Thank you for your efforts on this manuscript. The comments are valuable and very helpful for revising and improving this work. The detailed responses to the comments are given below one by one.

1) The simulation was done in the period of November 20, 2016 to January 20, 2017 but the comparison of simulations and observations was made only in December, 2016, why?

Response: According to the study of Yang et al. (2020), December 2016 was a month with severe $PM_{2.5}$ pollution, including a severe pollution process. Therefore, this study finally selected December 2016 for simulation and analysis. Because we did not know how long the spin-up time would take before the simulation experiments, we simulated forward 10 days as the spin-up time.

2) For model evaluation of meteorological fields, the comparison of simulated and observed wind speed and direction should be added.

Response: The comparison of simulated and observed wind speed has been done in the study of Yang et al. (2020). We used the same model configuration and monitoring sites. The simulation time period in the study of Yang et al. (2020) includes the simulation time period of this study. So the model performance was similar. As shown in Fig. 1 (Yang et al., 2020), the WRF detected the variation of meteorological elements in the pollution period. And the verification statistics were shown in Table 1(Yang et al., 2020). As the results show, the W10 is underestimated. The MB of the W10 is -0.14 m/s. The R of the W10 is 0.63, which indicates a good agreement between the observations and the model results. Owing to the particular topography of Xi'an, the wind speed is low at all times. The RMSE of the simulated W10 is only 0.96.

Due to the lack of wind direction data, and the large error in model performance of wind direction shown from previous studies, this study did not evaluate the wind direction.

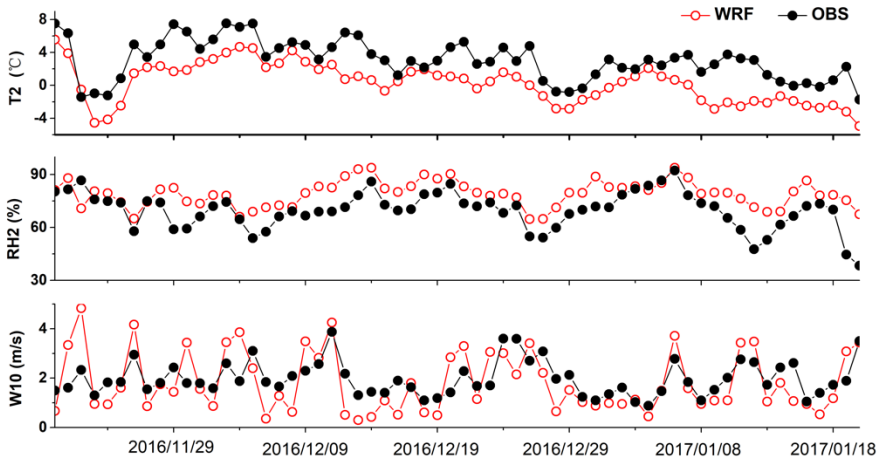

Figure 1. Time series of daily observed and simulated temperatures at a 2 m altitude (T2), relative humidity at a 2 m altitude (RH2) among nine monitoring sites, and daily average wind speed at a 10 m altitude (W10) in Xi'an station from 20 November 2016 to 20 January 2017. The black and red lines represent the observations and the values simulated by WRF. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.) (Yang et al., 2020)

Table 1. Verification statistics of daily temperature at a 2 m altitude (T2), relative humidity at a 2 m altitude (RH2), and wind speed at 10 m altitude (W10) from 20 November 2016 to 20 January 2017. ME, MB, R, and RMSE are abbreviations for mean error, mean bias, correlation coefficient, root-mean-square error, and accuracy rate, respectively. The units of the Mean, ME, and MB are for T2, % for RH2, and m s^{-1} for W10. (Yang et al., 2020)

Variable	Mean		ME	MB	R	RMSE
	Obs.	Sim.				
T2 (℃)	3.06	0.28	2.80	-2.77	0.82	3.14
RH2 (%)	69.80	79.05	9.91	9.26	0.71	11.87
W10 (m·s-1)	1.98	1.83	0.79	-0.14	0.63	0.96

3) The CAMx has only one domain . . .(section 2.3)? Actually, Fig.4 shows that the CAMx also has three nested domains.

Response: Fig. 4 shows the model domain in the WRF-CAMx modelling system. In this study, three-nest domains were designed for the WRF model, as shown in Figure 4. If the meteorological field only simulates one domain, directly interpolating from the initial field resolution to a 3 km-grid, the error will be relatively large. The CAMx in this study has only one domain and the settings are the same as those in the D3 domain.

4) Please explain why the model performance in the studied urban area is poorer than that in the studied counties (section 4.1.1).

Response: Thank you for your comment. We randomly selected three monitoring sites in the urban area for analysis shown in Fig 8, so the conclusion may not be rigorous enough. According to your suggestion, we checked again and verified the statistical parameters of all monitoring sites. As shown in Table 2, For correlation coefficients (R), the model performances of stations in suburban counties are better than that of some stations in urban city, indicating that the simulation of suburban counties peforms better in the aspect of trends. The reason may be that PM2.5 emissions in urban areas are affected by more complex factors, and it is more difficult to simulate weather in urban areas. For bias, the model performances of stations in urban cities are better than that of the suburbs, and the simulated $PM_{2.5}$ concentrations in the suburbs are obviously underestimated. We also used the fraction of predictions within a factor of two of observations (FAC2) to verified the model performances of all the monitoring sites. The FAC2 of the urban cities is 74%, while the FAC2 of the suburban counties is 62%, indicating that the model performance of the urban area is better than that of the suburban area in terms of total emissions. We will revise the manuscript based on the following results.

Table 2. Verification statistics of PM2.5 concentrations on December 2016 among all monitoring sites.

	Station	R	MB ($\mu g/m^3$)	ME ($\mu g/m^3$)	NMB %	NME %	RMSE	FAC2 %
	CT	0.70	-58.86	66.22	-38.08	42.84	89.99	64
	XQ	0.53	61.72	81.41	36.00	47.48	108.31	75
	JKQ	0.69	-76.00	84.91	-37.55	41.95	109.89	76
	TYC	0.66	75.46	89.90	45.98	54.78	108.49	77
urban	GYC	0.73	19.42	52.51	11.17	30.21	68.01	89
	QJ	0.59	14.79	71.73	8.06	39.11	89.90	76
	GYT	0.62	-31.46	69.01	-18.72	41.06	84.15	66
	FZC	0.67	3.11	60.68	1.79	35.01	76.42	78
	XZ	0.57	51.62	79.66	32.64	50.36	101.86	72
	GX	0.64	67.11	88.46	39.42	51.96	104.45	69
	CAQ	0.68	8.51	54.49	5.36	34.29	72.26	86
suburban	YLQ	0.60	-97.63	99.12	-57.67	58.55	120.78	31
	LTQ	0.57	-54.43	76.28	-33.53	46.98	97.49	69

5) Please use “simulation” instead of “mechanism” in the whole text.

Response: The authors thanks for your constructive suggestion. We will use “simulation” instead of “mechanism” in the revised version.

6) The writing English need more polishment.

Response: We will find the experts to help polish the writing English, and try our best to improve the level of writing English.