

The paper attempts to evaluate the performance of TM5 to simulate boundary layer heights and surface radon concentrations. Some biases are found that the authors link to some weaknesses in TM5. Overall, the paper is fairly well written but it is obvious that many people were involved in the analysis of data and model output which makes the paper appear 'fragmented' and, at times, unstructured and disorganized. Provided below are major and minor comments which also include some suggestions to improve the paper.

Major comments

- 1) Is Geosc. Model Dev. An appropriate journal for this type of paper? This paper addresses the evaluation of a model, not the development. A journal such as Atmospheric Chemistry and Physics or Boundary Layer Meteorology seems more appropriate to me.
- 2) The title is too broad and should be made more focused on those aspects that are actually studied in the paper, i.e. daytime and nocturnal boundary depths and ^{222}Rn -concentration. Boundary layer dynamics include the study of thermodynamical and dynamical processes in the boundary layer including e.g. winds, stability, entrainment, etc. These processes are not studied in this paper and the title is therefore misleading. The title should reflect that the analysis is only made over Europe.
- 3) The difficulty of a coarse model to represent a coastal zone has not only to do with the coarseness of the model, but also the horizontal spatial variability. Also in high resolution models the largest spatial variability for fluxes can be found in these regions. For CO_2 this has been addressed by Pillai et al. (2010).
- 4) There are some problems in the structure of the paper, and titles of sections are sometimes inappropriate/misleading. Also the introduction of some figures in the text is sometimes a bit strange. For example. Figure 2 is introduced very early, but is only discussed very late (much later than the discussion of Figs. 4 and 5). This must be resolved by either putting the discussion in the section where it is introduced, or before the model output is compared to observations. As for an example eof a misleading section titles, consider e.g. Section 4 which is entitled 'simulation setup'. Section 4.1 only addresses extraction of model output and no aspects of the simulation setup. These misleading/inappropriate titles should be corrected. It would be good to give subsection with appropriate titles in the Result section.
- 5) The ceilometer/lidar related part does not really fit in this paper. There are many issues with the comparability between radiosonde/lidar derived PBL heights as discussed in many papers (and also obvious from Fig. 2) and you don't want to include these issues and uncertainties in this paper. In fact, including these data makes some conclusions in the paper rather weak. Figure 6 and 7 (and stars in Fig. 11) which include the ceilometer data do not add anything new and can easily be removed.
- 6) The authors mention coastal and non-coastal stations as well as mountainous stations (that they have removed from the evaluation). It would be nice to include the IGRA stations in table (not just Radon stations as is currently done) and indicate what stations are in coastal and mountainous regions. It also seems important that the authors explain how they define a coastal or mountainous station.

7) The reader is overwhelmed with data and figures (not to speak of the supplemental figures!). Reduce the number of figures and also the number of subfigures with certain figures. Some of this could be addressed by removing lidar/ceilometer related data as indicated in major comment 5. In Figs. 4 and 5, not all stations need to be shown. Just pick a few that clearly show some points you are making in the paper. It would also be nice to see in the figures which stations are in coastal/non-coastal terrain, as this seems important in the analysis (see previous comment on coastal and non-coastal stations).

Minor comments

1. P2, general: The abstract is very long (almost longer than the introduction) and reads like a summary.
2. P2, line 4: “dynamics” should be “height”
3. P3, line 15: define BLH properly, is it above the surface (depth) or above sea level (height).
4. P4, line 11: Section title could also be depth, depending on definition
5. P4, line 19: The equation of bulk Richardson number should be introduced here and not on page 7.
6. P4, lines 19-22: There should be some more explanation on choices made and how to use the bulk Richardson number. For example, how is θ_v calculated from IGRA-soundings? The neglect of u^* is hardly explained, but this is stressed in the Seidel 2012-paper, a citation here would help.
7. P5, line 13: The introduction of this figure is very strange, as it is not discussed here.
8. Figure 2: Including the ceilometer data is not recommended as mentioned in the major comments. We see here clearly one of the issues in that ceilometer is underestimating blh from IGRA. A complicated issue that is not suitable for the current paper.
9. P6, line 5: unclear: +/- 10 to +/- 15% ? or does +/- means approximately?
10. P6, line 9: 15m inlet should with a space. The paper has many of these types of typos. Please check.
11. P6, section 3.1: the addition of a figure where vertical resolution of TM5 model and radiosonde are compared would be helpful. This would also make clear at what exact depths the TM5 model gives output. Then, as an example one could examine a typical boundary layer depth in this figure. Keep in mind that many readers of Geos. Mod. Dev. are probably not familiar with a concept like boundary layer height. See also major comment on appropriateness of journal.
12. P6, line 30: there are 60 vertical levels below 0.1 hPa and 25 layers below 0.2 hPa. How dense is the layering between 0.1 and 0.2 hPa? Or is it ECMWF and TM5 layering?
13. P7, line 5 . The idea of an “updated slopes scheme (treatment?)” is very unclear and should be clarified.
14. P7, line 19: Delete “vertical”. “aerosol” should be plural.
15. P7, line 20: All the observational devices..... are based on the search...” Not an accurate statement. For example, sometimes strongest gradients occur right at the surface.
16. P7, line 21: “can be either” should be “can be based either on”.
17. P7, line 42: m/s is m s^{-1} .
18. P7, line 44: Unclear/ambiguous sentence.
19. P8, line 1: Why is a value of Ri_c of 0.3 used in TM5 and not the more common value of 0.25? Should be an easy fix for the model developers.
20. P8, line 8 and 14: What is the difference between ‘ ^{222}Rn flux map’ and the ‘InGOS ^{222}Rn flux map’ one? Be sure that the ‘abbreviations’ are used properly throughout the text.
21. P8, line 18: $\text{mBqm}^{-2}\text{s}^{-1}$. Some spaces are lacking in the unit.
22. P8, line 30-32: How can the extraction (or calculation) of variables (model boundary layer heights) be a simulation set-up. See also one of the major comments.

23. P8, line 42: What does 2D interpolation exactly mean? Various 2D approaches exist. Be specific and more accurate here.
24. P8, section 4.1: Is it really valuable to have so many different definitions? Besides, in this section, I would expect some discussion about the representation of the grid points chosen with respect to reality of the stations as this seems important for your discussion later on (coastal and non-coastal).
25. P9, line 7: ECMWF can be added as a bullet point.
26. P9, section 4.2: it is very unclear what type of simulations have been done. Consider a table.
27. P9, line 31: for clarity, at least one bl-profile with the different calculations of bl-height could be shown. Here, also vertical resolution of both IGRA and models can be shown. Besides, you can point out the differences generally found for a nocturnal and daytime (a 00 and 12 UTC) bl- figure, for example.
28. P9, line 34: Which mountain stations, and how did you define a mountain station? You could add labels in Table 1. The same holds for coastal and non-coastal stations, it is not defined what they are, this could be labeled in Table 1 as well.
29. P10, line 4: "coastal sites". Why don't you show a map with the representation of these two stations in the several data points extraction?
30. P10, line 11: How are non-coastal sites defined?
31. P10, line 15: "probably". What makes you think probably and not certainly?
32. P10, line 25: "relatively" compared to what? And are you surprised by these results? It is well known that Sbls are very shallow, and often these are missed by the model anyway.
33. P10, line 27: costal should be coastal.
34. P10, line 31: As mentioned in previous comments, figure 2, and, in general, ceilometer related data, should be removed in this paper (the correlation is poor and subject to many discussions that are not appropriate to discuss in this type of paper). Furthermore, this figure that shows observations vs. observations does not fit in a section which is called simulations vs. observations?
35. P10, line 37: DeBilt should be De Bilt.
36. P10, Figures 6 and 7 are redundant, see one of the major comments.
37. p10, lines 29 to 45 should be removed. See previous comments on ceilometer data.
38. P11, line 4-5: About the timing of Rn-concentrations (05 and 14 UTC). Does this hold for both summer and winter?
39. P11, line 14: The list of coastal stations gets longer throughout the paper. Add it as labels in the paper. Is Cabauw really coastal station? I don't think locals would agree.
40. P11, line 24: This could be the start of an additional section.
41. P11, line 30: What do you mean by "Apparently"?
42. P11, line 31: What's a "model world"?
43. P11, line 44: "This finding..." What finding exactly?
44. P9-13, Section 5. what about a station selection for the figures? You seem to overwhelm the reader with graphs and bars, whereas only few things are to be highlighted. For example, the coastal and non-coastal zones are interesting, some are necessary to show due to later analysis with the Rn- and BLH combination. But certainly not all the stations are necessary. Then space would be saved, figures could be enlarged, these would be better readable and the article in general would be better appreciated. All the other redundant stations can then be stored in the supplemental material (which is very large as well).
45. P10-11, Section 5: Can be better divided in more sections.
46. P12, line 11: Add section 5.4 for this paragraph.
47. P12, line 38: "weather conditions", I suppose you mean "stability regimes"?

48. P12, line 42-43: About CB1 and CB4, see remarks about table.
49. P13, line 14 (section 6): this section feels more like a summary than a conclusion.
50. P13, line 16: "dynamics" should rather be "height". The dynamics are not evaluated.
51. P13, line 19: 10-20%, this is the first time I see a statistical value between TM5 and observations. That's very late.
52. P13, line 23: IGRA observations (not "data").
53. P13, line 26: moderate correlation or reasonable? In any case, it is not good and opens up the floor for many discussions that are not relevant to the topic of the paper. See comments before about removing ceilometer related analysis.
54. P13, lines 26-35: remove ceilometer related analysis/results.
55. P14, line 23-24: It is indeed difficult to draw conclusions. Try to be more quantitative, perhaps add more statistics, and this would make it easier to draw conclusions.
56. Table 1. Extend this table with labels as coastal and non-coastal stations. what is ANSTO? What is CB1? CB4? Probably different levels at the tower? The average readers do not have previous knowledge of the dataset and an attempt needs to be made to make it more readable for them. Maybe you want to change CB1 and CB4 to level 1 and level 2. Are the 'o' for latitude and longitude degree (°) signs? What does Altitude/Height exactly mean? Do you mean Terrain elevation (Mean Sea Level) and Height (Above Ground Level), respectively?
57. Figure 1. What are the abbreviations? The authors should refer to Table 1. Some letters are very hard to read, consider another color for either the names, or for the black dots. The black triangles and orange circles are barely visible. Where are the coastal stations exactly? What are the vertical and horizontal lines? Longitude and latitude? It should be indicated on the axis.
58. Figure 2. remove. See major/minor comments
59. Figure 3. Vertical and horizontal axis in the upper diagram?
60. Figure 4. These figures are very small. Can the whisker plots be centralized around the months to which they are concerning to? And maybe the spacing then between the whisker plots could be enlarged. The scaling on y-axis is not the same. It doesn't have to, but it should be stated in the caption. Although the scales are not very far apart, so probably same axis length would work. This is true for almost all figures.
61. In the text it is referred to coastal and non-coastal stations. It would be helpful to highlight that in the figure. Is it really necessary to show all stations? Maybe you could make coastal stations fig4a and continental stations in fig. 4b?
62. Figure 6/7. Redundant/remove
63. Figure 11: remove CEIL/LIDAR data from figure.
64. Figure 12. What is ratio? TM5 divided by IGRA?
65. Figure 13. Abbreviations are not explained well in caption (e.g what is CB1 and CB4?). "Mean diurnal variations..." should probably be "Monthly mean diurnal variations..."
66. Figure 13: Data points are outside the y-axis range. This should be corrected.
67. Figure 14. How can the R of the lower figure be almost the same as the R for the upper figure?

References:

Pillai, D., Gerbig, C., Marshall, J., Ahmadov, R., Kretschmer, R., Koch, T., & Karstens, U. (2010). High resolution modeling of CO₂ over Europe: implications for representation errors of satellite retrievals. *Atmospheric Chemistry and Physics*, 10(1), 83-94.