

RESEARCH

Open Access



Application of the WRF/Chem v.3.6.1 on the reanalysis of criteria pollutants over Metro Manila

Jacob Alberto Garcia^{1*} , Edgar Vallar¹, Maria Cecilia Galvez¹ and Gerry Bagtasa²

Abstract

Metro Manila, Philippines and other urban areas have reached internationally known unacceptable levels of pollution where about 80% can be attributed to vehicular emissions. The Weather Research and Forecasting model coupled with Chemistry v.3.6.1 was used in the reanalysis of pollutant concentrations for the year 2013. Initial results from the planetary boundary layer study suggested that the Yonsei University scheme provides a good estimate of the atmosphere's condition; hence, this setting was used for the succeeding simulations. The land coverage over Sangley point was not properly resolved by the model. This caused a cold bias for the station. Further evaluation of the model's sea level pressure output for all sites returned high correlations showing that modeled values are in phase with the observed time series; however, wind speed values did not correlate well with the observed values and were all overestimated. The low correlations found were a result of the incapability of the model to detect the urban canopy layer over Metro Manila. Pollutant concentrations were overestimated. The pollutant time series suggests that the model overestimates concentration values for PM₁₀, PM_{2.5}, and SO₂, while underestimating NO₂ and O₃ values. However, it does capture a significant 24-hourly cycle as seen in the time series' spectra in the frequency domain. Furthermore, through a student's t-test, the model also captures a significant difference in daytime and nighttime concentrations.

Keywords: WRF/Chem, Numerical modeling, Model bias, Planetary boundary layer, Spectral analysis, Pollutant concentrations

Introduction

Atmospheric models are complex mathematical and physical tools used for simulating or predicting meteorological phenomena. Controlled experiments cannot be performed in the atmosphere wherein a wide variety of processes occur. As expected, atmospheric scientists have developed a means of incorporating these processes in a representation of the atmosphere's state [1]. A scientific benchmark for understanding the outcomes of the universal physical laws is the ability to properly predict the result of an experiment. In the atmospheric sciences, this translates to the validity of numerical weather models (NWP) pertaining to their degree of accuracy and precision in creating forecasts of local, regional, and global scale meteorological events [2].

In line with this, various parameterizations are available to the users of atmospheric models that allow for different physical assumptions. These consider an assumption of an underlying process that cannot be resolved by the model grids. Some of the significant parameterizations include the following: cumulus, radiation, microphysics, and planetary boundary layer (PBL) parameterizations. For instance, cumulus parameterizations intend to explicitly resolve sub-grid scale clouds into the model. Similarly, microphysical parameterizations aim to include the effects of the microphysics in droplet formation and radiation parameterizations provide a method for estimating the total radiative flux at a given grid. One important configuration is that of the PBL. The PBL consists of the lower part of the atmosphere where most meteorological phenomena occur. This layer is directly influenced by the underlying surface, and its height is defined by the amount of

* Correspondence: jacob_garcia@dlsu.edu.ph

¹Physics Department, De La Salle University, 1004 Manila, Philippines
Full list of author information is available at the end of the article



turbulence present in the atmosphere from the bottom to the top. These turbulent fluxes of scalars such as heat, moisture, and momentum, allow for the diurnal evolution of the PBL. Moreover, these fluxes determine the amount of vertical mixing of boundary-layer variables due to eddy-transport and must, therefore, be properly estimated in numerical models [3, 4].

There are various applications of numerical weather prediction, and one of them is concerned with the estimation of pollutant concentrations and their dispersion over an area of interest [2]. As stated by the Department of Environment and Natural Resources - Environmental Management Bureau (DENR-EMB) [5], atmospheric pollution is the “alteration of the physical, chemical and biological properties of the atmosphere, or any discharge thereto of any liquid, gaseous, or solid substances that will or is likely to create or to render the air resources of the country harmful, detrimental, or injurious to public health, safety or welfare or which will adversely affect their utilization for domestic, commercial, industrial, agricultural, recreational, or other legitimate purpose”.

Emission sources vary by type and can be subcategorized as the following: stationary sources, mobile sources, and area sources. As documented by the European Environment Agency [6], area sources are a collective representation of individual stationary emission sources that are categorized due to their proximity relative to one another. Conversely, mobile sources are vehicles that are driven by the combustion of carbon-based fuels or other fuel types. The recent National Air Quality Status Report by the DENR-EMB [5] for the years 2008–2015 revealed that the majority of emissions stemmed from mobile sources (88%). Compared to stationary sources (10%) and area sources (2%), mobile sources make most of the emissions. As of 2015, there were a total of 93 monitoring stations across the Philippines with 27 of these being within Metro Manila. These monitoring stations are subdivided into different types as shown in Table 1 below.

The air quality monitoring stations present within Metro Manila obtain data based on the set National Ambient Air Quality Guideline Values (NAAQGV) for

criteria pollutants. These include TSP, PM₁₀, PM_{2.5}, SO₂, NO₂, O₃, CO, and Lead, all of which are emitted by the 3 main sources discussed previously [5]. Due to the spatial and temporal dependence of these emissions, it is important that these be documented in a spatial database known as an emissions inventory. With the availability of emissions inventories, it is possible to locate emissions hotspots within an area. Table 2 below presents a summary from a report by Biona et al. [7] showing the major sources of PM, SO₂, NO₂, and CO.

Atmospheric dispersion models are used in addressing the potential environmental and health-related risks. Epidemiological studies [8, 9] found that exposure to outdoor air pollution such as Particulate Matter (PM) enhanced the number of chromosome aberrations. The chemical composition of PM plays a big role in determining its carcinogenicity. Increased PM levels were revealed to induce mutations in the DNA of bacteria samples. This points out the relationship present with exposure from outdoor air pollution to genetic mutations, and changes in gene expressions; both of which contribute to an increased risk of cancer. Although the exact mechanism of action that allows for such an effect is still under investigation; it does suggest that the carcinogenicity of PM is linked to the suppression the DNA repair mechanism and the increase of DNA replication errors. Consequently, it is of utmost importance that dispersion models be investigated in terms of its performance in the reanalysis of surface pollutant concentration levels.

In Metro Manila and the surrounding areas, elevated levels of pollution have been observed; where Public Utility Jeepneys (20% of the total vehicular fleet) contribute to 94% of the total soot mass over the study site located in Metro Manila [10]. An overview of a scenario of increased pollutant concentration for PM₁₀ along the De La Salle University-Environmental Management Bureau (DLSU-EMB) station site during the northeast monsoon period from December to February is shown below through a set of bivariate plots in Figs. 1, 2. These were plotted using data from the DLSU-EMB station. The plots show the dependence of a pollutant on wind speed

Table 1 Types of air quality monitoring stations by the DENR

Air quality monitoring station types	Quantity
Manual/reference method – PM ₁₀	27
Manual/reference method – TSP	22
Continuous monitoring – DOAS (PM ₁₀ , PM _{2.5} , O ₃ , SO ₂ , NO ₂ , CO, and BTX)	14
Continuous monitoring van – 1 (CH ₄ , O ₃ , SO ₂ , NO ₂ , CO, and PM ₁₀)	3
Continuous monitoring – PMS (PM ₁₀ and PM _{2.5})	27
Total number of stations	93

DOAS Differential optical absorption spectroscopy

PMS Particle measuring system

Table 2 Pollutants and major sources

Pollutants	Source category with high shares	Major sources
PM	Area, stationary, and mobile	Cooking and other household activities; Road traffic
CO	Mobile	Road traffic
NO ₂	Mobile	Road traffic
SO ₂	Point	Power plant

(radial axis) and direction (polar axis) [11]. A map of the surrounding area is provided in Fig. 3.

An apparent dependence of high pollutant concentrations with wind speeds of around 4 to 8 m s⁻¹ from the easterly direction can be noticed. In addition, westerly wind speeds ranging from 2 to 6 m s⁻¹ also advect these pollutants toward the DLSU-EMB site, most likely due to the formation of a sea-breeze during the mid-day hours. Moreover, dividing the dataset into hourly intervals captures the daily high emission concentrations for PM₁₀ due to heavy traffic in Metro Manila. Further inspection of the subplots pertaining to hours 5 to 9 depicts the daily morning rush hour in Metro Manila. Additionally, it is also interesting to notice that the subplots during the nighttime hours do not capture the same emission concentration values. Incoming solar radiation decreases during these hours which hinders the formation of thermals and transport of pollutants.

These effects can be modeled through numerical methods with the accessibility to numerous dispersion modeling systems. Most systems require a precursor execution of a weather model before the input of

emissions; however, it would be best to incorporate the effects of gases and aerosols directly into the weather model to avoid loss of information due to the interrelation of the physical and chemical processes.

The combination of both meteorological and chemical processes is called coupling. In exchange, this produces an accurate form of an air pollution dispersion model where there is less loss of information within the chemical and physical processes that govern the pollutant species. A fully coupled online model that performs such calculations is the Weather Research and Forecasting (WRF) model coupled with Chemistry (WRF/Chem). Together, the WRF/Chem forms an online coupling of meteorological and chemical processes. This model is capable of simulating and predicting pollutant concentrations over an area or domain of study with the integration of various physical, and chemical options [3, 4].

In a study done by Grell et al. [12], the fully coupled online model was used within the WRF framework. The model configurations included the use of the transport scheme (mass and scalar preserving), the grid, and the physics schemes for sub-grid scale transport. This model configuration was used to evaluate and compare the WRF/Chem with the Fifth Generation Penn State/National Center for Atmospheric Research Mesoscale Model with offline chemistry (MM5/chem) along with photochemical data collected during the summer period of 2002. Statistical methods used to evaluate collected data ranged from the Pearson’s correlation factor (r) to the median error or bias. Data collected in all sites showed that the correlation coefficients for O₃ are higher for the WRF/Chem model compared with that of

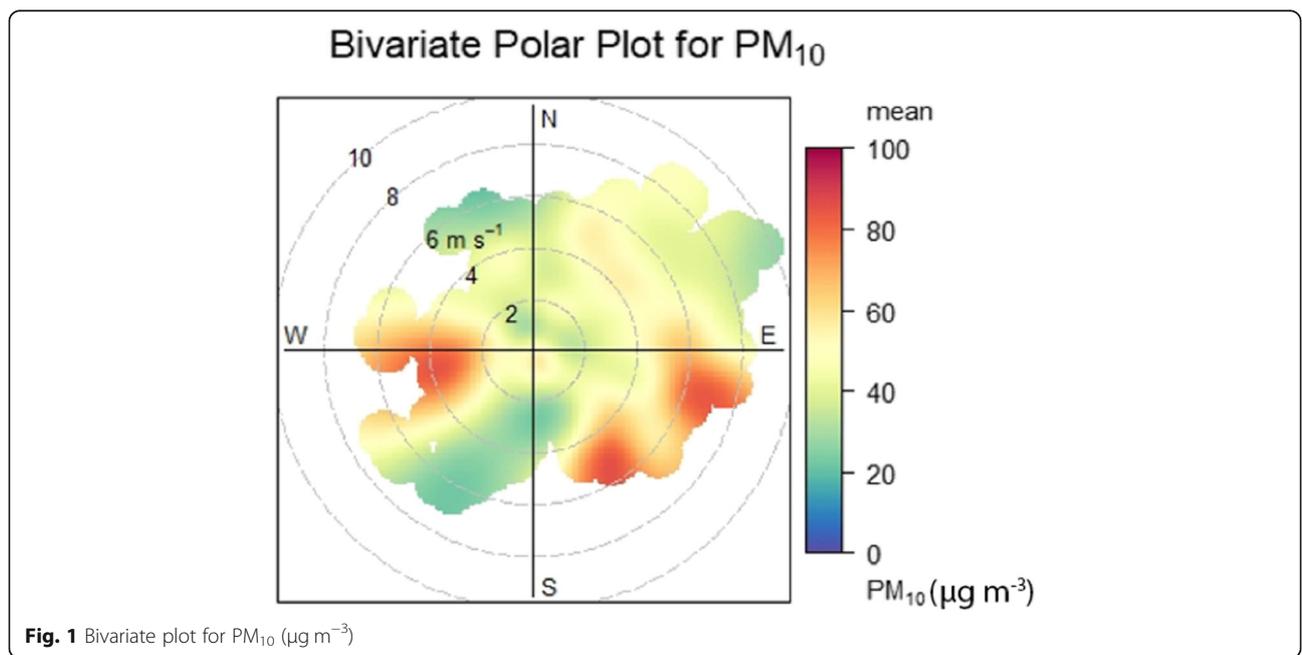
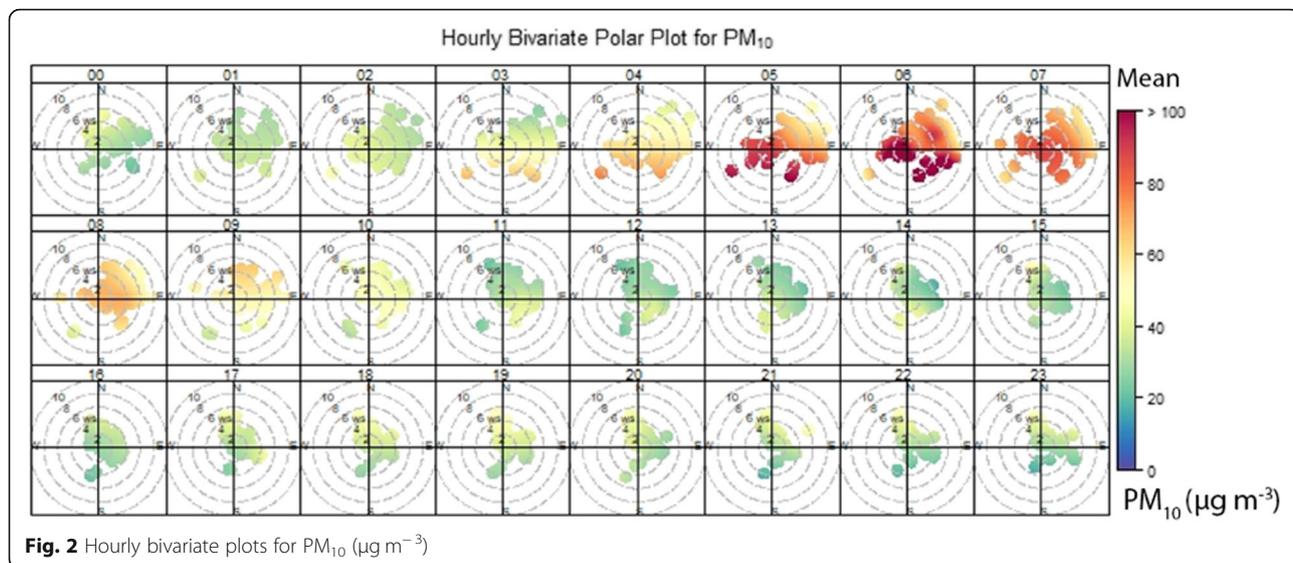


Fig. 1 Bivariate plot for PM₁₀ (µg m⁻³)



the MM5/Chem model. These results show that the WRF/Chem is proven to be better in comparison to the MM5/Chem model. With these methods, the WRF/Chem was seen to have better forecasting skill; however, the biases of ozone precursor forecasts seemed to be higher in comparison to that of the MM5/Chem.

Further studies [13, 14] utilized the WRF/Chem for applications in urban areas. As mentioned in these studies, urbanization greatly changes the underlying surface characteristics of an area; thus, also affecting boundary-layer variables. As a result, these would affect the dispersion and transport of pollutant fluxes in the vertical and horizontal directions.

Zhong et al. [14] discussed the individual and combined impacts of an urban heat island (UHI) case and an increase in anthropogenic emissions. The relationship between increased anthropogenic emissions and a change in land-use was found to affect 2-m temperature differently. Further investigation showed that the UHI was found to directly impact 2-m temperature with a greater degree than the case with an increase in anthropogenic emissions. Increased amounts of aerosols were seen to affect 2-m temperature indirectly through the radiative effects, which allowed less incoming solar radiation to reach the ground, thus, affecting 2-m temperature. The combined effect of both cases proved

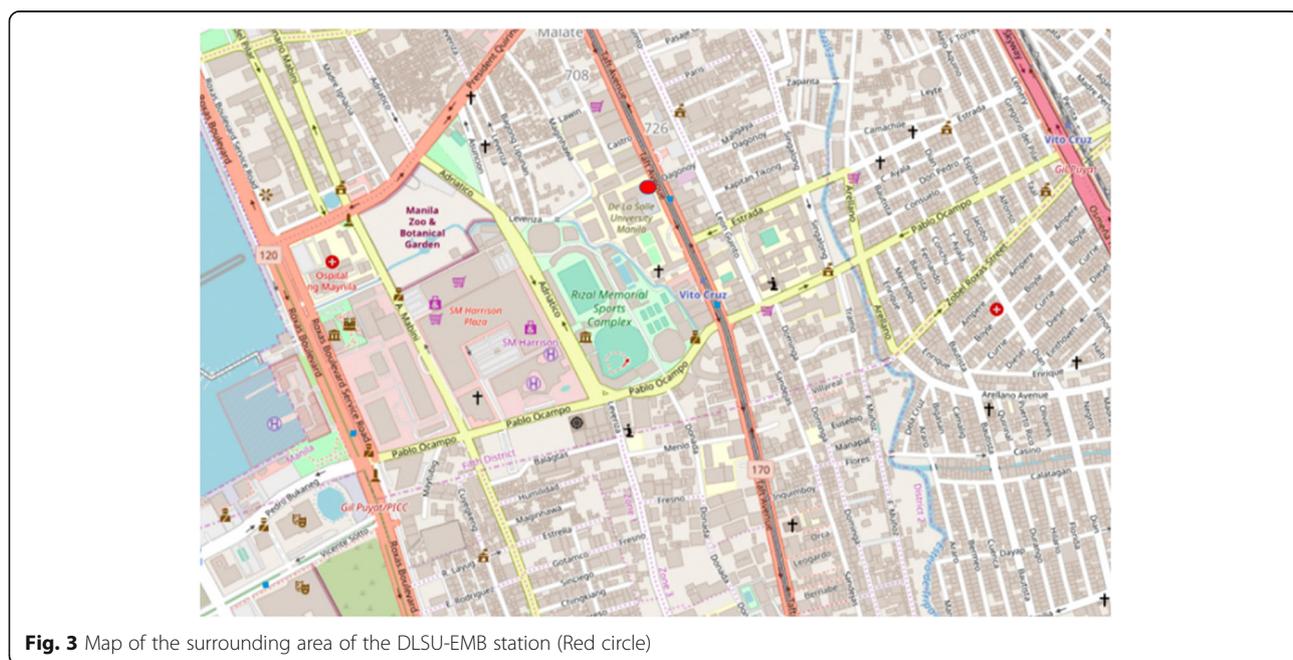


Table 3 Domain configuration

Domains	Domain 1	Domain 2	Domain 3
Horizontal resolution	36 km	12 km	4 km
Vertical resolution	51 levels	51 levels	51 levels
Grid points	45	67	52
Domains of integration	Latitude 7.53° N- 21.41° N	Latitude 11.05° N - 18.06° N	Latitude 13.62° N - 15.42° N
	Longitude 114.03° E-127.97° E	Longitude 117.43° E - 124.5675° E	Longitude 120.02° E - 121.85° E

that the local effects of a UHI yielded greater changes in 2-m temperature; however, indirect aerosol radiative effects tend to affect a greater area downwind the UHI.

In a study by Oliveros et al. [13], urbanization was quantified through changes in surface heat flux (SHF). The increase in SHF within Metro Manila was mainly attributed to the increase in heat capacity and decrease in surface albedo brought about by urbanization. The effects of the UHI only affected Metro Manila and areas within 25 km of the urban mega-city.

The WRF/Chem model was used in an unpublished study over Metro Manila by Ramos in 2015. The climatic effects of aerosols over the Metro Manila airshed were investigated in this study. The Morrison 2-moment microphysics scheme was used in the simulations. The study peered into the climatic effects of aerosols and these effects were quantified with rainfall and temperature from the model runs. Results showed that the direct effect of aerosols decreases the incoming shortwave radiation by 4.6 and 1.42 W m⁻² for December to February and June to August, respectively. In terms of the indirect effect of aerosols, a decrease in incoming solar radiation was noted for all seasons. These affected temperature and rainfall where an increase of 0.75, and 0.34 mm h⁻¹ were found for the cool-dry months and the warm-dry seasons respectively.

Minimal studies have been done on pollutant concentration simulations using the WRF/Chem in the Philippines, and it is important that such a model be investigated for its applicability over highly urbanized areas such as Metro Manila [15]. The city is located in a developing country and is an interesting area to investigate pollutant dispersion due to the high levels of emissions. Also, Metro Manila is located within the island of Luzon and poses an interesting area to study the effects of the local circulation on pollutant dispersion. In addition, the city has a high population density count; where a great percentage of the population is being exposed to harmful pollutants [16]. Thus, the study on air pollution dispersion modeling is timely and relevant to the Philippine society.

This study used the fully coupled “online” WRF/Chem model over the area of Metro Manila to simulate PM_{2.5}, PM₁₀, O₃, SO₃, and NO₂ mass concentrations over the

DLSU-EMB station and compared these with ground measurements. An optimization of the PBL settings available to the WRF/Chem was performed to enable its usage later in the reanalysis of meteorological data. The emissions database used was taken from the Emissions Database for Global Atmospheric Research (EDGAR) and REanalysis of Tropospheric chemical composition (RETRO) databases. The specific objectives were to (1) determine the appropriate PBL scheme that allows for the proper estimation of meteorological variables over the domain of study, (2) determine the performance of the model by establishing the correlation between the predicted and observed values of pollutant concentrations, and (3) investigate consistent differences between observed and modeled values.

Table 4 Emissions (mol km⁻² h⁻¹) by the EDGAR/RETRO emissions inventory

EDGAR-anthropogenic emissions	RETRO-anthropogenic emissions
CH ₄	Acids
CO ₂	Alcohols
CO	Benzene
N ₂ O	C ₂ H ₂
PM _{2.5} fossil	C ₂ H ₄
CO	C ₂ H ₆
NMVOC	C ₂ H ₈
OC	C ₄ H ₁₀
Agriculture emissions	C ₅ H ₁₂
Energy emissions	C ₆ H ₁₄
Residential emissions	Chlorinated hydrocarbons
Transport emissions	CO
Industry emissions	Esters
	Ketenes
	Other VOC
	Toulene
	Trimethylbenzenes
	Xylene
	CO

Table 5 Model configurations used in WRF/Chem PBL sensitivity simulations and the domain specifications

Shortwave radiation	Dudhia scheme
Longwave radiation	Rapid radiative transfer model
PBL process	Yonsei University Scheme/Mellor-Yamada-Janjic Scheme
Surface layer	MM5 surface layer scheme
Land-surface model	NOAH land surface model
Cumulus convection	Grell-Freitas scheme except on the fine nest
Cloud microphysics	WRF single-moment 3-class scheme
Photolysis scheme	Madronich F-TUV photolysis
Gas-phase mechanism	Regional acid deposition model version 2
Cloud chemistry	None (Innermost domain only)
Aerosol module	GoCart plus the Madronich F-TUV photolysis scheme

Methodology

The initial process began with a sensitivity analysis. This was done by varying the configuration of the PBL schemes within the model. With this, information on the key meteorological parameters was compared to determine the proper PBL scheme for the domain. Through the EDGAR and RETRO emissions inventories, the WRF/Chem model was used to simulate the chemical and physical processes within the domain of interest. The simulation area chosen is described in Table 3 and a table enumerating the EDGAR and RETRO emissions is included in Table 4.

The selected area was used to represent a simulation of chemical weather for the year 2013. The chosen year to conduct the simulations was based on the availability of meteorological data for model evaluation. Furthermore, the period included the enhanced “2013 Habagat” event that brought high amounts of rainfall in Metro

Manila. Initial and lateral boundary conditions required for the simulation were processed by the model’s pre-processing system using the National Centers for Environmental Prediction Final (NCEP FNL) Operational Global analysis data and the RETRO and EDGAR emissions inventories.

The modeling system

The WRF model is an atmospheric simulation system that has a broad range of applications. Several physical parameterizations and chemical parameterizations are present for the reconstruction of PBL processes. Each physical scheme differs in the manner flux is depicted within its framework [3]. The WRF/Chem model was released as part of the WRF modeling package. The model is highly dependent upon the WRF in making its results in air quality forecasting, aerosol simulation, and chemistry

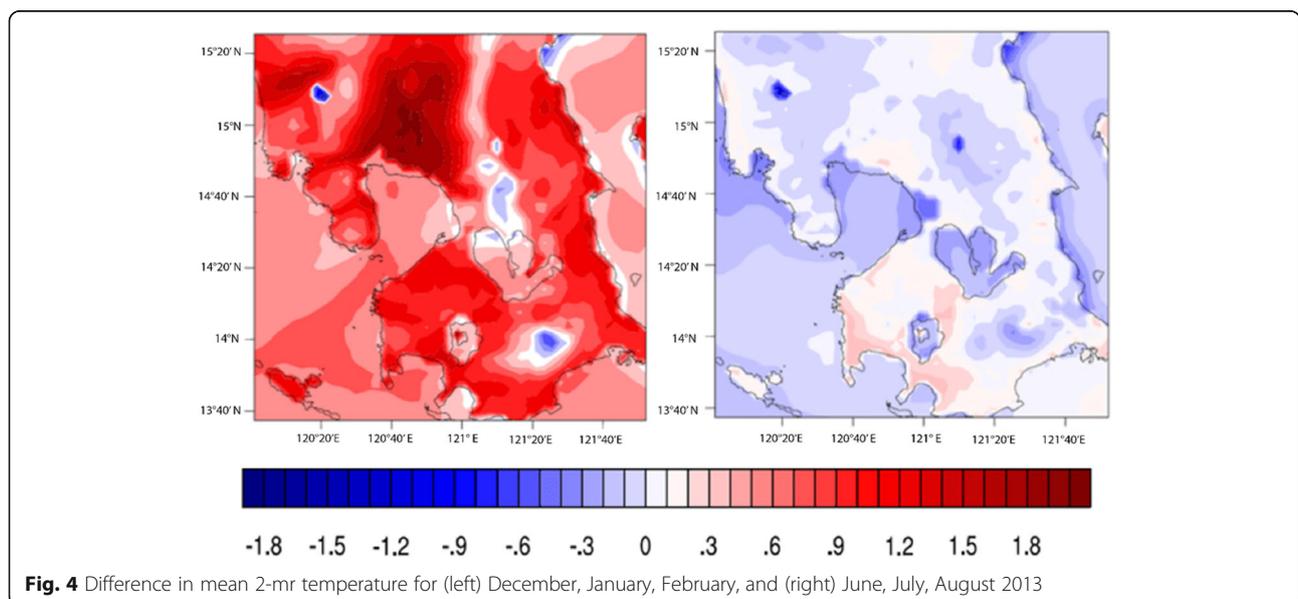


Fig. 4 Difference in mean 2-mr temperature for (left) December, January, February, and (right) June, July, August 2013

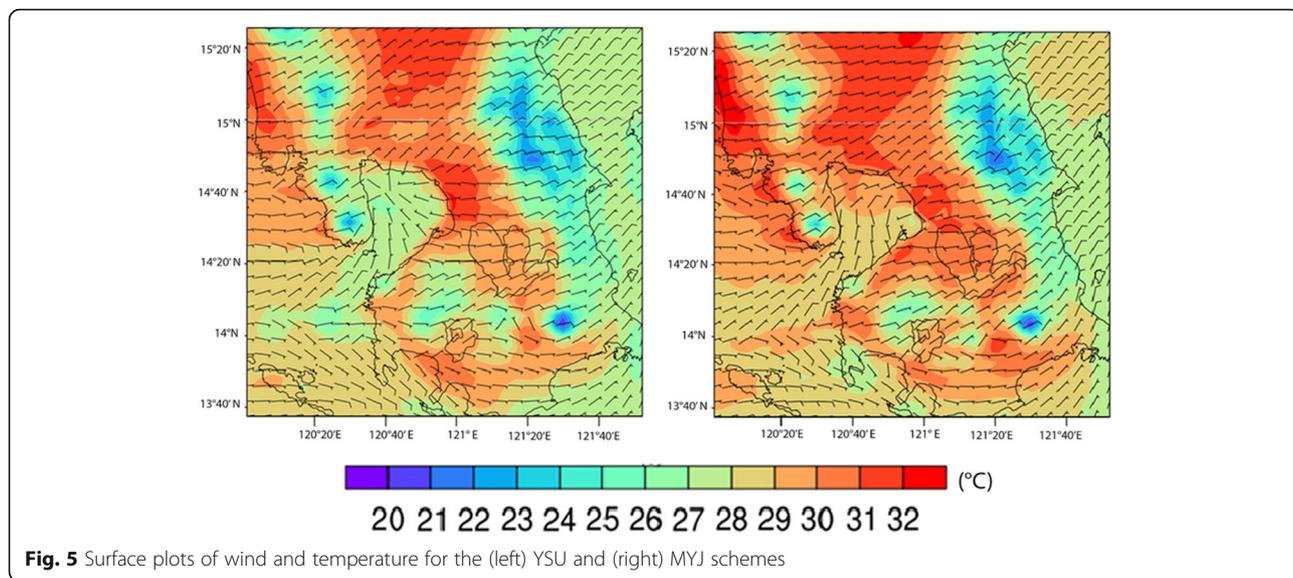


Fig. 5 Surface plots of wind and temperature for the (left) YSU and (right) MYJ schemes

consistent with the meteorological components. It is a numerical weather prediction and atmospheric simulation software that is predominantly used for research and operational applications. It is similar in structure to the WRF model; however, additional gridded inputs on emissions

data are to be utilized. The coupled system enables the WRF model to simulate the transport of various aerosols, and gaseous pollutants [12].

The model has four major programs that include the WRF pre-processing system, the WRF-DA, the Advanced

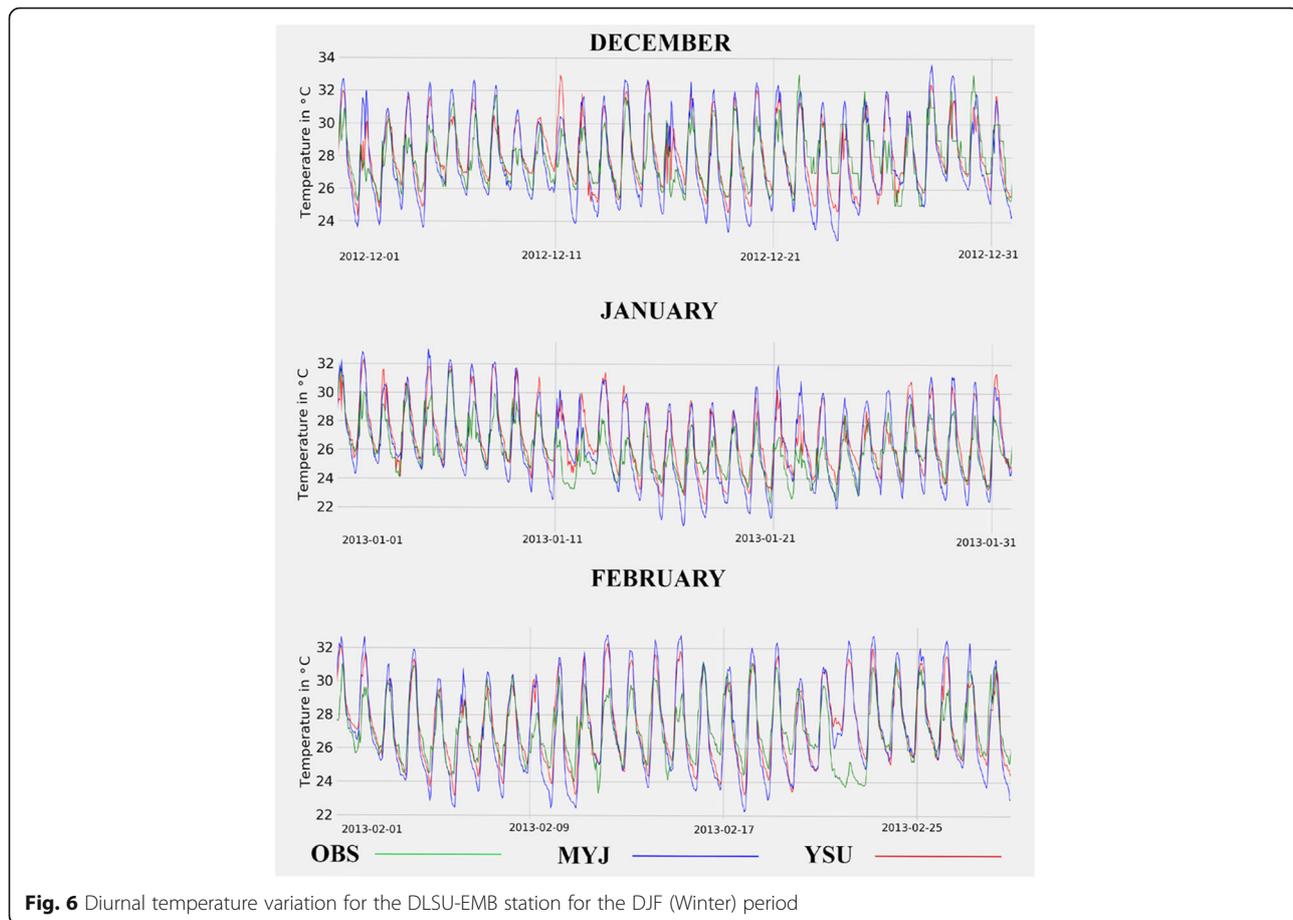


Fig. 6 Diurnal temperature variation for the DLSU-EMB station for the DJF (Winter) period

Research WRF Solver including Chemistry, and the Post-processing & Visualization Tools [3].

Description of input data

Meteorological data

Meteorological data were derived from the NCEP FNL Operational Global Analysis Data in a one by one-degree grid prepared with a temporal resolution of 6 h.

Geographical input data

The dataset used for land-use and topography was taken from the United States Geological Survey (USGS) database. The geographical dataset included the physical properties of the area such as soil category, land-use category, terrain height, annual mean deep-soil temperature, monthly vegetation fraction, monthly albedo, slope category, albedo, and surface roughness.

Emissions data

Emissions data were taken from the EDGAR and RETRO databases. These additional gridded datasets were used to include sources of anthropogenic and biogenic emissions within the modeled area. The EDGAR database has a resolution of 10 × 10 km. It utilizes

nationally reported emissions for CH₄, CO, SO₂, NO_x, NMVOC, NH₃, PM₁₀, PM_{2.5}, BC, and OC. In addition, chemical species not included, or missing from the EDGAR database were taken from the RETRO database. Table 4 lists the available emissions. There were no biogenic emissions included in the simulations.

Observational datasets

Observational datasets were compiled to evaluate the model's capability. The DLSU-EMB air quality monitoring station provided the hourly observed pollutant data for PM_{2.5}, PM₁₀, O₃, SO₃, and NO₂. Unfortunately, data from other EMB sites were not available at that moment due to technicalities. Meteorological data were derived from the DLSU-Davis weather station along with 6-hourly data from the Port Area, Sangley Point, Ninoy Aquino International Airport, and the Science Garden in Quezon City. Both the DLSU-EMB station and the DLSU-Davis weather station are within proximity of each other and will not significantly affect the results for model evaluation.

Setup and design

The sensitivity of the WRF/Chem v.3.6.1 to various PBL schemes was quantified using four basic statistical

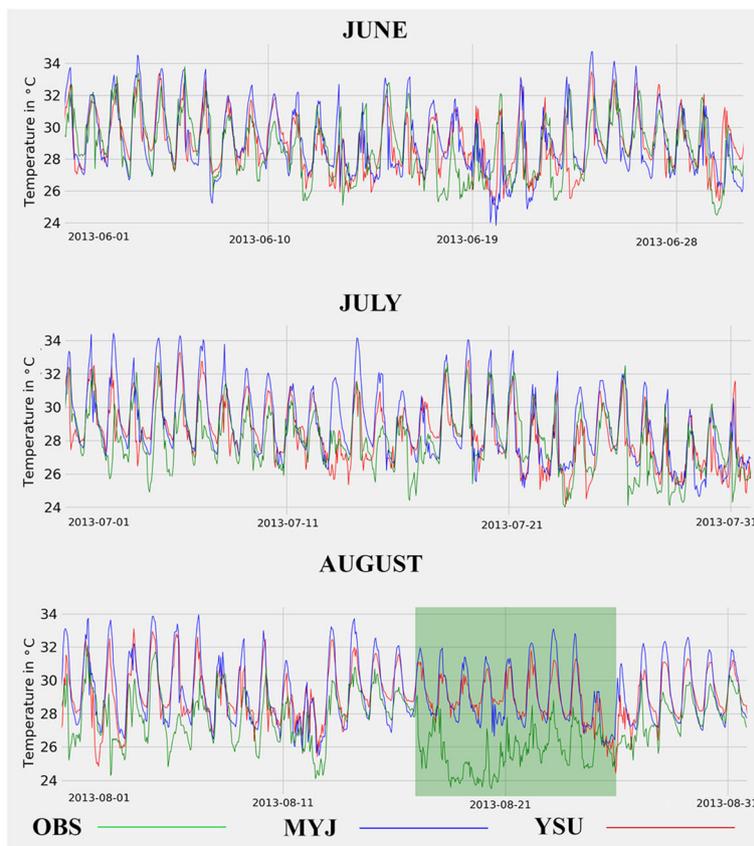


Fig. 7 Diurnal temperature variation for the DLSU-EMB station for the JJA (Summer) period

scores, namely bias, mean absolute error (MAE), root mean squared error (RMSE), and Pearson’s correlation coefficient (Pearson’s r). The sensitivity experiments were done for two seasons: the dry-cool season and the wet-warm season. For both seasons, PBL settings were varied between the Yonsei University (YSU) and Mellor Yamada-Janjic (MYJ) schemes. Results from the sensitivity study were then evaluated to determine the proper PBL configurations to be used in the estimation of pollutant concentration and transport.

Sensitivity experiments and computation of pollutant concentrations

The PBL parameterizations influence the simulations of the wind, turbulence, and other thermodynamic variables in the lower atmosphere. It is in this area where many processes that govern transport and dispersion of pollutants take place [17]. The uncertainty in parameter estimates arises from the different assumptions on the transport of mass, moisture, and energy. Consequently, this may lead to variations in the structure of the PBL along with the different boundary-layer parameters that are estimated by the model [18].

This sensitivity study was done to investigate two PBL schemes and their effect on simulating surface meteorological variables in the WRF/Chem model. The two PBL schemes used in the runs were the YSU and MYJ schemes. According to Yerramilli et al. [17], the YSU is a modification of the Medium Range Forecast scheme. It uses an explicit treatment of entrainment processes. Moreover, it is a non-local first-order scheme where vertical transfers are allowed for all grid levels. The MYJ scheme, on the other hand, has prognostic equations for the turbulent kinetic energy (TKE). It utilizes a 2.5 turbulence closure approximation to determine the eddy transfer coefficients, and it uses local vertical mixing to treat the PBL parameters. Since the YSU Scheme does not take TKE into account, it considers turbulent conditions by assuming an unstable mixed layer; where the potential temperature is constant. Moreover, an explicit term is present to treat the effects of entrainment above the mixed layer [4, 19, 20] This is shown in Eq. (1) below.

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial z} \left[K_c \left(\frac{\partial c}{\partial z} - \gamma_c \right) - \overline{(w'c)'}_h \left(\frac{z}{h} \right)^3 \right] \tag{1}$$

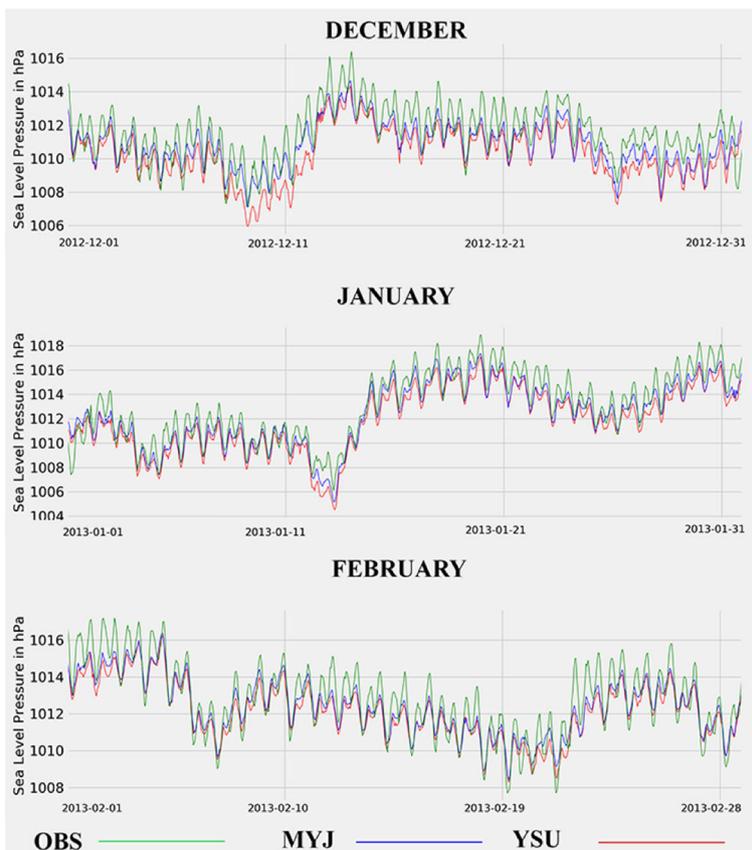


Fig. 8 Diurnal sea level pressure variation for the DLSU-EMB station for the DJF (Winter) period

where, C is the set of the prognostic variables $u, v, \theta,$ and q, u is the horizontal direction of wind, v is the vertical direction of wind, θ is the potential temperature, q is the specific humidity, K_c is the diffusivity coefficient, z is the vertical direction, h is the height of the mixed layer, and γ_c is the correction factor for the local gradient. The correction factor assists in the incorporation of large-scale eddy contributions to the total turbulent flux. The last term contains $\overline{(w'c')}_h$ which is the flux at the inversion layer and is directly responsible for the explicit treatment of entrainment into the mixed layer.

Model configuration and initialization The model was configured using a Lambert conformal conic projection having three domains of horizontal resolutions 36, 12, and 4 km centered at 14.58° N and 121.00° E. The grid sizes in the east-west and north-south directions for each domain are presented in Table 3.

A total of 51 vertical levels were considered in the domains. This vertical resolution was to ensure that the radiation calculation was not offset. Terrain, land-use, and soil data were interpolated to the model grids from the USGS geographical dataset. The model's initial and

lateral boundary conditions were initialized using the NCEP FNL data. Furthermore, anthropogenic emission sources were derived from the EDGAR/RETRO databases. The emissions were related to the speciation of the desired chemical mechanism. Finally, these were disaggregated to the model grids.

There are several physical parameterizations available to represent PBL turbulence, land surface fluxes, microphysics, atmospheric radiation, and cumulus convection. The physical parameterizations used in the model were the Grell-Freitas scheme for cumulus convection, the WRF Single-moment 3-Class scheme for the microphysics, the Noah scheme for land surface processes, and the Rapid Radiative Transfer Model and Dudhia shortwave radiation scheme for longwave and shortwave radiation, respectively. As shown in Table 5, no cumulus convection scheme was used for the innermost domain. Aside from this, chemical parameterizations were also included in the configuration. The chemical parameterizations considered in the model were the Regional Acid Deposition Model Version 2 Chemistry with GoCart Aerosol along with the Madronich F-TUV photolysis scheme.

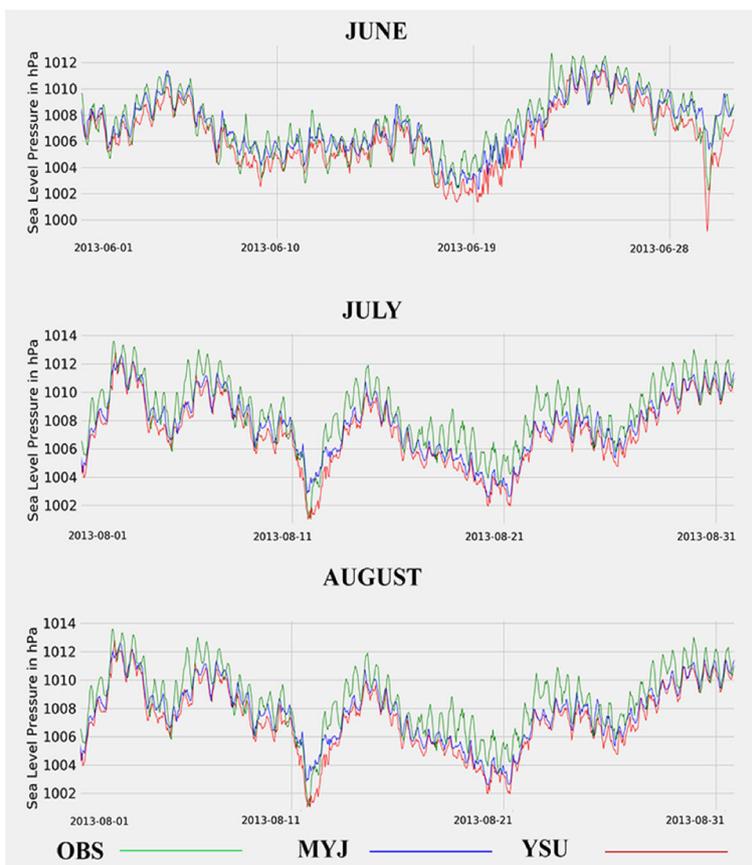


Fig. 9 Diurnal sea level pressure variation for the DLSU-EMB station for the JJA (Summer) period

Results and discussion

Various physical processes are interactive and play important roles in the simulation of the lower atmosphere. Given the many processes involved, the development of the PBL is paramount to simulating chemical weather. Thus, it is important to evaluate the model's performance first to be able to identify a reasonable PBL setting that can provide accurate simulations of meteorological variables.

Mean difference in spatial 2-m temperature

Terrain data was obtained through the USGS database. Terrain height is an important aspect of numerical models as it defines the height of the atmosphere above the ground, therefore affecting the amount of radiation received at the ground's surface [4].

Overall, temperatures simulated by all experiments showed that the YSU scheme produced higher temperatures compared to the MYJ scheme. Figure 4 depicts the difference in average spatial 2-m temperature. The figure depicts warmer PBLs simulated by the YSU scheme. Moreover, Fig. 5 shows two surface plots taken from the output at a random day within the North-East monsoon period. A noticeable sea breeze forms during this period

with the YSU scheme as a result of the temperature gradient between Metro Manila and Manila Bay. However, this is not the case with the MYJ scheme. This is consistent with the results of Hu et al. [21]; where the MYJ scheme produces a cold bias due to the lack of entrainment of warmer air. Garcia-Diez et al. [22] also show similar results for the YSU scheme. Cold biases were seen using the MYJ scheme during the daytime. The YSU scheme showed better forecast skill compared to the MYJ scheme, and it simulates warmer PBLs that are attributed to the scheme's ability for stronger entrainment processes. The YSU scheme is a non-local scheme. It contains counter-gradient terms that allow for the rise of thermals along with an explicit treatment for entrainment allowing the PBL to grow in depth and have higher simulated temperatures [17, 23, 24].

Temporal evolution of surface meteorological variables

Output taken from model runs was plotted against time along with observed data from the DLSU-EMB site.

Temperature

Figures 6 and 7 depict the variations in diagnosed 2-m temperature. The simulated temperature agrees well

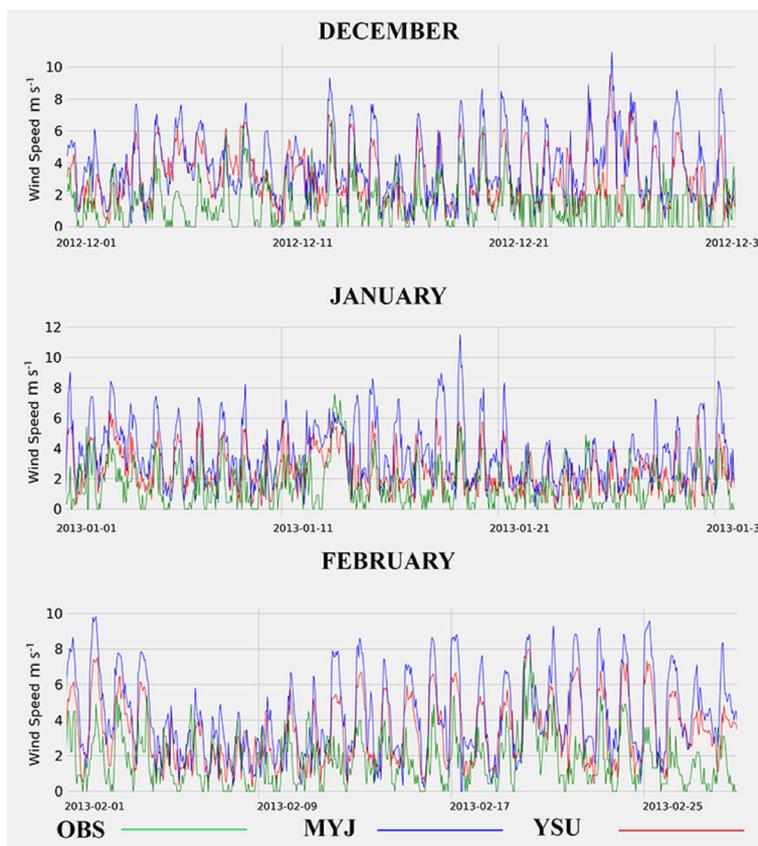


Fig. 10 Diurnal wind speed variation for the DLSU-EMB station for the DJF (Winter) period

with the diurnal variation of the observed counterpart. However, during the surge of the southwest monsoon from August 18 to August 26, 2013, the model was not able to catch the drop in observed temperatures as seen in Fig. 7 highlighted in green. Other sites also behaved the same way and did not capture the decrease in temperatures for the same period. Another feature that the model simulations show is the diurnal pattern of temperature simulated in Sangley Point. Sangley Point is not resolved in the model as a land area. Due to this, diurnal variations captured are similar to those above water.

During the surge of the monsoon, the flux of incoming shortwave radiation was inspected at the surface for the DLSU-EMB station. Downward flux of shortwave radiation mostly comes from solar energy. Accordingly, the output, shows that incoming shortwave radiation still reaches its maximum during these days. The values of the downward flux of shortwave radiation do not show any dramatic change during the “2013 Habagat” event. This then shows that the model might have under-predicted cloud cover over the station; hence, allowing solar heating of the surface.

Sea level pressure

Both YSU and MYJ schemes show good diurnal patterns for sea level pressure as shown in Figs. 8 and 9. For all other stations, sea level pressure is simulated well. Although the model captures the diurnal wavy pattern of sea level pressure, it does not capture sudden pressure changes.

Wind speed

The model over-predicts wind speed for both YSU and MYJ in all stations. The wind speed as simulated by the model does not have a significant correlation with the observed data based on the magnitudes of wind speed observed. Figures 10 and 11 show a plot of wind speed at the DLSU-EMB site as compared to the YSU and MYJ runs. As observed, most simulations overestimated values for the wind speed. The reason is due to unresolved urban canopy or topography.

Statistical analysis for sensitivity runs

Table 6 presents the summary of all metrics averaged over all sites. All periods with no data from the DLSU-EMB site were not considered in the calculations of the

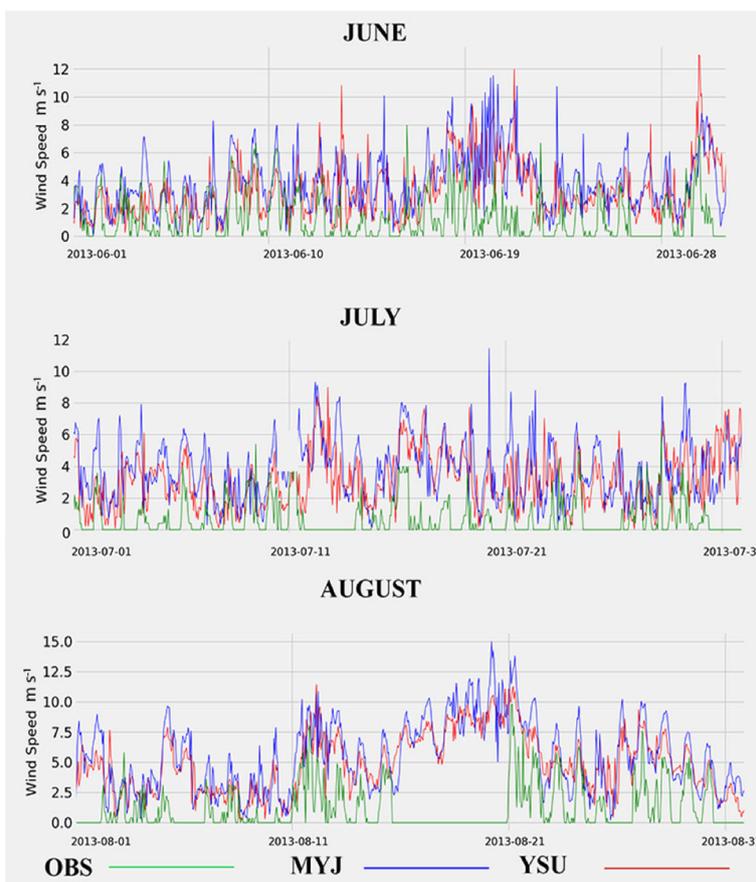


Fig. 11 Diurnal wind speed variation for the DLSU-EMB station for the JJA (Summer) period

Table 6 Metrics for meteorological parameters for the DLSU-EMB site

	YSU: 6Mos Ave.	MYJ: 6Mos Ave.	YSU_DJF Ave.	MYJ_DJF Ave.	YSU_JJA Ave.	MYJ_JJA Ave.
Surface temperature						
BIAS	0	0	0	0	1	1
MAE	2	2	2	1	2	2
RMSE	2	2	2	2	2	2
Pearson-r	0.70	0.72	0.77	0.79	0.64	0.69
Sea level pressure						
BIAS	0	0	0	0	0	1
MAE	1	1	1	2	1	1
RMSE	1	2	2	2	1	1
Pearson-r	0.92	0.92	0.86	0.87	0.86	0.87
Wind speed						
BIAS	2	2	2	1	3	2
MAE	3	2	2	2	3	3
RMSE	3	3	3	3	4	3
Pearson-r	0.32	0.26	0.35	0.31	0.35	0.28

$$BIAS: \frac{1}{N} \sum_{i=1}^n (f_i - o_i) = \bar{f} - \bar{o}$$

$$Mean\ Absolute\ Error\ (MAE): \frac{1}{N} \sum_{i=1}^n |f_i - o_i|$$

$$Root\ mean\ squared\ error\ (RMSE): \sqrt{\frac{\sum_{i=1}^n (f_i - o_i)^2}{N}}$$

where f_i is the modelled value, O_i is the observed values, and N is the total number of observations

metrics. The qualitative descriptions of both schemes of interest show that the YSU predicts warmer temperatures across the innermost domain for the dry season (DJF). Simulated diurnal temperatures followed the temporal variations in all stations except at Sangley point – an unresolved area as seen by the model. With respect to sea level pressure, all simulations show good performances on the daily variations; however, abrupt changes in pressure are not simulated in the model. On the other hand, wind speed was overestimated in all locations. This is possibly due to the inability of the model to capture the exact urban effect on wind speed.

The mean bias over all sites for the whole 6-month period presented a warm bias for both the YSU (YSU = 0.35) and the MYJ (MYJ = 0.32) Schemes. This is similar to the results found in studies that also used the WRF model [25, 26]. The evaluation of parameter estimations showed slight overpredictions with a cold bias seen for Sangley Point. The calculated average MAE scores for the entire 6-month period over all stations returned a value of 1.57 and 1.71, for the YSU and MYJ scheme, respectively. The same method was done for the RMSE. The YSU scheme provided a better score at 1.92 as compared to the MYJ scheme with a score of 2.06. The previous scores provided for the RMSE included the period of overestimation from June to August; consequently, the average RMSE scores over all sites improved upon

its removal from the calculations (YSU = 1.68 and MYJ = 1.95). Upon inspecting the correlation of modeled values with observed values, Sangley point resulted in the lowest correlation for both MYJ and YSU. With respect to the average score of correlation over all sites, the YSU Scheme returned the highest correlation.

Sea level pressure model outputs returned high correlations with observed data as seen in Table 6. The average correlation coefficients for the DLSU-EMB site showed that both YSU and MYJ performed well and that they are both in sync with the variations in daily pressure readings (MYJ = 0.92 and YSU = 0.92).

With regard to wind speed observations, the model showed over-predictions at all sites, and returned low correlations for all sites. This implies that the variations of wind speed over time for all stations are out of phase with observations. These statistical scores were calculated using the wind speed only. All stations used for evaluation are located within a highly urbanized area in Metro Manila. The flow of wind over such an area allows for turbulent conditions that are not resolved with the model’s grid spacing. This caused the deviation of simulated wind speed from the modeled values. In order to include the effects of urbanization on wind speed, data on sub-grid scale information such as detailed land use maps and detailed surface characteristics would need to be ingested in the model. In relation to this, the

researchers may apply an urban canopy layer over the model domains to capture the effects of the highly urbanized area. Many studies show that by using an urban canopy model, simulated urban canopy wind speed may be calculated based on the underlying characteristics of land as set by the urban land use dataset. A study done by Sharma et al. [27] showed that simulated surface wind speed and temperature were greatly influenced by the selection of urban parameters and land use in the configuration of the model. Both wind speed and near-surface temperature simulations improved.

Analysis of pollutant concentrations for the year 2013

The average bias for all pollutants presented in Table 7. SO₂ and O₃ returned the lowest values with an average value of 3.2 and -10.0 μg m⁻³, respectively. Upon inspecting the variations of these two pollutants, it seems that SO₂ is the most properly simulated with regard to its cyclic pattern. For O₃, the cyclic pattern is also seen, but an underestimation of NO₂ could be the source of error in the simulated O₃ fields. In terms of the correlation coefficient, PM_{2.5}, PM₁₀, and SO₂ show the highest correlation. Other than that, the rest of the pollutant species showed low values of correlation for the year 2013.

Table 8 shows that nighttime concentrations are overestimated compared to the daytime values. A similar trend between nighttime and daytime values was noticed by Cheng et al. [28] in a study focusing on the importance of the PBL in simulating air quality. Higher nighttime concentrations are noticed in the model due to the underestimation of the nocturnal boundary layer. Nocturnal boundary layers usually trap pollutants. Upon comparing the mean bias scores of this division to that of the entire period, daytime bias is significantly reduced. This points out the overestimation of nighttime values. Furthermore, overestimation of nighttime values may have arisen from the continuous and time-independent emissions data ingested into the system.

Table 9 presents the results from the t-test conducted. It presents that the pairs of mean values for the observed are significantly different from zero. Since this is also seen in the modeled data, this suggests that the model

Table 7 Metrics for the year 2013 pollutant concentrations for the DLSU-EMB station

	PM ₁₀	PM _{2.5}	O ₃	SO ₂	NO ₂
Mean obs	40	24	81	14	51
Mean mod	121	42	72	17	5
Bias	83	19	-10	3	-46
MAE	90	26	34	10	46
RMSE	125	38	43	14	59
Pearson-r	0.21	0.29	0.09	0.32	0.03

Table 8 Metrics for day/night pollutant concentrations for 2013 for the DLSU-EMB station

Day	PM ₁₀	PM _{2.5}	O ₃	SO ₂	NO ₂
Mean obs	42	23	99	13	54
Mean mod	85	31	74	11	4
Bias	45	8	-26	-2	-50
MAE	56	18	42	7	50
RMSE	85	29	53	10	61
Pearson-r	0.36	0.39	-0.17	0.4	0.24
Night	PM ₁₀	PM _{2.5}	O ₃	SO ₂	NO ₂
Mean obs	39	24	63	16	49
Mean mod	157	53	70	23	7
Bias	122	30	6	8	-42
MAE	125	35	25	13	42
RMSE	156	45	30	18	57
Pearson-r	0.17	0.25	-0.2	0.27	-0.06

simulates significantly different concentrations when one compares daytime and nighttime values.

The correlation factor accounts for the oscillations of the values. However, it is not yet clear whether the model captures certain oscillations that the observed counterparts have for the other species of pollutants.

To further inspect the model's performance in simulating the observed oscillations, a spectral analysis was done to look at the frequency domain of the time series and look for signals that happen at the same frequency as that of the observed. These are depicted in Fig. 12. This test was done only to see whether the modeled data also peaks at the same frequencies.

Since the values of the modeled data are greatly overestimated, the spectrum values for the modeled and observed periodograms may differ greatly. Also, since the periodogram contains numerous values of the spectra, a

Table 9 T-test for day/night concentrations for the DLSU-EMB Station

	Day/Night Modeled				
	PM ₁₀	PM _{2.5}	O ₃	SO ₂	NO ₂
t	-37.90	-37.03	12.23	-58.20	-44.00
p-value	2.2E-16	2.2E-16	2.2E-16	2.2E-16	2.2E-16
Mean Day	84	31	74	11	4
Mean Night	157	53	70	23	7
	Day/Night Observed				
	PM ₁₀	PM _{2.5}	O ₃	SO ₂	NO ₂
t	3.77	4.89	44.94	-8.79	5.86
p-value	1.7E-04	3.7E-04	2.2E-16	2.2E-16	4.7E-09
Mean Day	42	27	98	123	54
Mean Night	39	24	63	15	49

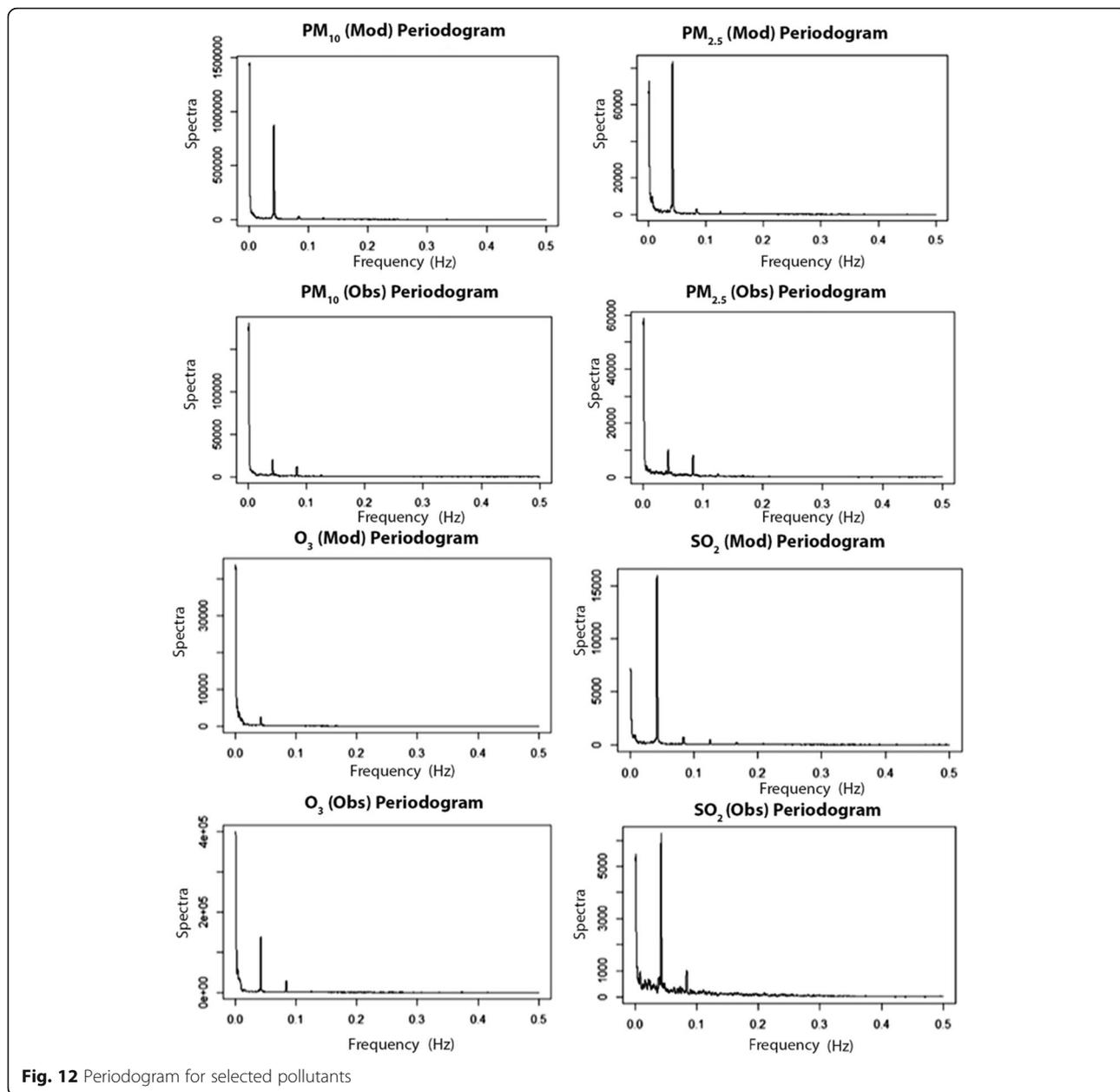


Fig. 12 Periodogram for selected pollutants

smoothing kernel was applied to the raw periodogram to give importance to the most dominant peaks. The frequencies of interest would be 0.04, 0.08, and 0.125 Hz which correspond to the daily, 12-hourly, and bi-annual cycles, respectively.

With respect to $PM_{2.5}$, the model showed a distinct power spectrum at the frequency of 0.04 Hz. Although spectra magnitudes differ, it is apparent that the model captures the same daily variation of $PM_{2.5}$ similar to the observed counterpart. Another peak was also seen in the observed time series and was traced to a peak at 0.08 Hz. This peak was seen in the $PM_{2.5}$ modeled periodogram but is less pronounced as compared to other peaks in

the spectrum. For both the modeled and the observed time series, a small peak was seen at frequency 0.125 Hz. Both the modeled and observed time series presented the yearly cycle observed at the lowest frequency. In the time series for both the modeled and the observed SO_2 values, two distinct peaks can be noticed. The peak at the lowest frequency represents the yearly cycle, while the second peak describes the daily cycle at a frequency of 0.04 Hz. Also, both contain peaks at the frequencies of 0.08 and 0.125 Hz but are less pronounced for modeled data. On the other hand, the only distinct peak for the O_3 time series is the yearly cycle, also seen for the observed. Observed O_3 values show a distinct peak at

the frequency of 0.04 and 0.08 Hz, which the modeled values do not capture. Lastly, for NO₂, it was seen that the model captures a repetitive cycle at the 0.04 Hz frequency. The observed time series also captures this peak, but it is less pronounced.

Conclusions

This work used the WRF/Chem Version 3.6.1. The YSU simulates the closest approximation to the observed atmosphere. An overview of selected time series showed that the model simulates temperature properly but is not able to capture extreme events such as the “Habagat” event during August 2013. Aside from this, the model’s horizontal grid spacing could not resolve the area of Sangley Point; hence temperature was not simulated properly in this area. The pollutant time series suggests that the model overestimates concentration values for PM₁₀, PM_{2.5}, and SO₂ while underestimating NO₂ and O₃ values. Although the model has some difficulties in simulating the proper concentration values, the model does capture a significant 24-h cycle supported by the inspection of the time series’ spectrum in the frequency domain. Furthermore, through a student’s t-test, the model also captures a significant difference in daytime and nighttime concentrations.

Further investigations may be done with the use of a high-resolution emissions inventory with temporal dependence. Aside from this, further simulations involving land use data that includes the urban canopy are encouraged to further characterize the surface characteristics and would probably enhance the simulation of 2-m temperature and wind speed.

Finally, more areas of comparison for modeled pollutant data would be of great assistance as pollutants do not only vary temporally but also spatially and it would validate the modeled data to a better degree.

Acknowledgements

This is to acknowledge the Commission on Higher Education of the Republic of the Philippines for supporting the participants of the 2018 International Conference on Sustainable Environmental Technologies. The support of the Physics Department of De La Salle University through the Environment and Remote Sensing research group is also highly appreciated.

Authors’ contributions

All authors have read and approved the final manuscript. GB provided his knowledge and expertise in the field of numerical weather modeling. EV and MCG provided observational data from the DLSU-EMB site for use in the model evaluation. The corresponding author was responsible for creating the methodology, proceeding with the numerical simulations, and the analysis of the results. The manuscript was written by the corresponding author and EV.

Funding

Not applicable.

Availability of data and materials

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon request.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Physics Department, De La Salle University, 1004 Manila, Philippines.

²Institute of Environmental Science and Meteorology, University of the Philippines, 1008 Quezon City, Philippines.

Received: 27 April 2019 Accepted: 31 October 2019

Published online: 17 December 2019

References

- Andrews DG. An introduction to atmospheric physics. 2nd ed. Cambridge: Cambridge University Press; 2010.
- Bauer P, Thorpe A, Brunet G. The quiet revolution of numerical weather prediction. *Nature*. 2015;525:47–55.
- Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Duda MG, et al. A description of the advanced research WRF version 3. Boulder: National Center for Atmospheric Research; 2008.
- Stensrud DJ. Parameterization schemes: keys to understanding numerical weather prediction models. 1st ed. Cambridge: Cambridge University Press; 2007.
- EMB. National Air Quality Status Report (2008-2015). Quezon City: Environmental Management Bureau; 2016.
- EEA. EMEP/EEA air pollutant emission inventory guidebook 2016. Copenhagen: European Environment Agency; 2016.
- Biona MJB, Montecastro D, Bedano JA, Belleza E, Dollete UG, Asuncion R, et al. Emission inventory of major air pollutants in Iloilo City. Bangkok: ASEAN-GIZ Project Office; 2015.
- Loomis D, Grosse Y, Lauby-Secretan B, El Ghissassi F, Bouvard V, Benbrahim-Tallaa L, et al. The carcinogenicity of outdoor air pollution. *Lancet Oncol*. 2013;14:1262–3.
- Gogna P, Narain TA, O’Sullivan DE, Villeneuve PJ, Demers PA, Hystad P, et al. Estimates of the current and future burden of lung cancer attributable to PM_{2.5} in Canada. *Prev Med*. 2019;122:91–9.
- Kecorius S, Madueno L, Vallar E, Alas H, Betito G, Birmili W, et al. Aerosol particle mixing state, refractory particle number size distributions and emission factors in a polluted urban environment: case study of metro Manila, Philippines. *Atmos Environ*. 2017;170:169–83.
- Carlsaw DC, Ropkins K. *openair* - An R package for air quality data analysis. *Environ Modell Softw*. 2012;27:52–61.
- Grell GA, Peckham SE, Schmitz R, McKeen SA, Frost G, Skamarock WC, et al. Fully coupled “online” chemistry within the WRF model. *Atmos Environ*. 2005;39:6957–75.
- Oliveros JM, Vallar EA, Galvez MCD. Investigating the effect of urbanization on weather using the weather research and forecasting (WRF) model: a case of metro Manila, Philippines. *Environments*. 2019;6:10.
- Zhong S, Qian Y, Zhao C, Leung R, Wang HL, Yang B, et al. Urbanization-induced urban heat island and aerosol effects on climate extremes in the Yangtze River Delta region of China. *Atmos Chem Phys*. 2017;17:5439–57.
- Boori MS, Choudhary K, Kupriyanov A, Kovelskiy V. Satellite data for Singapore, Manila and Kuala Lumpur city growth analysis. *Data Brief*. 2016;7: 1576–83.
- PSA. Population of the national capital region (based on the 2015 census of population). Quezon City: Philippine Statistics Authority; 2016.
- Yerramilli A, Challa VS, Dodla VBR, Myles L, Pendergrass WR, Vogel CA, et al. Simulation of surface ozone pollution in the Central Gulf Coast region during summer synoptic condition using WRF/Chem air quality model. *Atmos Pollut Res*. 2012;3:55–71.
- Hariprasad KBRR, Srinivas CV, Singh AB, Rao SVB, Baskaran R, Venkatraman B. Numerical simulation and intercomparison of boundary layer structure with different PBL schemes in WRF using experimental observations at a tropical site. *Atmos Res*. 2014;145–6:27–44.
- Hong SY, Noh Y, Dudhia J. A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon Weather Rev*. 2006;134:2318–41.
- Tyagi B, Magliulo V, Finardi S, Gasbarra D, Carlucci P, Toscano P, et al. Performance analysis of planetary boundary layer parameterization schemes in WRF modeling set up over southern Italy. *Atmosphere-Basel*. 2018;9:272.
- Hu XM, Nielsen-Gammon JW, Zhang FQ. Evaluation of three planetary boundary layer schemes in the WRF model. *J Appl Meteorol Clim*. 2010;49: 1831–44.

22. Garcia-Diez M, Fernandez J, Fita L, Yague C. Seasonal dependence of WRF model biases and sensitivity to PBL schemes over Europe. *Q J Roy Meteor Soc.* 2013;139:501–14.
23. Xie B, Fung JCH, Chan A, Lau A. Evaluation of nonlocal and local planetary boundary layer schemes in the WRF model. *J Geophys Res Atmos.* 2012;117: D12103.
24. Wallace JM, Hobbs PV. *Atmospheric science: an introductory survey.* 3rd ed. Burlington: Elsevier; 2006.
25. Kuik F, Lauer A, Beukes JP, Van Zyl PG, Josipovic M, Vakkari V, et al. The anthropogenic contribution to atmospheric black carbon concentrations in southern Africa: a WRF-Chem modeling study. *Atmos Chem Phys.* 2015;15: 8809–30.
26. Zhang Y, Sartelet K, Wu SY, Seigneur C. Application of WRF/Chem-MADRID and WRF/Polyphemus in Europe - part 1: model description, evaluation of meteorological predictions, and aerosol-meteorology interactions. *Atmos Chem Phys.* 2013;13:6807–43.
27. Sharma A, Fernando HJS, Hamlet AF, Hellmann JJ, Barlage M, Chen F. Urban meteorological modeling using WRF: a sensitivity study. *Int J Climatol.* 2017; 37:1885–900.
28. Cheng FY, Chin SC, Liu TH. The role of boundary layer schemes in meteorological and air quality simulations of the Taiwan area. *Atmos Environ.* 2012;54:714–27.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

