Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-238-AC1, 2018

Geoscientific Model Development Discussions

Interactive comment on "A spatial evaluation of high-resolution wind fields from empirical and dynamical modeling in hilly and mountainous terrain" *by* Christoph Schlager et al.

Christoph Schlager et al.

christoph.schlager@uni-graz.at

Received and published: 22 December 2018

We thank Referee #1 for the valuable and constructive feedback to our manuscript. We carefully considered all comments and made due effort to account for the concerns expressed, and we think it really helped improving the readability and quality of the text and how we convey the findings. We also would like to thank the referee for the care related to remaining typos and spelling mistakes.

Responses to Major comments:

1) The term "dynamical modelling" is repeated through the manuscript, and even in the

title. I think this expression it is not very common in the Regional Climate Modelling literature. This term seems to combine two more common expressions: "regional climate modelling" and "dynamical downscaling". Both are used in the literature more or less interchangeably, but I think "dynamical modelling" is not generally used. The reason for this is that, technically, a Global Circulation Model is also dynamical modelling, but I'm sure the authors do not mean this type of model. Therefore, I would advise to stick to one of the two aforementioned alternatives.

Answer: Thank you for this hint, we will carefully recheck our usage of the general term "dynamical modeling" (in the sense of empirical modeling vs. dynamical modeling) and will replace it by a more specific term where we see needed, such as "regional climate modeling" or "dynamical high-resolution climate modeling" and so.

2) The authors refer to two former publications (Schlanger et al. 2017, 2018) where the WPG seems to be further described. I acknowledge that I didn't read these publications, but it is not clear to me what this article improves or how it complements the formers. I think putting emphasis somewhere in the introduction on what new issues/questions this new article tries to address, compared to the formers, would help to frame this work and to better justify why it is necessary.

Answer: OK, we agree that the introduction about the ongoing work described in this article in relation to the two former articles gains from more context. Hence we will insert a relevant paragraph in the introduction to clarify how this article complements the formers, as follows.

Inserted paragraph (page 2, after line 26): "... For both networks, diagnostic wind fields at a high resolution grid of 100 m $\times$ 100 m are generated by a weather diagnostic application, the WegenerNet Wind Product Generator (WPG). Schlager et al. (2017) introduced the WPG and its performance evaluation for the WegenerNet FBR, which was then advanced by Schlager et al. (2018) to the WegenerNet JBT region and a longer-term evaluation in both the FBR and JBT regions. Jointly these studies established the level of quality of the empirical WPG wind fields. In this study, we now make use of these empirical high-resolution wind fields as reference data in order to intercompare them with empirical-dynamical wind field analysis data and . . .”

3) The INCA dataset assimilates observations. Then this dataset is compared/validated with respect to the WPG, which are also observations. Are they the same? Are WPG observations assimilated to produce INCA? I assume not, as otherwise there would be an important circularity issue.

Answer: No, INCA does not assimilated WegenerNet data; and indeed we intentionally keep them independent just to avoid such circularity issues, yes. Having said this and based on rechecking our related description we agree, though, that the description of which station measurements are used in which model is a bit vague. We therefore will improve this a bit to make clear that observations from ZAMG stations (the ones used in the INCA analyses) are not used as model input for the WPG and vice versa that the INCA just uses observations from ZAMG stations, but not from WegenerNet stations as model input. We will do this as follows.

Improved description in Section 2.2 (starting on page 4, line 31): “. . . again with a maximum latency of about 1-2 hours. To keep the meteorological input data of the WPG independent from the data pertaining to the other operational station networks, observations from the ZAMG stations (violet stars in Fig.1a and Fig.1b) and other external stations are not used as WPG input. For the WegenerNet FBR, the gridded wind fields. . .”

Improved description in 2.3 INCA and COSMO-CLM data (starting on page 5, line 18): “In the WegenerNet FBR area, INCA assimilates observations from the ZAMG Feldbach and Bad Gleichenberg station (violet stars in Fig.1b) to the NWP’s first guess, and within the WegenerNet JBT region, observations from the ZAMG Admont station are used. However, data from WegenerNet FBR and JBT stations are not used in INCA data assimilation and hence the WPG fields can be used for independent evaluation.”

4) I’m not sure what is meant by a "wind event". I understand that the criteria in Table 2 is applied on an hourly basis, right? Are then the events hourly-based, i.e. a given hour might be included as a calm event, while the next one might be included as strong? Or do the authors select for instance the whole day when at least a single hour within the day meet the criteria? Another way of posing this question is, are there as many events as hours within each period?

Answer: Thanks for this comment which led us to notice that some further information regarding data selection would be helpful. To implement this, we will therefore modify some text passages in section 3.1 (Events for wind field evaluation). For example, the number of events shown in Table 2 corresponds to hourly events, but the data selection method for thermally induced wind events differs from the method for selecting strong wind events. In general, the data selection for thermally induced wind events is based on daytime and nighttime mean values as indicated by subscripts (dm, nm) in Table 2, which are also explained in the footnotes of this table. If a day is selected as autochthonous day, all 24 hours from this day are used for evaluating thermally induced wind events, i.e., such a day contributions 24 hourly wind events. In case of strong wind events, we compared the hourly mean values from the datasets with the hourly thresholds defined in Table 2 (hm subscripts). If the hourly mean value is larger than the defined threshold, this event is used for evaluating strong wind speeds. We will explicitly include a further line of footnote to Table 2, making clear that the “Number of events” column denotes hourly wind events as the basis for the statistical analysis, and we will modify the text as follows.

Modifications (starting page 6, line 18): “For the estimation of autochthonous days, we compared the observed daytime mean values of wind speed (v) and relative humidity (rh) as well as the nighttime mean values of net radiation (Qn) from selected stations with the respective thresholds. . . . . .If all criteria are fulfilled for a given day, the data from the entire day are added to the thermally induced wind events dataset, leading to 24 hourly events (i.e., 24 hourly-mean wind speed values).” And starting page 6, line 31: “The strong wind events, caused by synoptic weather conditions such as cyclones and frontal system at larger scale, are selected on an hourly basis, by comparing hourly-mean values from the gridded reference datasets with defined minimum wind speeds (Table 2, v for strong wind speed cases). If the hourly-mean value of this reference dataset is larger than the defined minimum wind speed, the data of this reference dataset and the corresponding model dataset are used as part of the hourly event data for evaluating strong wind events.”

5) Another detail I could not understand is how the WFSS is calculated for different spatial scales. Is the data interpolated onto successive grids with coarser resolution?

Answer: The calculation for different spatial scales is performed for defined neighborhood sizes, which have to be odd integers. A neighborhood size (n) defines the side length of a square, which is moved as sliding window over the dataset (e.g., n=5 corresponds to a neighborhoods size of 500 m at a spatial resolution of 100 m, and the square hence contains 25 grid points). We calculated the WFSS values for neighborhood sizes from n=1 to n=2N - 1, where N is the number of grid points of the largest domain size from the WegenerNet FBR or the WegenerNet JBT. The maximum domain size of 2N - 1 was used to ensure that the sliding window is large enough to always encompass the whole domain at every position - as a consequence, the fractions inside the domain are guaranteed to be the same at all locations within the domain and further enlarging the neighborhood will not change the WFSS value.

We will add additional text to section 3.2 (Statistical evaluation methods) to explain the calculation for different neighborhood sizes, as follows.

Additional text (starting at page 7 line 14) “. . .for different neighborhood sizes. The neighborhood sizes must be defined as odd integers. A neighborhood size of n defines a square consisting of n x n grid points, i.e., it denotes the side length of the square (e.g., for n=5 the square contains 25 grid points). These squares of defined neighborhood sizes are moved as sliding windows over the datasets, centered successively at

each grid point, whereby the area outside the domain is assumed to contain no wind class. In terms of FSS value, . . .”

Responses to Minor comments (in order of appearance):

1) The abstract is in my opinion longer than necessary. For instance, between lines 5 and 10 a great amount of details are given about the datasets. This level of detail is overwhelming at this early point of the paper, and distracts the reader from the main conclusions of the manuscript.

Answer: Ok, we agree that the abstract gives too detailed information. We will therefore reduce the level of detail regarding the explanation of the two meteorological station networks at the beginning and also somewhat the discussion of the results at the end of the abstract.

2) Pag 2, Line 9: course-resolution → coarse-resolution. Answer: Ok, will be done

3) Pag 2, Line 15: "data fusion". I think a more precise term is "data assimilation" or "assimilation of observations".

Answer: Ok, we will use “data assimilation” instead of “data fusion”

4) Pag 2, Line 19: "dynamical regional climate models" → "regional climate models".

Answer: Ok, will be done

5) Pag 3, Lines 3-8 These two paragraphs read as a summary of the methodology. I do not think this is necessary in the introduction.

Answer: Thank you for this hint. We agree, that this information is also given in section “3 Evaluation events and methods” and we will therefore remove these two paragraphs from the introduction.

6) Pag 3, Line 10: I was not aware of the concept "two penalty problem". Therefore I was puzzled to read this without either a reference or a couple of lines that briefly

summarise what is the deal with this. It is explained later, so I would advise to bring those explanations already here.

Answer: Thank you for this hint. We will move the explanation regarding the “double penalty” to this paragraph to immediately explain this kind of penalty. 7) Pag 4, Line 6: "eleven"→11 (for consistency reasons with the way this is reported for FBR)

Answer: Ok, will be done

8) Pag 4: Lines 20-26: Is it really necessary this amount of detail about how the data about temperature and humidity is produced for this system, given that these fields are not used in the manuscript?

Answer: Thank you for this hint, we agree that the gridded fields of temperature, precipitation, and relative humidity are not so relevant for this manuscript. We therefore will remove the (too) specific description parts about the lapse rate and the different interpolation methods for the generation of these fields.

9) Pag 5, Lines 15-16: "Therefore the output shows errors in regions with low station density" The model resolution does not imply that there are larger errors in areas with low station density. Why would it be the case? The validation is more difficult, but it could be that the model does a good job. We just don’t know.

Answer: Thanks, we agree that this statement is not correct at this position of the text. We therefore will remove this sentence and modify the text in this paragraph to ensure, that the statement is related to the INCA analysis algorithm and not to the RCM’s first guess. We will modify as follows.

Modified Text (page 5, line 16): “The INCA wind fields have already been evaluated for a moderately hilly region in the north of Austria (47.8° N–49° N, 13.8° E–17° E), where the wind analyses show significantly higher errors compared to the statistical results from other meteorological variables. These higher errors mainly root in the limited representativeness of station data, as well as on the low station density, which can be

only partly compensated by INCA’s analysis algorithms (Haiden et al., 2011).”

10. Pag 5, Line 22: The number of vertical levels in the RCM (not only the driving dataset) is an important parameter worth to mention.

Answer: Ok, we will add additional text regarding the number of vertical levels. The COSMO-CLM simulations are provided for 40 vertical levels. The first level is simulated for 10 m above ground and the last level corresponds to the 100 hPa level, whereby the vertical resolution is higher for the boundary layer and decreases towards to the top level.

11) Pag 7, Line 23: the units (m s-1) should not be italic. This applies to several locations through the manuscript. Please review them.

Answer: Ok, will be corrected.

12) Pag 8, Line 15 says that wind speeds are systematically underestimated. This is curious, as normally models tend to overestimate wind speed. Indeed, in the conclusions (Page 13, Line 19) this is noted when it is stated that wind speed are overestimated in both types of events. Isn’t this contradictory? Please clarify the details.

Answer: Thank for noticing this. The statement “systematically underestimated” is not fully correct in the context of what we try to address in the corresponding section (Pag 9, not Pag 8). In this section we are explaining the behavior of the WFSS for selected wind events and not for event-averaged statistical results. Therefore, the underestimation by the COSMO-CLM model explained in the text refers to a single event. We will correct the corresponding sentence as follows. Modified Text (page 9, line 13): “This discrepancy leads to the smallest WFSS values at all neighborhood sizes for this region (Fig. 2e, COSMOvsWN_strong_FBR) and indicates that the dynamically modeled COSMO wind speeds are underestimated relative to the empirically diagnosed wind speeds for this hourly event.”

13) Page 10, Line 21: "fundamentally able". Do the authors mean "unable"?

Answer: Yes, we will change it to “unable”

14) A bottleneck of WFSS is that it does not allow to disentangle if low skill is driven by problems with wind speed or direction. However in Pag 10, from lines 29, this is somehow solved, and low skill is attributed to errors to these two variables separately. But it is not obvious how these conclusions can be drawn from the shown figures. Is this based on an analysis that is not shown in the manuscript?

Answer: It has to be noted that the WFSS can also be used to separately evaluate the two wind components, for example by classifying the datasets just based on wind direction. In general, the definition of the classes should reflect what a user wants to verify. We used the advantage of the WFSS and evaluated wind speed and wind direction in a combined way. Regarding the separate evaluation of both variables in relation to Fig. 3, we agree that this conclusion cannot be drawn by simply interpreting this figure and that additional information is needed. The behavior of the influence of wind speed or wind direction on the WFSS is indicated by the results of error measures additionally calculated by traditional statistical methods. These are summarized in Table 4 and the generated mean wind speed bias distribution map, illustrated in Fig. 4. Furthermore, we visually interpreted the windroses for most of the events (the windroses for all events are not shown in this manuscript, Fig. 2 just shows windroses for selected events, for good illustration). To make clear of how we draw this conclusion, we will modify the corresponding text passages and refer there to the results calculated by traditional methods. The spatial displacement and the biases for the INCAvsWN_therm_JBT case are mainly caused by the differences in wind directions for these thermally induced wind events, indicated by the large mean absolute error of wind direction (MAEdir) (Table 4). Modified text is as follows.

Modified text (page 10, line 32): “The INCAvsWN_strong_JBT case shows the lowest values at all neighborhood sizes, but this time caused by the differences in the wind speed categories. These low values are caused by the INCA-analyzed wind speeds, which, in case of strong winds, are overestimated in the summit regions and underestimated in the valley regions. Slightly overestimated WegenerNet wind speeds in the Enns valley are somewhat reinforcing the difference between the INCA and the WegenerNet wind speeds. These differences in wind speed especially in the valley and the summit regions become obvious from Fig. 4c and 4d and are discussed in further detail below.”

15) Page 12, Line 21: where→were.

Answer: Ok, will be done

16) The conclusions are overly long. They review every single detail of the results and after reading them is not obvious what are the take-home messages. I advise to summarise the conclusions to leave the most important and general conclusions, those that can be exported to other studies/regions

Answer: Thank you for your advice, we agree that the conclusion gives too detailed information, which especially applies to the discussion of the results. We will therefore summarize the explanation of the results and shortly discuss what’s relevant for ongoing next steps of work and other studies/regions.

17) This may seem as a tiny detail, but the fact that the panels in Fig. 1 do not follow the expected order (a, then b, finally c) puzzled me for a couple of minutes until I realised that FBR (labelled b, and firstly described in the text) is actually the last panel of the figure. Perhaps a trivial re-ordering of the panels following a more intuitive order might facilitate the reading.

Answer: Thank you, we agree that the panel sequence and the corresponding labeling is a bit confusing. We will therefore move the FBR panel to the top of Fig. 1 and label it with (a), and the JBT panel to the bottom and label it with (b). Furthermore, the overview in the middle of Fig. 1 will be denoted as middle panel in the text; the discuss-panels is (a) and (b) and so everything will be clear.

We thank Refereee #1 again very much for the valuable comments that will help to improve the manuscript.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-238, 2018.