

## Review of Feng et al. : LA megacity : a high-resolution land-atmosphere modelling system for urban CO2 emissions

*Overview:* The manuscript presents simulated carbon dioxide fields for 2 months centered over Los Angeles. The work demonstrates and tests the ability of a high-resolution meso-scale model to reproduce observed meteorological and carbon dioxide dynamics, with a focus on urban areas, LA in particular. The paper presents a valuable modelling approach in order to understand the temporal and spatial variability of weather variables and CO2 mixing ratio in urban and background sites. This work is appropriately placed in ACP, and contributes to the burgeoning area of studying carbon emissions from urban areas. I have some general and specific concerns delineated below, that need to be addressed before its publication.

*General Comments:* Overall things look quite nice and interesting, but I have a couple of reservations that require more explanation and must be addressed. There needs to be better presentation of modelled vs observed fields in terms of table of scores and 1:1 plots. As currently presented, it is difficult to assess model performance. The second point is that discussions on the physical reasons why a parametrized scheme is better, or on the performance of the modelling, are missing. The last parts that study correlations of the simulated CO2 fields with GHG measurements is interesting, and well oriented to further inverse modelling studies. I do not have specific remarks on this part.

- 1) **CO2 initial and boundary condition.** This is only briefly touched upon in section 2.1, and it is unclear. From what I understand the model is initialized and coupled with CO2 concentrations coming from observations. The simulations run for 36h. Do you use the predicted CO2 field from the end of the previous day to start the following day ? Or do you only use CO2 observations at the beginning of each run ? In the 2nd case, what is the spin-up time ? Is there a significant horizontal and vertical variability in the CO2 observations ? What impact do varying boundary condition choices make on simulations? We know that in regional studies boundary conditions play a tremendously important role (Lauvaux et al. TELLUS 2012). The authors must better describe what they've done for boundary conditions, and make quantitative assessments of impacts of boundary condition choices on simulations.
- 2) As a large part of the simulated domains is on the sea, and as LA is largely influenced by maritime air masses, is it not a problem to ignore ocean fluxes ? Classically, **oceanic CO2 fluxes** are parameterised following Takahashi et al. (1997). A sensitivity test with ocean parametrized fluxes would be appreciated.
- 3) One objective of the paper is to assess the **PBL schemes**, but they are not physically described and the differences between the schemes are not presented . Therefore the conclusions are only limited to WRF technical configuration and physical aspects are not addressed. The 3 PBL schemes have to be described properly (closure, mixing lengths ...) to highlight the differences. Then strengths and weaknesses of each scheme need to be highlighted relating to their characteristics.
- 4) In the same way, 2 **urban surface schemes** are tested without having presented their physical differences. The scientific interest is therefore limited. We need to know the scientific reasons why UCM seems better.

- 5) *In the **comparison to aircraft PBL height**, the method to determine PBL height is based on the vertical virtual potential temperature gradient. Among the existing methods to determine this parameter (Ri number, parcel method ...), none is perfect. What is the impact of the choice of the method on the results ? For the 3 PBL schemes, biases on PBL heights are significant : errors of 160m in PBL height are not small by any measure. You can see for instance Riette and Lac (2016) for evaluation of PBL height over 1 year with an operation NWP model, with more satisfying values. Qualitative statements should be toned down. What is the error standard deviation ? Figure 3 is not appropriate as only biases are represented without standard deviation, and without length scale. How do you also explain that biases are smaller at 4km than at 1.3km, and that the results are different than the comparison to ceilometer ?*
- 6) ***Dynamics** : why do you use one-way nested domains and not 2-way ? Advection and temporal schemes should be specified in Table 1, with the time steps for the different resolutions. Page 7 line 16 : what is the height of the 1<sup>st</sup> level ?*
- 7) ***Comparison to radar wind profiler** : what is the period of evaluation ? Is it 2 months ? Tables of scores for wind speed and duration would be useful and easier to read than scores included in the text. Also, in Fig.5, if it is related to a 2 months period, it would be better to normalize the vertical coordinate by the PBL height.*
- 8) ***Comparison to NWS surface stations** : all the stations are not represented on Fig.S1 and the domain is not the same. As a complement to Fig.6, a table with scores for MYNN\_UCM is necessary, not only with biases but also with rmse. As a complement to Fig.6, it would be useful to provide two figures with the orography and the urban fraction for 1.3km resolution, and to discuss if the scores are related to orography, urban area... At 1.3km, what is the resolution of the orography database ?*
- 9) ***Comparison to in-situ CO2** : once again, a table of scores (bias and rmse) with the 4 simulations, as a complement to Fig.7, is missing.*
- 10) *This study focuses only on **two months** of modelling and observations (May-June 2010). Conclusions thus must be quite limited, as one cannot extrapolate to generalized model performance from such a limited duration comparison, which could be particularly favourable or unfavourable. The limited duration of model/observations must be presented, and its impact on conclusions should be discussed. One element of this is discussing time/computation to simulate one-month, and whether the current model construct could be expected to run for years to compare w/ the observational record being recorded in LA & USA.*

Specific comments :

P8 line 5 : It can be added that the coupling between mesoscale meteorological model and lagrangian particle model can be used in an operational framework to deal with accidental release (Lac et al., 2008).

Table 1 : There could be probably a mistake for shortwave radiation scheme : does RRTMG deal with SW radiation ?

*Abstract : The acronym FFCO2 is used before being presented.*

References :

*Riette and Lac : A New Framework to Compare Mass-Flux Schemes Within the AROME Numerical Weather Prediction Model, Boundary-Layer Meteorology, 2016, 1—29.*

*Lac, C., F. Bonnardot, O. Connan, C. Camail, D. Maro, D. Hebert, M. Rozet, and J. Pergaud, Evaluation of a mesoscale dispersion modelling tool during the CAPITOUL experiment, Meteorol. Atmos. Phys., 102, 263-287, 2008.*