

A correlation-based selection of VOCs for emission estimates in Yangtze River Delta

Haoxing Ju¹ Lili Lei¹ Zhen Peng¹

¹School of Atmospheric Sciences, Nanjing University, Nanjing, China

Introduction

Volatile organic compounds (VOCs) have been listed by China's Ministry of Ecology and Environment (MEE) as key pollutants to be controlled in the series plans since 2013, for they play important roles in the formation of other pollutants, such as secondary PM_{2.5} and troposphere O₃. In most Chemical Transport Model (CTM), errors in emissions could have even larger impacts on forecast than initial conditions(IC)[Sandu and Chai 2011], and the uncertainty of VOC emissions can reach up to 78% [Li et al. 2017]. Thus, it can be expected that a better estimation of VOC emissions could help to improve CTM forecast results.

Data assimilation tries to get better estimates based on prior information and observations. Ensemble-based filter is one of the main approaches in atmospheric data assimilation, and many efforts have been devoted to seeking for better VOC estimates using ensemble-based filter(for example, [Tang, Zhu, Z. F. Wang, et al. 2011]; [Ma et al. 2019]; [Xing et al. 2020]). For lacking direct observations of VOC, the common approach is using other correlated observations (O₃ in most of studies) to update ICs/emissions of all VOC species via cross-variable assimilation. This kind of approach is convenient to realize, but the neglection of differences among VOC species could bring in extra errors. Also, this approach can hardly include other observations which show no obvious linear relations with overall VOC, since conducting cross-variable assimilation between variables with strong nonlinear relations could have negative effects on forecasts. (see, for example, [Tang, Zhu, Z. Wang, et al. 2016])

This work tries to make a quantified selection for VOC species to be updated in assimilation system based on correlations/potential influence. The 2022 summer heat wave event in the Yangtze River Delta(YRD) region is selected to conduct real case experiment. This selection will respect the different chemical properties of VOC species, and can include more 'not so correlated' observations into assimilation system. The estimated emissions have been evaluated by launching series of 72h forecasts using WRF-Chem V3.6.1.

Methodology for emission estimation

To get estimated emissions, the scaling factor λ is extended as part of WRF-Chem model state variables. The scaling factor is integrated using ensemble forecasts of chemical concentration produced by WRF-Chem. At each assimilation window, an EnSRF is performed to get scaling factor analysis λ^a , the estimated emission can then be calculated by multiplying prescribed emission E^p with λ^a . The flow chat below describes this framework.

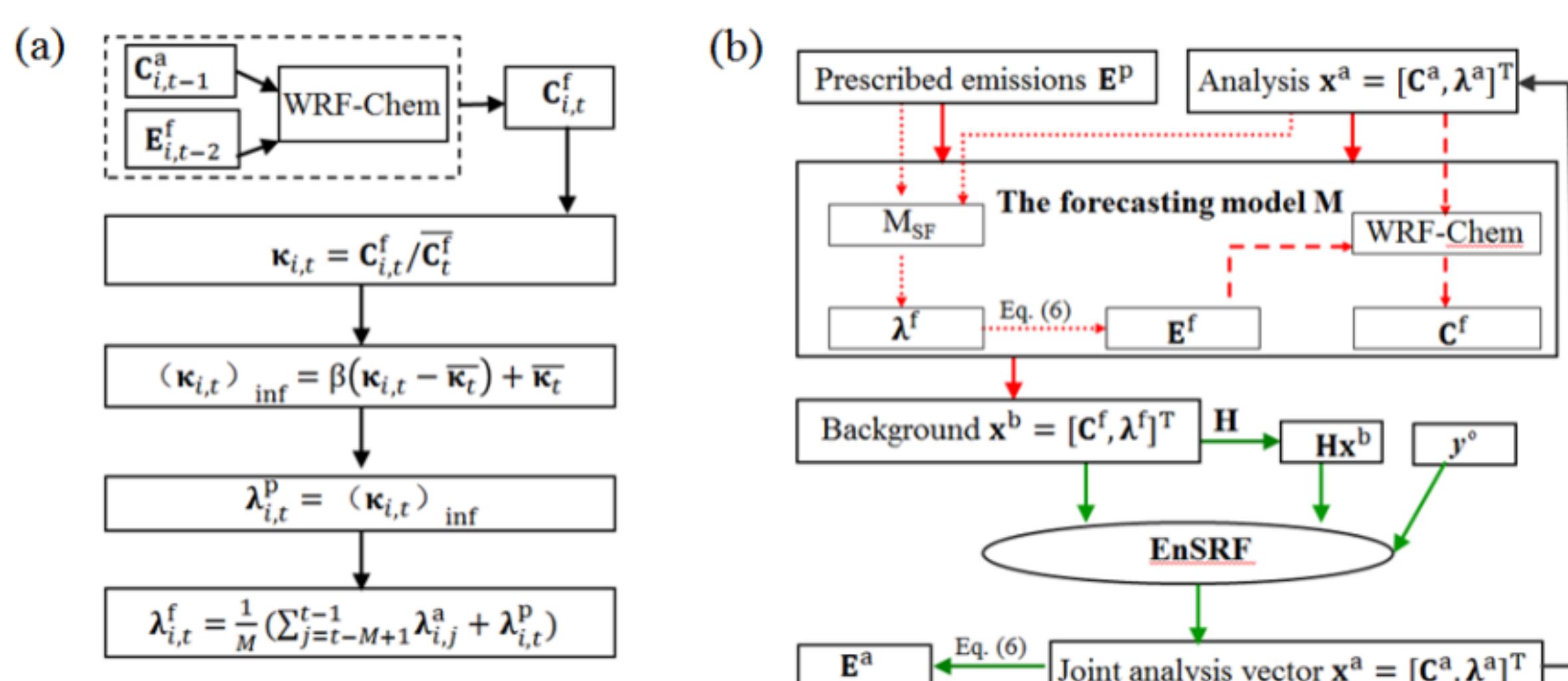


Figure 1. (a) describe the framework to integrate scaling factor λ . (b) flow chart of the data assimilation system that simultaneously optimizes the chemical initial conditions and emissions.

System settings and Observation

- simulation model:** WRF-Chem version 3.6.1. The model domain, together with Yangtze River Delta(YRD) region are shown in fig. 2, with a horizontal resolution of 18km. GOCART-RACM-KPP is selected as the aerosol-chemical scheme.
- assimilation system:** Ensemble Square Root Filter(EnSRF), with forward operator taken from Gridpoint Statistical Interpolation(GSI).
- meteorology observation:** NCEP GDAS. All in situ observations and cloud motion vectors are assimilated every 6 hours.
- surface chemical observation:** MEE in-situ observation of PM_{2.5}, PM₁₀, SO₂, NO₂, O₃, CO. A subset of 560 stations is selected to be assimilated, in order to avoid correlated observations. There are 95 validation station in YRD region. Details of station distribution can be seen in fig. 2.

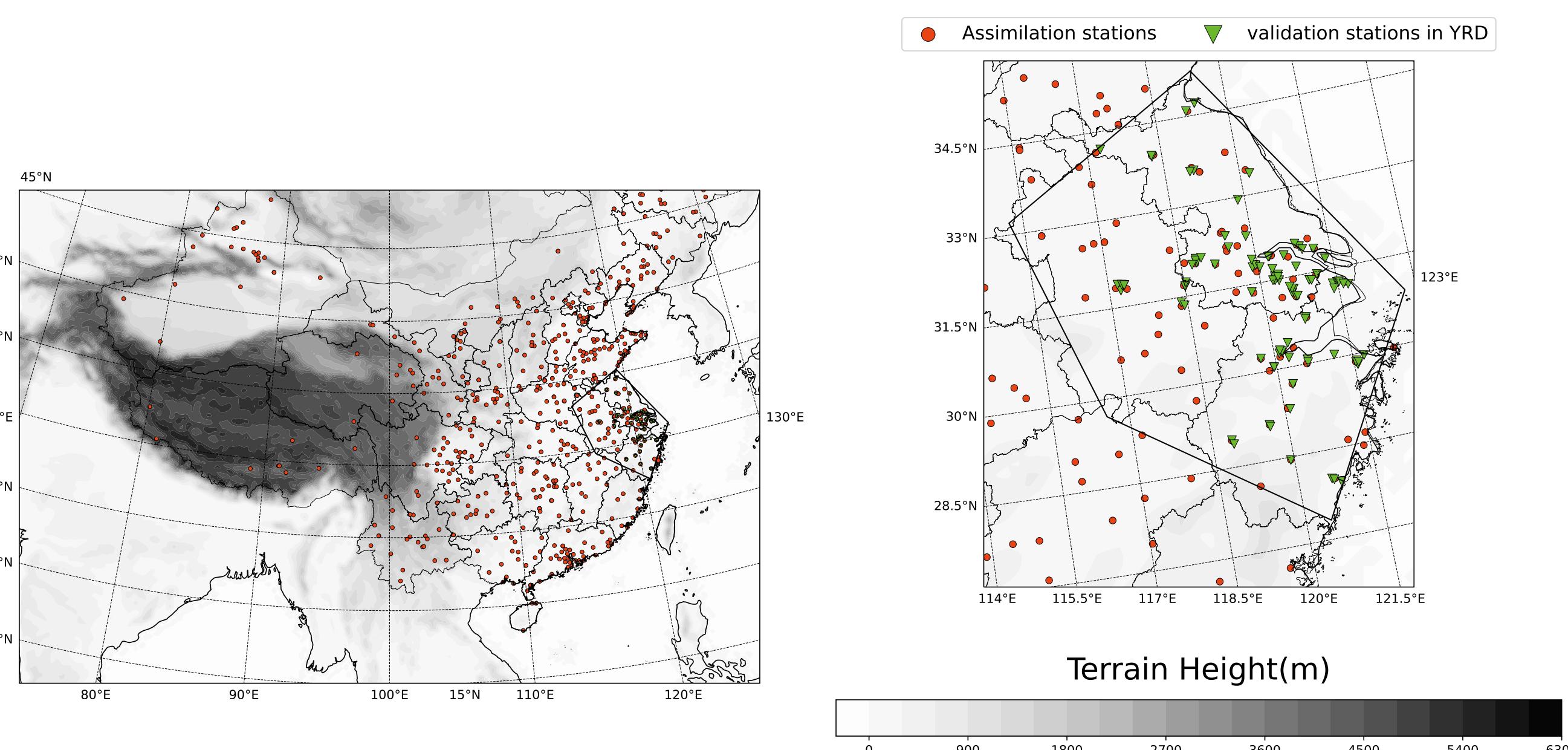


Figure 2. Domain setting and station distribution. The left subplot shows model domain in WRF-Chem, the shading indicates the terrain height. Red dots represent assimilation stations, and green triangles are validation stations in Yangtze River Delta(YRD) region selected by the black polygon. The right subplot shows more detailed station distribution in YRD region.

Quantified selection for VOC species

A quantified standard will make it clear to decide which VOC species should be updated by observations. Here, we propose a VOC factor to quantify which VOC species should be included in the assimilation system. The VOC factor considers two parts:

- The correlation between observation and VOC species.
- The potential influence of updating VOC species on other variables.

A 50-member ensemble of free forecasts without any assimilation is conducted from 2022-07-10:00 to 2022-08-10:00 to calculate the VOC factor(also provide spin-up IC for cycling assimilation). The potential influence of updating a VOC species is roughly quantified by prescribed emission and average concentration during free forecast period. More details can be found by scan the QR code in Appendix.

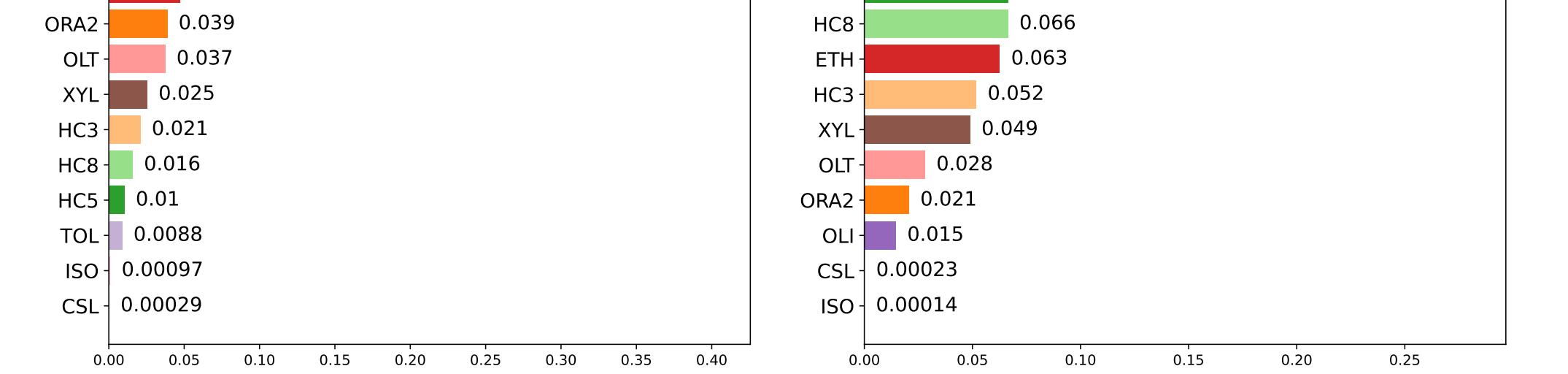


Figure 3. Normalized VOC factor for each observation kind. VOCs with a red name will be included in our assimilation system.

Experiment design

Four experiments are designed to explore the effects of VOC selection for emission estimates: NOVOC, O₃_AllVOC, O₃_CSVOC, AllOBS_CSVOC.

A cycling assimilation system is launched for each experiment from 2022-07-20:00 to 2022-07-30:00. The assimilation cycle period is 6h for meteorology variables and 1h for chemical variables/scaling factors. The meteorology variables and six main pollutions, together with their corresponding scaling factors, will be updated in all experiments, which is the same in [Peng et al. 2020]. The difference between four experiments is the strategy to update VOC species, which is shown in the table below.

Experiment	Observation	VOC species to be updated
NOVOC	N/A	N/A
O ₃ _AllVOC	O ₃	All VOC species
O ₃ _CSVOC	O ₃	KET/HCHO/OLI
AllOBS_CSVOC	O ₃	KET/HCHO/OLI
	PM _{2.5}	KET/HCHO/OLI
	NO ₂	KET/HCHO/TOL/ALD/HC8/HC5
	CO	KET/HCHO/TOL/ALD

Table 1. Description of VOC selections for each experiment. The VOC species and corresponding scaling factors in the 3rd column will be updated by observations in the 2nd column.

Since there is no direct emission observations, the estimated emissions will be evaluated by launching series of 72h forecasts. Starting at 2022-07-22:00, every 6h a deterministic forecast is launched with ensemble mean of cycling analysis as IC and emission estimates of last day as emission input.

It should be noted that all kinds of chemical observations have already been used to update their corresponding state variables, thus no new observation is introduced into assimilation system to update VOC species.

Emission estimates

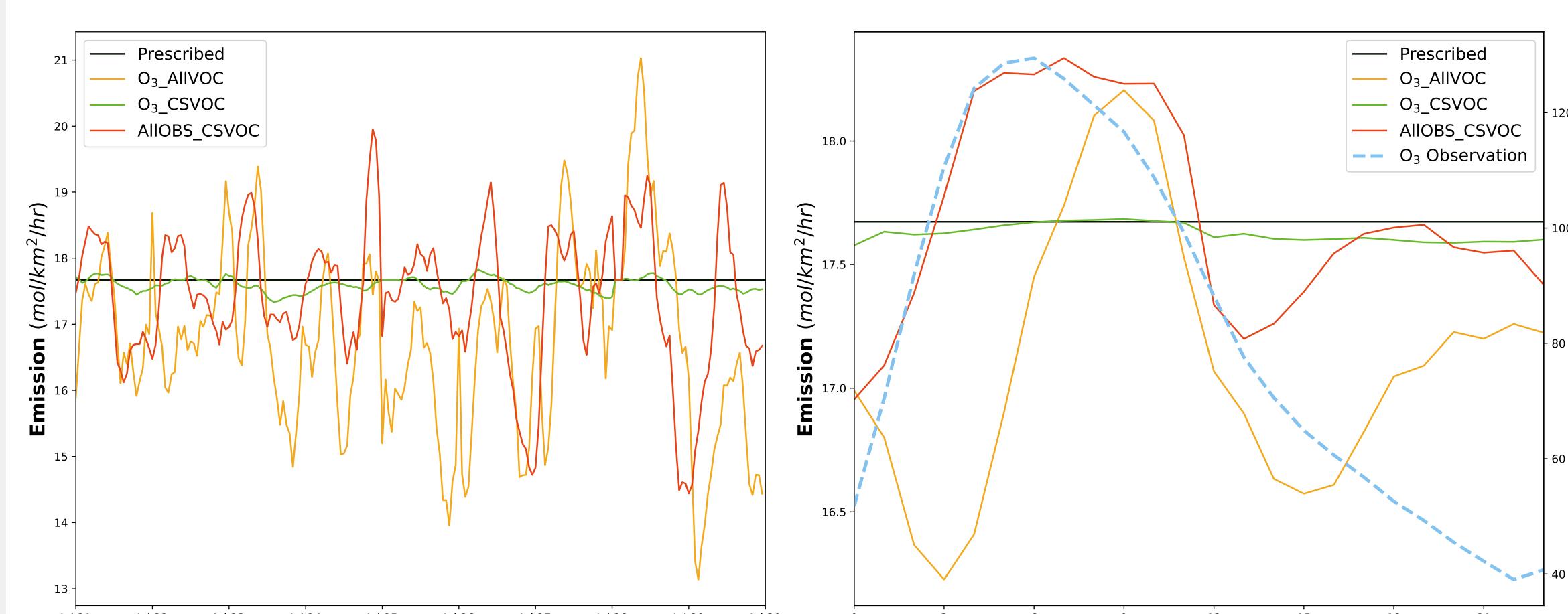


Figure 4. The left subplot shows series of overall VOC emission estimates in each experiment. The right subplot presents diurnal-variation of overall VOC emission estimates in left y-axis, and diurnal-variation of O₃ observation in right y-axis.

- O₃_AllVOC estimates have an obvious diurnal variation, but show a suspicious drop in UTC00-UTC03.
- O₃_CSVOC estimates fix the suspicious drop in UTC00-UTC03, but present little diurnal variation. This is because only 3 VOC emissions are updated in this experiment, others remain unchanged compared to the prescribed emissions.
- AllOBS CSVOC estimates also show an obvious diurnal variation, with a constant peak in daytime. This variation in daytime is similar to the YRD averaged O₃ observation. As important precursors of O₃, this peak in daytime could help the model to overcome the underestimation of O₃, as show in fig. 5.

72H O₃/PM_{2.5} forecast results

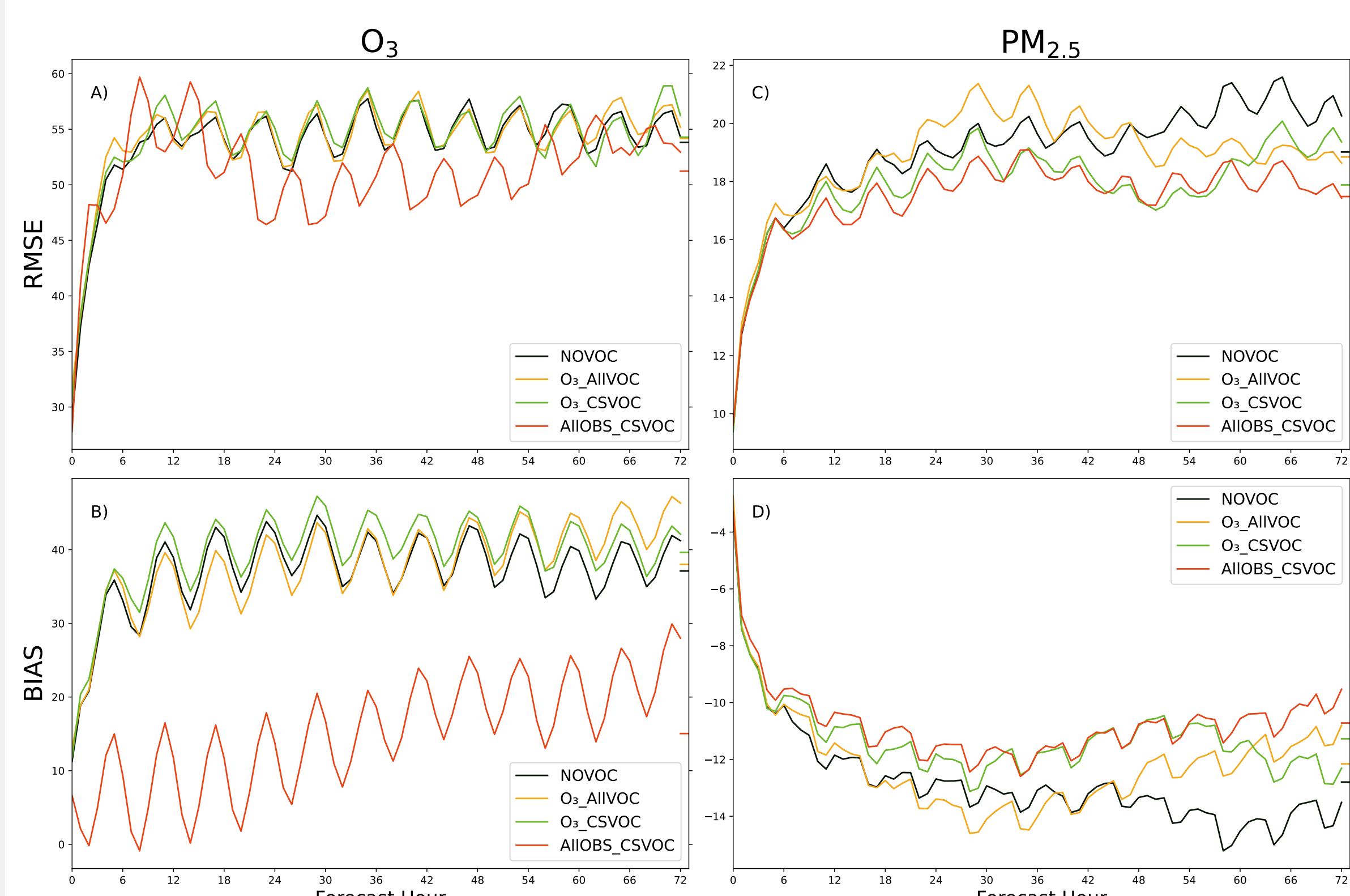


Figure 5. Averaged RMSE/BIAS value for O₃(1st column) and PM_{2.5}(2nd column). The bars on the rightside in each subplot represent the mean value . Units for all this variables: $\mu\text{g}/\text{m}^3$

- O₃_AllVOC(yellow) shows little RMSE/BIAS improvements compared with NOVOC(black).
- O₃_CSVOC(green) has obvious improvements in PM_{2.5} RMSE/BIAS, but failed to improve O₃ forecasts.
- AllOBS CSVOC shows advantage for both O₃ and PM_{2.5} forecasts, especially in reducing O₃ BIAS. Possibly because AllOBS CSVOC estimates have a constant peak in daytime, which compensates for the underestimation of O₃ and serves as a model error correction.