

In response to reviewer 1 about the LST bias in both URBClim and WRF as compared to MODIS, the explanation you provided to that comment also needs to be in the manuscript, rather than just the reply.

Reviewer 2 described your work as comparing an offline urban canopy model to a fully coupled meso-scale atmospheric model WRF. You disagree in this respect that UrbClim is not an offline model. However, I think you are mis-understanding the point made by the reviewer. In WRF, the inputs to the PBL scheme come from the LSM, SLC, as well as the dynamical core, which provides the prognostic variables. Yes, Urbclim parameterizes the PBL, but the inputs from ERA interim are fixed. This is not the case in WRF where the PBL scheme and the inputs to the PBL scheme have two-way interaction. In this sense, Urbclim is an offline model compared to WRF. You need to better explain the inner-workings of UrbClim and the key differences to WRF to make this clearer (not just refer to the paper which describes Urbclim).

Reviewer 2 asked for more physical explanation of the processes, which you have not carried out. However, since UrbClim also includes a land surface scheme, a comparison of energy balance from UC-ERA with WRF, e.g., sensible and latent heat, may help shed some light. I think you could try a bit harder in this respect. This is still plenty of room in this paper for a bit more analysis of the processes. I do not disagree that it is hard, but you should try nonetheless.

Reviewer 3 made a very good point about the use of statistical downscaling. Your reply needs to be incorporated within the manuscript, perhaps in the introduction.

Title: The main point here is that UrbClim is a “stand-alone” urban boundary layer model which is not coupled to a fully fledged atmospheric model, hence it is fast, which one would infer. I suggest to change the title to “Advantages of using a fast stand-alone urban boundary layer model as compared to a fully coupled mesoscale model to simulate .....”.

The abstract should generally be one paragraph.

Abstract, line 3, it is unclear what you mean by “these simulations”, as the previous lines do not refer to any simulations. I suggest “by high resolution (sub kilometer) fully coupled land-atmosphere simulations using urban canopy parameterizations” – This is more precise.

Abstract, line 4, “an urban” not “a urban”.

Why do you use upper case for “Urban Heat Island” throughout the abstract? Define the UHI acronym the first time, and use it in the rest of the abstract.

Abstract, line 10, as far as commonly used re-analysis products go (NNRP, FNL, ERA40, ERA-Interim, etc), 70 km resolution is pretty much as high as it gets! So your use of the term “relatively low resolution reanalysis (70 km)” is what the

rest of the climate community considers as high-resolution reanalysis. This needs to be changed.

Abstract, lines 11-12, simply change to “In addition, the effect of using driving data from a higher resolution forecast model (15 km) is explored in the case of UrbClim.”

Abstract, final paragraph. This is where you describe the main results and this lacks detail. I suggest to add 70 km in brackets after “reanalysis data” at line 14 at the end of the sentence. You need to state the actual “problem with the winds” this is too broad, and “day-to-day correlation” of what? I do not like the use of terms such as “the problem disappears” when referring to issues with models. Rather, “errors are substantially reduced from x units to y units” – is a lot more scientific and useful to a reader, rather than “the problem disappears”.

Additionally, provided it fits within the work limit for the abstract, a sentence or two on when it would be appropriate, and more importantly, when it would NOT be appropriate to use UrbClim would be good to mention. This is really important information I would expect in an abstract.

Page 2, line 31, you previously defined the UHI acronym on page 1, yet you use the full term. Please check this through the manuscript and be consistent.

Page 2, line 38, “Others have highlighted” rather than “Other authors”, put the reference at the end of the sentence.

You generally have too many one-sentence paragraphs, which makes the paper a bit “jumpy”. For example, page 3, lines 63, and 75, there are two paragraphs of one sentence each. You need to improve on the overall structure of your paper in terms of paragraphing.

Page 2, line 49, sentence starting with “Here we show”. This is a result and does not belong in an introduction.

Page 2, lines 54 to 56, where you describe the scope of UrbClim versus RCMs. Computational efficiency is only one aspect. You also need to mention that Urbclim, cannot be used to investigate the effect of UHI on atmospheric circulation, such as interactions with the sea breeze and convection/storm initiation, which has been shown by other studies. I think this is very important to make clear, rather than just the computational aspect.

Page 3 ,line 60, wrong cite command for Chen et al. (2011).

The description of the climatology of Barcelona etc, should be in the methods section, not the introduction.

Page 3 , second dot point, “of the UrbClim simulation” rather than “of this run”.

Page 5, line 98, what is a “well maintained” station is subjective, rather, is there any form of Quality control applied to this data before you used it? This may be more useful information to provide.

Page 6, line 131, “a minimum” rather than “the minimum”.

Page 6, line 134- “The boundary data needs to be read from a lower resolution model” – This is too broad and not detailed enough. State exactly what inputs UrbClim requires. Your use of the term “boundary data” is rather broad as well. WRF, as any RCM, uses input data at the lateral boundaries, and the influence of this data decreases within the sponge zone. I suspect UrbClim does not have an actual “boundary zone” as RCMs do?

Page 6, line 136, you state “Mesoscale models are tied to their driving models by the boundary conditions. Yet, they develop internal variability”. To me, this is stating the obvious, of course they will develop internal variability as the influence of the input boundary conditions decreases exponentially within the relaxation zone, and of course the driving data has a large influence.

Page 6, line 140, not just “wind” but as a consequence the advection of heat and moisture?

Page 6, line 141 to 143, I think the point you should make here, is that conventional RCMs have to use several nests to achieve high resolutions of 1 km or less, whereas this is not an issue with UrbClim.

Page 6, line 151, wrong cite command for Ridder and Schayes, (1997) and De Ridder et al. (2015). Please PROOF READ your manuscript.

Page 6, line 160. Again here, poor paragraph structure, you cannot start a paragraph with “In contrast”, it does not flow.

Page 7, line 178, it’s “Forecasting” and Not “Forecast”.

Page 8, line 196, “previous studies” rather than “previous works”.

Table 1 shows UC-ERA has consistently larger +ve biases and RMSE as compared to WRF across all stations. This is an important result I would expect to find in the abstract.

Page 9, line 205, “at some stations”, not “in some stations”.

Page 9, line 206, rather than “see below”, use “this is discussed later in the manuscript” or something along those lines.

Page 9, line 208, you state “Instead, UC-FC and WRF show similar, smaller scores, which indicate the good performance of these simulations”. You are stating the obvious here, of course smaller errors mean improved performance, one can assume the reader will make this connection. Rather, the point here is that UrbClim is very sensitive to the driving data, and you need driving data at 15 km

for the errors to be the same order of magnitude as WRF. This is significant as 15 km driving data is not routinely available over long time periods! This is far more important to discuss.

Page 10, line 214, wrong cite command again. Line 219, “at both stations” not “in both stations”.

Your supplementary figures have no captions, so I do not know what I have looking it.

Page 12, line 250, I don't really see the point of comparing MAE of 0.8 to 1.11, this is a difference of 0.3C, rather small. UC-FC is only marginally better than WRF.

Page 12, line 22, replace “more biased” with “has slightly larger biases” .

A number of studies have compared different PBL schemes in WRF against wind speed observations from Atmospheric soundings. You should actually reference these here and look at the RMSE, MAE and biases they report. I largely suspect that using different PBL and other schemes in WRF would result in changes in biases compared to observations which are larger than the differences you find between WRF and UC-FC, which would imply it would be entirely plausible that different configuration(s) of WRF could easily result in even larger or smaller errors than UC-FC. We cannot tell unless we do the simulations, but you should discuss this in more detail on page 12.

Page 13, below Fig 8 – again here, a one-sentence paragraph which does not really flow.

Page 13, lines 285-286 – You refer to biases in Figure 8, but figure 8 is not a bias plot? Not sure I follow here.

Page 15, the first dot point is not really a conclusion of your study as the aim was not to quantify the UHI of Barcelona. This should be removed as a conclusion.

Second dot point – What systematic biases? State them!

Third dot point – Of course WRF will provide less detailed spatial information, You ran it at a coarser resolution as compared to UC-ERA! This is to be expected and not a conclusion of the study. You fail to mention that UC-ERA had consistently larger biases than WRF, which is far more important, as well as the fact that UrbClim needs inputs at considerably higher resolution than routine available, i.e., 15 km, to really show a distinction from WRF. This is far more important.

In your conclusion, you state that this opens the door to running UrbClim with GCMs simulations of future climate. There are several problems here. Firstly, GCMs have much coarser resolution than 70 km, and you have clearly shown the resolution of input data has a large influence on UrbClim. Most GCMs have

greater than 150 km resolution. Secondly, you ran UrbClim with a re-analysis, which is completely different to GCM simulations of current and future climate. You are extrapolating too much here.

It would be much more useful to have a paragraph which objectively discusses, in which circumstances one should choose UrbClim over an RCM such as WRF, and what the user should be mindful of, rather than trying to extrapolate too much. Some of this needs to be reflected in the abstract.

Finally, in section 3.3, you provide a lot of detail about compilers etc. I do not think this adds much value at all to the paper. One expects that a stand-alone model such as UrbClim would require significantly less computational resources than an RCM such as WRF. You can simply state that given your WRF setup, UrbClim was about 130 times faster than WRF. All this detail on the OS, compiler versions, MPI stuff etc is not really adding much. This detail is only relevant when comparing slightly different versions of the same code, rather than two completely different codes. It would add more value to the paper, to spend more words on physical processes.