Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Improving PM$_{2.5}$ predictions during COVID-19 lockdown by assimilating multi-source observations and adjusting emissions

Liuzhu Chen$^a$, Feiyoue Mao$^{a,b}$, Jia Hong$^{a,*}$, Lin Zang$^{c,d}$, Jiangping Chen$^b$, Yi Zhang$^b$, Yuan Gan$^a$, Wei Gong$^d$, Houyou Xu$^e$

$^a$ State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China
$^b$ School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China
$^c$ Chinese Antarctic Center of Surveying and Mapping, Wuhan University, Wuhan, China
$^d$ Electronic Information School, Wuhan University, Wuhan, China
$^e$ Sinosteel Wuhan Safety & Environmental Protection Research Institute Co., Ltd., Wuhan, China

ARTICLE INFO

Keywords:
COVID-19
Data assimilation
PM$_{2.5}$ predictions
3D-var
WRF-Chem

ABSTRACT

The Coronavirus Disease 2019 (COVID-19) outbreak caused a suspension of almost all non-essential human activities, leading to a significant reduction of anthropogenic emissions. However, the emission inventory of the chemistry transport model cannot be updated in time, resulting in large uncertainty in PM$_{2.5}$ predictions. This study adopted a three-dimensional variational approach to assimilate multi-source PM$_{2.5}$ data from satellite and ground observations and jointly adjusted emissions to improve PM$_{2.5}$ predictions of the WRF-Chem model. Experiments were conducted to verify the method over Hubei Province, China, during the COVID-19 epidemic from Jan 21st to Mar 20th, 2020. The results showed that PM$_{2.5}$ predictions were improved at almost all the validation sites, and the benefit of data assimilation (DA) can last for 48 h. However, the benefits of DA diminished quickly with the increase of the forecast time. By adjusting emissions, the PM$_{2.5}$ predictions showed a much slower error accumulation along forecast time. At 48Z, the RMSE still has an 8.85 $\mu$g/m$^3$ (19.49%) improvement, suggesting the effectiveness of emissions adjustment based on the improved initial conditions via DA.

1. Introduction

In the past decades, rapid economic and industrial development resulted in enormous increases in energy consumption and a deteriorated air quality in China. Air pollution, especially particulate pollution, has long been a serious environmental problem in China. At the end of 2019, COVID-19 cases were first reported in Wuhan, Hubei province, and then found worldwide, causing millions of infections and deaths. Studies have shown that SARS-CoV-2 RNA may be present in PM$_{2.5}$ particles, and 10 $\mu$g/m$^3$ increase in PM$_{2.5}$ and PM$_{10}$ was associated with a 2.24% (95% CI: 1.02 to 3.46) and 1.76% (95% CI: 0.89 to 2.63) increase in the daily counts of COVID-19 confirmed cases, respectively (Zhu et al., 2020), indicating that fine-particle pollutants may promote the spreading of the virus (Copat et al., 2020). Therefore, accurate PM$_{2.5}$ predictions during the COVID-19 epidemic helps in decision-making in both environmental pollution control and epidemic prevention.

However, the unrealistic emission inputs of the regional chemistry transport model during the COVID-19 epidemic pose a great challenge to the accurate predictions of PM$_{2.5}$. To prevent the spreading of the coronavirus, the Chinese government took strong measures at the beginning of the epidemic outbreak, including a ban on unnecessary outdoor activities, suspension of public transport, and unimportant production activities. Human and industrial activities in Hubei were reduced to the minimum, with most transportation and almost all outdoor human activities paused, which led to a significant reduction of the national air pollutant emissions (Wang et al., 2020). However, existing emission inventory, such as the Multi-resolution Emission Inventory of China of 2016 (MEIC2016), cannot reflect the actual emissions. Chemistry initial and boundary conditions (IC/BC) from global models are also biased during the COVID-19 due to the reduced emissions (Muhammad et al., 2020), thus challenging the air pollution forecasting (Lee et al., 2008; Wu et al., 2015). If the regional air quality forecast model were directly driven by these unrealistic emissions and IC/BC, the PM$_{2.5}$ predictions of the model would be problematic.
Accurate IC/BC and emissions are the prerequisite for accurate PM$_{2.5}$ prediction. Data sets of energy consumption and emissions are usually available after 1–2 years or more (Zheng et al., 2021), thus cannot meet the requirement of current predictions. Previous studies estimated the anthropogenic emissions during the COVID-19 epidemic based on statistical data (Li et al., 2020; Wang et al., 2020) or data assimilation approaches (Peng et al., 2020; Xing et al., 2020), but the chemistry species mainly limited to NH$_4$, NOX, etc. Peng et al. (2017) developed an ensemble Kalman filter-based joint adjustment method of initial field and source emissions (Peng et al., 2017; Peng et al., 2018), which has been proven effective in improving model prediction. But for the predictions of PM$_{2.5}$ during COVID-19 prevention and control, the current researches are still not targeted enough.

This study proposed to improve PM$_{2.5}$ predictions by assimilating multi-source observations and joint adjusting emissions. A coarse correction experiment of emissions was initially implemented to minimize the influence of the emission reduction on PM$_{2.5}$ predictions, which was caused by the COVID-19 lockdown. A 3DVAR data assimilation approach was adopted to assimilate aerosol observations from satellite and ground sites and adjust the emissions using an emission scaling factor based on the coarsely corrected emissions. Five experiments were conducted, including a raw experiment (RAW), control experiment (noDa), two assimilation experiments that assimilated site observations only (siteDa), and assimilated both site and satellite observations simultaneously (allDa), and an emission adjustment experiment that adjusted the emissions daily based on the allDa experiment (allDaEa). Those four experiments were run with the study areas centered at Hubei province, China, from Jan 21st to Mar 20th, 2020. The proposed approaches will facilitate the PM$_{2.5}$ predictions during the COVID-19 epidemic and other events which affect pollution sources.

2. Data and methodology

2.1. Multi-source PM$_{2.5}$ data from satellite and ground observations

We assimilated PM$_{2.5}$ observations collected by the China Environmental Monitoring Center (CEMC) and ground-level PM$_{2.5}$ estimations derived from Fengyun-4 satellite using a random forest method. The local PM$_{2.5}$ observations from Hubei Environmental Monitoring Station (HEMS) were adopted to validate the PM$_{2.5}$ predictions. These PM$_{2.5}$ observations from CEMC and HEMS are recorded by automated monitoring systems, collecting air pollutants, including PM$_{2.5}$, PM$_{10}$, SO$_2$, NO$_2$, CO, and O$_3$. The distributions of these sites are shown in Fig. 1.

Satellite-based PM$_{2.5}$ estimating has been extensively studied in the past years. Estimating PM$_{2.5}$ concentration from satellite Top-Of-Atmosphere (TOA) reflectance using machine learning methods has been proven effective (Liu et al., 2019; Mao et al., 2020; Shen et al., 2018; Yin et al., 2021). Those machine learning methods directly train the model parameters based on a large quantity of data, then use these parameters to make predictions. Random forest is a decision tree-based ensemble machine learning algorithm. The algorithm creates decision trees for randomly selected samples from the training dataset. Each decision tree makes predictions independently, then combines these predictions to make a final prediction. This study estimates PM$_{2.5}$ concentration from the Fengyun-4A observations using a random forest algorithm (Breiman, 2001).

Fengyun-4 is the second generation of geostationary orbit meteorological satellite in China (Shang et al., 2018). The first satellite of the Fengyun-4 series, FY-4A, is a scientific research test satellite launched on Dec 11th, 2016. The Advanced Geosynchronous Radiation Imager (AGRI) onboard FY-4A has 14 bands. For the visible (VIS), near-infrared (NIR) and remaining infrared bands, the spatial resolutions are 1 km, 2 km and 4 km, respectively. In this study, the L1 dataset with a resolution of 4 km was used. Different bands target different observation objectives. We use the channels with wavelength ranging 0.45–0.49 μm, 0.55–0.75 μm, 2.10–2.35 μm, which have been typically used in aerosol optical properties retrieving (Kaufman et al., 1997; Sayer et al., 2014).

Fig. 1. Study area, model domains, and the distributions of the PM$_{2.5}$ monitoring sites.

2.2. WRF-chem model configuration

The Weather Research and Forecasting model with Chemistry (WRF-Chem) is a fully coupled “online” model that predicts meteorological parameters and atmospheric composition (Grell et al., 2005). This study uses WRF-Chem version 3.9.1 to predict the PM$_{2.5}$ concentration during the COVID-19 outbreak, with two nested domains (Fig. 1). The outer domain (D01) covers eastern China with a 27 km × 27 km resolution and 118 × 94 grids. The inner domain (D02) focuses on Hubei in central China, with a 9 km × 9 km resolution and 93 × 60 grids. There are 35 vertical layers from 50 hPa to the ground surface.

In this study, we utilized the Noah as the land surface model (Fei et al., 2001), YSU as the boundary layer physics scheme (Hong et al., 2005; Hong and Pan, 1996), Morrison 2-moment as the microphysics scheme (Morrison et al., 2009) and Grell-3d as the cumulus scheme (Grell et al., 2005). RRTM (Mlawer et al., 1997) and Dudhia (Dudhia and Jimy, 1989) were chosen as the longwave and shortwave radiation schemes. The gas-phase chemistry model used in this study is the Regional Acid Deposition Model, version 2 (RADM2) (Stockwell et al., 1990), coupled with the GO-CART (the Goddard Chemistry Aerosol Radiation and Transport scheme) aerosol model (Chin et al., 2002).

We used the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) (0.25° × 0.25°) as the initial and boundary
meteorological conditions to drive the WRF-Chem model. The Multi-resolution Emission Inventory of China of 2016 (MEIC, 2016)(Li et al., 2017) was used to provide anthropogenic emissions. MEIC is a high-resolution inventory of China’s anthropogenic air pollutants and carbon dioxide emissions, developed and maintained by Tsinghua University since 2010, providing basic emission data support for scientific research, policy evaluation and air quality management. The Model of Emissions of Gases and Aerosols from Nature inventory (MEGAN) (Guenther et al., 2006) was used to calculate biogenic emissions online.

2.3. GSI 3D-VAR data assimilation system

Data assimilation improves the initial conditions of numerical weather prediction by combining the observation and the background field. Many DA methods have developed in the past few decades and are widely applied in meteorology and chemistry-transport models, including optimal interpolation (Lorenc, 1981), 3-dimensional variational (3DVAR) (Courtier et al., 1998; Lorenc, 1986; Parrish and Derber, 1992) and 4DVAR (Gauthier et al., 2010; Rabier et al., 2000). The ensemble Kalman filtering method (Evensen, 1994) is also commonly used for data assimilation. Gridpoint Statistical Interpolation (GSI) system is a unified data assimilation (DA) system for both global and regional applications. It was initially developed by NCEP Environmental Modeling Center (EMC) as a next-generation analysis system based on the operational Spectral Statistical Interpolation (SSI) analysis system (Kleist et al., 2009). This study uses the GSI 3d-var assimilation system to assimilate PM2.5 observations into the WRF-Chem model. 3D-VAR transforms the assimilation problem into a minimization problem with the following objective function:

\[ J(x) = \frac{1}{2}(x - x_b)\text{T}B^{-1}(x - x_b) + \frac{1}{2}(Hx - y)\text{T}R^{-1}(Hx - y) \]

where \( x \) is the analysis variable, \( x_b \) is the background field of the model, and \( y \) are the observations. \( R \) is the covariance matrix of observation error, which is related to the error of observation data. \( B \) is the background error covariance matrix for the errors of the short-range model forecast valid at analysis time, usually obtained by the National Meteorological Center (NMC) method (Parrish and Derber, 1992; Rabier et al., 2000). The NMC method approximates the background error covariance from the difference between two forecasts of different duration valid at the same time. In this study, differences of the 24h and 12h WRF-Chem predictions are used in the NMC method to estimate the background error covariance.

2.4. Emission adjusting

A model usually uses existing emissions, some of which are outdated and could not reflect the current status, especially during the COVID-19 pandemic. Studies showed that during the COVID-19 pandemic in China, NOx decreased by 36% (Miyazaki et al., 2020), SO2, CO, NMVOCs, and primary PM2.5 decreased by 27%, 28%, 31%, and 24%, respectively (Zheng et al., 2021). Thus, adjusting the emissions is necessary to ensure a reliable model prediction.

We took a two-step method to update the emissions based on the MEIC2016 emission dataset developed by Tsinghua University. First, we updated the anthropogenic emission following Wang et al. (2020) to perform an overall scaling down of the transportation, industry and agriculture emissions by 80% in Hubei province; and by 80%, 20% for the transportation and industry emission for the rest part of China. As a coarse adjustment, such measures can bring down part of the overestimation of the PM2.5 predictions but cannot reflect the daily variation of the real emissions due to the varied human activity in the influence of the COVID-19 epidemics. So, second, we dynamically adjusted the emissions by adopting a daily varied scaling factor \( \lambda_t \). The equation is as follow:

\[ E_t = \lambda_t E^0 \]

where \( E_t \) is the estimated emissions at day \( t \), \( E^0 \) is the original emission. This experiment uses the updated MEIC2016 emission inventory as the original emissions. The scaling factors at day \( t \) were calculated according to

\[ \lambda_t = \frac{C_t^2}{C_t'} \]

where \( C_t^2 \) is the PM2.5 analysis field at day \( t \), generated by assimilating PM2.5 observations; and \( C_t' \) represents the PM2.5 background field at day \( t \), generated by the WRF-Chem model with no data assimilation. Note all these data are averaged throughout the entire domain and applied to each grid. It is a compromise strategy since grid-based scaling is challenging due to complex atmospheric transport processes.

A basic assumption of such a scaling approach is that, after data assimilation, model bias during COVID-19 epidemics mainly results from the uncertainties of emissions. However, multiple factors still bring in uncertainties, such as atmospheric transport processes, boundary layer height, and cloud-precipitation process. Inevitably, the scaling factors can be sudden increase or decrease, which may factually incorrect. Therefore, we further applied a median filter to eliminate such sudden variation of the scaling factors \( \lambda_t \). The window size of the median filter was set as 3 for this study.

2.5. Experimental design

Five experiments were conducted to evaluate the impact of assimilating multi-source observations and emission adjustment on PM2.5 predictions during the COVID-19 outbreak. The experiments are detailed as follows:

1. Raw experiment (RAW): Directly use the original MEIC2016 anthropogenic emission inventory without assimilating any observations.
2. Updated anthropogenic emission (noDa): Reduce transportation, industry and agriculture emissions by 80%, while residential and power emission follows the RAW experiment in Hubei province. For areas out of Hubei province, the transportation and industry emission were scaled down to 80% and 20%, while the rest sectors followed the RAW experiment.
3. Assimilating site observations and running with updated anthropogenic emission (siteDa): The emission scheme follows noDa, but simultaneously performs data assimilation of ground-based PM2.5 observation.
4. Assimilation of multi-source observations and running with the updated anthropogenic emission (allDa): The emission scheme follows noDa experiment, but simultaneously performs data assimilation of PM2.5 from satellite and ground site observations.
5. Assimilation of multi-source observations and adjusting the anthropogenic emission (allDaEa): The emission scheme was dynamically adjusted daily and simultaneously performs data assimilation of PM2.5 from satellite and ground site observations.

All the experiments were conducted from Jan 21st to Mar 20th, 2020, when most of the transportation and outdoor activities in Hubei were stopped. We run a 48 hours forecast for each day, and each forecast is independent. Each run is divided into the preparation phase (0–12h) and the prediction phase (12–48h). The preparation phase is for the spinning up of the model. For siteDa and allDa, the PM2.5 observations were assimilated at 03:00, 06:00 and 09:00UTC (11:00, 14:00 and 17:00 CST), respectively. The model prediction phase is started at 12 UTC and performs a 36 h prediction for each run.
3. Results and discussion

3.1. Spatial-temporal distribution of the PM$_{2.5}$ concentration from observations

The spatial-temporal distribution of the PM$_{2.5}$ concentration from satellite and ground site observations is shown in Fig. 2. Overall, the PM$_{2.5}$ retrievals show good agreement with the site observations (Fig. 2a), indicating the reliability of the satellite-based PM$_{2.5}$ estimation. The high PM$_{2.5}$ polluted regions are mainly located in the north of the domain and the center of the Hubei province (Fig. 2a). Passive satellite observations are highly influenced by weather, so the availability of satellite observations varies spatially and temporally. Overall, the spatial coverage rate is around 0.3–0.4 in most regions of the study area except for part of southwestern Hubei Province (Fig. 2b), with a coverage rate of ~0.1. In temporal, the satellite data coverage showed high heterogeneity in the influence of periodic atmospheric oscillation, ranging from 0 to 0.9 (Fig. 2c). During the study period, there are several PM$_{2.5}$ episodes, such as around Feb 4th and Feb 25th, 2020. Some of these periods have a high data coverage rate, some not (Fig. 2c), which may potentially affect the accuracy of PM$_{2.5}$ prediction, especially for the correlation coefficient in temporal. So, it is necessary to simultaneously assimilate multi-source data from satellite and ground sites since there is a continuous observation at ground sites.

3.2. The influence of DA on the initial condition

The accuracy of the PM$_{2.5}$ IC before and after DA was evaluated using the local PM$_{2.5}$ observations. The RMSE and $R^2$ of the PM$_{2.5}$ concentration of the noDa experiment were 39.23 μg/m$^3$ and 0.14. DA significantly improved the $R^2$ of the PM$_{2.5}$ IC to above 0.30, and reduced the RMSE to ~20 μg/m$^3$ for MAE, it was significantly reduced almost by half (Fig. 3a–c). The allDa run shows similar PM$_{2.5}$ IC accuracy with siteDa in $R^2$ (0.31 versus 0.32), but a smaller RMSE (21.28 μg/m$^3$ versus 22.11 μg/m$^3$) (Fig. 3b and c), indicating a wider spatial influence of the satellite observations in improving PM$_{2.5}$ predictions via data assimilation.

The spatial distribution of the PM$_{2.5}$ initial condition overlaid site observations is shown in Fig. 3d–f. It can be seen that noDa has a much higher PM$_{2.5}$ IC than observations, especially in the central-eastern Hubei Province where there are dense populations, industries, and meanwhile, the most affected area after the outbreak of the COVID-19. The result showed that after assimilating ground site observations, the concentration of PM$_{2.5}$ was significantly reduced and much closer to observations (Fig. 3e). However, the PM$_{2.5}$ concentrations are still overestimated in regions with sparse sites, such as central Hubei Province. The assimilation of satellite PM$_{2.5}$ further reduced the overestimation of the PM$_{2.5}$ IC in these areas, and showed a good agreement with the site observations throughout the entire domain (Fig. 3f). The significant overestimation of the initial field is caused by the unrealistic model anthropogenic emissions. Studies showed that PM$_{2.5}$ in Hubei Province decreased by 30%–60% during the COVID-19 epidemic compared to the same period of the past five years (Bai et al., 2021), which is similar to this study in proportion.

3.3. Daily variation of emission scaling factors

The daily variation of the emission scaling factors during the study period is shown in Fig. 4. Most of the scaling factors were in the range of 0.5–0.9. Still, the scaling factors showed a sudden increase at some moments, such as in Feb 13th, 15th, and Mar 9th, 2020, with the value high up to almost 1.5 (Fig. 4). It was unreasonable since actual emissions unlikely to rise that much in a short time. The sudden increase of the scaling factors shows an association with weather conditions. It can be seen that, around the date mentioned above, there was a sudden decrease in temperature, pressure, high relative humidity, and precipitation (Fig. 4b and c), indicating a high impact of meteorological factors on the calculation of the emission scaling factors. The median filtering with a three-day window size was applied to reduce this effect to deal with the spike of the scaling factors. Results show that most abnormal values were filtered and assigned a more reasonable value.

The COVID-19 outbreak led to dramatic changes in human activity, and subsequently, anthropogenic emissions. Multiple pollutants were reduced, affecting the primary or secondary formation of PM$_{2.5}$. Miya-zaki et al. (2020) showed a nationwide 23 μg/m$^3$ decrease in PM$_{2.5}$, and a 36% decrease in NOx. Zheng et al. (2021) showed a 27%, 28%, 31%, and 24% decrease in SO$_2$, CO, non-methane Volatile organic compounds (NMVOCs), and primary PM$_{2.5}$. The change of some pollutant types could be particularly significant. For example, Jia et al. (2021) showed a 70% drop in black carbon emissions in eastern China. Overall, these findings are quite similar to the emission scaling factors obtained here in this study. However, the reasonability of the scaled emissions is still hard to be verified since there is no true value. The scaled emissions will be verified using the WRF-Chem model in this study.
3.4. DA influence on PM$_{2.5}$ predictions

The spatial distribution of the averaged PM$_{2.5}$ predictions was drawn overlaid with site observations (Fig. 5). Results showed improved initial conditions via data assimilation delivered a more accurate PM$_{2.5}$ prediction. The model produced a much higher PM$_{2.5}$ prediction than observations when no data was assimilated, especially in central Hubei Province (Fig. 5a). By assimilating ground site observations, the overestimation of the PM$_{2.5}$ predictions was significantly improved (Fig. 5b). When satellite observations were further assimilated (allDa), the overestimation of the PM$_{2.5}$ predictions was further slightly brought down (Fig. 5c). However, the model still overestimated the PM$_{2.5}$ predictions in such areas were further improved (Fig. 5d). The temporal variations of the PM$_{2.5}$ predictions averaged at all sites and three typical cities (Wuhan, Xiangyang and Yichang) are shown in Fig. 6a. Overall, all experiments showed a
good ability to capture the PM$_{2.5}$ episodes. However, the DA run has a much higher R$^2$ and lower RMSE than noDa (Fig. 6a). In Wuhan, DA significantly improved the RMSE (R$^2$) by 10.08 μg/m$^3$ (0.24) after assimilated site observations (Fig. 6b). In Xiangyang, the RMSE (R$^2$) was significantly improved by 5.92 μg/m$^3$ (0.19) (Fig. 6c). The R$^2$ in Yichang did not change much, but the RMSE significantly decreased by 10.17 μg/m$^3$ (Fig. 6d). Simultaneous assimilation of the observations from sites and satellite further improved the RMSE of the PM$_{2.5}$ predictions by 1.55 μg/m$^3$ (Fig. 6a). Similarly, the advantage of simultaneous assimilation of satellite and ground site observations has been recognized by some studies (Chai et al., 2017; Hong et al., 2022; Schwartz et al., 2012). However, this study also showed that assimilating satellite observation
produced a slight decrease in $R^2$, which is largely due to the discontinuity of data availability. Satellite observations have an advantage in spatial coverage, but their availability is significantly affected by the weather. Therefore, when the satellite observations were available, the allDa showed more accurate PM$_{2.5}$ predictions than siteDa; when the satellite observations were unavailable, the allDa showed similar PM$_{2.5}$ predictions accuracy with siteDa. Adjusting emissions (allDaEa) further produced a more accurate PM$_{2.5}$ prediction than allDa. The results showed that the RMSE of allDaEa significantly reduced by 5.08 $\mu$g/m$^3$ (21.33%), and $R^2$ improved by 0.04 (Fig. 6a), indicating the importance of adjusting emission in improving PM$_{2.5}$ prediction.

Fig. 7 showed a further decrease of the RMSE resulting from the five experiments (RAW, noDa, siteDa, allDa, and allDaEa). The raw experiments were conducted using the conventional anthropogenic emissions without any modification, and no data assimilation was performed. With the improved anthropogenic emissions, the RMSE was reduced at all the sites in Hubei, especially at Wuhan, where the RMSE was reduced by $\sim$15 $\mu$g/m$^3$ (Fig. 7a), indicating an overall overestimation of the model due to unrealistic anthropogenic emissions. The siteDa significantly reduced the RMSE at all validation sites (Fig. 7b), with an average of about 10.64 $\mu$g/m$^3$ (23.0%). The allDa further reduced the RMSE of the PM$_{2.5}$ predictions, especially in central Hubei but the RMSE did not change much compared to that of the siteDa (Fig. 7c), indicating that the improvement of PM$_{2.5}$ forecasting by simply improving the accuracy of the initial field is limited. By daily adjusting the emissions, the PM$_{2.5}$ forecast accuracy was significantly improved at most sites with an overall reduction of RMSE by 4.41 $\mu$g/m$^3$ (15.14%) (Fig. 7d).

Fig. 8 showed the time series of the hourly averaged PM$_{2.5}$ predictions, correlation coefficient and the RMSE of the PM$_{2.5}$ predictions during the study period. It can be seen that DA significantly brought the PM$_{2.5}$ predictions closer to the observation at 03Z (Fig. 8b), from 70.91 $\mu$g/m$^3$ to 50.41 $\mu$g/m$^3$ and 47.61 $\mu$g/m$^3$ respectively for siteDa and allDa, much closer to observations (43.79 $\mu$g/m$^3$). The continuous assimilation of observations at 06Z and 09Z further reduced the absolute bias to 2.54 $\mu$g/m$^3$ and 1.39 $\mu$g/m$^3$ for siteDa, 1.3 $\mu$g/m$^3$ and 0.26 $\mu$g/m$^3$ for allDa, respectively. The benefits of the data assimilation can last up from 12Z to 48Z. However, with the increase of the forecast time, the benefits of DA diminished quickly. At 24Z, the RMSE of siteDa and allDa reduced by 7.53 $\mu$g/m$^3$ (15.3%) and 8.80 $\mu$g/m$^3$ (18.9%) compared to noDa. At 36Z, the RMSE of siteDa and allDa reduced by 5.48 $\mu$g/m$^3$ (11.4%) and 6 $\mu$g/m$^3$ (12.5%) compared to noDa, respectively (Fig. 8c). At 48Z, DA performed almost the same with noDa (Fig. 8d). DA benefit on PM$_{2.5}$ predictions generally can last for up to 48 h or more, largely depending on the severity of the pollution and the stability of the weather conditions (S. Feng et al., 2018; Kong et al., 2021). During the COVID-19 lockdown, the PM$_{2.5}$ concentrations in the study area were significantly reduced, and the weather conditions did not change to be more stable, thus leading to a relatively shorter DA benefits duration (36 h). However, by adjusting emissions, the PM$_{2.5}$ predictions showed a much slower error accumulation tendency along forecast time. At 48Z, the RMSE of allDaEa still has an 8.85 $\mu$g/m$^3$ (19.49%) improvement, suggesting the effectiveness of joint adjustment of the emissions using the proposed method. Previous studies have developed more sophisticated and mathematically more skilled approaches to jointly adjust the initial field and emissions, and yielded accurate PM$_{2.5}$ forecast (Peng et al., 2018; Peng et al., 2017). However, this study showed that, though simple, the proposed method can also produce a comparable PM$_{2.5}$ improvement in model accuracy but is less time-consuming. The benefits of adjusting emissions were particularly significant to improve PM$_{2.5}$ predictions in a metropolitan area like Wuhan (Fig. 9). Results showed that in Wuhan the RMSE of allDaEa still kept a 12.99 $\mu$g/m$^3$ (24.53%) improvement at 48Z, much higher than that in Xinyang (5.62 $\mu$g/m$^3$, 17.42%) and Yichang (9.70 $\mu$g/m$^3$, 22.18%), suggesting the dramatic changes of the emissions in metropolitan areas. However, results still

Fig. 7. Difference of RMSE of the 12–36h PM$_{2.5}$ predictions (a) between noDa and RAW experiment (b) between siteDa and noDa, (b) between allDa and siteDa, and (c) between allDaEa and siteDa of each observation site during the study period.
showed a high overestimation in PM$_{2.5}$ predictions even if joint adjusted the aerosol initial conditions and emissions (Fig. 9a–c), indicating that there is still room for a further improvement in emissions.

4. Summary and conclusion

This study proposed to assimilate multi-source observations and jointly adjust emissions to minimize the influence of the emission
reduction and improve PM$_{2.5}$ predictions during the COVID-19 epidemic. Five experiments were conducted, including a raw experiment (RAW), control experiment (noDa), two assimilation experiments that assimilated site observations only (siteDa), and assimilated both site and satellite observations simultaneously (allDa), and an emission adjustment experiment that adjusted the emissions daily based on the allDa experiment (allDaEa).

Results showed that current anthropogenic emissions produced a large systematic bias of PM$_{2.5}$ predictions (BIAS is $+35.43$ μg/m$^3$, RMSE is 54.55 μg/m$^3$). DA reduced the uncertainties of initial conditions and better reflected the spatial-temporal distribution of the actual particulate pollution. With the improved IC, the WRF-Chem model produced more accurate PM$_{2.5}$ predictions, which can last for 36 h. However, the benefits of DA diminished quickly with the increase of the forecast time. By adjusting emissions, the PM$_{2.5}$ predictions showed a much slower error accumulation along forecast time. At 48Z, the RMSE still has an 8.85 μg/m$^3$ (19.49%) improvement, suggesting the effectiveness of emissions adjustment based on the improved IC via DA.

The prediction skill of the model is determined by both the accuracy of the initial field and emissions, as well as related physicochemical processes. Although this study yielded a significant improvement of PM$_{2.5}$ predictions, results still showed a high overestimation even if joint adjusted the aerosol initial conditions and emissions, indicating that there is still room for further improvement. The epidemic is still spreading worldwide. Strict control and “zero Covid” strategy will probably remain in China before a significant change of the virus itself or the anti-virus measures. In the foreseeable future, every outbreak of the COVID-19 cases in China will significantly reduce human activities and anthropogenic emissions. How to accurately predict PM$_{2.5}$ under such situations is still challenging. Having more observations assimilated into the model and using grid-based emissions scaling factor would be feasible approaches.

Credit author statement

Liu Zhao Chen: Software, Validation, Writing – original draft, Visualization Feiyue Mao: Conceptualization, Resources, Writing – review & editing, Supervision. Jia Hong: Methodology, Writing – review & editing Lin Zhang: Methodology. Jiangping Chen: Resources. Yi Zhang: Methodology. Yuan Gan: Software. Wei Gong: Resources. Houyou Xu: Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This study is supported by the National Natural Science Foundation of China (41971285 and 41627804), the National Key Research and Development Program of China (2017YFC0212600), and the LIEMARS Special Research Funding. The numerical calculations in this paper were done on the supercomputing system in the Supercomputing Center of Wuhan University.

References

Bai, H., Gao, W., Zhang, Y., Wang, L., 2021. Assessment of health benefit of PM2.5 reduction during COVID-19 lockdown in China and separating contributions from anthropogenic emissions and meteorology. J. Environ. Sci. 115, 422–431. https://10.1016/j.jes.2021.01.022.

Breiman, 2001. Random forests. Mach. Learn. 45 (1), 5–32. https://10.1023/A:1010903340432.

Chai, T., Kim, H.C., Pan, L., Lee, P., Tong, D., 2017. Impact of moderate resolution imaging spectroradiometer aerosol optical depth and AIRNow PM2.5 assimilation on community multi-scale Air quality aerosol predictions over the contiguous United States. J. Geophys. Res. 122 (10), 5399–5415. https://10.1002/2016JD026295.

Chin, M., Ginoux, P., Kinne, S., Torres, O., Holben, B.N., Duncan, B.N., Nakajima, T., 2002. Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and sun photometer measurements. J. Atmos. Sci. 59 (3), 461–483. https://10.1175/1520-0469(2002)059<0461:TAOTFT>2.0.CO;2.

Copat, C., Nistadi, A., Fiore, M., Grasso, A., Zuccarello, S., Santo Signorelli, S., Ferrante, M., 2020. The role of air pollution (PM and NO2) in COVID-19 spread and lethality: a systematic review. Environ. Res. 110129. https://10.1016/j.envres.2020.110129.

Courtier, P., Anderson, E., Heckley, W., Vasilevich, D., Pailleux, J., 1998. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. Q. J. R. Meteorol. Soc. 124 (550), 1783–1807. https://10.1002/qj.4971245505.

Dudhia, Jimy, 1989. Numerical study of convection during the winter monsoon experiment using a mesoscale two-dimensional model. J. Atmos. Sci. 46 (20), 3077–3107. https://10.1175/1520-0469(1989)046<3077:NSOCDE>2.0.CO;2.

Evenens, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. 99 (CT), 10143–10162. https://10.1029/94JC00572.

Fei, Chen, Dudhia, Jimy, 2001. Coupling an advanced land surface-hydrology model with the penn state-ncar MAM5 modeling system. Part I: model implementation and sensitivity. Mon. Weather Rev. 129, 569–585. https://10.1175/1520-0493(2001)129<0569:CAMSLH>2.0.CO;2.

Geoffroy, G.A., Peckham, S.E., Schmitz, R., Mckeen, S.A., Frost, G., Skamarock, W.C., Eder, B., 2005. Fully coupled “online” chemistry within the WRF model. Atmos. Environ. 39 (37), 6957–6975. https://10.1016/j.atmosenv.2005.04.027.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P.L., Geron, C., 2006. Estimates of global terrestrial isoprene emissions using MEGAN (model of emissions from Gases and Aerosols from Nature). Atmos. Chem. Phys. 6 (11), 3181–3210. https://10.5194/acp-6-3181-2006.

Hong, J., Mao, F., Gong, W., Gan, Y., Zhang, L., Quan, J., Chen, J., 2022. Assimilating Fengyun-4A observations to improve WRF-Chem PM2.5 predictions in China. Atmosphere 13, 973–993. https://10.3390/atmos13110973.

Hong, S.Y., Noh, Y., Dudhia, J., 2005. A new vertical diffusion package with an explicit treatment of entrainment processes. Mon. Weather Rev. 134 (9), 2318. https://10.1175/MWR3394.1.

Hong, S.Y., Pan, H.L., 1996. Nonlocal boundary layer vertical diffusion in a medium-resolution atmospheric model. J. Atmos. Soc. 59 (3), 461–483. https://10.1175/1520-0469(1996)059<0461:NSOCDE>2.0.CO;2.

Kleist, D.T., Parrish, D.F., Derber, J.C., Treadon, R., Wu, W.S., Lord, S., 2009. 3DVAR to 4DVAR: implementation of 4DVAR at the meteorological service of Canada. Mon. Wea. Rev. 137 (6), 2339–2354. https://10.1175/1520-0493(2009)137<2339:3DVAR4DVAR>2.0.CO;2.

Kong, Y., Sheng, D., Liu, C., Chen, H., Zhang, W., Zhou, Y., Jiang, X., 2020. The role of anthropogenic aerosol emissions in delivery and lethality: a systematic review. Environ. Res. 110129 https://doi.org/10.1016/j.envres.2020.110129.

Ko, W., Noh, Y., Hong, S.Y., Jong, W., Lee, J., 2019. Impact of domain size on modeled ozone forecast for the northeastern United States. Atmos. Res. 248, 1–10. https://10.1016/j.atmosres.2020.105366.

Koizumi, H., Kondo, Y., Zhu, T., Zhang, X., Sudo, K., 2008. A global three-dimensional multivariable statistical interpolation scheme. Mon. Wea. Rev. 136 (4), 1222–1237. https://10.1175/2007MWR2440.1.
