



# Improving Spatiotemporal Fine Particulate Matter from a Data Assimilation Approach

**Xuguo Zhang<sup>1</sup>, Jimmy Fung<sup>1</sup>, Alexis Lau<sup>1</sup>,  
Shaoqing Zhang<sup>2</sup>, and Wei Huang<sup>3</sup>**

<sup>1</sup>The Hong Kong University of Science and Technology, Hong Kong, China

<sup>2</sup>Qingdao Pilot National Laboratory for Marine Science and Technology, China

<sup>3</sup>Southern University of Science and Technology, China

Abstract



This presentation participates in OSPP



Outstanding Student & PhD  
candidate Presentation contest

## ■ Research Motivation



**Part II:** Assimilated Model  
& Scenario Setting

**Part III:** Major  
Findings

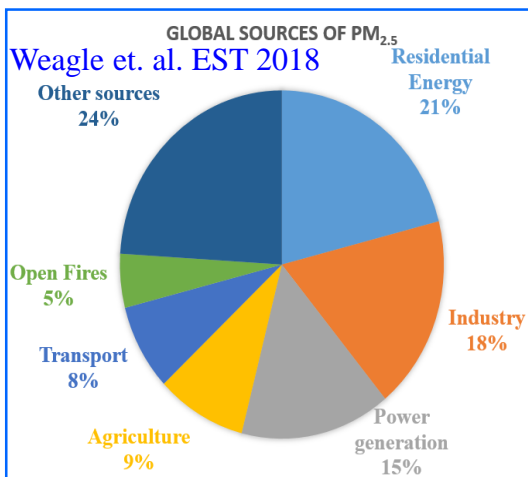
**Part I:** Research  
Framework

## ■ Conclusion

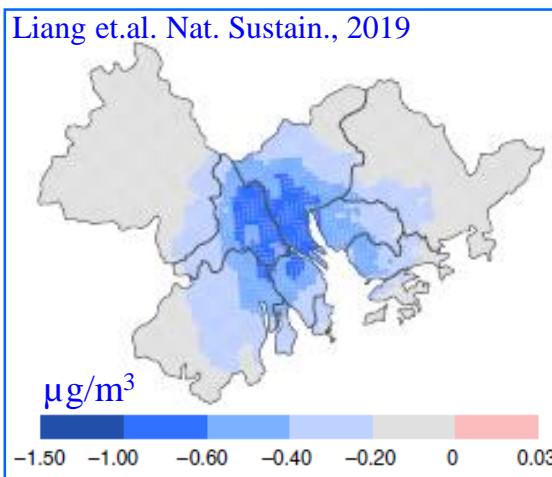


- Technical guide for **Source Apportionment** of ambient ozone pollution.
- Emission control policies in the **13<sup>th</sup> Five-Year-Plan** (2016-2020).
- Atmospheric pollution **prevention and control law** of China.
- Estimated **3 million deaths** in 2014 being attributable to PM<sub>2.5</sub>. (WHO report 2016)
- **87% of global people** reside in the areas in which the WHO air quality guideline (annual mean of 10 µg/m<sup>3</sup>) for PM<sub>2.5</sub> is exceeded. (Shaddick. et. al. JRSSAS-series C, 2018)
- Satellite retrieval or statistical method? (Bi et. al. RSE 2019)
- **Air Quality Modeling**

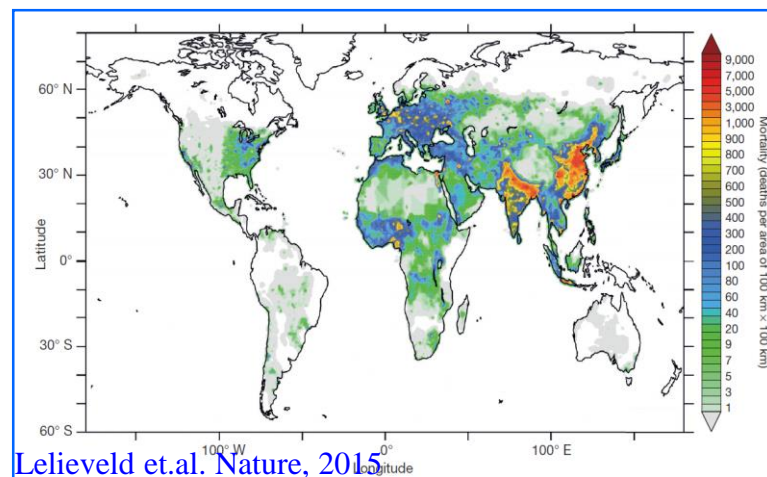
## Source Apportionment



## Policy Evaluation



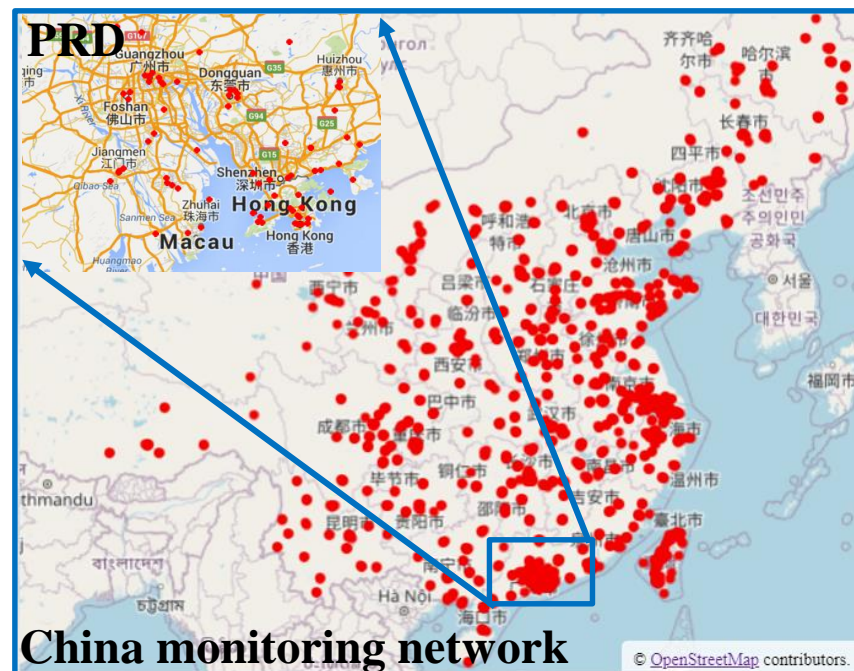
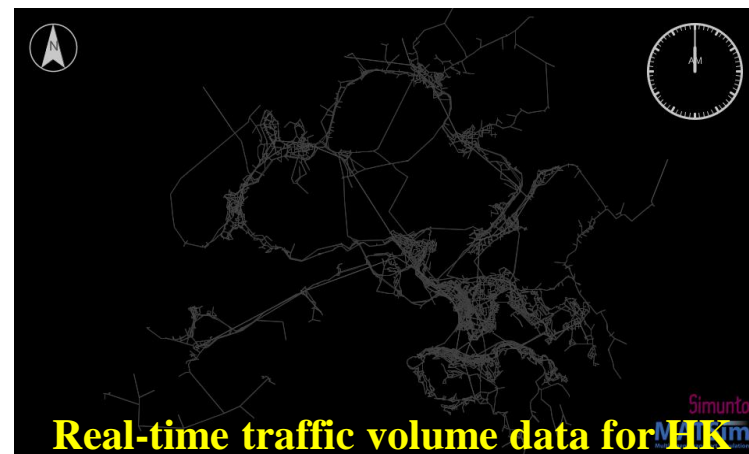
## Health Risks





- **Multi-source real-time data** offered an unprecedented opportunity to make the model validation and assimilation possible.

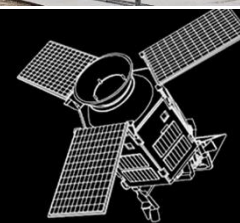
Work with Prof. Lo, H. K. in Civil, HKUST



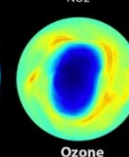
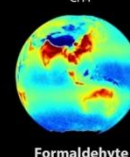
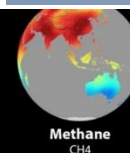
Super-site measurements

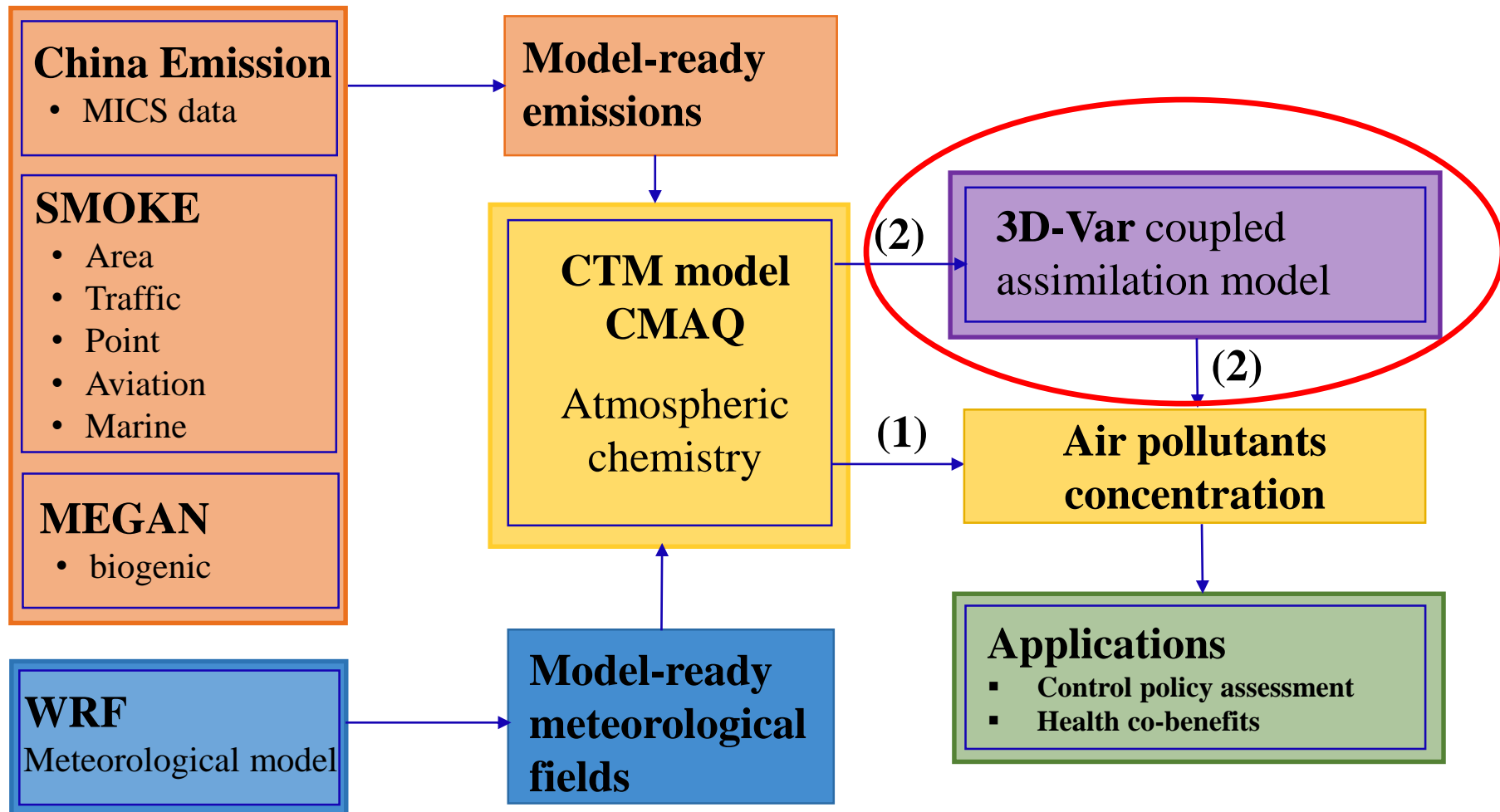


Air-born LIDAR



Sentinel-5 Precursor  
Air quality

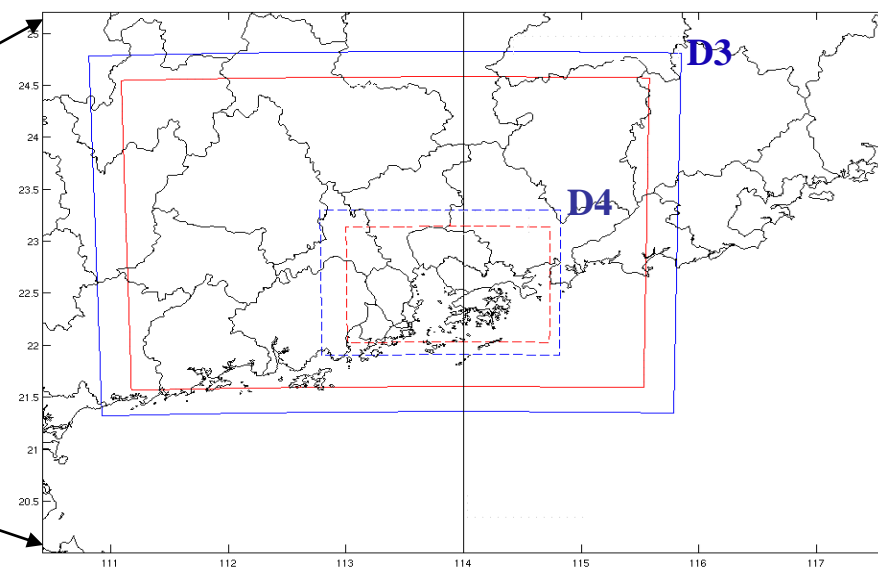
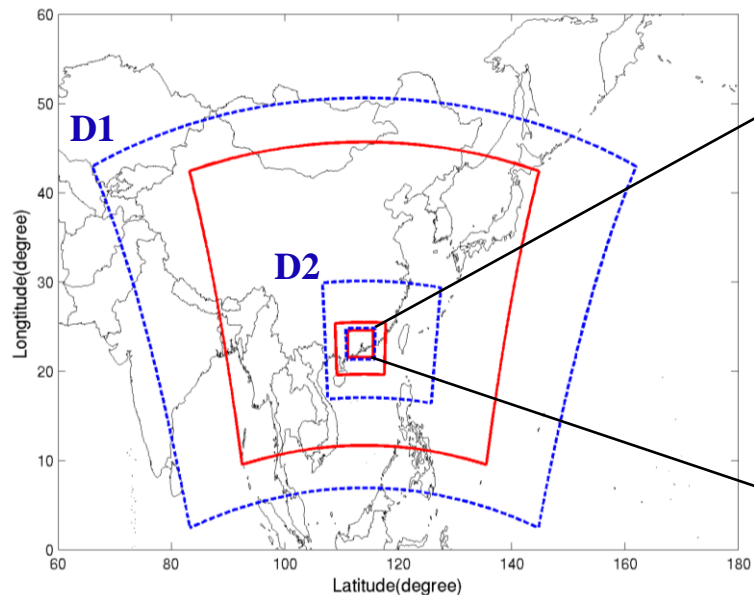








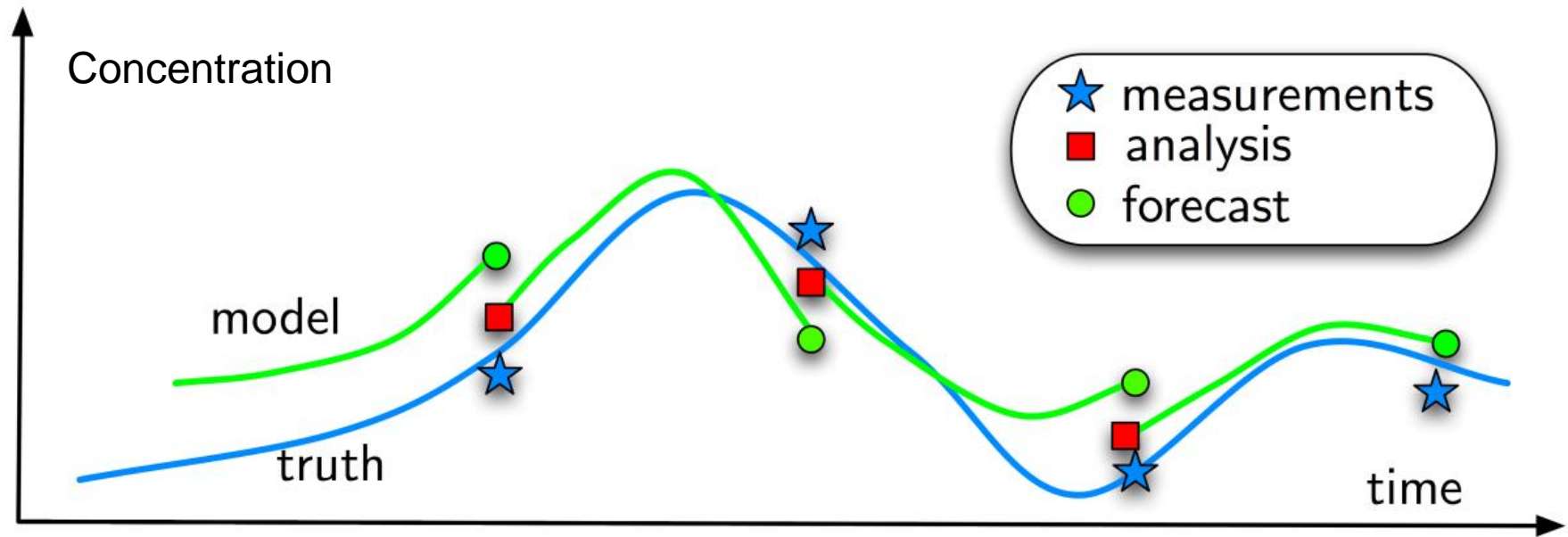
- Larger **WRF domain** could minimize the boundary effects of meteorological parameters on **CMAQ grid**.



Parameter	Value
Projection	Lamber-Conformal
Alpha	250°N
Beta	40°N
X center	114°E
Y center	28.5°N

Domain	Geographical coverage	WRF grid (km)	CMAQ grid (km)	Resolution (km)
D1	China and Japan	7641 x 4968	4914 x 3726	27
D2	Southeast China	2007 x 1467	882 x 666	9
D3	Guangdong	516 x 390	456 x 330	3
D4	Hong Kong	214 x 163	179 x 125	1

### Theory of the 3D-VAR model



$$J(\mathbf{x}_a) = \frac{1}{2}(\mathbf{x}_a - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}_a - \mathbf{x}_b) + \frac{1}{2}(\mathbf{y}_0 - \mathcal{H}(\mathbf{x}_a))^T \mathbf{R}^{-1}(\mathbf{y}_0 - \mathcal{H}(\mathbf{x}_a))$$

$\mathbf{x}_a$  is the vector of analysis;  $\mathbf{x}_b$  is the background;  $\mathbf{y}_0$  is the observation vector

$\mathbf{B}$  is the Background Error Covariance (BEC) matrix

$\mathbf{R}$  is the Observation Error Covariance (OEC) matrix

$\mathcal{H}$  is the observation operator



Scenarios	Assimilation cycles	Description
<b>NotDA</b>	No assimilation	Base control: spinning up for 10 days and digesting the initial conditions from the previous-cycle simulation.
<b>H24DAPM</b>	UTC 12:00	On the basis of the control case, assimilating PM <sub>2.5</sub> data once a day.
<b>H06DAPM</b>	UTC 12:00, 18:00, and 00:00, 06:00 the next day	On the basis of the control case, assimilating PM <sub>2.5</sub> data four times a day.
<b>H6DAPSN</b>	UTC 12:00, 18:00, and 00:00, 06:00 the next day	On the basis of the H06DAPM case, assimilating PM <sub>2.5</sub> , SO <sub>2</sub> , and NO <sub>2</sub> simultaneously.
<b>H6DAPSNO</b>	UTC 12:00, 18:00, and 00:00, 06:00 the next day	On the basis of the H06DAPM case, assimilating PM <sub>2.5</sub> , SO <sub>2</sub> , NO <sub>2</sub> , and O <sub>3</sub> simultaneously.
<b>H6DAALL</b>	UTC 12:00, 18:00, and 00:00, 06:00 the next day	On the basis of the H06DAPM case, assimilating PM <sub>2.5</sub> , SO <sub>2</sub> , NO <sub>2</sub> , O <sub>3</sub> , PM <sub>10</sub> and CO simultaneously.



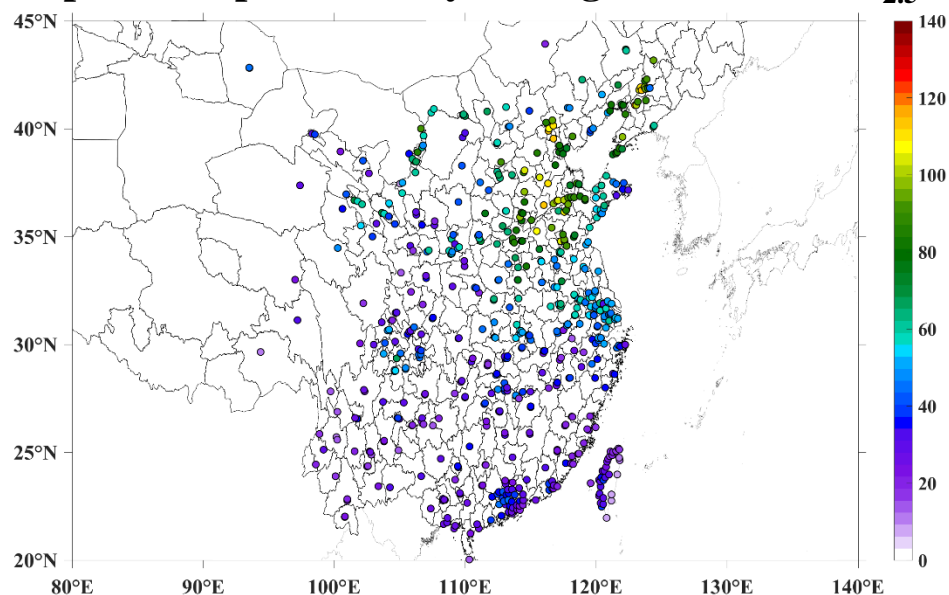


**China observation stations** (Observations of **584 stations** were assimilated after **cleaning >1600 stations** data.)

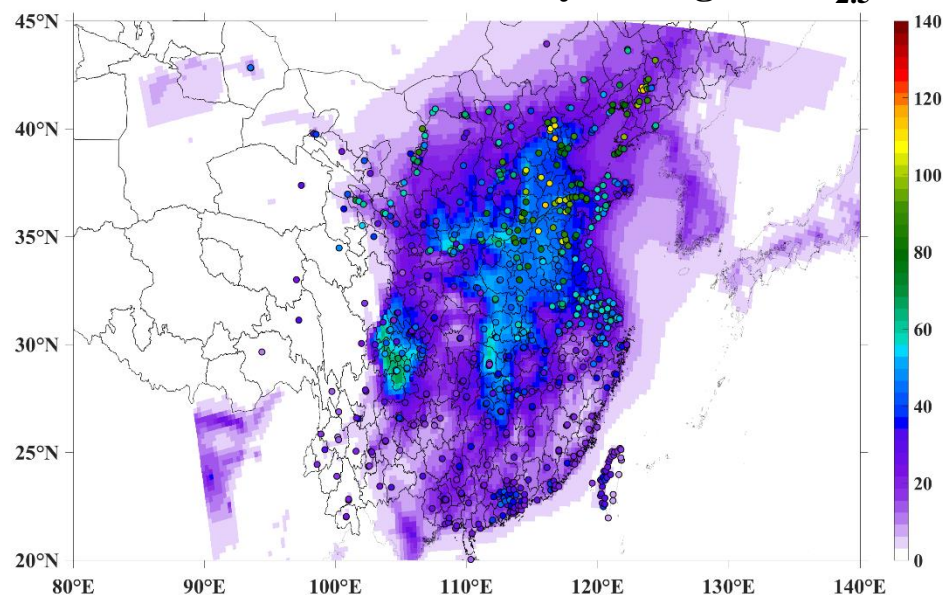
- Averaging the points within the same grid squares
- Setting concentration thresholds for different pollutants
- Checking data continuity

**Model simulation on top of the observations.**

**Spatial map of monthly averaged OBS for PM<sub>2.5</sub>**



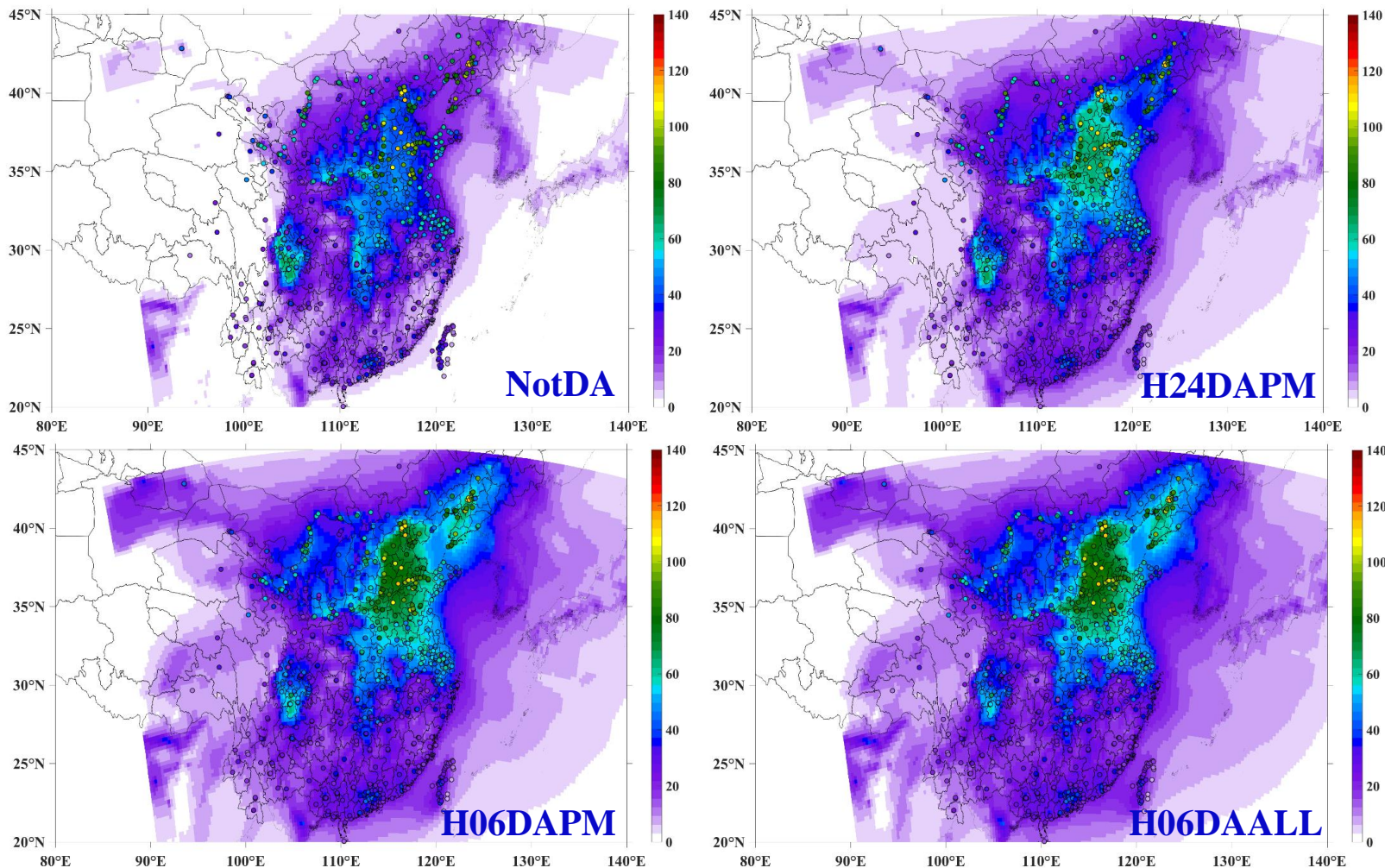
**Model simulated monthly averaged PM<sub>2.5</sub>**





- All the assimilation scenarios improved the model simulations.
- The H06DAPM outperformed the H24DAPM.
- The simulation in the PRD region differed less, possibly because the localized EI was adopted.
- The H06DAALL case is more or less the same with the H06DAPM case.

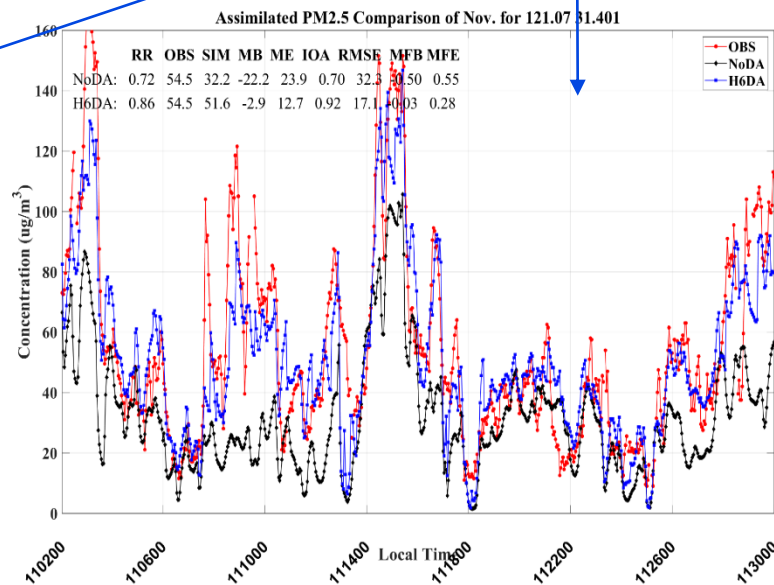
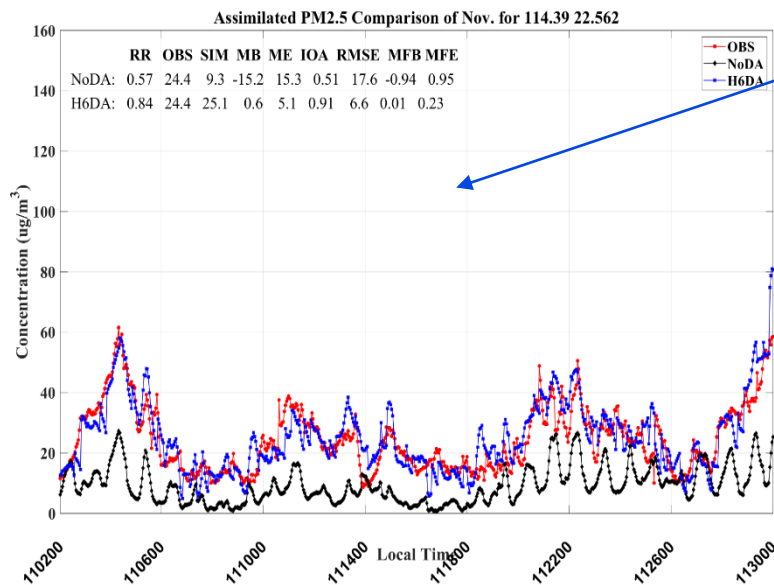
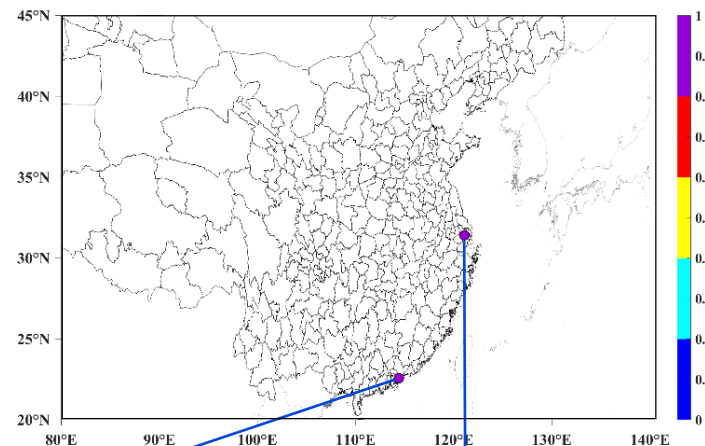
## November





## Monthly averaged statistical matrix (584 monitoring stations)

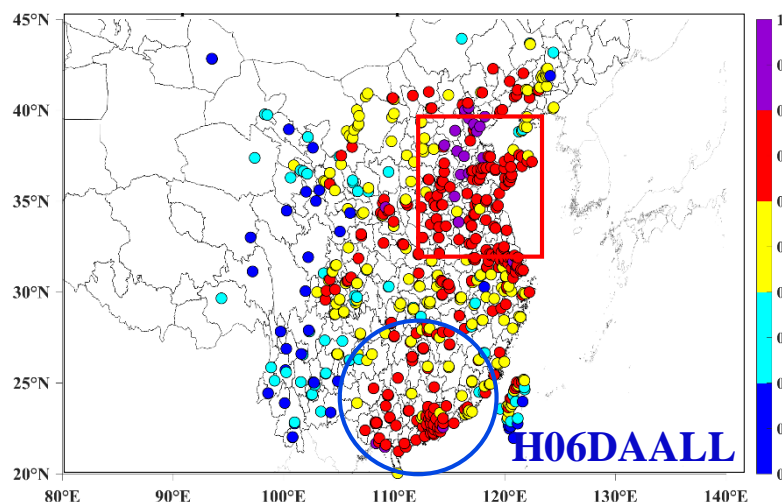
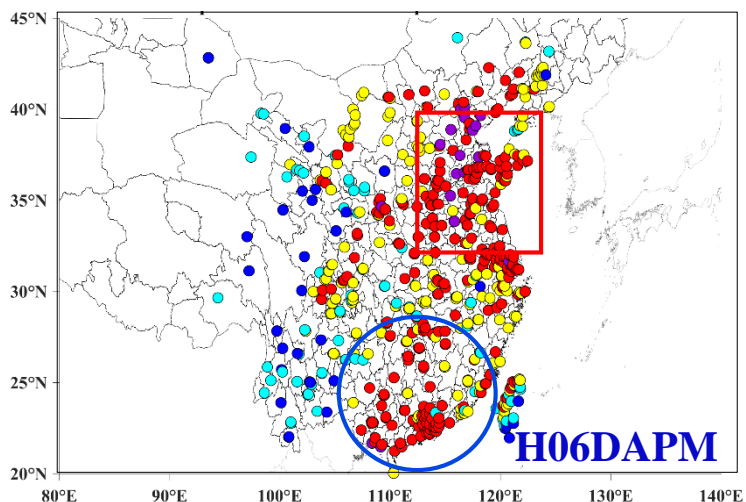
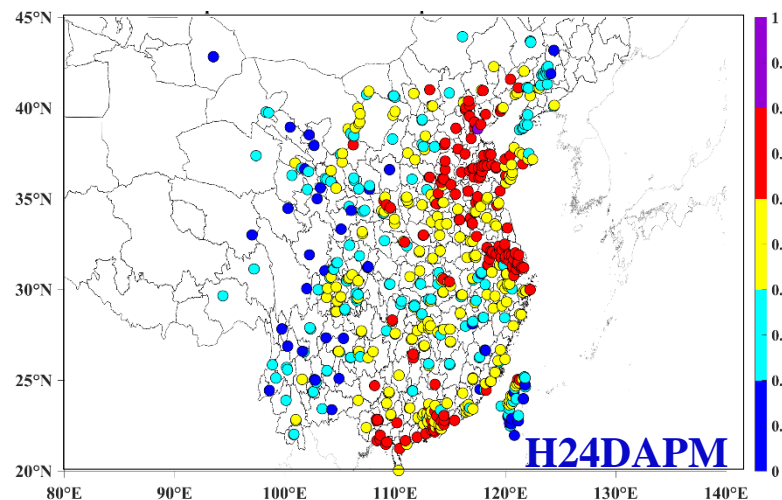
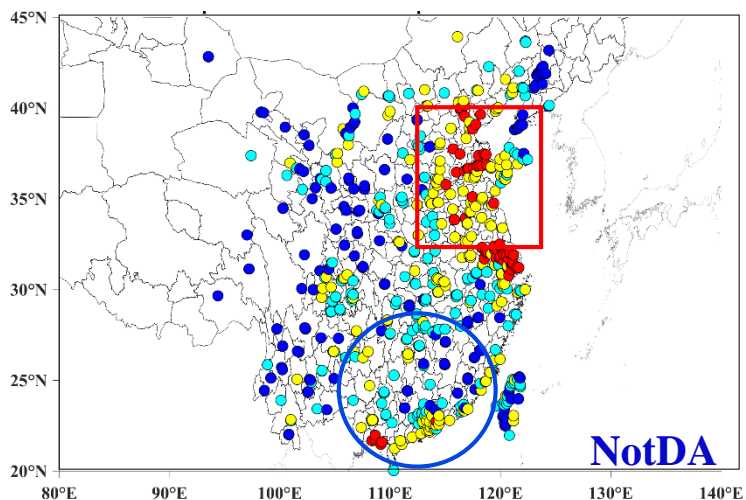
PM <sub>2.5</sub> November	CORR	MB	IOA	RMSE	MFB	MFE
NotDA	0.4	-21.18	0.54	37.52	-0.4	0.59
H24DAPM	0.52	-13.36	0.65	30.42	-0.23	0.45
H6DAPM	<b>0.61</b>	<b>-7.36</b>	<b>0.72</b>	<b>25.33</b>	<b>-0.1</b>	<b>0.37</b>
H06DAALL	<b>0.62</b>	<b>-7.03</b>	<b>0.73</b>	<b>24.9</b>	<b>-0.1</b>	<b>0.41</b>





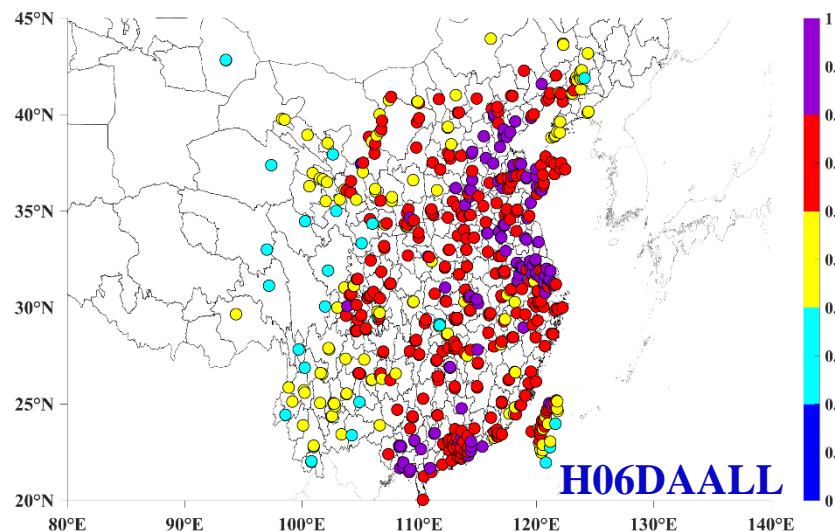
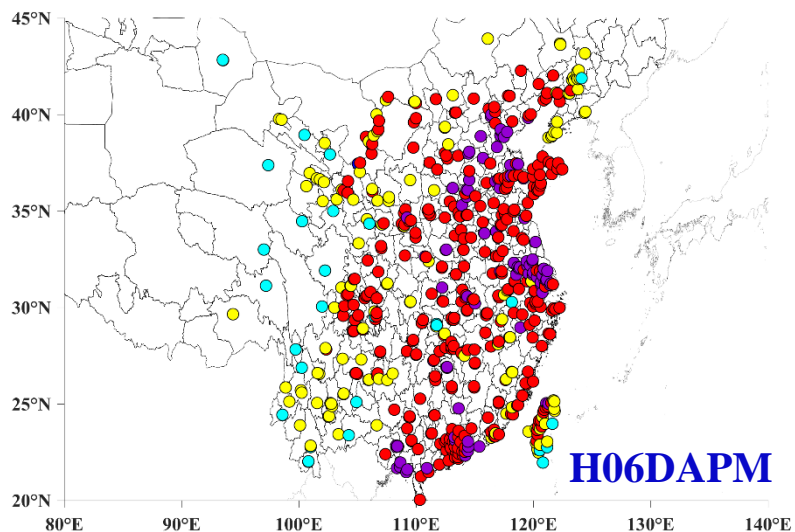
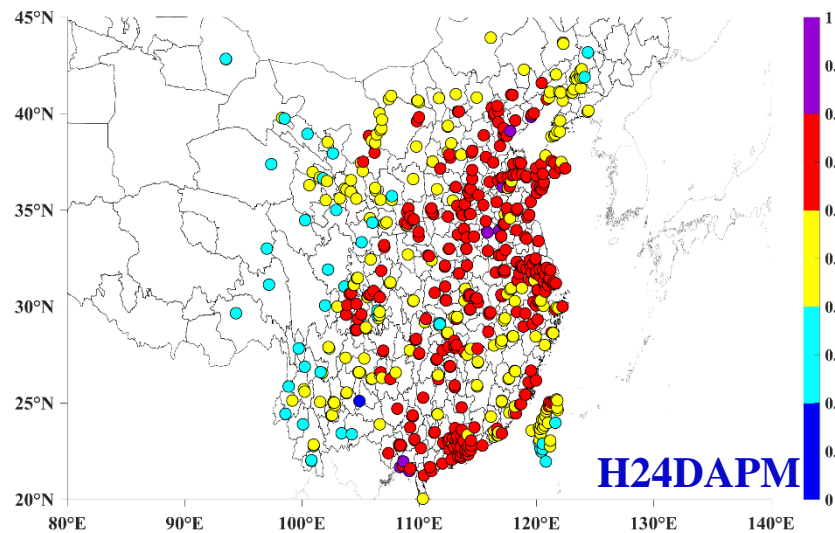
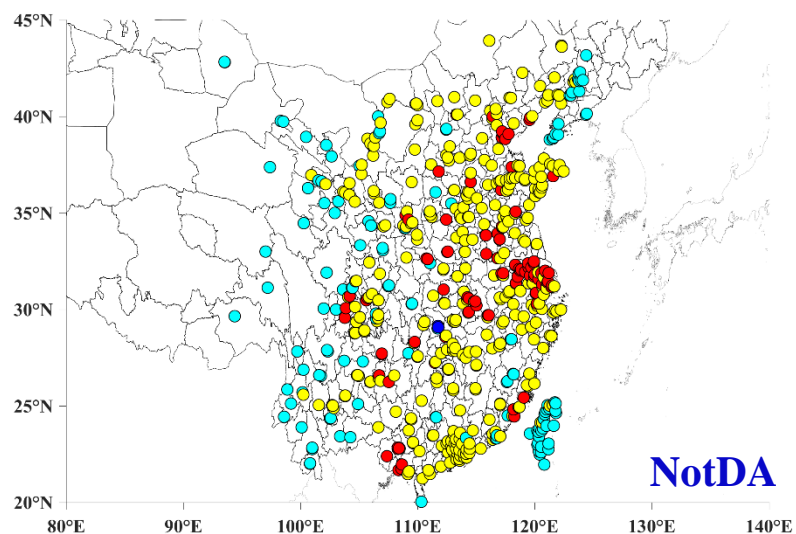


- The H06DAPM (red square) could improve model even the base model has already got a good correlation (0.6-0.8).
- The worse the base model performs (blue circle), the larger the improvement gained (0.4-0.6).
- The city-clusters tend to have a greater improvement than the boundary region.



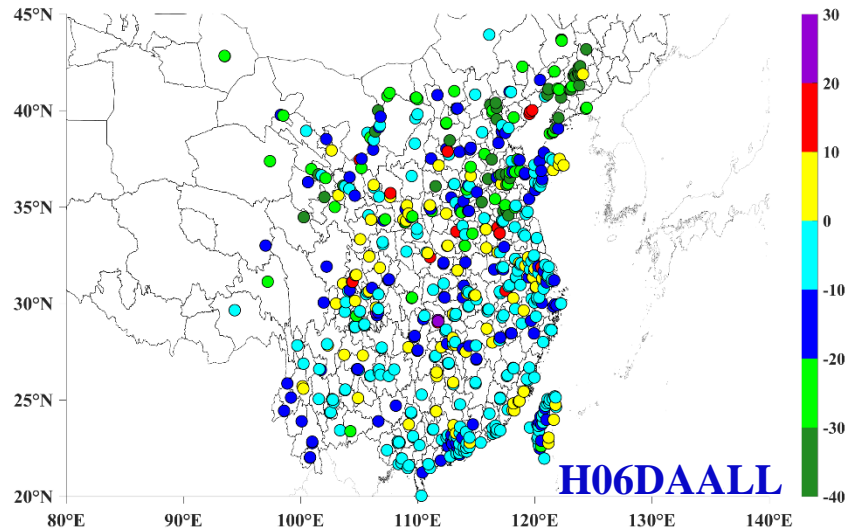
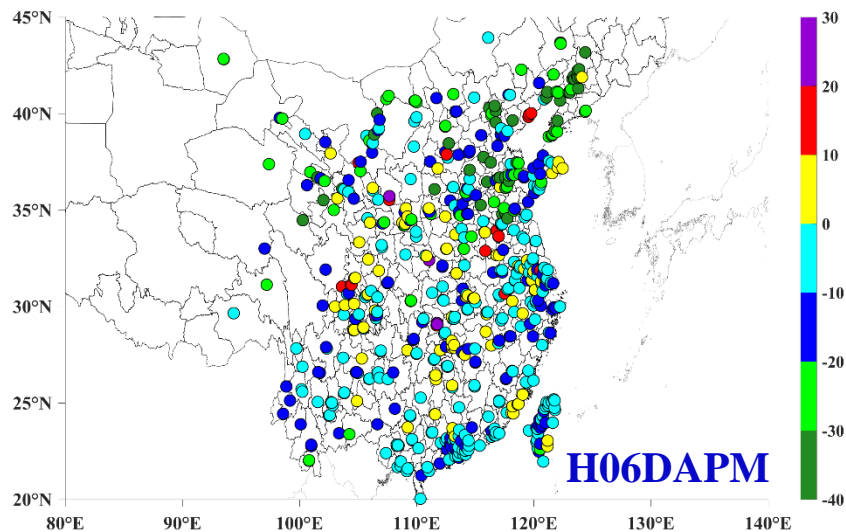
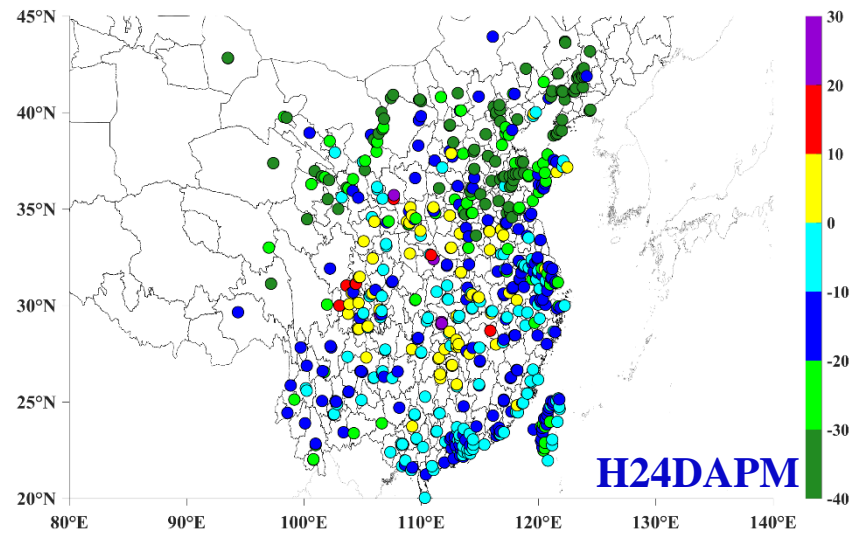
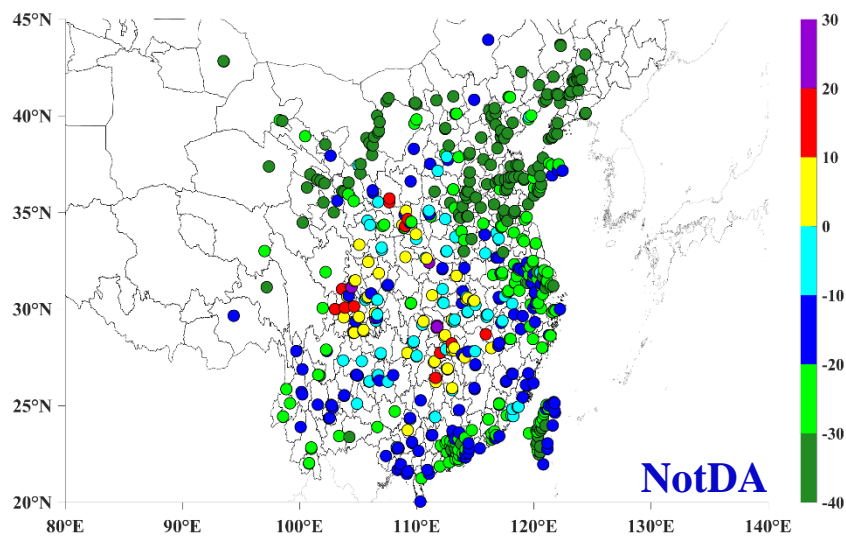


# Spatial distribution of the IOA for PM<sub>2.5</sub>





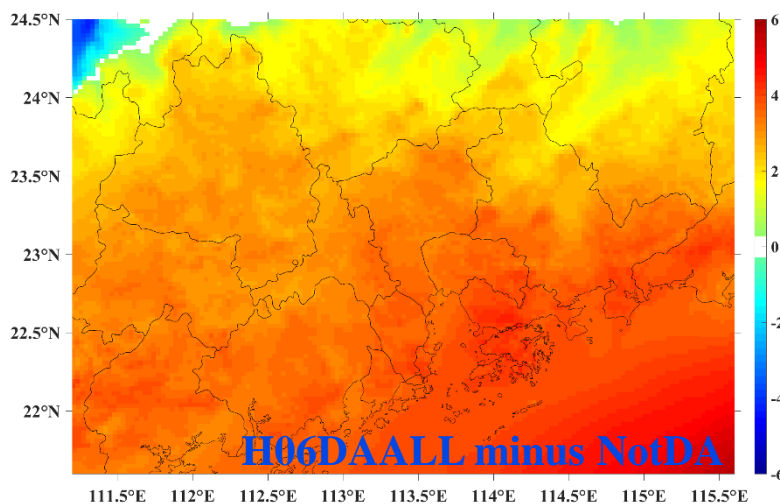
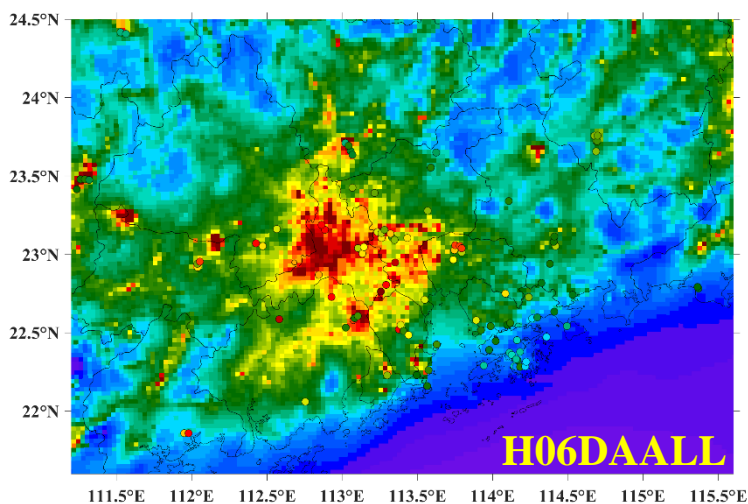
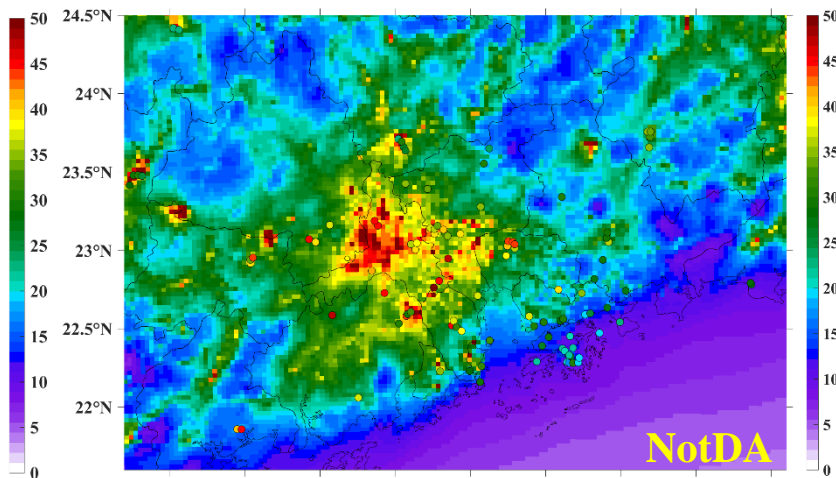
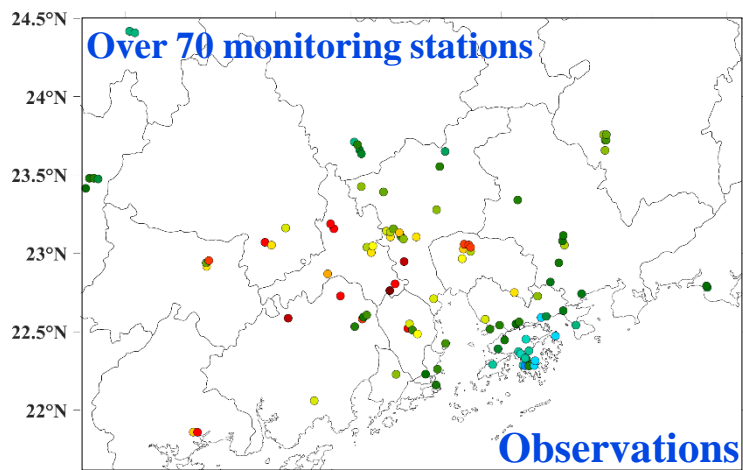
# Spatial distribution of the MB for $PM_{2.5}$







- Assimilating the air pollutants in the **outermost domain** (Domain 1) brought **substantial increases** in predicted  $\text{PM}_{2.5}$  in the **innermost domain** (Domain 3) **by around  $2\text{-}4\text{ }\mu\text{g}/\text{m}^3$**  for November in the PRD region.
- The potential reason is the north-easterly wind bring the pollutants along with the coastal line.





- **A novel 3D-Var approach** coupled with the CMAQ model assimilation system for improving model bias was constructed for the whole of China and the targeted PRD region.
- Sensitivity analysis scenarios were carried out to evaluate the **impact of the data assimilation frequency** on the benefits of assimilation.
- **Assimilating the air pollutants in the outermost domain** brought substantial improvement of the  $PM_{2.5}$  model simulations for the innermost domain, **offering an alternative method** to the existing domain-wide data fusion algorithms.
- **The north-easterly wind bring the air pollutants** along the **coastal line** to the PRD water, so as to lift up the domain-wide assimilated air pollutants during the domestic-heating season.
- **Multiple assimilations** of different pollutants, including  $PM_{2.5}$ ,  $SO_2$ ,  $NO_2$ ,  $O_3$ ,  $PM_{10}$ , and CO, confirm the effectiveness of the proposed data assimilation system in different geographic areas.



# Thank you!

zhangxuguo@ust.hk

- [1] **Zhang, X.**, Fung, J. C. H.\*, Lau, A. K. H., Zhang, S., & Huang, W. (2021). Improved modeling of spatiotemporal variations of fine particulate matter using a three-dimensional variational data fusion method. *Journal of Geophysical Research: Atmosphere*.