

Response

#Reviewer 2

Note on the organization of this response: Reviewer comments are shown in **blue**, while text quoted from the manuscript is shown in **red**. For each comment, we provide the reviewer's original comment, our response, the relevant original manuscript text, and the corresponding revised text. Because we have included both the original and revised manuscript passages to make all changes fully traceable, the response document is relatively long.

Summary:

The motivation of the paper is to estimate clear-sky fluxes during partly cloudy and overcast conditions for the observed CERES instantaneous footprint all-sky fluxes.

However, only clear-sky regions that have an area extent of at least a 20 km are used to estimate the clear-sky fluxes for partly cloudy and overcast footprint conditions. The population of clear-sky footprints is ~2% of all footprints and are not evenly distributed over the ocean as they occur over mostly the sub-tropics and rarely over the southern oceans and the ITCZ. The atmospheric profile conditions vary considerably if the clear-sky conditions are greater than 20-km than clear-sky conditions next to cloud edges, surrounded by clouds, or overcast conditions with no clear-sky. The near TOA clear-sky fluxes near clouds are colder and brighter than large domain clear-sky conditions (Eytan et al 2025). The clear-sky conditions within mostly cloudy conditions are less sampled than for dryer profiles. The clear-sky CSRFormer fluxes are then applied to overcast conditions, where the atmosphere is more humid and have greater aerosol reflectivity, for which it was not trained on and must extrapolate over. Maybe I do not fully grasp the capability of the CSRFormer methodology. Can a paragraph be added to address this.

Eytan, E., Gristey, J. J., & Feingold, G. (2025). The net radiative effect of the Ill-Defined clear-Sky in the vicinity of clouds. AGU Advances, 6, e2024AV001407. <https://doi.org/10.1029/2024AV001407>

If the goal of the Cloud Radiative Effect (CRE) is to derive the clear-sky flux under overcast conditions or the impact of instantaneously removing all clouds over the Earth, where the atmosphere still retains the cloudy atmospheric profile. I do not see how these unsampled clear-sky events can be estimated using CERES SSF L2 20-km clear sky fluxes that are based on clear-sky atmospheric profiles. It would be interesting to see a frequency plot similar to Fig. 4 of all SSF footprints during the training period.

Response:

We thank the reviewer for raising this important concern. We agree that the atmospheric states represented by large clear-sky CERES footprints do not necessarily span the full range of conditions encountered in partly cloudy and overcast scenes, and that this issue requires an explicit assessment rather than assuming that all cloudy-scene predictions are interpolations within the training distribution.

We first clarify the intended definition of the CSRFormer clear-sky reference. The prediction under cloudy conditions is not intended to represent the flux of a nearby naturally clear atmosphere. Instead, it represents a clear-sky state in which the instantaneous thermodynamic, aerosol, surface, and solar-geometry conditions of the cloudy footprint are retained while the cloud radiative contribution is removed. Thus, during application to cloudy scenes, CSRFormer is driven by the atmospheric state associated with the cloudy footprint itself. However, because the observational training targets are necessarily obtained from genuinely clear-sky CERES footprints, the atmospheric states encountered under clouds may extend beyond those represented in the training population.

Following the reviewer's suggestion, we therefore quantified the overlap between the 2019–2020 clear-sky training population and both the full CERES SSF ocean population and the overcast population in SST-TCWV space (Figure R3). Using $N_{\text{train}} \geq 50$ samples per SST-TCWV bin as an operational definition of a well-supported training regime, 91.41% of all SSF samples and 82.16% of overcast samples fall within the well-supported domain. An additional 2.00% of all-SSF samples and 5.38% of overcast samples occur in sparsely sampled bins containing 1-49 training samples. In contrast, 6.59% of all-SSF samples and 12.46% of overcast samples fall in SST-TCWV bins that contain no clear-sky training samples. Sensitivity tests using thresholds from $N_{\text{train}}=1$ to 500 lead to the same qualitative conclusion: the majority of cloudy atmospheric states overlap with the clear-sky training domain, but the overlap is systematically weaker for overcast conditions.



Figure R3. Training-domain coverage of CSRFormer in SST-TCWV space during 2019–2020. (a) Distribution of the clear-sky training samples; (b) distribution of all CERES SSF ocean samples; (c) distribution of overcast CERES SSF samples. Colors in (a)-(c) indicate $\log_{10}(N+1)$, where N is the number of samples in each SST-TCWV bin. (d) Training-

60 data support for the overcast population, where bins containing at least 50 clear-sky training samples ($N_{\text{train}} \geq 50$) are classified as well supported and bins with fewer than 50 training samples are classified as low-support regimes. Under this criterion, 91.4% of all SSF samples and 82.2% of overcast samples fall within the well-supported SST-TCWV domain.

65 The new frequency analysis confirms that the atmospheric states encountered in cloudy and overcast scenes are not fully covered by the clear-sky training population. In particular, a non-negligible fraction of overcast scenes falls outside the SST-TCWV regimes sampled by the clear-sky training data and therefore requires the model to generalize beyond the sampled thermodynamic domain. The geographically resolved analysis (Figure R4) further shows that these low-support conditions are not uniformly distributed and occur preferentially in humid tropical cloud regimes, including parts of the ITCZ and warm-pool regions.

70 regions. We have added the SST-TCWV frequency analysis to the main text and the geographical training-support analysis to the Appendix.

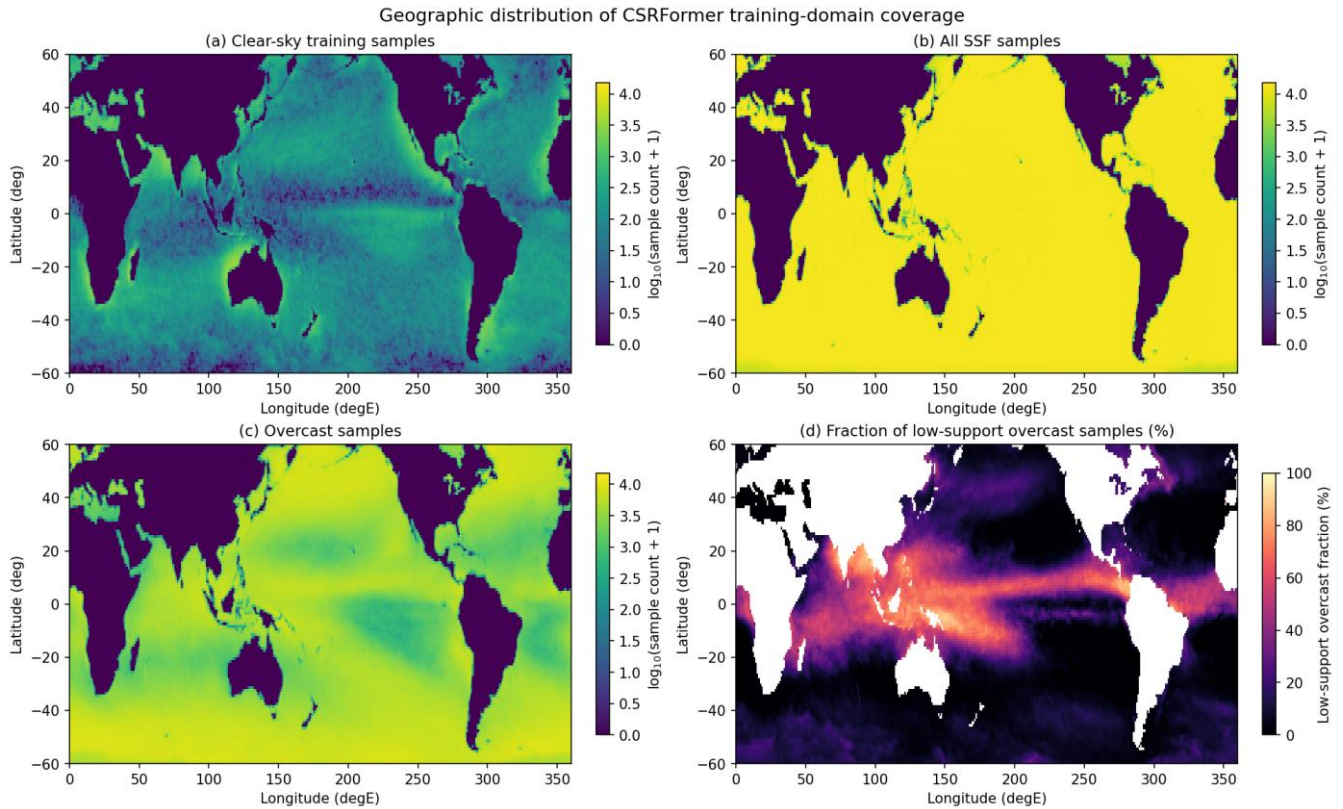


Figure R4. Geographic distribution of CSRFormer training-domain coverage during 2019–2020. (a) Spatial distribution of clear-sky training samples; (b) all CERES SSF ocean samples; and (c) overcast CERES SSF samples, expressed as $\log_{10}(N+1)$ within each geographic grid cell. (d) Geographic fraction of overcast samples falling within low-support SST-TCWV regimes, defined as bins containing fewer than 50 clear-sky training samples ($N_{\text{train}} < 50$).

Larger values indicate regions where overcast atmospheric states are less frequently represented in the clear-sky training population.

80 We further performed a perturbation-based stress test specifically targeting the low-support part of the training domain. Starting from independent test samples located in well-supported SST-TCWV bins, we increased either SST or both humidity-profile variables (q and RH) sufficiently to shift a subset of the samples into bins with $N_{\text{train}} < 50$. When SST was increased by 4 and 6 K, 41.5% and 59.2% of the perturbed samples, respectively, moved into the low-support domain. For those samples, the predicted clear-sky LW flux increased by 6.15 and 11.16 W m^{-2} , respectively, with the expected positive response in 100%
85 of cases. Similarly, increasing both q and RH by 50% and 100% shifted part of the samples toward more humid, weakly sampled conditions, and the predicted LW flux decreased by 5.99 and 9.26 W m^{-2} , respectively, with the expected negative response in approximately 99% of cases. These tests do not provide a direct validation of prediction accuracy in poorly sampled regimes, because corresponding observational clear-sky targets are unavailable, but they indicate that the model retains the expected qualitative LW response to surface warming and atmospheric moistening even when the perturbed states extend into
90 sparsely represented portions of the clear-sky training distribution. The LW hard constraint additionally prevents the predicted flux from exceeding the corresponding surface blackbody emission.

Therefore, we have revised the manuscript to explicitly distinguish well-supported interpolation from sparsely supported and extrapolative regimes, rather than implying that cloudy and clear atmospheric states are fully interchangeable. We also added
95 a training-domain support flag to the released dataset so that users can identify footprints falling in sparsely sampled or unsampled SST-TCWV regimes and treat those predictions with appropriate caution. We note that the SST-TCWV comparison provides only a simplified measure of training-domain coverage, because atmospheric states with similar SST and TCWV can still differ in their vertical temperature, humidity, and aerosol profiles.

100 Additionally, we examined the training-domain coverage for SW radiation using AOD and SZA, which are the two dominant predictors of clear-sky SW flux in the CSRFormer feature-importance analysis. Using the same $N_{\text{train}} \geq 50$ criterion, 96.4% of all SSF samples and 93.6% of overcast samples fall within well-supported AOD-SZA bins. This overlap is higher than that found in the SST-TCWV analysis, indicating that the two dominant controls on the shortwave prediction are represented relatively well in the clear-sky training population, including under overcast conditions. Nevertheless, low-support regions
105 remain, particularly toward less frequently sampled combinations of aerosol loading and solar geometry. We therefore regard this analysis as evidence of substantial, but not complete, training-domain coverage for the shortwave prediction. We also note that AOD-SZA provides only a simplified diagnostic and does not establish complete overlap in aerosol composition, vertical aerosol structure, or the full set of shortwave input variables.

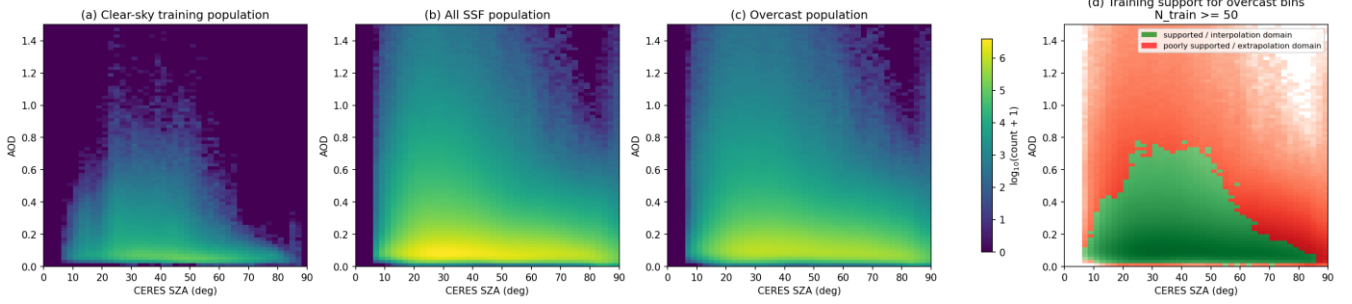


Figure R5. Training-domain coverage of CSRFormer in AOD–SZA space for clear-sky shortwave flux during 2019–2020. (a) Distribution of the clear-sky training samples; (b) distribution of all CERES SSF ocean samples; and (c) distribution of overcast CERES SSF samples. Colors in (a)–(c) indicate $\log_{10}(N+1)$, where N is the number of samples in each AOD–SZA bin. (d) Training-data support for the overcast population, with bins containing at least 50 clear-sky training samples ($N_{\text{train}} \geq 50$) classified as well supported and bins with fewer than 50 samples classified as low-support regimes. Under this criterion, 96.4% of all SSF samples and 93.6% of overcast samples fall within the well-supported AOD–SZA domain.

We also thank the reviewer for drawing our attention to the discussion of the “ill-defined clear sky” in Eytan et al. (2025). We have considered this study and cite it in both the Introduction, where we discuss the difficulty of defining and observing a clear-sky reference beneath clouds, and Sect. 4.1, where we discuss residual cloud contamination and near-cloud radiative effects in nominally clear CERES footprints.

We note, however, that the operational definition of clear sky used in the present study differs from the framework adopted by Eytan et al. (2025). Their study explicitly distinguishes cloudy, cloud-influenced clear-sky, and far-field clear-sky regimes in order to quantify the radiative influence of clouds on their surroundings. In contrast, we do not attempt to identify a separate cloud-transition or cloud-influenced clear-sky regime. Our training samples are defined operationally using the CERES Edition 4 clear fraction ($>99.9\%$) together with an additional MYD06_L2 1-km cloud-mask screening. The purpose of this screening is to construct a stringent observational clear-sky training population rather than to redefine the physical boundary between cloudy and clear-sky conditions.

This distinction also reflects the different objective of the present study. Our primary aim is to estimate instantaneous cloud radiative effects, which requires a clear-sky reference that is consistent with the atmospheric and surface state of the corresponding cloudy footprint. We therefore focus on constructing a stringent clear-sky training population and reconstructing this theoretical cloud-free reference, rather than on characterizing the radiative transition between cloudy and clear-sky regions. We recognize that the cloud–clear transition zone and the radiative influence of nearby clouds are important scientific issues,

as highlighted by Eytan et al. (2025), but a detailed classification and quantification of such transition regions are beyond the scope of the present study.

140 **Changes in the manuscript:**

Section 1:

Original:

The primary objective of this study is to develop and evaluate an instantaneous clear-sky radiative flux dataset over the global ocean using a physically constrained Transformer framework. In this study, cloud-free CERES clear-sky observations are used
145 as the observational anchor for training and evaluation, whereas the intended product application is the estimation of the theoretical clear-sky reference state beneath cloudy scenes, which cannot be observed directly. We focus on model construction, evaluation against CERES clear-sky observations, comparison with existing products, and application to instantaneous CRE analysis. The remainder of the manuscript is organized as follows. Section 2 describes the datasets. Section 3 presents the preprocessing, model architecture, and evaluation strategy. Section 4 reports the results, product comparisons, and CRE
150 application. Section 5 summarizes the main findings and limitations.

Revised:

The primary objective of this study is to develop and evaluate an instantaneous clear-sky radiative flux dataset over the global ocean using a physically constrained Transformer framework. Here, the theoretical clear-sky reference represents a state in
155 which the instantaneous thermodynamic, aerosol, surface, and solar-geometry conditions associated with the cloudy footprint are retained, while the radiative contribution of clouds is removed. No clear-sky atmospheric profile is spatially transferred or interpolated into cloudy footprints; instead, CSRFormer is driven by the atmospheric and surface state collocated with each individual footprint. We focus on model construction, evaluation against CERES clear-sky observations, comparison with existing products, and application to instantaneous CRE analysis. The remainder of the manuscript is organized as follows.
160 Section 2 describes the datasets. Section 3 presents the preprocessing, model architecture, and evaluation strategy. Section 4 reports the results, product comparisons, and CRE application. Section 5 summarizes the main findings and limitations.

Section 3.3:

Original:

165 In addition, as an indirect consistency test, based on the physical hypothesis that "similar meteorological states should have similar clear-sky radiation" (Allan et al., 2007; Prata, 2006; Koll and Cronin, 2018), this study searched for "cloudy-clear" pixel pairs with similar SST and TCWV within the dataset. By comparing the predicted clear-sky longwave radiative fluxes in the cloud-covered and clear-sky regions of these pairs and calculating their relative differences, we assessed the physical robustness of the model's inference in cloudy areas. Since TCWV data provided by ERA5 was used, to maintain consistency
170 between observed and model-determined cloud areas, the cloud areas selected in this evaluation adopted a dual screening

criterion of CERES clear fraction less than 0.1% and ERA5 cloud fraction greater than 99.9%. Cloudy-clear pairs were matched within SST bins of [0, 30] K and TCWV bins of [2, 52] kg m⁻², and a bin only was included in the statistics if it contains more than 2,500 pairs.

175 Revised and added:

In addition, because the intended application of CSRFormer includes cloud-covered scenes whereas the observational training targets are obtained from clear-sky CERES footprints, we first examined whether the atmospheric states encountered under cloudy conditions were represented in the clear-sky training population. The 2019-2020 clear-sky training samples were compared with all CERES SSF ocean footprints and overcast footprints from the same period in SST-TCWV space. TCWV was used here as a bulk measure of atmospheric moisture. For shortwave radiation, where aerosol loading and solar geometry are dominant controls, an analogous comparison was conducted in AOD-SZA space. For each bin in the corresponding two-dimensional state space, the number of clear-sky training samples (N_{train}) was calculated, and $N_{\text{train}} \geq 50$ was adopted as an operational criterion for a well-supported training regime. Sensitivity tests using different minimum sample-count thresholds were also conducted to assess the robustness of the inferred coverage.

185

As an indirect consistency test, based on the physical hypothesis that "similar meteorological states should have similar clear-sky radiation" (Allan et al., 2007; Prata, 2006; Koll and Cronin, 2018), this study searched for "cloudy-clear" pixel pairs with similar SST and TCWV within the dataset. By comparing the predicted clear-sky longwave radiative fluxes in the cloud-covered and clear-sky regions of these pairs and calculating their relative differences, we assessed the physical robustness of the model's inference in cloudy areas. Since TCWV data provided by ERA5 was used, to maintain consistency between observed and model-determined cloud areas, the cloud areas selected in this evaluation adopted a dual screening criterion of CERES clear fraction less than 0.1% and ERA5 cloud fraction greater than 99.9%. Cloudy-clear pairs were matched within SST bins of [0, 30] °C and TCWV bins of [2, 52] kg m⁻², and a bin only was included in the statistics if it contains more than 2,500 pairs.

195

Section 4.2:

Original:

Because clear-sky fluxes beneath clouds are not directly observed, evaluation in cloudy scenes is necessarily indirect. The key variables for predicting clear-sky shortwave radiative fluxes are SZA, AOD, SST, and wind speed; these variables are quantities that are not easily subject to direct variations caused by cloud cover. Therefore, the prediction of clear-sky shortwave radiative fluxes in cloudy areas should theoretically be as robust as in clear-sky areas (Zheng et al., 2025).

Regarding clear-sky longwave radiative fluxes, as demonstrated in Figures A4 and A5, clear-sky sample points near clouds possess temperature and humidity profiles similar to those in cloud-covered areas. Consequently, the features of such samples

205 are learned during the training process, thereby endowing the CSRFormer model with the potential to predict clear-sky longwave radiative fluxes within cloud-covered areas. To verify this idea, this study designed an indirect validation experiment. Specifically, we matched "cloudy-clear" pixel pairs with similar SST and TCWV on a global scale. By comparing the relative difference between the theoretical clear-sky longwave flux inferred by the CSRFormer model in the cloud-covered area and that in the clear-sky area, we assessed the generalization performance of the model. As shown in Fig. 4, CSRFormer
210 demonstrates stable generalization in cloud-covered regions with an average relative deviation of only -2.82%. The small mean deviation suggests that the model captures part of the thermodynamic control on clear-sky longwave flux beyond the exact clear-sky training samples, although this remains an indirect test. Notably, a slight negative bias was observed, where predicted clear-sky fluxes under clouds are marginally lower than in adjacent clear skies. This discrepancy likely reflects real thermodynamic differences rather than pure model error. Although SST and TCWV were controlled, cloudy regions typically
215 possess different temperature profiles—such as lower mid-to-upper tropospheric temperatures due to active vertical motion—which naturally leads to lower outgoing longwave radiation. While this indirect validation cannot serve as definitive proof, this physically explicable difference provides valuable supplementary evidence that the model remains sensitive to subtle variations in atmospheric vertical structures.

220 Revised and added:

(Since this section has been extensively revised and many new paragraphs have been added, completely replacing the original text, the paragraphs will not be highlighted in red here.)

Because clear-sky fluxes beneath clouds cannot be directly observed, the applicability of CSRFormer to cloudy conditions
225 must be evaluated indirectly. We first examined the extent to which atmospheric states encountered in cloudy scenes are represented in the clear-sky training population. Fig. 4 compares the 2019–2020 clear-sky training samples with all CERES SSF ocean footprints and overcast footprints in SST-TCWV space. Using $N_{\text{train}} \geq 50$ as an operational criterion for a well-supported training regime, 91.41% of all SSF samples and 82.16% of overcast samples fall within the well-supported domain. For overcast conditions, an additional 5.38% of samples occur in sparsely sampled bins containing 1-49 clear-sky training
230 samples, while 12.46% fall in SST-TCWV bins containing no clear-sky training samples. Sensitivity tests using different minimum sample-count thresholds yielded the same qualitative conclusion: most cloudy atmospheric states are represented to some extent in the clear-sky training distribution, but the overlap is incomplete and systematically weaker under overcast conditions. The geographical distribution of the training-domain coverage is shown in Fig. A6. The fraction of low-support overcast samples is spatially heterogeneous, with relatively high values occurring in several tropical regions, including parts
235 of the ITCZ and Indo-Pacific warm pool, consistent with the relatively limited sampling of warm and humid conditions in the clear-sky training population.

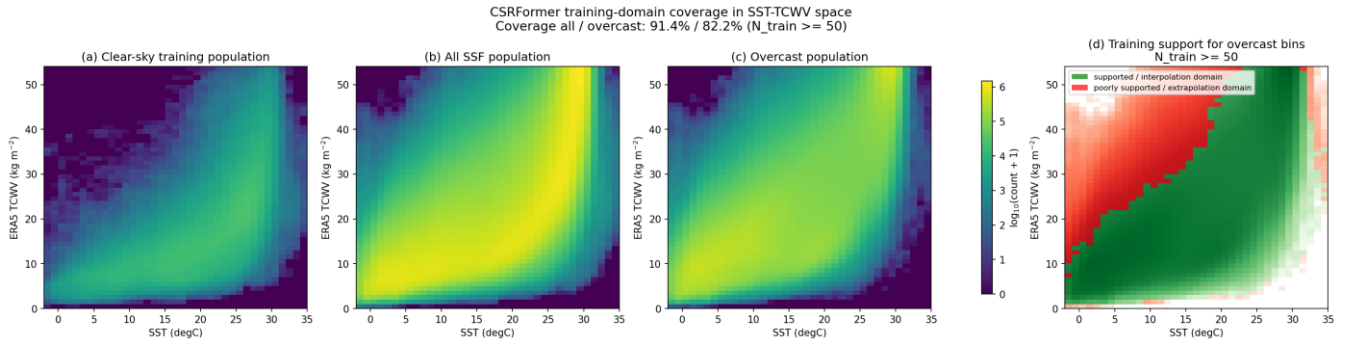


Figure 4. Training-domain coverage of CSRFormer in SST-TCWV space during 2019–2020. (a) Distribution of the clear-sky training samples; (b) distribution of all CERES SSF ocean samples; (c) distribution of overcast CERES SSF samples. Colors in (a)-(c) indicate $\log_{10}(N+1)$, where N is the number of samples in each SST-TCWV bin. (d) Training-data support for the overcast population, where bins containing at least 50 clear-sky training samples ($N_{\text{train}} \geq 50$) are classified as well supported and bins with fewer than 50 training samples are classified as low-support regimes. Under this criterion, 91.4% of all SSF samples and 82.2% of overcast samples fall within the well-supported SST-TCWV domain.

For SW radiation, we further examined the training-domain coverage in SZA-AOD space because aerosol loading and solar geometry are the two dominant controls identified by the feature-importance analysis. Under the same $N_{\text{train}} \geq 50$ criterion, 96.4% of all SSF samples and 93.6% of overcast samples fall within well-supported SZA-AOD bins. The substantially higher coverage indicates that the dominant aerosol-loading and solar-geometry conditions relevant to the SW prediction are comparatively well represented in the clear-sky training population, including under overcast conditions. Nevertheless, some low-support combinations remain. We interpret both the SST-TCWV and SZA-AOD analyses as simplified measures of training-domain coverage, because atmospheric states with similar bulk variables may still differ in their vertically resolved temperature, humidity, and aerosol profiles.

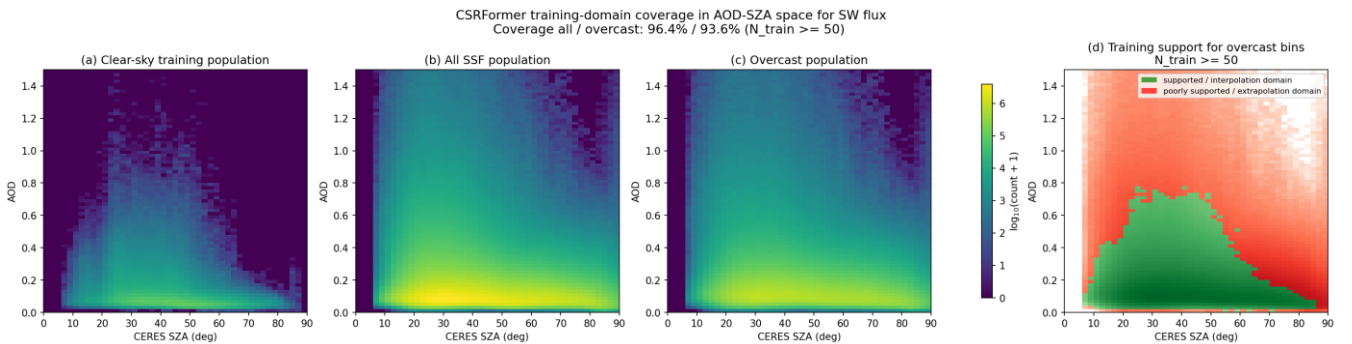


Figure 5. Training-domain coverage of CSRFormer in AOD–SZA space for clear-sky shortwave flux during 2019–2020. (a) Distribution of the clear-sky training samples; (b) distribution of all CERES SSF ocean samples; and (c) distribution of overcast CERES SSF samples. Colors in (a)–(c) indicate $\log_{10}(N+1)$, where N is the number of samples in each AOD–SZA bin. (d) Training-data support for the overcast population, with bins containing at least 50 clear-sky training samples ($N_{\text{train}} \geq 50$) classified as well supported and bins with fewer than 50 samples classified as low-support regimes. Under this criterion, 96.4% of all SSF samples and 93.6% of overcast samples fall within the well-supported AOD–SZA domain.

We further performed a perturbation-based stress test to examine the model response when atmospheric states were shifted from well-supported conditions toward low-support regions. For independent test samples that moved into $N_{\text{train}} < 50$ SST–TCWV bins after perturbation, increasing SST by 4 and 6 K increased the predicted clear-sky LW flux by an average of 6.15 and 11.16 W m^{-2} , respectively, with a positive response in 100% of cases. Similarly, increasing both the specific- and relative-humidity profiles by 50% and 100% decreased the predicted LW flux by 5.99 and 9.26 W m^{-2} , respectively, with the expected negative response in approximately 99% of cases. These experiments do not constitute a direct validation of prediction accuracy in low-support regimes because corresponding observational clear-sky targets are unavailable. They nevertheless indicate that CSRFormer retains the expected qualitative LW responses to surface warming and atmospheric moistening when shifted toward sparsely represented states. The LW hard constraint further prevents the predicted flux from exceeding the corresponding surface blackbody emission.

Having quantified the training-domain coverage, we further evaluated the conditional consistency of CSRFormer predictions between cloudy and clear-sky conditions within overlapping thermodynamic regimes. Based on the physical expectation that similar meteorological states should exhibit similar clear-sky radiation (Allan et al., 2007; Prata, 2006; Koll and Cronin, 2018), cloudy and clear-sky footprints with similar SST and TCWV were matched globally, and their CSRFormer-predicted clear-sky LW fluxes were compared within the same meteorological bins. As shown in Fig. 6, the mean relative difference between the cloudy and clear-sky predictions is only -2.82% , indicating generally consistent model responses under similar bulk thermodynamic conditions.

A small systematic negative difference remains, with the predicted clear-sky LW flux under cloudy conditions being slightly lower than that for matched clear-sky samples. This difference may partly reflect variations in the vertically resolved atmospheric structure that are not constrained by SST and TCWV alone. Even for samples with similar SST and column-integrated water vapor, the vertical distributions of temperature and humidity can differ between cloudy and clear-sky conditions, thereby affecting outgoing longwave radiation. Figures A4 and A5 provide illustrative examples of such similarities and differences in humidity-profile structure between selected cloudy and clear-sky footprints. Therefore, this comparison should be interpreted as a conditional consistency test rather than a direct validation of the unobservable clear-sky flux beneath

clouds. Together with the training-domain coverage and perturbation analyses described above, it provides complementary evidence for evaluating the applicability of CSRFormer under cloudy conditions.

295 **Section 4.5:**

Original:

The CSRFormer dataset provides instantaneous clear-sky top-of-atmosphere shortwave and longwave radiative fluxes over the global ocean, reconstructed at the CERES Aqua overpass time. The current release covers 2018-2020, with a native spatial sampling matched to the CERES SSF footprint (~20 km at nadir). The archived variables include clear-sky shortwave flux, clear-sky longwave flux and derived clear-sky albedo. The dataset is distributed in NetCDF format and is publicly available at 10.5281/zenodo.20046058 (Zheng et al., 2026). Each file contains time, latitude, longitude, clear-sky shortwave flux, clear-sky longwave flux and clear-sky fraction. The current v1.0 release is intended for instantaneous Aqua-overpass-time analyses over the global ocean and should not be used for long-term trend detection or climatological change assessment.

305 Revised and Added:

The CSRFormer dataset provides instantaneous clear-sky top-of-atmosphere shortwave and longwave radiative fluxes over the global ocean, reconstructed at the CERES Aqua overpass time. The current release covers 2018-2020, with a native spatial sampling matched to the CERES SSF footprint (~20 km at nadir). The archived variables include clear-sky shortwave flux, clear-sky longwave flux, derived clear-sky albedo and band-specific training-domain QC flags. The dataset is distributed in NetCDF format and is publicly available at 10.5281/zenodo.20046058 (Zheng et al., 2026). Each file contains time, latitude, longitude, clear-sky shortwave flux, clear-sky longwave flux, clear-sky albedo, clear-sky fraction and the corresponding SW and LW training-domain QC flags. The current v1.0 release is intended for instantaneous Aqua-overpass-time analyses over the global ocean and should not be used for long-term trend detection or climatological change assessment.

315 To facilitate the use of the dataset under atmospheric conditions that are less well represented in the clear-sky training population, two training-domain support flags are additionally provided. For LW flux, training support is evaluated in SST–TCWV space, whereas for SW flux it is evaluated in AOD–SZA space, following the analyses described in Sect. 3.3. For each flag, a value of 0 denotes a well-supported regime ($N_{\text{train}} \geq 50$), a value of 1 denotes a low-supported regime ($N_{\text{train}} < 50$). These flags are intended as quality-control indicators of training-data support rather than formally calibrated per-footprint prediction uncertainties. They therefore do not represent all possible sources of retrieval error, such as three-dimensional cloud-adjacency effects.

Section 5:

Original:

325 Overall, CSRFormer provides a practical framework for generating observation-time-matched clear-sky radiative flux
estimates over the global ocean. Its main strengths are the use of vertically resolved inputs, the alignment with CERES overpass
time, and the strong agreement with CERES SSF in clear-sky conditions. **The main caveats are that validation beneath clouds
is indirect, the model inherits limitations from its inputs and training target, and 3D radiative effects are not explicitly
represented.** With those limitations stated clearly, the dataset could be useful for future studies of CRE, cloud regimes, and
330 cloud-clear transition zones, and it may also be extendable to Terra in future work.

Revised:

Overall, CSRFormer provides a practical framework for generating observation-time-matched clear-sky radiative flux
estimates over the global ocean. Its main strengths are the use of vertically resolved inputs, the alignment with CERES overpass
335 time, and the strong agreement with CERES SSF in clear-sky conditions. **The main caveats are that validation beneath clouds
is indirect, the clear-sky training population does not fully cover all atmospheric regimes encountered under cloudy conditions,
the model inherits limitations from its inputs and training target, and three-dimensional radiative effects are not explicitly
represented.** With those limitations stated clearly, the dataset could be useful for future studies of CRE, cloud regimes, and
cloud-clear transition zones, and it may also be extendable to Terra in future work.

340

Comments 1

Line 265. It must be remembered that the CERES products do not use the Goddard cloud properties or mask but are based on
the CERES MODIS cloud retrievals designed to have as few non retrievals as possible. The cloud retrievals designate the
scene type required to convert nearly all footprint radiances into fluxes. The Goddard cloud retrievals are mostly from high
345 confidence clear-sky events disregarding the number of no retrievals from the cloud mask. From the CERES web site under
references for the SSF L2 product

<https://ceres.larc.nasa.gov/data/documentation/#ssf>

350 Trepte, Q. Z., P. Minnis, S. Sun-Mack, C. R. Yost, Y. Chen, Z. Jin, F.-L. Chang, W. L. Smith, Jr., K. M. Bedka, and T. L.
Chee, 2019: Global cloud detection for CERES Edition 4 using Terra and Aqua MODIS data. IEEE Trans. Geosci. Remote
Sens., 57, 9410-9449, doi: 10.1109/TGRS.2019.2926620.

Response:

355 Thank you for this important clarification. We agree that the original wording did not sufficiently distinguish between the
CERES-specific MODIS cloud retrievals and the independent MODIS cloud product used as an additional screening step in
our study. The CERES SSF cloud fraction is derived from the CERES Edition 4 MODIS cloud detection and retrieval system
developed specifically for CERES, rather than from the Goddard MODIS cloud mask.

360 In our workflow, the CERES SSF clear fraction was used as the primary clear-sky criterion, with footprints satisfying clear fraction > 99.9% retained. The 1-km MYD06_L2 cloud-mask information was then applied independently as an additional secondary filter to further reduce possible residual cloud contamination in the selected training samples. It was not used by CERES to define the SSF scene type or to convert CERES radiances to fluxes.

365 We have revised the relevant descriptions in Sections 2.1 and 3.1, as well as Table 1, to clearly distinguish the CERES-MODIS cloud retrieval from the additional MYD06_L2 screening used in this study, and we have added the CERES Edition 4 cloud-detection references suggested by the reviewer.

Changes in the manuscript:

370 **Section 2.1:**

Original:

The radiative flux data serving as the model training target and core ground truth are derived from the Single Scanner Footprint (SSF) Edition 4A product released by the Clouds and the Earth's Radiant Energy System (CERES) (Wielicki et al., 1996; Su et al., 2015b). The SSF product is a key dataset in the CERES project. It provides instantaneous, footprint-level (nadir
375 resolution of approximately 20 km) radiative observations by spatiotemporally matching broadband radiative fluxes measured by the CERES instrument with cloud and aerosol properties retrieved from Moderate High-Resolution Imaging Spectroradiometer (MODIS) hosted on the same platform (Terra and Aqua satellites). Specifically, this study selected the top-of-atmosphere (TOA) shortwave and longwave radiative flux data observed by the instrument onboard the Aqua satellite. Currently, the CERES SSF product has undergone rigorous validation and has been widely applied in numerous studies as a
380 benchmark for radiative fluxes or as a basis for product development (Smith et al., 2011; Loeb et al., 2018; Scott et al., 2022; Su et al., 2015a). Due to well-established use as a benchmark, we use its observed radiative fluxes as the training ground truth. In addition, this dataset provides detailed solar zenith angle (SZA), meteorological information from Goddard Earth Observing System (GEOS) Model 5.4.1, and cloud mask information retrieved based on MODIS observations. These provide the data foundation for accurately predicting clear-sky fluxes.

385

Revised:

The radiative flux data serving as the model training target and core ground truth are derived from the Single Scanner Footprint (SSF) Edition 4A product released by the Clouds and the Earth's Radiant Energy System (CERES) (Wielicki et al., 1996; Su et al., 2015b). The SSF product is a key dataset in the CERES project. It provides instantaneous, footprint-level (nadir
390 resolution of approximately 20 km) radiative observations by spatiotemporally matching broadband radiative fluxes measured by the CERES instrument with cloud and aerosol properties retrieved from Moderate High-Resolution Imaging Spectroradiometer (MODIS) hosted on the same platform (Terra and Aqua satellites). Specifically, this study selected the top-

of-atmosphere (TOA) shortwave and longwave radiative flux data observed by the instrument onboard the Aqua satellite. Currently, the CERES SSF product has undergone rigorous validation and has been widely applied in numerous studies as a benchmark for radiative fluxes or as a basis for product development (Smith et al., 2011; Loeb et al., 2018; Scott et al., 2022; Su et al., 2015a). Due to well-established use as a benchmark, we use its observed radiative fluxes as the training ground truth. In addition, the CERES SSF product provides solar zenith angle (SZA), auxiliary meteorological information from the Goddard Earth Observing System (GEOS) Model 5.4.1, and cloud properties derived from the CERES Edition 4 MODIS cloud detection and retrieval algorithms (Trepte et al., 2019; Minnis et al., 2021). These provide the data foundation for accurately predicting clear-sky fluxes.

Section 3.1:

Original:

To obtain training labels that are as close as possible to clear-sky conditions, we applied several screening steps. We first selected CERES SSF samples with Clear Fraction > 99.9%, following the spirit of the CERES SYN processing (Doelling et al., 2013). We then applied the 1 km MODIS cloud mask as a secondary filter to reduce residual cloud contamination (Frey et al., 2008; Platnick et al., 2017). Only ocean pixels were retained using the CERES SSF surface-type information, and sea-ice samples were excluded because they strongly affect surface albedo. Regions flagged for sunglint were also removed from the shortwave training sample to limit contamination by non-atmospheric reflection effects (Loeb et al., 2005; Su et al., 2015b; Platnick et al., 2017). Finally, all input variables were standardized using Z-scores before training with training set mean and variance.

Revised:

To obtain training labels that are as close as possible to clear-sky conditions, we applied several screening steps. We first selected CERES SSF samples with Clear Fraction > 99.9%, following the spirit of the CERES SYN processing (Doelling et al., 2013). The SSF clear fraction is derived from the CERES Edition 4 MODIS cloud detection algorithm developed specifically for the CERES processing system (Trepte et al., 2019), rather than from the standard Goddard MODIS cloud product. We then independently applied the 1-km cloud-mask information from MYD06_L2 as an additional secondary filter to further reduce residual cloud contamination in the training samples (Frey et al., 2008; Platnick et al., 2017). Only ocean pixels were retained using the CERES SSF surface-type information, and sea-ice samples were excluded because they strongly affect surface albedo. Regions flagged for sunglint were also removed from the shortwave training sample to limit contamination by non-atmospheric reflection effects (Loeb et al., 2005; Su et al., 2015b; Platnick et al., 2017). Finally, all input variables were standardized using Z-scores before training with training set mean and variance.

Tabel 1:

Original:

Cloud Mask	%	MODIS	1km (Planar)	Satellite	Screening
Sunlint Mask	%	(MYD06)		observation	Screening
Cloud Mask	%	CERES SSF	20km (Planar)	Satellite	Screening
				observation	

Revised:

Cloud Mask	%	CERES SSF	20km (Planar)	Satellite	Primary
				observation	Screening
Additional	%	MODIS	1km (Planar)	Satellite	Secondary
Cloud Mask		(MYD06)		observation	Screening
Sunlint Mask	%				Screening

430 **Comments 2**

Line 171 Where the clear-sky footprints for CSRFormer using both the Goddard and the CERES MODIS cloud mask? How many additional CERES clear-sky footprints were rejected due to the Goddard cloud mask?

Response:

435 Yes. The clear-sky samples used for CSRFormer were selected sequentially using both cloud-screening criteria. We first required the CERES SSF clear fraction to exceed 99.9%, where the clear fraction is determined by the CERES Edition 4 MODIS cloud-detection system. We then independently applied the 1-km MYD06_L2 cloud mask as an additional, more stringent screening step to reduce residual cloud contamination.

440 We have now quantified the additional sample rejection caused by the MYD06_L2 screening. For the independent 2018 test set, 4,595,180 ocean footprints satisfied the CERES clear-fraction criterion alone, whereas 3,728,769 remained after the additional MYD06_L2 screening. Thus, 866,411 footprints, corresponding to approximately 18.85% of the CERES-clear candidates, were additionally rejected by the MYD06_L2 cloud mask. The effect of omitting this secondary screening is also illustrated in Fig. A3.

445 We have added this information to the revised manuscript and clarified that the MYD06_L2 mask is an additional quality-control filter introduced in this study rather than part of the CERES SSF cloud-identification procedure.

Changes in the manuscript:

450 **Section 3.1 (Added in an edit to previous response revision):**

Revised:

To obtain training labels that are as close as possible to clear-sky conditions, we applied several screening steps. We first selected CERES SSF samples with Clear Fraction > 99.9%, following the spirit of the CERES SYN processing (Doelling et al., 2013). The SSF clear fraction is derived from the CERES Edition 4 MODIS cloud detection algorithm developed specifically for the CERES processing system (Trepte et al., 2019), rather than from the standard Goddard MODIS cloud product. We then independently applied the 1-km cloud-mask information from MYD06_L2 as an additional secondary filter to further reduce residual cloud contamination in the training samples (Frey et al., 2008; Platnick et al., 2017). For example, in the independent 2018 test set, 4,595,180 footprints satisfied the CERES clear-fraction criterion alone, while 3,728,769 remained after the additional MYD06_L2 screening, corresponding to an additional rejection of 866,411 footprints (18.85%). Only ocean pixels were retained using the CERES SSF surface-type information, and sea-ice samples were excluded because they strongly affect surface albedo. Regions flagged for sunglint were also removed from the shortwave training sample to limit contamination by non-atmospheric reflection effects (Loeb et al., 2005; Su et al., 2015b; Platnick et al., 2017). Finally, all input variables were standardized using Z-scores before training with training set mean and variance.

Comments 3

Figure 5.

I appreciate the authors including the important input features analysis, to see if the inputs align with physics. It is good to see the AOD is a primary component in clear-sky non-polar ocean SW fluxes. I wonder why SZA is a distant second. I believe this is due to the Aqua 1:30 sun-synch orbit where the SZA are constrained by LT and a bit by season. But why would SZA and cos(SZA) be treated separately, is this not redundant? I see wind speed contributions are small but impact the probability of glint. I see the glint is not considered. I see that SST is included in the SW contributions, is this to determine the near surface water vapor absorption along with specific or relative humidity? For the clear-sky LW, how is upper tropospheric humidity not considered in the feature importance?

475

Response:

Thank you for this insightful comment. We agree that treating SZA and cos(SZA) as independent variables in the original permutation analysis was not ideal because they represent strongly correlated descriptions of the same solar geometry. Incoming solar radiation is also closely related to these quantities. When these correlated variables are permuted separately, the predictive information associated with solar geometry can be distributed among them, leading to an underestimate of their collective importance.

We therefore repeated the feature-importance analysis by grouping SZA, cos(SZA), and incoming solar radiation into a single “solar geometry” feature group and perturbing them simultaneously. In the revised analysis, solar geometry remains the second most important contributor to clear-sky SW prediction, but its contribution increases substantially to approximately 25%,

485

compared with approximately 46% for AOD. Thus, the large separation seen in the original figure was partly caused by splitting highly correlated solar-geometry information among several predictors.

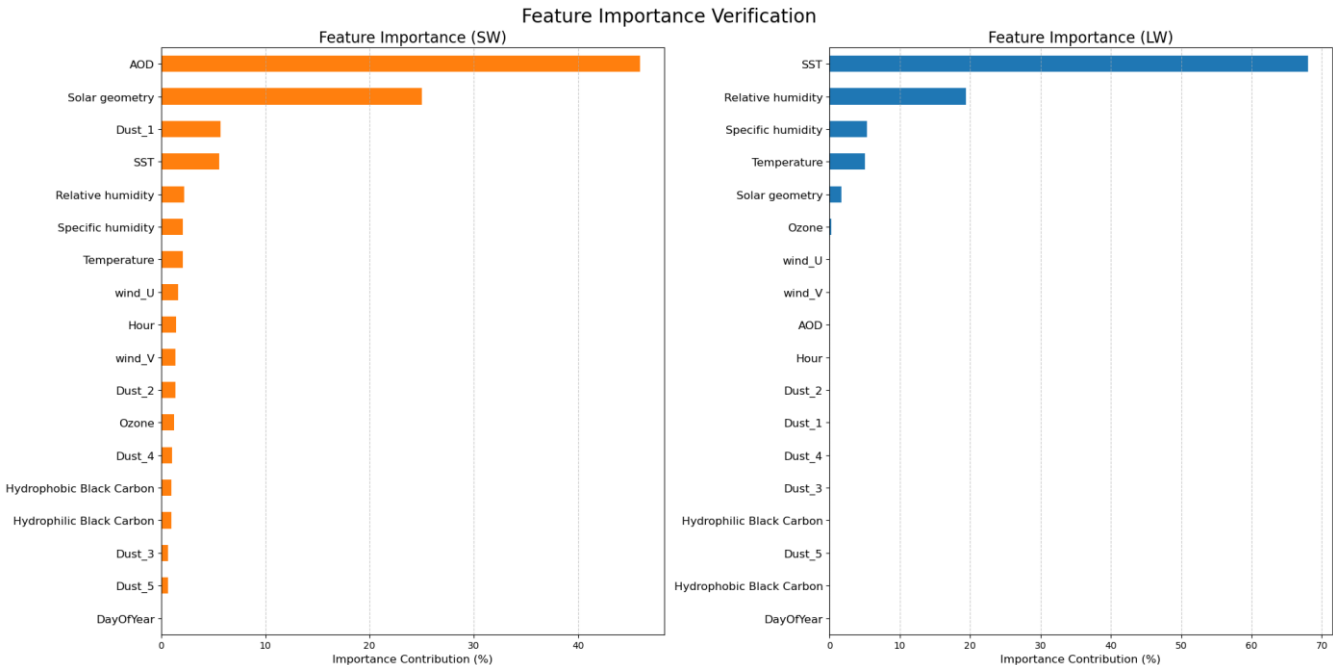


Figure R6. Permutation feature importance of CSRFormer input variables for clear-sky (a) SW and (b) LW fluxes. Importance is quantified by the increase in RMSE after perturbing each feature or physically related feature group while leaving the remaining inputs unchanged, and is normalized to sum to 100% within each radiative band. “Solar geometry” represents the grouped perturbation of SZA, cos(SZA), and incoming solar radiation. Temperature, specific humidity, and relative humidity represent grouped perturbations of their complete 37-level ERA5 vertical profiles.

The remaining difference between AOD and solar geometry is also physically reasonable for the Aqua sampling configuration. Aqua observes at a nearly fixed local overpass time of approximately 13:30, which constrains the variability of solar geometry relative to a dataset spanning the full diurnal cycle. In contrast, aerosol loading exhibits substantial spatial and temporal variability across the global ocean and therefore provides strong independent predictive information in the sampled population. SZA and cos(SZA) were both originally provided to the network to facilitate learning of the nonlinear dependence on solar geometry: SZA retains the angular coordinate itself, whereas cos(SZA) explicitly represents the projection of the direct solar beam. We agree, however, that they should not be interpreted as independent physical contributions in a feature-importance analysis. The revised grouped analysis avoids this ambiguity.

Sunglint is not explicitly used as a predictor in CSRFormer. Instead, CERES footprints flagged as affected by sunglint were excluded from the SW training samples. Wind speed was retained because it affects ocean-surface roughness and therefore

modulates surface reflectance even outside the explicitly flagged sunglint regions (Jin et al., 2004; Winter and Chýlek, 1997). We have clarified this distinction in the revised manuscript.

510 SST was not included in the SW prediction specifically to represent near-surface water-vapor absorption; this effect is represented directly by the vertically resolved humidity profiles. Rather, SST was retained as a surface-state variable. In addition to its direct role in the surface energy state, SST can covary with marine productivity, ocean color, and other ocean-surface conditions that influence surface optical properties (Jin et al., 2004). We therefore regard the small SW importance assigned to SST as an indirect empirical contribution rather than evidence of a direct causal control of SST on reflected shortwave flux.

515 Finally, upper-tropospheric humidity is included in the feature-importance analysis, but it is not shown as a separate predictor. The specific- and relative-humidity inputs consist of the full 37-level ERA5 profiles, and the entire profile is permuted jointly as one feature group. Therefore, the contribution of upper-tropospheric humidity is contained within the reported humidity-profile importance. The current analysis evaluates the importance of the full vertical humidity structure and does not isolate individual atmospheric layers. We have clarified this point in the revised text and figure caption.

Changes in the manuscript:

Section 3.2:

Original:

525 The weighting parameters α_n ($n = 1, 2, \dots, 5$) are set to 7, 3, 2, 0.5 and 3, respectively. The loss weights were manually selected based on preliminary experiments to balance the relative magnitudes of the individual loss components and to maintain stable joint optimization of the SW and LW objectives. The model was trained with a learning rate of 5×10^{-5} using the AdamW optimizer for 150 epochs and a batch size of 128, utilizing dropout (rate is 0.1) and weight decay ($1e^{-5}$) for regularization (Srivastava et al., 2014; Loshchilov and Hutter, 2017). The planar inputs are SST, AOD, wind speed, and SZA. The vertical-
530 profile inputs are ERA5 temperature, humidity, and ozone profiles (37 levels) and MERRA-2 aerosol profiles (72 levels). The shared encoder uses 128 embedding dimensions, 4 Transformer layers, and 8 attention heads.

Revised:

535 The weighting parameters α_n ($n = 1, 2, \dots, 5$) are set to 7, 3, 2, 0.5 and 3, respectively. The loss weights were manually selected based on preliminary experiments to balance the relative magnitudes of the individual loss components and to maintain stable joint optimization of the SW and LW objectives. The model was trained with a learning rate of 5×10^{-5} using the AdamW optimizer for 150 epochs and a batch size of 128, utilizing dropout (rate is 0.1) and weight decay ($1e^{-5}$) for regularization (Srivastava et al., 2014; Loshchilov and Hutter, 2017). The planar inputs include SST, AOD, wind speed, SZA, $\cos(\text{SZA})$, incoming solar radiation, date and time. Both SZA and $\cos(\text{SZA})$ are provided as alternative representations of solar geometry:

540 SZA retains the angular coordinate, whereas $\cos(\text{SZA})$ explicitly represents the nonlinear projection of the direct solar beam. The vertical-profile inputs are ERA5 temperature, humidity, and ozone profiles (37 levels) and MERRA-2 aerosol profiles (72 levels). The shared encoder uses 128 embedding dimensions, 4 Transformer layers, and 8 attention heads.

Section 3.3:

545 Original:

To deeply investigate the dependence of the CSRFormer model on different atmospheric and surface physical parameters during the retrieval of clear-sky radiative fluxes, and to verify whether the model has correctly learned the dominant physical mechanisms of radiative transfer, this study employed the Permutation Feature Importance evaluation method to conduct sensitivity analysis on key input variables (Cao et al., 2024; Wang et al., 2023). Specifically, we quantified permutation
550 importance as the increase in RMSE after randomly shuffling one feature or one profile group across samples; a more significant error increment indicates a stronger dependence of the model on that physical quantity, thereby providing a quantitative basis for the physical interpretability of the model.

Revised:

555 To deeply investigate the dependence of the CSRFormer model on different atmospheric and surface physical parameters during the retrieval of clear-sky radiative fluxes, and to verify whether the model has correctly learned the dominant physical mechanisms of radiative transfer, this study employed the Permutation Feature Importance evaluation method to conduct sensitivity analysis on key input variables (Cao et al., 2024; Wang et al., 2023). Specifically, we quantified permutation importance as the increase in RMSE after randomly shuffling one feature or one profile group across samples; a more
560 significant error increment indicates a stronger dependence of the model on that physical quantity, thereby providing a quantitative basis for the physical interpretability of the model. For vertically resolved variables, all levels belonging to the same physical profile were permuted jointly; therefore, the reported humidity importance represents the contribution of the full vertical profile rather than that of an individual atmospheric layer. Because SZA, $\cos(\text{SZA})$ and incoming solar fluxes are strongly correlated representations of the same solar geometry, they were treated as a grouped feature when interpreting their
565 permutation importance to avoid artificially splitting the contribution of solar geometry between redundant predictors.

For clear-sky longwave radiative flux, the distribution of feature importance highlights the dominance of thermodynamic laws and the key role of atmospheric vertical structure. SST is the most important predictor, which is expected because the ocean surface is the primary source of upward thermal emission under clear-sky conditions (Jin et al., 2004). Temperature and
570 humidity profiles rank next, indicating that the model also captures the vertical structure of atmospheric emission and water-vapor absorption (Allan et al., 2007; Prata, 2006; Koll and Cronin, 2018). Interestingly, evaluation results show that SZA and incident solar radiation also exhibit certain importance in longwave prediction. Although solar geometric parameters do not directly participate in longwave radiative processes, this correlation is physically explicable: SZA and incident solar radiation

are associated with atmospheric temperature and surface temperature, and this association indirectly influences the assessment results of feature importance (Gentemann et al., 2003; Trenberth et al., 2009). Overall, the importance pattern is consistent with the physical controls on clear-sky longwave radiation.

Section 4.2:

Original:

We next used permutation feature importance to assess which inputs most strongly influence the predictions (Fig. 7). For shortwave flux, AOD and SZA are the dominant predictors, consistent with aerosol extinction and solar geometry (Tegen et al., 1997; Payne, 1972; Taylor et al., 2006; Zhang et al., 2022). Absorbing aerosol profiles also contribute, indicating that vertically resolved aerosol information adds skill beyond column AOD alone (Samset et al., 2013; Bond et al., 2013). SST and wind speed are of secondary but non-negligible importance, consistent with their effects on ocean-surface reflectance (Hansen et al., 1983; Winter and Chýlek, 1997; Jin et al., 2004).

Revised:

We next used permutation feature importance to assess which inputs most strongly influence the predictions (Fig. 7). For clear-sky SW flux, AOD and solar geometry are the two dominant sources of predictive information (Tegen et al., 1997; Payne, 1972; Taylor et al., 2006; Zhang et al., 2022). AOD has the largest contribution, accounting for approximately 46% of the total permutation importance, consistent with the strong modulation of reflected shortwave radiation by aerosol extinction (Samset et al., 2013; Bond et al., 2013). The grouped solar-geometry variables, including SZA, $\cos(\text{SZA})$, and incoming solar radiation, contribute approximately 25%. The remaining difference between AOD and solar geometry is consistent with the sampling characteristics of Aqua. Because Aqua observes at a nearly fixed local overpass time ($\sim 13:30$ LT), the range of solar geometry represented in the dataset is more constrained than in a full diurnal sampling, whereas aerosol loading varies substantially across space and time. Wind makes a smaller but physically interpretable contribution through its effect on ocean-surface roughness and reflectance (Hansen et al., 1983; Winter and Chýlek, 1997; Jin et al., 2004). SST also provides a secondary contribution and is interpreted primarily as an indirect surface-state proxy, including possible covariance with marine productivity, ocean color, and other surface optical conditions (Hansen et al., 1983; Winter and Chýlek, 1997; Jin et al., 2004).

For clear-sky longwave radiative flux, the distribution of feature importance highlights the dominance of thermodynamic laws and the key role of atmospheric vertical structure. SST is the most important predictor, which is expected because the ocean surface is the primary source of upward thermal emission under clear-sky conditions (Jin et al., 2004). Temperature and humidity profiles rank next, indicating that the model also captures the vertical structure of atmospheric emission and water-vapor absorption (Allan et al., 2007; Prata, 2006; Koll and Cronin, 2018). The humidity importance reported here represents the complete 37-level ERA5 profiles because all pressure levels are perturbed jointly; upper-tropospheric humidity is therefore included within these profile-level contributions rather than evaluated as a separate predictor. Interestingly, evaluation results

show that solar geometry also exhibit certain importance in longwave prediction. Although solar geometric parameters do not directly participate in longwave radiative processes, this correlation is physically explicable: solar geometry parameters are associated with atmospheric temperature and surface temperature, and this association indirectly influences the assessment results of feature importance (Gentemann et al., 2003; Trenberth et al., 2009). Overall, the importance pattern is consistent with the physical controls on clear-sky longwave radiation.

Figure 7 caption:

Original:

Figure 7. Feature importance verification for CSRFormer input variables. Importance is quantified by the increase in RMSE when a specific feature is randomly permuted while others remain unchanged (Permutation Feature Importance). Features are sorted in descending order of importance; a larger increase in RMSE indicates a more critical contribution of that feature to the model's predictive accuracy. The importance in the chart has been normalised, with the total summing to 100%.

Revised:

Figure 7. Permutation feature importance of CSRFormer input variables for clear-sky (a) SW and (b) LW fluxes. Importance is quantified by the increase in RMSE after perturbing each feature or physically related feature group while leaving the remaining inputs unchanged, and is normalized to sum to 100% within each radiative band. “Solar geometry” represents the grouped perturbation of SZA, cos(SZA), and incoming solar radiation. Temperature, specific humidity, and relative humidity represent grouped perturbations of their complete 37-level ERA5 vertical profiles.

Comments 4

Figure 6. I would agree that the CSRFormer algorithm would achieve a near zero bias to the 20-km clear-sky observations, which is not the case for untuned Fu-Liou computed fluxes utilized by the CERES project, since the CSRFormer algorithm was trained with the observed clear-sky fluxes. Do the CSRFormer input datasets that differ from the CERES input dataset significantly improve the consistency of the CSRFormer and observed clear-sky fluxes or is most of the impact from CSRFormer algorithm?

Response:

We thank the reviewer for this important point. We agree that the close agreement between CSRFormer and CERES SSF in clear-sky conditions should first be interpreted in light of the fact that CSRFormer is trained directly against the observed CERES SSF clear-sky fluxes. Therefore, the near-zero mean bias in Fig. 8 is an expected consequence of the observation-anchored training objective, whereas the CERES CRS fluxes are produced independently using the Fu–Liou radiative-transfer model rather than being fitted directly to the observed footprint-level clear-sky fluxes.

At the same time, CSRFormer and CRS also use different atmospheric and surface inputs. CSRFormer is driven by ERA5 thermodynamic profiles, MERRA-2 aerosol information, NOAA OISST, and CERES-derived surface/solar-geometry variables, whereas CRS uses MODIS-derived aerosol and cloud properties together with GEOS meteorological fields. These input differences may also contribute to the pointwise differences between the two products, particularly for shortwave radiation where aerosol and surface properties are important.

For these reasons, we do not think the present comparison can objectively and completely separate the respective contributions of the input datasets and the model formulation. More fundamentally, CSRFormer and CRS follow different methodological pathways. CSRFormer treats CERES SSF clear-sky fluxes as the observational reference to be learned and explicitly optimizes its predictions toward that target, whereas CRS independently estimates radiative fluxes from a physically based radiative-transfer calculation using prescribed atmospheric and surface states. Differences between CRS and CERES SSF may therefore arise from multiple sources, including uncertainties or differences in the driving input fields as well as approximations inherent in the radiative-transfer calculation, such as those associated with finite-stream or other numerical approximations. The present analysis does not isolate the relative contribution of these factors.

Nevertheless, comparison with CERES SSF remains important because SSF provides one of the closest available observational references for instantaneous, footprint-scale TOA radiative fluxes. The stronger agreement of CSRFormer with SSF therefore demonstrates an important advantage of the observation-trained framework: its reconstructed clear-sky fluxes remain closely anchored to the satellite observations. At the same time, we have revised the manuscript to avoid interpreting the CRS-SSF differences simply as evidence that the CRS radiative-transfer algorithm is inaccurate, and instead discuss them as the combined result of differences in model formulation, driving inputs, and radiative-transfer approximations.

Changes in the manuscript:

Section 4.3.1:

Original:

In clear-sky regions, the joint comparison of CERES SSF, CERES CRS, and CSRFormer (Fig. 8) in clear-sky regions shows that CSRFormer remains closer to CERES SSF than CRS in both shortwave and longwave fluxes. These differences are consistent with the fact that CRS is based on radiative-transfer calculations with prescribed input fields, whereas CSRFormer is trained directly against CERES SSF observations. One possible contributor is the use of radiative-transfer approximations in CRS, although the present comparison does not isolate that effect explicitly. This mathematical approximation inevitably introduces truncation errors, making it difficult for simulated values to precisely restore complex instantaneous radiation fields under certain atmospheric states (Liou, 1993; Scott et al., 2022). To further clarify the sources of shortwave radiative flux

675 biases from a physical mechanism perspective, in addition to the limitations of the algorithm itself, it is also necessary to trace back to the key input data driving the model's operation.

When the comparison is extended to all-sky regions, including cloudy pixels where clear-sky fluxes are not directly observed, the differences between CSRFormer and CRS become more pronounced (Fig. 9). The larger difference may reflect the combined effects of radiative-transfer approximations and differences in input data sources. First, CRS's stream approximation methods may contribute to truncation errors under complex, multiple-scattering atmospheric scenarios, widening the flux deviation. Second, shortwave flux discrepancies are likely influenced by differences in aerosol inputs. While both models align in clear skies (CRS uses MODIS AOD; CSRFormer uses MODIS-assimilated MERRA-2), under clouds, CRS switches to MATCH AOD (Scott et al., 2022). Because MATCH AOD is reported to be higher than MERRA-2 in some high-aerosol regions, this may contribute to the higher CRS shortwave clear-sky fluxes found there (Fillmore et al., 2022). To ensure training-prediction stability and consistency, CSRFormer uses unified MERRA-2 aerosols rather than the mixed "observation + simulation" approach of CRS (Scott et al., 2022). Unlike MATCH, which relies solely on emission and transport simulations, MERRA-2's advanced assimilation of MODIS and AERONET data corrects model drift, ensuring greater spatiotemporal accuracy (Buchard et al., 2017; Gueymard and Yang, 2020; Su et al., 2023; Randles et al., 2017). Third, the models treat sub-cloud surface boundary conditions differently. Unable to directly observe the surface through clouds, CRS relies on static or empirical surface albedos that fail to capture instantaneous sea state changes, such as wind-driven roughness (Scott et al., 2022). Conversely, by inputting daily SST and hourly wind speed, CSRFormer implicitly learns these dynamic impacts on surface albedo. This may reduce some uncertainties associated with static parameterizations, although that effect is not isolated directly in the present comparison.

695

Revised:

In clear-sky regions, the joint comparison of CERES SSF, CERES CRS, and CSRFormer (Fig. 8) shows that CSRFormer agrees more closely with CERES SSF than CRS in both SW and LW fluxes. This strong consistency with SSF is an important characteristic of CSRFormer, because the model is trained directly against CERES SSF clear-sky observations and is therefore explicitly optimized to reproduce instantaneous footprint-scale observed radiative fluxes. As CERES SSF provides one of the closest available observational references for instantaneous TOA radiative fluxes at the footprint scale, this agreement demonstrates the ability of the observation-trained framework to remain closely anchored to satellite observations.

705 CRS follows a different methodological pathway, independently calculating radiative fluxes using the Fu–Liou radiative-transfer model and prescribed atmospheric and surface inputs. Differences between CRS and CERES SSF may therefore arise from multiple sources, including uncertainties and differences in the driving atmospheric and surface fields, as well as approximations inherent in the radiative-transfer calculation (Liou, 1993; Scott et al., 2022). In particular, finite-stream and

other numerical approximations may introduce errors in the representation of complex radiative-transfer processes. The present comparison does not isolate the relative contributions of input uncertainty and radiative-transfer approximation, and the CRS–SSF differences should therefore be interpreted as the combined outcome of these factors.

When the comparison is extended to all-sky regions, including cloudy footprints where the corresponding clear-sky flux cannot be directly observed, the differences between CSRFormer and CRS become more pronounced (Fig. 9). These differences likely reflect the combined effects of uncertainties and differences in the driving inputs and approximations associated with the respective modeling frameworks. Under cloudy conditions, differences in aerosol representation, meteorological fields, and surface boundary conditions may become particularly important, while radiative-transfer approximations may also contribute under more complex scattering environments. First, CRS's stream approximation methods may contribute to truncation errors under complex, multiple-scattering atmospheric scenarios, widening the flux deviation. Second, shortwave flux discrepancies are likely influenced by differences in aerosol inputs. While both models align in clear skies (CRS uses MODIS AOD; CSRFormer uses MODIS-assimilated MERRA-2), under clouds, CRS switches to MATCH AOD (Scott et al., 2022). Because MATCH AOD is reported to be higher than MERRA-2 in some high-aerosol regions, this may contribute to the higher CRS shortwave clear-sky fluxes found there (Fillmore et al., 2022). To ensure training-prediction stability and consistency, CSRFormer uses unified MERRA-2 aerosols rather than the mixed "observation + simulation" approach of CRS (Scott et al., 2022). Unlike MATCH, which relies solely on emission and transport simulations, MERRA-2's advanced assimilation of MODIS and AERONET data corrects model drift, ensuring greater spatiotemporal accuracy (Buchard et al., 2017; Gueymard and Yang, 2020; Su et al., 2023; Randles et al., 2017). Third, the models treat sub-cloud surface boundary conditions differently. Unable to directly observe the surface through clouds, CRS relies on static or empirical surface albedos that fail to capture instantaneous sea state changes, such as wind-driven roughness (Scott et al., 2022). Conversely, by inputting daily SST and hourly wind speed, CSRFormer implicitly learns these dynamic impacts on surface albedo. This may reduce some uncertainties associated with static parameterizations, although that effect is not isolated directly in the present comparison.

Section 4.4.2:

Original:

The spatial distribution of annual mean CRE differences is shown in Fig. 15. For SWCRE, the largest differences are concentrated in major high-aerosol regions, including the North African dust outflow, the Indian Peninsula, and eastern China. In these regions, CRS yields weaker-magnitude SWCRE than CSRFormer, implying a higher clear-sky reflected shortwave flux in CRS and therefore a brighter clear-sky background state. This pattern is consistent with the different aerosol inputs used in the two products (Fillmore et al., 2022). Meanwhile, the study by Scott et al. (2022) also pointed out that the clear-sky shortwave radiative flux of the CRS product has a positive bias compared to the CERES SSF product, which is consistent with the comparison results in Section 4.2 of this study. From this, we can see that the overestimation tendency of the CRS product for clear-sky shortwave radiative fluxes will consequently affect the estimation of SWCRE. For LWCRE, the largest

divergence occurs in the tropics and subtropics (40° N–40° S), where CRS yields systematically lower values, implying underestimated clear-sky longwave fluxes. **This negative bias primarily stems from CRS's meteorological driving fields (GEOS-5.4.1).** In some ocean areas, these inputs tend to be "too cold" or "too wet," excessively absorbing surface emissions and suppressing the outgoing flux at the top of the atmosphere (Ham et al., 2018).

Revised:

The spatial distribution of annual mean CRE differences is shown in Fig. 15. For SWCRE, the largest differences are concentrated in major high-aerosol regions, including the North African dust outflow, the Indian Peninsula, and eastern China. In these regions, CRS yields weaker-magnitude SWCRE than CSRFormer, implying a higher clear-sky reflected shortwave flux in CRS and therefore a brighter clear-sky background state. This pattern is consistent with the different aerosol inputs used in the two products (Fillmore et al., 2022). Meanwhile, the study by Scott et al. (2022) also pointed out that the clear-sky shortwave radiative flux of the CRS product has a positive bias compared to the CERES SSF product, which is consistent with the comparison results in Section 4.2 of this study. From this, we can see that the overestimation tendency of the CRS product for clear-sky shortwave radiative fluxes will consequently affect the estimation of SWCRE. For LWCRE, the largest divergence occurs in the tropics and subtropics (40° N–40° S), where CRS yields systematically lower values, implying underestimated clear-sky longwave fluxes. **This negative bias may be partly related to CRS's meteorological driving fields (GEOS-5.4.1).** In some ocean areas, these inputs tend to be "too cold" or "too wet," excessively absorbing surface emissions and suppressing the outgoing flux at the top of the atmosphere (Ham et al., 2018).

Comments 5

Figure 7. The figure caption says PDF of all-sky conditions, yet a) and b) indicate for clear-sky fluxes.

Response:

Thank you for pointing out this ambiguity. The figure uses all CERES SSF ocean footprints, including cloudy scenes, but the quantity being compared is the clear-sky radiative flux estimated by CSRFormer and CERES CRS for those footprints. Therefore, “all-sky” in the original caption referred to the scene/sample population rather than to the radiative quantity itself. We agree that the original wording was potentially confusing. We have revised the caption and panel descriptions to explicitly state that the figure compares estimated clear-sky fluxes under all-sky conditions, including cloudy footprints.

Changes in the manuscript:

Figure 9 caption:

Original:

Figure 9. Probability density distribution plot of the comparisons of the CERES CRS and CSRFormer in all-sky areas. The closer to the diagonal line represents the more similar the two products are. (a) Comparison of clear-sky shortwave fluxes between CRS and CSRFormer; (b) Comparison of clear-sky longwave fluxes between CRS and CSRFormer.

780 Revised:

Figure 9. Probability density distributions comparing CERES CRS and CSRFormer estimated clear-sky radiative fluxes for all-sky ocean footprints, including cloudy conditions. (a) Comparison of clear-sky shortwave flux estimates from CRS and CSRFormer; (b) comparison of clear-sky longwave flux estimates from CRS and CSRFormer.

785 **Comments 6**

Section 4.3.2 and Fig. 9 and 10. What SYN1deg clear-sky fluxes are compared here as well as only during Terra/Aqua overpass times? Please make this clear in the text and in the Fig. 9 and 10 captions

790 If it is the SYN1deg clear-sky observed flux (not computed). The SYN1deg product clear-sky observed fluxes are strictly based on CERES instrument observations and are temporally interpolated using albedo models and linear interpolation between 12-hourly CERES observations over oceans. These are based on the same 20-km SSF footprint but spatially averaged over 1° region. If the comparisons are made only at the CERES observation times from the SYN1deg product (not specifically mentioned in the text), then it is either based on the SSF L2 observed or the temporally interpolated. The temporally interpolate fluxes assume the same atmospheric conditions as during the observations as stated in the text (line 465).

795

Response:

Thank you for this clarification. We agree that the temporal sampling used in the SYN1deg comparison was not described clearly enough. We note that the original manuscript already stated in Section 2.3 that we used the SYN1deg “observed TOA fluxes”, rather than the Fu–Liou “computed TOA fluxes” also provided in the SYN1deg product. However, we did not
800 explicitly describe how the hourly SYN1deg fields were temporally matched to the Aqua observations.

In our comparison, the SYN1deg observed clear-sky TOA shortwave and longwave fluxes were used. For each Aqua CERES SSF footprint, we selected the SYN1deg hourly field corresponding to the same UTC hour, rather than using the complete 24-hour SYN1deg diurnal record. The CSRFormer/SSF samples were then aggregated to the 1° SYN1deg grid for comparison.

805

We also agree with the reviewer that the term “observed” requires careful interpretation. The SYN1deg observed clear-sky fluxes are anchored to CERES instrument observations, but the 1° , hourly fields are constructed through spatial averaging and temporal interpolation between available CERES observations. Therefore, depending on the availability of CERES

observations for a given grid cell and hour, the SYN1deg value may be directly constrained by contemporaneous observations
810 or may contain a temporally interpolated component. We have revised Section 3.3, Section 4.3.2, and the corresponding figure
captions to make both the flux definition and the temporal matching explicit.

Changes in the manuscript:

Section 3.3:

815 Original:

In addition to internal validation, this study also conducted a systematic cross-comparison between CSRFormer and
mainstream radiation products. Comparison targets included physically-based model products (CERES CRS), satellite
synoptic products (CERES SYN1deg), and reanalysis data (ERA5, MERRA-2). For the CRS product, we performed direct
pixel-to-pixel comparisons (Scott et al., 2022); for other gridded products, the CSRFormer results were resampled to a $1^\circ \times 1^\circ$
820 grid for comparison. To ensure fairness in the comparison, we strictly selected pure ocean grids free of sea ice and unaffected
by land/island edge effects using the ocean fraction and snow/ice fraction data from SYN1deg during the gridded comparison
(Doelling et al., 2013).

Revised:

825 In addition to internal validation, this study also conducted a systematic cross-comparison between CSRFormer and
mainstream radiation products. Comparison targets included physically-based model products (CERES CRS), satellite
synoptic products (CERES SYN1deg), and reanalysis data (ERA5, MERRA-2). For the CRS product, we performed direct
pixel-to-pixel comparisons (Scott et al., 2022); for other gridded products, the CSRFormer results were resampled to a $1^\circ \times 1^\circ$
grid for comparison. **For the SYN1deg comparison, the hourly SYN1deg field corresponding to the UTC hour of each Aqua
830 CERES SSF footprint was selected. The corresponding CSRFormer and CERES SSF samples were then aggregated to the 1°
SYN1deg grid before comparison. Thus, the comparison uses SYN1deg fields temporally matched to the Aqua observation
hour rather than the complete diurnal SYN1deg record.** To ensure fairness in the comparison, we strictly selected pure ocean
grids free of sea ice and unaffected by land/island edge effects using the ocean fraction and snow/ice fraction data from
SYN1deg during the gridded comparison (Doelling et al., 2013).

835

Section 4.3.2:

Original:

Comparing CSRFormer with the CERES SYN1deg product (Fig. 11) reveals that in clear-sky regions, SYN exhibits extremely
high consistency with the CERES SSF ground truth. This agreement stems directly from SYN's flux generation algorithms.
840 For shortwave radiation, SYN adopts a diurnal variation model based on SZA and surface type to precisely capture solar
geometric variations (Doelling et al., 2013) . For longwave radiation, given its gentle temporal variation, SYN employs linear
interpolation between adjacent CERES observations (Doelling et al., 2016) , making its clear-sky longwave flux virtually

identical to instantaneous SSF observations. The minor discrepancies observed are primarily attributed to spatial mismatches, as gridding Level 2 SSF footprint data into Level 3 SYN data ($1^\circ \times 1^\circ$) inevitably introduces smoothing errors.

845

Revised:

We compared CSRFormer with the CERES SYN1deg observed clear-sky TOA fluxes, using the SYN1deg hourly field corresponding to the UTC hour of each Aqua CERES SSF observation. Although these SYN1deg fields are observation based, they are not necessarily direct SSF measurements at every grid cell and hour because the product combines the underlying CERES observations through 1° albedo models, spatial averaging and temporal interpolation. Comparing CSRFormer with the CERES SYN1deg product (Fig. 11) reveals that in clear-sky regions, SYN exhibits extremely high consistency with the CERES SSF ground truth. This agreement stems directly from SYN's flux generation algorithms. For shortwave radiation, SYN adopts a diurnal variation model based on SZA and surface type to precisely capture solar geometric variations (Doelling et al., 2013). For longwave radiation, given its gentle temporal variation, SYN employs linear interpolation between adjacent CERES observations (Doelling et al., 2016), making its clear-sky longwave flux virtually identical to instantaneous SSF observations. The minor discrepancies observed are primarily attributed to spatial mismatches, as gridding Level 2 SSF footprint data into Level 3 SYN data ($1^\circ \times 1^\circ$) inevitably introduces smoothing errors.

850

855

Figure 11 caption:

860 Original:

Figure 11. Probability density distribution plot of the comparisons of the CERES SSF, CERES SYN and CSRFormer in clear-sky areas. The closer to the diagonal line represents the more similar the two products are. (a) Comparison of clear-sky shortwave fluxes between SYN and CSRFormer; (b) Comparison of clear-sky shortwave fluxes between SYN and SSF; (c) Comparison of clear-sky shortwave fluxes between SSF and CSRFormer; (d) Comparison of clear-sky longwave fluxes between SYN and CSRFormer; (e) Comparison of clear-sky longwave fluxes between SYN and SSF; (f) Comparison of clear-sky longwave fluxes between SSF and CSRFormer.

865

Revised:

Figure 11. Probability density distributions comparing CERES SYN1deg observed clear-sky TOA fluxes, CERES SSF observations, and CSRFormer predictions in clear-sky regions. The SYN1deg hourly field corresponding to the UTC hour of each Aqua CERES SSF observation is used, with the corresponding SSF and CSRFormer samples aggregated to the 1° SYN1deg grid. (a) Comparison of clear-sky shortwave fluxes between SYN and CSRFormer; (b) Comparison of clear-sky shortwave fluxes between SYN and SSF; (c) Comparison of clear-sky shortwave fluxes between SSF and CSRFormer; (d) Comparison of clear-sky longwave fluxes between SYN and CSRFormer; (e) Comparison of clear-sky longwave fluxes between SYN and SSF; (f) Comparison of clear-sky longwave fluxes between SSF and CSRFormer.

870

875

Figure 12 caption:

Original:

Figure 12. Probability density distribution plot of the comparisons of the CERES SYN and CSRFormer in all-sky areas. The closer to the diagonal line represents the more similar the two products are. (a) Comparison of clear-sky shortwave fluxes between SYN and CSRFormer; (b) Comparison of clear-sky longwave fluxes between SYN and CSRFormer.

Revised:

Figure 12. Probability density distributions comparing CSRFormer predictions with CERES SYN1deg observed clear-sky TOA fluxes for all CERES SSF ocean footprints, including cloudy conditions. The SYN1deg hourly field corresponding to the UTC hour of each Aqua observation is used. (a) Comparison of clear-sky shortwave fluxes between SYN and CSRFormer; (b) Comparison of clear-sky longwave fluxes between SYN and CSRFormer.

Comments 7

Conclusions. The authors trained 2019 and 2020 to compare against 2019 and 2020. If the authors were to process the whole 22-year SSF record is the training for 2019 and 2020 sufficient to apply to 2000 with confidence? Would CSRFormer fluxes capture the trends in the input datasets?

Response:

Thank you for raising this important point. The current CSRFormer v1.0 release covers only 2018–2020, and the present study does not attempt to construct or evaluate a multi-decadal clear-sky radiative flux record. As already stated in Section 4.5, the current dataset is intended for instantaneous Aqua-overpass-time analyses and should not be used for long-term trend detection or climatological change assessment.

For a potential future extension to the full CERES record, we agree that training on 2019–2020 alone would not automatically guarantee the same level of reliability for much earlier years such as 2000. Such an application would require the atmospheric, aerosol, and surface states in earlier years to remain sufficiently represented by the training distribution, and the learned relationship between the input states and clear-sky radiative fluxes to remain stable over time. If substantial distributional changes occurred over the two decades, additional training or temporal validation would be required.

Similarly, CSRFormer predictions would respond to long-term changes in the driving input datasets, but this alone would not guarantee that the resulting trends represent true radiative trends, because changes in the inputs themselves or extrapolation beyond the training distribution could also affect the predictions. Assessing such long-term trend fidelity would therefore require a dedicated multi-decadal evaluation, which is beyond the scope of the present 2018–2020 data release.

Changes in the manuscript:

Section 5:

Added:

915 Overall, CSRFormer provides a practical framework for generating observation-time-matched clear-sky radiative flux estimates over the global ocean. Its main strengths are the use of vertically resolved inputs, the alignment with CERES overpass time, and the strong agreement with CERES SSF in clear-sky conditions. The main caveats are that validation beneath clouds is indirect, the clear-sky training population does not fully cover all atmospheric regimes encountered under cloudy conditions, the model inherits limitations from its inputs and training target, and three-dimensional radiative effects are not explicitly
920 represented. The present model has been trained and evaluated only for the 2018–2020 release, and its transferability to substantially earlier portions of the CERES record has not yet been established. Extension to a multi-decadal record would require explicit temporal out-of-distribution evaluation and, where necessary, additional training; likewise, long-term trends in CSRFormer fluxes cannot be assumed to be reliable solely because the driving datasets contain corresponding trends. With those limitations stated clearly, the dataset could be useful for future studies of CRE, cloud regimes, and cloud-clear transition
925 zones, and it may also be extendable to Terra in future work.

930

935

940

Reference:

- Allan, R. P., Shine, K. P., Slingo, A., and Pamment, J. A.: The dependence of clear-sky outgoing long-wave radiation on surface temperature and relative humidity, *Quarterly Journal of the Royal Meteorological Society*, 125, 2103-2126, 10.1002/qj.49712555809, 2007.
- 950 Bond, T. C., Doherty, S. J., Fahey, D. W., Forster, P. M., Bernsten, T., DeAngelo, B. J., Flanner, M. G., Ghan, S., Kärcher, B., Koch, D., Kinne, S., Kondo, Y., Quinn, P. K., Sarofim, M. C., Schultz, M. G., Schulz, M., Venkataraman, C., Zhang, H., Zhang, S., Bellouin, N., Guttikunda, S. K., Hopke, P. K., Jacobson, M. Z., Kaiser, J. W., Klimont, Z., Lohmann, U., Schwarz, J. P., Shindell, D., Storelvmo, T., Warren, S. G., and Zender, C. S.: Bounding the role of black carbon in the climate system: A scientific assessment, *Journal of Geophysical Research: Atmospheres*, 118, 5380-5552, 10.1002/jgrd.50171, 2013.
- 955 Buchard, V., Randles, C. A., da Silva, A. M., Darmenov, A., Colarco, P. R., Govindaraju, R., Ferrare, R., Hair, J., Beyersdorf, A. J., Ziemba, L. D., and Yu, H.: The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part II: Evaluation and Case Studies, *Journal of Climate*, 30, 6851-6872, 10.1175/jcli-d-16-0613.1, 2017.
- Cao, Y., Zhu, Y., Wang, M., Rosenfeld, D., Zhou, C., Liu, J., Liang, Y., Huang, K. E., Wang, Q., Bai, H., Wang, Y., Wang, H., and Zhang, H.: Improving Prediction of Marine Low Clouds Using Cloud Droplet Number Concentration in a Convolutional Neural Network, *Journal of Geophysical Research: Machine Learning and Computation*, 1, 10.1029/2024jh000355, 2024.
- 960 Doelling, D. R., Sun, M., Nguyen, L. T., Nordeen, M. L., Haney, C. O., Keyes, D. F., and Mlynchzak, P. E.: Advances in Geostationary-Derived Longwave Fluxes for the CERES Synoptic (SYN1deg) Product, *Journal of Atmospheric and Oceanic Technology*, 33, 503-521, 10.1175/jtech-d-15-0147.1, 2016.
- 965 Doelling, D. R., Loeb, N. G., Keyes, D. F., Nordeen, M. L., Morstad, D., Nguyen, C., Wielicki, B. A., Young, D. F., and Sun, M.: Geostationary Enhanced Temporal Interpolation for CERES Flux Products, *Journal of Atmospheric and Oceanic Technology*, 30, 1072-1090, <https://doi.org/10.1175/JTECH-D-12-00136.1>, 2013.
- Eytan, E., Gristey, J. J., and Feingold, G.: The Net Radiative Effect of the Ill-Defined Clear-Sky in the Vicinity of Clouds, *AGU Advances*, 6, 10.1029/2024av001407, 2025.
- 970 Fillmore, D. W., Rutan, D. A., Kato, S., Rose, F. G., and Caldwell, T. E.: Evaluation of aerosol optical depths and clear-sky radiative fluxes of the CERES Edition 4.1 SYN1deg data product, *Atmospheric Chemistry and Physics*, 22, 10115-10137, 10.5194/acp-22-10115-2022, 2022.
- Frey, R. A., Ackerman, S. A., Liu, Y., Strabala, K. I., Zhang, H., Key, J. R., and Wang, X.: Cloud Detection with MODIS. Part I: Improvements in the MODIS Cloud Mask for Collection 5, *Journal of Atmospheric and Oceanic Technology*, 25, 1057-1072, <https://doi.org/10.1175/2008JTECHA1052.1>, 2008.
- 975 Gentemann, C. L., Donlon, C. J., Stuart-Menteth, A., and Wentz, F. J.: Diurnal signals in satellite sea surface temperature measurements, *Geophysical Research Letters*, 30, 10.1029/2002gl016291, 2003.
- Gueymard, C. A. and Yang, D.: Worldwide validation of CAMS and MERRA-2 reanalysis aerosol optical depth products using 15 years of AERONET observations, *Atmospheric Environment*, 225, 10.1016/j.atmosenv.2019.117216, 2020.
- 980 Ham, S.-H., Su, W., Smith, W. L., Huang, X., Doelling, D. R., Loeb, N. G., Thorsen, T. J., Rutan, D. A., Rose, F. G., and Kato, S.: Surface Irradiances of Edition 4.0 Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Data Product, *Journal of Climate*, 31, 4501-4527, 10.1175/jcli-d-17-0523.1, 2018.
- Hansen, J., Russell, G., Rind, D., Stone, P., Lacis, A., Lebedeff, S., Ruedy, R., and Travis, L.: Efficient Three-Dimensional Global Models for Climate Studies: Models I and II, *Monthly Weather Review*, 111, 609-662, 10.1175/1520-0493(1983)111<0609:ETDGMF>2.0.CO;2, 1983.
- 985 Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, *Quarterly Journal of the Royal Meteorological Society*, 146, 1999-2049, 10.1002/qj.3803, 2020.
- Jin, Z., Charlock, T. P., Smith, W. L., and Rutledge, K.: A parameterization of ocean surface albedo, *Geophysical Research Letters*, 31, 10.1029/2004gl021180, 2004.

- Koll, D. D. and Cronin, T. W.: Earth's outgoing longwave radiation linear due to H₂O greenhouse effect, Proceedings of the National Academy of Sciences, 115, 10293-10298, 2018.
- Liou, K.: Parameterization of the radiative properties of cirrus clouds, Journal of the atmospheric sciences, 50, 2008-2025, 1993.
- Loeb, N. G. and Davies, R.: Observational evidence of plane parallel model biases: Apparent dependence of cloud optical depth on solar zenith angle, Journal of Geophysical Research: Atmospheres, 101, 1621-1634, 10.1029/95jd03298, 1996.
- Loeb, N. G., Kato, S., Loukachine, K., and Manalo-Smith, N.: Angular Distribution Models for Top-of-Atmosphere Radiative Flux Estimation from the Clouds and the Earth's Radiant Energy System Instrument on the Terra Satellite. Part I: Methodology, Journal of Atmospheric and Oceanic Technology, 22, 338-351, <https://doi.org/10.1175/JTECH1712.1>, 2005.
- Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., Liang, L., Mitrescu, C., Rose, F. G., and Kato, S.: Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Top-of-Atmosphere (TOA) Edition-4.0 Data Product, Journal of Climate, 31, 895-918, 10.1175/jcli-d-17-0208.1, 2018.
- Loshchilov, I. and Hutter, F.: Decoupled weight decay regularization, arXiv preprint arXiv:1711.05101, 2017.
- Minnis, P., Sun-Mack, S., Chen, Y., Chang, F. L., Yost, C. R., Smith, W. L., Heck, P. W., Arduini, R. F., Bedka, S. T., Yi, Y., Hong, G., Jin, Z., Painemal, D., Palikonda, R., Scarino, B. R., Spangenberg, D. A., Smith, R. A., Trepte, Q. Z., Yang, P., and Xie, Y.: CERES MODIS Cloud Product Retrievals for Edition 4—Part I: Algorithm Changes, IEEE Transactions on Geoscience and Remote Sensing, 59, 2744-2780, 10.1109/TGRS.2020.3008866, 2021.
- Payne, R. E.: Albedo of the Sea Surface, Journal of Atmospheric Sciences, 29, 959-970, [https://doi.org/10.1175/1520-0469\(1972\)029<0959:AOTSS>2.0.CO;2](https://doi.org/10.1175/1520-0469(1972)029<0959:AOTSS>2.0.CO;2), 1972.
- Platnick, S., Meyer, K. G., King, M. D., Wind, G., Amarasinghe, N., Marchant, B., Arnold, G. T., Zhang, Z., Hubanks, P. A., Holz, R. E., Yang, P., Ridgway, W. L., and Riedi, J.: The MODIS Cloud Optical and Microphysical Products: Collection 6 Updates and Examples From Terra and Aqua, IEEE Transactions on Geoscience and Remote Sensing, 55, 502-525, 10.1109/TGRS.2016.2610522, 2017.
- Prata, A. J.: A new long-wave formula for estimating downward clear-sky radiation at the surface, Quarterly Journal of the Royal Meteorological Society, 122, 1127-1151, 10.1002/qj.49712253306, 2006.
- Randles, C. A., Da Silva, A. M., Buchard, V., Colarco, P. R., Darmenov, A., Govindaraju, R., Smirnov, A., Holben, B., Ferrare, R., Hair, J., Shinozuka, Y., and Flynn, C. J.: The MERRA-2 Aerosol Reanalysis, 1980 - onward, Part I: System Description and Data Assimilation Evaluation, J Clim, 30, 6823-6850, 10.1175/JCLI-D-16-0609.1, 2017.
- Samset, B. H., Myhre, G., Schulz, M., Balkanski, Y., Bauer, S., Bernsten, T. K., Bian, H., Bellouin, N., Diehl, T., Easter, R. C., Ghan, S. J., Iversen, T., Kinne, S., Kirkevåg, A., Lamarque, J. F., Lin, G., Liu, X., Penner, J. E., Seland, Ø., Skeie, R. B., Stier, P., Takemura, T., Tsigaridis, K., and Zhang, K.: Black carbon vertical profiles strongly affect its radiative forcing uncertainty, Atmospheric Chemistry and Physics, 13, 2423-2434, 10.5194/acp-13-2423-2013, 2013.
- Scott, R. C., Rose, F. G., Stackhouse, P. W., Loeb, N. G., Kato, S., Doelling, D. R., Rutan, D. A., Taylor, P. C., and Smith, W. L.: Clouds and the Earth's Radiant Energy System (CERES) Cloud Radiative Swath (CRS) Edition 4 Data Product, Journal of Atmospheric and Oceanic Technology, 39, 1781-1797, 10.1175/jtech-d-22-0021.1, 2022.
- Smith, G. L., Priestley, K. J., Loeb, N. G., Wielicki, B. A., Charlock, T. P., Minnis, P., Doelling, D. R., and Rutan, D. A.: Clouds and Earth Radiant Energy System (CERES), a review: Past, present and future, Advances in Space Research, 48, 254-263, 10.1016/j.asr.2011.03.009, 2011.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting, The journal of machine learning research, 15, 1929-1958, 2014.
- Su, W., Corbett, J., Eitzen, Z., and Liang, L.: Next-generation angular distribution models for top-of-atmosphere radiative flux calculation from CERES instruments: validation, Atmospheric Measurement Techniques, 8, 3297-3313, 10.5194/amt-8-3297-2015, 2015a.
- Su, W., Corbett, J., Eitzen, Z., and Liang, L.: Next-generation angular distribution models for top-of-atmosphere radiative flux calculation from CERES instruments: methodology, Atmospheric Measurement Techniques, 8, 611-632, 10.5194/amt-8-611-2015, 2015b.
- Su, X., Huang, Y., Wang, L., Cao, M., and Feng, L.: Validation and diurnal variation evaluation of MERRA-2 multiple aerosol properties on a global scale, Atmospheric Environment, 311, 10.1016/j.atmosenv.2023.120019, 2023.

- Taylor, J. P., Edwards, J. M., Glew, M. D., Hignett, P., and Slingo, A.: Studies with a flexible new radiation code. II: Comparisons with aircraft short-wave observations, *Quarterly Journal of the Royal Meteorological Society*, 122, 839-861, 10.1002/qj.49712253204, 2006.
- 1045 Tegen, I., Hollrig, P., Chin, M., Fung, I., Jacob, D., and Penner, J.: Contribution of different aerosol species to the global aerosol extinction optical thickness: Estimates from model results, *Journal of Geophysical Research: Atmospheres*, 102, 23895-23915, 10.1029/97jd01864, 1997.
- Tijhuis, M., van Stratum, B. J. H., and van Heerwaarden, C. C.: The impact of coupled 3D shortwave radiative transfer on surface radiation and cumulus clouds over land, *Atmospheric Chemistry and Physics*, 24, 10567-10582, 10.5194/acp-24-10567-2024, 2024.
- 1050 Trenberth, K. E., Fasullo, J. T., and Kiehl, J.: Earth's Global Energy Budget, *Bulletin of the American Meteorological Society*, 90, 311-324, 10.1175/2008bams2634.1, 2009.
- Trepte, Q. Z., Bedka, K. M., Chee, T. L., Minnis, P., Sun-Mack, S., Yost, C. R., Chen, Y., Jin, Z., Hong, G., Chang, F.-L., and Smith, W. L.: Global Cloud Detection for CERES Edition 4 Using Terra and Aqua MODIS Data, *IEEE Transactions on Geoscience and Remote Sensing*, 57, 9410-9449, 10.1109/tgrs.2019.2926620, 2019.
- 1055 Várnai, T. and Marshak, A.: Observations of Three-Dimensional Radiative Effects that Influence MODIS Cloud Optical Thickness Retrievals, *Journal of the Atmospheric Sciences*, 59, 1607-1618, [https://doi.org/10.1175/1520-0469\(2002\)059<1607:OOTDRE>2.0.CO;2](https://doi.org/10.1175/1520-0469(2002)059<1607:OOTDRE>2.0.CO;2), 2002.
- Wang, Q., Zhou, C., Letu, H., Zhu, Y., Zhuge, X., Liu, C., Weng, F., and Wang, M.: Obtaining Cloud Base Height and Phase From Thermal Infrared Radiometry Using a Deep Learning Algorithm, *IEEE Transactions on Geoscience and Remote Sensing*, 61, 1-14, 10.1109/TGRS.2023.3317532, 2023.
- 1060 Wielicki, B. A., Barkstrom, B. R., Harrison, E. F., Lee III, R. B., Smith, G. L., and Cooper, J. E.: Clouds and the Earth's Radiant Energy System (CERES): An earth observing system experiment, *Bulletin of the American Meteorological Society*, 77, 853-868, [https://doi.org/10.1175/1520-0477\(1996\)077<0853:CATERE>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0853:CATERE>2.0.CO;2), 1996.
- 1065 Winter, B. and Chýlek, P.: Contribution of sea salt aerosol to the planetary clear-sky albedo, *Tellus B: Chemical and Physical Meteorology*, 49, 10.3402/tellusb.v49i1.15951, 1997.
- Zhang, J., Zhou, X., Goren, T., and Feingold, G.: Albedo susceptibility of northeastern Pacific stratocumulus: the role of covarying meteorological conditions, *Atmospheric Chemistry and Physics*, 22, 861-880, 10.5194/acp-22-861-2022, 2022.
- Zheng, B., Cao, Y., Huang, K. E., Liu, J., Wang, Y., Zhu, Y., Wang, M., Rosenfeld, D., Zhou, C., and Huang, Y.: Estimation of Global Ocean TOA Instantaneous Clear-Sky Albedo From CERES for Shortwave Cloud Radiative Effect Analysis Based on a Deep Learning Model, *Journal of Geophysical Research: Atmospheres*, 130, 10.1029/2025jd044878, 2025.
- 1070 Zheng, B., Zhu, Y., Cao, Y., Huang, K.-E., Liu, J., Wang, Y., Shi, J., Wei, Y., Rosenfeld, D., Zhou, C., and Wang, M.: CSRFormer: An Instantaneous Global-ocean Clear-sky Radiative Flux Dataset Derived From CERES, Zenodo [dataset], 10.5281/zenodo.20046058, 2026.

1075