

Article Title: Development of a seamless forecast for solar surface irradiance using ANAKLIM++

Brief summary:

In this article, to provide the short-term forecast of solar surface irradiance, the authors suggested the blending method in which three types of input data (i.e., satellite-based nowcasting and the NWP forecast from ICON of DWD and IFS of ECMWF) are blended using ANAKLIM++ software. Though the aim of ANAKLIM++ algorithm is to assimilate the 2-D observation data with the forecast from the NWP system, in this study, this algorithm is used to blend various input data to provide more reliable forecast of solar surface irradiance for 1-5h forecast hours. For the verification with the SARA-2 data as a reference, it is shown that the solar irradiance forecast from this used algorithm is much better than other data products (e.g., NWP forecasts from ICON and IFS, nowcasting), thus reflecting the reliability of this algorithm.

I have just two major comments as follow. In particular, 2nd comment is the main question related to the forecast/assimilation concepts.

Major comments

1. Unclear description of the forecast methodology

I cannot make sure that I now understand the methodology using ANAKLIM++ algorithm to create the solar irradiance forecast over Europe. First, before the forecasts are made, three input data are ingested/blended with the adjusted weight value assigned to each input data, using ANAKLIM++ algorithm (shown Figure 1). And then, I cannot clearly understand how the forecast fields of solar irradiance (up to 5hours) are made.

According to what I understand from this article, the forecast fields are made by combining the fixed initial condition (produced by blending the nowcasting, two NWP products) at 1000UTC with the time-dependent NWP forecasts (1100, 1200, 1300, and 1400UTC), with the dynamically adjusted weight value assigned to each input data (shown in Table 2). Do I understand correctly? At this point, I would like to check that the nowcasting data, one of input data, are just adjusted with the assigned weight values at each forecast hour (1-4 hours), not evolved over time.

If possible, please explain the methodology in more detail for the potential readers who seem not to be familiar with this research field.

2. Question about the verification results

From Figure 4, the forecasts, made by blending the nowcasting and two NWP forecasts, show the best performance with the minimum errors (RMSE and MAE) as compared with the corresponding nowcasting and two NWP forecasts. But I have a question about this point. Considering that the forecasts are derived by just blending three input data, is it possible for the errors of the derived forecasts to be smaller than the errors of three input data? In the general concept of data assimilation, the model backgrounds (i.e., short-range forecasts) are updated by assimilating the ingested observation data with the embedded assimilation technique, and then the initial condition (called analysis) is finally made, which is used to initiate the NWP model. Generally, the accuracy of analysis field is position between the accuracy of model background and observations. In other words, the model background can be closer to the real field by blending more accurate observation data into the model field with the well-fitted error estimation but cannot be better than the observation.

In terms of this assimilation concept, in this study, the two NWP forecast can be considered “model background”, the satellite-based nowcasting is considered “observation”. Considering that two NWP forecasts are adjusted by being blended with the nowcasting, it is difficult for the derived forecast to be better than the accuracy of nowcasting, because the errors of NWP forecast are mixed. If the reference data (called SARA-H-2) are used as the input data for ANAKLIM++, the derived forecast can be better than the nowcasting, because the nowcasting is also made with the numerical algorithm even though the satellite data are employed. However, this reference data are just used to verify the derived forecast and other input data, not employed in the algorithm using ANAKLIM++. If you have another point able to explain this verification result, please add this point in the manuscript.