REVIEW OF MANUSCRIPT SUBMITTED TO Atmos. Meas. Tech. Discuss.

Authors: Stefano Federico, Rosa Claudia Torcasio, Paolo Sanò, Daniele Casella, Monica Campanelli, Jan Fokke Meirink, Ping Wang, Stefania Vergari, Henri Diémoz, Stefano Dietrich

Title: Comparison of hourly surface downwelling solar radiation estimated from MSG/SEVIRI and forecast by RAMS model with pyranometers over Italy

Date of review: 02 APR 2017

OVERALL EVALUATION

The manuscript presents a one-year comparison between satellite-estimated and numerical weather prediction model –forecasted solar radiation with ground-based measurements at Italian sites. The paper is of interest as it involves both weather forecasting and satellite algorithms, considering a topical subject connected to solar energy production.

The manuscript presents a fairly thorough evaluation of the performance of these two sources of solar radiation information for the chosen Italian sites. The figures presented and the used equations seem correct. However, the manuscript occasionally presents conclusions that are not supported by the evidence presented in this study, and furthermore suffers from unclear sentences that perhaps could be improved through a proper language check / proof reading.

I would recommend major revisions before accepting this manuscript. Note, however, that I do not believe that the actual scientific work will require a deep revision, but rather, that the authors need to pay attention to way things are expressed and what conclusions can be made based on the results presented in their manuscript.

SOMEWHAT GENERAL COMMENTS

- L59-61 and elsewhere: throughout the manuscript, it would be important to emphasize (and remind the reader of the fact) that RAMS is a forecast (for the day ahead), and MSG-GHI is a satellite-based estimate (available some time after the satellite observations have been made). This needs to always be kept in mind when comparing the performance of the two – the present manuscript is occasionally somewhat sloppy on this.
- L110—118: The RAMS model should be properly introduced before starting the paragraph on exchange between atmosphere and surface. What is the RAMS model?
- L151—164: The text seem somewhat unclear here, should be clarified. It seems to me that maybe there are two different groups of pyranometers used: (i) L151-157, and (ii) L158-164. The stability of the pyranometer in

Aosta is documented, but nothing is said about the other pyranometers. If sunshine duration is used (point 2), which stations have sunshine duration available? How is the Aosta check against libRadtran done, what are the criteria for data removal? Which institutes are responsible for the pyranometer measurements?

- Fig 6 + analysis: The manuscript up to this point has explained/speculated about the role of clouds in the performance. Now, Fig 6 finally shows quantitative results on how the performance behaves as a function of cloud classification. To me, it would make sense to bring this figure more toward the early part of the manuscript, so that the remaining analysis already could make use of these results as this would reduce the need for indirect determination.
- Conclusions -> (suggestion) Summary and Conclusion: The section is in its present form to a large extent repeating/summarizing the results presented in previous sections. It would be more interesting for the reader if some more discussion/conclusions would be added. Perhaps the section could also be shortened.
- L448: Are any evidence presented in the manuscript that show that the radiative scheme is inable to simulate cloudy conditions correctly? Where does this statement come from?

SPECIFIC SUGGESTIONS

- L25-26: this is not always true (see e.g. L445—446)
- L27: RMSE increases for Alpine stations (and similar statements elsewhere). This seems a bit misleading. I would suggest saying "is higher" or "is lower".
- L30: "RMSE ranges from 152 W/m2 to…" – here (and elsewhere) it should be defined that the RMSE has been calculated for hourly values.
- L36: "a reduction" -> "lower"
- L36: "of at least 10%" – here, it needs to defined what 10% means (10% of the base value, or a change corresponding to 10% steps/units (e.g. from 20 to 10%). Also elsewhere.
- L46: two specific papers are cited for a scope very wide. I would suggest removing the references.
- L51: Yes, PV can convert GHI to electricity, but much more commonly, they convert tilted GHI, or perhaps more generally, just solar radiation, to electricity
- L91: suggest to remove "So"
- L92: particle size is not an optical property
- L94: could you clarify the text here, is a mixture of ice/water clouds possible?
- L112: "most of Europe" seems a bit exaggerated
- L131-132: Somewhat unclear what exactly this means. Could be elaborated more.
- L133-136: Please clarify, are there any additional data assimilated into the RAMS model, e.g. weather observations, or is the RAMS model's initial state fully determined purely by ECMWF?

- L138-139: Seems contradicting that a weather forecasting model would need a spin up time of 12 hours.
- Section 2.3: could be separated into two: (i) surface observations / (ii) evaluation methodology
- L146: “Vigna di Valle is still” -> “Vigna de Valle is” (remove still)
- L165: “environmental characteristics” seem to actually mean cloud classification by the satellite method, is that correct? Please clarify text and use suitable terminology.
- L170: (language) “with the stations” -> “between different stations” ?
- L174-175: somewhat unclear sentence, please clarify
- L178-179 + L188-189: why not use equation numbers?
- Figure 3: I would suggest swapping the axes, so that pyranometer values are on the x-axis and estimated values on the y-axis. This makes values above the 1:1 line correspond to overestimation and vice versa, which is more logical. Also, I think it would be interesting to add this kind of figure for each station as a supplement or appendix as some readers will be interested in that information. Finally, the figure would be easier to read if grid lines and a legend were added, and if the point style would be modified so that points would not overlap (as much) in the busy areas of the plot.
- L196 / Fig. 3: clarify how the regression lines were determined
- L201-202 and L220: “it is apparent the larger scatter” -> (language) please rephrase (also similar sentence construction elsewhere)
- L229-230: in point b, it needs to be emphasized that RAMS is a forecast and thus not directly comparable to MSG
- L256-257: It is unclear to the reader how the conclusion about clouds being the main source of error was made. Could this be elaborated?
- L257-264: This seems to be mostly somewhat loose speculation, although things are expressed as hard facts. The evidence presented in the manuscript does not support all these statements. Therefore, I recommend rewriting, to use more careful statements.
- L268-274: On a general level, I believe the explanations presented here to be plausible. However, I also find that the authors focus too much on explanations that have to do with local orography and horizontal resolution. Could there be something else involved as well? For example, one factor that certainly plays a role here is the fact that mountain stations have more clouds and clouds are difficult for the satellite algorithm (as seen later on in Fig 6).
- L287: RMSE -> rRMSE?
- L288-289: unclear to me what is meant by “statistic shows more clearly the impact of …”
- L294 + Fig 5 caption: “RAMS-GHI one-day forecast”. Here (and elsewhere when mentioning RAMS one-day forecast) it would be important to emphasize that RMSE is based on hourly values of the day ahead forecast from RAMS. The present text leads me to think that values may be daily.
- L301: I find it odd to say “This result is caused by RMSE statistics”.
- L321: “all sky conditions, which showed” -> (suggest) “all sky conditions, which indirectly showed…”
- L329-331: Unclear how this explains the difference?

- L333: Not completely true, compare with L445-446
- L344-346: Please clarify how the persistence forecast is created. Are values hour by hour assumed to be the same between the two days?
- L345: I find the use of the short-version “1D” somewhat misleading (as it makes me think of one-dimensional) and therefore suggest writing it out: one-day.
- Section 3.3: split into two separate sections? (i) Daily evaluation / (ii) MOS application
- L381-383: There may also be other sources of MBE
- L384: “The MOS consists of” -> (suggestion) “The MOS used here consists of”
- L389-393: Unclear how exactly this works, please clarify.
- L402: Somewhat unclear how this conclusion can be made based on the above sentences
- L436-440 / L441-446: the latter paragraph presents discussion on performance as function of cloud classification, while the previous contains similar information, but indirectly. Maybe the paragraphs could be combined into one, or the order be changed, to make more effective and convincing communication.