Response to review comment 1

Manuscript: AMT-2023-26

Title: Validation of a camera-based intra-hour irradiance nowcasting model using synthetic cloud data

Authors: Philipp Gregor, Tobias Zinner, Fabian Jakub, and Bernhard Mayer

We thank the anonymous referee for the comments on the manuscript and suggestions to improve its quality. These are addressed in the following. The authors' answers are printed in blue. A version of the manuscript with tracked changes is provided along the updated manuscript.

Summary:

This study introduces a short-term nowcasting model that combines techniques such as machine learning and data assimilation in a novel fashion to help in predicting the direct normal irradiance. Validation of the models and methods is thorough and the authors take the time to explain the interpretation of their results. The addition of an "undecided" class to the training step is clever, especially when tied to their loss function which works to sidestep uncertainties in classification. Data assimilation from two separate imagers is used appropriately and adds an extra layer of context for the initial state. Overall, the paper advances the state-of-the-art of nowcasting by combining several innovative methods and could serve as a baseline for future research in the radiation or energy communities using such techniques.

Specific Comments:

1. The caption for Fig. 2 needs to be more descriptive. As well, Fig. 2 is not adequately described in the main text when it is first referenced although lines 251-253 do add more context. I would strongly suggest providing the reader with that context for the figure to start with by adding more information in the caption.
   For easier interpretation of Figure 2, we added the axis description (x, y) and extended the caption. Additionally, the main text was extended to introduce P0, P1, P2 and A1 along the reference of Fig. 2.

2. In Sect. 3.2, it would be helpful if a histogram of the retrieved cloud base height is added. This would allow a reader to quantify the performance of the derived base height for the entire set of scenes without limiting samples as has been done for Fig. 3.
   We changed Fig. 3. to two histograms of (a) height of matched pixels and (b) image average cloud base height to include all samples. The text in Sect. 3.2 was adapted accordingly.

3. Starting from line 340, the authors use a value within parentheses when describing the irradiance. It is unclear as to what these values are referring to, particularly as there is a preceding value before the parentheses as well. For instance, lines 343 - 344 say "Typical improvement over persistence for these longer lead times is thereby on the order of $50 Wm^{-2}$ ($50 Wm^{-2}$) and more" but both values being the same creates confusion. I would recommend

introducing the parameter within the parentheses first or explaining it at the top of the paragraph.
The first value gives the error for point DNI whereas the value in the brackets gives the error for area DNI. We tried to indicate this in line 335-336 (previous version, now 348-349): *In the following, error values are given for point DNI and in brackets for area DNI if not stated otherwise*.
For further clarification we extended this to: *Errors for point and area forecasts show similar characteristics. Therefore, it is discussed jointly in the following. If not stated otherwise, error values are given for point DNI and in brackets for area DNI.*

4. There are a number of grammatical errors overall that will need to be corrected before publication. For instance, in lines 20-21, the phrasing should be “Since direct irradiance can be blocked completely by clouds within seconds to minutes, knowledge of future direct irradiances is especially important for solar energy applications.”. The incorrect use of adverbs and articles in many places interrupts the flow and might particularly detract a reader from the point of the sentence or a paragraph which is why I am adding this issue as a major comment.
   We rechecked the manuscript completely and tried to correct grammatical errors and improve overall readability. We are thankful for your comment and hope that we were able to improve the manuscript.

5. In the appendix, it is mentioned that the ResNet encoder of the CNN uses pre-trained weights from ImageNet. This seems like an unnecessary step as the ImageNet classes are oriented at natural object detection (and not clouds) and transfer learning from a pre-trained ResNet would not necessarily reduce the convergence time on a task such as cloud detection. Could the authors clarify why a pre-trained model is better as opposed to training from scratch for this application?
   The ImageNet dataset is not focussed on cloud or sky images and does not feature according classes. Therefore we agree that transfer-learning from a pre-trained ResNet encoder may seem unnecessary. In our tests, however, we found the pre-trained weights to be helpful for faster convergence during training on the cloud segmentation data. Unfortunately we have not looked deeply into this difference in convergence. Arguing based on an intuitive understanding, we suggest that the reduced convergence time is due to the fact that especially the weights of the first convolution layers are pre-trained to focus on gradients in the input images. Although cloud boundaries are often fuzzy, there still are gradients present especially in the vicinity of cloud borders. Using a ResNet with randomly initialized weights, additional training is necessary for this focus on gradients. Due to the simple availability of ResNet weights pre-trained on ImageNet, we compared training with and without these and found improved convergence with pre-trained weights.

Technical Corrections:

1. The LaTeX equations have not rendered correctly in the preprint. For instance, line 207 has a question mark instead of equation numbers. This needs to be corrected.

The broken link to Equ. 4 was fixed in the LaTeX document. Apart from this, we found no other broken links.

2. The cloud optical depth threshold in line 273 should be reversed to say $\tau > \tau_{thresh}$ is classified as a cloudy region.
   We corrected it to $\tau \geq \tau thresh$.

3. In line 388, there is mention of mean absolute error. Since this metric is not presented in the main text or appendix or supplement, I would recommend removing this sentence as it is unnecessary. The RMSE and MBE already provide sufficient quantification.
   The reference to mean absolute error in line 388-389 was removed.