Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2020-52-RC2, 2020

Interactive comment on "BAYWRF: a convection-resolving,present-day climatological atmospheric dataset for Bavaria" *by* Emily Collier and Thomas Mölg

Michael Warscher (Referee)

michael.warscher@uibk.ac.at

Received and published: 9 July 2020

In their manuscript "BAYWRF: a convection-resolving, present-day climatological atmospheric dataset for Bavaria", the authors present a new high-resolution RCM simulation using WRF and ERA5 reanalysis data as boundary condition. They evaluate the performance for the target region of Bavaria using station observations.

General Comments

The manuscript is very well written and of high technical and scientific quality. It fits very well in the scope of ESSD. However, I have several issues, questions, and suggestions

which at large could lead to major revisions. However, I understand that the manuscript is mainly an overview of the presented data and thereby, the amount and detail of the analyses and following content has to be limited at certain points. I would gladly leave the decision to the editors on how much of and at what detail level my suggestions should be addressed. The dataset they produced is generally very valuable for the scientific community, as well as for many users in different sectors.

The authors have chosen a single year as specific validation period. In addition, they point out in L. 101 that it is not an average year in terms of seasonal climatology (record heatwave in 2018). The chosen year might therefore not be a representative period for the RCM performance in other years. However, an extension of the evaluation period seems just limited by a missing run using the NO_NUDGE configuration. As the whole exercise is a historic / present-day reanalysis driven simulation effort, the NUDGE setup was run for the whole 30-year period anyways and would be available for additional validation years. I would highly recommend to add at least one additional year of validation (the more the better) to strengthen the results under different conditions. The validation could potentially even be done for the whole 30-years (just being limited by available observations and – by now - the missing NO_NUDGE for more than one year). I am also quite sure that an extension of the analysis would not limited by available observation data. If no additional simulations (NO_NUDGE) can be performed, the authors might think of some additional validation using the 30-year NUDGE run only.

While I really acknowledge the direct comparison to station data, the study would highly benefit from a comparison to gridded observation data sets such as REGNIE (1 km, https://www.dwd.de/DE/leistungen/regnie/regnie.html) or HYRAS (5 km, https://www.dwd.de/DE/leistungen/hyras/hyras.html). Besides the correct representation of single stations, the real benefit of such a computationally expensive high-resolution simulation might or should be – besides the reproduction of observed station data - the resolving of spatial distributions.

The authors point out, that they use a convection-permitting resolution of 1.5 km. However, the topic of simulating convective precipitation is not referred to again in the manuscript. This is still a very important and relevant topic, and the presented data would be ideally suited to look into this. Several questions arise and could easily be tackled. E.g. the authors show an underestimation of precipitation at some point. Could this be explained by the 1.5 km still being too coarse to resolve all or enough convective events? Are the results of the KF parameterization in the 7.5 km (D1) similar or totally different? Are sub-daily precipitation dynamics captured? Some of these questions could quite easily be investigated by comparing your results of D1 (convection parameterized) and D2 (convection resolved) to gridded precipitation products such as REGNIE and maybe even to station data.

This leads to another question regarding the resolution. While I do not question the validity and satisfying performance of the presented simulation, I would be very glad to see more about the added value of such a high resolution. This could be done by a comparison of the performance between the results of Domain D1 and Domain D2. There are no analyses in the manuscript that try to address this important question.

Another important issue regarding the trend analysis can be found in the specific comments.

Specific Comments

L. 12: I suggest to remove the reference to the project here (and at other positions in the manuscript) and state the project name solely in the acknowledgement section.

L. 32 – 38: I see that the linkage to dendroclimatological studies refers to the research project, but in my opinion, this is not needed here. You don't show any further results regarding this topic, and the general effort and method of dynamical downscaling does not really need to be justified or explained within this manuscript.

L. 49 – 59: The same as the comments above: this is in general interesting information,

but not within this manuscript. The paragraph should be shortened, maybe only keep the last sentence: "High-temporal…"

L. 66 – 68: Please remove the sentence: "These data…" for the reason stated above.

L. 69: While your statement here is certainly true, I would prefer a more moderate phrasing, e.g., "These data has the potential to find…" instead of "These data will also find…".

L. 78: Could you please give some more information on how the WRF configuration was chosen? This should then also be added to the manuscript. You state that the setup is based on Collier et al. (2019) but the study seems to be located in completely different climate and terrain conditions (East Africa). It is widely shown in literature that the performance of the chosen configuration strongly depends on the region. I understand that it is not feasible to perform a full configuration optimization ensemble, but some more information on this issue should be added.

L. 122: It would be very interesting to see sub-daily results also for precipitation from such a simulation. By permitting convective events, this could potentially be one of the strong points of such a high-resolution simulation.

L. 123: What about all the cases where modeled precipitation > 0 but the observed precipitation = 0? These cases should somehow be analyzed too and not be neglected in the performance analysis.

L. 128: Why did you choose two hourly WRF output? I see that the output somehow has to be confined, but hourly values would also be very valuable! Do you still have these available? I think it is fine for the manuscript to keep two hourly results, but at least for the main surface variables, it would be very useful to have hourly values as well. If they are available upon request, you could add this information to the data availability section.

L. 147 - 157: I highly appreciate the very well investigated and documented error handling here!

L. 160: See comment above: why did you choose two-hourly output? Was it just to save storage or is there another reason?

L. 169 - 170: It is very valuable that you try to compare the results to other studies (here to the work by Warscher et al. 2019), but the numbers are not really comparable here (different investigation area, nudging strategy, stations, terrain, etc.). I would either keep your statement and add an explanation regarding the differences in the analyses or remove the statement or phrase it differently ("similar but lower" is quite inexplicit).

L. 180 - 181: To me, this is a strong hint that the used resolution is still not high enough to correctly simulate absolute convective precipitation amounts. That's one reason why it would be so valuable to analyze more than one year of data and to compare results between D1 and D2.

L. 213: You clearly show that the grid-nudged run is performing better than the "free" simulation, which again leads to the question of the benefits of the simulation. This result indicates that WRF adds biases compared to the ERA5 forcing simulation when not grid-nudged to them. The DWD stations you used for your validation might even have been assimilated in ERA5 which again questions to some point the added value of the simulation.

L. 215 - 216 Remove "(the temporal resolution of data available in BAYWRF)". This is not important here.

L. 228: Three typos: add spaces after "WRF:"

Sect. 3.5: Trend analysis: it is quite obvious that the trends are reproduced by the simulation when it is forced by a reanalysis product such as ERA5 (and grid-nudging is used). You could think about removing the whole section, as I do not see a value in this information. If the trends would not have been reproduced, it would be an argument that something goes wrong, but – the other way round - these results are not

proving a good performance of WRF (as stated in the paragraph). You simply see the overall dynamics of the forcing (which includes assimilations of historic observations and therefore reproduces historic trends).

L. 241 - 246: The paragraph falls a bit short compared to the other ones. The spatial distribution of trends could be more elaborated (if a trend section is kept). The fine scale spatial differences of trends is in the end the information that is produced by the RCM simulations (see the statements regarding trends and reanalysis above).

L. 261 - 263: The statement regarding grid-nudging may be true, but I do not see it as a success, as the forcing obviously includes assimilated observations (see comments above).

Fig. 7 c) and d): If I understand it right, the values in the legend should be reversed (wrong sign).

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2020-52, 2020.