Review of "Simulating the impacts of utility-scale photovoltaic installations with a physically based coupled WRF-PV model" by *Chen et al.* submitted to *Geoscientific Model Development*

This study couples a physically based PV-panel energy-balance model (an adaptation of Heusinger et al., 2020) into the Noah land-surface module in the WRF model, deriving a combined PV-ground energy-balance equation that ensures closure between panels, ground, and atmosphere. Coupled model simulations with and without PV modules are done over a major utility-scale PV deployment region in northwestern China for the summers of 2018-2024 and compared against available observations. The authors report a daytime cooling of skin temperature, a daytime warming of 2 m air temperature, enhanced sensible heat flux, weakened lower-tropospheric stability, increased low-level cloud, and a redistribution of summer precipitation toward extremes.

The manuscript addresses physically consistent representation of utility-scale PV installations in regional climate models, a topic of clear relevance to GMD readers. The energy-closed coupling formulation is a reasonable contribution to the modeling side of PV modules' climate effect. I have a few general concerns on incomplete implementation details described in the methodology section, significance of the model validation results, and how it differentiates from the previous modeling studies. Overall, I would recommend a major revision before publication.

Please find my suggestions listed below.

General Suggestions

1. Although I understand that this study incorporates two-way coupling and atmosphere-panel-surface energy closure, which is a valuable addition to the previous models, these changes seem incremental and are not shown to bring significant improvement to the model performance. Specifically:

- Section 4.1 serves as validation of the model implementation. However, the comparison mainly focuses on regional mean, and the difference between the PV run and the control run is so small that I doubt whether it is noise from internal variability. From my expectation, I would see larger difference when averaged only over PV regions. In this case, how is the model-simulated T and precipitation compared to the observed ones?
- While section 4.1.3 presents simulated skin temperature compared against MODIS-observed skin temperature only over the PV grid cells, the difference of the probability distribution is still very small, and there are still a substantial number of samples that have very large bias above 5 K.
- Sections 4.2~4.4 proceeds to analyze impacts on temperature, local energy budget, and atmospheric responses. I appreciate it to include physical interpretation of the model results. However, previous modeling work also has similar analyses, and such decomposition in this study is not accompanied by in-situ measurements as validation. Is the simulated PV module temperature consistent with the real world scenario?

- What is the expected primary benefit of introducing the two-way coupling mechanism that previous modeling studies do not possess? It would be interesting to compare the results of different implementations (including effective albedo method in Li et al., 2018) side by side and see how large the difference is.

2. Since this is a manuscript submitted to GMD, I would expect a complete and detailed model description. In its current form, the manuscript is missing a few important pieces. A rule of thumb is that readers should be able to reproduce the results from reading the methodology section, even when the model is open source. Some notable examples are listed below, while more can be found in the specific points:

- The authors in many cases point the readers to previous papers. All parameters used in this study are not specified. A table that lists all configurable parameters in this module and their values is desired. Furthermore, does the model take both fixed-tilt and single-axis tracking solar farms into account? Does the model account for azimuth angle of solar panel tilt and solar radiation? Do all the identified solar farms share the same panel configurations? Please discuss the sensitivity of the results to this assumption, or at minimum acknowledge it as a limitation.
- In addition, Section 2.2 derives Eqs. (9)~(14) but never explains where in the land model the new equations sit. At what stage of the Noah time step are the PV fluxes computed? Are SW_dir, SW_dif and LW_down obtained from the radiation scheme of the host atmospheric column? Is the panel temperature T_PV solved diagnostically (since C_module is set to zero, per L131) or with a small numerical inertia? How is the soil-temperature equation modified given that GRDFLX is now driven by the shaded ground rather than the unshaded surface? A short numerical-implementation subsection is required.

3. The motivation could be more sharply tied to GMD's scope. GMD is primarily a model-development journal. The introduction reads at present like a regional climate-impact paper that happens to develop a model along the way. I would recommend reframing the goal so that the model description and evaluation come first, with the regional impact analysis as a demonstration of capability.

Specific Points

L32~33: Please elaborate on specific "potential environmental impacts". There are papers you can refer to that specifically discuss such environmental impacts (e.g., Turney and Fthenakis, 2011). Also, it seems that this sentence seems fit better with the next paragraph.

L35~38: These two sentences about PV panels' shortwave effects should be joined together. All the cited studies cover both points (i.e., reduced surface albedo, greater shortwave absorption by PV panels, and reduced solar radiation reaching the ground) in the sentences.

L54: "Guoqing et al., 2021" the citation and the associated reference are not formatted correctly. It is using the first name instead of the last name.

L71~72: “advanced UCRC-Solar by introducing a PV canopy energy-exchange scheme” please briefly state what this scheme adds.

L79: How much “radiative forcing” would the utility-scale solar farms induce? In the manuscript, this is not addressed, either (see specific comments below).

L84~85: I suggest including the time period and the geographical region of the simulations to be more specific.

L92: “total absorbed shortwave radiation” => “total absorbed shortwave radiation by the panels”

L93: “net longwave radiation exchange” => “net longwave radiation flux into the panel” (indicating that positive values mean net energy gain in the panels)

L95: “Heusinger (2021)” => “Heusinger et al. (2021)”

Eq. (2): A physical term-by-term explanation is desired. For example, “the first term on the RHS is direct solar radiation, where the cosine term is […]”. I am confused about the third and the fourth terms. Are they absorption by the back panel?

L100~101: “f_PV is the fractional […]” first this sentence needs grammatical correction. Also, is the “fractional surface coverage” projected to the horizontal ground? why is shaded area fraction approximated here instead of calculated analytically from angles?

Eq. (4): Again, a brief explanation of the physical meaning of each term on the RHS would be appreciated. What is the assumption on the geometrical placement of the PV panels? Are PVF_PV the same for top panel surface and bottom panel surface?

Eq. (5): The authors justify the factor of two by stating that “the front and rear surfaces of the panel participate in convective heat exchange with the atmosphere”. This is reasonable for free-standing bifacial panels in still air, but it is not obvious that both sides should use the same heat transfer coefficient (the rear face is partly shielded by the support structure and stands in much weaker mean wind). Please justify in more detail or reference an experimental basis.

(following the comment above) Does the model consider dynamical effect of solar panels (i.e., modification to surface friction)?

Eq. (6): Here, the temperature derating uses a coefficient of 0.005 K^{-1}. Is it suitable for the PV panels examined in this study?

Section 3.1: Please provide more details on the observational datasets. It should at least cover the spatial and temporal resolution (which the authors have stated for CMFD and NMSDC datasets), overview of the instruments or algorithms, validation results, and uncertainty.

L149~151: Which MODIS LST product is exactly used in this study? MODIS provides two variants of the surface emissivity and LST products, MOD11/MYD11 and MOD21/MYD21. They employ very different retrieval algorithms. Also, what is the collocation strategy between MODIS pixels and model grid cells in derivation of Figure 5?

L150: MODIS derived LST also accompanies retrieved surface emissivities in two IR bands. Is the effective surface emissivity in the model simulations consistent with the one reported by MODIS? This could affect interpretation of the skin temperature.

Section 3.2: What is the temporal frequency and spatial resolution of these three datasets? If the spatial resolution is incompatible with the model grid, how are they incorporated into the model? Are these datasets consistent with each other so that at the transition points they do not introduce discontinuity issues? Specifically, the second and the third datasets have overlapping observations between 2019~2022. Do these two datasets agree? How many solar farms were selected as a result?

L170: Please specify the exact values of the pressure levels.

L179: "sea surface temperature […] updated daily in the WRF simulations". Please confirm, as ERA5 provides hourly fields of sea surface temperature.

L182: "annual evolution of PV distribution" is a bit confusing. Is the fractional PV assigned every day that gradually grows with the installation of new solar farms? Or is the PV distribution constant in the simulated period of each year?

L186: Is there any reason why this study focuses on the boreal summer seasons rather than including the boreal winter seasons?

Section 3.4: How many ensemble member runs are conducted in this study? How would internal variability affect the results?

Figure 1: What is the total area of solar panel region?

Figures 2 and 3: please add bias maps (model – CMFD) since the spatial pattern panels alone do not show where the errors lie.

L197: Figure 4 is introduced before Figure 3.

L225: How many valid MODIS samples are used in the analysis?

L229: What is the mean bias value of both simulations? Given the large RMSE of over 3 K in both simulations, it is hard to argue that WRF_PV showed improved skills over the control simulation. What's the difference of RMSE over surrounding non-PV grid cells (to see whether the improvement is local to PV pixels).

Figure 5: x-axis label and unit are missing. What does the blue vertical line mean?

Section 4.3: How are these local energy budget perturbations translated to the regional mean values (e.g., within the D3 domain)? This could partly answer the question of how large the radiative forcing of solar farms could impose on the regional climate system.

Section 4.3: It would strengthen the model-evaluation case to present a closed energy-balance budget table or schematic diagram (NetRad, SH, LH, GRDFLX, P_out — daytime mean, nighttime mean, 24-h mean) for both runs and verify closure to within a stated tolerance.

L268: "specular reflection from the PV panels at large solar zenith angles". Is the angular dependence of panel albedo parameterized? I suspect this is in fact a numerical artifact of Eq. (2) with a constant panel albedo value.

Figure 8: The figure does not have multiple panels, yet the caption includes "(a)".

L316: additional comma at the end of the sentence.

Figure 9: The SW_down decrease is attributed to "an increase in low cloud fraction" (L308~309) and Fig. 9 shows positive low-cloud anomaly as well as increasing updraft. It is also possible that increased atmospheric water vapour absorption from the enhanced surface-air mixing could also contribute. I suggest at least adding a decomposition of the SW_down anomaly into clear-sky and cloud components and, optionally, quantify the change in column-integrated water vapour and aerosol optical depth.

Figure 10: Explain the elements of the box plot. What is the red curve? Also it is important to note that all differences are statistically insignificant. Maybe examine the spatial pattern of the extreme-precipitation change so that we can see whether it is correlated with PV location, downstream of it, or essentially noise?

L355~366: Related to General Suggestion #1, the conclusion makes a reasonable case that the present model produces smaller anomalies than the effective-albedo Sahara studies because of correct treatment of inter-row ground and panel thermal capacity. It would be even better to demonstrate this directly: i.e., run a third experiment in which the WRF-PV scheme is replaced by an effective-albedo formulation in the same domain, and show that the over-estimation reproduces.

L367: "this study considered only the evaporation change". I assume the authors mean that this study only analyzes the latent heat and precipitation variables, not that soil moisture was held fixed, is that correct?

Section 5: The discussion of the implications from this study seems thin. Some ideas: how easily WRF-PV can be ported to other land-surface schemes and at what computational cost; what evaluation data would most help future development; how the SW_down feedback (-1%) translates into a reduction in power production.

References

Turney, D., & Fthenakis, V. (2011). Environmental impacts from the installation and operation of large-scale solar power plants. *Renewable and Sustainable Energy Reviews*, *15*(6), 3261–3270. https://doi.org/10.1016/j.rser.2011.04.023