



A correlation-based selection of VOCs for emission estimates

Haoxing Ju¹ Lili Lei¹ Zhen Peng¹

¹Nanjing University



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 secondary PM_{2.5} and troposphere ozone. 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 VOCs emission can reach up to 78% [Li et al. 2017]. Thus, it can be expected that a better inversion of VOC emissions could have great impacts on PM_{2.5} and ozone forecasts.

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 estimate 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(usually ozone) to update all VOCs IC/emission via cross-variable assimilation. This approach is convenient to realize, but ignore the differences among VOC species and can thus bring in extra errors in some cases. Also, this approach can hardly use other observations which show no obvious correlation with all VOCs, since conducting cross-variable assimilation between variables with nonlinear relations can be dangerous and can even make the assimilation system diverge (see, for example, [Tang, Zhu, Z. Wang, et al. 2016])

We are trying to make a selection for VOC species to be updated based on correlations and potential influence. This selection will respect the different characters among VOCs, and can include more 'not so correlated' observations into assimilation system. The emission estimation will be tested by launching series of 72h forecasts using WRF-Chem V3.6.1.

Observation and model setting

- **meteorology observation:** NCEP GDAS. All in situ observations and cloud motion vectors are assimilated every 6 hours.
- **surface chemical observation:** MEE in-situ observation for $PM_{2.5}$, PM_{10} , SO_2 , NO_2 , O_3 , CO . A subset of 560 stations is selected in order to avoid observation correlation. Chemical observation will be assimilated hourly.
- **simulation model:** WRF-Chem version 3.6.1, model domain refer to fig. 1. 50-member ensemble cycling assimilation starts at 2022-07-20:00 and ends at 2022-07-30:00.
- **assimilation system:** Ensemble Square Root Filter(EnSRF). GSI observation part as forward operator.

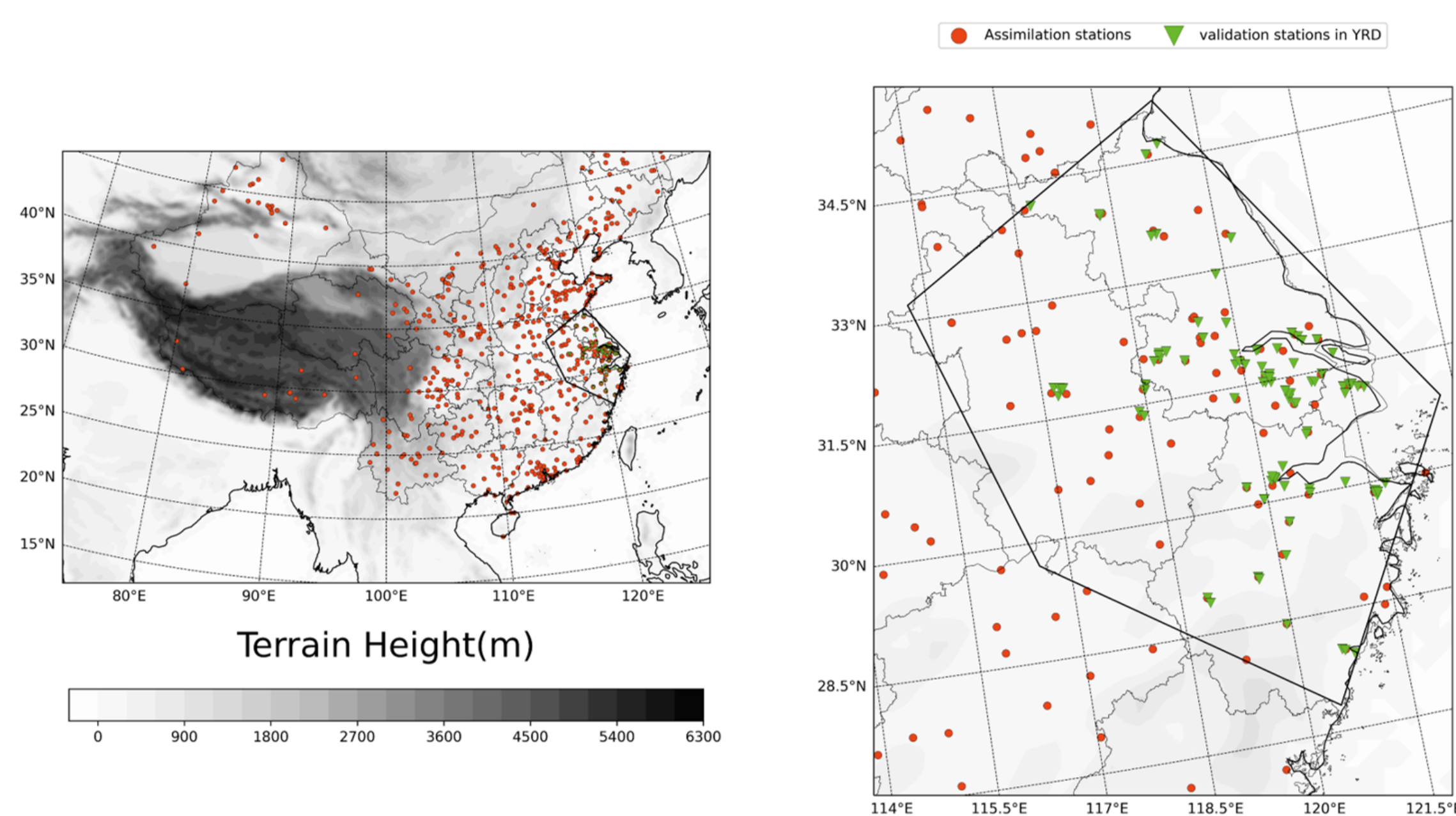


Figure 1. domain setting and station distribution. The left subplot shows model domain in WRF-Chem, shading shows the terrain height. Red dots represent assimilation stations, and green triangles are validation stations in Yangtze River Delta(YRD) region selected by the black polygon. The right subplot shows more detailed station distribution in YRD region.

Methodology for emission inversion

In order to get emission inversion, we extend emission scaling factors λ as part of WRF-Chem model state variables, and it will be integrated by using the ensemble forecast chemical concentration fields produced by WRF-Chem. The flow chat below describes this framework.

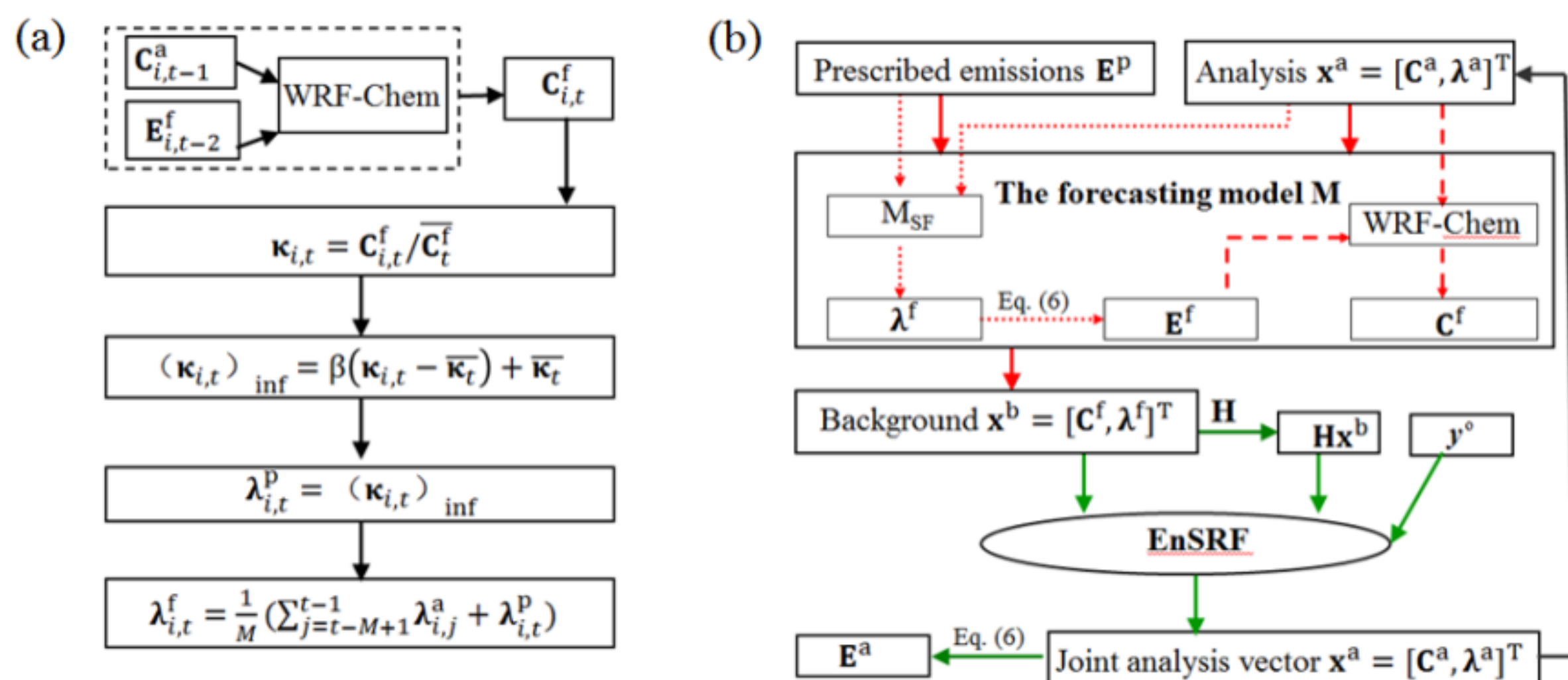


Figure 2. (a) describe the framework to integrate scaling factor λ . (b) flow chart of the data assimilation system that simultaneously optimizes the chemical initial conditions and emissions.

VOC selection

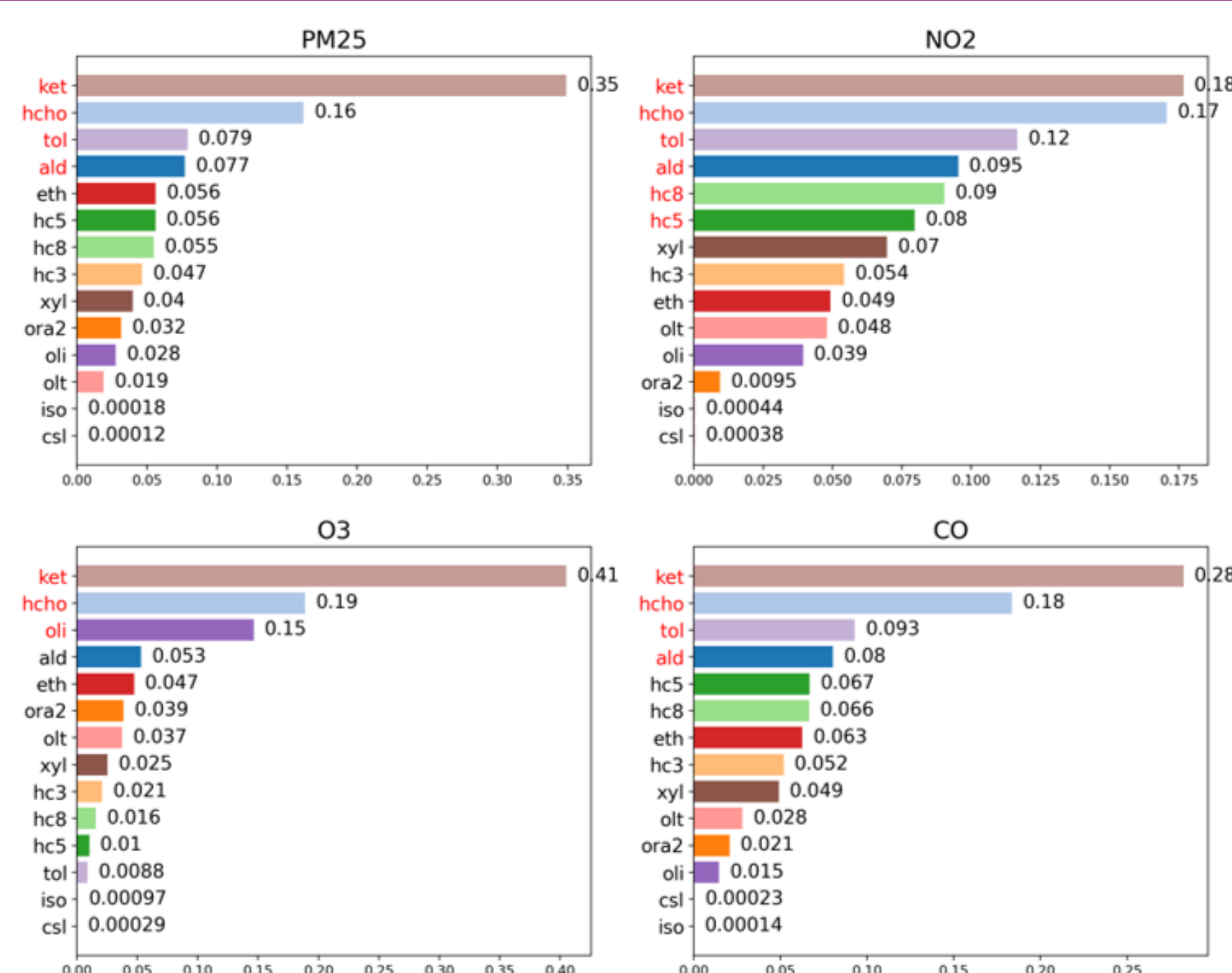


Figure 3. Normalized VOC factor for each observation kind. VOCs with a red name will be included in our assimilation system. A quantified standard will make it clear to decide which VOC species should be assimilated. Here, we propose a VOC factor to quantify which VOC species should be included in the assimilation system. The VOC factor considers two parts:

- 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 these factors(and also provide spin-up IC for cycling assimilation). Details about the calculation can be seen at the QR code in Appendix.

Experiment design

We designed four experiments to explore the effects of VOCs selection in assimilation: no_voc, o3_allvoc, o3_5pctvoc, allob_avgvoc. The meteorology variables and six main pollutions(including NOx), together with the emission scaling factors, will be updated by their corresponding observations in all experiments, which is the same in [Peng et al. 2020]. The only difference between the four experiments is the strategy to assimilate VOC species, which is shown in the table below.

	no_voc	o3_allvoc	o3_5pctvoc	allob_avgvoc
ket	x	o3	o3	pm25/no2/co/o3
hcho	x	o3	o3	pm25/no2/co/o3
tol	x	o3	x	pm25/no2/co
oli	x	o3	o3	o3
ald	x	o3	o3	pm25/no2/co
hc8	x	o3	x	no2
hc5	x	o3	x	no2
xyl	x	o3	x	x
eth	x	o3	x	x
hc3	x	o3	x	x
olt	x	o3	x	x
ora2	x	o3	x	x
iso	x	o3	x	x
csl	x	o3	x	x

Table 1. VOC selection table. The columns are experiments, the rows are 14 RACM VOC species. If the value is "x", this VOC will not be updated by any observation in the column experiment. If the value is observation type(s), this VOC will be updated by the given observation type(s) in the column experiment.

Starting on Jul 22nd, every UTC00, 06, 12, 18, a deterministic 72h forecast will be launched based on the ensemble mean of cycling analysis and the posterior emission to evaluate the assimilation system.

72H ozone/PM2.5 forecast results

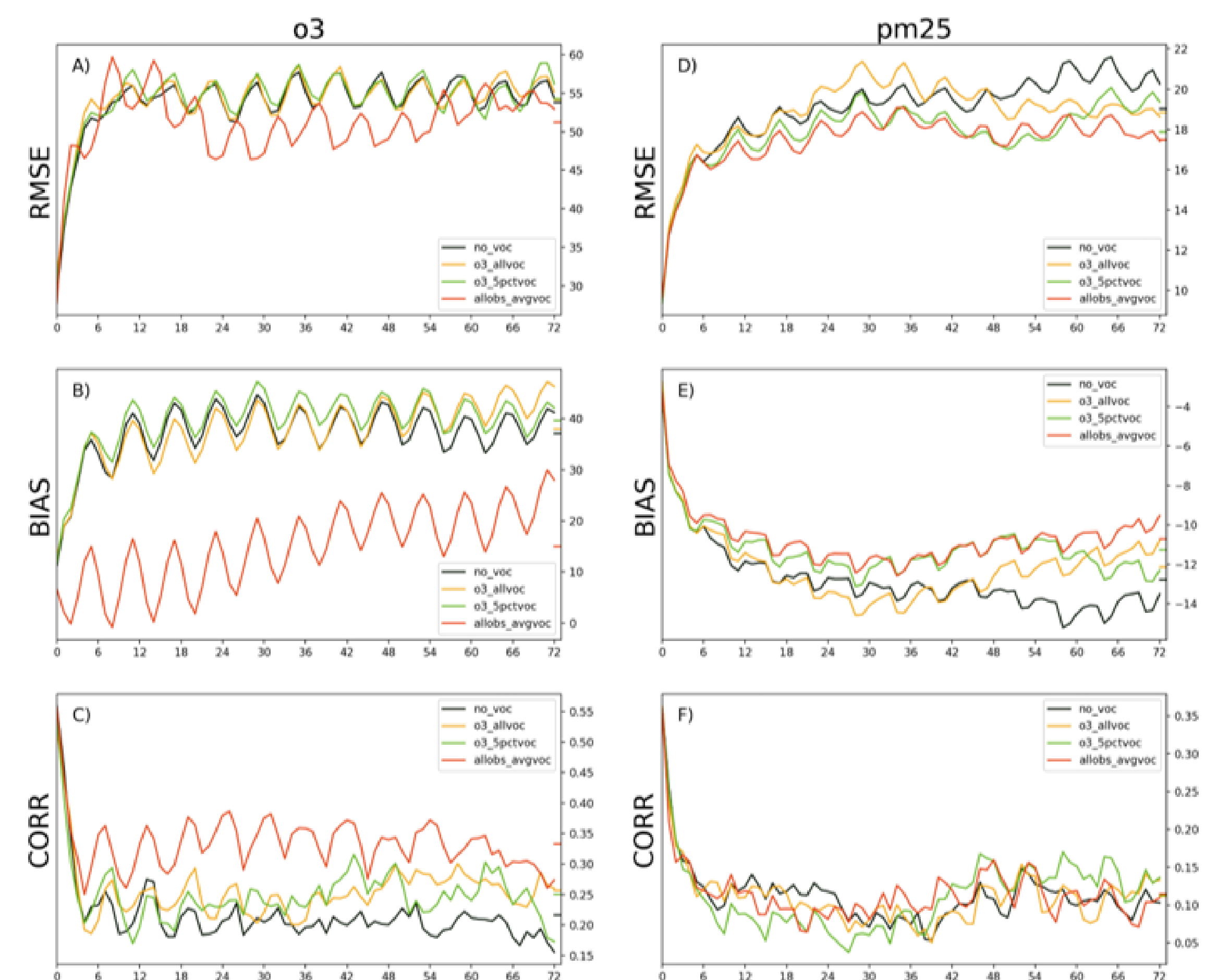


Figure 4. RMSE/BIAS/CORR values averaged over all sets of forecasts. Columns are o3 and PM2.5, and rows are corresponding RMSE/BIAS/CORR for o3 and PM2.5. The right bars in each subplot represent the averaged value of RMSE/BIAS/CORR.

- When using only ozone obs, a VOC selection results better RMSE and BIAS of PM2.5 forecasts.
- A VOC selection makes it possible to utilize 'uncorrelated' obs, and results better forecast performance for both PM2.5 and ozone.

estimated emission

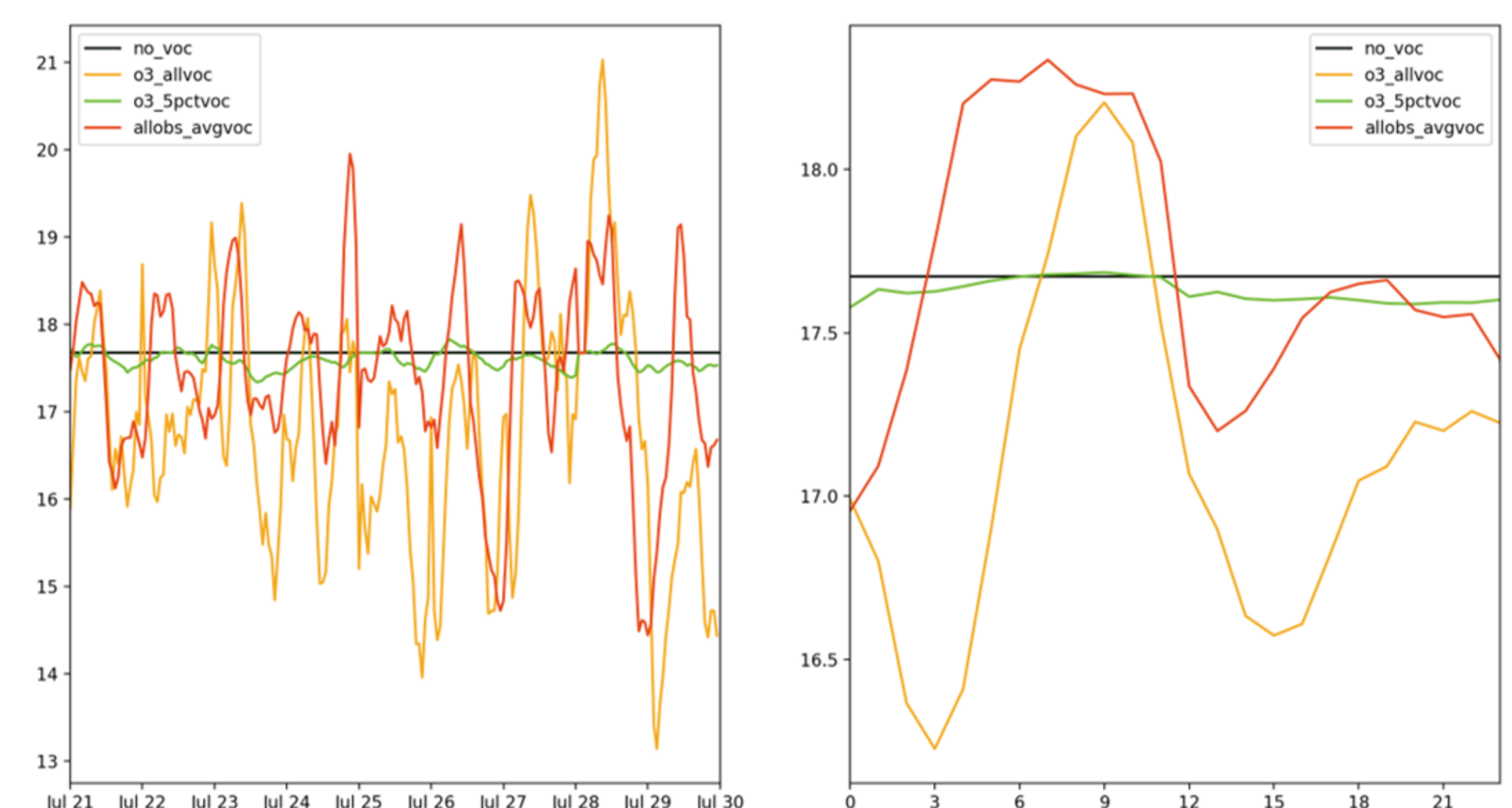


Figure 5. Time series of the ensemble mean of overall VOC emissions in each experiment. The left subplot shows the whole cycling period, and the right subplot shows the time series averaged every 24h.

- The "allob_avgvoc" emission shows an obvious daily change, with a peak during the daytime and a valley after sunset.
- The daily change of VOC emission after selected updates is consistent with Ozone's daily change, and the updated emission peak in daytime can help the model to overcome the underestimation of Ozone simulation shown in fig. 4