | 68.0     | 0.00  |
|                 |                     | CONFIG 3   | 9.38                | 5.44 | 2.21         | −15.00   | 33.0     | 73.0     | 0.06  |
| São Bento       | 5.35                | CONFIG 1   | 13.78               | 7.59 | 6.71         | 6.00     | 32.0     | 68.0     | 0.08  |
|                 |                     | CONFIG 2   | 17.23               | 6.31 | 11.35        | 13.00    | 27.0     | 57.0     | 0.11  |
|                 |                     | CONFIG 3   | 15.14               | 5.16 | 8.44         | 8.00     | 32.0     | 66.0     | 0.16  |
The first and perhaps the most important factor that impacts the correlation index is associated with the better representation of temporal variability of the emission sources. Although SO$_2$ vehicular emissions are small compared to fixed sources, and there is an hourly function for these emissions in the model, it is still an average representation, which does not include casual traffic jams, traffic diversions, and the difference between weekdays and weekends. Then, it is not surprising that there are lags between the maximum and minimum values and discrepancies in the intensity between the simulated and observed hourly concentration, as shown in Fig. 7. As for industrial sources, probably they are not responsible for the poor correlation found, since the oil refineries are the main sources of SO$_2$ in RJMA and operate at constant load (INEA 2010), with exceptions for eventual maintenance.

The second factor is related to pollutant transport. Based on WRF results, the advective transport was satisfactorily represented in the present study. The turbulent transport and the atmospheric stability were not evaluated in this study, but as the representation of vehicle emissions, they can be sources of uncertainty. However, the only available upper-air data, i.e., the soundings at SBGL, are inadequate to evaluate the atmospheric stability of the WRF simulations since the soundings were only launched at 9 am and 9 pm which is during the assembly and disassembly of the planetary
boundary layer. Thus, the investigation of these factors deserves further specific study using other data sources (e.g., SODAR, micrometeorological stations, etc.). The use of meteorological data with higher temporal resolution would allow the simulation of the transient effects of greater frequency (< 1 h), and more meteorological stations could accurately represent the spatial inhomogeneities of the area.

The bias of hourly wind speed and hourly SO₂ concentrations by observed wind speed categories (Fig. 8) reveals important information that can contribute to a better understanding of the model errors and consequently indicate possible paths to be taken in future studies. According to Fig. 8 (left), it is observed that the smallest errors in wind intensity occur between 0.5 and 5.7 m·s⁻¹, whereas high deviations occur for the calm winds and speeds above 5.7 m·s⁻¹. It is noteworthy that no significant differences were found between the three configurations evaluated. Note that for the wind class with the smallest bias (2.1–3.6 m·s⁻¹) (Fig. 8, left), the smallest deviations for SO₂ concentration also occur, mainly for CONFIG 1 and CONFIG 3 (Fig. 8, right). Despite the similar bias values regarding the wind speed among the three configurations (Fig. 8, left), CONFIG 2 presents the biggest discrepancies for SO₂ concentration (Fig. 8, right). These analyzes highlight the importance of a meteorological monitoring network representative for wind and micrometeorological conditions.

**Outlook for regulatory purposes**

For regulatory purposes, the air quality standards in Brazil for the sulfur dioxide pollutant are established for 24 h (125 μg.m⁻³) and annual sampling times (40 μg.m⁻³). Thus, in this section, the results are analyzed considering the comparison between the simulated and observed average concentrations of 24 h, for the entire period of analysis, aiming to establish the performance of the CAQMS for practical applications.

In general, from the perspective of the evaluation of meteorological configurations among the simulations for regulatory purposes (Table 5 and Fig. 9), a similar pattern of results was obtained for average hourly concentrations (Table 4). The best performances were obtained for configurations that used observed meteorological data (CONFIG 1 and CONFIG 3), with a slight superiority for CONFIG 3. However, it is notable that the results for 24-h averages (Table 5) are significantly better than those presented for hourly averages (Table 4). This corroborates with the previous argument that the difficulty of representing meteorological and emissions variations on minutes to hours scales impacts the results of the simulations.

As can be seen in Table 5 and Fig. 9, the simulation results overestimate the average concentrations over the entire period at all air quality monitoring stations. It is evident in Fig. 9 that the simulated concentrations rarely underestimate the observed concentrations. This result is reinforced by the overall positive bias index, whose all values are positive. Another information presented that is of paramount importance for regulatory studies is the model’s ability to represent the maximum concentrations. As shown in Table 5, all maximum concentrations simulated overestimated those observed, reinforcing the tendency for overestimations. These errors follow the concentration levels of the stations, that is, the higher the concentration level at the station, the greater the associated error. Despite the models showing a positive bias, the simulated concentrations did not exceed the national air quality standard, in agreement with what was observed (Table 5).

Depending on the intended objectives, the evaluation of an air quality model may vary in terms of metrics and ways of comparison. For example, in military applications, in general, the main objective is to evaluate the performance of a model to predict the location of the danger area (Chang and Hanna 2004). In regulatory applications, the focus is on evaluating the performance of a model in predicting
Table 5 Statistical results for CALPUFF’s daily averaged SO₂ concentrations (μg.m⁻³), data pairing in time, and space

| AQ stations       | Mean/max obs. (μg.m⁻³) | Runs          | Mean/max sim. (μg.m⁻³) | NMSE | Bias (μg.m⁻³) | FOEX (%) | FAC2 (%) | FAC5 (%) | R   |
|-------------------|------------------------|---------------|------------------------|------|---------------|----------|----------|----------|-----|
| Copacabana        | 3.83 / 14.08           | CONFIG 1      | 5.41 / 17.28           | 0.76 | 1.54          | 14.0     | 58.0     | 95.0     | 0.09|
|                   |                        | CONFIG 2      | 5.69 / 16.73           | 0.59 | 1.83          | 18.0     | 61.0     | 96.0     | 0.31|
|                   |                        | CONFIG 3      | 5.06 / 16.89           | 0.52 | 1.21          | 11.0     | 69.0     | 97.0     | 0.30|
| São Cristóvão     | 8.36 / 25.16           | CONFIG 1      | 9.49 / 36.09           | 0.79 | 1.09          | −2.0     | 56.0     | 90.0     | 0.03|
|                   |                        | CONFIG 2      | 21.64 / 61.71          | 1.44 | 13.24         | 44.0     | 40.0     | 79.0     | 0.16|
|                   |                        | CONFIG 3      | 11.27 / 37.53          | 0.72 | 2.67          | 9.0      | 64.0     | 93.0     | 0.03|
| Tijuca            | 4.45 / 14.55           | CONFIG 1      | 5.12 / 19.82           | 0.68 | 0.73          | 4.0      | 53.0     | 96.0     | 0.31|
|                   |                        | CONFIG 2      | 6.86 / 17.74           | 0.56 | 2.46          | 26.0     | 60.0     | 93.0     | 0.41|
|                   |                        | CONFIG 3      | 6.27 / 22.11           | 0.50 | 1.82          | 23.0     | 64.0     | 96.0     | 0.45|
| Jardim Primavera  | 6.47 / 21.39           | CONFIG 1      | 9.32 / 53.37           | 1.78 | 2.87          | −5.0     | 47.0     | 88.0     | 0.14|
|                   |                        | CONFIG 2      | 16.85 / 63.42          | 2.41 | 10.49         | 30.0     | 39.0     | 80.0     | −0.03|
|                   |                        | CONFIG 3      | 8.84 / 44.15           | 1.43 | 2.43          | −3.0     | 47.0     | 88.0     | 0.06|
| São Bento         | 6.19 / 27.56           | CONFIG 1      | 12.84 / 55.08          | 1.78 | 6.67          | 25.0     | 47.0     | 80.0     | 0.13|
|                   |                        | CONFIG 2      | 17.72 / 50.10          | 1.96 | 11.45         | 36.0     | 32.0     | 71.0     | 0.23|
|                   |                        | CONFIG 3      | 14.74 / 45.12          | 1.45 | 8.45          | 33.0     | 42.0     | 81.0     | 0.36|

Fig. 9 Timeseries of daily averaged SO₂ concentrations (μg.m⁻³) between all simulated (dashed lines) and observed (shaded lines) sites.
quantitative, specifically, the maximum concentrations, regardless of location and period of occurrence (USEPA 1992). In this sense, since the 1980s, several scientific studies have been dedicated to discussing the best metrics and ways of evaluating air quality models (Fox 1984; Hanna 1989; Cox and Tikvant 1990; USEPA 1992; Hanna et al. 1991, 1993); Mosca et al. 1998; Arya 1999; ASTM 2000; Chang and Hanna 2004). However, the work of Chang and Hanna (2004) stands out in the literature for presenting a detailed review about different methods for evaluating the performance of air quality models and proposing the best methods to be used. According to this study, the recommended statistical indices are fractional bias (FB), the geometric mean bias (MG), the normalized mean square error (NMSE), the geometric variance (VG), the Spearman correlation coefficient (Rspm), and the fraction of predictions within a factor of two of observations (FAC2). Based on the above, to present an evaluation of the simulations for regulatory purposes, the next results (Table 6) are presented according to the indices cited by Chang and Hanna (2004), considering the pairing only in the space between the observed and simulated data, that is, arranged in descending order, regardless of time. Since the classic experiments considered in most studies to evaluate air quality models consist of uniformly distributed receivers, it was decided to keep the pairing in space for the present study, due to the low density and non-uniform distribution of the air quality monitoring network.

In general, the results presented for the regulatory assessment (Table 6) indicate the same pattern as previously described, that is, the CONFIG 1 and 3 configurations present superior performance than the CONFIG 2. About meeting the typical criteria and limits exposed by the scientific community, USEPA (1992) proposed that the first criterion to be met by a regulatory model is to present FB values between −0.67 and 0.67, that is, within a factor of 2. From this criterion, most of the results fit, except for the results for the São Bento station and the result of CONFIG 2 for the Jardim Primavera station.

The FAC2 index can be considered the most robust statistical measure among those presented in Table 6 since it is not excessively influenced by outliers. For this index, the percentages obtained are above or close to 50% for the most part (Table 6), fitting within the typical satisfactory percentages according to Chang and Hanna (2004). It is noteworthy that the results paired in space and time presented in Table 5 also fit for the most part within the 50% limit for FAC2.

A perfect model will have values for MG and VG equal to 1.00. According to the study by Hanna et al. (2001), who evaluated three regulatory models under five classic experiments, that is, under emission conditions much more controlled than those considered in the present study, the average performance for the three models varied from 0.7 to 1.77 for MG, and from 2.4 to 7.7 for VG. From these reference values, it can be seen in Table 6 that the results fit for most stations considering CONFIG 1 and CONFIG 3. Regarding the VG index, it is highlighted that the results of the present study are better than those obtained by Hanna et al. (2001).

For the NMSE index, it appears that not only the results presented in Table 6 but also those in Table 4 and 5 are compatible or even better than the results presented in several other studies. (Hanna 1989; Cui et al. 2011; Ghannam and El-Fadel 2013; Hernández-Garcés et al. 2021).

Among the statistical indices cited by Chang and Hanna (2004), the correlation index is perhaps the one that most generates discussion about its application in regulatory assessments. The correlation coefficient, $R$, reflects the linear relationship between two variables but does not consistently

| AQ stations          | Runs | FB   | FAC2 | MG   | VG   | NMSE | Rspm |
|----------------------|------|------|------|------|------|------|------|
| Copacabana CONFIG 1 | −0.33| 1.00 | 0.74 | 1.10 | 0.18 | 1.00 |
| CONFIG 2             | −0.38| 1.00 | 0.73 | 1.17 | 0.23 | 1.00 |
| CONFIG 3             | −0.27| 1.00 | 0.80 | 1.07 | 0.12 | 1.00 |
| São Cristóvao CONFIG 1 | −0.12| 1.00 | 0.90 | 1.02 | 0.05 | 1.00 |
| CONFIG 2             | −0.88| 0.05 | 0.35 | 3.05 | 1.09 | 1.00 |
| CONFIG 3             | −0.27| 0.93 | 0.72 | 1.16 | 0.11 | 1.00 |
| Tijuca CONFIG 1      | −0.15| 1.00 | 0.95 | 1.03 | 0.10 | 1.00 |
| CONFIG 2             | −0.43| 1.00 | 0.64 | 1.25 | 0.24 | 1.00 |
| CONFIG 3             | −0.34| 1.00 | 0.71 | 1.13 | 0.16 | 1.00 |
| Jardim Primavera CONFIG 1 | −0.36| 0.77 | 1.03 | 1.40 | 0.96 | 1.00 |
| CONFIG 2             | −0.90| 0.27 | 0.46 | 2.29 | 1.68 | 1.00 |
| CONFIG 3             | −0.32| 0.80 | 1.02 | 1.40 | 0.57 | 1.00 |
| São Bento CONFIG 1  | −0.70| 0.36 | 0.50 | 1.71 | 0.87 | 1.00 |
| CONFIG 2             | −0.96| 0.16 | 0.35 | 3.45 | 1.54 | 1.00 |
| CONFIG 3             | −0.81| 0.26 | 0.44 | 2.18 | 1.12 | 1.00 |
relate to forecasting accuracy. Furthermore, this index is sensitive to extreme data pairs and can be significantly influenced by a few outliers. Therefore, when applying such an index, the use of Spearman’s correlation coefficient is recommended, which considers the ranks rather than the values themselves (Chang and Hanna 2004). Another point reported in the study by Chang and Hanna (2004) is that in classical experiments, it is common to have concentration data measured along concentric arcs at various uniform distances from the source. In this case, the value of $R$ can be deceptively high, mainly reflecting the fact that the concentration decreases with distance downwind, which any reasonable dispersion model will simulate. On the other hand, when considering both the matching in time and space, the correlation can be very low. Since small deviations in the direction of the simulated winds and approximations assumed regarding the variability of emissions, the plumes are positioned incorrectly, consequently, simulating maximum and minimum concentration at incorrect times and places. This condition was evident in the results of Table 5. This situation can be even more inconsistent when non-homogeneous and steady-state models are submitted to such evaluations, which is the case of pure Gaussian models, the majority among regulatory models. For these and other reasons, the use of the correlation coefficient is discouraged in regulatory assessments (Willmott 1982); however, from a scientific perspective, it is useful to indicate limitations, consequently, it makes it possible to apply improvements. As shown in Table 6, the results for the Spearman correlation coefficient are all equal to 1, however, it should be noted that this evaluation does not consider the pairing in time; therefore, it is not useful for the investigation of the best meteorological configuration in the CALPUFF.

Figure 10 reinforces what has already been shown in Fig. 9. In the study region, SO$_2$ levels remain predominantly below 6 μg·m$^{-3}$ in 24 h (~65%), that is, in the two lowest concentration classes of the graph (Fig. 10). As for the bias of the simulations by frequency class, in general, the model errors more towards extreme concentrations, that is, tending to zero and >12 μg·m$^{-3}$ (Fig. 10). However, it should be noted that CONFIG 3 has the smallest errors for the extremes, while CONFIG 1 has the best performance for the intermediate classes. In all classes, CONFIG 2 has the worst performance with errors twice as large as CONFIG 1 and CONFIG 3 in most concentration frequency classes.

**Conclusions**

The study demonstrates that the configuration of the WRF model for the RJMA was successful in the representation of the near-surface wind field. An evaluation was carried out for a robust period of 5 months that comprises the main meteorological systems, in which a satisfactory agreement was demonstrated between the prevailing winds and the simulated winds by frequency distribution analysis. In a complementary way, it was found that the results from the statistical analysis showed a model performance comparable to and even superior to other studies, including for the same region using a higher resolution of terrain databases and horizontal grid resolution. It is worth remembering that RJMA is recognized for its high physiographic complexity, which makes the results found even more relevant.

The WRF model represented the wind regime with characteristics of the breeze systems and the most frequent wind directions as the observations, in most of the sites, however, the percentage of the north and east components of the wind were overestimated. As the SASA induces the near-surface winds to blow from the first quadrant (0°–90°), probably the simulations are representing a greater influence of this system on the local circulation. These discrepancies in the representation of the local winds may be related to the synoptic-scale influence through the atmospheric boundary conditions, underestimation of the local land-sea thermal gradient through the ocean and land boundary conditions, and a miss-representation of the terrain effects on the wind. The indication that the mesoscale forcing (land/sea breeze) is underestimated in relation to the synoptic forcing (SASA), or the overestimation of the latter, is an important highlight that remains for future studies.

Regarding the best meteorological configuration for the CAQMS, CONFIG 3 stands out from the other configurations, with a positive influence when considering
observational meteorological information as well as the simulated ones, and thus a better physical representation of the SO₂ dispersion process in the RJMA. Unlike CONFIG 2, which only includes WRF model data, this result reinforces the importance of investing in full-time meteorological monitoring.

As for the skill of the CAQMS to simulate hourly SO₂ concentrations in RJMA, in general, the concentrations simulated overestimated the observed data. The results for the R index exposed the challenge in the temporal representation of the observed concentrations (on the hourly scale). The hypothesis is that the main source of uncertainties is the limited representativeness of the emissions regarding time variability. Besides, the unavailability of high-frequency meteorological data that records the transients in intra-hourly variability could also bring gains in transport representation. This theme merits a deep investigation, also expanding the analysis to other variables such as atmospheric stability. To encourage future studies, recently, two SODAR were installed in strategic sites in RJMA that could contribute significantly to the atmospheric conditions’ assessment on pollutant concentration fields. However, the present results are very promising, since the region has a rough terrain setting, non-homogenous soil use, and occupation, presence of water bodies, time–space-varying emissions originated from several source types, besides the casualties of a highly urbanized region. The performance of the present study is comparable to other studies, which developed a study with controlled emissions, different from the current study that tried to represent the main anthropogenic emissions.

From a regulatory point of view, there was a significant improvement in the results towards those obtained for hourly averages. This result indicates a great potential of the WRF/CALMET/CALPUFF modeling system for regulatory purposes. Since RJMA has a high degree of physiographic complexity and consequently a heterogeneous wind field, it is recommended that the competent environmental agencies review the current policies, which end up recommending the use of steady-state Gaussian models indiscriminately without specific legislation. The WRF/CALMET/CALPUFF air quality regulatory modeling system has proven to be a potential tool for environmental licensing studies and subsidized land use policies in Brazilian Metropolitan Regions aiming to achieve the UN 2030 Agenda goals.

Acknowledgments The INEA (State Environmental Institute) and SMAC (Municipal Secretary of the Environment of Rio de Janeiro) by the air quality monitoring data. We would also like to thank the three anonymous reviewers for providing insightful comments that led to improvements in the article, as well as the Editor and the editorial staff of Environmental Science and Pollution Research for their support.

Author contribution Mauricio Soares da Silva: conceptualization, methodology, software, formal analysis, investigation, data curation, writing—original draft, writing—review and editing and visualization. Luiz Claudio Gomes Pimentel: conceptualization, methodology, validation, formal analysis, investigation, resources, writing—original draft, writing—review and editing, supervision, project administration. Fernando Pereira Duda: conceptualization, formal analysis, and writing—review and editing. Leonardo Aragão: methodology, validation, formal analysis, investigation, resources, writing—review and editing. Corbiniano Silva: resources, writing—original draft, writing—review and editing, visualization. Ian Cunha D’Amato Viana Dragaú: methodology, investigation, writing—review and editing. Pedro Caffaro Vicentini: data curation, methodology, investigation, writing—review and editing.

Funding This work was conducted during a scholarship supported by the CAPES (Brazilian Federal Agency for Support and Evaluation of Graduate Education within the Ministry of Education of Brazil) and CNPQ (National Council for Scientific and Technological Development) at the Federal University of Rio de Janeiro.

Data availability The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare no competing interests.

References

Abdul-Wahab AS, Ali S, Sardar S, Irfan N, Al-Damkhin A (2011a) Evaluating the performance of an integrated CALPUFF-MMS modeling system for predicting SO₂ emission from a refinery. Clean Techn Environ Policy 13:841–854. https://doi.org/10.1007/s10098-011-0360-6
Abdul-Wahab S, Sappurd A, Al-Damkh A (2011b) Application of California Puff (CALPUFF) model: a case study for Oman. Clean Techn Environ Policy 13:177–189. https://doi.org/10.1007/s10098-010-0283-7
Acuti D, Bellucci M, Manetti G (2020) Company disclosures concerning the resilience of cities from the Sustainable Development Goals (SDGs) perspective. Cities 99:102608
Ajtai N, Stefanie H, Botezan C, Ozunu A, Radovici A, Dumitrache R, Iriza-Burcă A, Diamandi A, Hirtl M (2020) Support tools for land use policies based on high resolution regional air quality modelling. Land Use Policy 95:103909. https://doi.org/10.1016/j.landusepol.2019.03.022
Aragão LFS, Di Sabatino S, Pimentel LCG, Duda FP (2017) Analysis of the internal boundary layer formation on tropical coastal regions using SODAR data in Rio de Janeiro (Brazil). Int J Environ Pollut 62:136–154. https://doi.org/10.1504/IJEP2017.089403
Yang D, Wang A, Zhang R (2008) Estimating air quality impacts of elevated point source emissions in Chongqing, China. Aerosol Air Qual Res 8:279–294. https://doi.org/10.4209/aair.2008.02.0005
Arino O, Ramos Perez JJ, Kalogirou V, Bontemps S, Defourny, P, Van Bogaert E (2012) Global Land Cover Map for 2009 (GlobCover 2009). European Space Agency (ESA) & Université catholique de Louvain (UCL), PANGAEA. https://doi.org/10.1594/PANGAEA.787668
Arya SP (1999) Air pollution meteorology and dispersion. Oxford University Press, New York, p 310
from agglomeration in the property market. Land Use Policy 84:177–191

Hohes NS, Morawska L (2006) A review of dispersion modelling and its application to the dispersion of particles: an overview of different dispersion models available. Atmos Environ 40:5902–5928. https://doi.org/10.1016/j.atmosenv.2006.06.003

Holnicki P, Nahorski Z (2015) Emission Data Uncertainty in Urban Air Quality Modeling—Case Study. Environ Model Assess 20(6):583–597. https://doi.org/10.1007/s10666-015-9445-7

Hong S-Y, Dudhia J, Chen S-H (2004) A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation, monthly weather review. Mon Weather Rev 132:103–120. https://doi.org/10.1175/1520-0493(2004)132%3C1016:ARATIM%3E2.0.CO;2

Hong S-Y, Noh N, Dudhia J (2006) A new vertical diffusion package with an explicit treatment of entrainment processes. Mon Weather Rev 134:2318–2341. https://doi.org/10.1175/MWR3199.1

Iacono MJ, Delamere JS, Mlawer EJ, Shephard MW, Clough SA, Collins WD (2008) Radiative forcing by long–lived greenhouse gases: calculations with the AER radiative transfer models. J Geophys Res 113:D13103. https://doi.org/10.1029/2008JD009944

IBGE (2020) Territorial units at the metropolitan region level. Instituto Brasileiro de Geografia e Estatística. https://sbida.ibge.gov.br/territorio/#/N7. Accessed 11 June 2020. (In Portuguese)

INEA (2010) Annual air quality report for the state of Rio de Janeiro - 2009. Instituto do Meio Ambiente do Estado do Rio de Janeiro (State Environmental Institute), Rio de Janeiro, p 108. (In Portuguese)

INEA (2015) Atlas of Protected Areas in Rio de Janeiro State. Instituto do Meio Ambiente do Estado do Rio de Janeiro (State Environmental Institute). (In Portuguese). https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwj36tZ9Z75AhXZtCkhGHeqQDeEQFnoECCAQAQ&url=http%3A%2F%2Fwww.ineajrj.gov.br%2Fcss%2Fisd pl%3DService%3DGET_FILE%3D%26id%3D133385%26Do cName%3D%26i%253D0%253C0170%2F%253D01251%2F1&usg=AOvVaw089hnN5cvex7MBpk1T7F1U2. Accessed 29 July 2022

Jackson B, Chau D, Gurer K, Kadiuvela A (2006) Comparison of ozone simulations using MM5 and CALMET/MM5 hybrid meteorological fields for the July/August 2000 CCOS episod. Atmos Environ 40:2812–2822. https://doi.org/10.1016/j.atmosenv.2006.01.020

Jiménez PA, Dudhia J (2012) Improving the representation of resolved and unresolved topographic effects on surface wind in the WRF model. J Appl Meteorol Climatol 51:300–316. https://doi.org/10.1175/JAMC-D-11-084.1

Jiménez PA, Dudhia J, Gonzalez Rouco JF, Navarro J, Montavez JP, Garcia-Bustamante E (2012) A revised scheme for the WRF surface layer formulation. Mon Weather Rev 140:898–918. https://doi.org/10.1175/MWR-D-11-00056.1

Kain JS (2004) The Kain-Fritsch convective parameterization: an update. J Appl Meteorol Climatol 43:170–181. https://doi.org/10.1175/1520-0450(2004)043%3C0170:TFCPA%3E2.0.CO;2

Kjer JFS, Beike CH, Dias GTM, Filippo AM, Da Silva QV (1997) Oceanographic characteristics of an impacted coastal bay: Baia de Guanabara, Rio de Janeiro, Brazil. Cont Shelf Res 17:1609–1643. https://doi.org/10.1016/S0278-4343(97)00028-9

Kjerfve B, Dias GTM, Filippo A, Geraldes MC (2021) Oceanographic and environmental characteristics of a coupled coastal bay system: Baia de Ilha Grande-Baía de Sepetiba, Rio de Janeiro, Brazil. Reg Stud Mar Sci 41:101594. https://doi.org/10.1016/j.rsma.2020.101594

Klausmann AM, Groot W, Phadnis M, Scire JS (2003) A real-time meteorological analysis and dispersion prediction system for
Soares da Silva M, Cunha BD, Pimentel LCG, Pérez-Guerrero JS, Vicentini PC (2014) Comparative analysis of sulfur dioxide concentration by AERMOD and CALPUFF regulatory models and monitoring data near the emission sources at air basin III - MRJ. Revista Brasileira de Meteorol 29(2):281–298

Soares da Silva M, Pimentel LCG (2017) Desafios e estratégias de controle da qualidade do ar em regiões metropolitanas. Diversidade e Gestão, 1:107–126. Volume Especial. Gestão Ambiental: Perspectivas, Conceitos e Casos. (in Portuguese) https://doi.org/10.29327/538126

Sobral BS, de Oliveira Júnior JF, Gois G, de Bodas Terassi PM, Pereira CR (2018) Regime de Vento na Serra do Mar – Rio de Janeiro, (in Portuguese) Brasil. Rev Bras Meteorol 33:441–451. https://doi.org/10.1590/0102-7786333004

Steck JL, Lorenzetti JA (1992) The response of the South Brazil Bight to the passage of wintertime cold fronts. J Geophys Res 97:9507–9520. https://doi.org/10.1029/92JC00486

Sykes RE, Lewellen W, Parker S (1984) A turbulent-transport model for concentration fluctuations and fluxes. J Fluid Mech 139:193–218. https://doi.org/10.1017/S002211208400032X

Sykes RI, Parker SF, Henn DS, Gabrús RS (1996) SCIPUFF — A generalized dispersion model. In: Air pollution modeling and its application XI. Springer US, Boston, MA, pp 425–432. https://doi.org/10.1007/978-1-4615-5841-5_45

Thunis P, Miranda A, Baldasano JM, Dourou J, Graff A, Janssen S, Juda-Rezlerk K, Karvonenja N, Maffeis G, Martilli A, Rasolohariramea M, Reul E, Vane E, Volta M, White I (2016) Overview of current regional and local scale air quality modelling practices: assessment and planning tools in the EU. Environ Sci Pol 65:13–21

UN (2015) Resolution adopted by the general assembly on 25 September 2015. In: Transforming our world: The 2030 agenda for sustainable development. United Nations. http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E. Accessed 29 July 2022

UN (2018) Department of economic and social affairs, population division. In: The World’s Cities in 2018 – Data Booklet (ST/ESA/SER.A/417). United Nations. https://www.un.org/en/events/citie Asset/pdf/the_worlds_cities_in_2018_data_booklet.pdf. Accessed 21 April 2021

USEPA (1992) Protocol for determining the best performing model. United States. EPA-454/R-92-025. Environmental Protection Agency, North Carolina

USEPA (1995) User’s guide for the industrial source complex (ISC3) Dispersion Models (revised) Volume II – Description of model Algorithms. EPA-454/b-95-003b, United States Environmental Protection Agency, North Carolina

USEPA (2004) AERMOD: description of model formulation. EPA-454/R-03-004. United States Environmental Protection Agency, North Carolina

Vicentini PC (2011) Assessing the PROCONVE effect in Rio de Janeiro metropolitan area since 2008 to 2020 by means of air quality models. Dissertation, Federal University of Rio de Janeiro. (in Portuguese)

Vicentini PC, Pimentel LCG, Soares da Silva M, Souza Pontes A (2011) Effects of PROCONVE in the Metropolitan Region of Rio de Janeiro between 2008 and 2020 - evaluation using an emissions inventory and an air quality model. In: XIX Simpósio Internacional de Engenharia Automotiva - SIAMEA, v. 1 (in Portuguese)

Wang S, Hao J (2012) Air quality management in China: issues, challenges, and options. J Environ Sci 24(1):2–13. https://doi.org/10.1016/S1001-0742(11)60724-9

Willmott CJ (1982) Some comments on the evaluation of model performance. Bull Amer Meteor Soc 63:1309–1313

WHO (2016) Ambient air pollution: a global assessment of exposure and burden of disease. World Health Organization. https://apps.who.int/iris/handle/10665/250141

WHO (2022) Air pollution data portal. World Health Organization. https://www.who.int/data/gho/data/themes/air-pollution. Accessed 19 April 2022

WMO (2018) Guide to meteorological instruments and methods of observation, World Meteorological Organization. Volume 1 – Measurement of Meteorological Variables WMO-No. 8, 2018 edn. p 548. https://library.wmo.int/index.php?id=12407&lvl=notice_display. Accessed 29 July 2022

WMO (2020) The gaps in the Global Basic Observing Network (GBON). In: World Meteorological Organization, SOFF Series- No. 02. p 6. https://library.wmo.int/index.php?lvl=notice_display&idi=21771&YEN74begzaU1. Accessed 29 July 2022

WMO (2021a) Systematic observations financing facility (SOFF): draft terms of reference. World Meteorological Organization, 137 pages. https://library.wmo.int/index.php?lvl=notice_display&idi=21945#.YUMeEZWbtTX2. Accessed 29 July 2022

WMO (2021b) Ambient (outdoor) air pollution. World Meteorological Organization. https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health. Accessed 19 April 2022

Wu H, Zhang Y, Yu Q, Ma W (2018) Application of an integrated WRF/CALPUFF modeling tool for source apportionment of atmospheric pollutants for air quality management: a case study in the urban area of Benxi, China. J Air Waste Manage Assoc 2247. https://doi.org/10.1080/10473289.2017.1391009

Yadav R, Trivedi P, Sahu LK, Beig G, Tripathi N (2020) Air pollution modeling. In: In Air Pollution and Environmental Health. Springer, Singapore, pp 37–55

Yang Z, Wang J (2017) A new air quality monitoring and early warning system: air quality assessment and air pollutant concentration prediction. Environ Res 158:105–117

Yim SHL, Fung JCH, Lau AKH (2010) Use of high-resolution MM5/CALMET/CALPUFF system: SO2 apportionment to air quality in Hong Kong. Atmos Environ 44:4850–5858. https://doi.org/10.1016/j.atmosenv.2010.08.037

Yver CE, Graven HD, Lucas DD, Cameron-Smith PI, Keeling RF, Weiss RF (2013) Evaluating transport in the WRF model along the California coast. Atmos Chem Phys 13:1837–1852. https://doi.org/10.5194/acp-13-1837-2013

Zhang Y, Cheng S-H, Chen Y-S, Wang W-X (2011) Application of AERMOD: description of model formulation. EPA-454/R-03-004. United States Environmental Protection Agency, North Carolina

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under applicable law. This journal article may not be reproduced, adapted or transmitted by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without the written permission of Springer Nature or its licensor.

This article is protected by copyright. All rights reserved.