Response to Referee #2

The authors wish to thank Reviewer #2 for the comments, which greatly contribute to an improvement of the paper.

In the following, we address the issues raised by the reviewer:

Q2.1: This study only focuses on surface temperature (T2). How about water vapor and winds? You may improve temperature performance at the cost of worse performance for other variables. One cannot simply focus on one variable but ignore other variables during model calibration/optimization.

A2.1: The main objective of the present study is to estimate the influence caused by the change of Kzmin on the prediction of the near-surface temperature. Thus, we focused more on the temperature prediction in the manuscript. Actually, we have also performed comparisons of other meteorological parameters such as the wind speed with the observational data, which is presented below.

The values of statistical parameters for simulations of 2-m temperature (^{o} C), 10-m wind speed (m s^{-1}) and 2-m specific humidity (i.e. Q2) (g kg^{-1}) are listed in Tab. A1 of this rebuttal. It can be seen that the changes of the wind speed and the specific humidity caused by the altering of Kzmin are significantly smaller than that of the 2-m temperature. It means that the simulation accuracy of the wind speed and the specific humidity is not heavily influenced by the change of Kzmin. Thus, we paid more attention to the influence on the temperature prediction brought about by the change of Kzmin.

Table A1 Statistical parameters for simulations of 2-m temperature (^{o} C), 10-m wind speed (m s^{-1}) and 2-m specific humidity (i.e. Q2) (g kg^{-1}).

	T2				W10				Q2			
Kzmin	RMSE	IOA	R	MB	RMSE	IOA	R	MB	RMSE	IOA	R	MB
0.01	2.79	0.84	0.94	-2.49	3.21	0.26	0.64	2.51	0.21	0.85	0.73	0.02
0.2	2.32	0.88	0.93	-2.00	3.27	0.26	0.62	2.60	0.21	0.84	0.72	0.01
0.5	2.01	0.90	0.92	-1.58	3.31	0.25	0.63	2.70	0.21	0.85	0.73	0.01
0.8	1.88	0.91	0.90	-1.27	3.38	0.25	0.62	2.77	0.20	0.85	0.74	0.01
1.0	1.82	0.91	0.89	-1.08	3.39	0.25	0.62	2.79	0.20	0.86	0.75	0.00

The time series of the 10-m wind speed and the 2-m specific humidity (i.e. Q2) simulated in five scenarios using different constant Kzmin values are shown in Fig. A1 of this rebuttal. It can be seen that the increase of Kzmin exerts a small influence on the change of the wind speed and the specific humidity. Moreover, from Fig. A1(a), it was found that the increase of Kzmin induces a stronger 10-m wind speed. In contrast, the specific humidity Q2 was found lower under a higher Kzmin (see Fig. A1(b)). It is because that a larger Kzmin causes a stronger mixing. As a result, the momentum is transported downwards from higher altitudes and the moisture is transported upwards from altitudes near the surface. Thus, a larger 10-m wind speed and a lower Q2 were obtained in the present study when a higher Kzmin is used.

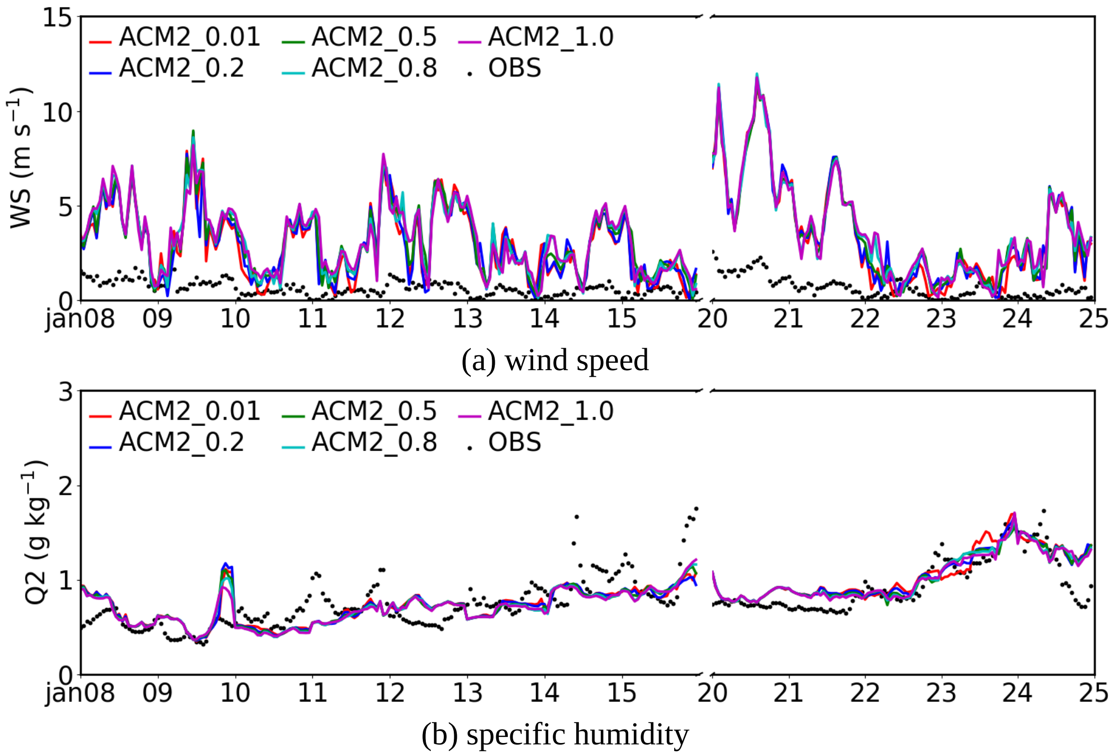

Figure A1 Time series of 10-m wind speed (WS) and 2-m specific humidity (Q2) simulated by five scenarios with different Kzmin values.

Although the impact on 10-m wind speed and 2-m specific humidity caused by the change of Kzmin is relatively minor, based on the suggestions of the reviewer, we still added the related content about the validations of the simulated wind speed and the specific humidity in the revised supplementary material. Please see Sect. 4 of the supplementary material and the corresponding discussions.

Q2.2: Only data at one urban site are used in model evaluation in this study. The cold bias at this site may not be representative. Normally, regular soundings are used for PBL scheme calibration. Why don't use soundings?

A2.2: Thanks a lot for the suggestion of the reviewer. We added the content about the evaluation of 2-m temperature and 10-m wind speed at five observation stations in the revised manuscript in Section 3.1. These five observation stations include the IAP, CAS station and four AWS stations (Automatic Weather Stations). The information of the IAP station has been given in the original manuscript. Regarding to the four AWS stations, the No. 54433 AWS station is located in urban area of Beijing, similar to the IAP station, while the other three AWS stations (No. 54406, 54419, 54501) are located in rural or suburban areas of Beijing.

Table A2. Values of statistical parameters measuring the model performance in simulating the 2-m temperature (T2) and the 10-m wind speed (W10) at five observation stations.

	T2				W10			
Station	RMSE	IOA	R	MB	RMSE	IOA	R	MB
IAP	2.79	0.84	0.94	-2.49	3.21	0.26	0.64	2.51
54406	2.84	0.88	0.83	1.06	3.08	0.44	0.50	2.21
54419	3.16	0.85	0.86	1.11	2.28	0.30	0.35	1.49
54433	2.17	0.92	0.91	-1.38	2.40	0.62	0.65	1.26
54501	4.85	0.76	0.78	2.75	2.06	0.52	0.36	0.94

The values of statistical parameters measuring the model performance are listed in Tab. A2 of this rebuttal. From a global view, the model behavior in capturing the 2-m temperature is satisfying. The correlation coefficients between the simulated temperature and the observations at these five stations reside in a value range of 0.78-0.95. Moreover, the index of agreement (i.e. IOA) also possesses a value above 0.75 for all these five stations. It was also found that the model performs better at the two urban stations (IAP and No. 54433) than at the other three rural stations, denoted by a relatively smaller RMSE and a higher R (see Tab. A2).

In contrast to that, the model tends to predict a higher wind speed at all five stations (see RMSE of W10 in Tab. A2). The deviation between the simulations and the observational data is more obvious at the IAP station, as it possesses the largest RMSE of 3.21 m/s. Moreover, from the values of the correlation coefficient R, it was found that the simulated trend of the 10-m wind speed at two urban stations (IAP and No. 54433) is more consistent with the observations than that at the rural stations, as the correlation coefficient R at these two urban stations reaches a value above 0.6.

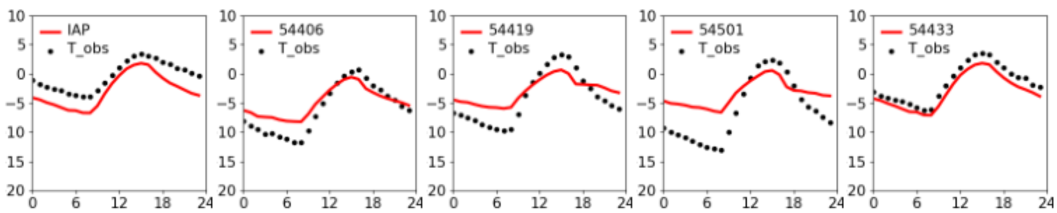

Figure A2. Diurnal mean time series of the 2-m temperature (T2) at five stations.

Figure A2 of this rebuttal also shows the diurnal change of the averaged T2 at these five stations as well as the observations. It can be seen that at these stations, the highest T2 appears at approximately 15:00 local standard time (LST), while the lowest T2 appears at around 8:00 LST. This is also the reason why we focused on these two time points in the present study. Moreover, it was found that at the three rural stations (No. 54406, 54419 and 54501), the simulated 2-m temperature at these stations is higher than the observational value in the nighttime, while it is lower than the observation during the daytime when the temperature is high. As a result, the simulated diurnal variation of T2 is weaker than the observations. On the contrary, at two urban stations, the cold bias appears during the whole day.

With respect to the 10-m wind speed, Fig. A3 shows the spatial distribution of the time-averaged 10-m wind speed over the nighttime. It was found that a larger wind speed is estimated in mountain areas than in plain areas, leading to a stronger wind shear in mountain areas than that in plain areas.

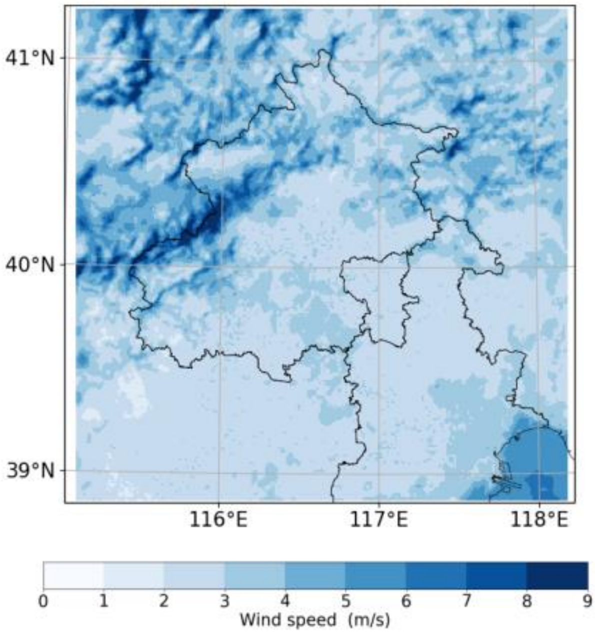

Figure A3. The spatial distribution of the mean 10-m wind speed over the nighttime.

Evaluations of the 2-m temperature and the 10-m wind speed at these five observation stations are added in the revised manuscript as a section "3.1 Model Evaluations". Please see lines 211-237 in the revised manuscript.

Currently we did not use the data from regular soundings in the present study, because the vertical observational data from the tower has a relatively high time resolution of 10 minutes. In contrast, the time resolution of regular soundings is lower, i.e. 12 hours. Aside from that, the observation tower is located next to the surface observation stations, while the locations of regular soundings are more remoted. Thus, the vertical observational data obtained from the tower are more consistent with the surface data achieved from the surface stations, compared with those from regular soundings. Thus, we prefer the observational data from the tower. However, if the reviewer insists on adding the comparison with regular soundings, we are happy to perform this comparison if the submission deadline of the manuscript revision can be extended further. Thanks again for the valuable suggestion.

Q2.3: Cold bias seen at the urban site in this study may be due to model errors associated with urban scheme and anthropogenic heat, as well as errors in initial model states particularly soil moisture. So, this study may be attributing other model errors to PBL scheme.

A2.3: Thanks for the suggestion. The reviewer possibly thought that we were looking for a better PBL scheme in capturing the change of near-surface temperature over Beijing area in the present study. However, the focus of this study is to figure out the effect of using various Kzmin on the prediction of the temperature, instead of finding an improved PBL scheme. We are clearly aware that the deviation between the simulation results and the observational data can be caused by many other factors rather than the modification of the PBL scheme as the reviewer pointed out. Moreover, even we found the simulation results using Kzmin =1.0 closer to observations, we cannot guarantee that the scheme using Kzmin =1.0 is an improved scheme which is more realistic, because the better performance of the modified scheme can be caused by the offset of biases introduced by other factors. In order to avoid the misunderstandings that the biases all come from the modifications of the PBL scheme, we have revised all the related contents throughout the manuscript. Please see lines 108-109, 230-234, 435-436 and 494-495 of the revised manuscript.

Minor comments:

Q2.4: LN27-30 These are old PBL schemes, not including modern PBL schemes, for example scale-aware schemes and mass flux schemes

A2.4: Thanks a lot for the information. We have added the contents introducing some modern PBL schemes in the revised manuscript. Please see lines 43-54 in the revised manuscript.

Q2.5: LN128-130

Aerosol effect may impact temperature, which is not considered by the model simulation in this study. So, this study ascribes the potential bias of not considering aerosol effects to vertical mixing.

A2.5: This question is similar to the question **Q2.3** above. The reviewer is correct saying that the existence of aerosols is able to significantly affect the behavior of the model and thus the deviation between the simulation results and the observations. However, as we stated in the answer **A2.3** that the main objective of the present study is not to find a more appropriate PBL scheme to reduce the deviation. Instead, in the present study, we actually tried to assess the influence on the change of temperature caused by using different Kzmin values. In order to indicate this issue more clearly, we added the information about the treatment of aerosols in the revised manuscript. Please see lines 145-146.

Q2.6: LN166, what is the Landusef value used in the simulation

A2.6: The Landusef value as well as PURB used in this study are given in Fig. A4 of the rebuttal. We have also included these figures in the revised supplementary material (see Fig. S2 of the supplements). Moreover, the related description is also added in the revised manuscript. Please see line 186 of the revised manuscript.

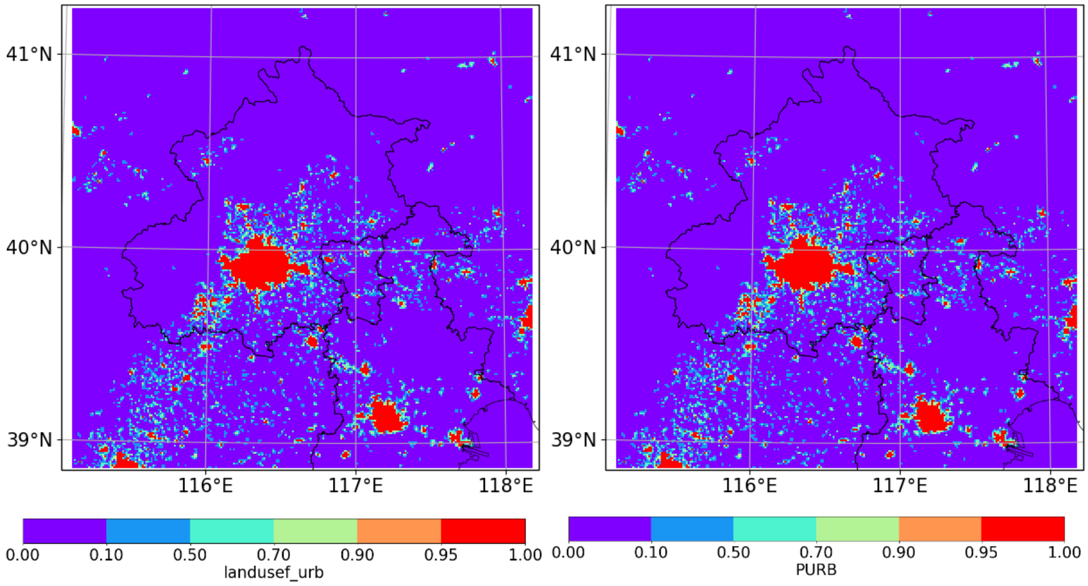

Figure A4. The spatial distributions of the land-use fraction of the urban category and the calculated PURB.

Q2.7: LN255

I don't think Kzmin would affect daytime prediction. It must be the residual effects from the nighttime impact

A2.7: This question is actually similar to the question **Q1.11** raised by Reviewer #1. The reviewer is correct saying that the nighttime bias can accumulate so that it would impact the daytime simulation. However, in ACM2, the implementation of Kzmin can be simply represented as follows:

$$Kz' = Kz + Kzmin,$$

which means that Kzmin is added to Kz to constitute a total vertical turbulent diffusivity Kz'. As a result, even though in the daytime when Kz is relatively large, Kzmin is still able to affect Kz' in areas that the turbulent mixing is relatively weak. We have added this information as well as the explanations in lines 319-323 of the revised manuscript.

Based on the information given above, we can conclude that the difference in the simulated temperature during the daytime brought about by the change of Kzmin is caused by the combined effect of the large temperature difference during the nighttime and the different turbulent mixing intensity during the daytime. To clarify it, we performed another numerical experiment. In this experiment (named AC_night_0.01), Kzmin was set to 0.01 during the nighttime (same as ACM2_0.01), but 1.0 during the daytime (same as ACM2_1.0). By doing that, the contributions to the difference of the temperature by these two processes can be assessed separately.

The time-averaged vertical profiles of the potential temperature at 8 LST and 15 LST are shown in Fig. A5. In Fig. A5(a), potential temperature profiles belonging to AC_night_0.01 and ACM2_0.01 were found close to each other, due to the same values of the nighttime Kzmin used in these two scenarios. In contrast to that, in the daytime (see Fig. A5(b)), AC_night_0.01 was found predicting a higher temperature than ACM2_0.01, which is caused by the increase of Kzmin during the daytime and the stronger turbulent mixing.

In contrast, AC_night_0.01 was also found giving a lower temperature than ACM2_1.0, although a same Kzmin (=1.0) is used during the daytime in these two scenarios. Thus, the difference between AC_night_0.01 and ACM2_1.0 denotes the residual effect caused by the bias of the temperature during the nighttime.

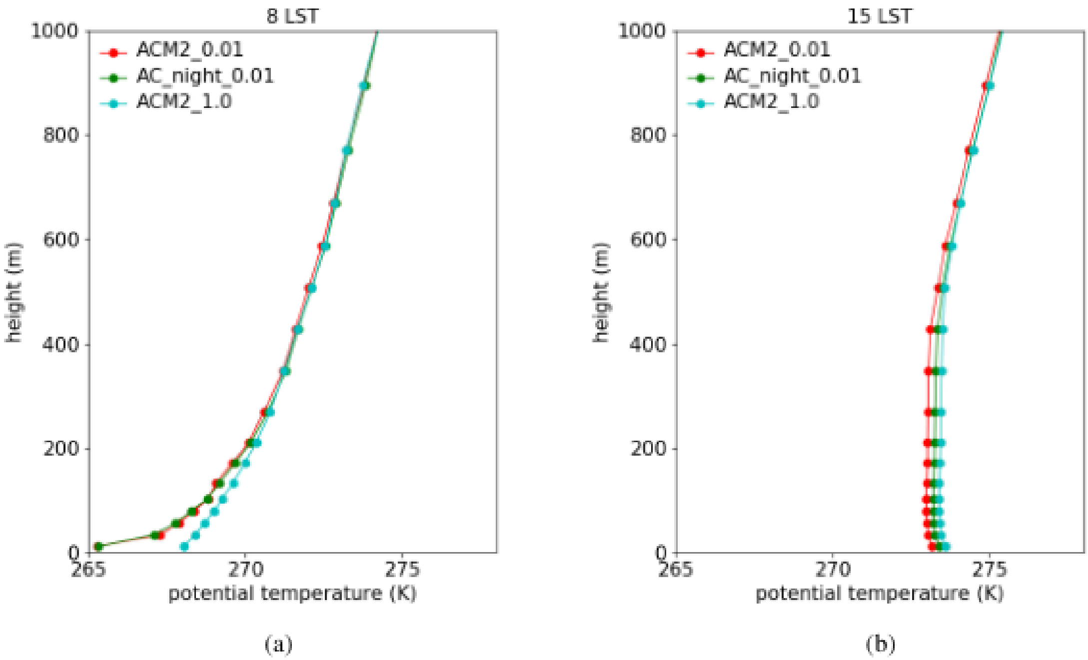

Figure A5. The vertical profiles of the potential temperature predicted by ACM2_0.01, ACM2_1.0 and AC_night_0.01 at (a) 8 LST and (b) 15 LST, averaged over the simulated days.

Thus, according to this numerical experiment, we confirm that there are two primary processes causing the temperature difference during the daytime. One is the change of Kzmin during the daytime. When Kzmin is increased, the vertical mixing in the boundary layer is enhanced, which causes a stronger transport of the air from the upper layer into the boundary layer, resulting in a warmer boundary layer during the daytime. The other process is the residual effect caused by the change of Kzmin in the nighttime. It is because that different Kzmin results in a large deviation in the near-surface temperature during the nighttime. This deviation would maintain until the daytime comes so that the prediction of the daytime temperature would also be affected.

The corresponding results and the related discussions are added in the revised manuscript. Please see lines 327-355.

Q2.8: In Figure 6.
Actually K=0.01 gives better temperature variation/cooling rate during nighttime. This simulation just has some systematic bias, which may not be due to PBL schemes, but due to other model/inputs bias/errors, for example, the systematic bias from urban scheme and uncertainties in initial land properties especially soil moisture.
A2.8: This question is similar to the question **Q2.3**. The reviewer is correct saying that K=0.01 predicts a better trend of the temperature change during the nighttime. However, the magnitude of the simulated temperature obtained by assuming K=1.0 was found closer to observations in the present study, compared to those using other Kzmin values. However, the objective of the present study is actually to find out the influence of the change of Kzmin on the estimation of temperature, instead of figuring out an improved PBL scheme. We are clearly aware that the deviation between the simulation results and the observational data can be caused by other factors as the reviewer pointed out. Even we found the simulation results using Kzmin =1.0 closer to observations, we cannot guarantee that the scheme using Kzmin =1.0 is an improved scheme which is more realistic, because the better performance of the modified scheme can be caused by the offset of biases introduced by other factors. In order to clarify the objective of the present study more clearly, we have revised our manuscript and made related modifications in the revised version of the paper, to avoid possible misunderstandings. Please see lines 230-234 of the revised manuscript.

Q2.9: ACM2_CMAQ should be identical as K=1 since urban fraction is 1.
A2.9: Thanks. In completely non-urban areas, ACM2_CMAQ is identical to K=0.01, while in completely urban areas, it is identical to the case K=1.0. We have made the statement more clearly in the manuscript, please see lines 189-190 of the revised manuscript.

Q2.10: Fig. 7, there is only one profile. Is this a profile averaged over several days or just at a specific time of a day? Please be explicit.
A2.10: This a profile averaged over the time period under the investigation of the present study. In order to avoid the misunderstanding of the reviewer, we refined the captions of the related figures in the revised manuscript. Thanks for the comment.