

**Interactive comment on “Evaluation of the boundary layer dynamics of the TM5 model”  
by E. N. Koffi et al.**

**Anonymous Referee #1**

Received and published: 29 April 2016

We thank the reviewer for his/her constructive review. In what follows, the comments of the reviewer are in italic and our reply in normal face.

**General comments**

*I think this paper could be significantly enhanced by including some further discussion or even recommendations on estimating model transport errors based on the model-observation comparisons of  $^{222}\text{Rn}$  and BLHs. As the authors already point out, transport errors are a substantial source of uncertainty in the fluxes estimated in atmospheric inversions. There are already a number of groups using TM5 in atmospheric inversions, but such recommendations need not be limited only to TM5 but in general the use of the new  $^{222}\text{Rn}$  emission map and the IGRA BLH dataset for assessing model transport errors.*

We have added several recommendations at the end of the conclusions.

*The paper includes many detailed figures of the comparison of BLHs and  $^{222}\text{Rn}$  but I think a couple of figures that summarize (i.e. give a more immediate indication of) the comparison between the model and observations and of the seasonal and diurnal cycles could be very helpful. Then some of the detailed figures could be moved into the supplement.*

We have reduced the number of figures, and reduced the number of scenarios shown both for the comparison of the BLH and  $^{222}\text{Rn}$  activity concentrations (as described in more detail in reply to reviewers #2 and #3).

***Specific comments***

*P3, L21: Here the authors mention only surface monitoring stations in regional inversions but not aircraft data, which are often used (e.g. the Kort et al. 2008 study cited here). Model representation of aircraft observations will be also affected by errors in BLH and simulations of boundary layer dynamics. Perhaps this should be mentioned.*

We have modified the text, mentioning explicitly the use of aircraft data in the study of Kort et al. (2008). Furthermore, we added a reference (Miller et al., 2013), which also use aircraft data for their flux inversion.

*P10, L26-27: It is interesting that the modelled nocturnal BLHs tend to be higher than observed in summer but that this is not the case in winter? Can the authors comment on this?*

Obviously, the model has in particular difficulties to simulate the very shallow nocturnal BLH, which is often observed at continental stations in summer. This is partly due to the fixed lower limit of 100m for the BLH in the model (see Figures S12 and S13 in the revised version of the Supplement).

*P10, L41: The authors do not discuss comparison of the modelled and observed (at IGRA sites) nighttime BLHs for Cabauw or Trainou.*

We had not discussed in detail the comparison of modelled and observed nighttime BLHs at Cabauw and Trainou due to the limitations of the ceilometer / LIDAR measurements during night (see section 2.1.2).

*P11, L7: Please give a quantitative estimate of “better agreement” either stating the improvement in the RMSE or correlation.*

The statement refers primarily to Figures 8 and 9 (Figures 6 and 7 the revised version), which shows the seasonal variation of observed and simulated <sup>222</sup>Rn activity concentration. In addition, the improvement is also clearly visible in the overall statistics, shown in Figure 11 (Figure 8 in the revised version), which shows the improvement both in the RMSE values and correlation coefficients. This is briefly discussed later (in Section 4.2 of the revised manuscript).

*P11, L15: Please delete “apparently” – either the InGOS 222Rn flux maps give better agreement or they don’t, so “apparently” is not appropriate here.*

Deleted as suggested

*P11, L38-39: The authors state that the mismatch between the observed and modelled 222Rn activity concentrations cannot be due to the modelled BLH because this matches the observed BLH well. However, I understand that the modelled BLH is determined by vertical interpolation, therefore, I wonder if the vertical resolution in TM5 may be a possible reason for the mismatch?*

The dependence of the TM5 BLH on the vertical resolution has not been investigated. However, we note that the TM5 BLH (evaluated in the model version with 25 vertical layers) is in general very close to the ECMWF ERA-Interim BLH (60 vertical layers).

*P13, L1-11: I think this section should be expanded to discuss the influence of compensating errors in the 222Rn fluxes (in the constant versus InGOS flux maps) and in the BLHs and how this might explain the fact that the simulations with the constant fluxes lead to a better comparison with the observations.*

This is a good point. We have added a short statement that this could point to partially compensating systematic errors (See Section 4.2 in the revised version).

*Technical comments*

*P4, L46: “as” should be replaced by “compared to”*

We have slightly modified the sentence to: "attribute the height of the residual layer of aerosol ... as height of the real mixed layer". The suggested "compared to" would change the content of the sentence

*P10, L36: delete “also” after “In addition”.*

Deleted as suggested

**Interactive comment on “Evaluation of the boundary layer dynamics of the TM5 model”  
by E. N. Koffi et al.**

**Anonymous Referee #2**

**Received and published: 2 May 2016**

We thank the reviewer for his/her constructive review. In what follows, the comments of the reviewer are in italic and our reply in normal face.

**General comments**

*This study reports on a thorough evaluation of TM5 to describe the boundary layer dynamics, comparing various parameterization settings of the BL and extraction methods height to radiosonde, lidar and ceilometer observations. Furthermore simulations of  $^{222}\text{Rn}$  using two different emissions and various settings for advection and convection in TM5 are compared. The study draws potentially important conclusions regarding uncertainties due to convection parameterization in TM5, relevant for GHG emission studies, and is therefore well suited for publication in GMD.*

*While this study is certainly thorough, in its current shape the manuscript is merely a report on the numerous sensitivity runs that have been executed. A more rigorous selection of sensitivity experiments to be presented, along with a more selective presentation of observational data could largely improve the readability of the manuscript. Also the abstract is currently too elongated.*

*For instance, both presenting ‘TM5’ and ‘TM5-IGRA’ and likewise ‘TM5-INGOS’ and ‘TM5-INGOS-INGRA’ in figures 4-7 seems not necessary, as ‘differences are usually very small’ (p.10, l.13)’, and furthermore such differences cannot be explained in terms of sensitivity of the parameterization, but rather reflect a representation error. Therefore I believe these simulation results are even a bit confusing and should be removed from the figures.*

We have significantly reduced the number of sensitivity experiments shown in the main paper: For the TM5 boundary layer heights we show now in the revised version only the boundary layers heights evaluated with the InGOS definition (consistent with the definition used for the IGRA radiosondes), evaluated both at the InGOS stations and the adjacent IGRA stations (see Section 3.2 in the revised version). The additional evaluations of the BLH are now shown only in the supplementary material. For  $^{222}\text{Rn}$  activity concentrations, we show now only 3 cases (FC\_CT, FI\_CT, FI\_CU; see Section 3.4 in the revised version) in the main Figures. Also the abstract has been significantly shortened.

*In Figures 8 and 9 a clear improvement with the revised  $^{222}\text{Rn}$  emission map is visible, but differences between various convection/advection parameterizations is less obvious, which*

*makes me wonder if presentation of all these results could not be more condensed, or moved to the supplementary material.*

We have condensed the presentation of the various convection/advection parameterizations, and show now in the revised version only the simulations with the combined 'revised slopes scheme' and ECMWF ERA-Interim convection. Furthermore, we removed the paragraph on this issue from the abstract

*I believe the figures 4-9 benefit from presenting only seasonal mean statistics, rather than monthly means: The same messages can be conveyed with much condensed use of figures.*

We had deliberately chosen to show the monthly means and would like to keep this presentation, since it gives more detailed information (more precise representation of the seasonal evolution) than the seasonal means. As already explained above, to render more readable the different graphs, we show now in the revised version only the more relevant model experimental settings in these Figures.

*Also the authors put large emphasis on the improvement in the comparison to  $^{222}\text{Rn}$  observations when using the new flux map. However, the purpose of this paper is rather the evaluation of the boundary layer dynamics in TM5, by performing sensitivity runs. While many figures are presented, in the end it remains unclear to me how the parameterizations quantitatively compare, presented preferably in a Table. E.g. the statistics of the analyses given in Figures 11 and 12 could be averaged over the different stations, while excluding coastal stations hampered by representation errors and excluding the results obtained with the simplified flux map.*

Although the evaluation of the new  $^{222}\text{Rn}$  flux map is not the primary objective of this paper, realistic  $^{222}\text{Rn}$  emissions are an essential prerequisite for the model validation.

We prefer to keep the presentation of the statistics per station (Figure 11, now Figure 8 in the revised version), because of (1) considerable differences also among the non-coastal sites, and (2) the limited number of stations.

*On the abstract, I believe the authors should condense this strongly, by reporting only the key findings of this study, which I believe are the performance of TM5 to represent BLH (l.14-l.17), and the achievements and limitations of the comparison against the new  $^{222}\text{Rn}$  flux map.*

We have condensed the abstract significantly and deleted the paragraph on the different convection/advection parameterizations.

### ***Detailed comments***

*Abstract*

See our reply above

*Please consider to condense especially lines 3-12.*

We shortened this part of the abstract.

*Also I suggest to remove the conclusion regarding the improvement with the new Karstens et al. emissions from the abstract because, even though interesting, it is not essential to the subject of this manuscript.*

We think that this conclusion is important because realistic  $^{222}\text{Rn}$  emissions are an essential prerequisite for the model validation. Therefore, we would like to keep this conclusion in the abstract.

*Consider re-formulation of sentence on l. 21-24, which is difficult to grasp.*

The sentence has been slightly rephrased

*Also lines 37-42 read a bit confusing: while ECMWF convection results in much lower  $^{222}\text{Rn}$  activity than TM5 the authors cannot conclude if this is an improvement or not, which in its current formulation, does not appear a useful finding.*

We have deleted this paragraph from the abstract.

### *Introduction*

*I expect a few more references to studies to previous work that have considered the relevance of boundary layer dynamics for trace gas distributions, (and inversions), e.g. Locatelli et al., GMD 2015. How does this new work relate to that study?*

We have included the suggested reference Locatelli et al. (2015).

### *Section 2*

*Page 5, l. 14: You introduce a figure where you compare the Cabauw ceilometer BL with IGRA data. I expect some discussion and interpretation of this result at this point.*

We have moved Figure 2 (submitted version) to the Supplement. The scatter plots of Cabauw ceilometer BLHs at 00 and 12 UTC are now shown in a single Figure S1 in the Supplement (and the figure caption updated accordingly).

### *Section 3*

*Here, and at several points throughout the manuscript, you mention the issues associated to the resolution of TM5 (1x1 horizontally over Europe, 25 vertical layers, 3-hourly surface meteorological data, 6-hourly 3D fields). Considering its apparent relevance it would have been interesting to see a sensitivity study at different model resolutions. Could you specify the temporal resolution of the ECMWF convection fields in your sensitivity study? Is this 6 hour?*

The temporal resolution of the ECMWF ERA-Interim convective fields is 3 hours. However, in the TM5 version used in this study, 6 hourly 3D meteo fields were applied (See Section 3.1).

*P 8, l 14: 'Noah soil moisture data' : Do you have a reference here?*

The following reference of Rodell et al. (2004) has been added:

Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll, 2004. The Global Land Data Assimilation System, Bulletin of the American Meteorological Society, 85(3): 381-394

*Considering it's apparent sensitivity to soil moisture, why didn't you consider use of the ERA-Interim reanalysis? This would be more consistent with the 222Rn atmospheric model simulations, or?*

Karstens et al. (2015) recommended the use of the new emission maps derived from the Noah reanalysis. The authors found that “comparison with observations suggests that the flux estimates based on the GLDAS Noah soil moisture model on average better represent observed fluxes”. We included the conclusion from Karstens et al. (2015) in the text to explain our choice of the Noah data based <sup>222</sup>Rn flux map.

Furthermore, we simulated the <sup>222</sup>Rn activity concentrations also using the ERA-Interim based <sup>222</sup>Rn flux map (not shown). These additional sensitivity runs showed overall poorer agreement with 222Rn observations than the Noah data based <sup>222</sup>Rn simulations, confirming the conclusion of Karstens et al. (2015).

#### *Section 4*

*“We extract. . .”: Are TM5 simulations of 222Rn collocated in time and space with respect to the observations? Please be more specific here on horizontal, vertical and time interpolation.*

We apply 3 dimensional interpolation (i.e., horizontal and vertical interpolation) using the <sup>222</sup>Rn activity concentrations of the neighboring grid cells. The model output are hourly averaged concentrations (which are directly compared to the hourly averaged observations)

#### *Section 5.*

*Pp 10, l 14 – 119: So do you have any indication that ECMWF treatment of the BLH is better than the one currently used in TM5, based on this? Please provide more quantitative conclusions.*

No. We only stated that the differences we sometimes observe between TM5 BLHs and ECMWF BLHs at some coastal sites may be attributed to i) the relatively finer spatial resolution of ECMWF (~80 km in horizontal on 60 levels) and ii) to the different treatment of BLH in the

two models. The ECMWF BLHs are not discussed anymore in the text, hence this sentence has been deleted

*Pp 11, l 38: “The mismatch (. . .) cannot be explained by the modelled BLH”. This statement seems inconsistent with Figure 12, where ‘potential shortcomings of TM5 to correctly simulate the vertical  $^{222}\text{Rn}$  activity concentration gradients’ are illustrated. Please explain this apparent inconsistency.*

This is maybe confusing. No, Figure 12 (submitted version; Figure 10 in the revised version) shows the ratios of boundary layer heights (modelled BLH versus observed BLH) at noon along with the ratios of  $^{222}\text{Rn}$  activity concentrations (observed versus simulated) at 12, 13, 14, 15 LT for different seasons. We found that at most of the studied stations, the modelled BLHs compare well with observed BLHs, while the differences between the simulated  $^{222}\text{Rn}$  activity concentrations and observed ones can be larger. This result points to potential shortcomings of TM5 to correctly simulate the vertical mixing of  $^{222}\text{Rn}$  activity concentrations within the boundary layer. The text has been updated.

*P12, l3-17: Karstens et al. pointed out that the uncertainty averaged over the footprint might be smaller than 50*

The uncertainty averaged over the footprint could be smaller. However, as discussed in the paper, the uncertainties of neighboring pixels in the  $^{222}\text{Rn}$  flux map are likely strongly correlated, and therefore the reduction of the relative uncertainty (integrated over a typical footprint on the order of 50-200km) is probably relatively small.

*P12, l22: the authors suggest that the GHG-emissions derived in inverse modeling change by the same order of magnitude as  $^{222}\text{Rn}$ , i.e. 10-30*

Yes, this is correct. As mentioned in the paper, this has also been confirmed by first GHG inversions with the new ECMWF based convection (not shown).

*P12, l35-39: Please provide a short interpretation of this sensitivity analysis.*

We analyzed the ratios of both boundary layer heights and  $^{222}\text{Rn}$  activity concentrations as shown in Figure 12 (submitted version; Figure 10 in the revised version) for the 3 main stability regimes (stable, neutral, unstable or fair). We used the modelled Richardson number obtained at the first level of the model to discriminate between the 3 stability regimes. Results for the three stability regimes are similar and similar to those obtained when considering sample covering all the stability regimes shown in Figure 10 (revised version). A limitation of this exercise was that for both stable and neutral stability regimes, we had at most stations, only few cases by seasons. The text has been revised

*P12, 143: “tower height of 20m is within the first model layer 200m is within layer 3.”: Considering it’s sensitivity, how did you treat the model sampling? Did you apply vertical interpolation? Do you expect any sensitivity to vertical model resolution?*

As mentioned above, <sup>222</sup>Rn activity concentrations are 3-D interpolated, i.e. including vertical interpolation. Yes, we expect some dependence on the vertical resolution of the model. However, this has not yet been analyzed in detail. The 3-D interpolation is now stated (see Section 3.2)

*P13, 111: In this section, and in Figure 13, I miss results from the FI-CE run using the ECMWF meteo. Or are differences marginal? Please comment.*

We now present the simulations of <sup>222</sup>Rn activity concentrations by using convection scheme based on ECMWF reanalysis (FI-CE) combined with the “revised slopes scheme (FI\_CS). The differences between FI-CT and FI-CS are marginal (Figures S14-S24 in the revised version of the Supplement), hence the differences between FI\_CT and FI\_CU are dominated by FI-CE

*Figure 11, right panels and Figure 12: Which parameterizations are used for the computation of the TM5 boundary layer height? Standard TM5 or ECMWF convection? Or is the difference in BLH for the two parameterizations marginal?*

In Figure 11 (submitted version), the TM5 default boundary layer was shown in the submitted version. We now show the boundary layer height extract at the closest IGRA station associated to the InGOS measurement sites (acronym TM5\_INGOS\_IGRA; Figure 8 in the revised version)

### *Conclusions*

*P 14, 117 “The updated slopes treatment”: This is jargon. Please reformulate to something more generic, e.g. “the revised advection parameterization”.*

We use now the term 'revised slopes scheme' throughout the paper.

*Could you indicate the importance of this study for GHG inversions based on TM5? Is this study a ground for replacing the convection treatment in TM5, or is it merely useful in providing a constraint on the uncertainty estimate of the GHG emission inversions?*

Since we did not find a significant difference / improvement of the <sup>222</sup>Rn simulations with the new ECMWF convection, this study does not provide enough evidence, which would justify the replacement of the convection scheme. Further studies are currently performed within the TM5 modelling community (including the use of further tracers), however at this stage no clear conclusion can be drawn.

**Interactive comment on “Evaluation of the boundary layer dynamics of the TM5 model”  
by E. N. Koffi et al.**

**Anonymous Referee #3**

We thank the reviewer for his/her constructive review. In what follows, the comments of the reviewer are in italic and our reply in normal face

**Received and published: 3 June 2016**

*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.*

Yes, we believe that GMD is appropriate:

(1) GMD lists under 'Aims and scope' explicitly 'full evaluations of previously published models', see:

[http://www.geoscientific-model-development.net/about/aims\\_and\\_scope.html](http://www.geoscientific-model-development.net/about/aims_and_scope.html)

(2) Moreover, the referee #2 explicitly stated that the paper is "well suited for publication in GMD".

*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  $^{222}\text{Rn}$ -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.*

The paper investigates the boundary layer heights as well as the processes in the boundary layer (including dynamic and thermodynamic, etc...). When simulating the  $^{222}\text{Rn}$  activity concentration in TM5, all the thermodynamic and dynamic processes in the model are relevant. Thus, the evaluation of the model simulations of  $^{222}\text{Rn}$  activity concentrations implicitly includes the evaluation of the whole boundary layer dynamics.

We agree with the second statement regarding the focus of the paper on Europe and have updated the title accordingly: “Evaluation of the boundary layer dynamics of the TM5 model 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<sub>2</sub> this has been addressed by Pillai et al.(2010).*

We agree that also spatial variability of <sup>222</sup>Rn fluxes close to the coast may also play a role for the simulation of <sup>222</sup>Rn activity concentrations at stations close to coast. For the <sup>222</sup>Rn fluxes, the largest effect should be related to the variability / gradient of the water table close to the coast. In principle, this should be covered by the <sup>222</sup>Rn flux map (within the horizontal resolution of the input data sets)

*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.*

We have revised the structure of the paper. Figure 2 has been moved to the Supplement. Several titles have been updated and subtitles have been added.

*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 heighths 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.*

Despite the discussed limitations (especially for the ceilometer at Cabauw during night), we consider the ceilometer/lidar data useful as they provide information about the dynamic evolution, which is not well resolved by the IGRA data with only 2 measurements per day.

*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.*

We indicate now the chosen IGRA stations (which are closest to the InGOS stations) in the updated Figure 1.

We consider stations as 'coastal' (in a strict sense) if they are located at the coast (as e.g, Mace Head). However, when comparing with model simulations, also the horizontal model resolution has to be taken into account. Thus, model representation errors (related to the land / sea gradients) arise, if the model grid cell, in which the station is located, covers also a significant sea fraction. For stations, for which this is relevant (but which are not coastal stations in a strict sense) we choose now the term 'close to the coast'

*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).*

We have reduced the number of figures both in the main paper and in the Supplement. Furthermore, we have reduced significantly the number of scenarios. However, we would like to show the full set of InGOS stations, since this paper also aims to support the further analysis of the GHG flux inversions (CH<sub>4</sub>, N<sub>2</sub>O) performed within the InGOS project (manuscripts in preparation). We consider the link between these studies very important, since potential systematic errors in the simulation of the BLH dynamics (discussed in the present paper) could directly translate into systematic errors in the derived fluxes.

### ***Minor comments***

*1. P2, general: The abstract is very long (almost longer than the introduction) and reads like a summary.*

The abstract has been revised (and significantly shortened)

*2. P2, line 4: “dynamics” should be “height”*

We would prefer to keep the term “dynamics” (see also our reply to reviewer's major comment (2))

*3. P3, line 15: define BLH properly, is it above the surface (depth) or above sea level (height).*

The BLH is defined with reference to surface elevation, and not to sea level (Seidel et al., 2012). This has now been added in section 2.1.

4. P4, line 11: Section title could also be depth, depending on definition

We prefer to keep the title "Boundary layer height"

5. P4, line 19: The equation of bulk Richardson number should be introduced here and not on page 7.

Yes. The paragraph has been moved to Section 2.1 and updated

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  $\theta_y$  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.

This part of the text has been revised and the paper of Seidel et al. (2012) quoted again there

7. P5, line 13: The introduction of this figure is very strange, as it is not discussed here.

This Figure has been moved to the Supplement as Figure S1

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.

See our reply to the reviewer's major comments (5)

9. P6, line 5: unclear: +/- 10 to +/- 15% ? or does +/- means approximately?

We have corrected this to '10-15%'

10. P6, line 9: 15m inlet should with a space. The paper has many of these types of typos. Please check.

The GMD convention seems to be not to use a space before 'm' (meter)

[http://www.geoscientific-model-development.net/for\\_authors/manuscript\\_preparation.html](http://www.geoscientific-model-development.net/for_authors/manuscript_preparation.html)

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.

We agree that such a figure would be useful. However, it would increase the number of figures that the reviewer asked to decrease. We refer the reader to the paper of Seidel et al. (2012), where the method is nicely illustrated in their Figure 1.

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?*

The 25 vertical layers of the TM5 model version used in this study are defined as a subset of the 60 vertical layers of the ECMWF ERA-Interim reanalysis. The text has been updated.

13. P7, line 5 . *The idea of an “updated slopes scheme (treatment?)” is very unclear and should be clarified.*

We have updated the short description of the “revised slopes scheme” (and use this term now throughout the text). For further details the reader is referred to van der Veen (2013).

14. P7, line 19: *Delete “vertical”. “aerosol” should be plural.*

Has been corrected as suggested (moved to section 2.1)

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.*

This should exclude indeed the gradients right at the surface.

16. P7, line 21: *“can be either” should be “can be based either on”.*

Corrected as suggested

17. P7, line 42: *m/s is m s<sup>-1</sup>.*

Corrected as suggested

18. P7, line 44: *Unclear/ambiguous sentence.*

First, we computed the Richardson number  $R_{ib}$  at each of the model levels by using the equation (1). To determine the boundary layer height, the vertical profile of  $R_{ib}$  is interpolated linearly between consecutive levels. The BLH is defined as the height, where  $R_{ib}$  reaches the critical value  $R_{ic}$ . The text has been updated.

19. P8, line 1: *Why is a value of  $R_{ic}$  of 0.3 used in TM5 and not the more common value of 0.25? Should be an easy fix for the model developers.*

$R_{ic}$  of 0.3 is the default value for the BLH determination in TM5, but there is no publication about this specific aspect. However, as discussed in Seidel et al. (2012), the choice of  $R_{ic}$  close to 0.25 does not introduce large uncertainty. Moreover, the differences between BLHs determined by using  $R_{ic}$  of 0.25 and  $R_{ic}$  of 0.3 are very small (see e.g., Figures S2-S11 in the Supplement; acronyms TM5 and TM5\_InGOS).

20. P8, line 8 and 14: *What is the difference between ‘222Rn flux map’ and the ‘InGOS 222Rn flux map’ one? Be sure that the ‘abbreviations’ are used properly throughout the text.*

It is the same flux map. This has been clarified throughout the text by using “InGOS 222Rn flux map”

21. P8, line 18: *mBqm-2s-1. Some spaces are lacking in the unit.*

In accordance with our response above, we keep it as it

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.*

The extraction of the BLH according to the INGOS definition required some specific modification of the TM5 source code.

The text has been updated and “Simulation set up” has been deleted

23. P8, line 42: *What does 2D interpolation exactly mean? Various 2D approaches exist. Be specific and more accurate here.*

It is linearly interpolated between the grid cell and its closest neighbor (both along longitude and latitude). The text has been updated

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).*

As already mentioned in our reply to the reviewer's major comments (item 7), we now consider only two definitions: “TM5\_INGOS” and TM5\_INGOS\_IGRA” that use the same expression of Bulk Richardson number, as performed for the IGRA data. TM5\_INGOS stands for the boundary layer heights (BLH) extracted at InGOS stations, while TM5\_INGOS\_IGRA is the BLH of the closest IGRA stations. The other model experimental settings are now defined in the Supplement and the relevant results are also shown in the Supplement

25. P9, line 7: *ECMWF can be added as a bullet point.*

This has been deleted in the main paper and put in the Supplement.

26. P9, section 4.2: *it is very unclear what type of simulations have been done. Consider a table.*

As already described above, now we use only two definitions of boundary layer height in the paper. The figures have been revised accordingly and are more readable

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.*

As mentioned above, we refer the reader here to Seidel et al. (2012) (and their Figure 1, which clearly illustrates the method). Finally, we do not think that illustrating the vertical bl-profiles for both nocturnal and daytime will help in our discussions on IGRA BLHs and modelled BLHs

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.*

For mountain stations, we excluded InGOS measurement sites such as e.g., Jungfrauoch and Schauinsland in the analysis. About the coastal sites, see our reply to the reviewer's major comments (item 6)

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?*

We now show the locations of the IGRA stations associated to InGOS in Figure 1.

Regarding “coastal sites” see our reply to the reviewer's major comments (item 6)

30. P10, line 11: *How are non-coastal sites defined?*

We have not defined objectively “coastal or non-coastal stations” as explained above or in our responses to the reviewer's major comments in item 6

31. P10, line 15: *“probably”. What makes you think probably and not certainly?*

This sentence has been deleted when revised the text

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.*

“Relatively” here was about a comparison between nighttime and daytime BLHs. “Relatively” has been deleted

33. P10, line 27: *costal should be coastal.*

Corrected

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?*

As discussed in our reply to the reviewer's major comments (see item 5), we prefer to keep the analysis based on the ceilometer/lidar data, but moving the Figure 2 (submitted version of the paper) in the Supplement (Figure S1 in the revised version). The text has been revised accordingly

*35. P10, line 37: DeBilt should be De Bilt.*

Corrected

*36. P10, Figures 6 and 7 are redundant, see one of the major comments.*

We prefer to keep these two Figures, but they are now in a single Figure (Figure 5 in the revised version). For more detail, see our reply to the reviewer's major comments (item 5)

*37. p10, lines 29 to 45 should be removed. See previous comments on ceilometer data.*

In accordance with our reply to the reviewer's major comments (see item 5) about the use of ceilometer/lidar in this paper, we have kept the content of this part of the text

*38. P11, line 4-5: About the timing of Rn-concentrations (05 and 14 UTC). Does this hold for both summer and winter?*

The chosen reference times should be reasonable for both summer and winter seasons

*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.*

See our reply to the reviewer's major comment (6)

*40. P11, line 24: This could be the start of an additional section.*

Yes. We add a new sub-section: "Relationship between 222Rn activity concentrations and BLHs"

*41. P11, line 30: What do you mean by "Apparently"?*

The term "apparently" has been deleted

*42. P11, line 31: What's a "model world"?*

The sentence has been changed as follows: "The sharp changes in BLHs and 222Rn activity concentrations are due to the relatively coarse temporal resolution of ECMWF meteorological data (3-hourly for surface data (e.g., BLHs) and 6-hourly for 3D fields (temperature, wind, humidity, and convection); see Section 3.1)"

43. P11, line 44: “This finding...” What finding exactly?

This “finding” refers to the fact during daytime, the TM5 BLHs are close to IGRA measurements at most stations, while larger differences are observed between <sup>222</sup>Rn activity concentration simulated and observed. The text has been clarified

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).

See our reply to the reviewer's major comments (item 7). We have considered in the revised version of the paper only few relevant experimental settings. This makes the figures more readable

45. P10-11, Section 5: Can be better divided in more sections.

Yes. We add four more sub-sections as follows:

- Relationship between <sup>222</sup>Rn activity concentrations and BLHs
- Sensitivity of simulated <sup>222</sup>Rn activity concentrations to convection scheme
- Comparison of simulated and observed <sup>222</sup>Rn activity concentrations: Impact of sampling time
- Vertical gradients of <sup>222</sup>Rn activity concentrations in the boundary layer at Cabauw

46. P12, line 11: Add section 5.4 for this paragraph.

Yes. See our responses above

47. P12, line 38: “weather conditions”, I suppose you mean “stability regimes”?

Yes. Corrected and suggested

48. P12, line 42-43: About CB1 and CB4, see remarks about table.

The text has been revised. The Table 1 is quoted there, where CB1 and CB4 are defined

49. P13, line 14 (section 6): this section feels more like a summary than a conclusion.

This has been revised

50. P13, line 16: “dynamics” should rather be “height”. The dynamics are not evaluated.

See our reply to the reviewer's major comments (item 2). Therefore, we prefer to keep the term “dynamics”

*51. P13, line 19: 10-20%, this is the first time I see a statistical value between TM5 and observations. That's very late.*

Figure 11 (submitted version; Figure 8 in the revised version) that summarizes the statistics on BLHs and  $^{222}\text{Rn}$  activity concentrations is now commented earlier (see Section 4.2)

*52. P13, line 23: IGRA observations (not “data”).*

Changed as suggested

*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.*

We have kept the ceilometer/lidar observations in the analysis, but this part has slightly been revised

*54. P13, lines 26-35: remove ceilometer related analysis/results*

See our reply to the reviewer's major comments (see item 5)

*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.*

As clearly shown in Figure 11 (submitted version; Figure 9 in the revised version), the performance of the model simulations compared to the  $^{222}\text{Rn}$  activity concentration observations is very similar in term of e.g., root mean square and correlation coefficient for both convection schemes. The statistics are shown in Figure 11 (submitted version; Figure 8 in the revised version)

*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 (o) signs? What does Altitude/Height exactly mean? Do you mean Terrain elevation (Mean Sea Level) and Height (Above Ground Level), respectively?*

ANSTO stands for Australian Nuclear Science and Technology Organisation and was already defined in the submitted version of text (Section 2.2). The Table 1 (including its caption) has been revised (including definition of ANSTO) to clarify the points mentioned in the comments. However, as already mentioned, we do not qualify the stations as “coastal or non-coastal sites”

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.*

Figure 1 (including the caption) has been revised. Table 1 is now referred

58. *Figure 2. remove. See major/minor comments*

Figure 2 (submitted version) is now put in the Supplement (Figure S1 in the revised version)

59. *Figure 3. Vertical and horizontal axis in the upper diagram?*

They are latitude and longitude.

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.*

Figures 4-5 (submitted version; Figures 3-4 in the submitted version) have been revised. The scaling on y-axis is now the same for nocturnal BLHs and different, but the same for daytime BLHs. In general most of the figures of the paper have been revised, as mentioned above

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?*

Figure 1 in the revised version helps to have an idea about coastal and non-coastal stations, as already stressed on our reply to the reviewer's major comments (see item 6)

62. *Figure 6/7. Redundant/remove*

See our reply to the reviewer's major comments (5)

63. *Figure 11: remove CEIL/LIDAR data from figure.*

See our reply to the reviewer's major comments (item 5).

64. *Figure 12. What is ratio? TM5 divided by IGRA?*

Yes. This has been clarified. It is now Figure 10 in the revised version of the paper

65. *Figure 13. Abbreviations are not explained well in caption (e.g what is CBI and CB4?).*

*“Mean diurnal variations...” should probably be “Monthly mean diurnal variations...”*

Corrected by “Monthly mean diurnal variations” as suggested.

The acronyms CB1 and CB2 already defined in Table 1. These acronyms are now defined in the caption of Figure 13 (submitted version; Figure 11 in the revised version)

*66. Figure 13: Data points are outside the y-axis range. This should be corrected.*

Yes, but we did so because we wanted to focus on the variations of the vertical gradients of  $^{222}\text{Rn}$  activity concentrations around noon. We now increase a bit the y-axis, but some data points during night and early in the morning are still outside

*67. Figure 14. How can the R of the lower figure be almost the same as the R for the upper figure?*

This has been verified and it is correct. This is fortuitous. This is certainly due to the large values for daytime

***References:***

*Pillai, D., Gerbig, C., Marshall, J., Ahmadov, R., Kretschmer, R., Koch, T., & Karstens, U. (2010). High resolution modeling of CO<sub>2</sub> over Europe: implications for representation errors of satellite retrievals. Atmospheric Chemistry and Physics, 10(1), 83-94.*

## Evaluation of the boundary layer dynamics of the TM5 model over Europe

E. N. Koffi<sup>1</sup>, P. Bergamaschi<sup>1</sup>, U. Karstens<sup>2,3</sup>, M. Krol<sup>4,5,6</sup>, A. Segers<sup>7</sup>, M. Schmidt<sup>8,11</sup>, I. Levin<sup>8</sup>, A. T. Vermeulen<sup>9,3</sup>, R. E. Fisher<sup>10</sup>, V. Kazan<sup>11</sup>, H. Klein Baltink<sup>12</sup>, D. Lowry<sup>10</sup>, G. Manca<sup>1</sup>, H. A. J. Meijer<sup>13</sup>, J. Moncrieff<sup>14</sup>, S. Pal<sup>15</sup>, M. Ramonet<sup>11</sup>, H.A. Scheeren<sup>1,13</sup>, A. G. Williams<sup>16</sup>

1 European Commission Joint Research Centre, Institute for Environment and Sustainability, Ispra (Va), Italy

2 Max-Planck-Institute for Biogeochemistry, Jena, Germany

3 ICOS Carbon Portal, ICOS ERIC at Lund University, Sweden

4 SRON Netherlands Institute for Space Research, Utrecht, Netherlands,

5 Institute for Marine and Atmospheric Research Utrecht, Utrecht, University, Utrecht, Netherlands

6 MAQ, Wageningen University and Research Centre, Wageningen, Netherlands,

7 Netherlands Organisation for Applied Scientific Research (TNO), Utrecht, Netherlands

8 Institut für Umweltphysik, Heidelberg University, Germany

9 Energy research Center Netherlands (ECN), Petten, Netherlands

10 Royal Holloway, University of London (RHUL), Egham, UK

11 Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France

12 Royal Netherlands Meteorological Institute (KNMI), Netherlands

13 Centrum voor Isotopen Onderzoek (CIO), Rijksuniversiteit Groningen, Netherlands

14 Atmospheric Chemistry Research Group, University of Bristol, UK

15 Department of Meteorology, Pennsylvania State University, State College, PA, USA

16 Australian Nuclear Science and Technology Organisation (ANSTO) Environment Research Theme, Locked Bag 2001, Kirrawee DC, NSW 2232, Australia

26 July2016

Revised version for Geosci. Model Dev.

## Abstract

We evaluate the capability of the global atmospheric transport model TM5 to ~~reproduce observations of~~ simulate the boundary layer dynamics and ~~the~~ associated variability of trace gases close to the surface, using radon ( $^{222}\text{Rn}$ ), ~~which is an excellent tracer for vertical mixing owing to its short lifetime (half life) of 3.82 days~~. Focusing on the European scale, we compare the boundary layer height (BLH) in the TM5 model with observations from the NOAA Integrated Global Radiosonde Archive (~~IGRA~~) and ~~also in addition~~ with ceilometer ~~measurements at Cabauw (The Netherlands)~~ and lidar BLH retrievals at two stations. ~~Trainou (France)~~. Furthermore, we compare TM5 simulations of  $^{222}\text{Rn}$  activity concentrations, using a novel, process-based  $^{222}\text{Rn}$  flux map over Europe (Karstens et al., 2015), ~~with~~ harmonized quasi-continuous  $^{222}\text{Rn}$  measurements ~~from at~~ 10 European monitoring stations.

The TM5 model reproduces relatively well the daytime BLH (within ~~~~~10-20% for most of the stations), except for coastal sites, for which differences are usually larger due to model representation errors. During night, however, TM5 overestimates the shallow nocturnal BLHs, especially for the very low observed BLHs (< 100 m) during summer.

The  $^{222}\text{Rn}$  activity concentration simulations based on the new  $^{222}\text{Rn}$  flux map show significant improvements especially regarding the average seasonal variability, compared to simulations using constant  $^{222}\text{Rn}$  fluxes. Nevertheless, the (relative) differences between simulated and observed daytime minimum  $^{222}\text{Rn}$  activity concentrations are larger for several stations (on the order of 50%) ~~than compared to~~ the (relative) differences between simulated and observed BLH at noon. Although the nocturnal BLH is often higher in the model than observed, simulated  $^{222}\text{Rn}$  nighttime maxima are actually larger at several continental stations. This counterintuitive behaviour, ~~which~~ points to potential deficiencies of TM5 to correctly simulate the vertical gradients within the nocturnal boundary layer, limitations of the  $^{222}\text{Rn}$  flux map, or issues related to the definition of the nocturnal BLH.

At several stations the simulated decrease of  $^{222}\text{Rn}$  activity concentrations in the morning is faster than observed. In addition, simulated vertical  $^{222}\text{Rn}$  activity concentration gradients at Cabauw decrease faster than observations during the morning transition period, and are in general lower than observed gradients during daytime. Although these effects may be partially due to the slow response time of the radon detectors, they clearly ~~which~~ points to too fast vertical mixing in the TM5 boundary layer during daytime. Furthermore, the capability of the TM5 model to simulate the diurnal BLH cycle is limited ~~by~~ due to the current coarse temporal resolution (3hr/6hr) of the TM5 input meteorology.

~~Additionally, we analyze the impact of a new treatment of convection in TM5, based on the ECMWF reanalysis, leading to overall significantly lower (on the order of ~20%) surface  $^{222}\text{Rn}$  activity concentrations during daytime compared to the current default convection scheme based on Tiedtke (1989). However, the performance of the model simulations compared to the  $^{222}\text{Rn}$  observations is very similar in terms of root mean square and correlation coefficient for both convection schemes.~~

## 1. Introduction

The boundary layer, being the lowest portion of the atmosphere, is largely affected by the Earth's surface forcing. This layer is usually separated from the free troposphere (where the surface effects are weak) by a thin and strongly stable layer (capping inversion) that traps turbulence, moisture, and trace gases below in the boundary layer. The thickness of the boundary layer is variable in space and time and can range from tens of meters to 4 km, depending on both the synoptic and local meteorological conditions (Stull, 1988). The height of the boundary layer is ~~an~~ essential a critical parameter in atmospheric transport models, since it controls the extent of the vertical mixing of trace gases emitted near the surface. Previous studies that evaluated The the ability of global-atmospheric transport models to reproduce the boundary layer dynamics demonstrated the importance of temporal resolution of meteorological data, horizontal and vertical model resolutions, and parameterizations of vertical mixing has been investigated earlier (e.g., Denning et al., 1999; Dentener et al., 1999); The authors have recommended the use of both high temporal resolution of meteorological data within the lower levels (Dentener et al., 1999) and fine horizontal and vertical resolutions (Krol et al., 2005; Locatelli et al., 2015) for a better reproduction of the meso-scale processes in the model. The realistic simulation of ~~the~~ boundary layer height (BLH) is crucial, especially for inverse modelling simulations that aim at estimating surface fluxes from observed concentrations. This is the case in particular for regional flux inversions which make use of regional networks of surface and tower-based trace gas concentration measurements to that capture the signals offrom regional sources (and sinks). Regional inversions of greenhouse gases (GHG) (CO<sub>2</sub>, CH<sub>4</sub>, N<sub>2</sub>O, halocarbons) were reported especially for Europe and North America, making use of the increasing number of regional monitoring stations in these areas (e.g., Gerbig et al., 2003; Carouge et al., 2008; ~~Kort et al., 2008;~~ Bergamaschi et al., 2010; Corazza et al., 2011; Manning et al., 2011; Broquet et al., 2013; Bergamaschi et al., 2015; Ganesan et al., 2015) as well as aircraft observations (e.g., Kort et al., 2008; Miller et al. 2013).

In order to evaluate the quality of such flux inversions, a thorough validation of the applied atmospheric transport models is essential. In this study, we present a detailed evaluation of the boundary layer dynamics of the TM5 model (Krol et al., 2005), which is the global transport model used in the TM5-4DVAR inverse modelling framework-system (Meirink et al., 2008), applied in several of the European inversions mentioned above (Corazza et al., 2011; Bergamaschi et al., 2010; 2015). AsIn a first step, we compare the model BLH with the sounding-derived BLH of the NOAA Integrated Global Radiosonde Archive (IGRA) (Seidel et al., 2012) at European scale. Radiosonde data have been considered to give the most accurate BLHs (Collaud Coen et al., 2014). The model BLHs are also compared to those derived from the ceilometer and lidar measurements at two European stations (Cabauw and Trainou). As a second step, we compare TM5 simulations of <sup>222</sup>Rn activity concentrations with measurements at 10 European stations. <sup>222</sup>Rn is an excellent tracer for boundary layer mixing due to its short lifetime (half-life) of 3.82 days and has been widely used for model validation (e.g., Jacob and Prather, 1990; Jacob et al., 1997; Dentener et al., 1999; Chevillard et al., 2002; Taguchi et al., 2011) and mixing studies (e.g., see reviews in Zahorowski et al., 2004; Chambers et al., 2011; Williams et al., 2011, 2013). However, the use of <sup>222</sup>Rn for this purpose has been limited by the simplified

1 assumption of constant  $^{222}\text{Rn}$  fluxes over land used in most  $^{222}\text{Rn}$  validation studies published so  
2 far. It has also been limited by the fact that the observed  $^{222}\text{Rn}$  activity concentrations from  
3 different stations were not harmonized.

4 Here, we make use of a novel detailed  $^{222}\text{Rn}$  flux map over Europe (Karstens et al., 2015) based  
5 on a parameterization of  $^{222}\text{Rn}$  production and transport in the soil as well as improved observed  
6  $^{222}\text{Rn}$  activity concentrations obtained through a detailed comparison study (Schmithüsen et al.,  
7 2016). The development of this  $^{222}\text{Rn}$  flux map has been performed within the European project  
8 InGOS ('Integrated non-CO<sub>2</sub> Greenhouse gas Observing System'), including also a comparison  
9 of different transport models (including TM5). While this model comparison will be published  
10 elsewhere (Karstens et al., 2016, manuscript in preparation), we present here the analysis for the  
11 TM5 model aiming at the identification and quantification of potential systematic errors in the  
12 simulation of the BLH dynamics, which could directly translate into systematic errors in the  
13 derived surface fluxes. Our study also includes the evaluation of a new parameterization of  
14 convection in TM5, based on ECMWF (re)analysis, compared to the default convection scheme  
15 used so far, based on the parameterization of Tiedtke (1989).

## 17 2. Observations

### 20 2.1. Boundary layer height

21  
22 Vertical mixing in the atmospheric boundary layer is mostly turbulent. The BLH is confined by a  
23 thin vertical layer where steep vertical gradients of meteorological variables, pollutants, trace  
24 gases, and aerosols occur. Consequently, all the observational devices built for the retrieval of  
25 BLH are based on the search of the height at which the strongest gradients occur. These  
26 gradients can be based either on the atmospheric potential temperature profile, the wind profile,  
27 or the aerosol backscatter profile. For meteorological data sets and atmospheric transport models,  
28 the bulk Richardson number, a dimensionless parameter defined as the ratio between the of  
29 turbulence due to buoyancy consumption by thermal stability and the mechanic generation of  
30 turbulence by wind shear, has been widely used to determine BLHs (e.g., Vogelezang and  
31 Holtslag, 1986; Seibert et al., 2000; Seidel et al., 2012). Thus, the BLH is the vertical level at  
32 which the bulk Richardson number ( $R_{ib}$ ) computed from the ground reaches a critical value  $R_{ic}$   
33 characterizing the passage of turbulent fluid flow to laminar one. The general expression of  
34 Vogelezang and Holtslag (1986) used to compute  $R_{ib}$  is given as follows:

$$36 R_{ib} = \left( \frac{g}{\theta_{vs}} \right) \frac{(\theta_{vh} - \theta_{vs})(z_h - z_s)}{(u_h - u_s)^2 + (v_h - v_s)^2 + bu_*^2} \quad (1)$$

37  
38 where  $g$  is the gravitational acceleration ( $9.81 \text{ m s}^{-2}$ ),  $h$  the geopotential height of the model,  $\theta_{vh}$ ,  
39  $\theta_{vs}$  the virtual potential temperature,  $z$  the geopotential height, at the surface and  $\theta_{vh}$  the virtual  
40 potential temperature at the model level  $h$ .  $u_s$  denotes the zonal wind speed, and at the surface and  
41  $u_h$  the zonal wind speed at the model level  $h$ .  $v_s$  the meridional wind speed. The indices  $h$  and  $s$   
42 denote the vertical layer, and the surface, respectively. at the surface and  $v_h$  the meridional wind  
43 speed at the model level  $h$ .  $bu_*^2$  depicts the turbulence production due to the surface friction, a  
44 term which also prevents an undetermined  $R_{ib}$  in case of uniform high wind speeds relevant for  
45 neutral boundary layers.  $b$  is a coefficient estimated to be 100 (Vogelezang and Holtslag, 1986)

1 and  $u^*$  is the surface friction velocity. The geopotential height  $z_s$   ~~$z_h$  and  $z_s$  are~~ is expressed in m.  
2 The virtual potential temperature  $\theta_v$  is in K, and the velocities are in  $\text{ms}^{-1}$ .  
3 The vertical profile of  $R_{ib}$  is linearly interpolated between consecutive vertical layers. The BLH  
4 is defined as the height, where  $R_{ib}$  reaches the critical value  $R_{ic}$ . Commonly, a  $R_{ic}$  value of 0.25  
5 has been used (e.g., Vogelezang and Holtslag, 1986; Seibert et al., 2000; Seidel et al., 2012). The  
6 boundary layer height is defined with reference to surface elevation, and not to sea level (Seidel  
7 et al., 2012).

### 10 2.1.1. IGRA data

12 We use BLHs of the NOAA IGRA database, which covers the 1990-2010 period (Seidel et al.,  
13 2012). The IGRA data is based on radiosonde measurements that are usually released at 00 and  
14 12 UTC. The IGRA radiosonde network over Europe is ~~presented~~ shown in **Figure 1**. The  
15 dynamic (wind speed and direction) and thermal (temperature and humidity) profiles from the  
16 radiosondes are utilized to compute BLHs using the bulk Richardson number method [Eq.1;  
17 Section ~~3-22.1~~]. In these BLH calculations both the surface wind (i.e.,  $u_s$  and  $v_s$  in Eq.1) and the  
18 surface friction velocity ( $u^*$ ) are unknown and set to zero. The critical value of the bulk  
19 Richardson number ( $R_{ic}$ ) is set to 0.25 (instead of 0.3 as used in TM5; see Section 3.2). Further  
20 details on the choice of the settings as well as the vertical profiles of the dynamic,  
21 thermodynamic, and bulk Richardson number quantities are described in Seidel et al. (2012).  
22 These settings for the IGRA database were also adopted in the InGOS protocol for the evaluation  
23 of the transport models involved in InGOS inverse modelling analyses (Karstens et al., 2016,  
24 manuscript in preparation). The methodological uncertainties in the IGRA BLH data were  
25 evaluated based on paired soundings released at the same site (Seidel et al., 2012). Results show  
26 that the choice of  $R_{ic}$  does not introduce large uncertainty, but other methodological choices  
27 (including surface wind speed estimates and vertical interpolation of the bulk Richardson number  
28 profile) as well as the vertical resolution of the sounding data are larger sources of uncertainty in  
29 the derived BLHs (Seidel et al., 2012). The authors reported relative uncertainties in the IGRA  
30 BLHs that can be large (>50%) for shallow BLHs (< 1 km; mainly observed during night or  
31 early in the morning), but much smaller (usually <20%) for deep BLHs (> 1 km) during daytime.

### 33 2.1.2. Lidar and ceilometer data

35 The principle of LIDAR (LIght Detection And Ranging; hereafter lidar) is based on a pulsed  
36 laser light emitted into the atmosphere which is back-scattered by aerosol particles and  
37 molecules. The lidar algorithms derive the BLHs by searching the location of the strongest  
38 aerosol gradient in the vertical dimension (e.g., Haeffelin et al., 2011; Pal et al., 2012; Griffiths et  
39 al., 2013; Pal et al., 2015). A ceilometer is a 'low-cost lidar' which was initially used for the  
40 detection of cloud base heights. However, since the backscatter signal of aerosols is lower than  
41 that of clouds, the sensitivity of ceilometers in retrieving the boundary layer height is much less  
42 than that of lidar instruments (Pal, 2014). In contrast to IGRA data (i.e., radiosonde based BLH),  
43 the ceilometer and lidar allow measurements of the diurnal BLH cycle. However, the algorithms  
44 of both lidar and ceilometer have some difficulties to assign the BLH during night and tend to  
45 wrongly attribute the height of the residual layer of aerosol (often with larger signal) as height of  
46 the real mixed layer (e.g., Angevine et al., 1998; Eresmaa et al., 2006; Hajj et al., 2006).

1 Lidar/ceilometer nocturnal BLHs are also higher due to the fact that their overlap height can be  
2 above the nocturnal shallow BLH (Pal et al., 2015). Uncertainties in lidar retrieved BLHs were  
3 assessed based on a comparison between radiosonde based BLHs and wavelet derived BLH  
4 estimates from lidar and found to be about 60 m (Pal et al., 2013).

5  
6 We use the BLHs retrieved from lidar and ceilometer measurements at Trainou and Cabauw,  
7 respectively (see **Figure 1** for their locations). The lidar (ALS-300) measurements at Trainou are  
8 described by Pal et al. (2012). The ceilometer at Cabauw is part of the network of the Vaisala  
9 LD-40 ceilometer in the Netherlands operated by the Royal Netherlands Meteorological Institute  
10 (KNMI; Haij et al., 2006). We analyze the ceilometer measurements at Cabauw for 2010 and the  
11 lidar data at Trainou for 2011. For Cabauw we compare the ceilometer based BLH for 2010 with  
12 the BLH data from the closest IGRA station (De Bilt), with results ~~at 12 UTC~~ shown in the  
13 Supplement (Figure S1).Figure 2.

## 17 2.2. Observed $^{222}\text{Rn}$ activity concentrations

19 The observed  $^{222}\text{Rn}$  activity concentrations are obtained from 2 different measurement methods:  
20 (1) The 'two-filter' method developed by the Australian Nuclear Science and Technology  
21 Organisation (ANSTO) (Whittlestone and Zahorowski, 1998; Chambers et al., 2011). After  
22 drawing the sampled air continuously through a delay volume to let all short-lived  $^{220}\text{Rn}$  (thoron)  
23 gas in the sampled air decay, it passes through a first filter that removes all ambient  $^{222}\text{Rn}$  and  
24  $^{220}\text{Rn}$  decay products. Filtered air then enters in a delay chamber in which new  $^{222}\text{Rn}$  progeny  
25 ( $^{218}\text{Po}$  and  $^{214}\text{Po}$ ) are produced. An internal-second flow loop within the delay chamber passes the  
26 air through a second filter, which collects the new  $^{222}\text{Rn}$  progeny formed under controlled  
27 conditions. Hence, in the ANSTO system  $^{222}\text{Rn}$  activity concentration in the sampled air is  
28 measured directly through its newly formed progeny within the controlled environment of the  
29 delay chamber in the sampled air (Whittlestone and Zahorowski, 1998; Zahorowski et al., 2004;  
30 Chambers et al., 2011). In routine operation, ANSTO monitors are calibrated monthly by  
31 injecting  $^{222}\text{Rn}$  from a well characterized (to about  $\pm 4\%$ )  $^{226}\text{Radium}$  source. For ambient air  
32 measurements at  $1 \text{ Bq m}^{-3}$  activity concentration, the total uncertainty of hourly measurements is  
33 of order 10%, which includes uncertainty in flow rate as well as counting statistics. ANSTO two-  
34 filter detectors have a response time of around 45min, and are quite bulky ( $\sim 3\text{m}^2$ ) which can  
35 hinder their deployment in constricted locations.

37 (2) The one-filter methods used at the European stations are all based on the direct collection and  
38 counting of the short-lived ambient  $^{222}\text{Rn}$  and  $^{220}\text{Rn}$  ( $^{212}\text{Pb}$ ) decay products, which are attached to  
39 aerosols in the sampled air. These decay products are accumulated on either static or moving  
40 aerosol filters and measured by  $\alpha$  or  $\beta$  spectroscopy (see references given in **Table 1**). In order to  
41 derive the atmospheric  $^{222}\text{Rn}$  activity concentration, this method requires corrections for the  
42 atmospheric radioactive disequilibrium between the measured  $^{222}\text{Rn}$  daughters ( $^{214}\text{Po}$  and/or  
43  $^{218}\text{Po}$ ) and  $^{222}\text{Rn}$  (e.g., Levin et al., 2002).

45 We use  $^{222}\text{Rn}$  activity concentration measurements from 10 European stations over the 2006-  
46 2011 period (**Figure 1** and **Table 1**). The data from the different stations have been harmonized

1 based on an extensive comparison study performed within the InGOS project (Schmithuesen et  
2 al., 2016). Based on the tall tower measurements at Cabauw and Lutfjewad conducted at different  
3 heights above ground level as well as on an earlier comparison at Schauinsland station (Xia et  
4 al., 2010) and new comparison measurements in Heidelberg with an ANSTO system, correction  
5 factors for disequilibrium have also been estimated (Schmithuesen et al., 2016). All data used in  
6 the present study have been corrected accordingly and brought to a common ANSTO scale. A  
7 typical uncertainty of  $^{222}\text{Rn}$  data from the different one-filter systems, including the uncertainty  
8 of the disequilibrium is estimated to 10-15%±10 to 15%.

9  
10 At the monitoring station Ispra,  $^{222}\text{Rn}$  activity concentration has been measured using an ANSTO  
11 instrument, sampling air at an inlet positioned at 3.5m above the ground, close to the GHG-  
12 sampling mast with a height of 15m. Recent additional  $^{222}\text{Rn}$  measurements using the 15m inlet  
13 of the GHG mast (employing an Alphaguard PQ2000 (Genitron) instrument, calibrated against  
14 the ANSTO monitor) revealed significant differences of the  $^{222}\text{Rn}$  activity at the two sampling  
15 heights during periods with low wind speeds. These differences showed that there are significant  
16 vertical  $^{222}\text{Rn}$  gradients close to the ground. Based on the comparison of the two sampling  
17 heights during a 3-month period, we derive a wind-speed dependent correction, in order to  
18 'normalize' the entire time series of the ANSTO measurements (at 3.5m above ground) to the  
19 15m inlet, which is considered to be more representative. The uncertainty of this wind-speed  
20 dependent correction (based on the  $1\sigma$  standard deviation during the 3-month comparison) is  
21 included in the time series shown in the Supplement (**Figure S248**).

### 22 23 24 **3. Models simulations**

#### 25 26 3.1. TM5 Model

27  
28 TM5 is a global chemistry transport model, which allows two-way nested zooming (Krol et al.,  
29 2005). In this study we apply the zooming with  $1^\circ \times 1^\circ$  resolution over Europe, while the global  
30 domain is simulated at a horizontal resolution of  $6^\circ$  (longitude)  $\times$   $4^\circ$  (latitude). TM5 is an offline  
31 transport model, driven by meteorological fields from the European Centre for Medium-Range  
32 Weather Forecasts (ECMWF) Integrated Forecast System (IFS) ERA-Interim reanalysis (Dee et  
33 al., 2011). The spatial resolution of this data set is approximately 80 km (T255 spectral) on 60  
34 vertical levels from the surface up to 0.1 hPa. We employ the standard TM5 version with 25  
35 vertical levels, defined as a subset of the 60 We use 25 vertical layers of the ERA-Interim  
36 reanalysis (extending up to 0.2 hPa). The extraction of the meteorological fields is performed  
37 through a pre-processing software, which supplies fully consistent meteorology data with those  
38 of ECMWF at the different spatial resolutions of TM5 (Krol et al., 2005). The boundary layer,  
39 the free troposphere, and the stratosphere are represented by 5 (up to 1 km), 10, and 10 layers,  
40 respectively. The temporal resolution of the data is 3-hourly for near surface data (e.g., BLHs)  
41 and 6-hourly for 3D fields (e.g., temperature, wind, humidity, and convection).

42  
43 Tracers in TM5 are transported by advection (in both horizontal and vertical directions), cumulus  
44 convection, and vertical diffusion. Tracer advection is based on the so-called "slopes scheme"  
45 which considers a tracer mass within a grid cell as a mean concentration and the spatial gradient  
46 of the concentration within the grid box (Russel and Lerner, 1981), which is caused by the

1 motion of the tracer into and out of the grid box. Non-resolved transport by shallow cumulus and  
 2 deep convection in TM5 is parameterized by a bulk mass flux approach originally described in  
 3 Tiedtke (1989). Such convective clouds are described by single pairs of entraining/detraining  
 4 plumes representing the updraft/downdraft motion. The parameterization of the vertical turbulent  
 5 diffusion in the boundary layer is based on the scheme of Holtslag and Moeng (1991), while the  
 6 formulation of Louis (1979) is considered in the free troposphere. The BLH is computed by  
 7 using the expression of Vogelezang and Holtslag (1986), as described in Section 3.2.2.1. The  
 8 exchange coefficients from the vertical diffusion are combined with the vertical convective mass  
 9 fluxes to calculate the sub-grid scale vertical tracer transport. After redistributing the tracer mass  
 10 by convection and diffusion, the slopes are updated.

11 ~~Recently, van der Veen (2013) proposed a revised scheme to update the slopes. This “revised~~  
 12 ~~slopes scheme” results in~~ Since in convective areas, transport in the vertical can be more efficient  
 13 ~~than in the horizontal, van der Veen (2013) enhanced horizontal transport in TM5 by increasing~~  
 14 ~~the horizontal diffusivity of the numerical scheme of the convection routine~~ decreased the vertical  
 15 ~~slopes (called “updated slopes treatment” in Section 4) through an adjustment scheme. van der~~  
 16 ~~Veen (2013)The author~~ found an improvement of the inter-hemispheric mixing gradient in TM5,  
 17 which was initially underestimated as reported in e.g., Patra et al. (2011). This “revised updated  
 18 slopes ~~treatmentscheme~~” has been used for the sensitivity tests described below. Furthermore, we  
 19 performed sensitivity tests using directly the convection fields from the ECMWF IFS model,  
 20 instead of the default convection scheme based on Tiedtke (1989). The ECMWF convection  
 21 scheme includes several improvements of the parameterizations of deep convection, radiation,  
 22 clouds and orography, introduced operationally since ECMWF ERA-15 analyses (e.g., Gregory  
 23 et al., 2000; Jakob and Klein, 2000; Morcrette et al., 2001). Finally, we evaluate the  
 24 combination of the “revised updated-slopes scheme” and the use of ECMWF convection fields.

### 3.2. TM5 Boundary layer height scheme

28 ~~Vertical mixing in the atmospheric boundary layer is mostly turbulent. The BLH is confined by a~~  
 29 ~~thin vertical layer where steep vertical gradients of pollutants, trace gases, and aerosol occur.~~  
 30 ~~Consequently, all the observational devices built for the retrieval of BLH are based on the search~~  
 31 ~~of the height at which the strongest gradients occur. These gradients can be in either the~~  
 32 ~~atmospheric potential temperature profile, the wind profile, or the aerosol backscatter profile. For~~  
 33 ~~meteorological and atmospheric transport models, the bulk Richardson number, a dimensionless~~  
 34 ~~parameter defined as the ratio between the buoyant consumption by thermal stability and the~~  
 35 ~~mechanic generation by wind shear, has been widely used to determine BLHs (e.g., Vogelezang~~  
 36 ~~and Holtslag, 1986; Seibert et al., 2000; Seidel et al., 2012). Thus, BLH is the vertical level at~~  
 37 ~~which the bulk Richardson number ( $R_{ib}$ ) computed from the ground reaches a critical value  $R_{ic}$~~   
 38 ~~characterizing the passage of turbulent fluid flow to laminar one. In the TM5 model, the~~  
 39 ~~expression of Vogelezang and Holtslag (1986) is used to compute  $R_{ib}$ , as follows:~~

$$R_{ib} = \frac{g}{\theta_{vs}} \frac{(\theta_{vh} - \theta_{vs})(h - z_s)}{(u_h - u_s)^2 + (v_h - v_s)^2 + b u_*^2} \quad (1)$$

42 where  $g$  is the gravitational acceleration ( $9.81 \text{ m s}^{-2}$ ),  $h$  the geopotential height of the model,  $\theta_{vs}$   
 43 the virtual potential temperature at the surface and  $\theta_{vh}$  the virtual potential temperature at the  
 44 model level  $h$ .  $z_s$  corresponds to the surface geopotential height.  $u_s$  denotes the zonal wind speed  
 45

1 at the surface and  $u_h$  the zonal wind speed at the model level  $h$ .  $v_g$  denotes the meridional wind  
2 speed at the surface and  $v_h$  the meridional wind speed at the model level  $h$ .  $bu_*^2$  depicts the  
3 turbulence production due to the surface friction, a term which also prevents an undetermined  $R_{ib}$   
4 in case of uniform high wind speeds relevant for neutral boundary layers.  $b$  is a coefficient  
5 determined to be 100 (Vogelezang and Holtslag, 1986) and  $u_*$  is the surface friction velocity.  
6 The geopotential heights  $h$  and  $z_g$  are expressed in m. The potential temperature is in K and the  
7 velocities are in m/s.

8  
9 In the TM5 model, the full expression of Vogelezang and Holtslag (1986) is used to compute  $R_{ib}$ ,  
10 (Eq. 1). First,  $R_{ib}$  is computed at each model level by using the equation (1). The vertical profile  
11 of  $R_{ib}$  is then linearly interpolated between consecutive levels of the model. The BLH is defined  
12 as the height, where  $R_{ib}$  reaches the critical value  $R_{ic}$ . In TM5,  $R_{ic}$  is set to 0.3. As discussed in  
13 Seidel et al. (2012), the choice of  $R_{ic}$  value close to 0.25 does not introduce large uncertainty.  
14 Moreover, the differences between BLHs determined by using  $R_{ic}$  of 0.25 and  $R_{ic}$  of 0.3 are very  
15 small as shown in this study (see e.g., **Figures S2-S11** in the Supplement). In the TM5 model,  
16 the minimum BLH is set to 100 m.

17  
18 ~~The vertical profile of  $R_{ib}$  is linearly interpolated from the first layer of the model until  $R_{ib}$~~   
19 ~~reaches its critical value  $R_{ic}$ . Commonly, a  $R_{ic}$  value of 0.25 has been used (e.g., Vogelezang and~~  
20 ~~Holtslag, 1986; Seibert et al., 2000; Seidel et al., 2012) while in TM5 a  $R_{ic}$  value of 0.3 has been~~  
21 ~~applied. Moreover, the minimum BLH in TM5 is set to 100 m.~~

22 In addition, we calculate the BLH in TM5 based on the definition of Seidel et al. (2012) as used  
23 in the InGOS model validation exercise (i.e.,  $R_{ic} = 0.25$  and both surface wind and friction  
24 velocity are set to zero in Eq.1; see Section 2.1). Furthermore, because InGOS and IGRA sites  
25 are not co-located, we extract the BLH in the model both at the location of the InGOS station and  
26 at the location of the nearest IGRA station, resulting two sets of modeled BLHs labelled by the  
27 following acronyms:

- 28 • 'TM5 INGOS': BLHs extracted at InGOS station
- 29 • 'TM5 INGOS IGRA': BLHs extracted at IGRA station, which is closest to the selected  
30 InGOS station.

31  
32 In both cases, we use a 2-dimensional interpolation (longitude / latitude) to the location of the  
33 (InGOS or IGRA) station.

34 Furthermore, we extract also the default TM5 BLH (both at the InGOS and IGRA station) and  
35 the BLHs from ECMWF reanalyses. In general, the difference between the BLH based on Seidel  
36 et al. (2012) and the TM5 default and ECMWF BLHs are very small. Therefore, the latter are  
37 only shown in the Supplement (**Figures S2-S11**).

### 40 3.3 InGOS $^{222}\text{Rn}$ flux map

41  
42 We use the new  $^{222}\text{Rn}$  flux map developed by Karstens et al. (2015) within the InGOS project  
43 (called hereafter 'InGOS  $^{222}\text{Rn}$  flux map'). This map is based on a parameterization of  $^{222}\text{Rn}$   
44 production and transport in the soil, using a deterministic model based on the equations of

1 continuity and diffusion (Fick's 1<sup>st</sup> law) to compute the transport of the <sup>222</sup>Rn flux from the soil  
2 to the atmosphere. The modelled radon flux is dependent on soil porosity and moisture, with the  
3 latter obtained from two different soil moisture data sets, i.e., from the Land Surface Model  
4 Noah (driven by NCEP-GDAS meteorological reanalysis and part of the Global Land Data  
5 Assimilation System (GLDAS); Rodell et al., 2004), and from the ERA-Interim/Land reanalysis,  
6 respectively. Karstens et al. (2015) found that the flux estimates based on the GLDAS Noah soil  
7 moisture model on average better represent observed fluxes. Therefore, we apply in this study  
8 ~~this study we apply~~ the <sup>222</sup>Rn flux map version based on the Noah soil moisture data set.  
9 Furthermore, the <sup>222</sup>Rn flux map considers the water table (from a hydrological model  
10 simulation), the distribution of the <sup>226</sup>Ra content in the soil, and the soil texture. For comparison,  
11 we apply also the commonly used constant emission maps with uniform continental <sup>222</sup>Rn  
12 exhalation of 21.98-mBqm<sup>-2</sup>s<sup>-1</sup> between 60°S and 60°N; uniform continental <sup>222</sup>Rn emissions of  
13 11.48 mBqm<sup>-2</sup>s<sup>-1</sup> between 60°N and 70°N (excluding Greenland); and zero flux elsewhere (Jacob  
14 et al., 1997). The InGOS <sup>222</sup>Rn flux map provides monthly <sup>222</sup>Rn fluxes over the 2006-2011  
15 period, aggregated to 0.5°×0.5° grid for Europe and complemented by the constant emissions for  
16 the regions outside Europe. **Figures 23a** and **23b** illustrate the spatial and mean seasonal  
17 variations of the <sup>222</sup>Rn fluxes from the InGOS <sup>222</sup>Rn flux map over Europe. The modelled <sup>222</sup>Rn  
18 flux is found to be larger in the areas where the <sup>226</sup>Ra activity concentration in the upper soil is  
19 very high, such as the Iberian Peninsula, areas in Central Italy and the Massif Central in Southern  
20 France (**Figure 23a**). The mean seasonal variations of the <sup>222</sup>Rn fluxes are mainly driven by the  
21 soil moisture. On average, the InGOS <sup>222</sup>Rn emissions over Europe are smaller than the constant  
22 emission (except July - September; **Figure 23b**).

## 23 ~~4. Simulation setup~~

### 24 ~~4.1. Model boundary layer heights~~

25  
26  
27  
28 ~~We extract the TM5 BLHs using either the TM5 default expression of  $R_{ic}$  (Section 3.2),~~  
29 ~~representing the effective BLH in the TM5 simulations, or based on Seidel et al. (2012) used in~~  
30 ~~the InGOS model validation exercise (i.e.,  $R_{ic} = 0.25$  and both surface wind and friction velocity~~  
31 ~~are set to zero in Eq.1; see Section 3.2). Furthermore, because InGOS and IGRA sites are not co-~~  
32 ~~located, we extract the BLH in the model both at the location of the InGOS station and at the~~  
33 ~~location of the nearest IGRA station, resulting in a set of four different modeled BLHs labelled~~  
34 ~~by the following acronyms:~~

- 35 ~~• 'TM5': TM5 default version (Eq.1 in Section 3.2 with  $R_{ic} = 0.3$ ); extracted at InGOS~~  
36 ~~stations by using 2D interpolation~~
- 37 ~~• 'TM5\_IGRA': As 'TM5', but extracted at IGRA station, which is closest to the selected~~  
38 ~~InGOS station~~
- 39 ~~• 'TM5\_INGOS': BLHs computed in TM5 model adopting the InGOS definition of the~~  
40 ~~BLH (i.e.,  $R_{ic} = 0.25$  and both surface wind and stress velocity are set to zero in Eq.1),~~  
41 ~~extracted at InGOS station.~~
- 42 ~~• 'TM5\_INGOS\_IGRA': As 'TM5\_INGOS', but extracted at IGRA station, which is closest~~  
43 ~~to the selected InGOS station~~

44  
45 ~~Furthermore, we evaluate the BLHs as provided by ECMWF analyses and interpolated to TM5~~  
46 ~~grids (labelled 'ECMWF'). The values of these BLHs are extracted only at the InGOS stations.~~

1 | ~~The ECMWF BLH is determined using an entraining parcel method, selecting the top of~~  
2 | ~~stratocumulus, or cloud base in shallow convection situations (Dee et al., 2011).~~

#### 5 | ~~4.2.~~ 3.4. Simulated $^{222}\text{Rn}$ activity concentrations

7 | We simulate  $^{222}\text{Rn}$  activity concentrations using either the InGOS  $^{222}\text{Rn}$  flux map based on Noah  
8 | soil moisture data, or constant  $^{222}\text{Rn}$  fluxes (see Section 3.3). Furthermore, we apply also the  
9 | “revised slopes scheme” and the updated convection scheme based on ECMWF reanalyses (see  
10 | section 3.1) ~~four different convection schemes in the TM5 model~~ (for the InGOS  $^{222}\text{Rn}$  flux map  
11 | based simulations only). These different simulations are labelled by the following acronyms:

- 12 | • FC\_CT: constant  $^{222}\text{Rn}$  fluxes, and default convection scheme in TM5 based on Tiedtke  
13 | (1989)
- 14 | • FI\_CT: InGOS  $^{222}\text{Rn}$  flux map, and default convection
- 15 | • ~~FI\_CS: InGOS  $^{222}\text{Rn}$  flux map and updated treatment of slopes in the TM5 convection~~  
16 | ~~scheme (see Section 3.1)~~
- 17 | • ~~FI\_CE: InGOS  $^{222}\text{Rn}$  flux map and the updated convection scheme based on ECMWF~~  
18 | ~~reanalyses (see Section 3.1).~~
- 19 | • FI\_CU: InGOS  $^{222}\text{Rn}$  flux map by using both the “revised updated treatment of slopes  
20 | scheme” and the updated convection scheme based on ECMWF reanalyses

22 | We also analyzed the use of “revised slopes scheme” and the updated convection scheme  
23 | independently (see Supplement; Figures S14-S24)

25 | The model simulations are 3-D linearly interpolated (i.e. horizontally and vertically) to the  
26 | location of the station, and averaged over 1 hour.

## 28 | 45. Results

### 30 | 45.1. Simulated boundary layer heights versus observations

32 | We focus the analysis on the InGOS stations (measuring  $\text{CH}_4$  and  $\text{N}_2\text{O}$ , and / or  $^{222}\text{Rn}$  activity  
33 | concentrations; **Figure 1**) at low altitudes (i.e., excluding mountain stations) and compare the  
34 | modelled BLHs with observations at the closest IGRA stations. **Figures 34** and **45** show the  
35 | mean seasonal variation for the nocturnal (00 UTC) and daytime (12 UTC) BLH, respectively  
36 | (2006-2010 average). The nocturnal BLHs show a clear seasonal cycle at most stations, with  
37 | typically higher nocturnal BLHs during winter (but also larger range between 25% and 75%  
38 | percentile) compared to summer. This seasonal pattern is very consistent between measurements  
39 | and model simulations. However, at some continental stations (e.g. Heidelberg, Gif-sur-Yvette)  
40 | the IGRA data show very low nocturnal BLHs (median value below 100m) during summer,  
41 | which are not reproduced by the models ~~(in particular not by the TM5 default BLH, which has an~~  
42 | ~~algorithmic internal lower limit of 100m)~~. In general, the Whisker plots (**Figure 34**) show a  
43 | skewed (non-normal) distribution for most monthly data (observations and model simulations)  
44 | with the median value being usually significantly lower than the mean. The daytime BLHs show  
45 | a very pronounced seasonal cycle at most continental stations (opposite in phase with the  
46 | seasonal cycle of the nocturnal BLH), with typical values around 500m during winter, and

1 ~1000-2000m during summer. The daytime BLH is in general relatively well simulated at most  
2 stations, as further illustrated by the ratios between modelled and observed BLHs, which are  
3 close to 1 (see **Figure 8-S13 in the Supplement**). An exception, however, are coastal sites (e.g.,  
4 Angus, Mace Head), where apparently the model representation errors (e.g., transition between  
5 land and sea) are a limiting factor. In general, it should be expected that the model BLH  
6 extracted at the location of the IGRA station should agree better than that extracted at the InGOS  
7 station (see Section 3.2.4.1 for the definition of the model BLHs). However, e.g. at Egham the  
8 opposite is the case, since the IGRA station (Herstmonceaux) is closer to the coast, and the  
9 corresponding model BLH has more 'marine' character (and the transition zone between sea and  
10 land is not resolved by the model). For most ~~stations far from the coast'non-coastal' sites,~~  
11 however, the difference between the BLH at the InGOS station and the IGRA station, ~~as well as~~  
12 ~~the difference between the TM5 default and 'TM5\_INGOS' BLH~~ is usually very small (**Figures**  
13 **34 and 45** and **Figures S2-S11 and S13** in the Supplement). ~~The ECMWF BLH is in some~~  
14 ~~cases slightly different compared to the TM5 or 'TM5\_INGOS' BLH, especially at coastal sites,~~  
15 ~~probably partly also due to model representation errors (different horizontal grids of the~~  
16 ~~ECMWF IFS model and TM5 (see Section 3.1), and different methods of BLH computation (see~~  
17 ~~Section 4.1)).~~ Compared to the data for the nocturnal BLH, the daytime BLHs show much  
18 smaller difference between median and mean value, indicating a less skewed frequency  
19 distribution (**Figures 3 and 4-S12 and S13 in the Supplement**).

20 In the supplement (**Figures S2 to S11**) we show the full time series for the 10 stations in 2009,  
21 illustrating that also the synoptic variability of the BLH is relatively well reproduced by the  
22 models (for both nocturnal and daytime BLH). Furthermore, we extend the analysis by using all  
23 IGRA stations over Europe (about 130 stations; see **Figure 1** and **Figures S124 and S135** in the  
24 Supplement). This extended analysis confirms the major findings discussed above, especially (1)  
25 the ~~relatively~~-good agreement between simulated and observed BLH during daytime, (2) the  
26 tendency for the simulated nocturnal BLHs to be too high during summer, and (3) larger  
27 differences between TM5 and IGRA BLHs for stations located ~~close to the coasts in coastal zones~~.

28  
29 In the following we include the ceilometer and lidar derived BLH at Cabauw and Trainou,  
30 respectively, in the analysis. As clearly visible from the correlation plot between ceilometer and  
31 IGRA data for Cabauw (**Figure S12**), the ceilometer BLHs during midday are usually lower than  
32 the IGRA data (especially for the period March to September), while modelled BLHs fall in  
33 between the two observational datasets (**Figure 56**). Part of this difference is likely due to the  
34 different methodologies. Hennemuth and Lammert (2006) pointed out that inconsistencies  
35 between the atmospheric thermal profile and the aerosol concentration profile can result in  
36 differences between radiosonde and lidar/ceilometer BLH retrievals. In addition, ~~also~~ the spatial  
37 separation between Cabauw and De Bilt (~23 km) combined with different surface  
38 characteristics (wetter soils in Cabauw and different large scale surface roughness) may play  
39 some role. While the correlation between IGRA BLHs and the ceilometer BLH retrievals at  
40 Cabauw is reasonable ( $r=0.63$ ) during daytime (**Figure 2**), it is very poor during night (**Figure**  
41 **S1**), probably due to the issues of ceilometers to detect the shallow nocturnal BLH, as mentioned  
42 in Section 2.1.2. The lidar daytime data at Trainou for 2011 agree relatively well with the model  
43 BLHs (except May) (**Figure 57**). While no IGRA data are available for this period, the  
44 comparison between model simulations and IGRA for 2006-2010 at Trainou (**Figure 5**) shows  
45 similar (or slightly better) agreement as the comparison between lidar and model for 2011.

46

## 45.2. Simulated $^{222}\text{Rn}$ activity concentrations versus observations

Figures 68 and 79 show the mean seasonal variations of observed and simulated  $^{222}\text{Rn}$  activity concentrations at each of the studied InGOS sites at 05 UTC (time around which typically the daily maximum  $^{222}\text{Rn}$  activity concentration occurs) and at 14 UTC ( $^{222}\text{Rn}$  daily minimum), respectively. For most stations, TM5 simulated  $^{222}\text{Rn}$  activity concentrations based on the InGOS  $^{222}\text{Rn}$  flux map show significantly better agreement with observations than the simulations based on the constant  $^{222}\text{Rn}$  flux, especially regarding the average seasonal variations. The improvement is largest during winter months, when TM5 simulations based on the constant  $^{222}\text{Rn}$  fluxes often overestimate observations, while simulated concentrations based on the InGOS  $^{222}\text{Rn}$  flux map are significantly lower owing to the lower  $^{222}\text{Rn}$  fluxes (Figure 6 and 73b). This, in turn, is driven mostly by the higher soil moisture and consequently lower permeability of the soil in winter. Furthermore, large differences are visible at many North European sites close to the coast (Angus, Lutjewad, Mace Head, Cabauw), where the water table can be very shallow, significantly reducing the  $^{222}\text{Rn}$  fluxes (Karstens et al., 2015). Apparently, Model simulations based on the InGOS  $^{222}\text{Rn}$  flux map (which include modelled water table in the parameterization of  $^{222}\text{Rn}$  fluxes) agree much better with observations than the control runs with constant  $^{222}\text{Rn}$  fluxes. Despite the larger  $^{222}\text{Rn}$  fluxes during summer, daily minimum  $^{222}\text{Rn}$  concentrations in model and observations are usually lower at continental stations (e.g. Heidelberg, Gif-sur-Yvette) due to the much higher daytime boundary layer in summer compared to winter.

Figures S148 to S248 in the supplement show the full time series of simulated and observed  $^{222}\text{Rn}$  concentrations at the 10 studied InGOS stations (with  $^{222}\text{Rn}$  activity concentration observations available) for 2009.

### Relationship between $^{222}\text{Rn}$ activity concentrations and boundary layer heights

In the following, we analyze the relationship between  $^{222}\text{Rn}$  activity concentration and BLH in more detail. Figure 910 shows the mean seasonal diurnal cycle of observed and simulated  $^{222}\text{Rn}$  activity concentration and BLH for the four seasons at different sites. The figure illustrates the very strong anti-correlation between simulated BLH and  $^{222}\text{Rn}$  activity concentration: The modelled BLHs increase sharply between 9:00 and 10:00 UTC (10:00/11:00 and 11:00/12:00 LT), resulting in an immediate decrease of modelled  $^{222}\text{Rn}$  concentrations. In contrast, the  $^{222}\text{Rn}$  activity concentration measurements show a slower decrease over several hours. Although this slow decrease may be partially due to the slow (45min) response time of the two-filter detectors, it is clear that the sharp changes in simulated BLHs and  $^{222}\text{Rn}$  activity concentrations are due to the relatively coarse temporal resolution of ECMWF meteorological data (3-hourly for surface data (e.g., BLHs) and 6-hourly for 3D fields (temperature, wind and humidity); see Section 3.1). Apparently the sharp changes in the 'model world' are due to the relatively coarse temporal resolution of ECMWF meteorological data (3-hourly for surface data (e.g., BLHs) and 6-hourly for 3D fields (temperature, wind and humidity); see Section 3.1). Because the ceilometer data at Cabauw during night might be questionable, we included in Figure 910 only the lidar measurements at Trainou (TR4). These that shows a much slower growth of the BLH, starting in the morning and reaching its maximum in the late afternoon, as also illustrated in Pal et al. (2012, 2015). Despite in spite of the obvious issue of the temporal resolution of the model, however, inspection of Figure 910 also indicates significant illustrates that the mismatches

1 between simulated and observed  $^{222}\text{Rn}$  activity concentrations that cannot be explained wholly  
2 by problems with the modeled BLH (even accounting for possible instrumental response time  
3 effects). Especially during daytime, the TM5 BLHs are close to the IGRA measurements at most  
4 stations (as also illustrated by the ratios of BLHs in Figure 8), whereas ~~while~~ larger differences  
5 are observed between the simulated and measured  $^{222}\text{Rn}$  activity concentrations ~~simulations and~~  
6 ~~measurements~~ at several stations. This is further illustrated in **Figure 10**, where we compare  
7 the ratio of simulated ~~to~~ ~~and~~ observed BLH with the ratio of observed to simulated ~~and observed~~  
8  $^{222}\text{Rn}$  activity concentration during daytime; ~~and in Figure 12 where these ratios are shown~~ for  
9 the different seasons. If the  $^{222}\text{Rn}$  activity concentration errors were purely due incorrect  
10 dilutions resulting from errors in the modeled BLH at a given station, the two ratios would be  
11 similar. This is clearly not the case, however, and the modelled afternoon concentration ratios  
12 range widely (from 0.2 to 1.8) from station to station. These mismatches between observed  
13 and simulated  $^{222}\text{Rn}$  activity concentrations may be related to finding points to potential  
14 shortcomings of TM5 ~~into~~ correctly simulatinge the vertical  $^{222}\text{Rn}$  activity concentration  
15 gradients within the boundary layer (see below). Furthermore, it is important to consider the  
16 uncertainties of the InGOS  $^{222}\text{Rn}$  flux map. Karstens et al. (2015) estimated that the most  
17 important uncertainty in the InGOS  $^{222}\text{Rn}$  flux is due to the uncertainties in the soil moisture  
18 data. Altogether, the uncertainties in modelled  $^{222}\text{Rn}$  fluxes for individual pixels ( $0.083^\circ \times$   
19  $0.083^\circ$ ) are estimated to be about 50%. Karstens et al. (2015) pointed out that the uncertainty of  
20 the  $^{222}\text{Rn}$  fluxes averaged over the footprint of the measurements might be smaller. However, the  
21 uncertainties of neighboring pixels in the InGOS  $^{222}\text{Rn}$  flux map are likely to be strongly  
22 correlated, and therefore the reduction of the relative uncertainty (integrated over a typical  
23 footprint on the order of 50-200km) is probably relatively small. Assuming an overall  
24 uncertainty of ~50% of the regional  $^{222}\text{Rn}$  fluxes, the model simulations could be considered  
25 broadly consistent with observations at most sites.

### 27 Sensitivity of simulated $^{222}\text{Rn}$ activity concentrations to convection scheme

28  
29 The use of the new ECMWF based convection combined with the “revised updated treatment of  
30 slopes scheme” (i.e., FI\_CU acronym in Section 3.4.4.2) results in a small decrease of simulated  
31  $^{222}\text{Rn}$  concentrations at most stations, typically on the order of ~10-30% (**Figures 6-9 see**  
32 **Figures S31 to S41 in the Supplement**). However, root mean square (RMS) and correlation  
33 coefficients are very similar at most sites for both convection parameterizations (**Figure 8**).  
34 Hence, no clear conclusions can be drawn, which parameterization is more realistic. At the same  
35 time, **Figure 8** demonstrates again the improvement using the InGOS  $^{222}\text{Rn}$  flux map,  
36 resulting in (1) ratios between simulated and observed  $^{222}\text{Rn}$  activity concentration closer to one,  
37 (2) lower RMS, and (3) higher correlation coefficients at several stations, compared to the model  
38 simulations using constant  $^{222}\text{Rn}$  fluxes. This highlights the challenge to validate model  
39 simulations. The difference of ~10-30% of  $^{222}\text{Rn}$  activity concentrations using a different  
40 convection parameterization is expected to result in a difference of similar order of magnitude  
41 for the GHG emissions derived in inverse modelling. First GHG inversions with the new  
42 ECMWF based convection confirm that derived emissions change significantly (not shown).

### 44 Comparison of simulated and observed $^{222}\text{Rn}$ activity concentrations: Impact of sampling 45 time

1 | **Figure 1012** illustrates further that the ratio between observed and simulated daytime  $^{222}\text{Rn}$   
2 activity concentration also depends on the exact hour, decreasing significantly between 12:00  
3 and 15:00 UTC at several stations (very pronounced at Trainou and Ispra). This is clearly due to  
4 the shortcomings of TM5 to simulate the diurnal cycle in the BLH discussed above (owing to the  
5 coarse temporal resolution of the meteorological data). In the current TM5-4DVAR system the  
6 average (observed and simulated) concentrations between 12:00 and 15:00 LT are used to derive  
7 emissions (Bergamaschi et al., 2010; 2015). Given the too fast increase of the BLH and  
8 consequently too fast decrease of simulated mixing ratios in the morning transition period, the  
9 choice of the assimilation time window may introduce some systematic errors in the flux  
10 inversions.

11  
12 | In the analyses shown in **Figure 102**, the data include all stability regimes meteorological  
13 conditions. In addition, we performed this analysis separately for unstable, neutral, and stable  
14 vertical mixing conditions. We used, based on the bulk Richardson number calculated at the first  
15 level of the near the surface in TM5 model. This extended analysis, however, showed relatively  
16 similar model performance for these different weather conditions (results not shown). -A  
17 limitation of this exercise is that for both stable and neutral stability regimes, we had at most  
18 stations only few cases by seasons.

### 20 | Vertical gradients of $^{222}\text{Rn}$ activity concentrations in the boundary layer at Cabauw

21  
22 | Finally, we explore the vertical gradients of TM5 simulated  $^{222}\text{Rn}$  activity concentrations at  
23 Cabauw where measurements are available at two vertical levels (20-m [CB1] and 200-m [CB4]  
24 height [CB4]; Table 1). The measurement tower height of 20-m is within the first model layer,  
25 while 200-m is within layer 3. **Figure 113** shows the monthly mean mean-diurnal variation  
26 of modeled and observed vertical gradients of  $^{222}\text{Rn}$  activity concentrations for each month for  
27 2009. Although the InGOS  $^{222}\text{Rn}$  flux based model simulations agree better with observations (in  
28 terms of  $^{222}\text{Rn}$  activity concentrations; see **Figures 6, 7, and 88 and 9**) compared to the model  
29 simulations based on constant fluxes, this is not the case for the  $^{222}\text{Rn}$  gradients for some months  
30 (€ between June and November the modelled gradients based on the constant fluxes agree better  
31 with observations, which could point to partially compensating systematic errors (e.g. too high  
32  $^{222}\text{Rn}$  fluxes might be compensated by too fast vertical mixing)). During large parts of the year,  
33 the InGOS  $^{222}\text{Rn}$  flux based model simulations underestimate the observed gradients. This is  
34 further illustrated in the scatter plots shown in **Figure 124** (separately for 00 and 12 UTC). For  
35 inverse modelling, especially the underestimated vertical gradient during daytime is critical and  
36 could lead to biases in the GHG inversions. Furthermore, **Figures 113** shows that during the  
37 transition phase in the morning the modelled  $^{222}\text{Rn}$  activity concentration vertical gradient  
38 decreases faster than the observed gradient, which is again probably largely due to the coarse  
39 time resolution of the meteorological data in TM5 together with the slow response time of the  
40 two-filter radon measurements, although it may also indicate that, but could point in addition  
41 also to too fast vertical mixing is proceeding too rapidly in the model.

## 44 | **56. Conclusions**

1 In the first part of this study, we evaluated the boundary layer dynamics of the TM5 model by  
2 comparison with BLHs from the NOAA IGRA radiosonde data as well as with BLH retrievals  
3 from a ceilometer at Cabauw and lidar at Trainou.

4 TM5 reproduces reasonably well the IGRA BLHs during daytime within 10-20% (which is  
5 within the uncertainty of the IGRA data) for continental stations at low altitudes. During night,  
6 the model overestimates the shallow nocturnal BLHs, especially for very low BLHs (<100 m)  
7 observed during summer time. At coastal sites, the differences between simulated BLH and  
8 IGRA observations data (both day and nighttime) are usually larger due to model representation  
9 errors (since the transition zone between the marine boundary layer over sea and the continental  
10 boundary layer over land is not resolved by the model).

11 The BLH retrievals at Cabauw show a reasonable moderate correlation with IGRA data from De  
12 Bilt at 12 UTC, but are systematically lower. During night (00 UTC), however, the two data set  
13 show only a very poor correlation. Besides the fundamental differences in the BLH retrieval  
14 methods, however, also the spatial separation between Cabauw and DeBilt (~23 km) probably  
15 contributes to the differences in the derived BLH. For the lidar BLH data from Trainou, no direct  
16 comparison with the IGRA data is available (due to different time periods), but the comparison  
17 with the modelled BLH show similar agreement with the two different observational datasets  
18 [IGRA: for 2006-2010; lidar: 2011]. For the better exploitation of ceilometer / lidar data in the  
19 future, the further development of BLH retrievals is essential to ensure consistency between the  
20 different methods.

21  
22 In the second part of this study, we compared TM5 simulations of  $^{222}\text{Rn}$  activity concentrations  
23 with quasi-continuous  $^{222}\text{Rn}$  measurements from 10 European monitoring stations.

24 The  $^{222}\text{Rn}$  activity concentration simulations based on the new  $^{222}\text{Rn}$  flux map show significant  
25 improvements compared to  $^{222}\text{Rn}$  simulations using constant  $^{222}\text{Rn}$  fluxes, especially regarding  
26 the average seasonal variability and generally lower simulated  $^{222}\text{Rn}$  activity concentrations at  
27 North European sites close to the coast. These improvements highlight the benefit of the process-  
28 based approach, including a parameterization of water table (Karstens et al., 2015). Nevertheless,  
29 the (relative) differences between simulated and observed daytime minimum  $^{222}\text{Rn}$   
30 concentrations are larger for several stations (on the order of 50%) than compared to the (relative)  
31 differences between simulated and observed BLH at noon. This is probably partly related to the  
32 uncertainties in the  $^{222}\text{Rn}$  flux map (estimated to be on the order of 50%). In addition, however,  
33 also potential shortcomings of TM5 to correctly simulate the vertical  $^{222}\text{Rn}$  activity concentration  
34 gradients are likely to play a significant role, which may be caused by the vertical diffusion  
35 coefficients and/or the limited vertical resolution in the model.

36 The comparison of simulated  $^{222}\text{Rn}$  activity concentrations with measurements at Cabauw (20 m  
37 versus 200 m) shows that the model underestimates the measured vertical gradient (i.e.,  
38 differences of concentrations between 20m and 200m levels) at this station. Furthermore, the  
39 sharp increase of the modeled BLH in the morning transition period results in a rapid decrease of  
40 the simulated  $^{222}\text{Rn}$  activity concentrations, while  $^{222}\text{Rn}$  measurements show a slower decrease at  
41 many stations. Although this latter timing effect may be partially due to the slow (45min)  
42 response time of the two-filter radon detectors, it is clear that the current coarse temporal  
43 resolution of the TM5 meteorological data (3-hourly for surface data and 6-hourly for 3D fields)  
44 limits the capability of simulating the diurnal cycle realistically. ~~The sharp increase of the~~  
45 ~~modeled BLH in the morning transition period results in a rapid decrease of the simulated  $^{222}\text{Rn}$~~   
46 ~~activity concentrations, while  $^{222}\text{Rn}$  measurements show a slower decrease at many stations.~~

1 | ~~These~~ issues probably leads to systematic biases in inversions of GHG emissions. An updated  
2 | TM5-4DVAR system is currently under development with increased temporal resolution of the  
3 | meteorological data (3-hourly ECMWF data, interpolated to observational data time).

4 |  
5 | Finally, we evaluated the "revised updated slopes scheme" ~~treatment~~ and the new ECMWF  
6 | based convection scheme in the TM5 model. The results show a relatively small impact of the  
7 | new slopes treatment, but a significant impact of the new ECMWF convection scheme, leading  
8 | to significantly lower  $^{222}\text{Rn}$  activity concentrations (about 20%) during daytime, especially in  
9 | winter. While this is expected to have a significant impact on derived emissions in GHG  
10 | inversions, the comparison with the available European  $^{222}\text{Rn}$  activity concentration observations  
11 | showed very similar performance. Hence, no clear conclusion about which parameterization is  
12 | more realistic can be drawn from this study—. These findings highlight the challenges of  
13 | validating atmospheric transport models with the accuracy required to better evaluate and  
14 | improve the quality of GHG flux inversions. In order to improve the validation capabilities it  
15 | would be important (1) to increase the number of  $^{222}\text{Rn}$  monitoring stations, (2) to perform  
16 | vertical  $^{222}\text{Rn}$  activity concentration profile measurements at tall towers and also from aircraft  
17 | (e.g. Chambers et al., 2011; Williams et al., 2011, 2013), (3) to extend the validation of the  $^{222}\text{Rn}$   
18 | inventories by local/regional  $^{222}\text{Rn}$  flux measurements, (4) to further develop the BLH retrievals  
19 | from ceilometer / LIDAR instruments and (5) to further extend the ceilometer / LIDAR network.  
20 | More work is also needed to improve the representation of the nocturnal boundary layer in global  
21 | and regional models. The use of  $^{222}\text{Rn}$  in the diagnosis of the nocturnal mixing effects is one area  
22 | showing promise in this regard (Williams et al., 2013).

### 23 | 24 | 25 | 26 | 27 | **Code availability**

28 | Further information about the TM5 code can be found at <http://tm5.sourceforge.net/>. Readers  
29 | interested in the TM5 code can contact Maarten Krol ([maarten.krol@wur.nl](mailto:maarten.krol@wur.nl)), Arjo Segers  
30 | ([arjo.segers@tno.nl](mailto:arjo.segers@tno.nl)) or Peter Bergamaschi ([peter.bergamaschi@jrc.ec.europa.eu](mailto:peter.bergamaschi@jrc.ec.europa.eu))

### 31 | 32 | 33 | **Acknowledgment:**

34 | This work has been supported by the European Commission Seventh Framework Programme  
35 | (FP7/2007–2013) project InGOS under grant agreement 284274. We thank Juha Hatakka for  
36 | providing  $^{222}\text{Rn}$  data from Pallas. Furthermore, we are grateful to Clemens Schlosser from the  
37 | German Federal Office for Radiation Protection for the  $^{222}\text{Rn}$  data from Schauinsland, which  
38 | were used for additional analyses. ECMWF meteorological data has been preprocessed by  
39 | Philippe Le Sager into the TM5 input format. We are grateful to ECMWF for providing  
40 | computing resources under the special project 'Global and Regional Inverse Modeling of  
41 | Atmospheric  $\text{CH}_4$  and  $\text{N}_2\text{O}$  (2012-2014)' and 'Improve estimates of global and regional  $\text{CH}_4$  and  
42 |  $\text{N}_2\text{O}$  emissions based on inverse modelling using in-situ and satellite measurements (2015-  
43 | 2017)'.  
44 |  
45 |  
46 |

1 **References**

2  
3 Angevine, W., Grimsdell, A., Hartten, L. M., and Delany, A. C.: The Flatland Boundary Layer  
4 Experiments, *Bulletin of the American Meteorological Society* 79, 419–431, 1998.

5  
6 Bergamaschi, P., Krol, M., Meirink, J. F., Dentener, F., A. Segers, A., van Aardenne, J., Monni,  
7 S., Vermeulen, A., Schmidt, M. Ramonet, M., Yver, C., F. Meinhardt, F., Nisbet, E. G., R.  
8 Fisher, R., S. O'Doherty, S., and Dlugokencky, E.J.: Inverse modeling of European CH4  
9 emissions 2001-2006, *J. Geophys. Res.*, 115(D22309), doi:10.1029/2010JD014180, 2010.

10  
11 Bergamaschi, P., Corazza, M., Karstens, U., Athanassiadou, M., Thompson, R. L., Pison, I.,  
12 Manning, A. J., Bousquet, P., Segers, A., Vermeulen, A. T., Janssens-Maenhout, G., Schmidt,  
13 M., Ramonet, M., Meinhardt, F., Aalto, T., Haszpra, L., Moncrieff, J., Popa, M. E., Lowry, D.,  
14 Steinbacher, M., Jordan, A., O'Doherty, S., Piacentino, S., and Dlugokencky, E.: Top-down  
15 estimates of European CH4 and N2O emissions based on four different inverse models, *Atmos.*  
16 *Chem. Phys.*, 15, 715-736, doi:10.5194/acp-15-715-2015, 2015

17  
18 Biraud, S., Ciais, P., Ramonet, M., Simmonds, P., Kazan, V., Monfray, P., O'Doherty, S., Spain,  
19 T. G., and Jennings, S. G.: European greenhouse gas emissions estimated from continuous  
20 atmospheric measurements and radon 222 at Mace Head, Ireland, *J. Geophys. Res.*, 105(D1),  
21 1351–1366, 2000

22  
23 Broquet, G., Chevallier, F., Bréon, F.-M., Kadygrov, N., Alemanno, M., Apadula, F., Hammer,  
24 S., Haszpra, L., Meinhardt, F., Morguí, J. A., Necki, J., Piacentino, S., Ramonet, M., Schmidt,  
25 M., Thompson, R. L., Vermeulen, A. T., Yver, C., and Ciais, P.: Regional inversion of CO2  
26 ecosystem fluxes from atmospheric measurements: reliability of the uncertainty estimates,  
27 *Atmos. Chem. Phys.*, 13, 9039-9056, doi:10.5194/acp-13-9039-2013, 2013.

28  
29 Carouge, C., Bousquet, P., Peylin, P., Rayner, P. J., and Ciais, P.: What can we learn from  
30 European continuous atmospheric CO2 measurements to quantify regional fluxes – Part 1:  
31 Potential of the network, *Atmos. Chem. Phys. Discuss.*, 8, 18591-18620, doi:10.5194/acpd-8-  
32 18591-2008, 2008

33  
34 [Chambers, S., Williams, A.G., Zahorowski, W., Griffiths, A. and Crawford, J.: Separating](#)  
35 [remote fetch and local mixing influences on vertical radon measurements in the lower](#)  
36 [atmosphere. \*Tellus\* 63B, 843–859. doi:10.1111/j.1600-0889.2011.00565.x, 2011](#)

37  
38  
39 Chevillard, A., Ciais, P., Karstens, U., Heimann, M., Schmidt, M., Levin, I., Jacob, D., Podzun,  
40 R., Kazan, V., Sartorius, H., and Weingartner, E.: Transport of 222 Rn using the regional model  
41 REMO: a detailed comparison with measurements over Europe, *Tellus B*, 54, 850–871, 2002

42  
43 Collaud Coen, M., Praz, C., Haeefe, A., Ruffieux, D., P. Kaufmann, P. and B. Calpini B.:  
44 Determination and climatology of the planetary boundary layer height above the Swiss plateau  
45 by in situ and remote sensing measurements as well as by the COSMO-2 model, *Atmos. Chem.*  
46 *Phys.*, 14, 13205-13221, 2014

1  
2 Corazza, M., Bergamaschi, P., Vermeulen, A. T., Aalto, T., Haszpra, L., Meinhardt, F.,  
3 O'Doherty, S., Thompson, R., Moncrieff, J., Popa, E., Steinbacher, M., Jordan, A., Dlugokencky,  
4 E., Brühl, C., Krol, M., and Dentener, F.: Inverse modelling of European N<sub>2</sub>O emissions:  
5 assimilating observations from different networks, *Atmos. Chem. Phys.*, 11, 2381-2398,  
6 doi:10.5194/acp-11-2381-2011, 2011.  
7  
8 Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U.,  
9 Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L.,  
10 Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy,  
11 S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally,  
12 A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C.,  
13 Thépaut, J.-N. and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the  
14 data assimilation system. *Q.J.R. Meteorol. Soc.*, 137: 553–597. doi: 10.1002/qj.828, 2011  
15 Denning, A. S., Holzer, M., Gurney, K. R., Heimann, M., Law, R.M., Rayner, P. J., Fung, I. Y.,  
16 Fan, S.-M., Taguchi, S., Friedlingstein, P., Balkanski, Y., Taylor, J., Maiss, M., and Levin, I.:  
17 Three-dimensional transport and concentration of SF<sub>6</sub>: A model intercomparison study  
18 (TransCom 2). *Tellus* 51B: 266-297, 1999  
19  
20 Dentener, F., Feichter, J., and Jeuken, A: Simulation of the transport of Rn222 using on-line and  
21 off-line global models at different horizontal resolutions: a detailed comparison with  
22 measurements, *Tellus*, 51B, 573-602, 1999.  
23  
24 Eresmaa, N., Karppinen, A., Joffre, S. M., Räsänen, J. and Talvitie, H: Mixing height  
25 determination by ceilometers, *Atmos. Chem. Phys.*, 6, 1485–1493, 2006.  
26  
27 Ganesan, A. L., Manning, A. J., Grant, A., Young, D., Oram, D.E., Sturges, W. T., Moncrieff, J.  
28 B., and O'Doherty, S.: Quantifying methane and nitrous oxide emissions from the UK and  
29 Ireland using a national-scale monitoring network, *Atmos. Chem. Phys.*, 15, 6393-6406,  
30 doi:10.5194/acp-15-6393-2015, 2015.  
31  
32 Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bakwin, P.  
33 S., and Grainger, C. A.: Toward constraining regional-scale fluxes of CO<sub>2</sub> with atmospheric  
34 observations over a continent: 1. Observed spatial variability from airborne platforms, *J.*  
35 *Geophys. Res.*, 108(D24), 4756, doi:10.1029/2002JD003018, 2003  
36  
37 Gregory, D., Morcrette, J. -J., Jakob, C., Beljaars, A. M., and Stockdale, T.: Revision of  
38 convection, radiation and cloud schemes in the ECMWF model, *Q. J. R. Meteorol. Soc.*, 126,  
39 1685–1710, 2000  
40  
41 Griffiths, A. D., Parkes, S. D., Chambers, S. D., McCabe, M. F., and A. G. Williams, A. G.:  
42 Improved mixing height monitoring through a combination of lidar and radon measurements,  
43 *Atmos. Meas. Tech.*, 6, 207–218, 2013  
44  
45 Haeffelin, M., Angelini, F., Morille, Y., Martucci, G., Frey, S., Gobbi, G. P., Lolli, S., O'Dowd,  
46 C. D., Sauvage, L., Xueref. Remy, I., Wastine, B., and Feist, D. G.: Evaluation of mixing-height

1 retrievals from automatic profiling lidars and ceilometers in view of future integrated networks  
2 in Europe, *Bound.-Layer Meteorol.*, 143, 49–75, doi:10.1007/s10546-011-9643-z , 2012  
3

4 Haij, M. J. de., Wauben, W. M. F., and Baltink, H. K.: Determination of mixing layer height  
5 from ceilometer backscatter profiles, *Proceedings of SPIE*, Vol. 6362, 63620R (2006);  
6 doi:10.1117/12.691050, 11/9/2006-14/9/2006, James R. Slusser, Klaus Schäfer, Adolfo Comerón  
7 (Ed), 2006, Stockholm, Zweden, SPIE., 2006  
8  
9 |

10

11 Hatakka, J., Aalto, T., Aaltonen, V., Aurela, M., Hakola, H., Kompula, M., Laurila, T.,  
12 Lihavainen, H., Paatero, J., Salminen, K., and Viisanen, Y.: Overview of the atmospheric  
13 research activities and results at Pallas GAW station, *Boreal Environ. Res.*, 8, 365–383, 2003  
14

15 Hennemuth, B. and Lammert, A: Determination of the convective boundary layer height from  
16 radiosonde and lidar backscatter *Boundary-Layer Meteorology*, 120, 181-209, 2006  
17

18 Holtslag, A. A. M., and Moeng, C. H.: Eddy diffusivity and counter-gradient transport in the  
19 convective atmospheric boundary layer, *J. Atmos. Sci.*, 48, 1690-1698, 1991  
20

21 Jacob, D. J., and Prather, M. J.: Radon-222 as a test of convective transport in a general  
22 circulation model, *Tellus B*, 42, 118–134, 1990  
23

24 Jacob, D. J., Prather, M. J., Rasch, P. J., Shia, R.-L., Balkanski, Y. J., Beagley, S. R., Bergmann,  
25 D. J., Blackshear, W. T., Brown, M., Chiba, M., Chipperfield, M. P., de Grandpré, J., Dignon, J.  
26 E., Feichter, J., Genthon, C., Grose, W. L., Kasibhatla, P. S., Köhler, I., Kritz, M. A., Law, K.,  
27 Penner, J. E., Ramonet, M., Reeves, C. E., Rotman, D. A., Stockwell, D. Z., Van Velthoven, P.  
28 F. J., Verver, G., Wild, O., Yang, H., and Zimmermann, P.: Evaluation and intercomparison of  
29 global atmospheric transport models using <sup>222</sup>Rn and other short-lived tracers, *J. Geophys. Res.*,  
30 102, 5953–5970, doi: 10.1029/96JD02955, 1997  
31

32 Jakob, C., and Klein, S. A.: A parametrization of cloud and precipitation overlap effects for use  
33 in General Circulation Models, *Quart. J. Roy. Meteorol. Soc*, 126, 2525-2544, 2000  
34

35 Karstens, U., Schwingshackl, C., Schmithüsen, D., and Levin, I.: A process-based <sup>222</sup>radon flux  
36 map for Europe and its comparison to long-term observations, *Atmos. Chem. Phys.*, 15, 12845-  
37 12865, doi:10.5194/acp-15-12845-2015, 2015  
38

39 Krol, M., Houweling, S., Bregman, B., van den Broek, M., Segers, A., van Velthoven, P., Peters,  
40 W., Dentener, F., and Bergamaschi, P.: The two-way nested global chemistry-transport zoom  
41 model TM5: algorithm and applications, *Atmos. Chem. Phys.*, 5, 417-432, doi:10.5194/acp-5-  
42 417-2005, 2005  
43

44 Kort, E. A., Eluszkiewicz, J., Stephens, B. B., Miller, J. B., Gerbig, C., Nehr Korn, T., D. B. C.,  
45 Kaplan, J. O., Houweling, S., and Wofsy, S. C.: Emissions of CH<sub>4</sub> and N<sub>2</sub>O over the United

1 States and Canada based on a receptor-oriented modeling framework and COBRA-NA  
2 atmospheric observations, *Geophys. Res. Lett.*, 35, L18808, doi: 10.1029/2008GL034031, 2008  
3

4 Levin, I., Born, M., Cuntz, M., Langendörfer, U., Mantsch, S., Naegler, T., Schmidt, M.,  
5 Varlagin, A., Verclas, S., and D. Wagenbach, D.: Observations of atmospheric variability and  
6 soil exhalation rate of Radon-222 at a Russian forest site: Technical approach and deployment  
7 for boundary layer studies, *Tellus*, 54B, 462-475, 2002  
8

9 Lopez, M., Schmidt, M., Yver, C., Messenger, C., Worthy, D., Kazan, V., Ramonet, M.,  
10 Bousquet, P., and Ciais, P.: Seasonal variation of N<sub>2</sub>O emissions in France inferred from  
11 atmospheric N<sub>2</sub>O and Rn-222 measurements, *J Geophys. Res.-Atmos.*, 117, D14103,  
12 doi:10.1029/2012jd017703, 2012  
13

14 [Locatelli, R., Bousquet, P., Hourdin, F., Saunois, M., Cozic, A., Couvreux, F., Grandpeix, J.-Y.,](#)  
15 [Lefebvre, M.-P., Rio, C., Bergamaschi, P., Chambers, S. D., Karstens, U., Kazan, V., van der](#)  
16 [Laan, S., Meijer, H. A. J., Moncrieff, J., Ramonet, M., Scheeren, H. A., Schlosser, C., Schmidt,](#)  
17 [M., Vermeulen, A., and Williams, A. G.: Atmospheric transport and chemistry of trace gases in](#)  
18 [LMDz5B: evaluation and implications for inverse modelling, \*Geosci. Model Dev.\*, 8, 129-150,](#)  
19 [doi:10.5194/gmd-8-129-2015, 2015](#)  
20

21 Louis, J. F. : A parametric model of vertical eddy fluxes in the atmosphere, *Boundary Layer*  
22 *Meteorology*, 17, 187-202, 1979  
23

24 Manning, A. J., O'Doherty, S., Jones, A. R., Simmonds, P. G., and R. G. Derwent, R. G.:  
25 Estimating UK methane and nitrous oxide emissions from 1990 to 2007 using an inversion  
26 modeling approach, *J. Geophys. Res.*, 116(D02305), doi:10.1029/2010JD014763, 2011  
27

28 Meirink, J. F., Bergamaschi, P., and Krol, M. C.: Four-dimensional variational data assimilation  
29 for inverse modelling of atmospheric methane emissions: method and comparison with synthesis  
30 inversion, *Atmos. Chem. Phys.*, 8, 6341-6353, doi:10.5194/acp-8-6341-2008, 2008  
31

32 Miller, S. M., Wofsy, S. C., Michalak, A. M., Kort, E. A., Andrews, A. E., Biraud, S.,  
33 Dlugokencky, E., Eluszkiewicz, J., Fischer, M. L., Janssens-Maenhout, G., Miller, B. R., Miller,  
34 J. B., Montzka, S., Nehrkorn, T., and Sweeney, C.: Anthropogenic emissions of methane in the  
35 United States, *Proceedings of the National Academy of Sciences*, doi:10.1073/pnas.1314392110,  
36 2013  
37

38 Morcrette, J.-J., Mlawer, E. J., Iacono, M. J., and Clough, S. A.: Impact of the radiation-transfer  
39 scheme RRTM in the ECMWF forecast system, Technical report in the ECMWF Newsletter, No.  
40 91, 2001  
41

42 Pal, S., Xueref-Remy, I., Ammoura, L., Chazette, P., Gibert, F., Royer, P., Dieudonné, E., J.-  
43 C. Dupont, J.-C., Haeffelin, M., Lac, C., Lopez, M., Morille, Y., Ravetta, F.: Spatio-temporal  
44 variability of the atmospheric boundary layer depth over the Paris agglomeration: An assessment  
45 of the impact of the urban heat island intensity, *Atmospheric environment*, Elsevier, 2012, 63,  
46 pp.261-275. <10.1016/j.atmosenv.2012.09.0, 2012

1  
2 Pal, S., Haeffelin, M., and E. Batchvarova, E.: Exploring a geophysical process-based attribution  
3 technique for the determination of the atmospheric boundary layer depth using aerosol lidar and  
4 near-surface meteorological measurements, *J. Geophys. Res. Atmos.* 118, 9277–9295,  
5 doi:10.1002/jgrd.50710, 2013  
6  
7 Pal, S.: Monitoring Depth of Shallow Atmospheric Boundary Layer to Complement LiDAR  
8 Measurements Affected by Partial Overlap, *Remote Sensing*, 6(9), 8468-8493, 2014  
9  
10 Pal, S., Lopez, M., Schmidt, M., Ramonet, M., Gibert, F., Xueref-Remy, I., and Ciais, P.:  
11 Investigation of the atmospheric boundary layer depth variability and its impact on the <sup>222</sup>Rn  
12 concentration at a rural site in France, *J. Geophys. Res. Atmos.*, 120, 623–643, doi:  
13 10.1002/2014JD022322, 2015  
14  
15 Patra, P. K., Houweling, S., Krol, M., Bousquet, P., Belikov, D., Bergmann, D., Bian, H.,  
16 Cameron-Smith, P., Chipperfield, M. P., Corbin, K., Fortems-Cheiney, A., Fraser, A., Gloor, E.,  
17 Hess, P., Ito, A., Kawa, S. R., Law, R. M., Loh, Z., Maksyutov, S., Meng, L., Palmer, P. I.,  
18 Prinn, R. G., Rigby, M., Saito, R., and Wilson, C.: TransCom model simulations of CH<sub>4</sub> and  
19 related species: linking transport, surface flux and chemical loss with CH<sub>4</sub> variability in the  
20 troposphere and lower stratosphere, *Atmos. Chem. Phys.*, 11, 12813-12837, doi:10.5194/acp-11-  
21 12813-2011, 2011  
22  
23  
24 [Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K.,](#)  
25 [Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll,](#)  
26 [D.: The Global Land Data Assimilation System, \*B. Am. Meteorol. Soc.\*, 85, 381–394,](#)  
27 [doi:10.1175/BAMS-85-3-381, 2004](#)  
28  
29 Russell, G. L., and Lerner, J. A.: A new finite-differencing scheme for the tracer transport  
30 equation, *Journal of Applied Meteorology*, 20, 1483-1498, 1981  
31  
32 Scheeren, H. A. and Bergamaschi, P.: First Three Years of CO<sub>2</sub>, CH<sub>4</sub>, N<sub>2</sub>O, and SF<sub>6</sub>  
33 Observations, and <sup>222</sup>Radon-Based Emission Estimates from the JRC Monitoring Station at  
34 Ispra (Italy): What Have We Learned So Far?, proceeding of the 16th WMO/IAEA Meeting on  
35 Carbon Dioxide, Other Greenhouse Gases and Related Measurement Techniques (GGMT-2011),  
36 Wellington, New Zealand, World Meteorological Organization, 2012  
37  
38 Schmithüsen, D., Chambers, S., Fischer, B., Gilge, S., Hatakka, J., Kazan, V., Neubert, R.,  
39 Paatero, J., Ramonet, M., Schlosser, C., Schmid, S., Vermeulen, A., and Levin, I: A European-  
40 wide <sup>222</sup>Radon and <sup>222</sup>Radon progeny comparison study, to be submitted to *Tellus B*, 2016  
41  
42  
43 Seibert, P., Beyrich, F., Gryning, S. E., Joffre, S., Rasmussen, A., and P. Tercier, P.: Review and  
44 intercomparison of operational methods for the determination of the mixing height, *Atmos.*  
45 *Environ*, 34, 1001–1027, 2000  
46

1 Seidel, D. J., Zhang, Y., Beljaars, A., Golaz, J. -C., Jacobson, A. R., and Medeiros, B.:  
2 Climatology of the planetary boundary layer over the continental United States and Europe, J.  
3 Geophys. Res.- Atmos, 117, 2012  
4

5 Smallman, T. L., Williams, M., and Moncrieff, J. B.: Can seasonal and interannual variation in  
6 landscape CO<sub>2</sub> fluxes be detected by atmospheric observations of CO<sub>2</sub> concentrations made at a  
7 tall tower?, Biogeosciences, 11, 735–747, doi:10.5194/bg-11-735-2014, 2014  
8

9 Stull, R. B.: An Introduction to Boundary Layer Meteorology, Kluwer Academic Publishers,  
10 666p, 1988  
11

12 Taguchi, S., Law, R. M., Rödenbeck, C., Patra, P. K., Maksyutov, S., Zahorowski, W., Sartorius,  
13 H., and Levin, I.: TransCom continuous experiment: comparison of <sup>222</sup>Rn transport at hourly  
14 time scales at three stations in Germany, Atmos. Chem. Phys., 11, 10071-10084,  
15 doi:10.5194/acp-11-10071-2011, 2011  
16

17 Tiedtke, M.: A comprehensive mass flux scheme for cumulus parameterisation in large scale  
18 models, Mon. Wea. Rev, 177, 1779–1800, 1989  
19

20 van der Laan, S., Karstens, U., Neubert, R. E. M., van der Laan-Luijkx, I. T., and Meijer, H. A.  
21 J.: Observation-based estimates of fossil fuel-derived CO<sub>2</sub> emissions in the Netherlands using  
22 Delta 14C, CO and <sup>222</sup>Radon, Tellus B, 62, 389–402, doi:10.1111/j.1600-0889.2010.00493.x,  
23 2010  
24

25 van der Veen, E., Optimizing transport properties in TM5 using SF<sub>6</sub>, Wageningen University,  
26 University of Twente, Master's thesis, 2013  
27

28 Vermeulen, A. T., Hensen, A., Popa, M. E., van den Bulk, W. C.M., and Jongejan, P. A. C.:  
29 Greenhouse gas observations from Cabauw Tall Tower (1992–2010), Atmos. Meas. Tech., 4,  
30 617–644, doi:10.5194/amt-4-617-2011, 2011  
31

32 Vogelesang, D. H. P., and Holtslag, A. A., M.: Evaluation and model impacts of alternative  
33 boundary-layer height formulation, Boundary Layer Meteorol., 81, 245–269, 1986  
34

35 Whittlestone, S., and W. Zahorowski, W.: Baseline radon detectors for shipboard use:  
36 Development and deployment in the First Aerosol Characterization Experiment (ACE 1), Journal  
37 of Geophysical Research, 103(D13), 16, 743-751, 1998  
38

39 [Williams, A. G., Zahorowski, W., Chambers, S. D., and Griffiths, A.: The vertical distribution of](#)  
40 [radon in clear and cloudy daytime terrestrial boundary layers. J Atmos Sci 68, 155-174, 2011](#)  
41

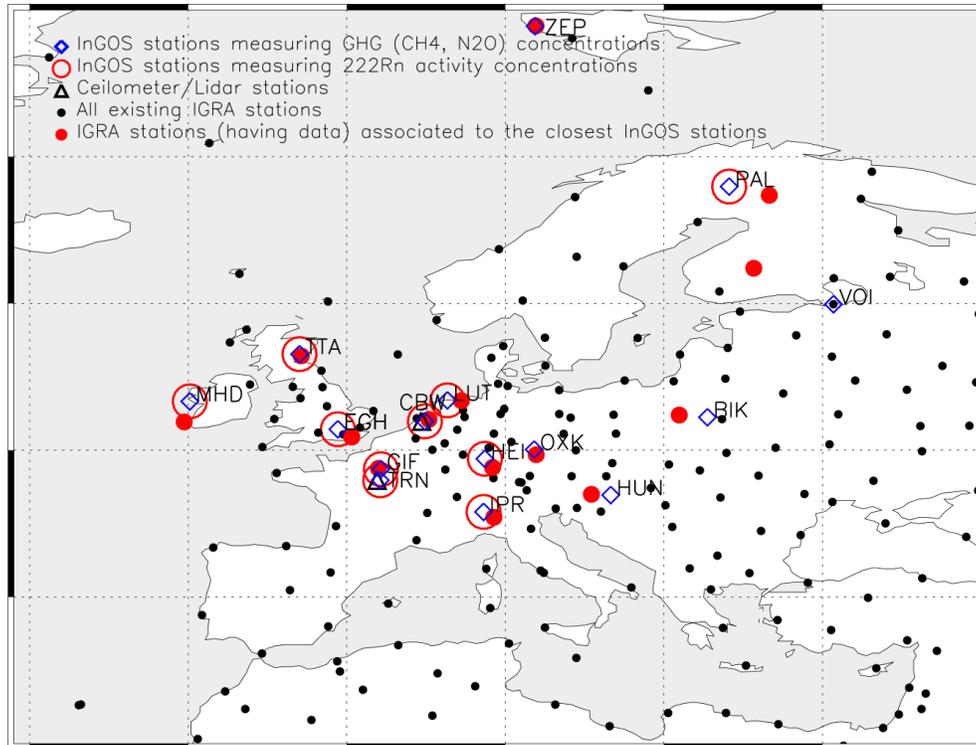
42 [Williams, A.G., Chambers, S.D. and Griffiths, A.D.: Bulk mixing and decoupling of the](#)  
43 [nocturnal stable boundary layer characterized using a ubiquitous natural tracer. Boundary-Layer](#)  
44 [Meteorol. 20 \(149\), 381-402. <http://dx.doi.org/10.1007/s10546-013-9849-3>, 2013](#)  
45  
46

- 1 Xia, Y., Sartorius, H., Schlosser, C., Stohlker, U., Conen, F., and Zahorowski, W.: Comparison  
2 of one- and two-filter detectors for atmospheric  $^{222}\text{Rn}$  measurements under various  
3 meteorological conditions, *Atmos. Meas. Tech.*, 3, 723–731, doi:10.5194/amt-3-723-2010, 2010  
4
- 5 Yver, C., Schmidt, M., Bousquet, P., Zahorowski, W., and Ramonet, M.: Estimation of the  
6 molecular hydrogen soil up-take and traffic emissions at a suburban site near Paris through  
7 hydrogen, carbon monoxide, and radon-222 semi-continuous measurements, *J. Geophys. Res.*,  
8 114, D18304, doi:10.1029/2009JD012122, 2009  
9
- 10 Zahorowski, W., Chambers, S. D., and Henderson-Sellers, A.: Ground based radon-222  
11 observations and their application to atmospheric studies, *J. Environ. Radioactivity*, 76, 3–33,  
12 2004

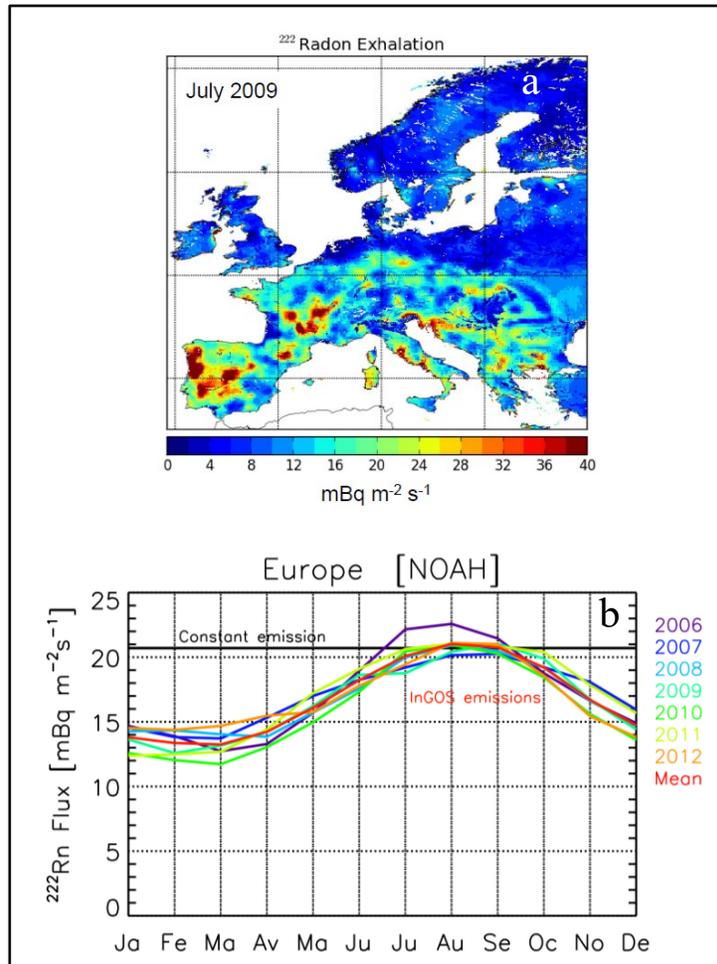
**Table 1:** Description of the different surface stations measuring  $^{222}\text{Rn}$  activity concentrations. See Figure 1 for the locations of the stations shown by their ID. ANSTO stands for Australian Nuclear Science and Technology Organisation. The locations of the stations are shown in Figure 1. CB1 and CB4 are the level 1 and level 2 of the Cabauw tower, respectively. Altitude stands for terrain elevation referred to the mean sea level and Height is the above ground level

Station ID	Name	Country	Latitude (degree)	Longitude (degree)	Altitude/Height (m)	$^{222}\text{Rn}$ instrument	Reference
PAL	Pallas	Finland	67.97	24.12	572/7	one-filter method	Hatakka et al. (2003)
TTA	Angus	United Kingdom	56.55	-2.98	363/50	ANSTO	Smallman et al. (2014)
LUT	Lutjewad	Netherlands	53.40	6.35	61/60	ANSTO	van der Laan et al. (2010)
MHD	Mace Head	Ireland	53.33	-9.90	40/15	one-filter method	Biraud et al. (2000)
CBW (CB1)	Cabauw	Netherlands	51.97	4.93	19/20	one-filter method	Vermeulen et al. (2011)
CBW (CB4)	Cabauw	Netherlands	51.97	4.93	199/200	ANSTO	Vermeulen et al. (2011)
EGH	Egham	United Kingdom	51.43	-0.56	<del>4570</del> /10	one filter method	Levin et al. (2002)
GIF	Gif-sur-Yvette	France	48.71	2.15	167/7	one-filter method	Lopez et al. (2012), Yver et al. (2009)
HEI	Heidelberg	Germany	49.42	8.71	146/30	one-filter method	Levin et al. (2002)
TRN (TR4)	Trainou	France	47.95	2.11	311/180	ANSTO	Schmidt et al. (2014)
IPR	Ispra	Italy	45.80	8.63	223/3.5 (15) <sup>1</sup>	ANSTO	Scheeren and Bergamaschi (2012)

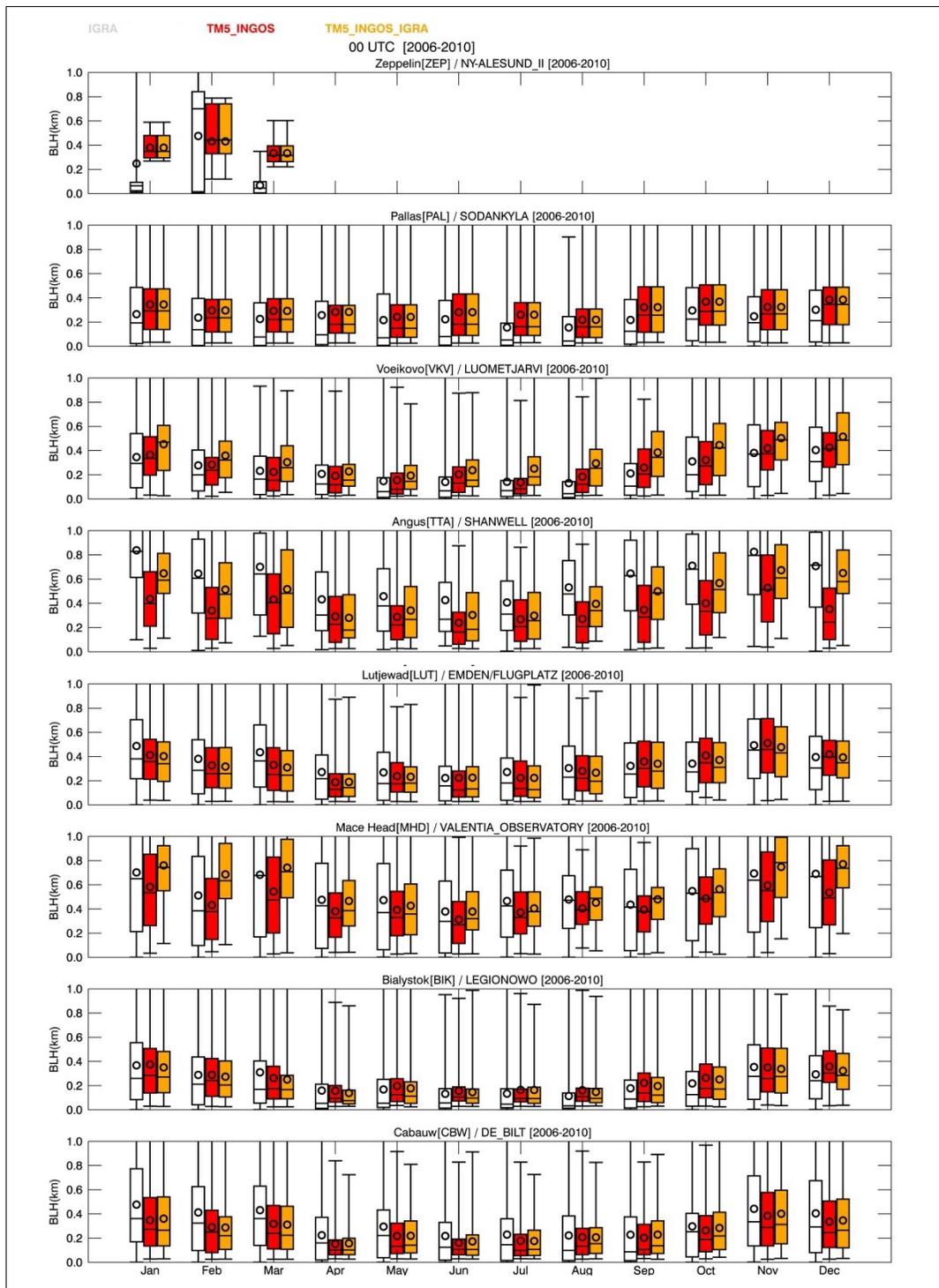
<sup>1</sup>measurements at 3.5m 'normalized' to sampling height of 15\_m based on wind-speed dependent correction (see Section 2.2)



**Figure 1:** Observational network of InGOS greenhouse gas (CH<sub>4</sub>, N<sub>2</sub>O) and radon (<sup>222</sup>Rn) concentration measurements and boundary layer height observations, blue diamonds (◇):InGOS stations that measure CH<sub>4</sub> and/or N<sub>2</sub>O concentrations; red circles (○):InGOS stations that measure radon (<sup>222</sup>Rn) activity concentrations; black dots (●): aAll existing IGRA stations; red dots (●): IGRA station closest to InGOS station; triangles (Δ):ceilometer/lidar measurement sites (i.e., Cabauw/Trainou); (●) IGRA stations that have data for the period under study and associated to the closest InGOS stations. - The acronyms for the stations measuring <sup>222</sup>Rn activity concentrations are compiled in Table 1.



**Figure 2:** Radon (<sup>222</sup>Rn) emissions used for the model simulations. (a) spatial distribution of InGOS emissions over Europe during July 2009. (b) seasonal and inter-annual variations of InGOS emissions (in different colors for different years; mean in red) and the commonly used constant emissions (black). The mean seasonal variations are averaged over the geographic domain between 10°W and 30°E longitude and between 35°N and 70°N latitude.



**Figure 3:** Observed (IGRA; [blank](#)) and modelled (TM5\_INGOS; [red](#) and TM5\_INGOS\_IGRA; [orange](#)) BLHs for InGOS stations at 00 UTC (2006-2010). The titles of each panel show the names and acronyms of the InGOS station, and the names of the nearest IGRA station used for comparison. The Whisker plots show the monthly minimum and maximum values (bars), and the 25% and 75% percentiles (boxes). The median values are given by the horizontal line and the mean values by the open circles in the boxes. ~~The IGRA data are in blank and the various colors represent the modelled BLHs.~~ The different acronyms of the model data are defined in Section [3.24](#) of the text

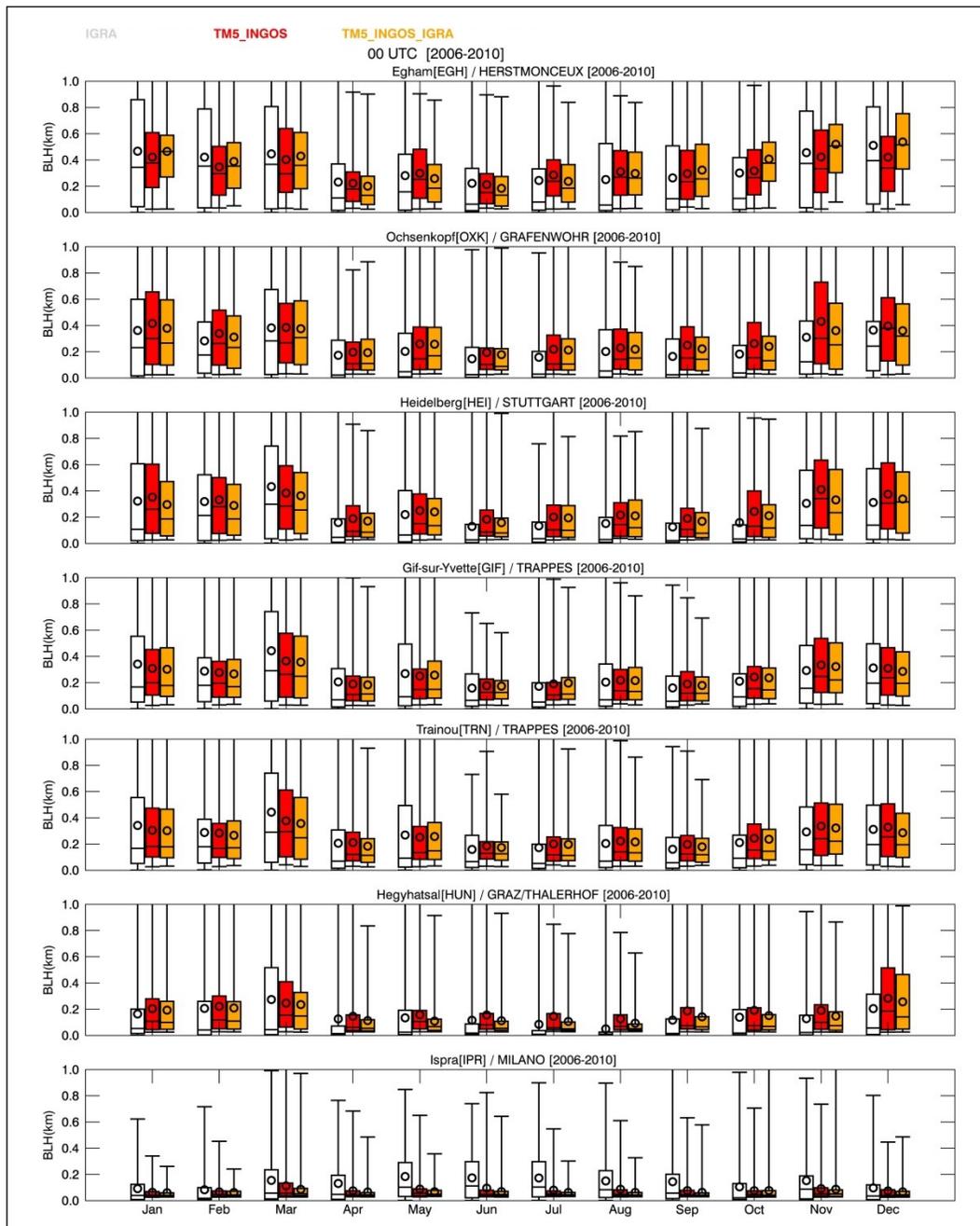


Figure 3: continued

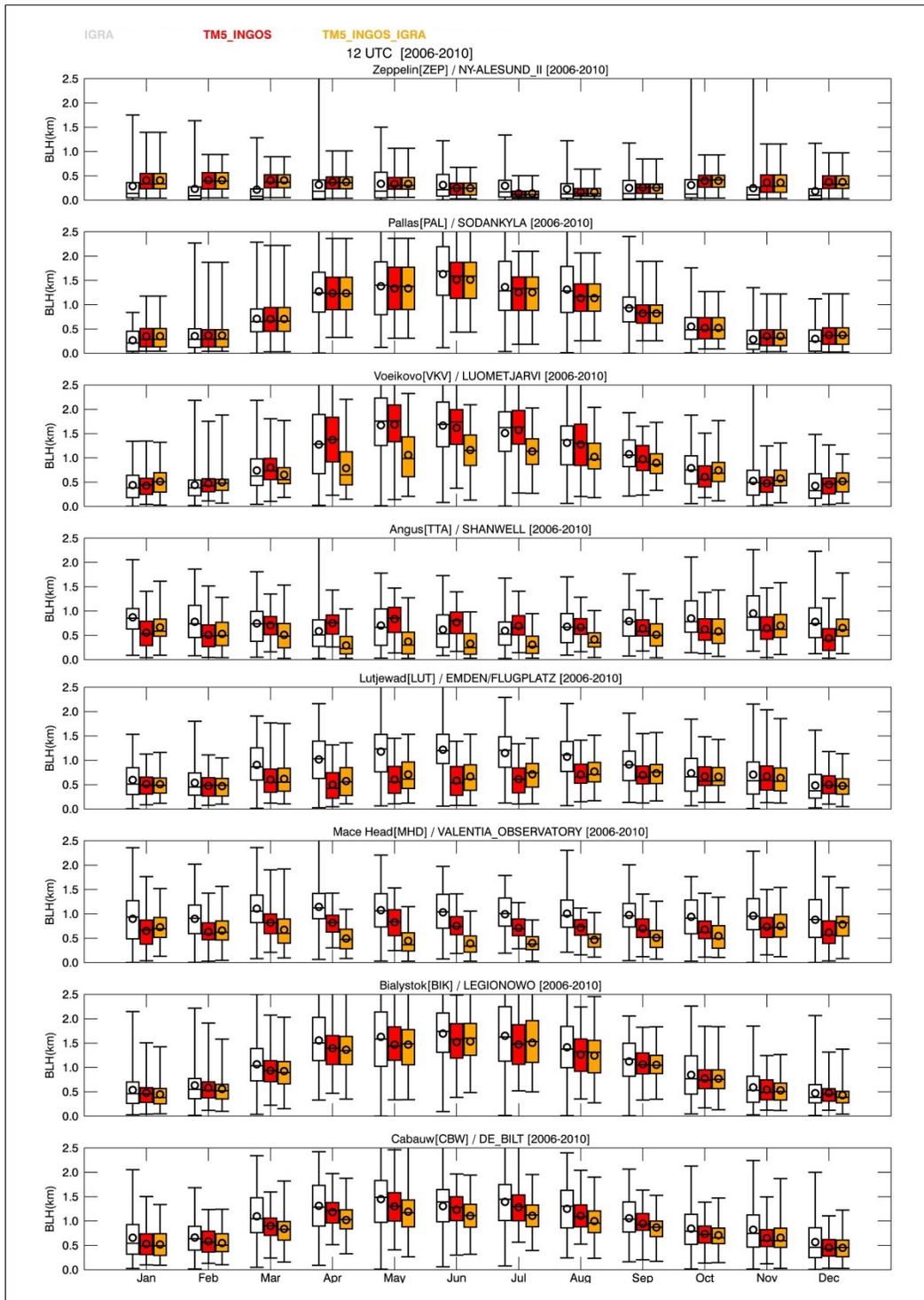


Figure 4: As Figure 3, but at 12 UTC

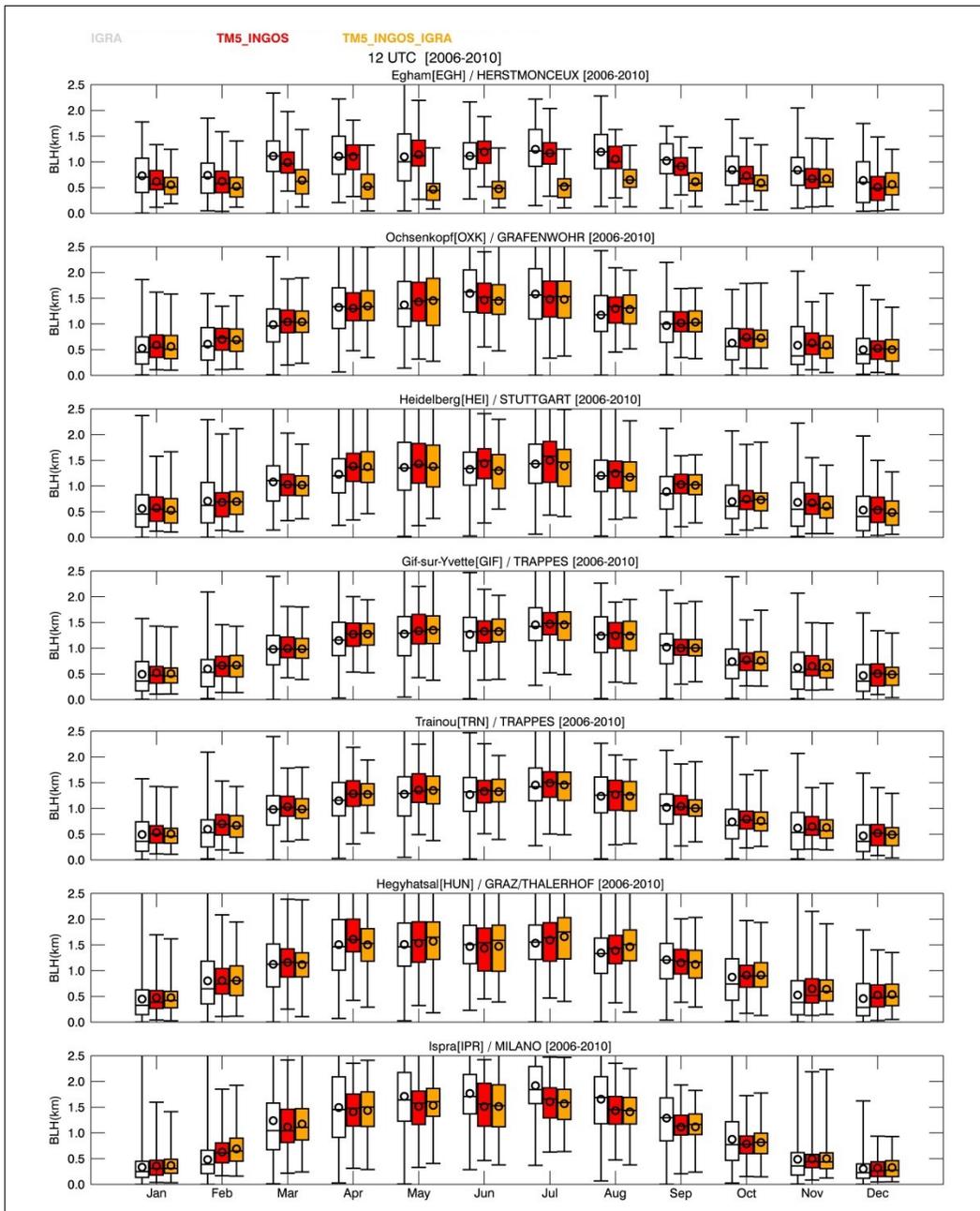
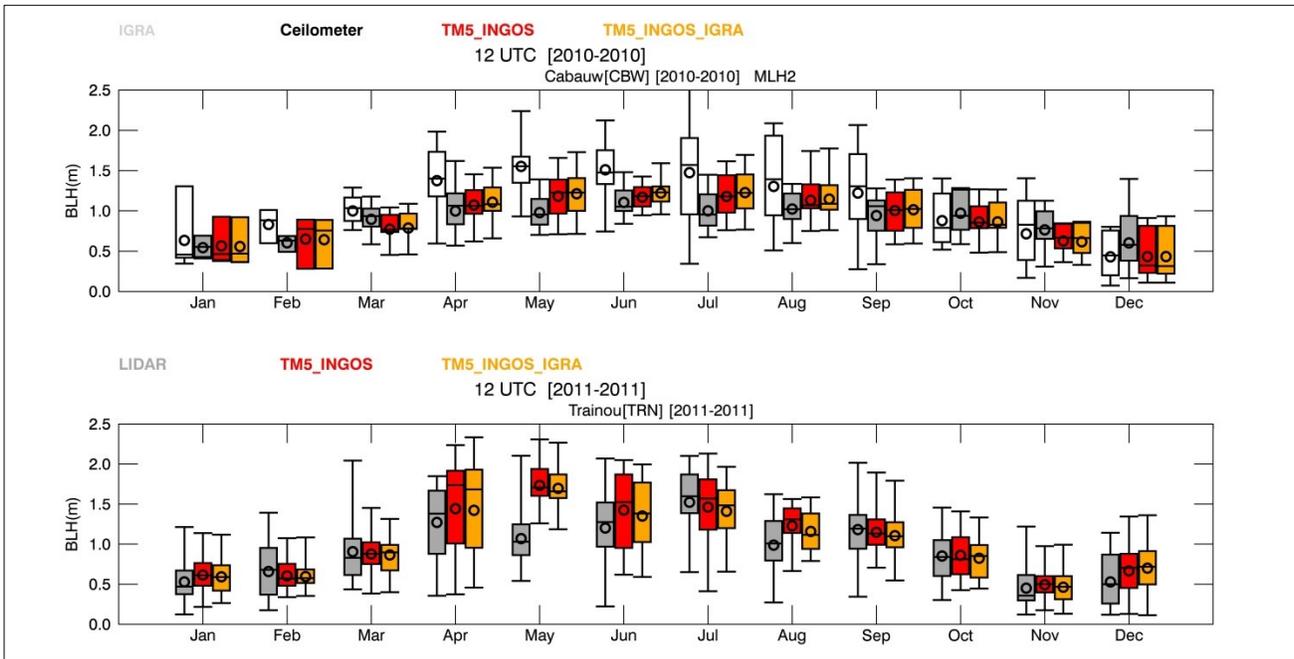
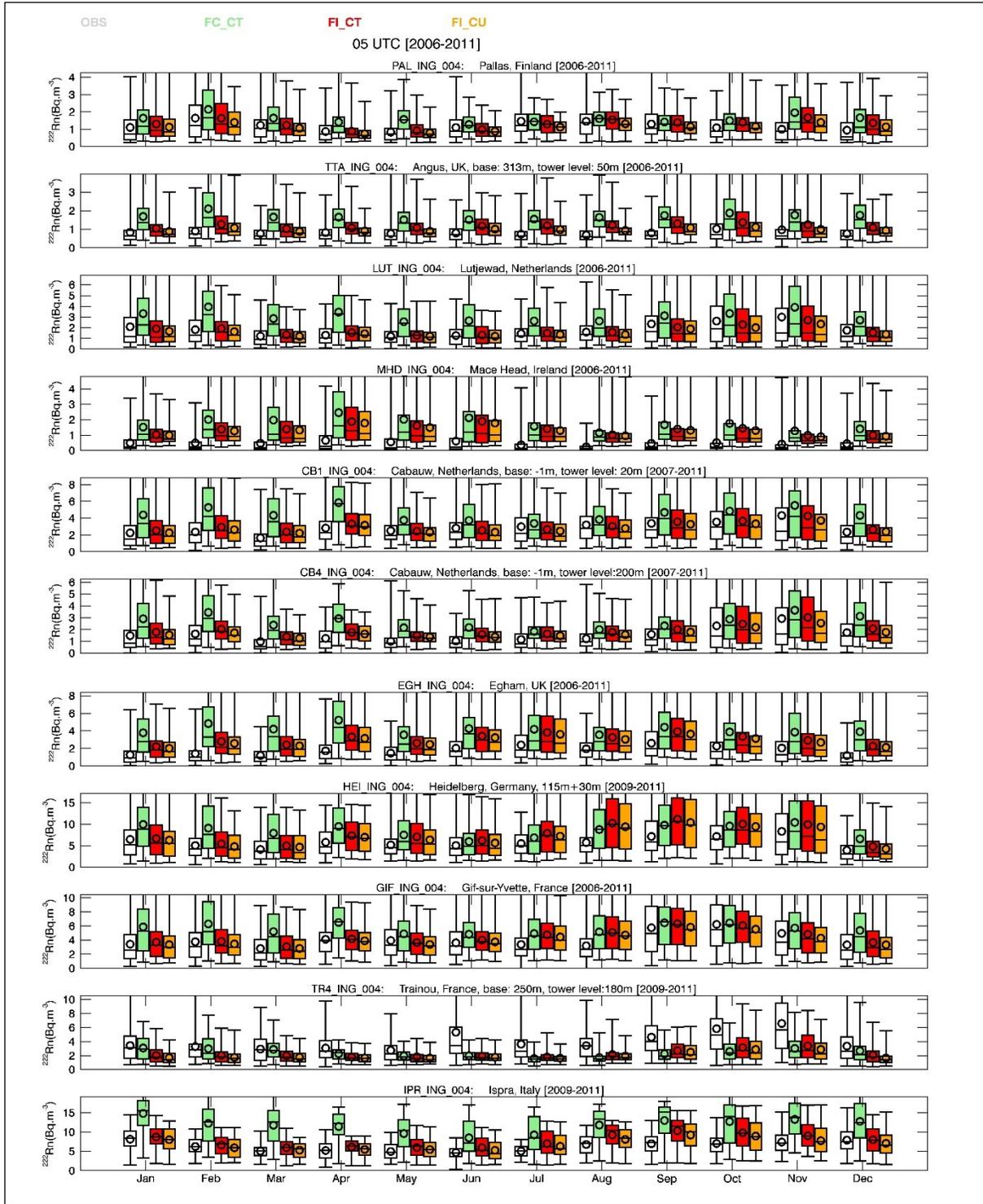


Figure 4: continued



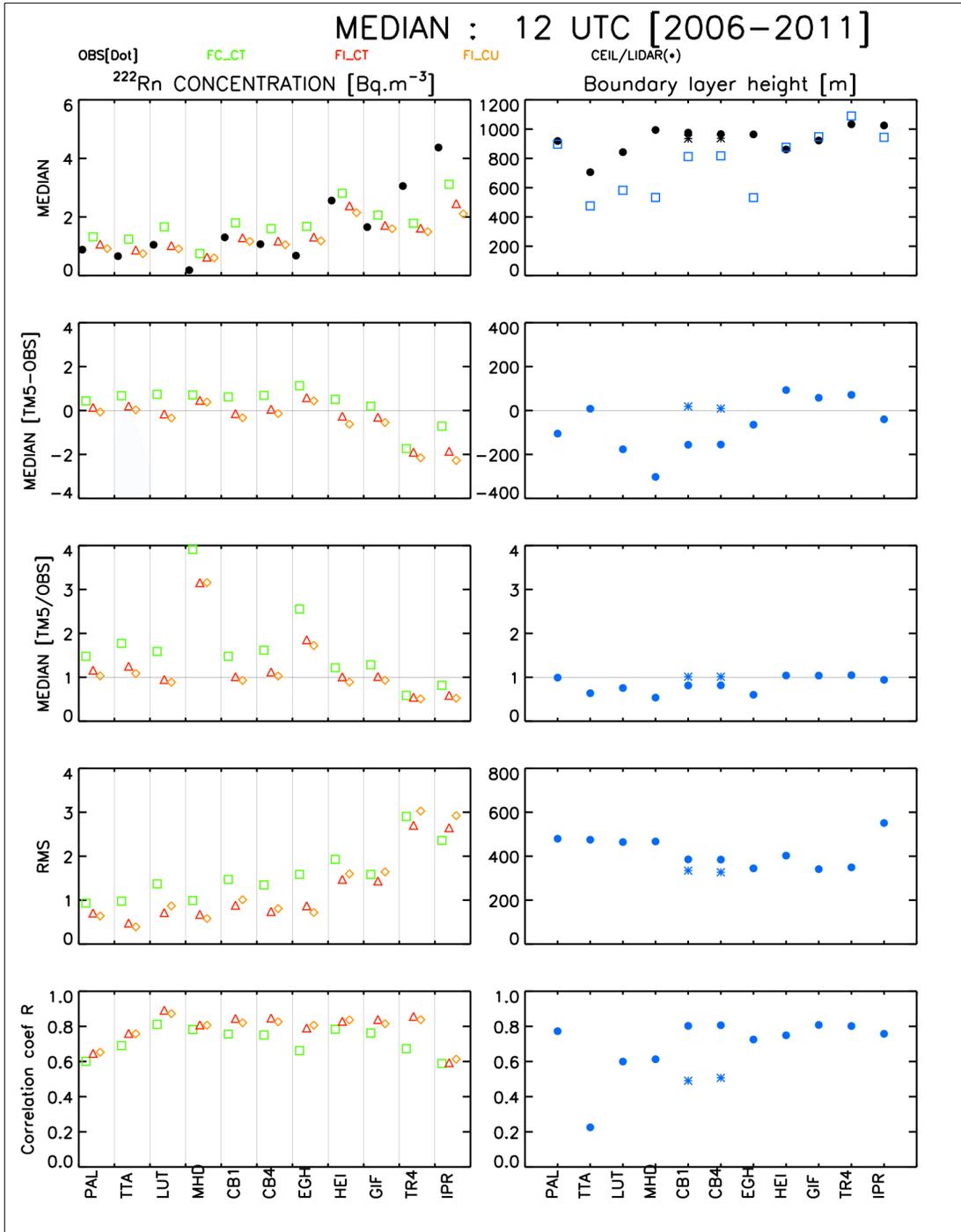
**Figure 5:** As Figure 3, but on the top) Cabauw (CBW) where both ceilometer and nearby IGRA observations (from De Bilt) are available. Observed (IGRA in blank; ceilometer in dark grey) and simulated (colors) boundary layer heights at 12 UTC and for 2010 are shown. IGRA data and ceilometer data are shown in blank and dark grey, respectively. On the bottom, Trainou (TRN) lidar based boundary layer heights (dark grey) at 12 UTC during 2011 are shown. The model boundary layer heights are represented by the colored boxes. The different acronyms are defined in Section 3.24 of the text



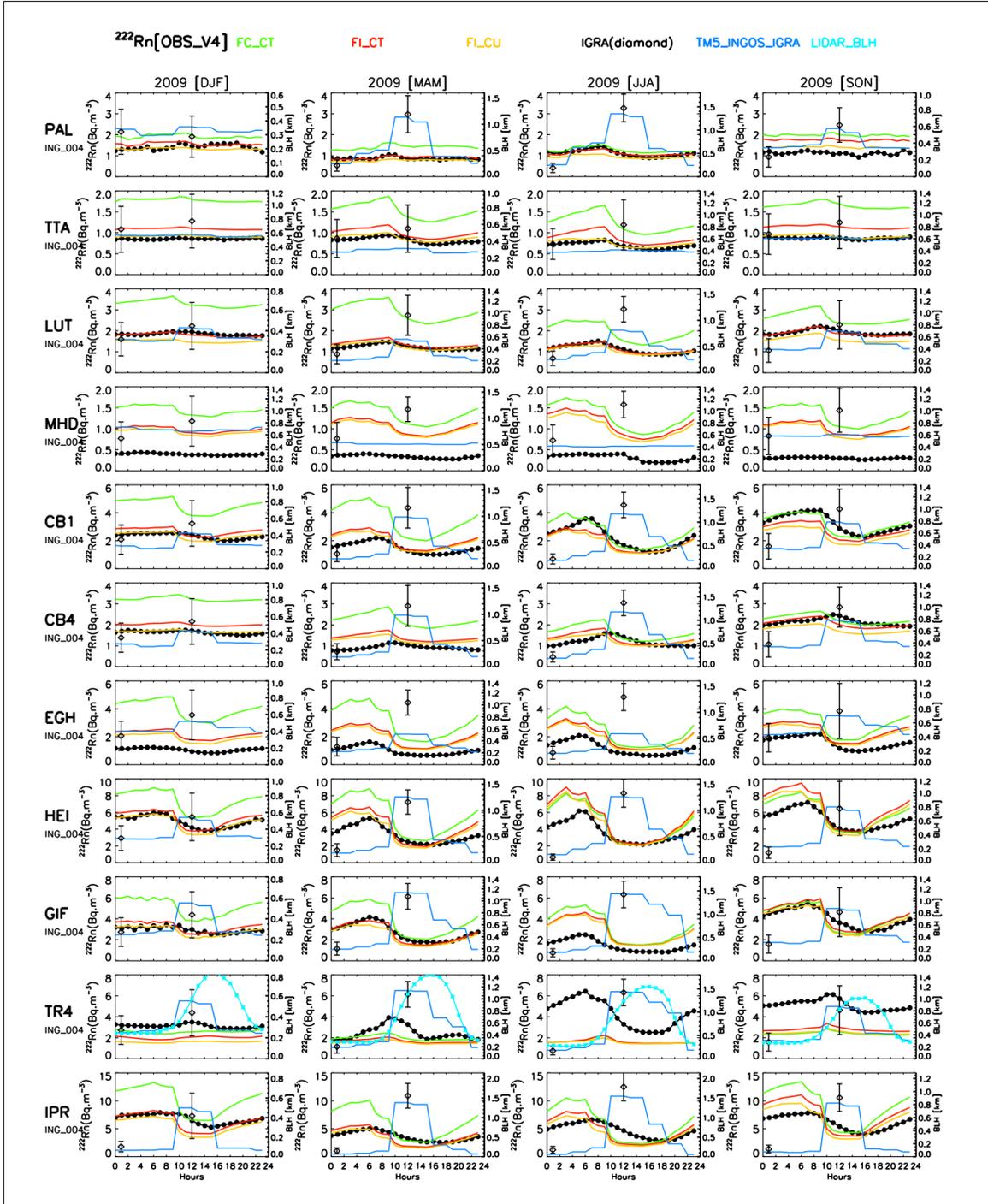
**Figure 6:** Seasonal variations of daily maximum of observed and simulated radon ( $^{222}\text{Rn}$ ) activity concentrations at InGOS sites at 05 UTC (2006-2011). The Whisker plots show the monthly minimum and maximum values (bars), and the 25% and 75% percentiles (boxes). The median values are given by the horizontal line and the mean values by the open circles in the boxes. The observed radon activity concentrations are shown in black, and the model simulations are represented by the colored boxes (the different acronyms are defined in Section 3.4). FC uses constant  $^{222}\text{Rn}$  fluxes and FI the InGOS flux map



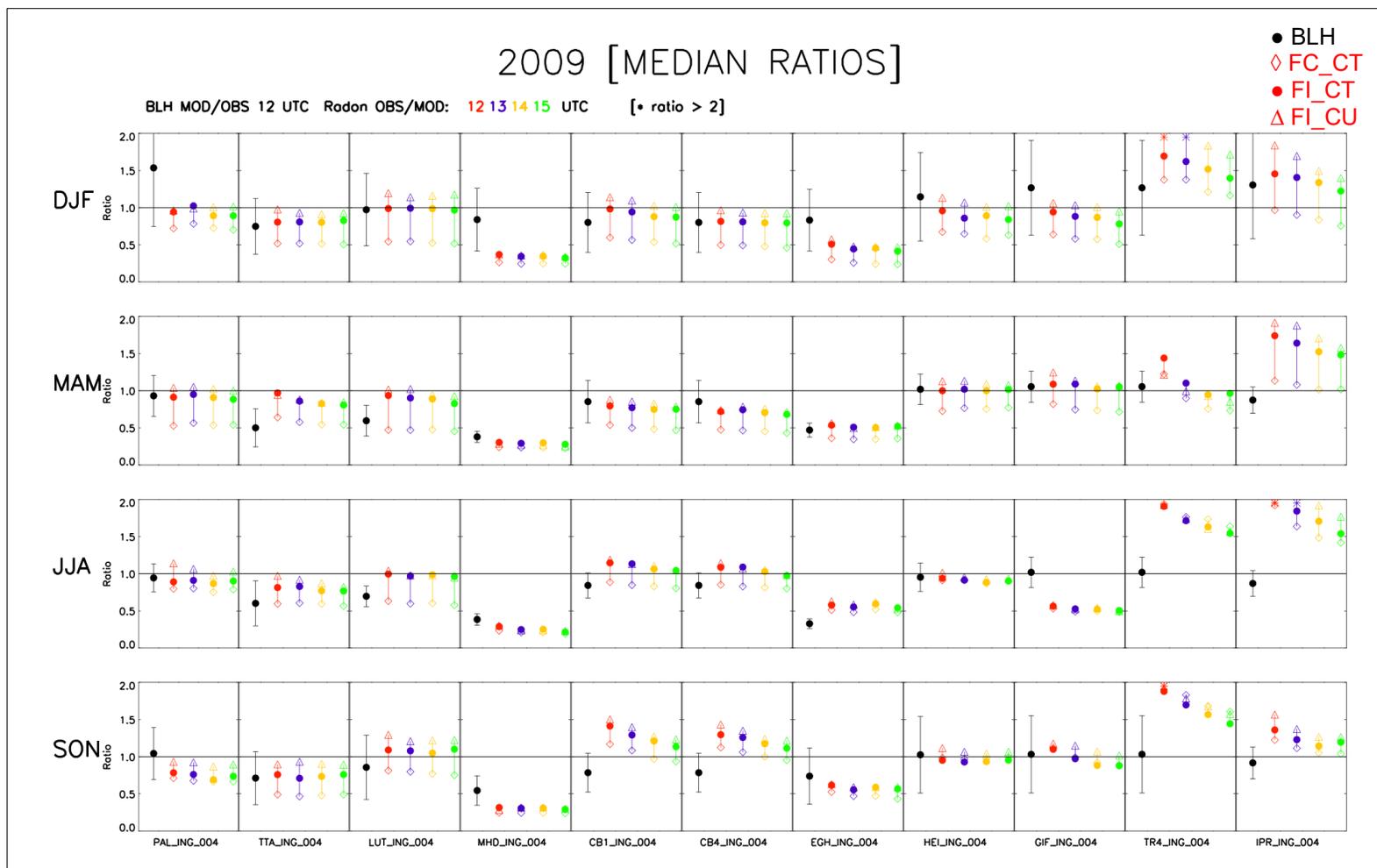
**Figure 7:** As for Figure 6, but at 14 UTC illustrating the seasonal variations of daily minimum of radon ( $^{222}\text{Rn}$ ) activity concentrations



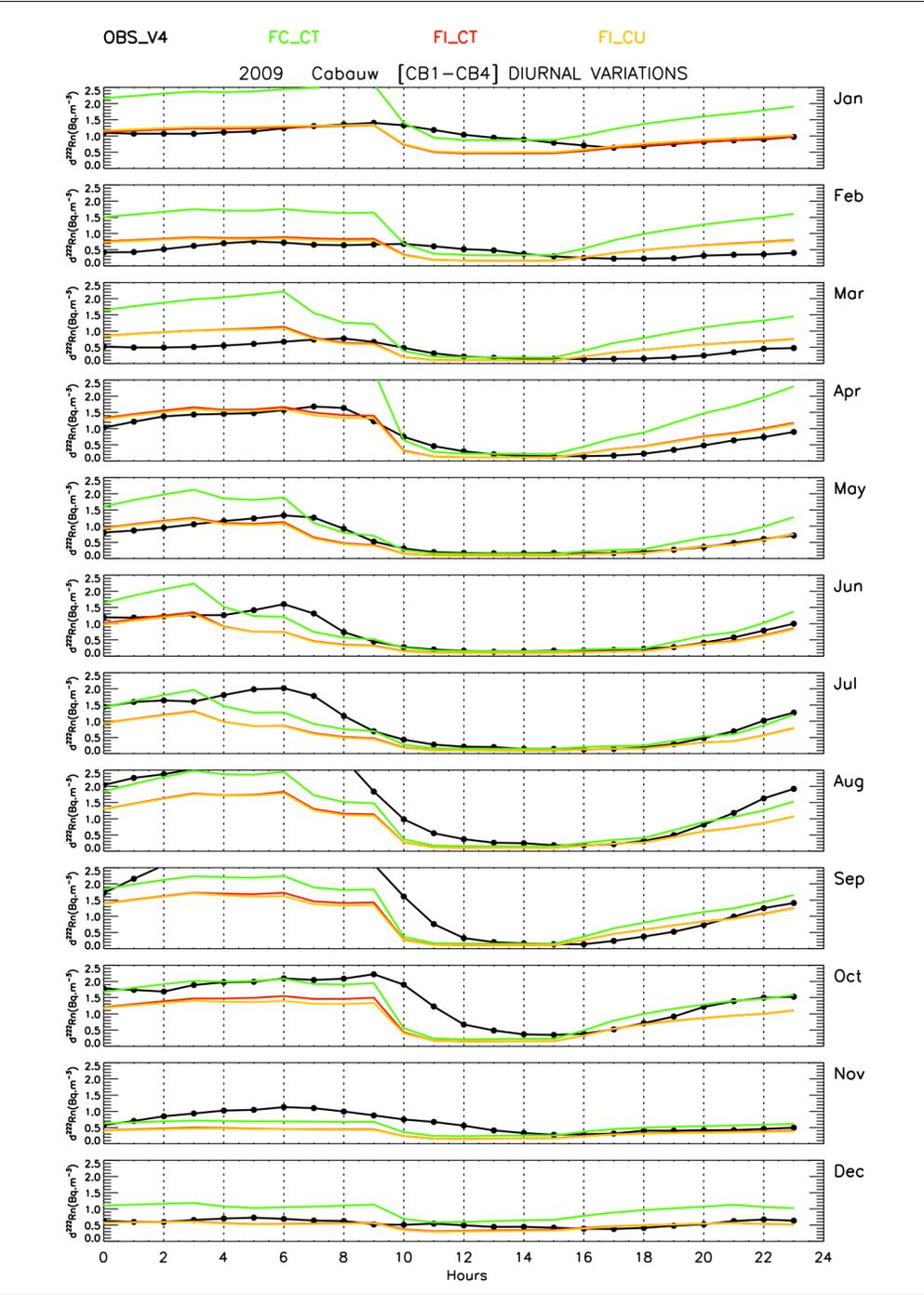
**Figure 89:** Left: statistics of observed vs. simulated  $^{222}\text{Rn}$  activity concentrations for the different stations (12 UTC). Right: statistics of observed (IGRA (●) and ceilometer (CEIL)/LIDAR (\*) vs. simulated boundary layer heights (TM5\_INGOS\_IGRA). (12UTC). The acronyms of the stations (x-axis) are given in Table 1. For the median and RMS values, the unit of the y-axis is given on the top of the relevant graphs. The different model settings are given on the top of the graphs. The number of pair of data for each station is larger than 500



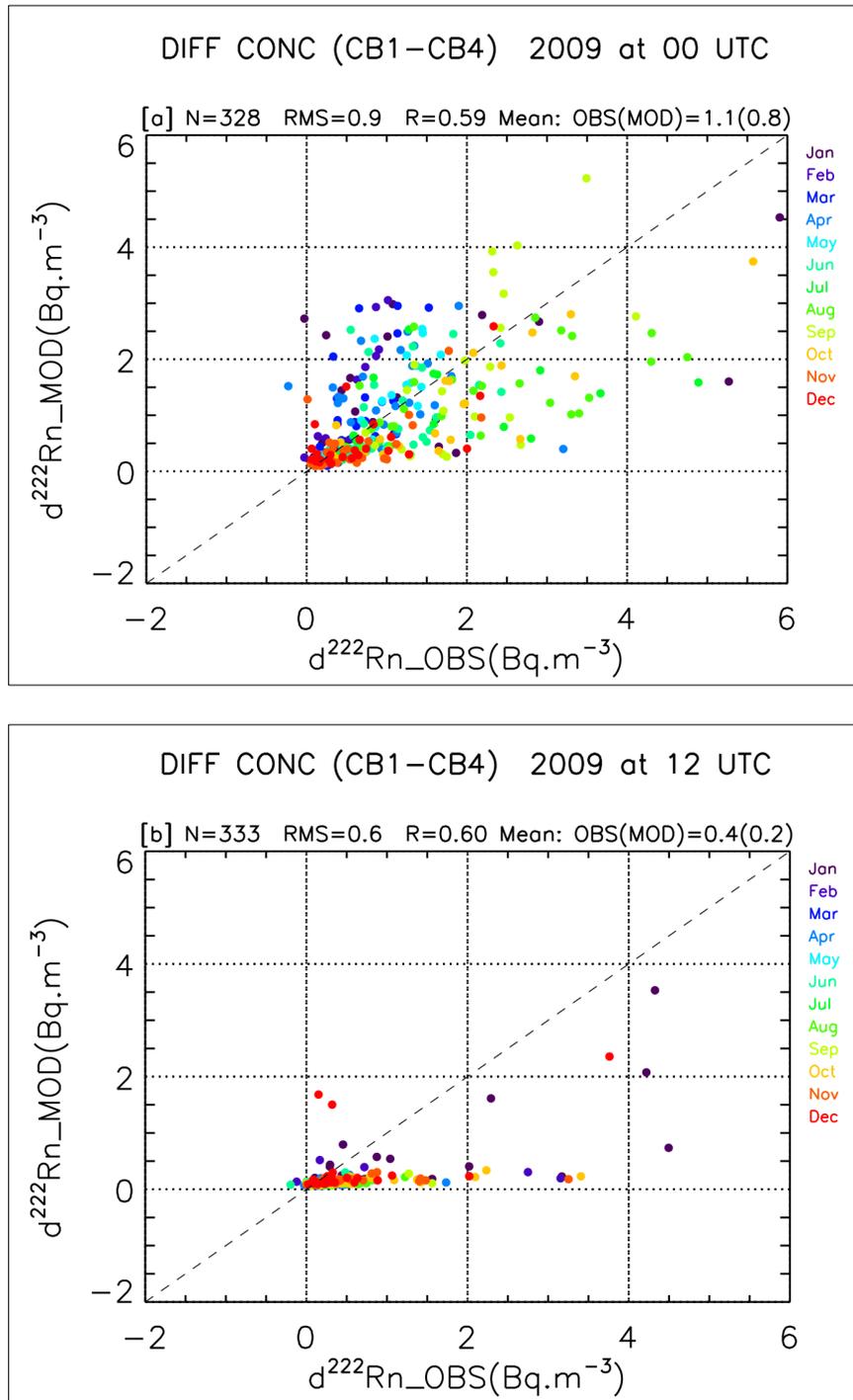
**Figure 9:** Seasonal variations of  $^{222}\text{Rn}$  activity concentrations and boundary layer heights (BLH) at the InGOS stations that measure  $^{222}\text{Rn}$  activity concentrations. The observed concentrations are represented by the black solid line with dots. Three model simulations are considered: FC\_CT, the model simulations using constant emissions, FI\_CT using the InGOS emissions and the default convection scheme of TM5; FI\_CU using the ~~the~~ InGOS emissions and the combination of the “revised slopes scheme” and the new convection scheme based on ECMWF reanalyses. The BLHs of TM5 (TM5\_INGOS\_IGRA) are in dark blue, while observed IGRA BLHs at 00 and 12 UTC are shown by the black diamonds together with their uncertainties. The lidar BLHs at Trainou (for 2011) are shown by the light blue line



**Figure 10:** The seasonal variations of the ratios of BLHs (TM5/IGRA; black dot with solid line) at 12 UTC and the ratios of  $^{222}\text{Rn}$  activity concentrations (OBS/TM5) at 12, 13, 14, and 15 UTC for the 4 seasons [DJF; MAM; JJA; SON] of the year 2009 and for ~~all~~ each InGOS  ~~$^{222}\text{Rn}$  radon~~ measurement sites. The closest IGRA station to the radon measurement site is considered ([see Figure 1](#)). Three TM5 simulations are shown here: The model simulations using the constant emissions (~~f~~FC\_CT; colored diamond), InGOS emissions and using the default convection scheme of TM5 (FI\_CT; colored ~~filled circles~~ ~~big dots~~) and using the new convection scheme (FI\_U; colored triangles)



**Figure 11:** Mean diurnal variations of the radon activity concentration differences between the two measurement levels at Cabauw (20\_m [CB1], 200\_m [CB4]). The observed gradient is shown by the black solid line with dots (for each month of the year 2009), and the modelled gradient by the solid greenblue line for the constant emissions (FC\_CT), by the solid red line for the InNGOS emissions (FI\_CT), and for the solid orangegreen-line for the simulations using the InNGOS emissions andwith the combination of the “revised slopes scheme” and the new convection scheme based on ECMWF reanalyses (FI\_CU), respectively.



**Figure 12:** Correlation plots between the simulated and observed vertical  $^{222}\text{Rn}$  activity concentration gradients (difference between 200m (CB4) and 20m (CB1)) at Cabauw at 00 UTC (top) and 12 UTC (bottom). Model simulations using InGOS emissions (FI\_CT termed as MOD) are shown. Each color indicates the month at which the data are obtained.

## Evaluation of the boundary layer dynamics of the TM5 model over Europe

E. N. Koffi<sup>1</sup>, P. Bergamaschi<sup>1</sup>, U. Karstens<sup>2,3</sup>, M. Krol<sup>4,5,6</sup>, A. Segers<sup>7</sup>, M. Schmidt<sup>8,11</sup>, I. Levin<sup>8</sup>, A. T. Vermeulen<sup>9,3</sup>, R. E. Fisher<sup>10</sup>, V. Kazan<sup>11</sup>, H. Klein Baltink<sup>12</sup>, D. Lowry<sup>10</sup>, G. Manca<sup>1</sup>, H. A. J. Meijer<sup>13</sup>, J. Moncrieff<sup>14</sup>, S. Pal<sup>15</sup>, M. Ramonet<sup>11</sup>, H.A. Scheeren<sup>1,13</sup>, A. G. Williams<sup>16</sup>

1 European Commission Joint Research Centre, Institute for Environment and Sustainability, Ispra (Va), Italy

2 Max-Planck-Institute for Biogeochemistry, Jena, Germany

3 ICOS Carbon Portal, ICOS ERIC at Lund University, Sweden

4 SRON Netherlands Institute for Space Research, Utrecht, Netherlands,

5 Institute for Marine and Atmospheric Research Utrecht, Utrecht, University, Utrecht, Netherlands

6 MAQ, Wageningen University and Research Centre, Wageningen, Netherlands,

7 Netherlands Organisation for Applied Scientific Research (TNO), Utrecht, Netherlands

8 Institut für Umweltphysik, Heidelberg University, Germany

9 Energy research Center Netherlands (ECN), Petten, Netherlands

10 Royal Holloway, University of London (RHUL), Egham, UK

11 Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France

12 Royal Netherlands Meteorological Institute (KNMI), Netherlands

13 Centrum voor Isotopen Onderzoek (CIO), Rijksuniversiteit Groningen, Netherlands

14 Atmospheric Chemistry Research Group, University of Bristol, UK

15 Department of Meteorology, Pennsylvania State University, State College, PA, USA

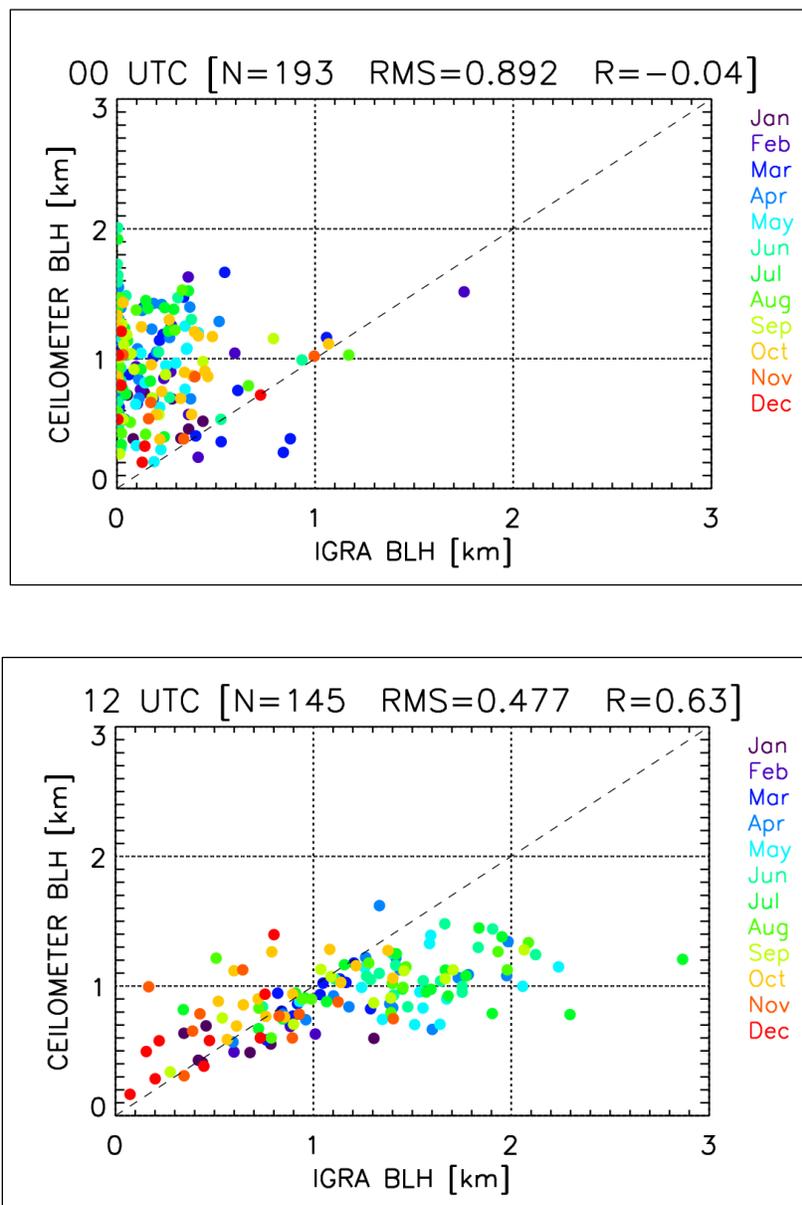
16 Australian Nuclear Science and Technology Organisation (ANSTO) Environment Research Theme, Locked Bag 2001, Kirrawee DC, NSW 2232, Australia

26 July 2016

Revised version for Geosci. Model Dev.

## Supplementary material

## Comparison of IGRA boundary layer heights to ceilometer observations at Cabauw



**Figure S1:** The Cabauw ceilometer boundary layer heights versus IGRA (De Bilt station) data for the year 2010 at 00 (on the top) and 12 (bottom) UTC are shown. The statistics (RMS in km and correlation coefficient  $R$ ) are indicated as well as the number of pair of data ( $N$ ) used to compute these metrics. The different colors indicate the months at which the data were obtained

# Comparison of TM5 boundary layer heights to observations

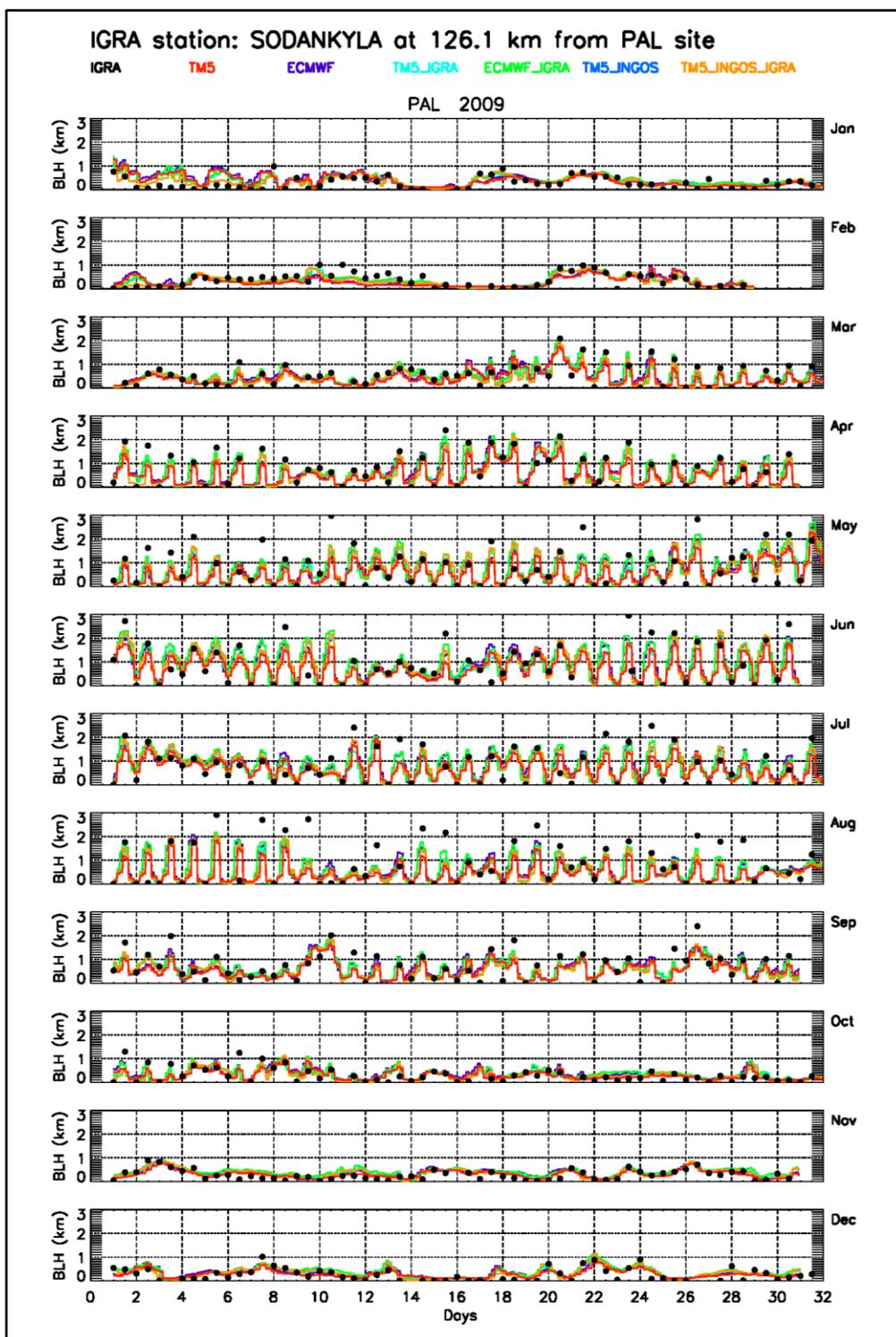
We extract the BLH in the model both at the location of the InGOS station and at the location of the nearest IGRA station, resulting in a set of four different modeled BLHs labelled by the following acronyms:

- 'TM5': TM5 default version (Eq.1 in Section 2.1 with  $R_{ic} = 0.3$ ); extracted at InGOS stations by using 2D interpolation
- 'TM5\_IGRA': As 'TM5', but extracted at IGRA station, which is closest to the selected InGOS station
- 'TM5\_INGOS': BLHs computed in TM5 model adopting the InGOS definition of the BLH (i.e.,  $R_{ic} = 0.25$  and both surface wind and stress velocity are set to zero in Eq.1), extracted at InGOS station. The BLH of the closest model grid point to the selected station is considered.
- 'TM5\_INGOS\_IGRA': As 'TM5\_INGOS', but extracted at IGRA station, which is closest to the selected InGOS station

Furthermore, we evaluate the BLHs as provided by ECMWF analyses and interpolated to TM5 grids (labelled 'ECMWF'). The values of these BLHs are extracted only at the InGOS stations.

For other details see Section 3.2 of the text of the main paper

We show hereafter only the data for the closest IGRA stations to InGOS ones and for year 2009



**Figure S2:** Time series of observed (IGRA in black dot) and modelled (TM5, TM5\_IGRA, TM5\_INGOS, and TM5\_INGOS\_IGRA in colors) boundary layer heights relevant for InGOS station Pallas (PAL) are shown. The closest IGRA station to Pallas is Sodankyla. The distance between IGRA and InGOS stations is given of the top of graph. The different model acronyms are defined at page 3 in this document

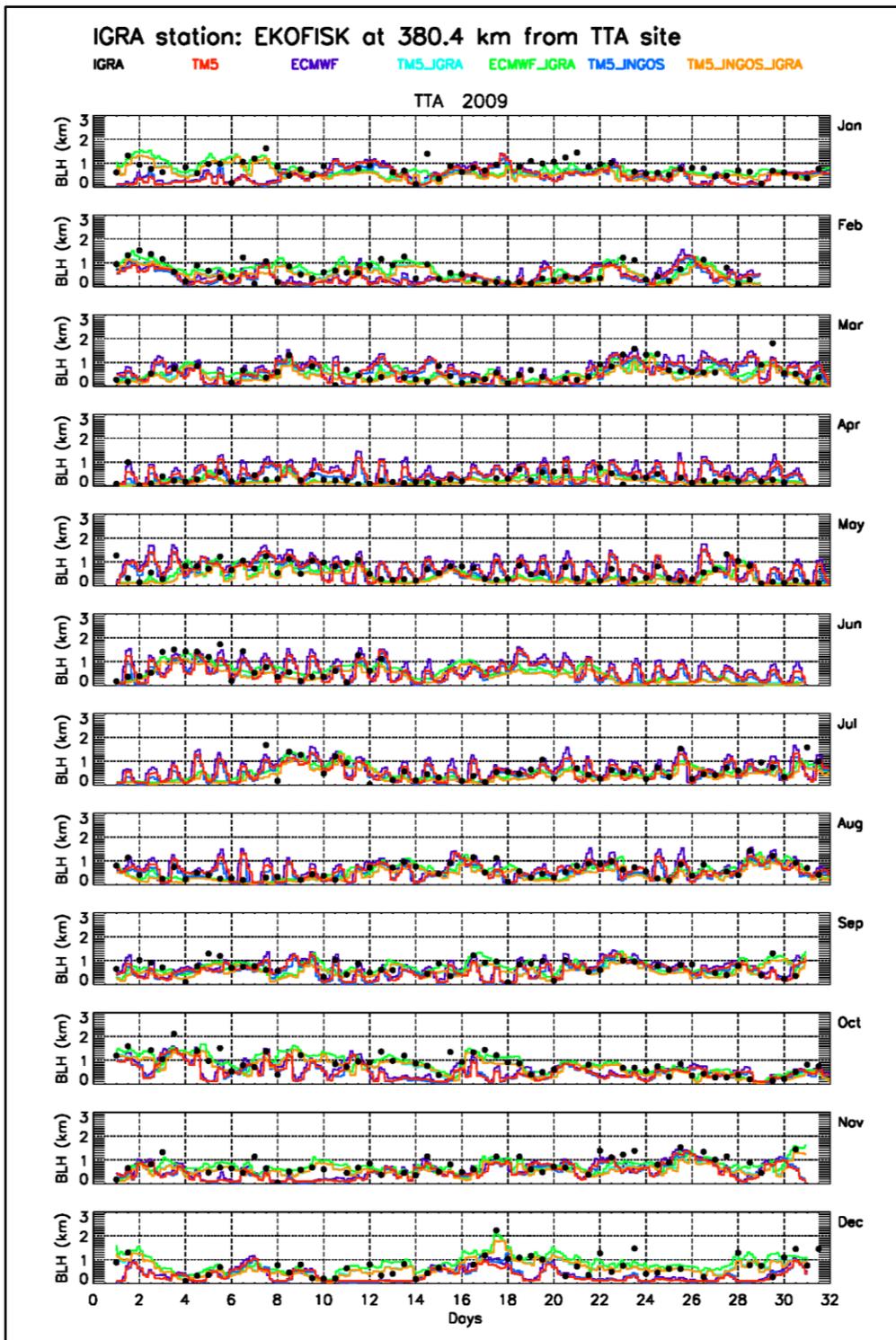


Figure S3: As Figure S2, but for the InGOS station Angus (TTA)

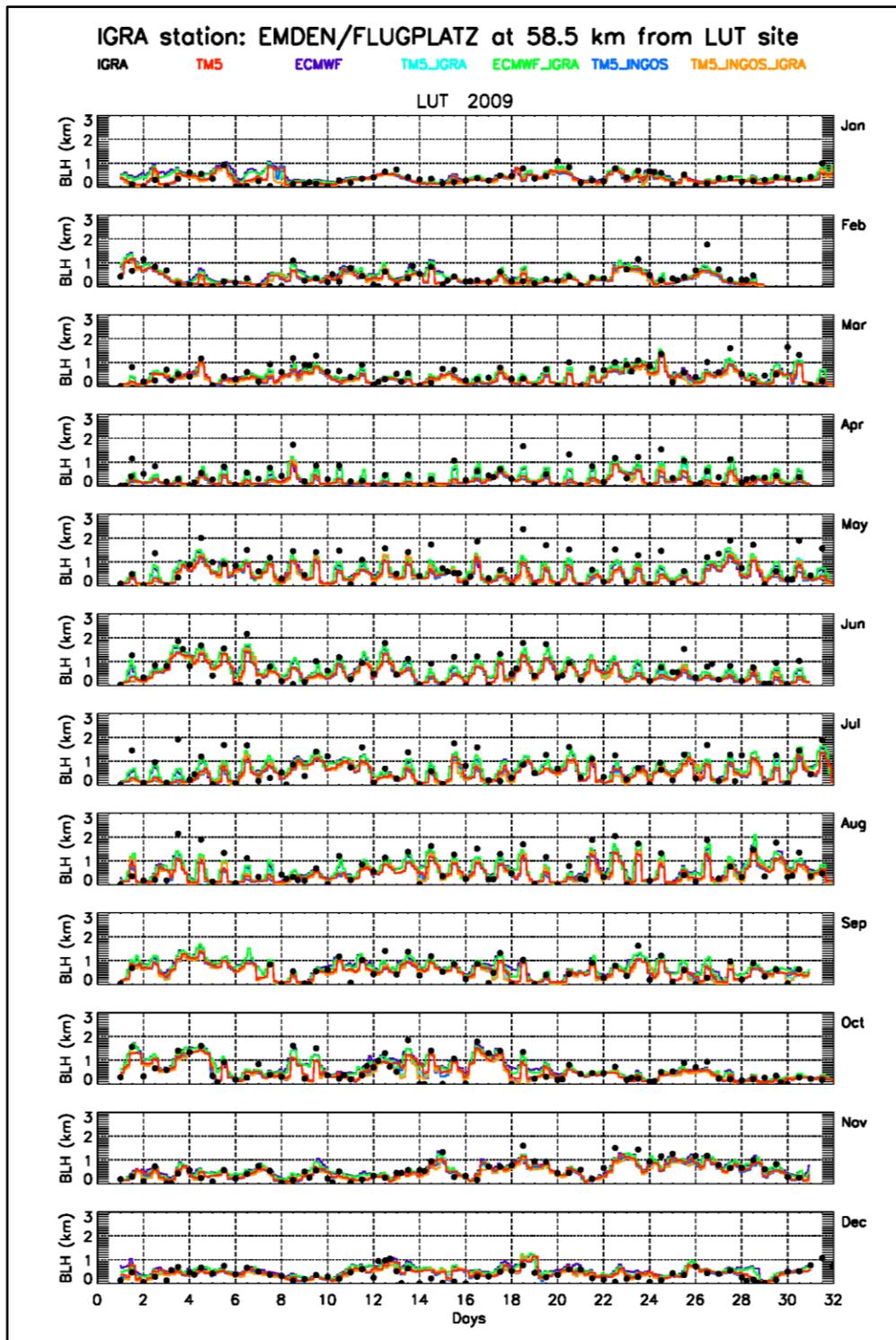


Figure S4: As Figure S2, but of the InGOS station Lutjewad (LUT)

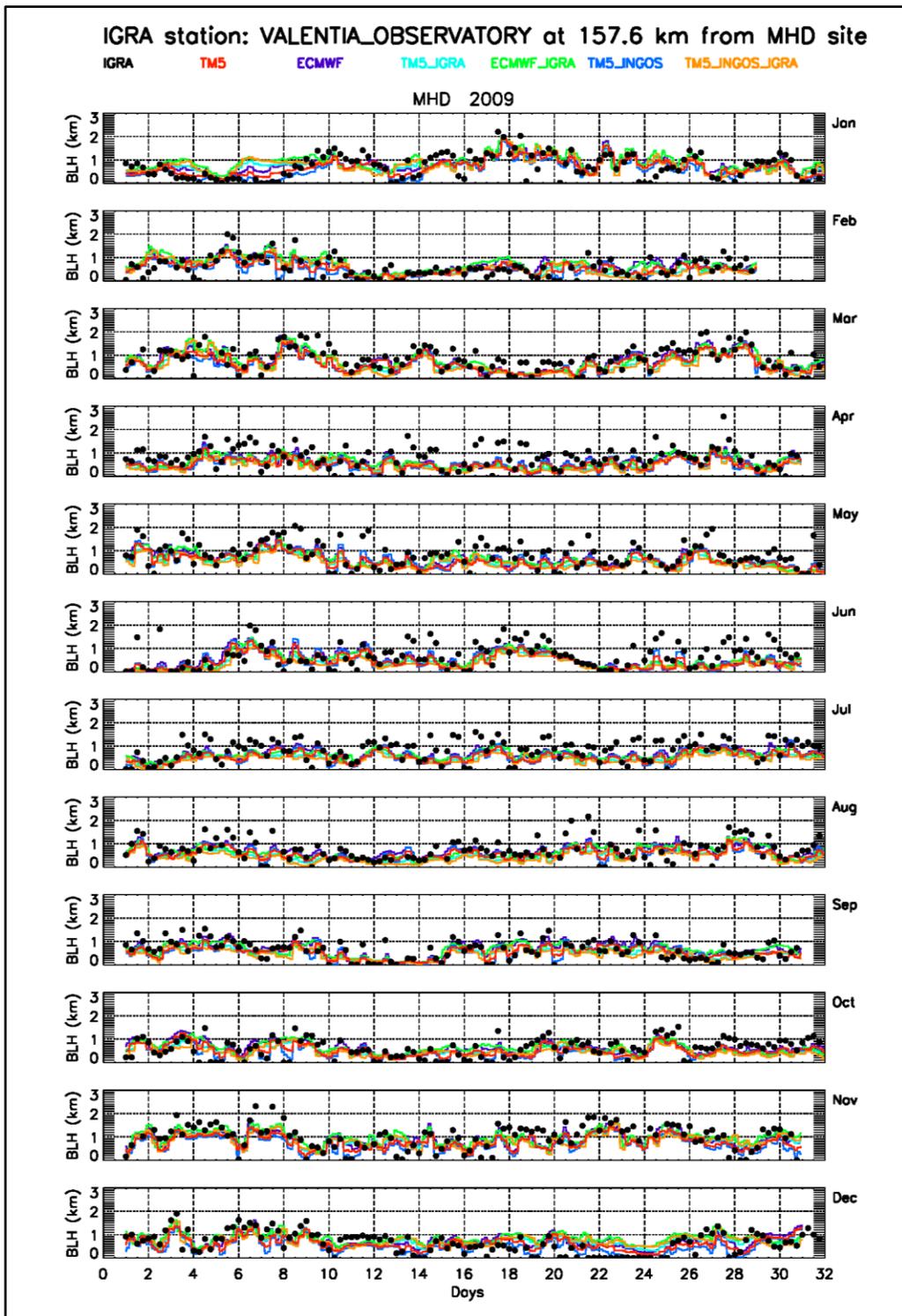
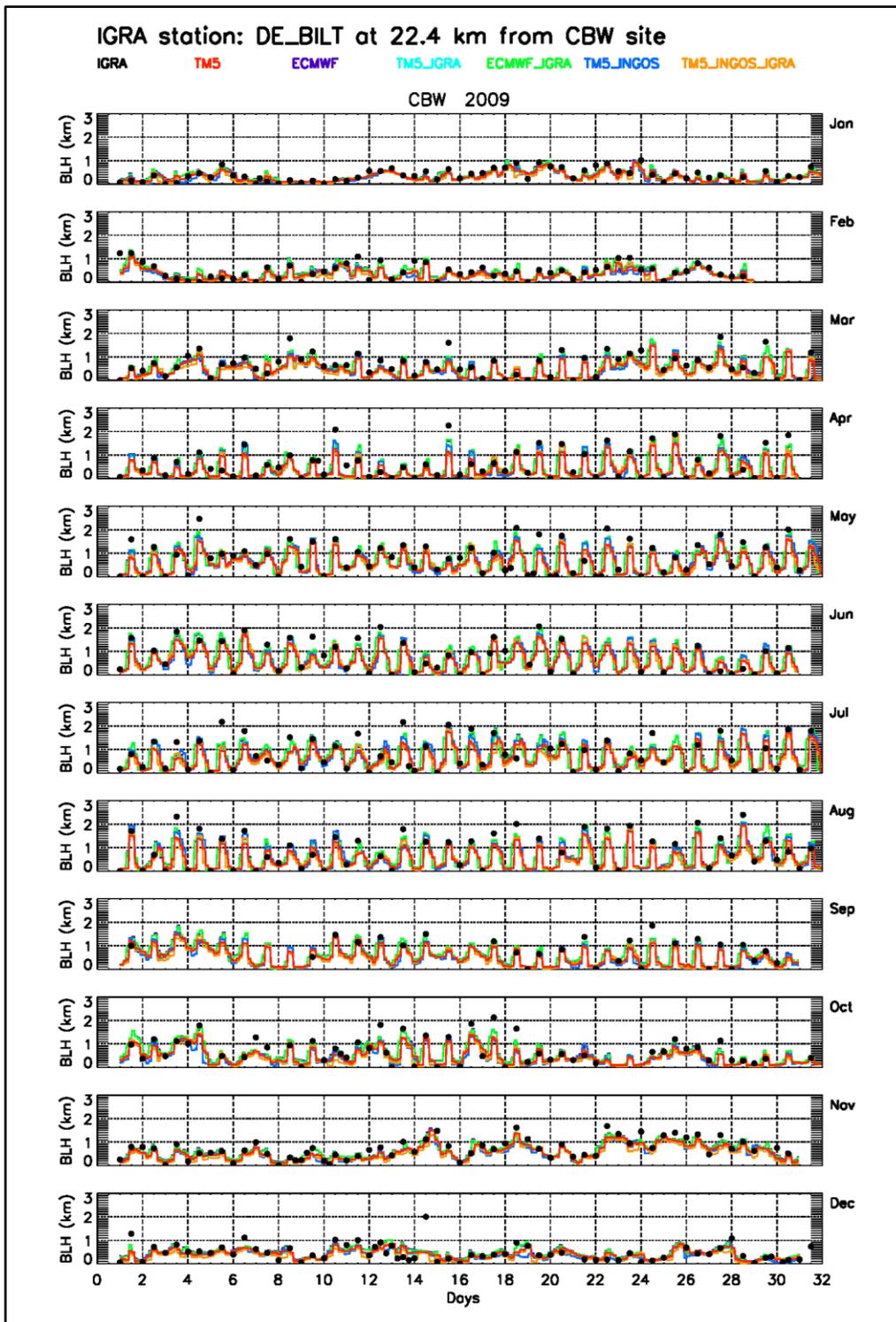
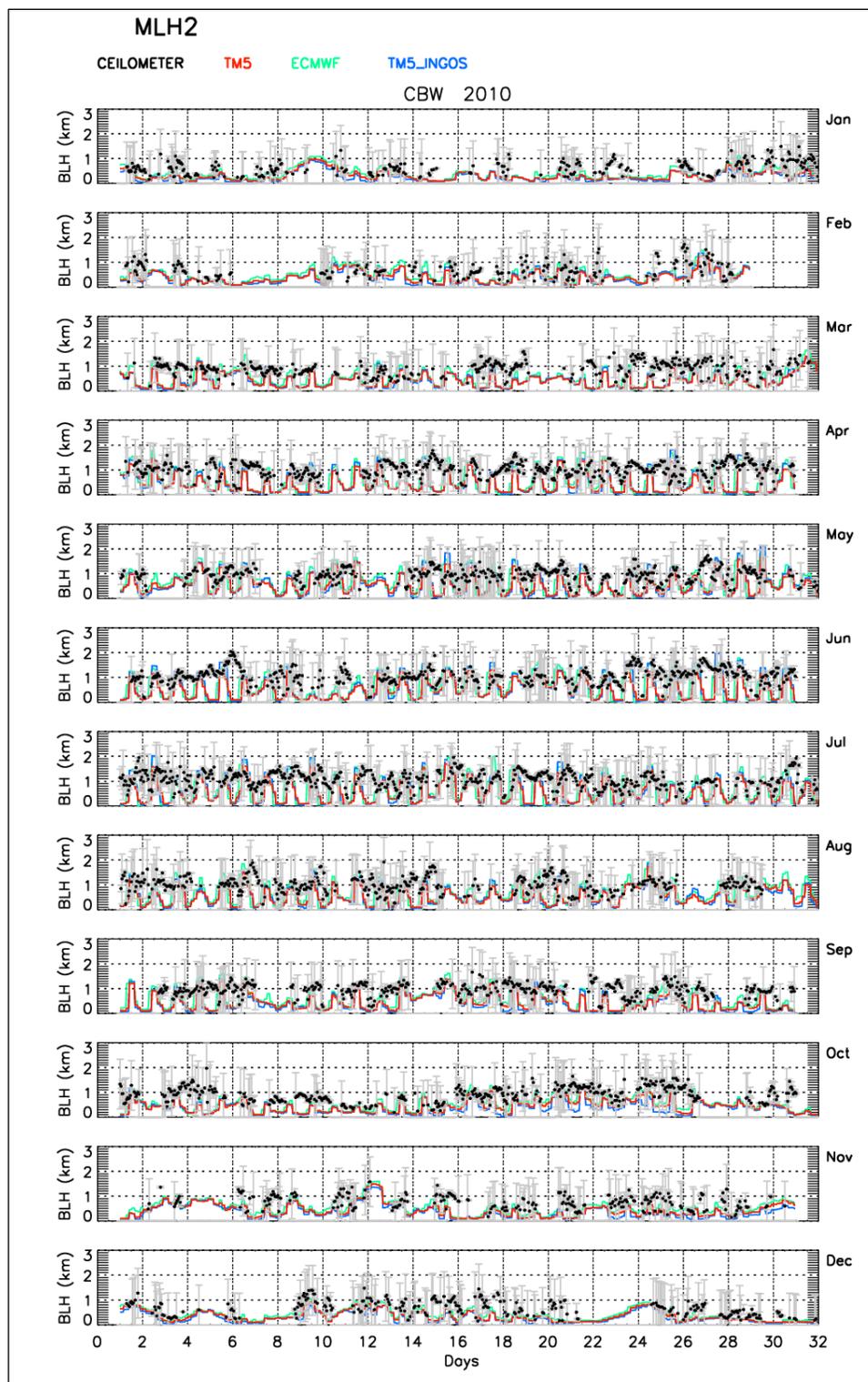


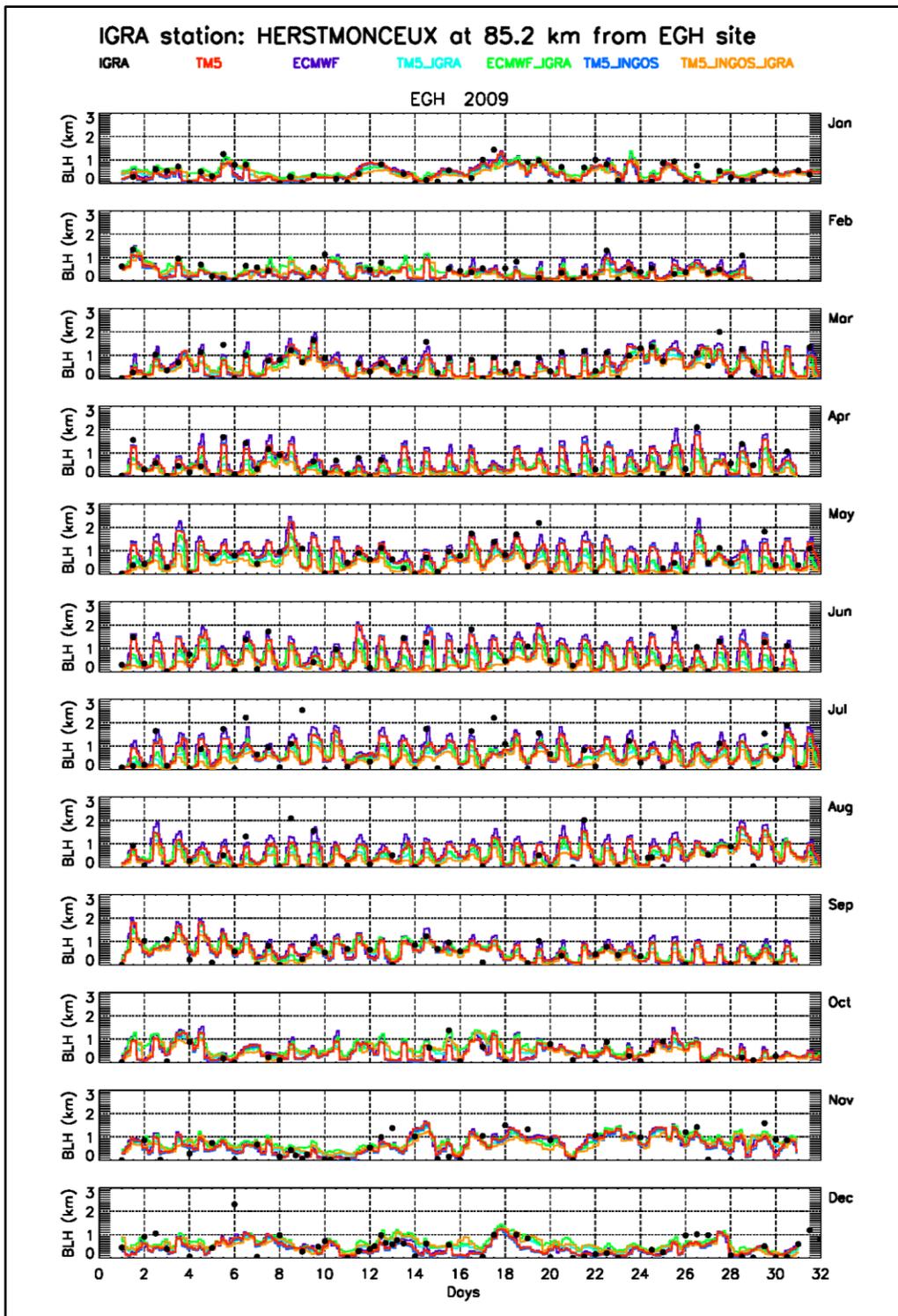
Figure S5: As Figure S2, but for the InGOS station Mace Head (MHD)



**Figure S6:** As Figure S2, but for the InGOS station Cabauw (CBW or CB1/CB4)



**Figure S6b:** As for Figure S6, but the observations for the year 2010 are from the ceilometer measurements at Cabauw (CBW). The standard deviations over 1h period are shown in grey



**Figure S7:** As Figure S2, but for InGOS station Egham (EGH)

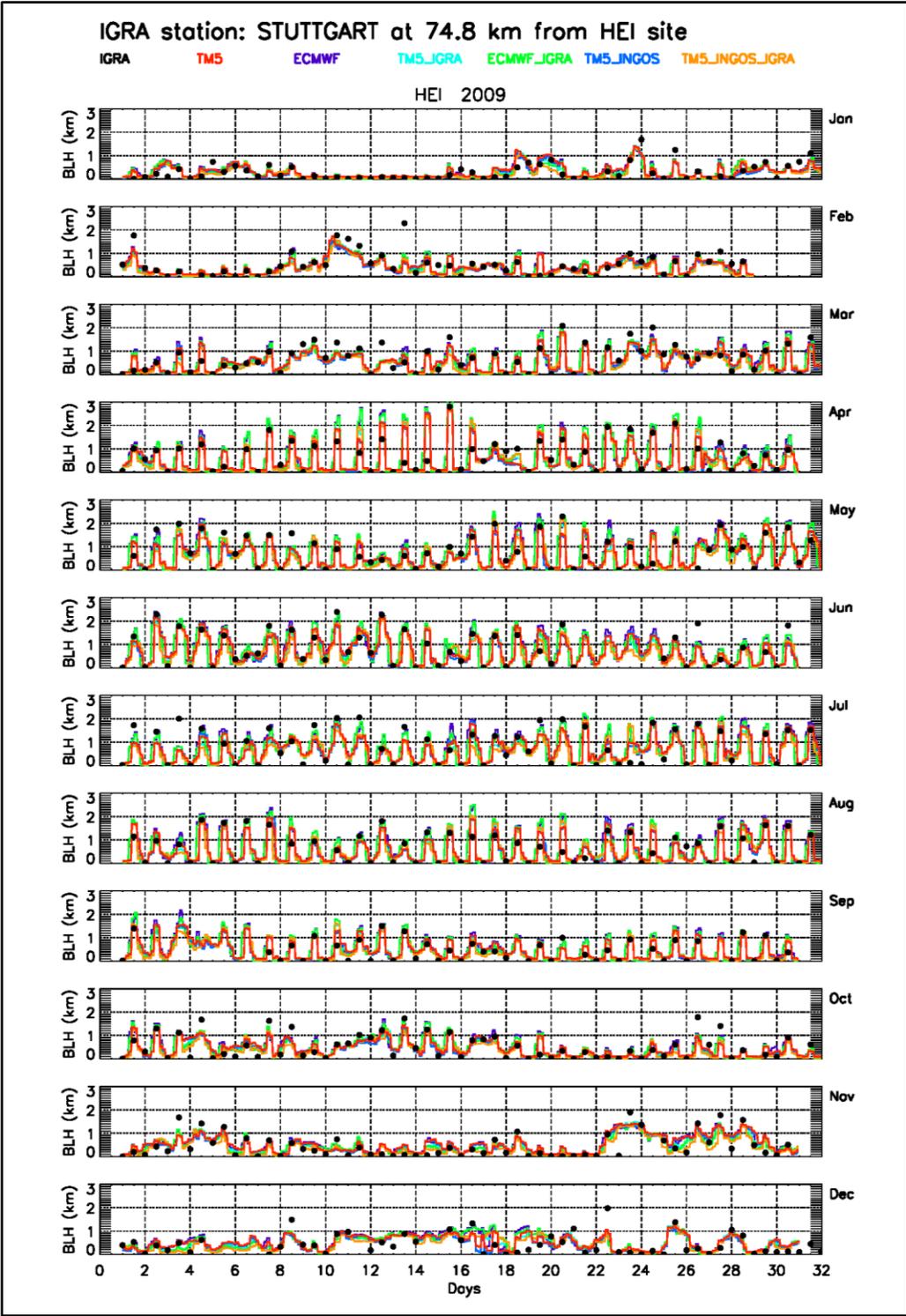
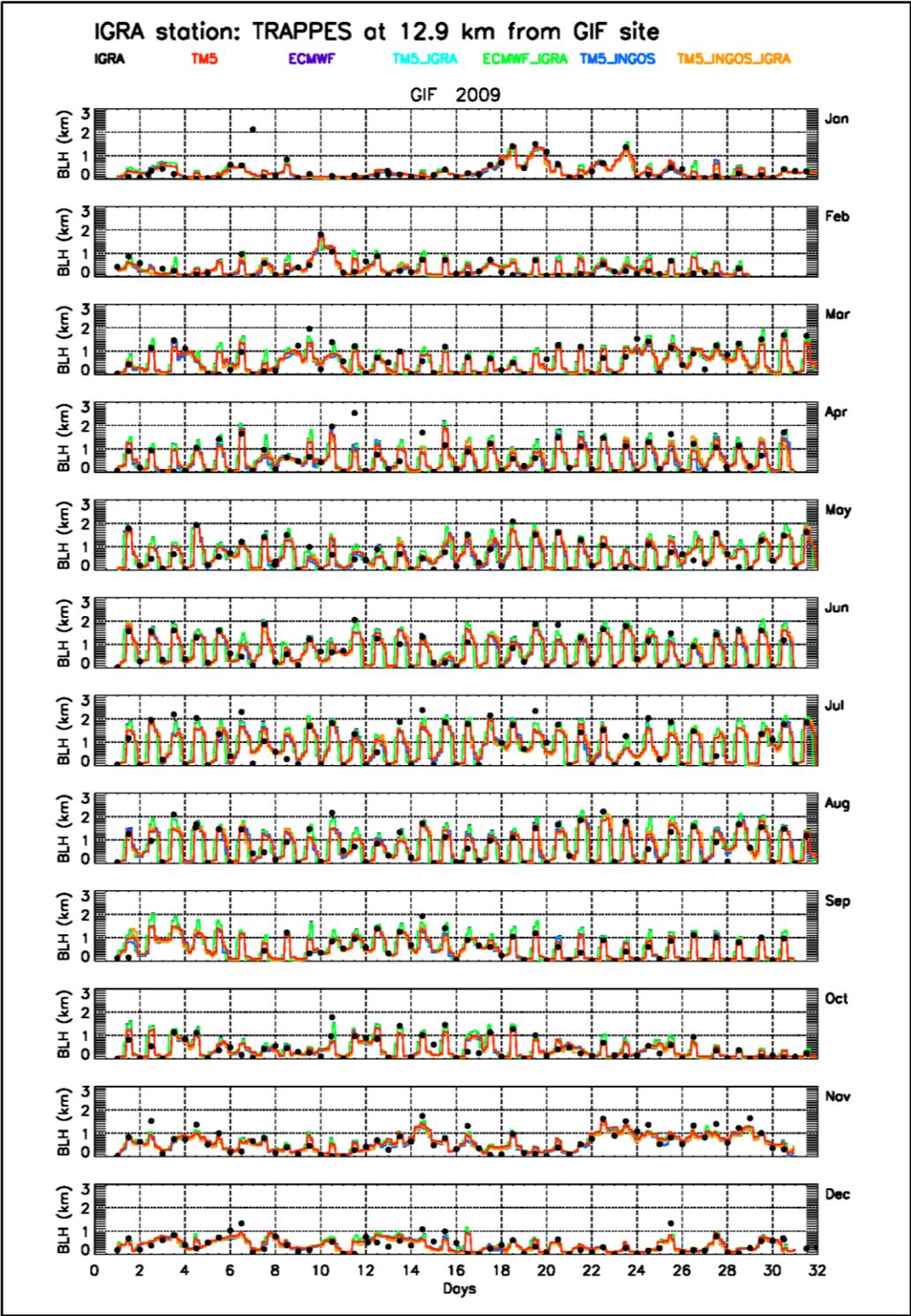


Figure S8: As Figure S2, but for InGOS station Heidelberg (HEI)



**Figure S9:** As Figure S2, but for InGOS station Gif sur Yvette (GIF)

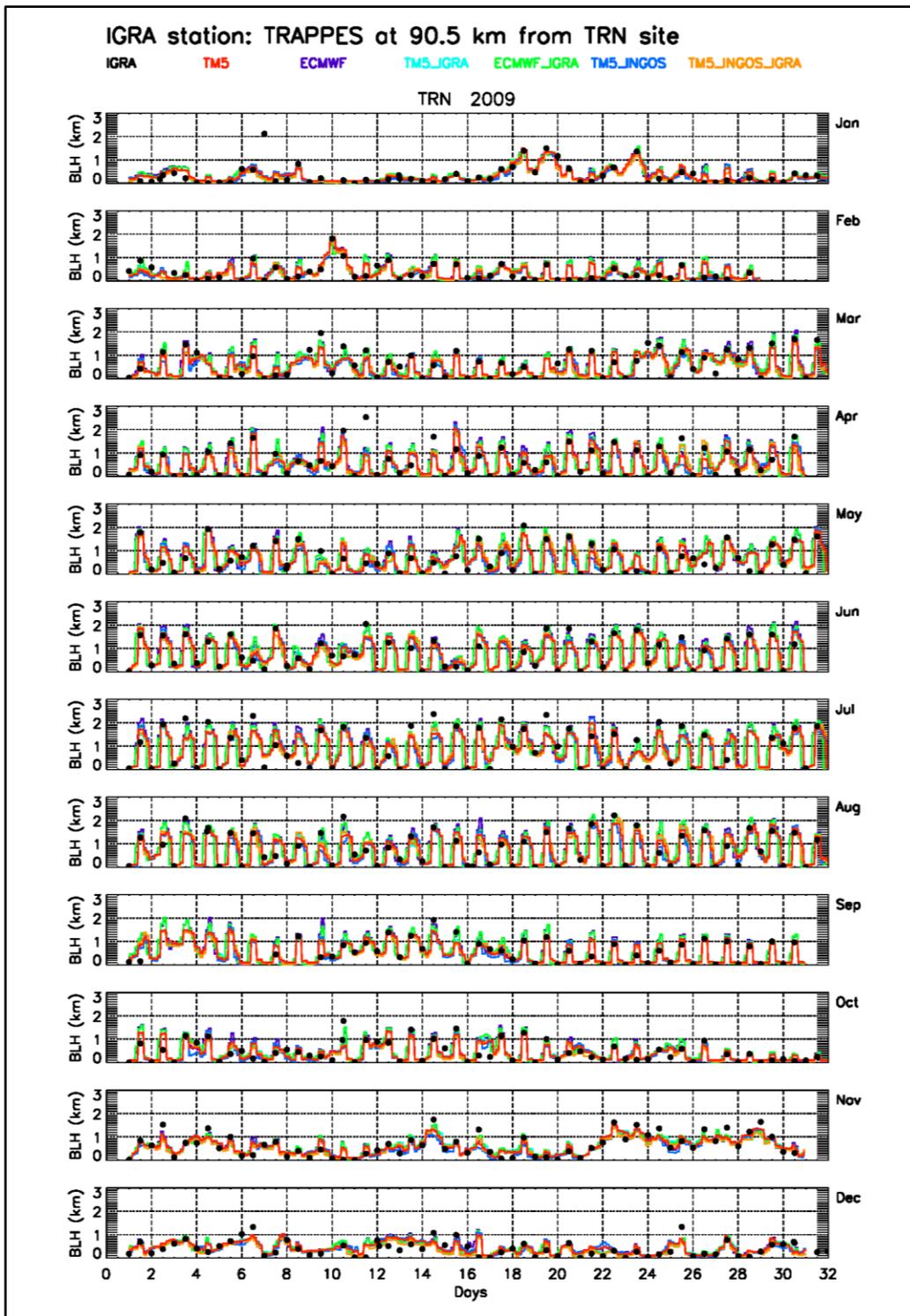
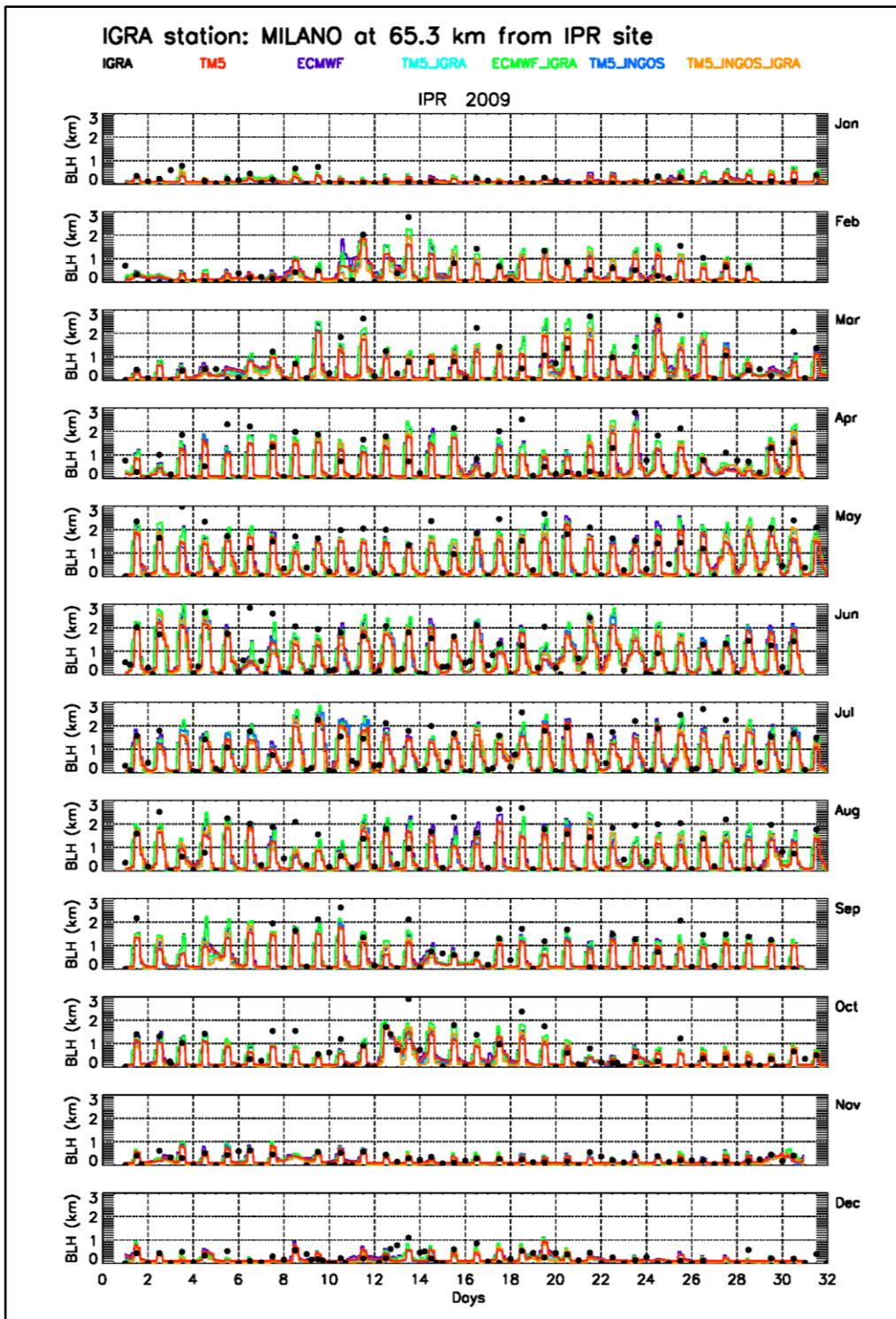
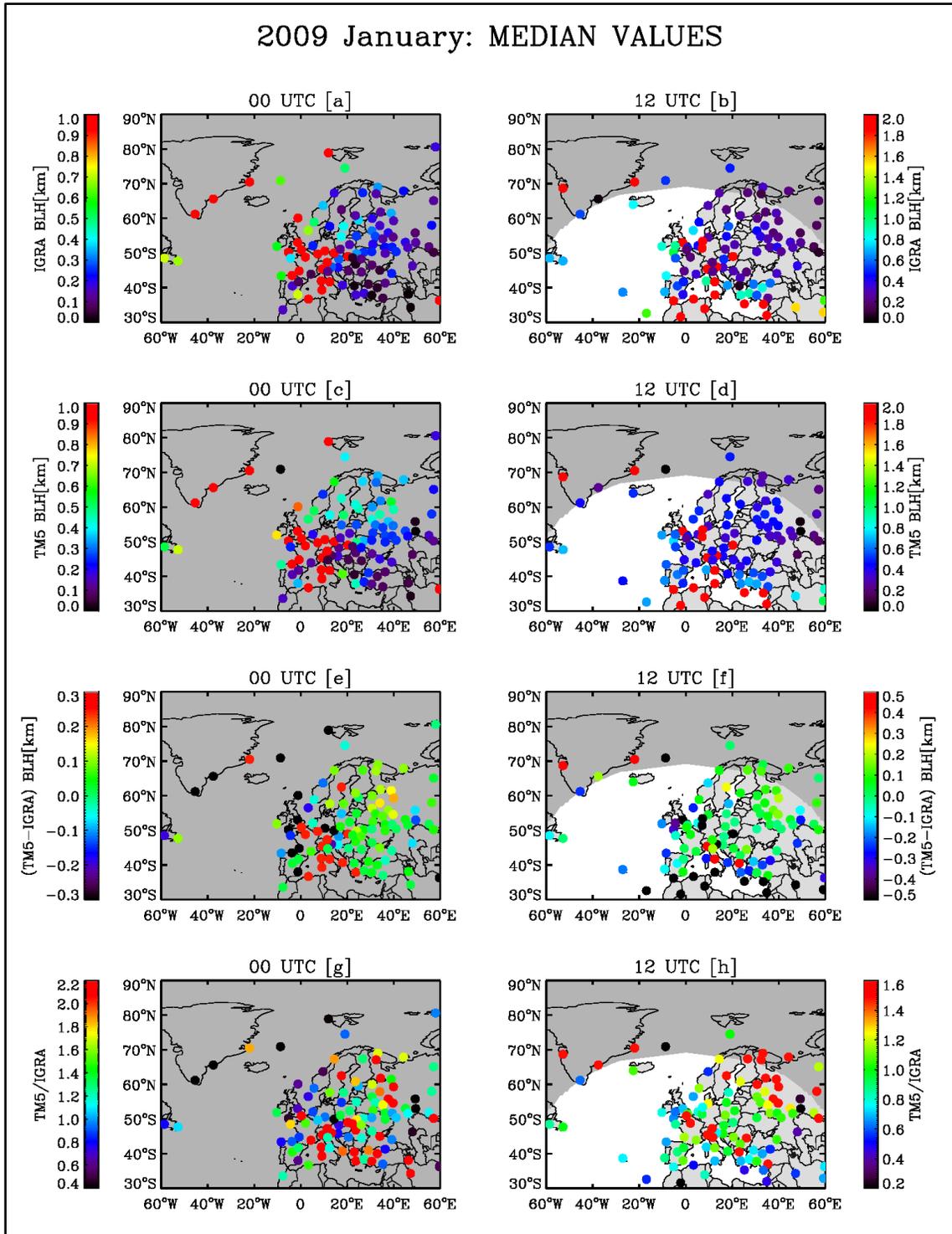


Figure S10: As Figure S2, but for the InGOS station Trainou (TRN or TR4)



**Figure S11:** As Figure S2, but for the InGOS station Ispra (IPR)



**Figure S12:** Median values of IGRA and TM5 boundary layer heights [a, b, c, and d] together with their differences (TM5-IGRA) [e and f] for January 2009 are shown. The median values of the ratios between TM5 and IGRA boundary layer heights (TM5/IGRA) [g and h] are also displayed. Left: 00UTC; right: 12 UTC.

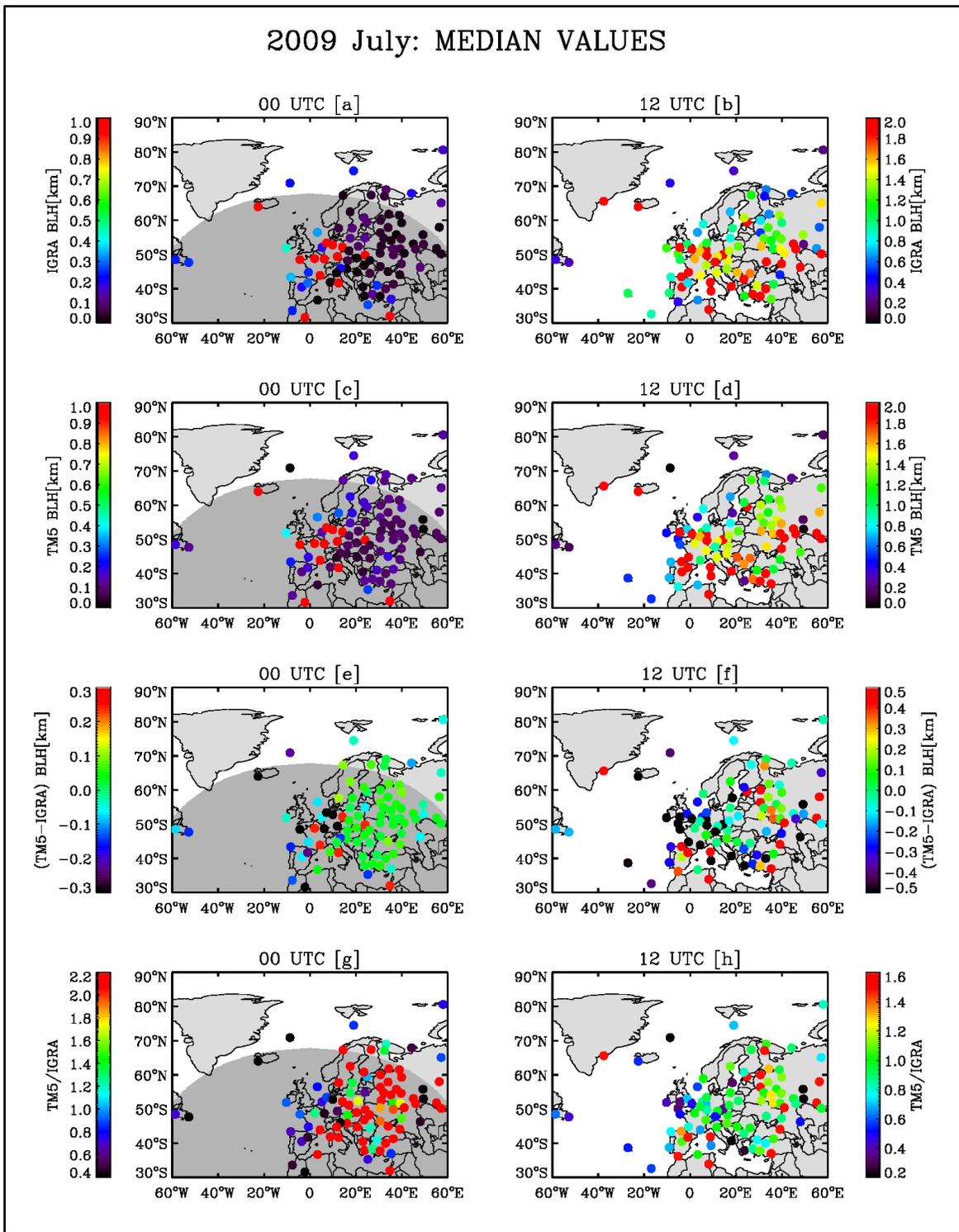


Figure S13: As for Figure S12, but for July 2009

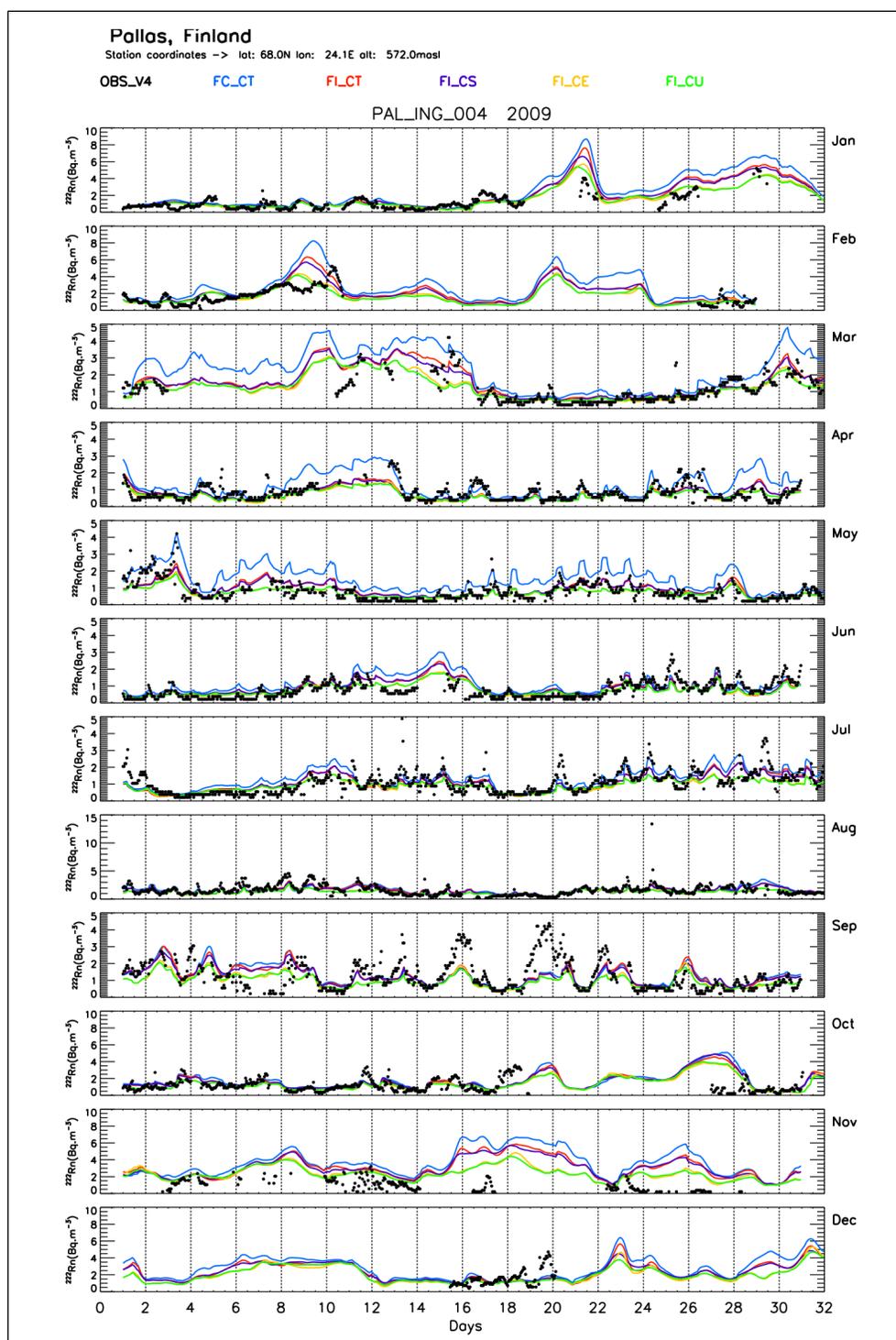
## Time series of $^{222}\text{Rn}$ activity concentrations

We simulate  $^{222}\text{Rn}$  concentrations using either the InGOS  $^{222}\text{Rn}$  flux map, or constant  $^{222}\text{Rn}$  fluxes (see Section 3.3). Furthermore, we apply four different convection schemes in the TM5 model (for the InGOS  $^{222}\text{Rn}$  flux map based simulations only). These different simulations are labelled by the following acronyms:

- FC\_CT: constant  $^{222}\text{Rn}$  fluxes, and default convection scheme in TM5 based on *Tiedtke* [1989]
- FI\_CT: InGOS  $^{222}\text{Rn}$  flux map, and default convection
- FI\_CS: InGOS  $^{222}\text{Rn}$  flux map and revised slopes scheme (see Section 3.1 of the text)
- FI\_CE: InGOS  $^{222}\text{Rn}$  flux map and the updated convection scheme based on ECMWF reanalyses (see Section 3.1 of the text)
- FI\_CU: InGOS  $^{222}\text{Rn}$  flux map, updated treatment of slopes and updated convection scheme based on ECMWF

For other details see Section 3.4 of the text

We show hereafter only the data for year 2009



**Figure S14:** The time series of the observed and simulated radon activity concentrations at Palas [PAL] for 2009. The observed radon concentrations are in big dots (●). The model simulations are obtained from constant emissions (solid blue line; FC\_CT) and four from INGOS emissions (solid red line: FI\_CT, solid violet line: FI\_CS, solid orange line: FI\_CE, and solid green line: FI\_CU). The different acronyms are defined in the previous page or in Section 3.4 of the text

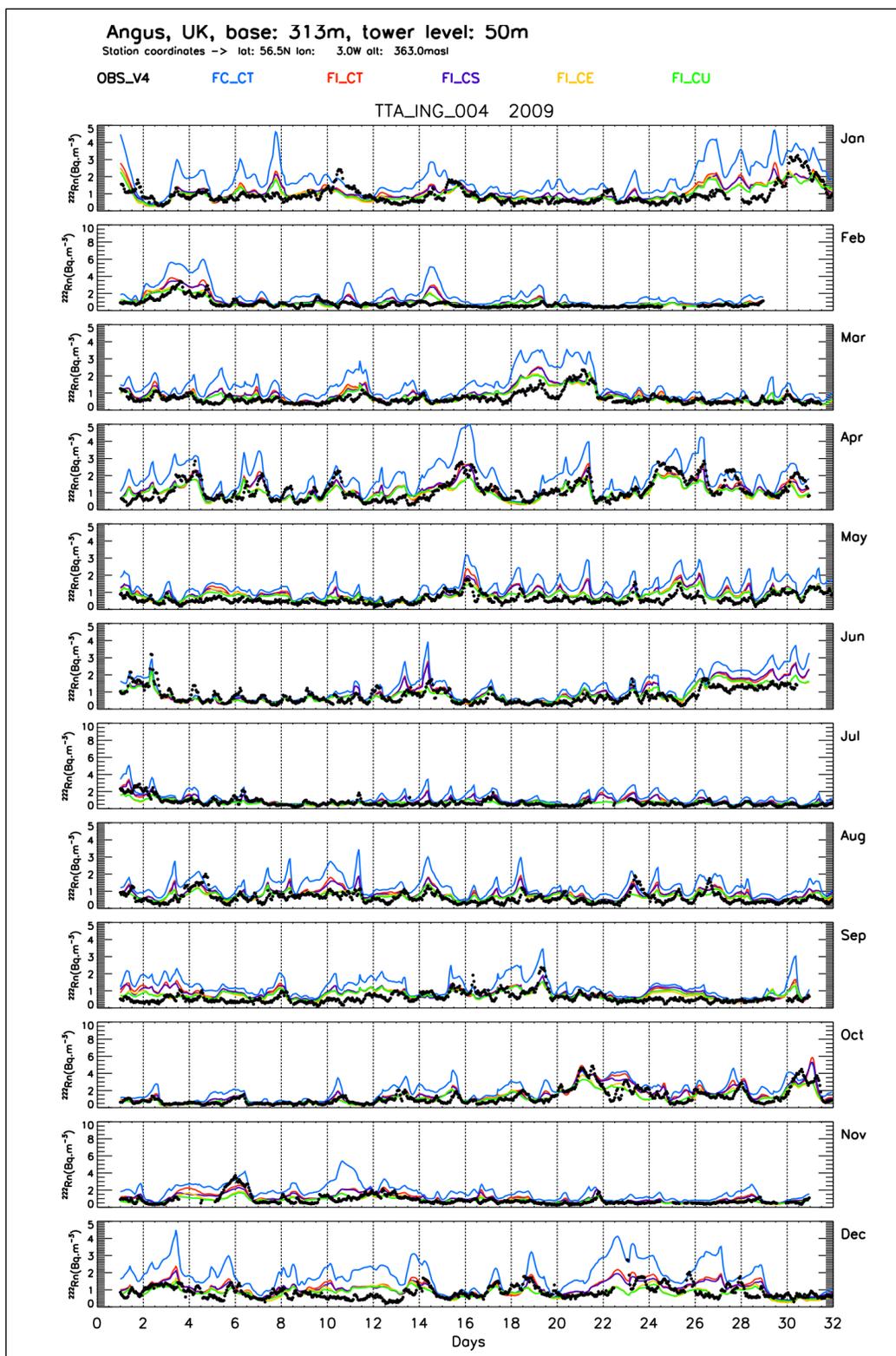
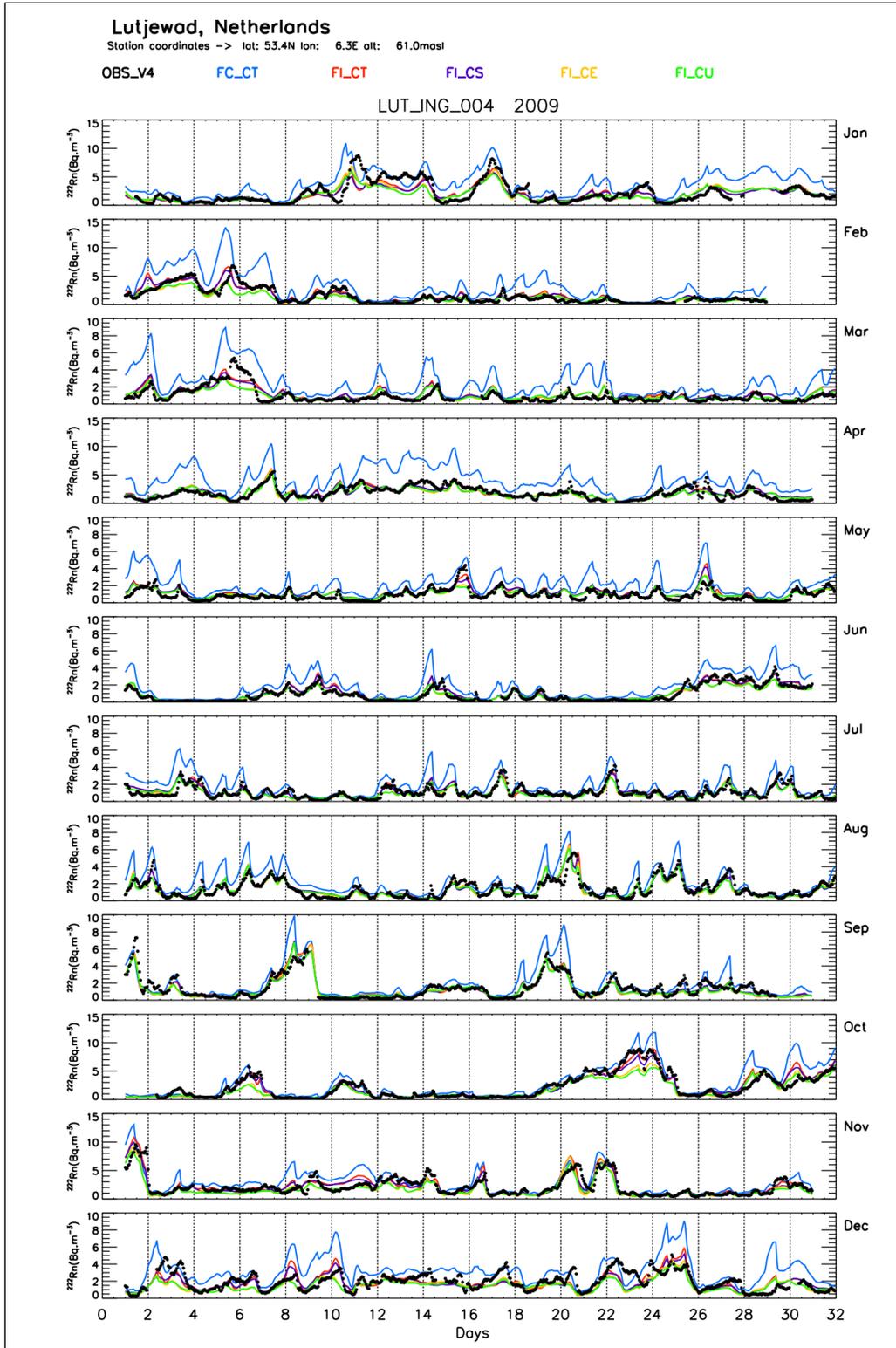


Figure S15: As Figure S14, but at Angus (TTA)



**Figure S16:** As Figure S14, but at Lutjewad (LUT)

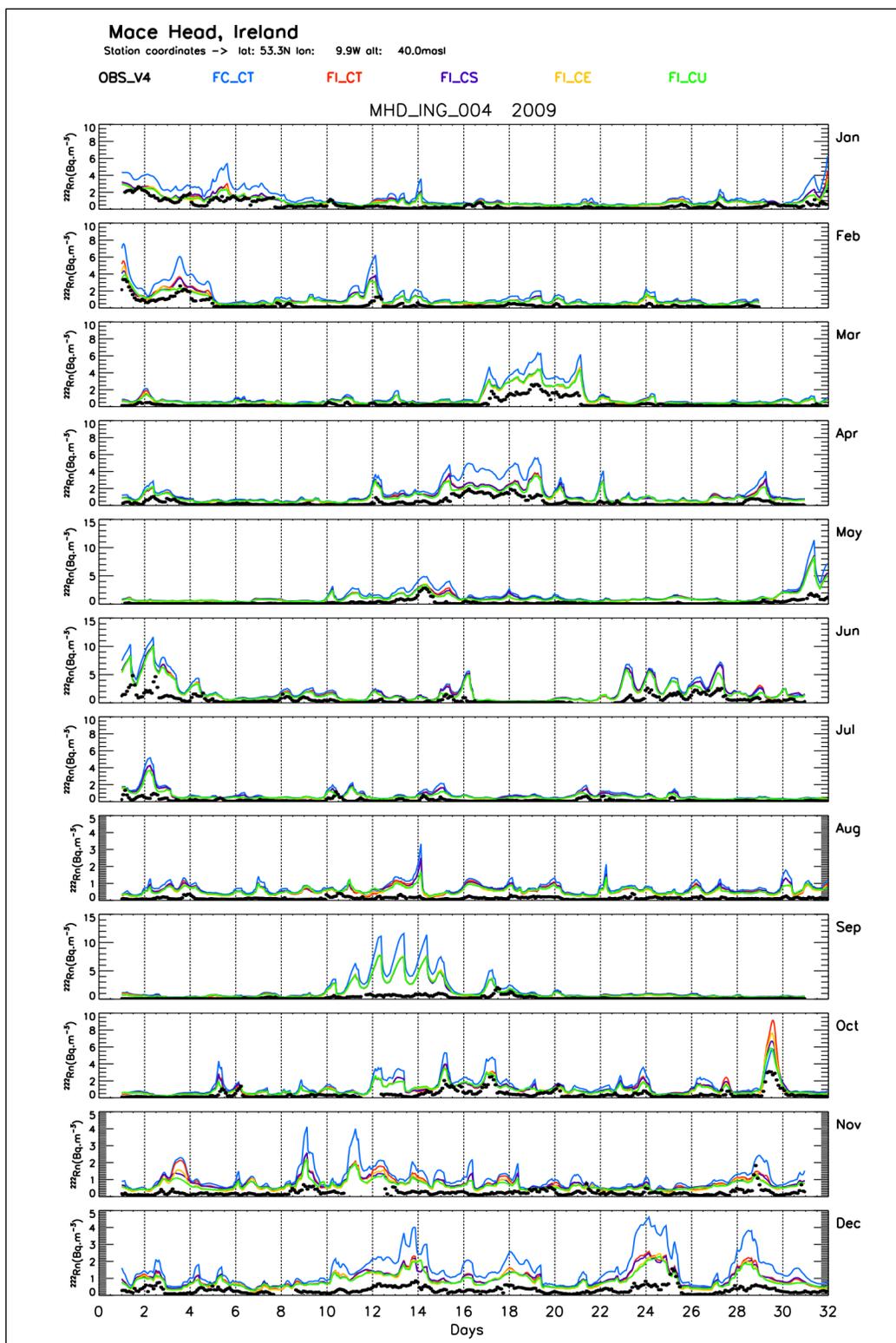
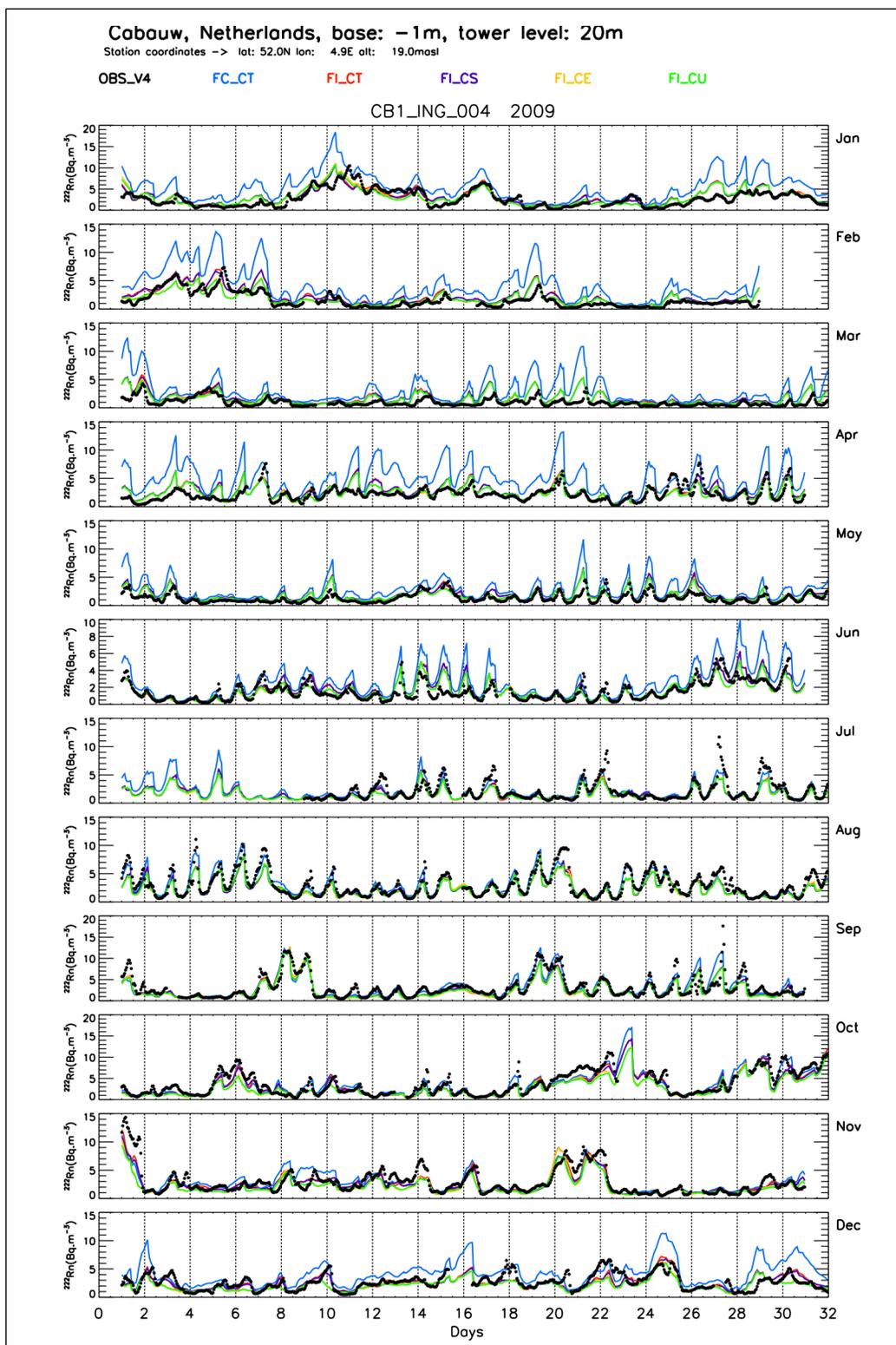
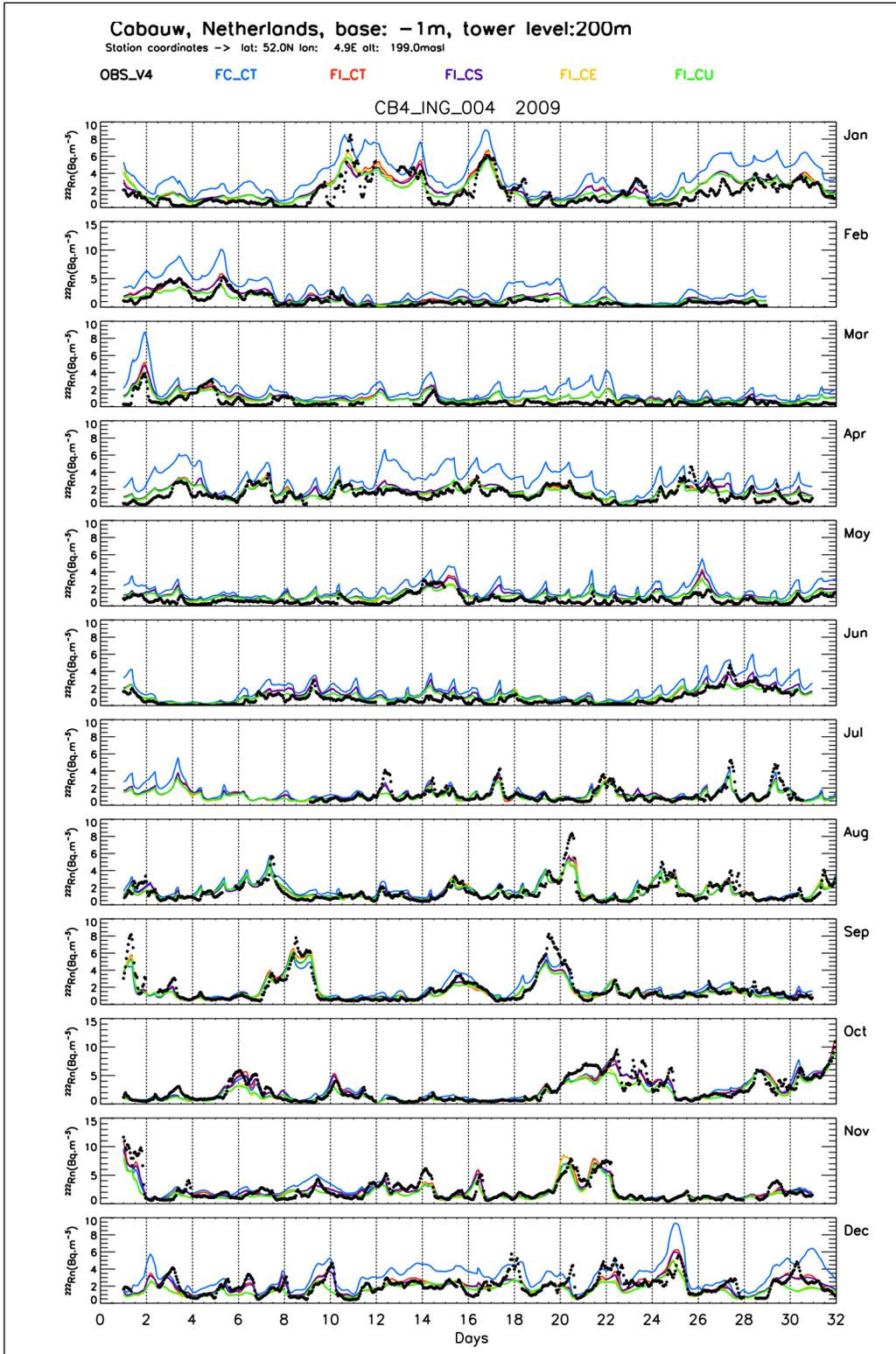


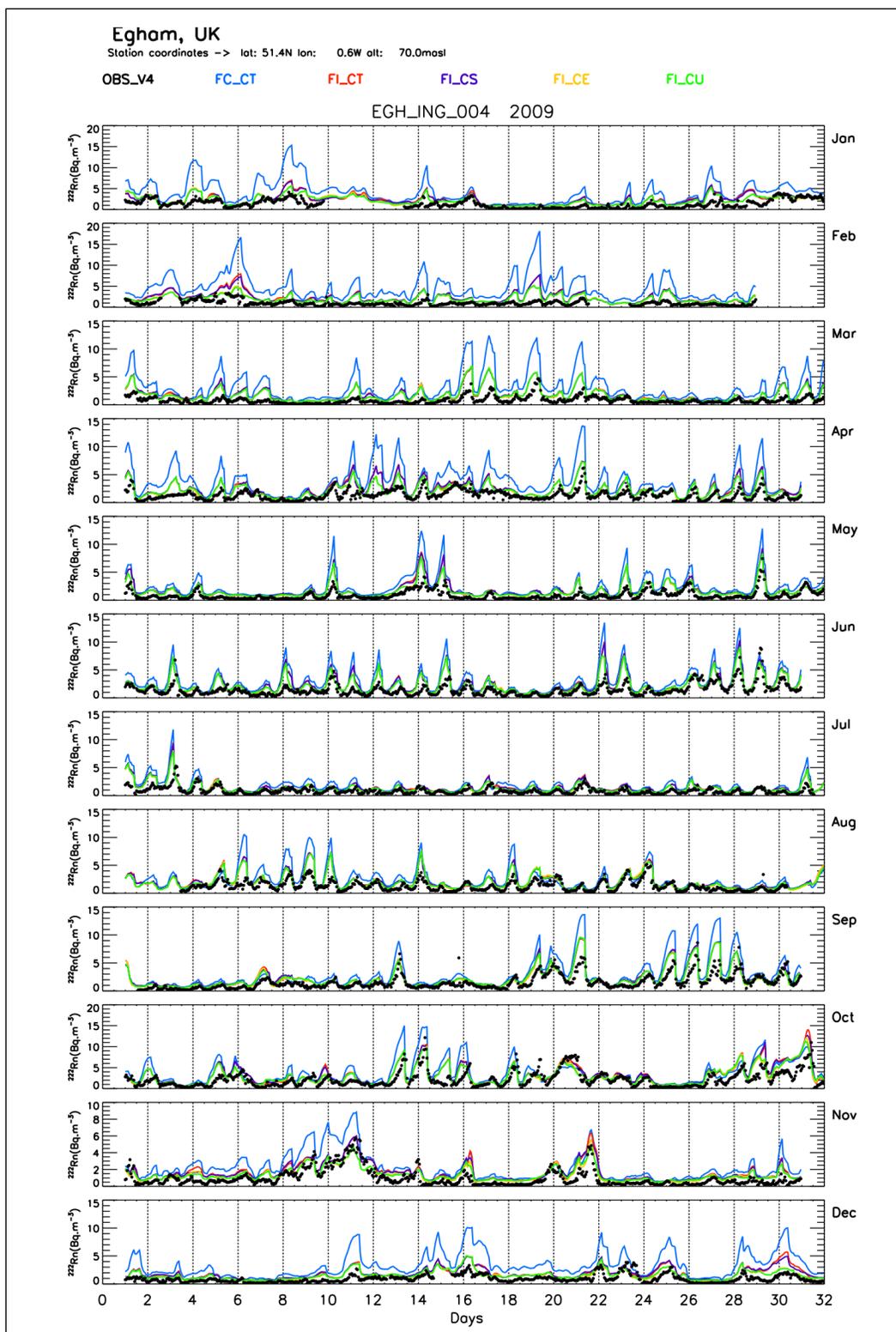
Figure S17: As Figure S14, but at Mace Head (MHD)



**Figure S18:** As Figure S14, but for Cabauw at 20 m [CB1]



**Figure S19:** As Figure S14, but for Cabauw at 200 m height [CB4]



**Figure S20:** As Figure S14, but at Egham (EGH)

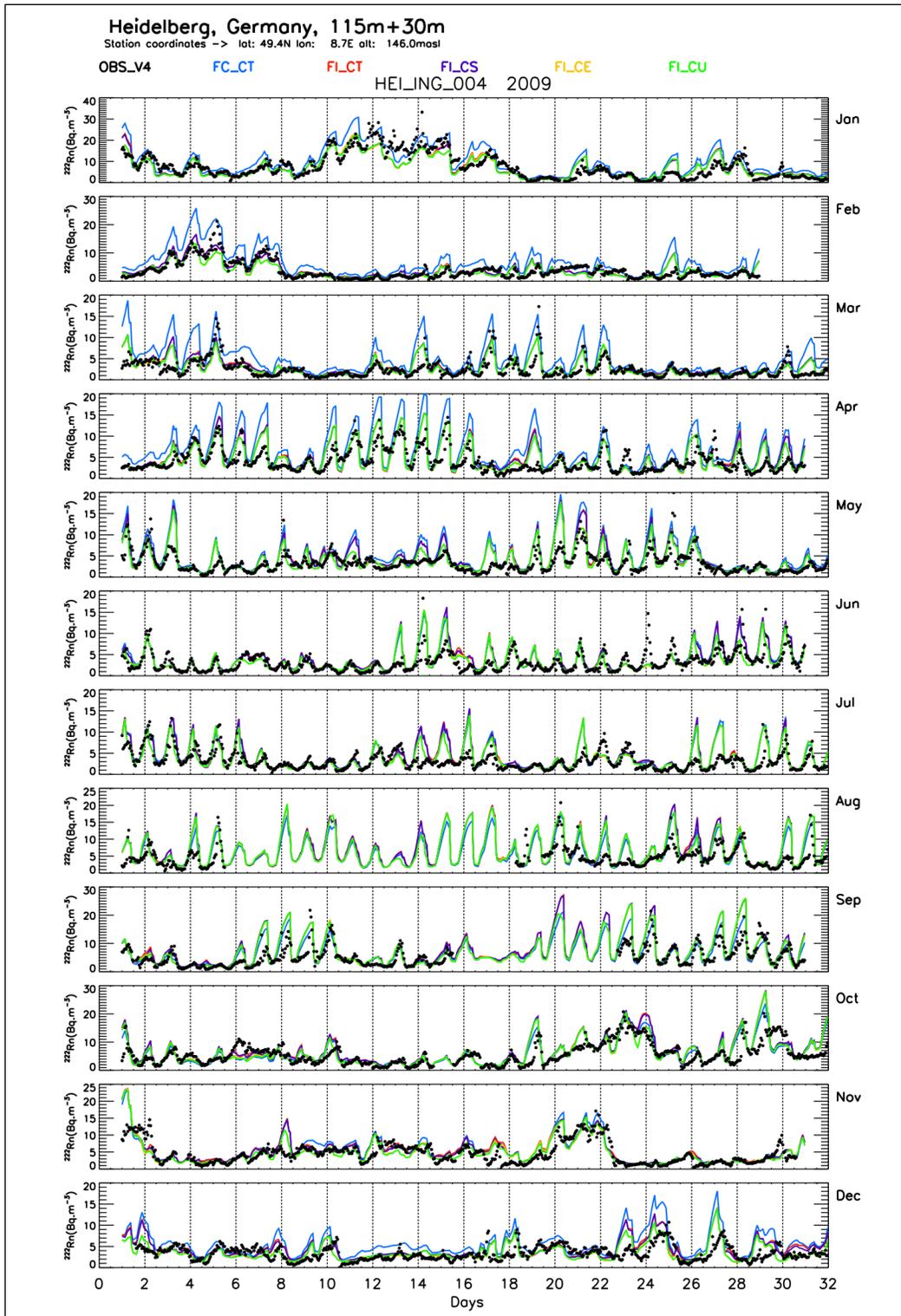
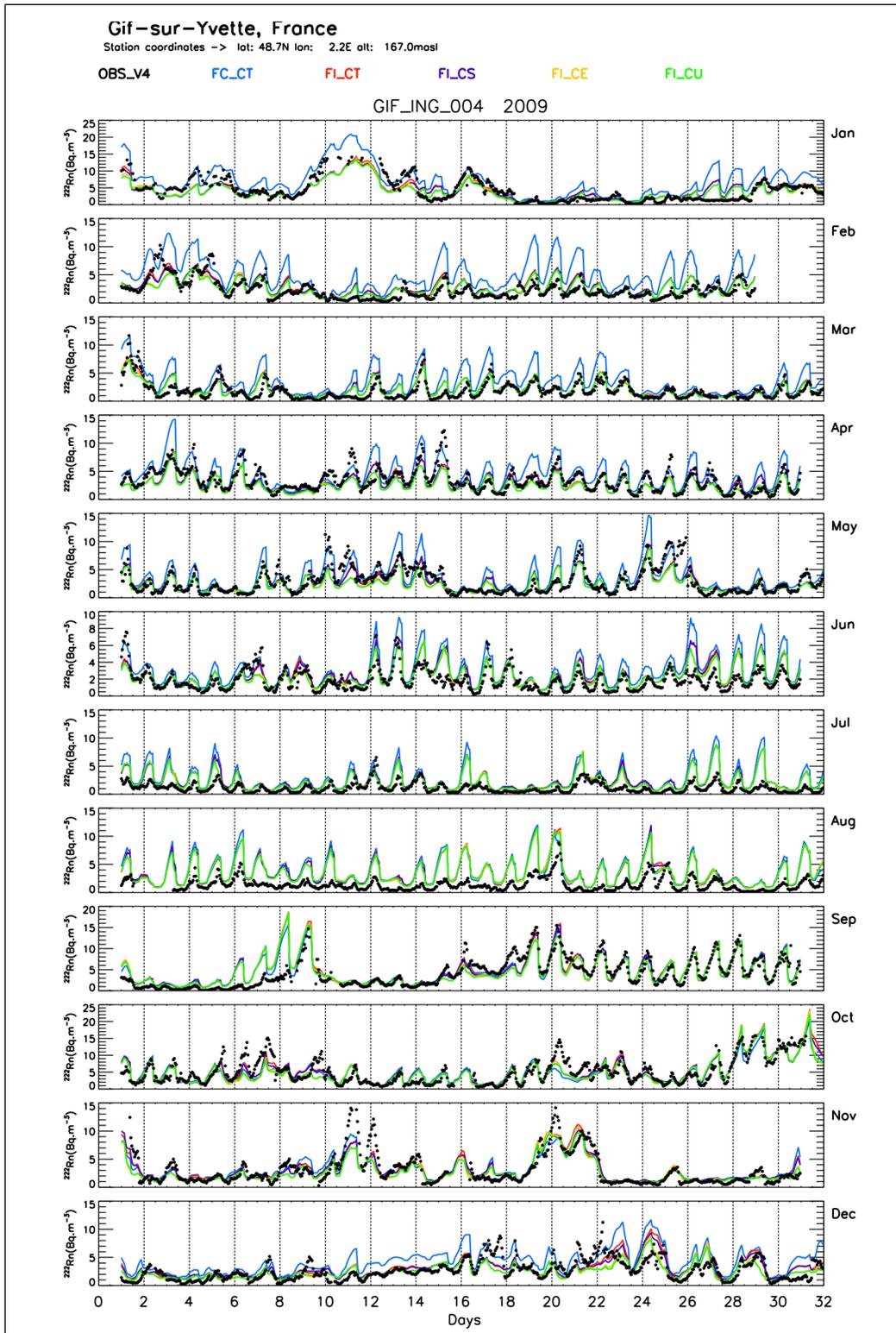


Figure S21: As Figure S14, but at Heidelberg (HEI)



**Figure S22:** As Figure S14, but at Gif sur Yvette (GIF)

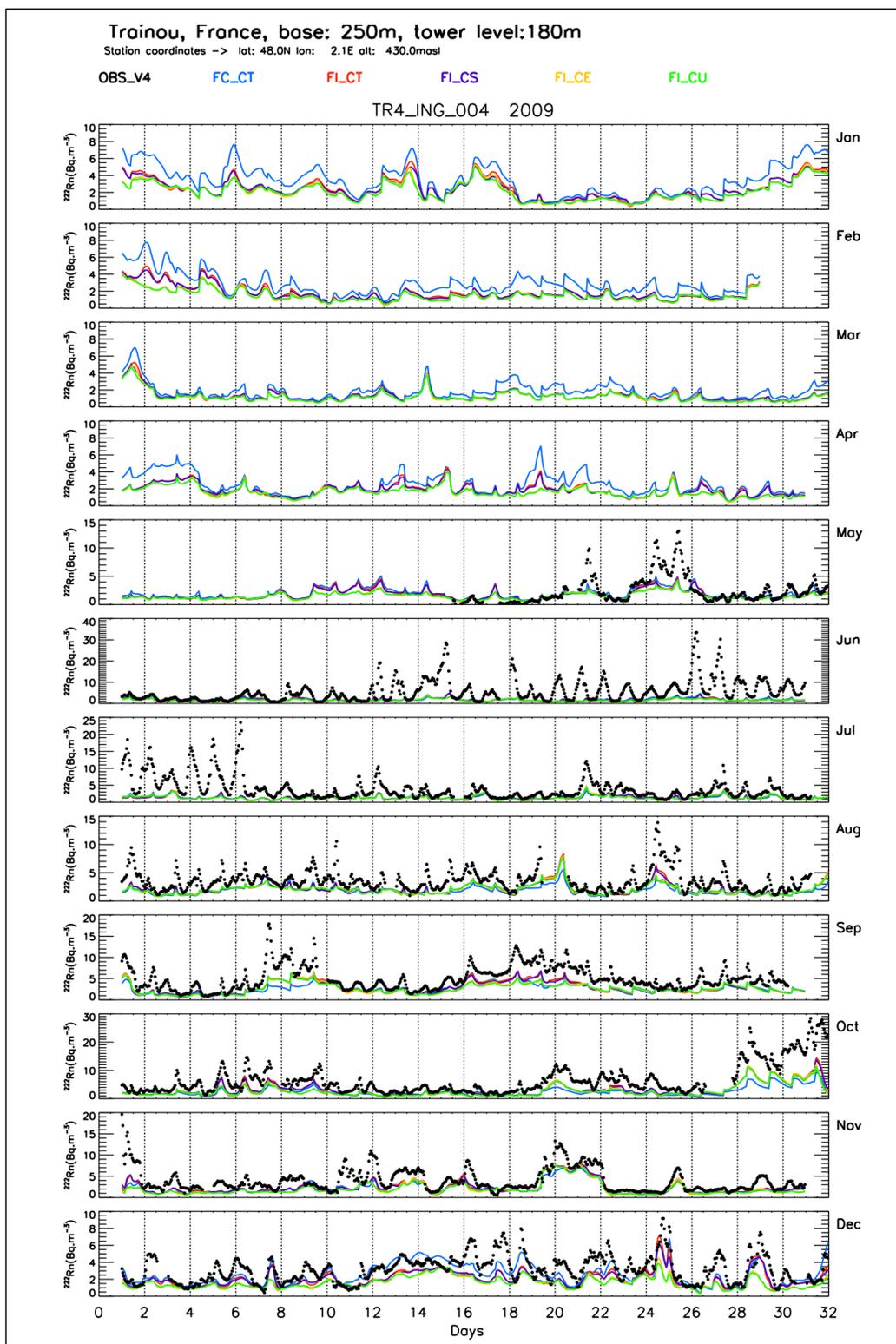
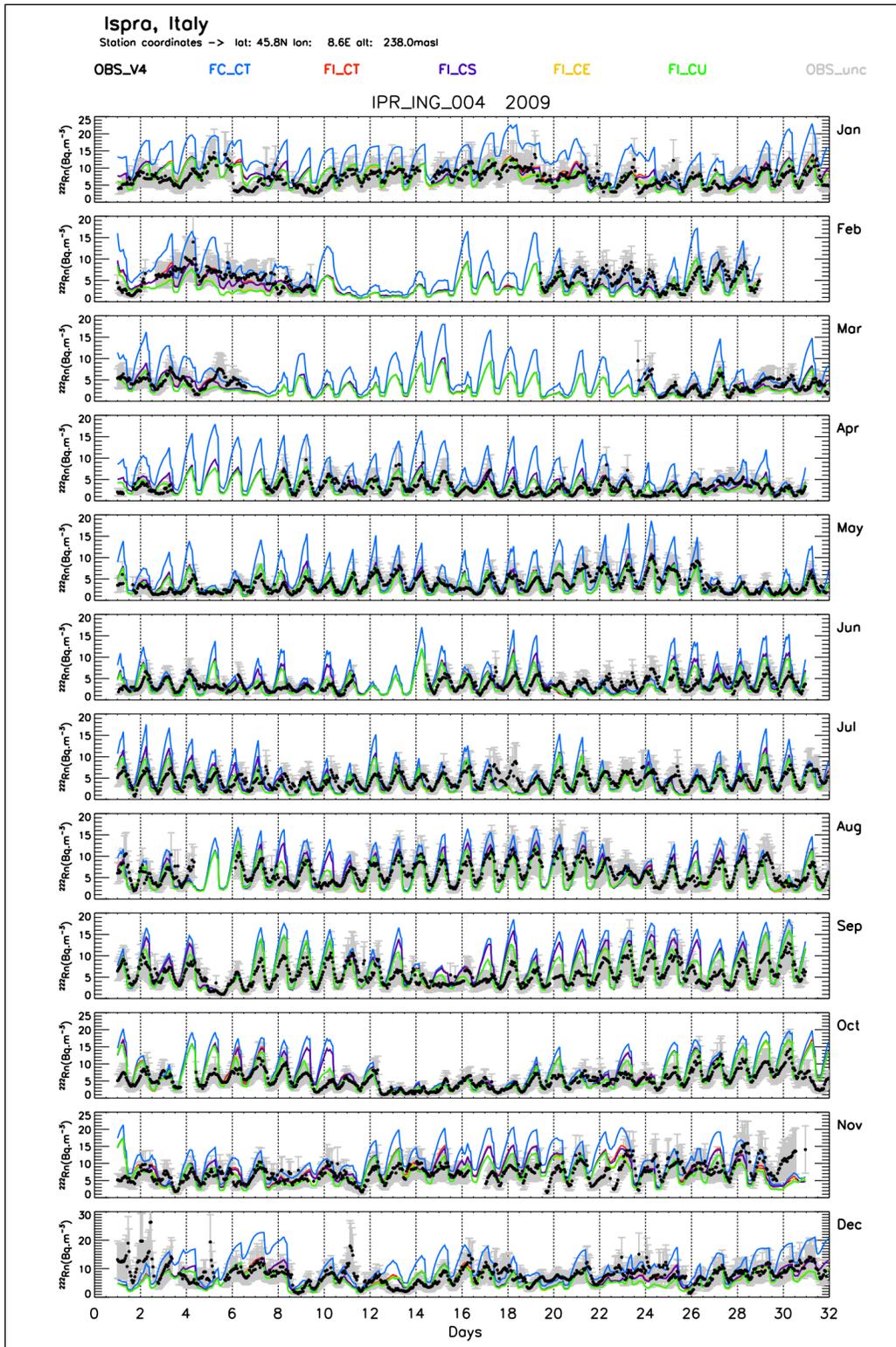


Figure S23: As Figure S14, but for Trainou at 180 m [TR4]



**Figure S24:** As Figure S14, but at Ispra (IPR). This figure includes also the uncertainties in the  $^{222}\text{Rn}$  activity concentrations (grey shaded area) from the wind-speed dependent correction of the measurements ('normalization' to 15m inlet height; see section 2.2).