IMPROVING CO FORECASTING AND EMISSION ESTIMATION BY COUPLING A DIMENSION–REDUCTION DATA ASSIMILATION METHOD WITH THE COMMUNITY MULTISCALE AIR QUALITY MODEL

LU, L. J.[1,2] – WANG, J. P.[3*]

[1]*College of Resource and Environment, Anhui Science and Technology University, Bengbu 233100, China*
(e-mail: lulj@ahstu.edu.cn)

[2]*Key Laboratory of Digital Rural Construction and Governance, Anhui Provincial Department of Science and Technology, Bengbu 233000, China*

[3]*Shandong Marine Resources and Environment Institute, Yantai 264000, China*

**Corresponding author*
e-mail: jianping_wang2010@163.com

(Received 13th Jul 2023; accepted 1st Sep 2023)

Abstract. Carbon monoxide (CO) is one of the six major pollutants and mainly comes from anthropogenic emissions. CO has active chemical properties, and its concentrations vary distinctly in the atmosphere, thus creating great uncertainty in forecasts. In this study, we conducted a data assimilation forecast system, i.e., CO-FCST, by coupling the proper orthogonal decomposition (POD) technique with an atmosphere transport model, i.e., CMAQ (Community Multiscale Air Quality model). The CO-FCST was then used in two districts, i.e., Shangdianzi (SDZ, located 40 km northeast of Beijing) and Waliguan (WLG, located in Qinghai Province, China), in China for CO concentration forecasts with a couple of continuous hourly surface CO observation datasets. The results show that after optimization by the CO-FCST system, CO concentrations had significant improvements in forecasting in both areas. In addition, the two posteriors were much higher than the two priors, i.e., 31.7% higher in SDZ and 62.3% higher in WLG.
Keywords: *data assimilation, proper orthogonal decomposition (POD), CMAQ, CO, forecast*

Introduction

Air pollution incidents, which are always caused by excess anthropogenic emissions in the atmosphere, have attracted increasing attention in recent years. Most air pollution, e.g., CO and NO_x, is generally considered to be greatly influenced by anthropogenic activities such as the burning of fossil fuels, i.e., petrol, coal, and natural gas (Kalogridis et al., 2017; Turnbull et al., 2011). The air pollution has a significant impact on the environment, public health and our lives (Aziz, 2007; Chen et al., 2007). It is urgent to diagnose the spatiotemporal distribution and variability of air pollutant emissions, based on which the government can take effective measures for emission reduction and concentration forecast. CO emissions can be obtained through "bottom-up" statistical methods, although these methods are labor intensive and time consuming (Zhao et al., 2011). The "top-down" method, which greatly saves time and costs, has been used by many researchers in recent years, and a couple of emission fluxes have been established in China at different scales, ranging from urban scales to national levels (Detlef et al., 2009; Vaughn et al., 2018). Unfortunately, these inventories mostly have low spatiotemporal distribution resolutions, and due to the large uncertainties in these inventories, they greatly hamper our ability to diagnose and forecast air quality (Flerlage et al., 2021; P Cheewaphongphan et al., 2019).

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2202_15091522

Many mathematical methods, such as data assimilation and Kalman filtering, have been coupled with the "top-down" approach to reduce the uncertainties in the estimates of emissions (Ma et al., 2019; Lopez-restrepo et al., 2020). Many emission inventories, ranging from global to regional scales, have been inverted by these two kinds of algorithms. Among the data assimilation methods, four-dimensional variational assimilation (4DVar) has been used frequently in recent years because of its advantage of fully absorbing physical models and observations (Yi et al., 2018; He et al., 2022). To calculate the gradient of the cost function, iterative integration must be conducted in the simulation procedure. However, one of the largest problems for using 4D-VAR with the CMAQ model is the large dimension of the control space, which can reach an order of 10^7–10^8. Therefore, massive computational resources are required to obtain integration results and the adjoint model for the CMAQ with highly nonlinear characteristics (Resler et al., 2007; Andronopoulos et al., 2011). The proper orthogonal decomposition (POD) is designed for reducing the dimension of the control space, which can be used for solving the massive computational problem (Cao et al., 2007; Siade et al., 2010). Many experiments have shown that if the samples are chosen appropriately, the new space generated by the POD method can represent the control space well (Tian et al., 2015; Jin et al., 2018).

In order to improve the accuracy of CO concentration forecast, the prior flux must be optimized in both spatial and temporal dimensions. In this paper, a forecast system, CO-FCST, was constructed by incorporating the POD4DVAR method into the CMAQ model, and the system was then used in two districts of China, i.e., SDZ and WLG, for CO concentration forecasting based on hourly surface CO concentration observations. WLG is located in the middle-western region of China with a relatively underdeveloped industrial foundation, and a large portion of atmospheric CO is affected by biogenic activities. In contrast, SDZ is located in the Beijing–Tianjin–Tangshang area, which is one of the most industrialized regions of China, and its atmospheric CO is dominated by anthropogenic activities such as the burning of fossil fuels (i.e., petrol, coal, and natural gas). The forecast results were validated with CO observations, and a comparison between the posteriors and priors was conducted.

Data and method

Study area

China has a vast land area covering approximately 9,600,000 km^2 with unbalanced development in different regions, which leads to a great difference in anthropogenic CO emissions between the western and eastern parts of the country. Two different districts, i.e., WLG (Waliguan, located in Qinghai Province, China) and SDZ (Shangdianzi, located 40 km northeast of Beijing) (*Fig. 1*), can perfectly represent two areas with different levels of industrialization.

SDZ is located in northern China, which has many highly industrialized cities, e.g., Beijing, Tianjin, and Tangshan, discharging a large amount of CO every year through people's daily activities. WLG, unlike SDZ, lies in northwestern China. This area contains many small cities, e.g., Xining and Gannan, and most cities in this area do not have a good industrial foundation. In this area, atmospheric CO is influenced simultaneously by anthropogenic and natural emissions. To obtain CO concentrations, 22 and 9 atmospheric monitoring sites were established in SDZ and WLG, respectively.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2202_15091522

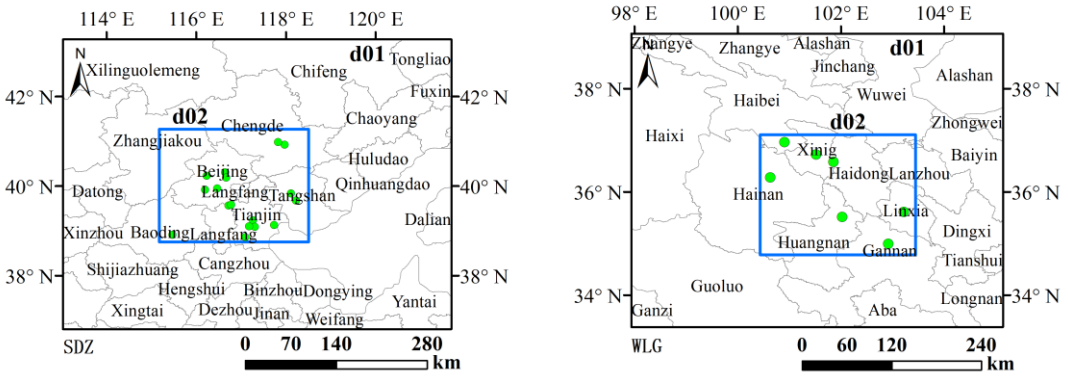

Figure 1. *Two study areas. d02 is the inner area, and d02 is the boundary layer area for providing boundary fields for d01. Green dots are the CO monitoring sites used in this paper*

Framework of CO-FCST

Currently, most air quality models, e.g., CMAQ, CMax, require an initial grid emission flux, which we called a prior. It is a prerequisite for model running and can be obtained from some organizations and institutions, e.g., EDGAR (https://edgar.jrc.ec.europa.eu/) and MEIC (http://meicmodel.org.cn/). These fluxes always have a coarse resolution and cannot meet the needs for high-resolution forecasts. The purpose of our research is to iteratively optimize the prior flux through the data assimilation methods, ultimately improving the accuracy of CO concentration forecasts. In this paper, we coupled the POD method, the 4DVar method with the CMAQ model, where CMAQ serves as the transport model containing a chemical mechanism module. The 4DVar method is used to fix the prior flux for optimized posterior fluxes, ultimately obtain accurate forecasts of CO.

Atmospheric pollutant simulation always has a high dimension of approximately 10^7–10^8, which makes further analysis and examination difficult. The POD method can characterize the spatiotemporal evolution of atmospheric pollutants with fewer variables, and then the solution can be carried out in a lower dimension. In this paper, a system, i.e., CO-FCST, for CO concentration forecasting was established by integrating the CMAQ and the POD4DVar techniques.

Suppose a forecast window (0, *W*) has *V* steps, e.g., 4 steps in one day with a step length of 6 h, then CO concentrations can be forecasted as follows:

(1) A cost function is defined as follows:

$$J(x) = (x - x_b)^T F^{-1}(x - x_b) + \sum_{i=0}^{n} \left(y_i - H_i\left[M_{t_0 \to t_k}(x)\right]\right)^T O_i^{-1} \left(y_i - H_i\left[M_{t_0 \to t_k}(x)\right]\right) \qquad \text{(Eq.1)}$$

where *M* is the forecast model CMAQ, *T* is the transpose, x_b is the first guess field, *k* is the observation timestep, *H* is the observation operator, and *F* and O_i are the background and observational error covariances, respectively.

(2) *N* first guesses of emissions are generated, $x_{0,n}(n = 1, \ldots, N)$, by using the Monte Carlo sampling method from the historical estimations.

(3) The CMAQ model is integrated with the *N* first guesses for *N* state series $x_{k,n}(k = 1, \ldots, V)$, and then then the sample ensemble can be obtained $X_n = (x_{0,n}, \ldots x_{v,n})^T$.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2202_15091522

(4) Matrix $A = (X_1, X_2, \ldots, X_N)$ of ensemble perturbations is reconstructed by X_n and the base vectors, P_n (PODs), are obtained from the ensemble space by using the proper orthogonal decomposition technique. The size of n, generally, is much smaller than *N*, and the optimal emission estimation can be expressed linearly by the PODs.

(5) The CMAQ model is reintegrated for optimal forecasts of CO. The CMAQ model is run with the P_n vectors from the previous step throughout the forecast time window; then, *n* simulations of CO concentrations are obtained.

(6) Steps (1)-(5) are repeated until all CO forecasts in the time window are completed. The CO-FCST system was run in the Lambert conformal projection with a two-nested mode (*Fig. 1*), and the d02 domain had a resolution of 3 km. The d01 had a resolution of 9 km, which was used for providing the boundary layer. In our study, the CO-FCST system was used to assimilate the observed concentrations in each cycle window. CO emissions were re-evaluated in two study areas, which were used for forecasting CO concentrations in every forecast time window.

Data

The meteorological field used in the CO-FCST system was the FNL reanalysis, and the dataset can be downloaded from the National Centers for Environmental Forecast (NCEP) with spatiotemporal resolutions of one degree and 6 h, respectively. The CO priors were derived from the Multiresolution Emission Inventory for China (MEIC, http://www.meicmodel.org/index.html), and the priors contain most major atmospheric pollutants of Asia in 2010, including CO, CO2, SO2, PM2.5, PM10, BC, NH3, NOx, NMVOC, OC, etc. The priors have a spatial resolution of 0.25 degrees for five sectors, i.e., residential, transport, power, industry and agriculture (Li et al., 2017). The in situ CO observations were derived from a data website of The Ministry of Environment Protection of the People's Republic (http://106.37.208.233:20035/). Hourly average CO observations of over 496 sites in China have been published on this website. In this paper, the CO baseline was extracted by the following steps: (1) Abnormal data were eliminated. Abnormal data are mainly caused by instrument failure or other emergencies. Such observation data cannot reflect the real CO concentration; therefore, they need to be deleted before use. (2) Pollution observations were deleted. The threshold value (observation value obviously containing incremental concentration) was set and observation data exceeding the threshold value were deleted. (3) In the remaining CO concentration observation data, 72 h was taken as the smooth window to remove observation values that exceeded ± 3 times the standard deviation of the mean value. (4) Circular filtering was performed. Step (3) was repeated until the number of remaining datasets tended to be stable. Fitting the remaining observation data provided the final CO background concentrations.

Sensitivity experiments

Forecasts of CO could be influenced by many factors, e.g., the boundary layer, chemical mechanisms, lag window, perturbing samples, etc. These parameters need to be reasonably optimized in realistic applications, and errors might increase significantly if these factors are not considered. In this paper, some sensitivity experiments were conducted for the CO-FCST system.

The sample number is a very important factor for forecasting CO in the CO-FCST system. In this paper, several experiments were conducted with different perturbing

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2202_15091522

samples, i.e., 10, 20, …, 200. The CO simulations based on the inversion fluxes were compared to the observations, and the results showed that the CO-FCST system was able to perform very well when the number of perturbing samples exceeded 136.

Generally, boundary layers can provide boundary status continuously for the inner area, thereby improving the simulation accuracy of CO concentrations. A realistic simulation process can lead to significant errors if the boundary layers are ignored. In this paper, a nested mode including d01 and d02 (*Fig. 1*) was adopted, and d02 could provide a boundary field for d01 in all timesteps.

In the simulations of CO_2, the lag window (lw) can directly affect the accuracy because it represents the amount of data involved in the assimilation. In our study, several experiments for diagnosing the influence of the lag window of CO were conducted before realistic applications. The results showed no significant differences in these experiments. In this paper, the lag window was set to 0 in the CO-FCST system for all simulation processes.

Ultimately, for realistic applications in SDZ and WLG, the ensemble size was set as 166, and two nested modes were adopted. The chemical mechanism was set as CB05, and the lag window was adopted as 0.

Indices for evaluations

Two indices were used for assessing the performance of forecasts by the CO-FCST system, i.e., mean error (*ME*) and correlation coefficient (*R*). These indices are illustrated as follows:

$$ME = (e_1 + e_2 + \cdots + e_n)/n \quad \text{(Eq.2)}$$

$$R = \frac{\sum_{i=1}^{n}(o_i - \bar{o})(f_i - \bar{f})}{\sqrt{\sum_{i=1}^{n}(o_i - \bar{o})^2}\sqrt{\sum_{i=1}^{n}(f_i - f)^2}} \quad \text{(Eq.3)}$$

n represents the size of the forecast–observation pairs, *e* represents the difference between a pair of forecasts and observations, and *o* and *f* represent the observations and forecasts, respectively.

Results and discussion

Sensitivity experiment for the forecast time window

The priors in SDZ and WLG were provided with a monthly resolution, which greatly hampered our ability to describe the diurnal variation in CO, thus creating significant uncertainties in CO concentration forecasting. In this paper, the priors were optimized before each forecast cycle by the CO-FCST system, which could greatly improve the forecast accuracy. However, the forecast accuracy based on the optimized priors fluctuated with variation in the forecast time window. In this paper, several experiments were conducted to diagnose the detailed influence of different forecast time windows. The results showed that the forecast accuracy of the system decreased significantly with the increase in the forecast time window, especially when the forecast time window exceeded 2 days (*Fig. 2*). In SDZ, there was no significant difference in forecast accuracy between 6 h and 18 h within a day, but after the forecast time window exceeded 2 days, the forecast error increased sharply from 0.09 mg/m^3 to 0.27 mg/m^3

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2202_15091522

(*Fig. 2*). In WLG, the forecast accuracy had a similar performance; when the forecast time window was shorter than one day, i.e., 6 h and 18 h, the forecast accuracy had no significant differences, i.e., 0.07 mg/m^3 and 0.05 mg/m^3, respectively (*Fig. 2*). When the forecast time window was changed to 15 days and 30 days, the forecast accuracy decreased sharply, and the forecast error increased to 0.13 mg/m^3 and 0.14 mg/m^3 in SDZ and WLG, respectively.

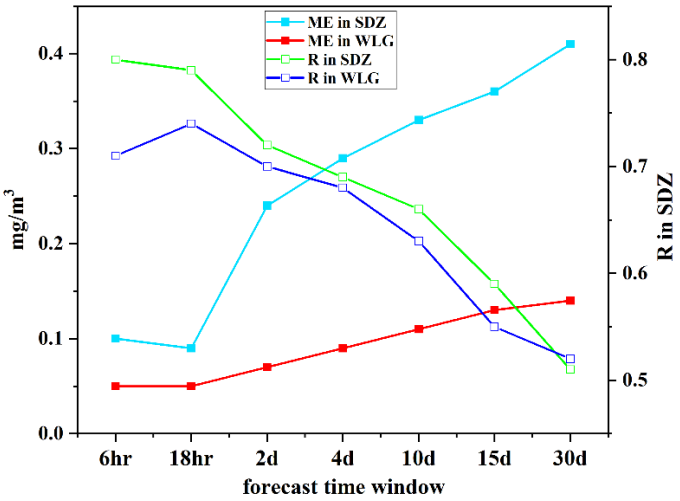

Figure 2. *ME (mean errors) and correlation R with different forecast time windows*

The performance of the CO-FCST system had significant differences in SDZ and WLG (*Table 1*). The average forecast error in SDZ for the whole forecast cycle window was 0.27 mg/m^3, which was significantly greater than that in WLG, i.e., 0.09 mg/m^3. This was mainly due to the great difference between the intensity of anthropogenic activities and emissions in the two regions. The CO emission intensity in SDZ is several times that in WLG, and the CO concentration changes forcefully in SDZ, which creates a great challenge for CO concentration simulations and forecasts, ultimately resulting in significant errors.

However, if the forecast time window is too small, it needs many computational resources; in the realistic application in SDZ and WLG, a forecast time window of 18 h was adopted.

Table 1. *MEs and Rs of different forecast time windows in SDZ and WLG*

	SDZ		**WLG**	
	ME (mg/m^3)	**R**	**ME (mg/m^3)**	**R**
6 h	0.16	0.74	0.07	0.71
18 h	0.09	0.79	0.05	0.74
2 d	0.24	0.72	0.07	0.70
4 d	0.29	0.69	0.09	0.68
10 d	0.33	0.66	0.11	0.63
15 d	0.36	0.59	0.13	0.55
30 d	0.41	0.51	0.14	0.52
Ave	0.27	0.67	0.09	0.65

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2202_15091522

Evaluation for forecasts

With the optimized priors, i.e., the posteriors, the CO-FCST system was reran for forecasts in every forecast cycle, and the forecasts were collected for comparison with observations in the two study areas (*Figs. 3* and *4*).

From the perspective of the whole year, regardless of SDZ or WLG, the CO concentration had significant seasonal variation; that is, the concentration began to decline at the end of spring, decreased to the lowest point at the end of autumn, and then rose sharply in winter until it reached the highest point.

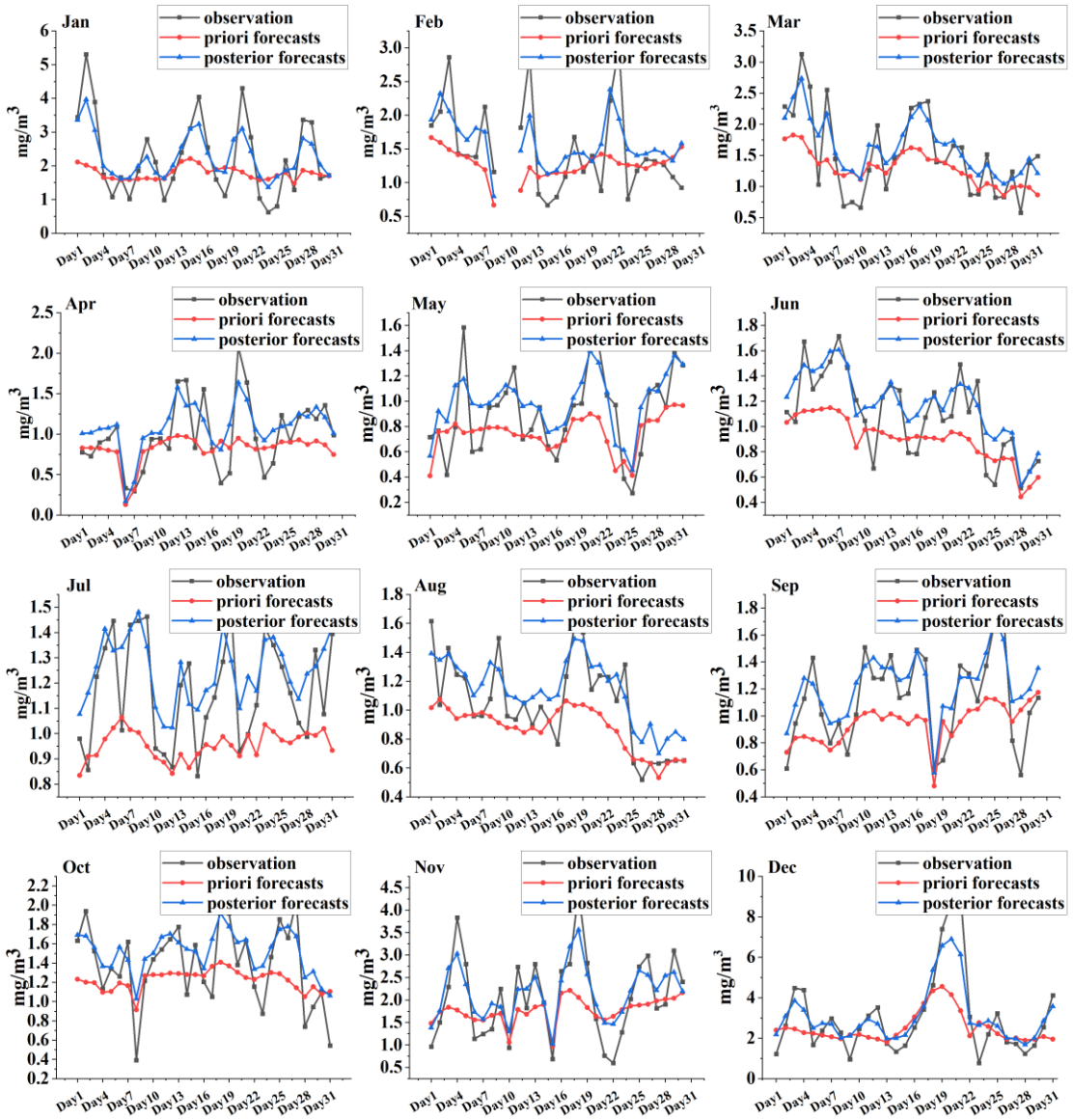

Figure 3. *Differences in CO observations, prior-based forecasts and posterior-based forecasts in SDZ*

The results showed that CO forecasts based on the priors could not grasp well the basic variation trend, and there was poor agreement between the CO forecasts and observations. In SDZ, the means of the forecast errors in winter, spring, summer, and

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2202_15091522

autumn were -0.32 mg/m^3, -0.17 mg/m^3, -0.18 mg/m^3 and -0.17 mg/m^3 respectively. It was difficult to forecast CO concentrations in winter because of the great variation, and the forecast accuracy in other seasons was much higher than that in winter. In WLG, the mean errors of prior-based forecasts were -0.14 mg/m^3, -0.16 mg/m^3, -0.22 mg/m^3 and -0.31 mg/m^3 in spring, summer, autumn and winter, respectively, which were significantly lower than those in SDZ, but CO concentrations in winter were still hard to forecast (*Table 2*). The main reason is that the prior flux resolution is too low, which fails to capture the rapid changes in pollutant concentration. In this study, the spatial-temporal resolution of the prior flux were improved obviously to better capture these variations in CO concentration.

Table 2. *Errors of the prior-based forecasts and posterior-based forecasts in SDZ and WLG*

	Time	**Evaluation Indices**	**SDZ (mg/m^3)**	**WLG (mg/m^3)**
Prior-Obs	Spring	ME	-0.17	-0.14
		R	0.57	0.54
	Summer	ME	-0.18	-0.16
		R	0.63	0.57
	Autumn	ME	-0.17	-0.22
		R	0.63	0.63
	Winter	ME	-0.32	-0.31
		R	0.62	0.64
	Annual	ME	-0.21	-0.21
		R	0.61	0.6
Posterior-Obs	Spring	ME	0.09	0.08
		R	0.73	0.68
	Summer	ME	0.09	0.09
		R	0.72	0.74
	Autumn	ME	0.11	0.03
		R	0.76	0.76
	Winter	ME	0.09	0.06
		R	0.82	0.78
	Annual	ME	0.09	0.06
		R	0.76	0.74

After the priors were optimized by the CO-FCST system, the forecast accuracy of CO was significantly improved, and the annual forecast error was reduced from -0.21 mg/m^3 to 0.09 mg/m^3 in SDZ and from -0.21 mg/m^3 to 0.06 mg/m^3 in WLG. From a seasonal perspective, forecast errors decreased in spring, summer, autumn and winter and they were also different. In SDZ, the means of forecast errors in spring, summer, autumn and winter were 0.09 mg/m^3, 0.09 mg/m^3, 0.11 mg/m^3 and 0.09 mg/m^3, respectively. The CO concentration in winter was still difficult to forecast, and the forecast accuracy in other seasons was still higher than that in winter. In WLG, the mean forecast error in spring, summer, autumn and winter was 0.08 mg/m^3, 0.09 mg/m^3, 0.03 mg/m^3 and 0.06 mg/m^3, respectively, which was significantly lower than that in SDZ (*Table 3*).

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2202_15091522

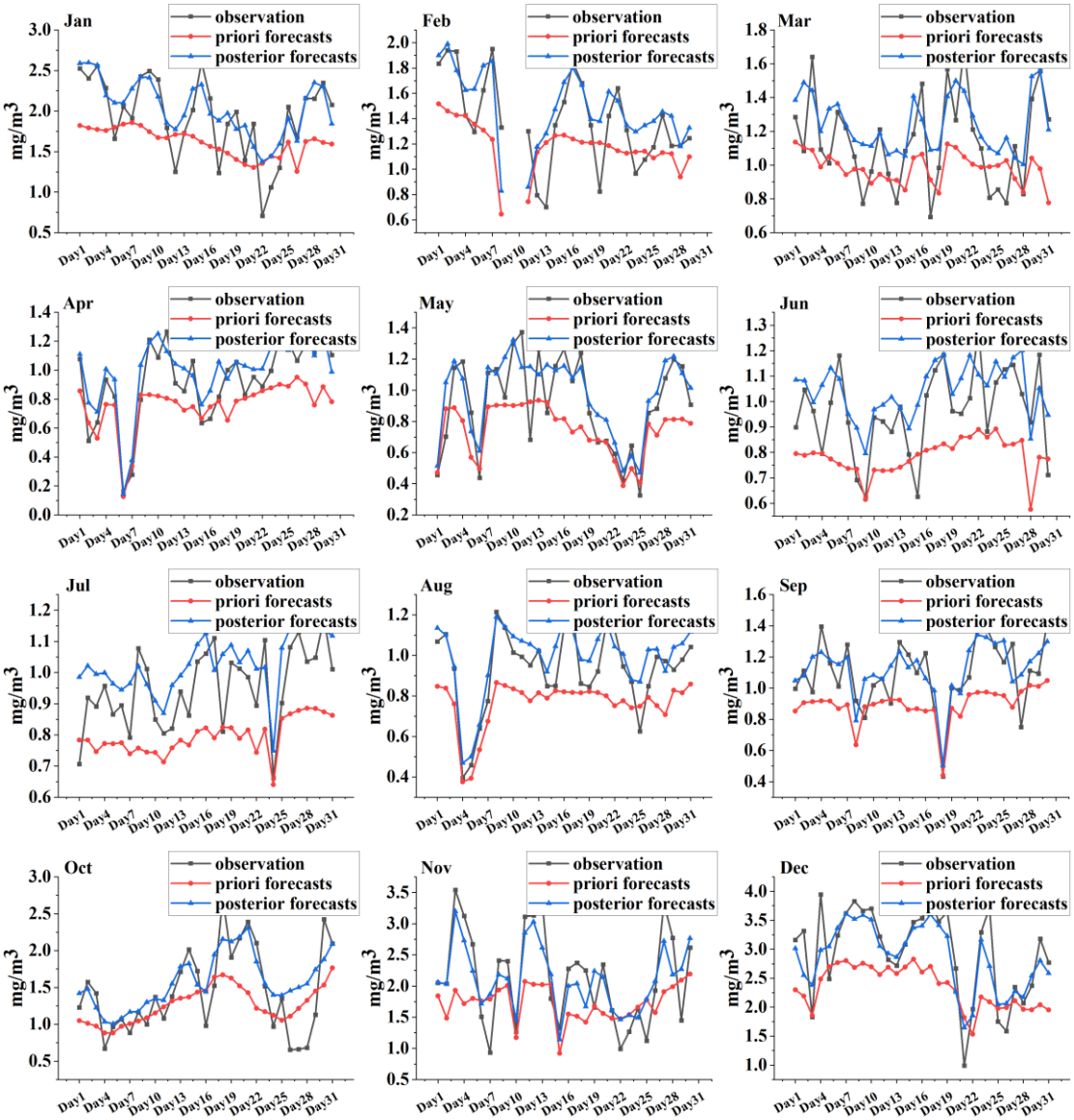

Figure 4. *Differences in CO observations, prior-based forecasts and posterior-based forecasts in WLG*

At the same time, we can see that the correlation between the forecasts and observations has also been significantly improved in both areas. In SDZ, the correlation of the annual simulation increased from 0.61 to 0.76, and the correlation in spring, summer, autumn and winter also increased from 0.57 to 0.73 in spring, 0.63 to 0.72 in summer, 0.63 to 0.76 in autumn, and 0.62 to 0.82 in winter, respectively. In WLG, the correlation of the annual simulation increased from 0.6 to 0.74, and the correlation of spring, summer, autumn and winter also increased from 0.54 to 0.68 in spring, 0.57 to 0.74 in summer, 0.63 to 0.76 in autumn, and 0.64 to 0.78 in winter, respectively.

Posteriors

As mentioned above, the forecast accuracy had significant improvement because of the optimized posteriors from the CO-FCST system. The two posteriors in SDZ and

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2202_15091522

WLG are shown in *Figures 5* and *6*. The results show that the priors were underestimated in both areas, and the posteriors were higher than the priors in most periods. In SDZ, the posteriors increased by 38.6%, 33.3%, 34.2%, 31.9%, 27.8%, 27.6%, 30.8%, 25.8%, 29.9%, 26.1%, 36.4% and 37.9% from January to December, respectively, compared with the priors, and the annual increasing ratio was 31.7% (*Table 3*). From a monthly perspective, the posteriors increased the most in January and December and the least in August and October.

Table 3. *Priors and increases of the posteriors in SDZ and WLG*

	SDZ			WLG		
	Pri-Avg (mole/s)	**Increase (mole/s)**	**Increase (%)**	**Pri-Avg (mole/s)**	**Increase (mole/s)**	**Increase (%)**
Jan	0.58	1.65	38.6	0.09	0.15	64.6
Feb	0.52	1.44	33.3	0.08	0.14	62.0
Mar	0.37	1.22	34.2	0.07	0.12	59.5
Apr	0.29	1.04	31.9	0.04	0.07	56.5
May	0.27	1.03	27.8	0.02	0.05	54.4
Jun	0.25	1.00	27.6	0.02	0.05	54.5
Jul	0.28	0.99	30.8	0.02	0.05	54.6
Aug	0.25	0.98	25.8	0.02	0.05	54.6
Sep	0.24	0.87	29.9	0.02	0.05	56.0
Oct	0.24	0.91	26.1	0.04	0.07	57.9
Nov	0.42	1.20	36.4	0.08	0.14	61.2
Dec	0.50	1.36	37.9	0.08	0.15	62.6
Annual	0.35	1.14	31.7	0.05	0.09	62.3

From an annual perspective in SDZ, the annual average of the prior was 1.14 mol/s, the minimum of which was 0.87 mol/s appearing in September, and the maximum was 1.65 mol/s appearing in January. The annual average of the posterior was 1.49 mol/s, which was approximately 1.3 times that of the prior. The minimum was 1.11 mol/s occurring in September, and the maximum was 2.23 mol/s occurring in January (*Table 3*). Overall, the variation trend of the posterior flux was basically consistent with that of the prior.

In WLG, the posteriors were also higher than the priors in the whole year, but the increasing ratios were much greater than those in SDZ. From January to December, the increased ratios of the posteriors were 64.6%, 62.0%, 59.5%, 56.5%, 54.4%, 54.5%, 54.6%, 54.6%, 56.0%, 57.9%, 61.2% and 62.6%, respectively, compared with the priors. The overall increasing ratio, i.e., 62.3%, was approximately 1.6 times that in SDZ, and the main reason for this might be the low value in the priors. WLG had a relatively high increase in the whole area, but the net increase of the posteriors in WLG (0.05 mol/s) was much smaller than that in SDZ (0.33 mol/s), which was only approximately one seventh of that in SDZ.

From the perspective of the annual average in WLG, the mean of the prior was 0.09 mol/s, the minimum of which was 0.05 mol/s appearing from May to September, and the maximum was 0.15 mol/s appearing in December and January. The mean of the posteriors was 0.14 mol/s, which was approximately 1.6 times that in the priors. The

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2202_15091522

minimum was approximately 0.07 mol/s, occurring from May to September, and the maximum was 0.24 mol/s, appearing in January. In addition, the overall trend of the posteriors was also basically consistent with the priors.

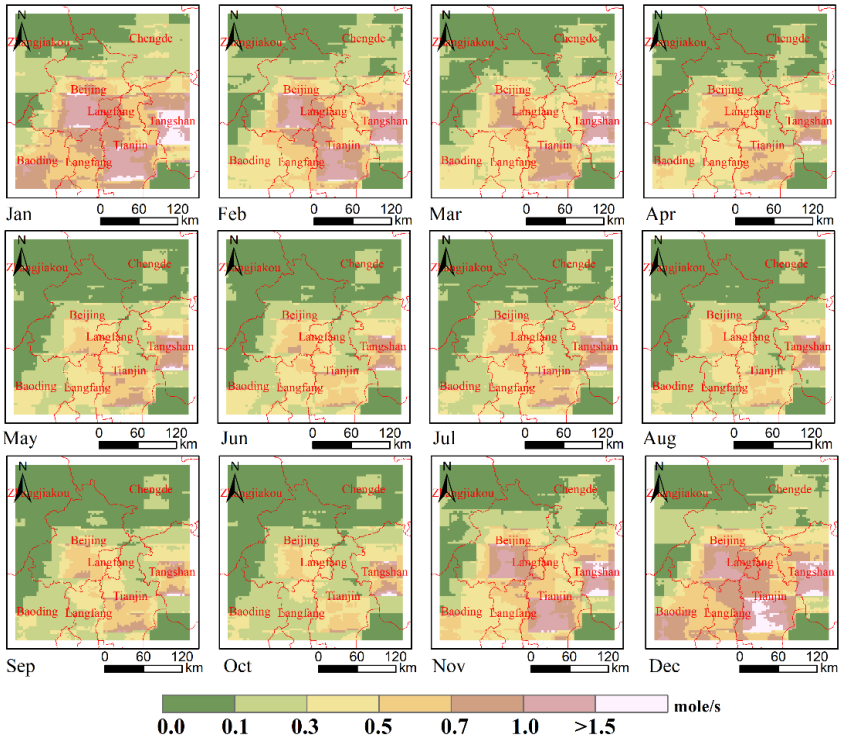

Figure 5. *Monthly posterior optimization by the CO-FCST system in SDZ*

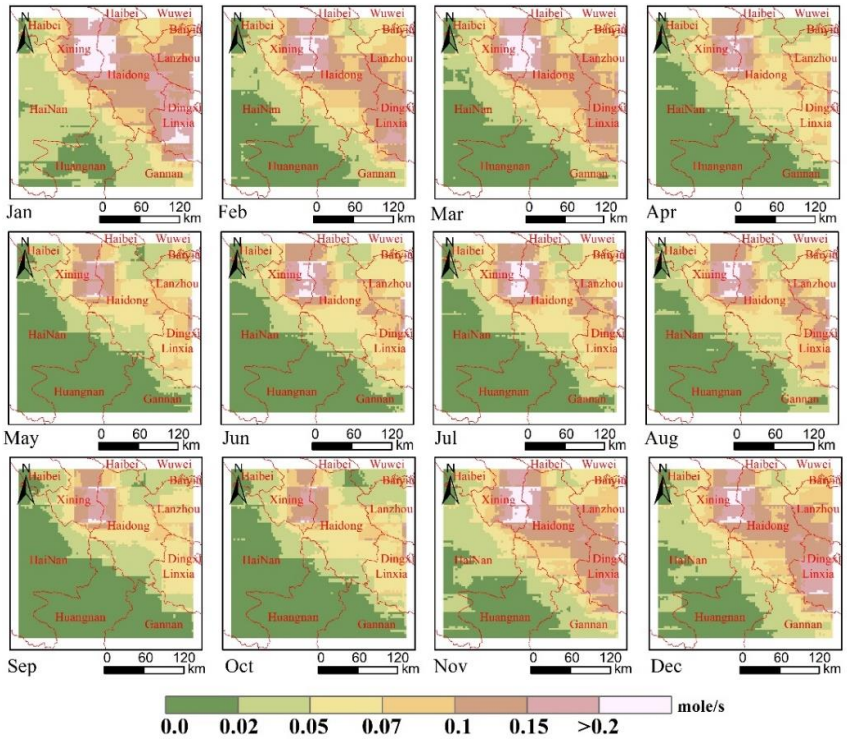

Figure 6. *Monthly posterior optimization by the CO-FCST system in WLG*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2202_15091522

The results also showed that there were significant differences between the increase and increasing ratio of the posteriors in both areas. In SDZ, the area with the highest increasing ratio appeared in Tangshan city, followed by Beijing and Tianjin, and Chengde city had the lowest increasing ratio. In WLG, the area near Xining city had an obvious increase in the posteriors, followed by the residential and surrounding areas of Linxia. In contrast, Hainan, Haidong and Haibei had much lower increases. Regardless of SDZ or WLG, most increases of the posteriors were concentrated in residential and nearby areas, where most anthropogenic activities are concentrated. Examining the increasing ratio, the performance of the two areas was also different. In SDZ, there was no obvious trend in the spatial distribution of the increase in growth. In WLG, although the regions with large increases were located in the urban and suburb areas, the growth rate of nonurban areas was significantly higher than that of urban areas. The main reason for this was that the priors of nonurban areas in WLG were very low, which was also the reason why the posterior growth ratio of WLG was significantly higher than that of SDZ.

Conclusions

The priors in SDZ and WLG were provided with a monthly resolution, which greatly hampered our ability to describe the diurnal variation in CO, thus creating significant uncertainties in CO concentration forecasting. In this paper, the priors were optimized before each forecast cycle by the CO-FCST system, which could greatly improve the forecast accuracy. However, the forecast accuracy Overall, this POD method can effectively eliminate the uncertainty in the priors, thus improving the accuracy for CO concentration forecasting. First, the experimental results in this paper showed that the priors were the key factor for CO concentration forecasting; however, it was difficult to obtain priors with a high resolution without uncertainties in a specific study area; therefore, this is still one of the most effective forecasting methods for CO based on currently available priors with a coarse resolution. Second, CO observation information could be effectively absorbed by the POD4DVar method, and uncertainties in the priors could be significantly eliminated after optimization by the CO-FCST system; ultimately, the accuracy of CO forecasts was largely improved in both study areas. However, hampered by the uncertainties in the priors, the forecasting accuracy in both areas was still unsatisfactory. Finally, the forecasting accuracy varied greatly in different seasons, different from common recognition. The CO forecasting accuracy was higher in winter than in other seasons in this paper, and the CO-FCST system could grasp the main trend of the CO variation. During the period of low CO concentrations in spring and summer, the CO forecasting accuracy was relatively low. At the same time, we could see that it was extremely difficult to forecast sharp CO variations, which is also one of the thorniest problems in forecasting. Due to the significant gap between the mathematical algorithm and the realistic situations, forecasts by the CMAQ model generally had gentle variations, which was the main reason for the significant uncertainties in the CO forecasts. To obtain high-accuracy forecasts in all periods, deeper research should be conducted in the future. CO is mainly generated due to insufficient combustion of fossil fuels, e.g., oil and coal. In northern China, e.g., SDZ and WLG, the main fuel for winter heating is coal. Therefore, the most effective way for reducing CO emissions is to replace coal with other fuels for heating, e.g., natural gas, which will greatly reduce the CO emissions. In recent years, the Chinese government has vigorously implemented the “coal to gas” project in northern China, which will greatly change the situation of the air quality in the future.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2202_15091522

Acknowledgements. The work was supported by the Talent Project of Anhui Science and Technology University (BSWD202102) and Research Project of Anhui Provincial Department of Education (2023AH051878, 2023AH040280).

REFERENCES

[1] Siade, A. J., Putti, M., Yeh, W. W.-G. (2010): Snapshot selection for groundwater model reduction using proper orthogonal decomposition. – Water Resources Research 46(8). https://doi.org/10.1029/2009WR008792

[2] Aziz, A., Bajwa, I. U. (2007): Minimizing human health effects of urban air pollution through quantification and control of motor vehicular carbon monoxide (CO) in Lahore. – Environmental Monitoring and Assessment 135(1): 459-464. DOI: 10.1007/s10661-007-9665-7

[3] Ma, C., Wang, T., Mizzi, A. P., Anderson, J. L., Zhuang, B., Xie, M., Wu, R. (2019): Multiconstituent data assimilation with WRF-Chem/DART: Potential for adjusting anthropogenic emissions and improving air quality forecasts over eastern China. – Journal of Geophysical Research: Atmospheres 124(13): 7393-7412. https://doi.org/10.1029/2019JD030421

[4] Cheewaphongphan, P., Chatani, S., Saigusa, N. (2019): Exploring gaps between bottom-up and top-down emission estimates based on uncertainties in multiple emission inventories: a case study on CH_4 emissions in China. – Sustainability 11(7): 2054. DOI: 10.3390/su11072054

[5] Van Vuuren, D. P., Hoogwijk, M., Barker, T., Riahi, K., Boeters, S., Chateau, J., Scrieciu, S:, Van Vliet, J., Masui, T., Blok, K., Blomen, E., Kram, T. (2009): Comparison of top-down and bottom-up estimates of sectoral and regional greenhouse gas emission reduction potentials. – Energy Policy 37(12): 5125-5139. https://doi.org/10.1016/j.enpol.2009.07.024

[6] Flerlage, H., Velders, G. J. M., de Boer, J. (2021): A review of bottom-up and top-down emission estimates of hydrofluorocarbons (HFCs) in different parts of the world. – Chemosphere 283: 131208. https://doi.org/10.1016/j.chemosphere.2021.131208

[7] Resler, J., Eben, K., Juruš, P., Krč, P. (2007): Inverse modeling of emissions using the CMAQ adjoint model. – Atmos. Chem. Phys 7: 3749-3769.

[8] Kalogridis, A. C., Vratolis, S., Liakakou, E., Gerasopoulos, E., Mihalopoulos, N., Eleftheriadis, K. (2017): Assessment of wood burning versus fossil fuel contribution to wintertime black carbon and carbon monoxide concentrations in Athens, Greece. – Atmospheric Chemistry & Physics 1-20. DOI: 10.5194/acp-2017-854.

[9] Yi, L., Zhang, W., Wang, K. (2018): Evaluation of heavy precipitation simulated by the WRF model using 4D-Var data assimilation with TRMM 3B42 and GPM IMERG over the Huaihe River Basin, China. – Remote Sensing 10(4): 646. https://doi.org/10.3390/rs10040646

[10] Li, M., Liu, H., Geng, G., Hong, C., Liu, F., Song, Y., Tong, D., Zheng, B., Cui, H., Man, H., Zhang, Q., He, K. (2017): Anthropogenic emission inventories in China: a review. – National Science Review 4(6): 834-866. https://doi.org/10.1093/nsr/nwx150

[11] Lopez-Restrepo, S., Yarce, A., Pinel, N., Quintero, O. L., Segers, A., Heemink, A. W. (2020): Forecasting PM10 and PM2.5 in the Aburrá Valley (Medellín, Colombia) via EnKF based data assimilation. – Atmospheric Environment 232: 117507. https://doi.org/10.1016/j.atmosenv.2020.117507

[12] Andronopoulos, S., Sfetsos, A., Vlachogiannis, D., Yiotis, A., Gounaris, N. (2011): Application of adjoint CMAQ chemical transport model in the Athens greater area: sensitivities study on ozone concentrations. – International Journal of Environment and Pollution 47(1): 193-206. https://doi.org/10.1504/IJEP.2011.047352

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2202_15091522

[13] Turnbull, J. C., Tans, P. P., Lehman, S. J., Baker, D., Conway, T. J., Chung, Y. S., Gregg, J., Miller, J. B., Southon, J. R., Zhou, L.-X. (2011): Atmospheric observations of carbon monoxide and fossil fuel CO_2 emissions from East Asia. – Journal of Geophysical Research Atmospheres 116(D24): D24306. DOI: 10.1029/2011JD016691

[14] Vaughn, T. L., Bell, C. S., Pickering, C. K., Nummedal, D. (2018): Temporal variability largely explains top-down/bottom-up difference in methane emission estimates from a natural gas production region. – Proceedings of the National Academy of Sciences 115(46): 11712-11717. https://doi.org/10.1073/pnas.1805687115

[15] Chen, T. M., Kuschner, W. G., Gokhale, J., Shofer, S. (2007): Outdoor air pollution: nitrogen dioxide, sulfur dioxide, and carbon monoxide health effects. – The American Journal of the Medical Sciences 333(4): 249-256. https://doi.org/10.1097/MAJ.0b013e31803b900f

[16] Tian, X., Feng, X. (2015): A non-linear least squares enhanced POD-4DVar algorithm for data assimilation. – Tellus A: Dynamic Meteorology and Oceanography 67(1): 25340.

[17] Jin, X., Kumar, L., Li, Z., Feng, H., Xu, X., Yang, G., Wang, J. (2018): A review of data assimilation of remote sensing and crop models. – European Journal of Agronomy 92: 141-152. https://doi.org/10.1016/j.eja.2017.11.002

[18] Cao, Y., Zhu, J., Navon, I. M., Luo, Z. (2007): A reduced-order approach to four-dimensional variational data assimilation using proper orthogonal decomposition. – International Journal for Numerical Methods in Fluids 53(10): 1571-1583. https://doi.org/10.1002/fld.1365

[19] He, Z., Yang, D., Wang, Y., Jin, B. (2022): Impact of 4D-Var data assimilation on modelling of the East China Sea dynamics. – Ocean Modelling 176: 102044. https://doi.org/10.1016/j.ocemod.2022.102044

[20] Zhao, Y., Nielsen, C. P., Lei, Y., McElroy, M. B., Hao, J. (2011): Quantifying the uncertainties of a bottom-up emission inventory of anthropogenic atmospheric pollutants in China. – Atmospheric Chemistry and Physics 11(5): 2295-2308. DOI: 10.5194/acp-11-2295-2011.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2202_15091522