## **Reply to Referee #1**

We appreciate the important comments made by the reviewer and we hope that our manuscript has improved.

As the authors repeated their simulation with a more recent version of WRF my main objections are eliminated.

A more detailed comments about the potential effect of the use of different land surface data sets for MM5 and WRF should be added. Which influence has this choice on the initial conditions etc?

#### Answer:

We added the following statement in Chapter 1 "Introduction":

"The land surface is a key component of meteorology, and air quality models. In meteorology modeling land surface exchange process are based on land cover categories within each modeling grid. It controls the partitioning of available energy at the surface between sensible and latent heat, and it controls the partitioning of available water between evaporation and runoff. In air quality modeling chemical surface fluxes are modeled based on different land cover categories. In this work the Noah land surface model (LSM) was used for both MM5 and WRF. The main objective of this scheme is to provide four parameters to the meteorological model: surface sensible heat flux, surface latent heat flux, upward longwave radiation, and upward shortwave radiation. This scheme is important because these variables represent the redistribution of energy at the surface atmosphere interface, and consequently impacts other variables such as PBL evolution, temperature, etc. The Noah scheme requires three input parameters: vegetation type, soil texture, and slope. All other parameters. Different land surface data sets can present distinct results for the energy redistribution in the model, consequently impacting the dynamic characteristics of simulation."

Are the PBL heights estimated in the same way for the YSU and the MRF scheme? How is the PBL height calculated in models? Please add this Information.

#### Answer:

#### We added the following information in Chapter 2.5:

"The YSU and MRF PBL schemes use nonlocal closure and rely heavily on Ri to compute PBL height for different regimes (e.g., stable, unstable, and neutral PBLs). Both of these PBL schemes essentially define PBL height as the height at which a critical Ri is reached 0.5 for the MRF scheme and 0.0 for the YSU scheme (Skamarock et al. 2008). For unstable conditions the PBL height in the YSU scheme is determined to be the first neutral level based on the bulk Richardson number calculated between the lowest model level and the levels above (Hong et al. 2006)."

What is the reason for the extreme spikes of SWUP for MM5? Please discuss the effect on the other variables. How well is cloudiness represented by WRF and MM5, respectively?

#### Answer:

We do not know the reason. It is not a consistent feature. It predominantly occurs on the two prefrontal days. Also, the occurrence of spikes of SWUP in MM5 are accompanied by concurrent underprediction of SWDOWN in MM5; often these underpredictions in SWDOWN are of the similar magnitude as the overpredictions in SWUP, which points to a deficient energy distribution in MM5 on these days. Regardless of the nature of this deficiency, its impact is visible in corresponding underpredictions in the fluxes of sensible and latent heat (Figure 5), and to some extent it is also reflected in the temperature time series (Figure 3). The significant PBL underprediction in MM5 on August 29 is likely due to the underprediction of sensible and latent heat in MM5 on that day (Figure 5), which in turn might be associated with the deficient simulation of incoming and outgoing shortwave radiation by MM5 on the same day. We added corresponding findings in the text in the section "Incoming/outgoing shortwave radiation".

Cloudiness may have a large impact on SWDOWN and related subsequent processes. Cloudiness represented by corresponding schemes in WRF and MM5 is likely associated with some higher degree of uncertainty. For instance, cloud fraction in a grid box is assumed either 0 or 1 (Dudhia, 1989), which is certainly not precise enough. Unfortunately, in our study we did not have quantitative information about observed cloudiness und thus we were not able to perform validation studies for these schemes. We added a corresponding statement in the chapter 2.5 Differences between the WRF and MM5 configuration.

A minor point: Why is are MM5 results labeld as 'EDAS' in the figures?

## Answer:

The abbreviation EDAS is now explained in the text. We replaced the label EDAS by MM5 in the plots.

## **Reply to Referee #2**

We appreciate the important comments made by the reviewer and we hope that our manuscript has improved.

The authors have been responsive to the comments of the reviewers. However, after simulations were done, several issues in the results and discussion part need to be clarified before the paper can be published. These issues are listed in the following:

Major issues or questions:

\* The authors estimate the roughness length near the observational location to be about 0.05m. Is a similarly low value also used for the simulations and does the value agree between both used models? What values of vegetation parameters are used in the simulation (such as vegetation fraction, leaf area index, ...), are they similar in both model runs and do they agree with the parameters of the impact area of the observations? If the surface parameters used in the model runs do not agree, which part of the differences between both models can be attributed to surface parameter differences and which to model differences?

## Answer:

We appreciate this comment. We set the roughness length to the same value as used in both models, which is 0.07 m. The impact distance was recalculated to be 80-98 m (previously: 85-100 m) and the 90% impact boundaries is 219-270 m (previously: 230-285 m).

We included the vegetation parameters, which are the same in both models, in a new table (table 3):

Parameter	Meaning	Value	Category
LU_INDEX	Land use index	2	Dryland cropland and pasture
IVGTYP	Vegetation type		
ISLTYPE	Soil type	12	Clay
VEGFRAC	Vegetation fraction	55.20	
SHDFAC	Green vegetation fraction	0.80	
LAI	Leaf Area Index	4.96	
EMISS	Emissivity	0.97	
ALBEDO	Albedo	0.18	
Z0	Roughness length	0.07	

Table 3. Noah LSM parameters used for the UH-CC site in the MM5 and WRF simulations.

\* The values of the reflected shortwave radiation with MM5 look very strange during the first two days. The pattern does not follow the incoming shortwave radiation and corresponds to albedo values of almost up to 50%. What is the reason for this?

#### Answer:

Referee #1 had a similar question. Therefore, we reiterate our statement here:

We do not know the reason. It is not a consistent feature. However, the occurrence of spikes of SWUP in MM5 are accompanied by concurrent underprediction of SWDOWN in MM5, often these underpredictions in SWDOWN are of the similar magnitude as the overpredictions in SWUP, which points to a deficient energy distribution in MM5. Regardless of the nature of this deficiency, its impact is seen in corresponding underpredictions in the fluxes of sensible and latent heat flux (Figure 5), and to some extent it is also reflected in the temperature time series (Figure 3). The significant PBL underprediction in MM5 is likely due to the underprediction of sensible and latent heat in MM5 on that day (Figure 5), which in turn might be associated with the deficient simulation of shortwave radiation by MM5 on the same day. We added corresponding findings in the text in the section "Incoming/outgoing shortwave radiation".

\* I find the discussion of nighttime vs daytime behaviour of the solar radiation not useful. During nighttime, there is no solar radiation so it is trivial that models work better during nighttime (mostly 0 W/m2) than during nighttime.

#### Answer:

We removed references to nighttime data, both in text and table 5.

\* Line 399-401: The models overestimate solar daytime insolation. Where is the underestimation you use to explain a cool bias?

## Answer:

We removed this part.

\* The discussion of the sensible heat flux focusses much on the bias suggesting a better performance of MM5, while r2 and the RMSE are clearly worse for MM5. Please extend your analysis. Check the values in the sentence starting in line 447.

#### Answer:

Thank you for this comment! We checked and corrected the values starting in line 447 and added a statement with regard to r2 and RMSE.

Minor issues:

\* define CST

## Answer:

#### Corrected.

\* Still, I think the outgoing longwave radiation can be easily calculated from the other model output, so at least the sentence in line 338 ("[outgoing longwave radiation] cannot be adequately compared with surface observations") has to be formulated much weaker.

# Answer:

We rephrased this statement.

\* Line 368: "The bias is positive" or "the flux is overestimated".

# Answer:

Corrected.

\* Line 442: WRF had higher overall, nighttime, and daytime biases than MM5E.

Answer:

Corrected

\* Line 520: replicate

Answer:

Corrected