

The authors are grateful to the reviewer's valuable comments that improved the manuscript.

Anonymous Referee #1

A very complete set of observations at the surface and upper air are used to evaluate three different land-surface parameterizations implemented in the mesoscale model WRF in the urban area around Houston. The intercomparison is very complete and well written, but it fails short in provide clear explanations and potential improvements in the parameterization performance. In my opinion, these are the three main aspects that require further elaboration and clarifications:

a) Nocturnal boundary layer. At different parts of the manuscript (abstract, section 4.3, conclusion) is found that the modeled night conditions are characterized by excessive turbulent mixing. This is a key aspect of the research since it involves the interaction between the land surface representation with the boundary layer scheme. In my opinion, this research needs to include how to improve this fundamental aspect in WRF. There has been already previous research on this subject and I will encourage the authors to implement it and discuss if it improves their results and the model performance at night (see for instance Steeneveld et al., 2008, Journal of Applied Meteorology, 869-887).

We tested the Mellor-Yamada-Janjic (MYJ) scheme, which is a local TKE-based scheme, to address the performance of urban parameterizations with another turbulence scheme as well as to investigate nocturnal representation of PBL structures. Fig. A1 shows vertical profiles of the observed and simulated wind speed and directions at the La Porte site around midnight (00:30 LST) on 12-14 August when the strong nocturnal low-level jets (LLJs) occurred. The flat vertical profile of wind speed in the YSU simulation explains the overestimation of surface wind speed (Fig. 6) and the underestimation of wind speed in upper layer (Fig. 11) during the nighttime on the days with LLJs. This fact was described in the manuscript. The sensitivity test with the MYJ scheme shows better agreement with the observed wind fields, specifically in vertical gradient of wind speed, than that of the YSU scheme (Fig. A1), by which comparisons of surface and upper layer wind speeds during the nighttime (Figs. 6 and 11) were also improved on the days (not shown). However, the use of MYJ scheme does not change our conclusions of this study.

Recent studies (Lee et al., 2006; Steeneveld et al., 2008) on stable boundary layer (SBL) parameterization in mesoscale meteorological models were mentioned in the revised manuscript because of their potential for improving nocturnal boundary layer simulations. The implementation of the SBL parameterizations suggested by Steeneveld et al. (2008) to WRF may be considered for future work.

b) Interaction mesoscale phenomena and boundary layer dynamics. In section 4.3 it is mentioned that the model is able to reproduce well the influence of the sea breeze. The influence of the sea breeze on the development of the boundary layer is very vaguely described. I have the following question related to this issue: What is the impact of the sea breeze on the different land-surface schemes understudy? Do all react in a similar way? Is the boundary layer still growing? I believe that in addition of the sea breeze, there is a urban breeze driven by the different thermal capacities. What is the influence of the breeze on the boundary layer dynamics? In that respect, it will be very interesting to determine the impact of horizontal

resolution (for instance by doing an extra numerical experiment imposing 2 x 2 km²) in their results (see also point 6 in the specific comments).

As briefly described in the manuscript, the model simulations well reproduced sea/bay breezes intrusion based on the evaluation of wind profilers at La Porte (Figs. 10 and 11). To address reviewer's concerns, we examined the sea breeze development in detail according to different land surface parameterizations tested in this study. Fig. A2 shows the observed and simulated surface wind fields in the afternoon on 16 August when the sea breeze was well developed. It also shows realistic reproduction of sea/bay breezes in both the LSM and UCM simulations. One of distinctive differences between the LSM and the UCM simulations is that sea/bay breezes simulated by the LSM are stronger than those by the UCM mainly as a consequence of larger sensible heat flux in LSM than in UCM.

For the issue on the spatial resolution, instead of looking into heterogeneity within urban patches, we attempted to understand the local sea-land circulations in Houston-Galveston area in which the model domain is characterized by major land-use categories such as urban, natural, and water areas (see Fig. 2). It is also expected that morphological variability in urban patches can have little impact on the simulation of local sea-land circulations. The authors think that the configuration with a 4 km horizontal resolution is sufficient to simulate local circulations. Within the Houston metropolitan area, most of the grids are characterized as urban land use. Thus, increasing the model horizontal resolution to 2 km does not change the land use characteristics.

What is the impact of the sea breeze on the different land-surface schemes understudy?

Different thermal forcing from land-surface schemes affects the development of sea/bay breezes. As described above, the stronger thermal contrast in the LSM simulation induced stronger sea/bay breezes than the UCM simulation as a dynamic adjustment of the thermal forcing. This was discussed in section 4.3 in the revised manuscript.

Do all react in a similar way?

Overall pattern of sea/bay breezes development is similar between the simulations, but the magnitudes in strength and depth are different.

Is the boundary layer still growing?

Yes. Fig. 12 shows that the boundary layer height grows and decays in a similar pattern for all the simulations through the day.

What is the influence of the breeze on the boundary layer dynamics?

Boundary layer dynamics in Houston-Galveston area can be affected by multi-scale forcing such as synoptic, local sea-land breeze, and urban breeze. To examine boundary dynamics and structure, 3-dimensional observations of temperature, moisture, wind, and boundary height are necessary. More intensive observations than those in TexAQS 2006 will be helpful to address this research topic.

c) Aerosols. There is hardly any discussion on the effects of aerosols in the development of the boundary layer. I should expect that in urban areas they exert an influence on the radiation and therefore on surface forcing, but also in stabilizing the upper part of the boundary layer by absorption and scattering (see for instance Yu et al., 2002, Journal of Geophysical Research 107, D12,4142). How is the interaction between the radiation schemes and the land-surface model understudy? Does it influence the boundary layer characteristics in the studied situation? The reviewer's comment is right. We did not discuss aerosol impact (mainly through a radiative transfer process) on the boundary layer evolution in this study. According to a few journal

articles including Yu et al. (2002), the aerosol can affect the PBL dynamics. Wang et al. (2004) showed that the aerosol effect can reduce the surface radiative energy by 30–40 W m² during the day and increase by 10 W m² during the night, resulting in near surface air temperature change up to ±0.5°C. Based on this result, the PBL height may be reduced roughly by about 125 m using Eq. (6.14) (encroachment rate) in Garratt (1992). For the calculation, $H = 40 \text{ W m}^2$, $h = 1500 \text{ m}$, $\Gamma = 0.005 \text{ K m}^{-1}$. Judging from this simple calculation, it is expected that the consideration of aerosol effects do not change the conclusions of this study. This topic is very interesting and is worth more investigation. According to the previous studies, the aerosol effects can depend on several factors such as aerosol optical properties, surface radiative properties, and local time.

The authors hope to address the aerosol effects in Houston in our ongoing coupled meteorology-chemistry simulations using the WRF-Chem model. This discussion was included in section 5 in the revised manuscript.

References (not in the revised manuscript)

Garratt, J. R. (1992), *The atmospheric boundary layer*, Cambridge University Press, 316 pp.
Wang, J., U. S. Nair, and S. A. Christopher (2004), GOES 8 aerosol optical thickness assimilation in a mesoscale model: Online integration of aerosol radiative effects, *J. Geophys. Res.*, 109, D23203, doi:10.1029/2004JD004827

Specific comments

1- What are the assumptions behind equations (4) and (5)?

When a model grid includes both urban and natural surfaces with coverage fractions, the resultant flux for the grid cell is calculated as an average with their fraction. This is a methodology commonly used in mesoscale and global models for better representation of sub-grid heterogeneity, so called, tile (mosaic) approach. Avissar and Pielke (1989) and Chen et al (2004) can be referred for this approach (see references in the revised manuscript). No interaction between the two components is assumed in calculating the radiative and turbulent fluxes. The same is applied to Eqs. 4 and 5 of the manuscript. Avissar and Pielke (1989) was added to the references in the revised manuscript.

2- How is the entrainment flux estimated in equation (6)? Is this term included in all the thermodynamic variables?

3- Shear is a very local process. In equation (8), it is only included the shear at the surface. What about the contribution of shear in the inversion?

4- In equation (8), why the convective scale is multiplied by the von Karman constant?

5- In equation (9), I think they should use the virtual potential temperature

The section 2.3 was removed in the revised manuscript according to the suggestion of the reviewer 2. For the reviewer's questions 2-5, two papers (Hong et al. 2006; Hong 2010 in references in the revised manuscript) can be referred. These are original papers regarding the YSU PBL scheme.

6- In the majority of the figures, they are comparing a single point measurement with the 4 x 4 km² grid. The urban area is highly heterogeneous and I think they need to justify the assumptions in comparing observations and WRF model results.

The authors agree with the reviewer because the subgrid-scale heterogeneity in urban patches can directly affect the measurements below the roughness sublayer which is, in general, defined as 2-3 times of mean obstacle height. However, this problem may be inevitable in comparing observations with mesoscale model simulations. The authors would like to point out that the

surface measurements used in this study are hourly averaged data, thus having large footprint. Therefore, the problem caused by subgrid heterogeneity may be alleviated to some extent. We also emphasize that the measurement data sets used in this study are made from the multiple platforms during the intensive field campaign, which made these data unique and invaluable. The fact that surface measurements are hourly averaged data was explicitly described in section 3 in the revised manuscript. Several reference articles regarding the measurements were included in the revised manuscript. Those references describe more details including uncertainties and instrumentation.

The authors are grateful to the reviewer's valuable comments that improved the manuscript.

Anonymous Referee #2

This paper describes the performance of an urban canopy parameterization using measurements collected during a recent air quality field campaign in Houston. Overall, the paper is well written and the results are clearly presented. However, there are several points that need to be clarified to improve the manuscript.

Major Comments:

1) The abstract could be improved by including some specific numbers on the performance of the urban surface parameterizations. This material would add to the length, but some details (e.g. specifics on the measurement platforms) that are not needed could also be removed. It seems that one of the advantages of testing an urban canopy parameterization for a field campaign period is availability of a wide range of data that is normally not available. Yet, the abstract is worded to basically just list the instruments used even though the most important aspect is having a wide range of measurements aloft that characterize the boundary layer and not just the surface quantities.

The abstract was rewritten according to the reviewer's suggestion.

2) Figure 3 shows a commercial/industrial region, but Figure 2 has no grid cells with this type of urban classification (denoted in black). Figure 3 shows that Houston is either low or high density residential. Yet subsequent figures show results for the commercial/industrial regions. Either the plot is mislabeled or I am missing something. The test starting on page 24052 then goes on to show results of the model for the commercial/industrial regions. Please clarify.

The urban classes of 'commercial/industrial', 'high density residential', and 'low density residential' are shown as 'dark brown', 'red', 'light red' colors in Figure 2, respectively. Figure 2 was redrawn for clarity in the revised manuscript by reducing the symbol size for each surface station.

3) It would also seem that the performance of the model depends on both the urban canopy and boundary layer parameterizations since they are coupled. It would have been useful to see how UCM performs with another PBL parameterization. At least the authors should comment somewhere in the text regarding what they expect the performance would be had the urban canopy parameterization was used with a different boundary layer parameterization. Another area of discussion are the parameters used for both the LSM and UCM. It was shown that the LSM could be made somewhat better by adjusting parameters. Presumably the UCM could be similarly "tuned"; however, adjusting parameters needs to be based on some type of data rather than just trying to fit the prediction of atmospheric quantities.

We conducted two simulations (LSM and UCM) with the MYJ scheme, which is a local TKE-based scheme. The MYJ simulations tend to show cooler temperature and lower ABL height than the YSU simulations, which is an identical result in previous studies (e.g., Hu et al., 2010). Fig. A3 shows statistical comparison results from the simulations with the MYJ scheme. With the MYJ scheme, the UCM performs better than the LSM in simulating boundary layer height. Similar to the case of the YSU scheme, the improvement is attributed to the realistic simulation of surface energy balance. The authors do not put the model's performance with the MYJ scheme to focus on evaluation of the urban surface parameterizations.

The parameters used in urban surface parameterizations were independently determined based on available data before conducting the model simulation, which are explicitly given in Tables 2

and 3 in the manuscript for reproducibility. Additional references (Liu et al., 2006; Chen et al., 2010; Loridan et al., 2010) were included in the revised manuscript. Those papers describe the parameters of the urban surface parameterizations.

Reference (not in the revised manuscript)

Hu, X.-M., Nielsen-Gammon, J. W., and Zhang, F.: Evaluation of three planetary boundary layer schemes in the WRF model, *J. Appl. Meteorol. Clim.*, 49, 1831-1844, 2010.

4) Some additional discussion is needed on measurements for two key analyses made in this paper. First, surface flux measurements were not collected in Houston proper. I understand that modelers use what is available, but they need to discuss that data in Houston would have provided the most useful data to evaluate the parameterization.

Second, discussion is needed regarding the lack of boundary layer depth and other boundary layer properties at night, which are critical to evaluate the performance of the model at this time. Or am I missing some information regarding the measurement strategy during the field campaign?

Unfortunately, we do not have the measurement of surface energy balance fluxes in Houston. Because the Brenham site is characterized as low density residential urban land use, the data at this site were used to evaluate the simulated urban surface energy balance. The simulated surface energy balance fluxes for low-density residential areas in Houston are similar to those at the Brenham site. In addition, the turbulent energy partitioning (daytime Bowen ratio of about 1) in Houston in the UCM simulation is similar to recent measurements at an urban site in Houston during summertime (Boedeker et al., 2008). Discussions on the surface energy balance fluxes were added in section 4.2 in the revised manuscript.

Because of the limited data sets, the model performance on the nighttime atmospheric boundary layer in Houston was evaluated using surface meteorological stations only. More evaluation of the model nocturnal urban boundary layer will be important. This discussion is included in section 5 in the revised manuscript.

5) The last paragraph in the summary section seems out of place. Since the paper is focuses on urban canopy parameterizations, this material should be moved earlier, so that the paper ends on conclusions regarding the primary focus of research. Also, the implications of their findings on the urban canopy parameterization should be discussed. For example, how will these results affect air quality predictions? Or are the changes (that are modest for some parameters) going to be significant in terms of other applications?

The last paragraph was deleted in the revised manuscript. Instead, brief discussions on limitation, implication, and future work were added in the revised manuscript.

Secondary Comments:

Title: "in the WRF model" could be removed from the title since it is merely the host model for the urban surface parameterizations that are the focus of this study.

There exist several urban surface parameterizations with different complexities (Grimmond et al., 2010, see references in the manuscript). Therefore, the authors think it would be better to keep the phrase as it is to clarify the fact that the tested urban surface parameterizations are implemented in the WRF model.

Page 25034, line 24: Change "Model showed" to "The model showed".

The sentence was deleted in the revised manuscript.

Page 24036, line 25-20: This sentence is a bit misleading. Some of the previous studies were performed before more sophisticated urban parameterizations were available or fully tested in

mesoscale models. The sentence makes it sound like a conscience decision not to use a more advanced urban canopy parameterization.

The sentence was rewritten to avoid misleading as follows.

There were several numerical simulations conducted for the Houston area (e.g. Bao et al., 2005; Fast et al., 2006; Cheng and Byun, 2008), but the performance of an urban canopy model (UCM) in predicting the urban boundary layer evolution has not been tested with high fidelity.

Section 2.1, first paragraph: Most of this text seems unnecessary. The authors should just include a sentence with a reference for the WRF model that contains these details and the version number, and combine that sentence with the next paragraph. Readers unfamiliar with WRF will not know what the “Advanced Research version” is compared to the other version that is available.

The paragraph was removed in the revised manuscript following the reviewer’s suggestion. A short description of the WRF model was included in section 1 with a reference for the WRF model (Skamarock et al. 2008).

Page 25038, line 23: 35 levels is rather coarse. Please include some information on how many nodes are located within the lowest few kilometers of the atmosphere. Since strong vertical gradients in temperature, humidity and wind can occur this region, particularly associated with the sea/bay breeze, it is important to know whether the current model configuration is sufficient to resolve the phenomena the model is intended to simulate.

15 full sigma levels were included below 2 km. This description was included in section 2.1 of the revised manuscript. This vertical grid configuration is sufficient to resolve the sea/bay breeze intrusion as well as low-level jets. Please see Fig. 10 in the manuscript and Fig. A1.

Section 2.2: It would be useful to provide sub-section headings for the LSM and UCM descriptions to better differentiate the two.

As suggested, the two urban surface parameterizations were described in separate subsections (Sect. 2.2.1 and 2.2.2) in the revised manuscript.

Page 25040, line 23: Are there other studies that evaluate the UCM parameterization in WRF? It would be useful to cite those if possible. Also, has the UCM parameterization in WRF been applied to the Houston area?

A reference (Chen et al., 2010) regarding WRF urban modeling system was included in the revised manuscript for readers. The paper summarizes the model development, evaluation, and some applications (including an application to Houston) for the WRF urban surface parameterizations. To authors’ knowledge, this paper may be the first peer-reviewed article evaluating the UCM for Houston with the extensive observations.

Page 24053, lines 21-22: What were the vegetation fractions based on?

The values for three urban classes were estimated based on the National Land Cover Dataset and Burian et al (2003) (see references in the manuscript). It was explicitly described in the revised manuscript.

Section 2.3: Much of the discussion on the details of the YSU scheme seem unnecessary in this paper since the focus is on urban canopy parameterizations. There is only one other short section that discusses how this scheme affects the simulation results, in terms of the nocturnal boundary layer. It would seem that only the introductory material and the last sentence is important.

Section 2.3 in the manuscript was removed, and a short description of the YSU scheme was added in section 1 in the revised manuscript.

Page 25044, line 20: Why only use 10 stations and why these particular ones?

All available data at 24 surface sites were included in the revised manuscript. Figures 6 and 7 were redrawn based on the available surface stations.

Page 25044, line 24: Change “in the northwestern area and northeastern area from Houston” to “northwest and northeast of Houston”. Please comment that it would have been useful to have surface flux sites in Houston where the importance of the urban canopy parameterization is the greatest.

The sentence was corrected in the revised manuscript. A brief discussion on surface energy balance flux comparison was added in section 4.2 in the revised manuscript.

Page 24045, line 6: Sentence “Figure 2 shows : : :” seems out of place. Figure 2 shows more than just the location of the surface site. Perhaps moving this phrase earlier in the paragraph.

It was corrected following the reviewer’s suggestion.

Page 25045, second paragraph: Are there are any papers that can be cited regarding the aircraft flights and instrumentation?

Three references (Alvarez et al., 2008; Langford et al., 2010; Senff et al., 2010) were additionally included for TOPAZ instrument and aircraft flights in the revised manuscript. Overall information on TexAQS 2006 including aircraft flights can be referred by Parrish et al. (2009) that is included in the manuscript.

Page 25046, line 4: As stated in my major comment, I don’t see which model cells are classified as commercial/industrial region.

The commercial/industrial area is represented by ‘dark brown’ in Figure 2.

Page 25046, line 10: Change “formation” to “characteristics” or “structure”.

As suggested, “formation” was changed to “characteristics” in the revised manuscript.

Page 25046, lines 21-22: This sentence correctly shows that the LSM does better than the UCM for this classification, yet it seems to be downplayed in the manuscript. Overall UCM does better than LSM overall, but still requires some improvement.

Agreed. The observed peak temperatures in commercial/industrial urban land use are better reproduced by the LSM simulations. The UCM does need further improvement. These discussions were added in section 4.1 in the revised manuscript.

Page 25048, line 7: I cannot really tell which stations in Figure 2 are the wind stations.

All available 24 stations were used for both wind and temperature fields in the revised manuscript. Therefore, symbols (pink dots) for 10 selected wind stations in Fig. 2 were removed in the revised manuscript.

Page 25049, line 3: This sentence is awkward. Change “northeastern forest area from Houston” to “forest northeast of Houston”.

The sentence was corrected in the revised manuscript as suggested.

Page 25049, line 14: Why show these days as opposed to other days? Please clarify in the text.

The local wind fields of the first three days (12–14 August) of the simulation period can be characterized by southerly winds throughout the days with strong nocturnal low-level jets, and the last three days (15–17 August) can be characterized by the development of sea/bay breezes under a relatively weak synoptic forcing. The two days (12 and 16 August) selected

explained both the local wind patterns, respectively. A brief description on the local wind patterns of the selected days was included in section 4.3 in the manuscript.

Page 25050, line 1-3: What is described is difficult to see in the plot. Also, it would be useful to reiterate that the stronger land-sea thermal contrast is related to changes in temperature over Houston (Table 4, etc.).

Fig. 10 was redrawn with including the difference plot in the revised manuscript. The surface temperature and surface energy balance comparisons were reiterated for clarity.

Page 25050, line 14: The surface fluxes in Fig. 8 are not over the Houston metropolitan area as stated in text. Based on the figure I would say the site is in a small city outside of the metropolitan area.

The sentence was rewritten for clarity by removing “shown in Fig. 8” in the revised manuscript.

Page 25050, line 17: Change “for lower” to “in the lower”.

The words were corrected in the revised manuscript.

Page 25050, line 24: The predicted boundary layer depths at night are lower than 300 m for some nights, so the explanation provided may not entirely explain the problem with the nocturnal low level jet. The relevance of the whole discussion regarding the wind evaluation is not clearly articulated. I would expect that the urban boundary layer would have a larger impact on surface temperatures and boundary layer temperature profiles (and consequently PBL depth), than on winds. The results seem to confirm this, and one would not expect the urban canopy parameterization to change the predictions of the low-level jet.

Fig. A4 shows the comparison of observed and simulated wind fields at 200 m AGL at the La Porte site. Overall results are similar to the comparison of wind fields at 300 m (Fig. 11). Fig. A1 shows vertical profiles of the observed and the simulated wind speed and directions at the La Porte site around midnight (00:30 LST) on 12-14 August when the strong nocturnal low-level jets occurred. This figure shows more clearly that the flat vertical profile of wind speed in YSU simulation explains the overestimation of surface wind speed (Fig. 6) and underestimation of wind speed in upper layer (Fig. 11) during the nighttime on the days with LLJs. This fact was described in the manuscript. As the reviewer mentioned, the prediction of nocturnal LLJs is more related to the PBL scheme. The sensitivity test with the MYJ scheme shows better agreement with the observed wind fields, specifically in vertical gradient of wind speed, than that of the YSU scheme (Fig. A1). Consequently, with MYJ scheme, the simulations of surface and upper layer wind speeds during the nighttime (Figs. 6 and 11) were also improved on the days (not shown).

Page 25051, section 4.4: Please describe the uncertainties associated with the estimates of PBL depth derived from the aircraft. Has this technique been described in a paper that can be cited?

Some references (White et al., 1999; Alvarez et al., 2008; Nielsen-Gammon et al., 2008) were included in section 3. Those describe instrumentation, ABL estimation technique, and associated errors and potential biases in the lidar ABL retrieval.

Page 25053, line 4: Would be useful to add what the wind direction is (presumably westerly) for the warm advection. One has to piece this information together from multiple plots to visualize what is going on. Are the winds over-predicted for the warm advection?

The wind direction (westerly winds) was mentioned in the revised manuscript. Unfortunately, there were no measurement data needed for thermal advection estimation.

Page 250053: I do not think how water temperatures were prescribed in the model. Databases may have good estimates for ocean temperatures, but the estimates over bay may be off. Some additional discussion on the bay temperatures is needed in the model configuration section. [Sea surface temperatures in Galveston Bay and the Gulf of Mexico are taken from the GFS data, and remain constant in time during the simulation period. This sentence was added in section 2.1 in the revised manuscript.](#)

Page 25055, line 14: Change “great potential to accurately” to “great potential to more accurately”. The application in this study certainly shows an improvement, but there is room for more improvement. [It was changed in the revised manuscript as suggested.](#)

Figure 2: I do not see “dots” in this figure. There are 2 solid circles for the surface flux sites, but are they different than dots? [Figure 2 was redrawn with a clear figure caption in the revised manuscript.](#)

Figure 8: Explicitly state that the site is a low density residential area. [It was explicitly described in the figure caption in the revised manuscript.](#)

Figure 9: Explicitly state that the site is in a forest region. [It was explicitly described in the figure caption in the revised manuscript.](#)

Figure 10: It is hard to tell the differences in the plot. I suggest changing it to a color plot, with color denoting speed and using the same vector length for the wind direction. [The figure was redrawn including the difference plot in the revised manuscript.](#)

Figure 15: Include specific time for (a) and (b). [Specific times were included in the figure of the revised manuscript.](#)

Figure 16: Please label Houston Ship Channel and Barbour's Cut. Most readers will not know where these locations are. [The figure was redrawn including the labels of the locations.](#)

The authors are grateful to the reviewer's valuable comments that improved the manuscript.

Anonymous Referee #3

The study presents a very extensive evaluation of two of the current urban parameterizations implemented in WRF, while introducing a third alternative based on several parameter changes. Model runs for a 6 day period corresponding to the TexAQS 2 measurement campaign are performed to enable comparison with observations of 2 m temperatures, 10 m wind speed and direction, surface energy balance fluxes, wind profiles and atmospheric boundary layer height. Results underline some improvement when using the more sophisticated Urban Canopy Model (UCM) scheme. The major reason identified is the ability of the UCM to represent (1) the impact of sub-grid scale vegetation on energy partitioning and (2) heterogeneities between urban grid cells. The study is very well presented and results are of great interest for the urban-WRF user community. Several clarifications are however needed to facilitate the understanding of the model settings (and their reproducibility). Some adjustments on the "surface energy balance fluxes" section (4.2) are also suggested. I would therefore suggest acceptance of the manuscript after some "minor revisions".

General comments:

* Given that the version of WRF used in the study is very recent (WRF 3.1, March 2009) the results presented here are of great interest for the WRF/urban user community. The urban modelling options were revisited in this 3.1 release (Chen et al., 2010) and this paper is probably one of the first evaluating the updated system with such an extensive set of observations. It is however not always very clear which schemes/settings are used for the simulations. I therefore suggest more extensive references to the paper by Chen et al. (2010) which provides an overview of WRF 3.1 and its urban modeling options.

[Based on the reviewer's comments, three references \(Chen et al. 2010; Loridan et al. 2010; Liu et al. 2006\) that describe the urban parameterizations were included in the revised manuscript.](#)

1) Some references are needed with regards to the "original" urban LSM: is it the "bulk urban parameterization" as described in Chen et al. (2010)? In that case a reference to the work of Liu et al. (2006) is probably needed.

[Yes. Liu et al. \(2006\) was included in section 2.2 in the revised manuscript.](#)

2) A similar comment holds for the UCM; the author only reference Kusaka et al. (2001) but the version of the UCM implemented in WRF 3.1 has been modified (see Loridan et al., 2010). The modifications mainly focussed on a simplification of input parameter list which led to the one presented in Table 3. Most of these modifications are actually listed at the end of section 2.2 (i.e. roughness length and displacement height calculated via morphometric method, internal calculation of aspect ratios and sky view factors) - a reference to the Loridan et al. (2010) paper would be appropriate.

[Loridan et al. \(2010\) was included in section 2.2 in the revised manuscript.](#)

3) The UCM from WRF 3.1 has the ability to include an anthropogenic heat contribution to the sensible heat flux. Has this option been activated in the runs presented? In any case can the author comment on their choice (i.e. if included: what profile/magnitude? and if not: is it neglected? or left out for better comparison with the LSM?).

[No. The anthropogenic heat flux was not included in the UCM simulation in this study. It was briefly discussed in section 5 in the revised manuscript.](#)

4) In several places the authors state that “the urban parameterizations implemented in WRF” are evaluated. I believe it is important to specify that only two of the three options implemented in WRF are evaluated: the more sophisticated Multi-layer urban canopy model from Martilli et al. (2002) is also available as an option in WRF v3.1 (see Chen et al., 2010) but not used in the study.

A brief description on the three available urban surface parameterizations in WRF was added in section 1 in the revised manuscript.

* The comparison of modelled and observed surface energy balance fluxes is very interesting and useful to understand the differences in performance from the 3 schemes. However, more details are needed with regard to the measurement system (unless a reference to another study where it is fully described is provided): for instance information on the height at which the fluxes are measured would be of interest to compare to the height of the lowest WRF level (20 m) at which they are simulated. Some comments on the measurement footprint, and comparison of its scale to the 4 km x 4 km horizontal resolution would also improve the evaluation. More generally, is there a reference for the TexAQS 2 campaign?

More description about the surface flux measurements was added in section 3 of the revised manuscript. Two references (Zamora et al., 2003, 2005) describing the measurement system were included in the revised manuscript. The manuscript had a reference (Parrish et al., 2009) for the TexAQS 2006 field campaign. The model-observation comparison in turbulent heat fluxes assumes that both the simulated and measured turbulent fluxes represent the characteristics in the inertial sublayer (Kastner-Klein and Rotach, 2004; Grimmond, 2006). This was stated in section 3 of the revised manuscript. In addition, the surface flux data we used were calculated with a time window of every 30 minutes, having enough footprint to compare with our simulations with a 4-km horizontal resolution. The flux data intervals were also included in section 3 in the revised manuscript.

* The authors compare modelled and observed incoming radiative fluxes but I believe the outgoing components and/or the net all wave radiation would teach more about possible model deficiencies. Were these fluxes observed? Could they be included in the analysis? In particular the ability of the UCM to represent radiation trapping within the canyon should be better reflected in the outgoing components.

As the reviewer’s comment, the observed net all-wave radiation was compared with the simulation for both Brenham and Kerbyville sites (Figs. 8 and 9). According to this, Figs. 8 and 9 were redrawn and comparison results for radiation in section 4.2 were revised.

* An interesting feature from Fig. 8 and 9 is that the UCM is the only scheme catching the dip in incoming shortwave radiation on the 12th of August when more cloud cover was present. Could the author comment briefly on that aspect? I suppose it is directly linked to the fact that the UCM generates higher turbulent heat fluxes but any further comment would be of interest.

Rapid reduction in the simulated radiation was caused by simulated clouds. As shown in Fig. 8, the UCM simulation has smaller (larger) sensible heat flux (latent heat flux) than the LSM simulation in urban areas, probably being able to alter dynamic and thermodynamic structures and form clouds at the top of convective boundary layer. On 12th August, clouds near ABL top formed above the flux sites. The cloud fields were also simulated by different urban surface parameterizations, but the spatial locations of the clouds are slightly different among the simulations. However, it is not conclusive that the UCM is better than the others in terms of the cloud simulations.

* The English in a number of places needs to be improved
English was further improved in the revised manuscript.

* There are a lot of places where references are needed
Proper references were included in the revised manuscript. Please see the revised manuscript and the replies to the reviewers' comments.

* Acronyms should be given in full when they first appear.
It was checked in the revised manuscript.

Specific comments:

p25036, l11: "the urban parameterizations implemented in WRF" → only 2 out of 3, see comment above.

This issue was mentioned in section 1 in the revised manuscript.

P25038, l7: "The LSM" → "Noah LSM" is the complete name of the scheme.

It is clearly stated that "the LSM" denotes "the Noah LSM" in the manuscript. Therefore, "LSM" was used throughout the manuscript for simplicity.

P25038, l23: The authors refer to the "vegetation fraction" but I think what is meant here is "green vegetation fraction" (or shading factor as it is called in WRF: parameter SHDFAC). This distinction is important in order to avoid confusion with the "vegetation fraction" mentioned in p P25042, l21 which refers to (1-fU) in eq. 4. One is a plan area of vegetation cover (i.e. 1-fU) and the other is defined as the part of the vegetation that is photosynthetically active (green vegetation fraction). Same comment applies to Table 2. I think the distinction is important because as it stands it appears that the LSM is actually ran using a tile approach (i.e. as for eq. 4).

It was corrected throughout the revised manuscript (including Table 2) by using 'green vegetation fraction' instead of 'vegetation fraction' for the bulk urban parameterization.

P25038, l23: the original value for this green vegetation fraction is said to be 10% in the text but is 5% in Table 2. From WRF tables it seems to be 10%

The value in the text was corrected as 5%, identically in Table 2, in the revised manuscript.

P25039: reference to Liu et al. (2006) needed somewhere.

Liu et al. (2006) was included in section 2.2 in the revised manuscript.

P25040, l11: → "green" vegetation fraction

This was corrected as suggested.

P25042, l11-l18 → this section is a description of the Loridan et al. (2010) work; a reference to that study is needed.

Loridan et al. (2010) was included in section 2.2 in the revised manuscript.

P25043, l17: "determined a method" → "determined from a method"

This sentence (section 2.3 in the submitted manuscript) was removed in the revised manuscript following the suggestion of the reviewer #2.

P25044: Observations section. Is this the first time the TexAQS 2 campaign is presented? If not I think a reference to where more detailed can be found is needed.

A reference for TexAQS 2006 (Parrish et al., 2009) was included in the manuscript and it is shown again at this section for clarity.

P25049, l4,5: this is an interesting point, could the authors provide more details? Is warmer air being advected from the urban cells? Is there a dependency on wind direction?

The nocturnal boundary layer in Houston on 12 August can be characterized by nocturnal low-level jets. The overestimation of nocturnal ALB height by model simulations is largely associated with the performance of turbulence parameterization (YSU scheme) under the condition of low-level jets as shown in Fig. A1. The urban impact on the nocturnal boundary layer evolution was not significant even though southwesterly winds existed on the day. The sentence was rewritten clearly in the revised manuscript.

P25048: as mentioned in the general comments, analysis of upward components of the radiative fluxes and/or net all-wave radiation would help to evaluate the ability of the schemes to simulate the radiation balance (see for instance Grimmond et al. 2010).

Following the reviewer's comment, the observed net radiation was compared with the simulations. Refer Figures 8 and 9 in the revised manuscript.

P25052 and Fig 15: simulated → simulated from UCM only should be clarified

It was clearly described in the figure caption.

p25054, l3: "this may be due that ..." → "due to the fact that"

It was corrected in the revised manuscript.

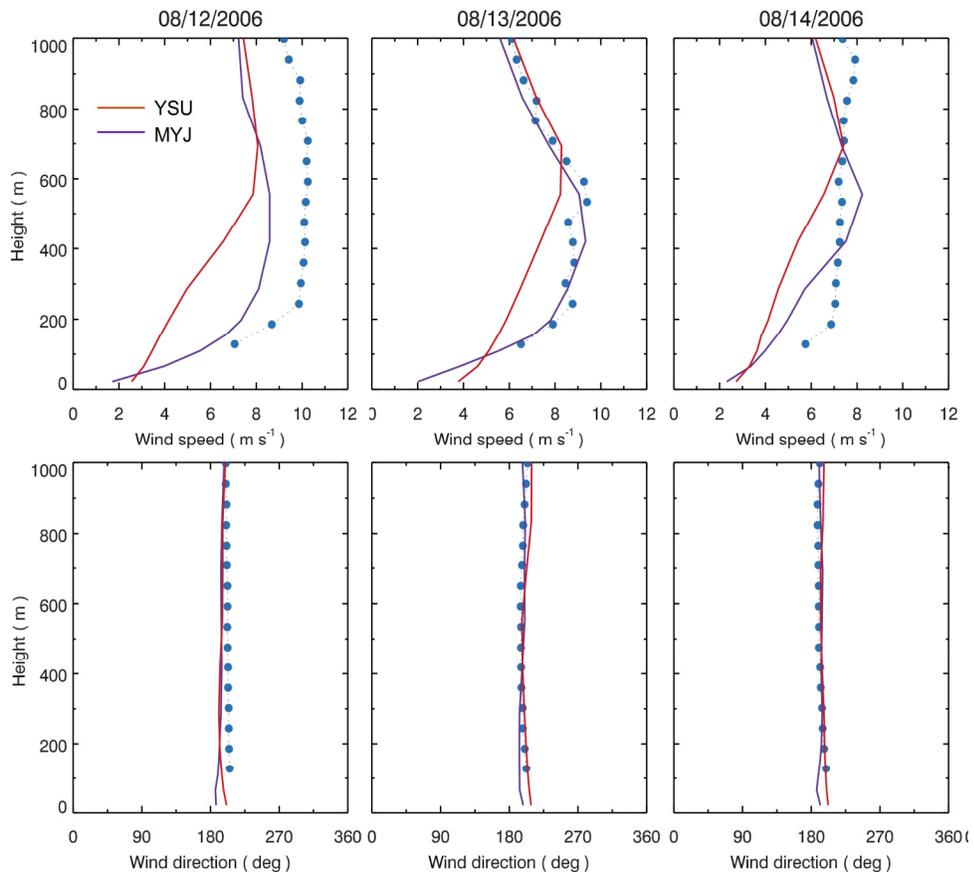


Fig. A1. Vertical profiles of observed and simulated wind speeds and directions at La Porte at 00:30 LST on 12-14 August.

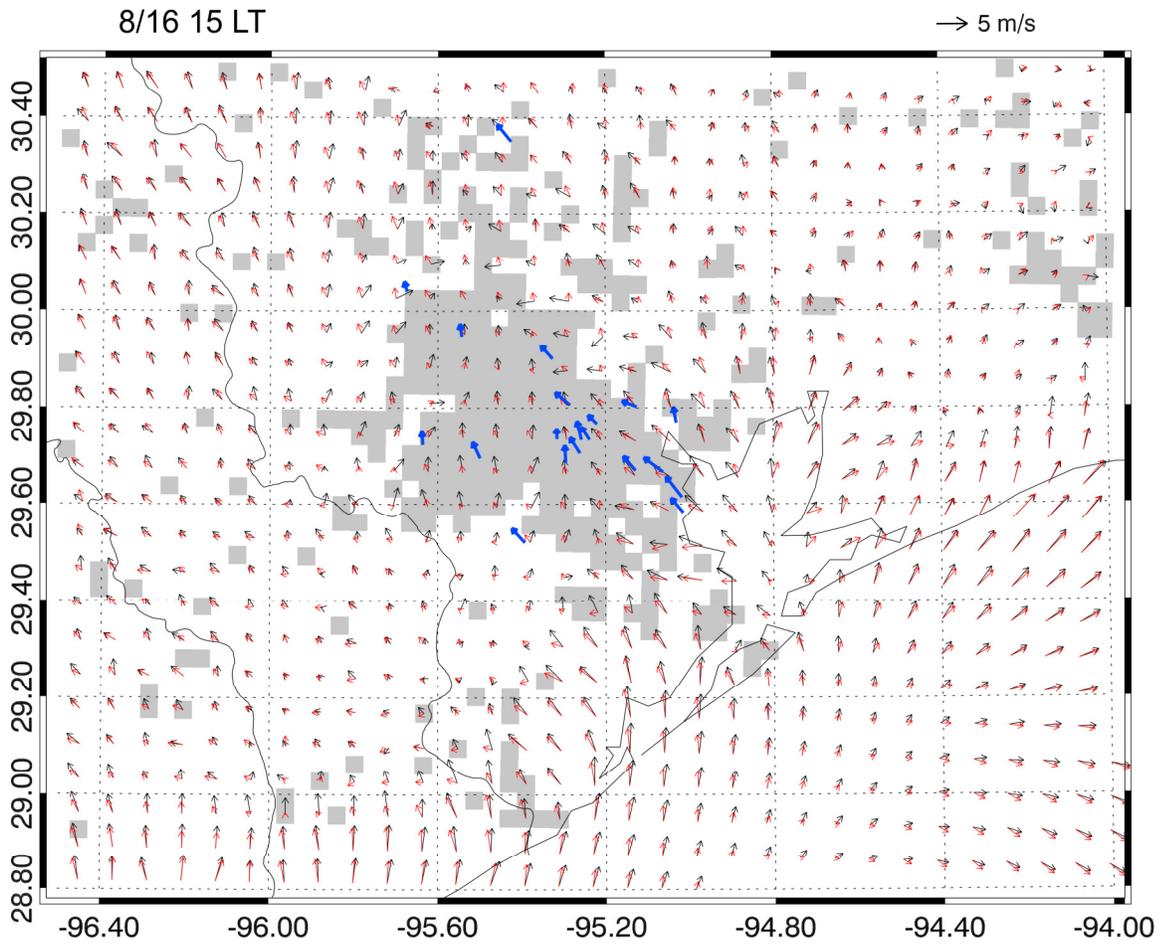


Fig. A2. Observed and simulated surface wind fields in 15:00 LST on 16 August. Blue arrows are observations, and black and red arrows represent the original LSM and the UCM simulations, respectively.

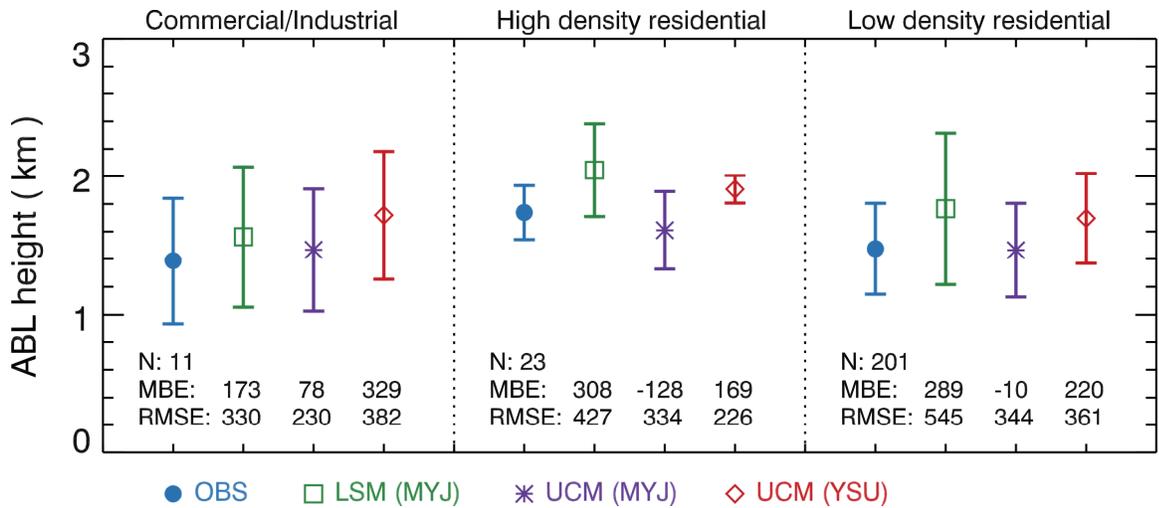


Fig. A3. Statistics from observed and simulated ABL heights in urban patches. The same as in Fig. 14 in the manuscript except that two simulations with the MYJ scheme were used.

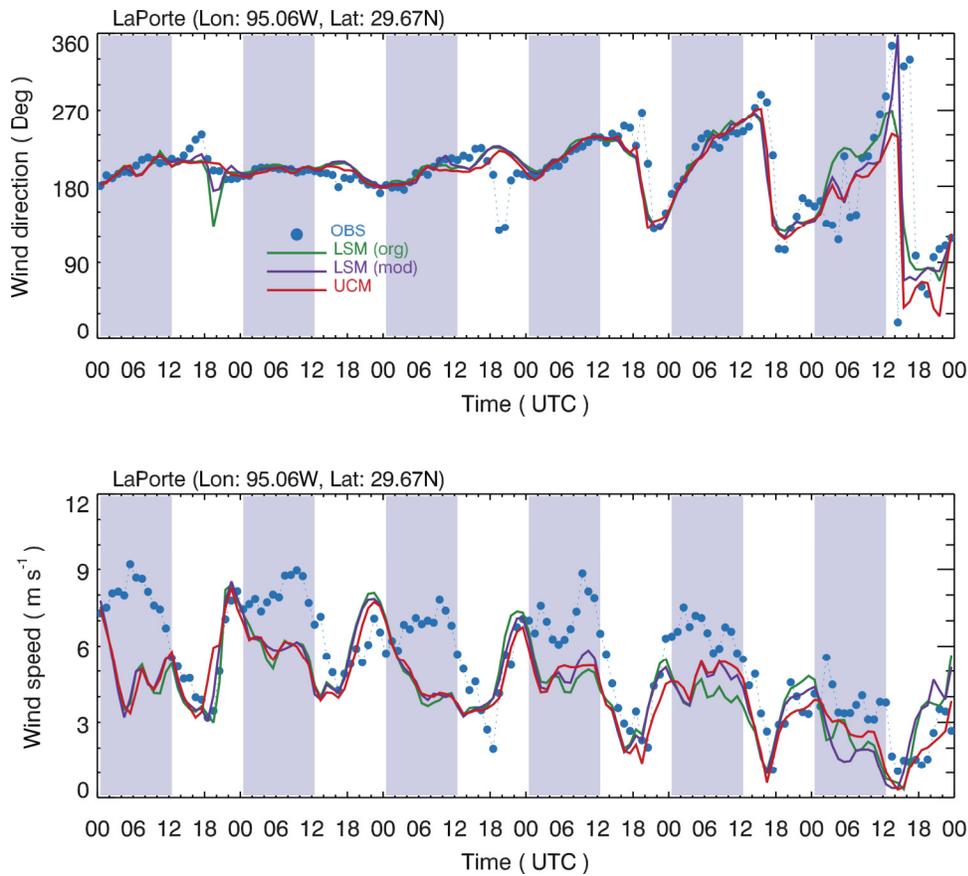


Fig. A4. The same as in Fig. 11 of the manuscript except for the comparison at 200 m AGL.