<u>Response to Community comments:</u>

<u>General comments:</u>

The authors present an approach to refine the ground meteorological stations in order to improve the regional air quality forecasts. Based on the sensitive areas for targeted observation, the authors first identified the essential stations, then scattered the other stations with different distances, finally a cost-effective observation network is provided. The refinement of the ground stations is a desirable one and the method is described in a clear and logical manner. Apart from the comments posted on the website by two referees, I only have a few minor comments.

Response: We thank your appreciations.

<u>Specific comments:</u>

1. ***What are the compositions of PM2.5 matter considered in the model?***

Response: The components of PM2.5 simulation here include black carbon, organic carbon, secondary inorganic aerosol (sulfate, nitrate, ammonium) and primary PM2.5 emitted directly from various sources. The compositions of PM2.5 matter considered in the model will be added in the revised manuscript.

2. ***Is it the domain setup in WRF the same with NAQPMS? Please clarify.***

Response: The WRF model is configured with the same horizontal and vertical grid structure with the NAQPMS model. The details will be added in the revised manuscript.

3. ***The boundary layer height is a key meteorological variable that affects the regional PM2.5 concentration, but the authors do not consider it in the definition of CNOP-type errors. The reasons why the BLH is not considered should be mentioned.***

Response: We agree with the reviewer that the boundary layer height is a key meteorological variable that affects the regional PM2.5 concentration forecasts. The CNOP in the present study only considers the sensitivity from initial uncertainties. Since the boundary layer simulation is more influenced by the parameterization in the WRF model (Chen et al., 2017; Mohan and Gupta, 2018), to study the role of boundary layer uncertainties in yielding the PM2.5 forecast uncertainties, an extension of the CNOP method, CNOP-parametric perturbation (CNOP-P; Mu et al., 2010) or nonlinear forcing singular vector (NFSV, Duan and Zhou, 2013), can be used to identify the sensitivities of boundary layer uncertainties. The related discussions will be added in the revised manuscript.

4. ***I am curious about how much improvements will be when the cost-effective stations are removed from the all the constructed stations. It is suggested to add some experiments in the manuscript. At least the authors could take some forecasts as examples to show the differences.***

Response: We thank the reviewer's suggestions. The CNOP-type error represents the initial error that results in the largest forecast error in the verification area at the verification time. The CNOP-type error considers the interaction among the errors on spatial grid points and in this situation, the errors on the grid points with large amplitude of the CNOP-type error contribute much more to the final prediction error. When we sort the spatial grid points with a decreasing order according to the amplitude of the error and choose the first 3% grid points as the essential grid points, the interactions between these grid points are remained, so that it is assumed that assimilating the observations on these grid points may contribute more to the improvements of forecast skills. Based on a series of OSSEs, it is verified that assimilating the essential or cost-effective observations can indeed improve greatly the PM2.5 forecasts. Specifically, when the 279 cost-effective station observations are assimilated for the AFs, they achieve an overall 41.11% the improvement of PM2.5 forecasting skills, which explains 99% the improvement when assimilating constructed station observations; furthermore, when the cost-effective station observations are removed from all the constructed station observations, the number of the rest station observations is 77 smaller than that of the cost-effective station observations and the assimilation of these observations explains much less, which is 70% the improvement obtained by assimilating all constructed station observations. To be specially emphasized, for the DFs, when the simulated observations from the 241 cost-effective station observations are assimilated, it results in an improvement of 47.55% of PM2.5 forecasting skills, even 1.7% higher than the improvement of assimilating all constructed station observations; however, when the cost-effective station observations are removed, assimilating the rest 240 station observations would only result in an improvement of 22.60% PM2.5 forecasting skill. Obviously, although the number of rest station observations is almost the same with the cost-effective station observations, the improvement of PM2.5 forecasting skills is less than half of the improvements obtained by assimilating the cost-effective station observations.

Totally, assimilating the cost-effective station observation will lead to much higher PM2.5 forecasting skills than assimilating the rest observations, which emphasizes the important role of the cost-effective station observations in improving the PM2.5 forecast skills. The relevant experimental results and discussions will be added in the revised manuscript.

5. ***A series of OSSEs is designed to verify the effectiveness of refined stations, due to the unavailable of real meteorological observations. It is suggested to add more discussions on how future work could use real observations.***

Response: We thank the reviewer's suggestion. We will add discussions in the revised manuscript.

As we showed on Lines 295-315 in the manuscript, to identify the sensitive area of the ground meteorological field in each forecast, we adopt the idea of Lorenz (1965) and take the better simulation initialized by ERA5 as "truth run" and the simulation initialized by GFS forecast data as "control run", where these two simulations have the same emission inventory and use the same model; so the difference between them reflect the sensitivities of forecast uncertainties of PM2.5 concentrations on the accuracy of initial meteorological field. When we compute the CNOP-type initial perturbation superimposed on the better simulation initialized by ERA5, it can be regarded as an approximation to the most sensitive initial error and the sensitive area identified by such CNOP-type error can be regarded as an approximation to the real sensitive area. If the approximate sensitive area is valid, assimilating the additional observations in the sensitive area of control forecast will make the updated forecasts approach to the truth run.

Although the present study is associated with hindcasts of PM2.5, it is still difficult to obtain the meteorological observations from the Monitor Center; therefore, we can only assimilate the simulated observations (i.e. the ERA5 data) to the control run to show the effectiveness of the cost-effective observation network. If the cost-effective station network is useful along this thinking, it can be inferred that assimilating real observations from the cost-effective stations to the initial field of the meteorological of the control forecast would improve the meteorological field forecasting and then the PM2.5 forecasting greatly against the observations.

References:

Chen, D., Xie, X., Zhou, Y., Lang, J., Xu, T., Yang, N., Zhao, Y., Liu, X., 2017. Performance evaluation of the wrf-chem model with different physical parameterization schemes during an extremely high PM2.5 pollution episode in Beijing. Aerosol Air Qual. Res. 17 (1), 262–277.

Duan, W., and Zhou, F., 2013. Non-linear forcing singular vector of a two-dimensional quasi-geostrophic model. Tellus, 65(18452), 256-256.

Mohan, M. and Gupta, M., 2018. Sensitivity of PBL parameterizations on PM10 and ozone simulation using chemical transport model WRF-Chem over a sub-tropical urban airshed in India. Atmospheric Environment, 185, 53-63.

Mu, M., Duan, W. S., Wang, Q., and Zhang, R., 2010. An extension of conditional nonlinear optimal perturbation approach and its applications, Nonlin. Processes Geophys., 17(2), 211-220.