Authors' response to reviewers: Optimization of weather forecasting for cloud cover over the European domain using the meteorological component of the Ensemble for Stochastic Integration of Atmospheric Simulations version 1.0

Yen-Sen Lu[1,4], Garrett Good[2], and Hendrik Elbern[3]

[1]Institute of Energy and Climate Research – Troposphere (IEK-8), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany
[2]Fraunhofer Institute for Energy Economics and Energy System Technology IEE, Königstor 59, 34119 Kassel, Germany
[3]Rhenish Institute for Environmental Research at the University of Cologne, Cologne, Germany
[4]Jülich Supercomputing Centre, Forschungszentrum Jülich, Jülich, Germany

Correspondence: Yen-Sen Lu (ye.lu@fz-juelich.de)

First of all, we authors thank the more comments from both reviewers for improving this manuscript. We also thank Editor Prof. Ullrich for editing our manuscript. All the questions and comments are replied to point by point and we also improved the manuscript over these points.

1 Reply to the first reviewer's comments

1.1 Reply to the *Major Comments*

Line 98 – The authors state that they could not evaluate Eta or Morrison microphysics due to a requirement of higher resolution; however, microphysics parameterization is necessary at all NWP resolutions, and these schemes can certainly be run at coarser scales than convection allowing resolutions (for instance, the Eta scheme is run in the NAM model). Note that the WRF documentation states that different Eta microphysics settings should be chosen based on resolution, not that the Eta scheme shouldn't be used at all for specific resolutions: "e. Eta microphysics: The operational microphysics in NCEP models. A simple efficient scheme with diagnostic mixed-phase processes. For fine resolutions (< 5 km) use options (5) and for coarse resolutions use option (95)." Please consider revising this text.

We agree that this is not because those two microphysics can not perform under this setup. In our simulations, we did not use the correct option (95) for the Eta microphysics. And therefore we do not evaluate these simulation results over the original 674 combinations. We change the sentence to: *We thus evaluate the ten microphysics which are limited to no cloud-resolving simulations (UCAR, 2015).*

1.2 Reply to the *Minor Comments*

Line 39 – I would define what specific physics combination results in biased predictions of insolation over Germany, as this is most likely a physics-related problem.

We extend the sentence to include those specific combination: *It is for example the authors' experience that WRF, using the combination of microphysics Kessler, WSM5, or WSM6 with planetary boundary layer physics YSU, MYJ, or MYNN3 Berndt (2018), typically results in biased predictions of the solar resource in Germany, which results in overestimation of solar energy in the simulation case.*

Line 63 – Please define a value or range for "moderate resolution"

We modify it as *(at moderate resolution, between convection-permitting and global high-resolution (Kitoh and Endo, 2016))*

Line 96 – A reference to Figure 2 here would be helpful.

We add the reference length accordingly.

Line 97 – "uneven spacing near the boundary layer" – Are the first eleven vertical layers more concentrated near the surface?

Yes, and thus we modify it as *uneven spacing through the vertical direction, especially near the surface*

Line 100 – The reference to the Stergiou et al. (2017) paper may be better suited for the results section where it can be compared to the findings from this manuscript.

We extend the discussion with some comparison to Stergiou et al. (2017) at line 435: *Our results show some agreement with Stergiou et al. (2017), that WSM6 and Goddard was performing well (as TOPSIS Ranking as 5th and 7th) for the precipitation in July and CAM5.1 performs good for the temperature in January.*

Line 189 – "applied to the simulation" – used as initial conditions?

We modify our sentence to: *is also used as initial conditions*

Figure 5 – The caption should say "different PBL configuration"

Yes we fix this typo from *microphysics* to *PBL.*

Line 276 – Ukrain -> Ukraine

We fix this typo.

Figure 11 (and others) – Units appear to be missing for some axis titles

For Kernel Density and cloud fraction it's unitless, which is marked as (-)

Line 306 – . . . the Goddard scheme works best. . . ?

Yes. we modify it to: *By row, the Goddard works better with the Tiedtke and Grell-3D cumulus ... .*

Line 309 – "different microphysics" – it would be good to specify which.

We modify the sentence as: *perform better with WSM5, WSM6, and Goddard microphysics*

Figure 18 – X-axis title should by "Month of Year"

We fix this typo.

Line 407 – "insufficiencient"

We fix this typo to: *insufficient*

Line 424 – "... comprehensive insights into other spatial scales, other meteorological variables, further physics configurations..."

We fix this to ... *comprehensive insights into other spatial scales* ...

2 Reply to the second reviewer's comments

2.1 Reply to the *General Comments*

My only general comment regards the new and final figure, which for some reason plots individual values of the matching rates instead of utilizing the kappa score, which was used in all of the other analyses. As it is now, Figure 18 is not very useful because there are just too many points for the eye to easily interpret. With these longer simulations, the authors should consider compositing the hourly matching rates based on environmental factors (local time of day?, week of the year?, synoptic setting?, etc.) and present the kappa values computed in each of the bins. Doing so would provide a much more convincing argument that the simulations with these four models are similar.

We change the original dot plot into the boxplot that shows the variation within a month for better interpretation. Also we use change to use the kappa to show a more comprehensive result over the comparison between simulated and observed cloud mask.

2.2 Reply to the *Specific Comments*

138: How are they biased?

We believe you asked the same question with the first reviewer for Line 39, and thus we've already extended this sentence to clarify.

Eq 3: I think this should be pi instead of pj because the outer-sum in Eq 2 is over i and not j and also the sum in Eq 3 is already over j so there are no remaining j-specific variables. But then in Line 160 it says "pj over j" so I am very confused. I also do not understand why the final n in Eq 2 and Eq 3 is in parentheses. It also seems that I do not understand the difference between n and nij. The description in the "2.3.2 Kappa score" section seems like a textbook description but it would be more useful for the reader if the variables were introduced in relation to how they are used in this specific work, for example by relating them to the elements in Table 5.

We have re-worked on the formula to make it more clear for reader. We also add the description connecting to our simulation work as: $\bar{P} - \bar{P}_e$ *is the actual degree of agreement between raters and* $1 - \bar{P}_e$ *is the degree of agreement when matching correctly. For number of* n *raters,* N *subjects will be rated into* k *categories. Each* n_{ij} *represents the number of raters agreeing on the* j*-th category. In our work, the five possible outcomes in Table 5 are the categories for the two raters, the simulation and observation results for the gridcells as subjects.*

Figure 18: Why did you choose to examine the matching rate instead of the kappa score like was done in the rest of the analyses?

We change it to the kappa score for consistency. Using matching rate is convenient when the cloud mask data shows only two data type including cloudy and clear sky. Using matching rate can be straight forward for showing only 3 categories (match, miss, and false).

2.3 Reply to the *Technical Comments*

160: I suggest moving Eq 2 and Eq 3 up to directly follow Eq 1 because the text on lines 159-160 includes variables that have not yet been shown in an equation.

Yes, I've follow this suggestion for working on the specific comment.

167: The variable n is introduced in a different section than where it is first used (Eq 2).

We have re-worked on the equations, as your specific comment.

Figure 3: Increase the font size in this and many of the other figures.

We modify the figures accordingly, including Figure 3, 4, 5, 8, 9, 10, 11, 14, 16, and 17.

242: There is no Figure 4c or Figure 5c. I think you mean Figure 4a and Figure 5a but I'm not completely sure.

Yes, it's 4b and 5b, we fix this typo.

Figure 5 caption: This says microphysics configuration but it should be PBL configuration, according to lines 232-233.

Yes, we fix this typo.

References

Berndt, J.: On the predictability of exceptional error events in wind power forecasting–an ultra large ensemble approach–, Ph.D. thesis, Universität zu Köln, 2018.

Kitoh, A. and Endo, H.: Changes in precipitation extremes projected by a 20-km mesh global atmospheric model, Weather and Climate Extremes, 11, 41–52, https://doi.org/https://doi.org/10.1016/j.wace.2015.09.001, https://www.sciencedirect.com/science/article/pii/S2212094715300219, observed and Projected (Longer-term) Changes in Weather and Climate Extremes, 2016.

Stergiou, I., Tagaris, E., and Sotiropoulou, R.-E. P.: Sensitivity Assessment of WRF Parameterizations over Europe, Proceedings, 1, https://doi.org/10.3390/ecas2017-04138, 2017.

UCAR: User's Guide for the Advanced Research WRF (ARW) Modeling System Version 3.7, https://www2.mmm.ucar.edu/wrf/users/docs/user_guide_V3/user_guide_V3.7/users_guide_chap5.htm, 2015.