https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

Open Access Earth System Science Data Discussions

CSRFormer: An Instantaneous Global-ocean Clear-sky Radiative Flux Dataset Derived From CERES

Boyang Zheng[1,2], Yannian Zhu[1,2,*] , Yang Cao[1,2,4,5], Kang-En Huang[1,2], Jihu Liu[1,2], Yichuan Wang[1,2], Jun Shi[1,2], Yicheng Wei[1,2], Daniel Rosenfeld[1,3], Chen Zhou[1,2], Minghuai Wang[1,2]

[1] School of Atmospheric Sciences, Nanjing University, Nanjing, China.
[2] Joint International Research Laboratory of Atmospheric and Earth System Sciences & Institute for Climate and Global Change Research, Nanjing University, Nanjing, China.
[3] Institute of Earth Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel.
[4] Centre for Climate Change and Environmental Health, Nanyang Technological University, Singapore, Singapore 639798
[5] Earth Observatory of Singapore, Nanyang Technological University, Singapore, Singapore 639798

Correspondence to: Yannian Zhu (yannian.zhu@nju.edu.cn)

Abstract. Accurate estimates of the Cloud Radiative Effect (CRE) require clear-sky radiative fluxes that are temporally consistent with the corresponding all-sky observations. Because passive satellite sensors cannot directly observe the sub-cloud clear-sky state, instantaneous clear-sky fluxes in cloudy regions are commonly approximated using radiative transfer calculations or temporally interpolated products, both of which introduce limitations. Here we present CSRFormer, a physically constrained deep-learning framework for estimating instantaneous clear-sky shortwave and longwave radiative fluxes over the global ocean. CSRFormer uses ERA5 atmospheric profiles, MERRA-2 aerosol properties, and NOAA sea surface temperature as inputs, and is trained against CERES Single Scanner Footprint (SSF) clear-sky observations. Over independent clear-sky samples, the model shows strong agreement with CERES SSF, with R^2 values of 0.81 for shortwave flux and 0.99 for longwave flux, corresponding to estimated uncertainties (RMSE) of ~3.2 W m^{-2} and ~2.0 W m^{-2}, respectively. Because the predictions are matched to the observation time of the CERES overpass, the resulting dataset avoids the temporal mismatch inherent in hourly mean products. Application of CSRFormer to cloudy scenes provides an observation-trained estimate of the theoretical clear-sky reference state beneath clouds for CRE analyses over the global ocean. The dataset is therefore a useful framework for examining the radiative effects of clouds and their covariation with cloud properties at the satellite overpass time. The CSRFormer v1.0 dataset is available at https://doi.org/10.5281/zenodo.20046058.

1 Introduction

Cloud shortwave and longwave radiative effects are central to the Earth's energy budget. Clouds reflect incoming solar radiation and absorb and emit terrestrial radiation, so the net Cloud Radiative Effect (CRE) depends on the balance between these opposing processes (Ramanathan et al., 1989; Trenberth et al., 2009). Accurate CRE estimates are also important for studying aerosol-cloud interactions, because changes in cloud amount, phase, and microphysics can alter both shortwave cooling and longwave warming (Allan, 2011; Bony et al., 2015; Bellouin et al., 2020; Forster, 2021).

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

Estimating CRE is challenging because the required clear-sky reference state is not directly observed beneath clouds (Eytan et al., 2025). Passive instruments such as CERES provide broad spatial coverage, but they cannot observe the sub-cloud clear-sky flux at the instant of the all-sky observation. Products such as CERES EBAF provide robust monthly mean CRE estimates (Loeb et al., 2018), whereas higher-temporal-resolution products such as SYN1deg rely on diurnal models or interpolation between nearby clear-sky observations (Doelling et al., 2013; Doelling et al., 2016). Those approaches are useful, but they are not strictly matched to the overpass time of the instantaneous cloud-property retrievals used in many process studies. Active sensors such as CloudSat and CALIPSO provide valuable vertical information, yet their narrow swaths limit their ability to provide spatially continuous instantaneous clear-sky reference fields (Henderson et al., 2013).

Model-based approaches are widely used to estimate clear-sky and all-sky radiative fluxes, but they also have limitations. Climate models inherit errors from cloud simulations themselves, including longstanding biases in marine cloud amount and optical properties (Nam et al., 2012; Konsta et al., 2022). Stand-alone 3D radiative transfer calculations can be highly accurate, but they are computationally expensive for global-scale, footprint-level applications (Pincus and Evans, 2009). Operational products therefore often use more efficient radiative transfer approximations; for example, CERES Cloud Radiative Swath (CRS) is based on the Fu-Liou radiative transfer code (Scott et al., 2022). These products are valuable, but they may differ from instantaneous observations because of approximations in the forward model and uncertainties in the input fields.

Machine-learning methods offer a complementary approach because they can learn nonlinear relationships directly from large observational or reanalysis datasets while remaining computationally efficient at inference time (Gentine et al., 2018; Grundner et al., 2022). When trained against observations, such models may also learn to compensate for some persistent input biases, although that compensation is empirical rather than guaranteed (Krasnopolsky, 2007; Rasp and Lerch, 2018; Reichstein et al., 2019). For that reason, machine learning is increasingly being explored for atmospheric prediction and radiative applications (Wang et al., 2023; Cao et al., 2024).

Despite these overarching advantages, applying generic ML architectures to atmospheric radiation may not always fully capture the inherent physical geometry and vertical thermodynamic structure of the atmosphere. For instance, Zheng et al. (2025) successfully employed a Multilayer Perceptron (MLP) model to calculate clear-sky albedo for estimating shortwave cloud radiative effects. However, MLPs are inherently limited in their ability to process continuous vertical structures. As a result, they are typically restricted to using data from a single specific altitude or column-integrated quantities as inputs, making it difficult to capture the complex, altitude-dependent physical mechanisms within atmospheric layers. To address this and more deeply embed physical processes into the learning framework, our work advances the data-driven paradigm by designing a sophisticated, physically-constrained model—the Clear-Sky Radiation Transformer (CSRFormer). Unlike MLPs, the Transformer architecture is uniquely equipped to handle sequential data (Vaswani et al., 2017), allowing it to explicitly capture

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

Open Access Earth System Science Data Discussions

the physical order of atmospheric vertical stratification (such as temperature, humidity, and aerosol extinction profiles) and their layered radiative interactions. Furthermore, to ensure thermodynamic rationality and physical consistency across varying radiative baselines, CSRFormer integrates physical hard constraints (e.g., the Stefan-Boltzmann law) alongside multi-task loss constraints (Beucler et al., 2021; Karniadakis et al., 2021). Guided by these physical principles, CSRFormer is designed to estimate both shortwave and longwave clear-sky radiative fluxes simultaneously.

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 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 application. Section 5 summarizes the main findings and limitations.

2 Data

The datasets used in this study fall into three categories: CERES observations used as the training target, meteorological and aerosol inputs used to drive the Transformer model, and external products used for evaluation and comparison.

2.1 Observational data from CERES

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 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, this dataset provides detailed solar zenith angle (SZA), meteorological information from Goddard Earth Observing

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

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.

2.2 Input variables

We selected input variables that describe both surface conditions and the vertical atmospheric state. For convenience, we group them into planar inputs and vertical-profile inputs. Clear-sky samples from 2019 and 2020 were divided into a training set and a validation set at a ratio of 7:3, while data from 2018 were used as the test set.

Planar Inputs are primarily used to describe surface properties and the aerosol load of the entire atmospheric column. Sea Surface Temperature (SST) is an important factor affecting the ocean surface radiation balance; on one hand, it determines the baseline for surface longwave flux emission, and on the other hand, it implicitly contains partial ocean surface optical properties (such as ocean color changes related to ocean productivity) (Jin et al., 2004). In this study, SST data are obtained from the high-resolution SST product of the National Oceanic and Atmospheric Administration (NOAA): the NOAA 1/4° Daily Optimum Interpolation Sea Surface Temperature (OISST) (Huang et al., 2021). Aerosol Optical Depth (AOD) reflects the extinction effect of aerosols and has a significant impact on clear-sky radiative fluxes (Tegen et al., 1997). Since satellite-based AOD products inherently suffer from missing data in cloud-covered areas, we chose to use the AOD product provided by the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) to achieve clear-sky flux prediction in cloud cover regions (Randles et al., 2017; Gelaro et al., 2017). MERRA-2 provides spatially complete aerosol fields and assimilates observational constraints such as MODIS, while providing physically consistent AOD simulations in cloud-covered areas, thereby filling observational gaps and making it possible to consider aerosol extinction effects under all weather conditions (Su et al., 2023; Gueymard and Yang, 2020). In addition, we introduced wind speed as auxiliary variables, since sea surface wind speed alters sea surface roughness, thereby affecting sea surface shortwave albedo; this data is provided by the CERES SSF product (Hansen et al., 1983; Winter and Chýlek, 1997; Jin et al., 2004). Furthermore, although SZA belongs to geometric parameters, it plays a decisive role in albedo; this study directly uses the SZA data included in CERES SSF to ensure precise temporal matching (Payne, 1972; Taylor et al., 2006; Zhang et al., 2022).

Vertical Inputs aim to capture the thermodynamic and compositional structure of the atmosphere. The primary meteorological field inputs in this study are derived from the European Centre for Medium-range Weather Forecasts (ECMWF) Reanalysis v.5 (ERA5) (Hersbach et al., 2020). ERA5 is based on the Integrated Forecasting System (IFS) Cy41r2 and provides a comprehensive description of the global atmospheric state by assimilating massive amounts of historical observational data. Since longwave radiative transfer is highly dependent on temperature stratification and the vertical distribution of greenhouse gases, this study uses vertical profile data of temperature, humidity, and ozone provided by ERA5, thereby enabling the precise capture of the impact of the entire vertical air column on longwave radiative transfer, rather than just considering the effects

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

of certain key layers. Given the distinct roles that absorbing and scattering aerosols play in radiative transfer, relying solely on AOD is insufficient to characterize their radiative effects (Samset et al., 2013; Bond et al., 2013). Thus, this study also utilized the vertical profile data of absorbing aerosols provided by MERRA-2, which includes two major absorbing aerosols: black carbon (BCPHILIC and BCPHOBIC) and dust (DU001, DU002, DU003, DU004 and DU005) (Gelaro et al., 2017; Randles et al., 2017). This data is generated based on the GEOS-5 Earth System Model and the Goddard Chemistry Aerosol Radiation and Transport (GOCART) aerosol module, capable of providing three-dimensional distribution information of various aerosol components including black carbon and dust, which is of great significance for capturing the radiative forcing effects of aerosols.

2.3 Validation data

To comprehensively and objectively evaluate the accuracy and generalization capability of the CSRFormer model, we selected several radiative products widely used as validation benchmarks, covering physically-based model products, satellite observation Level-3 products, and reanalysis data products. The CERES Cloud Radiative Swath (CRS) product is calculated using the NASA Langley Fu-Liou radiative transfer code (Scott et al., 2022). It utilizes cloud and aerosol properties from MODIS and meteorological fields from GEOS reanalysis data to simulate multilayer radiative fluxes from the surface to the top of the atmosphere. CERES Synoptic 1-degree (SYN1deg) aims to provide a Level-3 radiative product with high temporal resolution (1 hour) (Doelling et al., 2013; Doelling et al., 2016). In this research, we used SYN1deg's "observed TOA fluxes", which generates clear-sky shortwave fluxes based on diurnal variation models dependent on SZA and surface type, and generates clear-sky longwave fluxes using temporal interpolation techniques. Additionally, the SYN1deg product includes "computed TOA fluxes" calculated using the Fu-Liou radiative transfer model. Since this methodology is highly similar to that of the CERES CRS product, the comparison results for this model-based approach will be represented by our evaluation of the CRS product. As representatives of two mainstream methods for estimating clear-sky radiative fluxes, these two products can effectively verify the accuracy and reliability of the CSRFormer model. In addition, we also compared the radiative fluxes provided by ERA5 and MERRA-2 reanalysis data (Gelaro et al., 2017; Hersbach et al., 2020). The radiative fluxes in these two datasets are typically calculated by internal forecast models based on physical parameterization schemes (e.g., ERA5 uses the RTTOV model), and they are also widely used radiative products.

Table 1. Summary of Variables Used in This Research.

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

Variable	Unit	Data source	Resolution	Data type	Role
Aerosol Optical Depth (AOD)	-	MERRA-2	0.625°×0.5° (Vertical)	Reanalysis	Aerosol effect
Aerosol Mixing Ratio	kg kg^{-1}				Aerosol effect
Meteorological Fields	K, % or kg kg^{-1}	ERA5	0.25°×0.25° (Vertical)	Reanalysis	Thermodynamic state
Sea Surface Temperature (SST)	K	NOAA OISST	0.25°×0.25° (Planar)	Reanalysis	Surface state
Wind Speed	m s^{-1}	CERES SSF	20km (Planar)	Reanalysis	Surface roughness
Solar Zenith Angle (SZA)	°			Satellite observation	Solar geometry
Cloud Mask	%	MODIS (MYD06)	1km (Planar)	Satellite observation	Screening
Sunglint Mask	%				Screening
Cloud Mask	%	CERES SSF	20km (Planar)	Satellite observation	Screening
Clear-sky fluxes	W/m^2	CERES SSF	20km (Planar)	Satellite observation	Target
Clear-sky fluxes	W/m^2	CERES SYN	1°×1°(Planar)	Satellite observation synthesis	Benchmark
	W/m^2	CERES CRS	20km (Planar)	Radiative transfer model	Benchmark
	W/m^2	ERA5	0.25°×0.25°(Planar)	Reanalysis	Benchmark
	W/m^2	MERRA-2	0.5°×0.5°(Planar)	Reanalysis	Benchmark

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

3 Methodology

This study develops CSRFormer, a deep-learning framework with physical constraints for reconstructing clear-sky shortwave and longwave radiative fluxes from surface and atmospheric state variables. The model uses a Transformer architecture to encode vertical profile information (Vaswani et al., 2017). Additional output constraints and loss terms are introduced to reduce physically implausible predictions and to improve interpretability.

3.1 Data Pre-processing

Before training, all datasets were collocated to the CERES SSF footprint in space and time. ERA5 (0.25° × 0.25°) and MERRA-2 (0.5° × 0.625°) fields were linearly interpolated to the center of each CERES SSF footprint. Because the native temporal resolution of the MERRA-2 aerosol product is 3 h, we linearly interpolated it to 1 h before matching to CERES. All datasets were temporally collocated based on Coordinated Universal Time (UTC).

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.

3.2 CSRFormer model framework

Based on the traditional Transformer Encoder-Decoder architecture, CSRFormer has been specifically optimized for handling "planar-vertical" coupled data and specifically optimized for the physical characteristics of radiative transfer. We treat the vertical profiles (temperature, humidity, ozone, aerosols) from ERA5 (37 levels) and MERRA-2 (72 levels) as sequence data and introduce pressure positional encoding, enabling the model to explicitly perceive the upper and lower stratification structure of the atmosphere. These features pass through a 4-layer Transformer encoder utilizing 8-head self-attention mechanisms to capture radiative interactions between atmospheric layers (such as the re-absorption of lower-layer emission by upper-layer water vapor). Surface parameters (SST, wind speed) and columnar parameters (AOD, solar zenith angle, etc.) constitute the boundary conditions for radiative transfer. To achieve deep coupling between surface boundary conditions and atmospheric layers, the model first represents planar and aggregated information through fusion feature embedding tokens, and subsequently utilizes an encoder module to encode the fused sequence, thereby achieving the deep integration of spatial

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

Open Access Earth System Science Data Discussions

features and vertical structures. The "Fused Memory" generated by this design is fed into the Shortwave (SW) and Longwave (LW) decoders respectively, ensuring that the model can extract key features specific to the physical mechanisms of different wavebands. Our model structure is shown in Fig. 1.

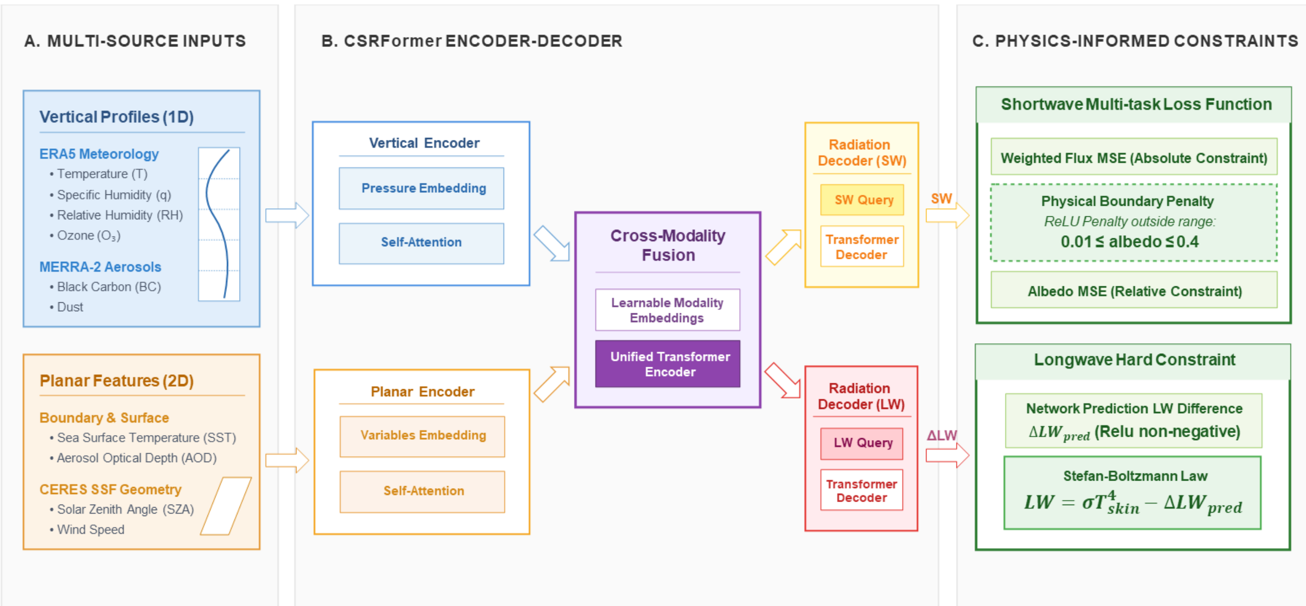


Figure 1. Schematic diagram of the CSRFormer model structure. Our main features lie in the use of vertical height encoding and the incorporation of physical constraints.

To reduce physically implausible output, we introduced output constraints and a multi-task loss function. For longwave radiation, the model predicts an atmospheric attenuation term rather than directly predicting the top-of-atmosphere longwave flux. The final longwave estimate is then expressed as a constrained function of sea-surface temperature and the predicted attenuation term:

$$LW = \sigma T_{skin}^4 - \Delta LW_{pred} \quad (1)$$

where σ is the Stefan-Boltzmann constant and Tskin is the NOAA sea-surface temperature used as an approximation to the surface skin temperature. The ReLU function constrains the attenuation term to be non-negative, which prevents the predicted longwave flux from exceeding the corresponding surface blackbody emission in this formulation.

For shortwave radiation, the large dynamic range associated with solar zenith angle and season makes a purely absolute-error loss less effective across the full domain. We therefore added an albedo-based auxiliary loss so that the optimization remains sensitive to relative errors under both weak and strong illumination. We also penalized clear-sky albedo values outside a prescribed range (0.01-0.4) and applied additional weight to high-flux samples. The complete form of the loss function is as follows:

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

$$Loss = \alpha_1 MSE(LW) + \alpha_2 MSE(SW) + \alpha_3 MSE(albedo) + \alpha_4\ Albedo\ range\ loss + \alpha_5\ High\ SW\ weighted\ loss \quad (2)$$

The weighting parameters $\alpha_n (n = 1, 2, \ldots, 5)$ are set to 7, 3, 2, 0.5 and 3, respectively. The model was trained with a learning rate of in a 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-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.

3.3 Validation and comparison

To comprehensively evaluate the accuracy of the CSRFormer model and its generalization capability in non-observed regions (i.e., cloud-covered areas), this study adopted a hierarchical validation strategy. First, direct validation was performed on an independent clear-sky test set, calculating the Root Mean Square Error (RMSE) and Coefficient of Determination (R^2) between model predictions and CERES SSF observational ground truth to quantify the reconstruction accuracy of the CSRFormer model under real clear-sky conditions. Furthermore, we further calculated the clear-sky albedo, which is obtained by dividing the clear-sky SW fluxes by the CERES SSF incoming solar radiation fluxes. Comparing the difference between model-retrieved clear-sky albedo and observed clear-sky albedo provides a more rigorous test for the physical consistency of shortwave radiation.

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 Total Column Water Vapor (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] 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.

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

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

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.

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°×1° 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).

3.4 Application

Based on the all-sky theoretical clear-sky fluxes generated by the CSRFormer model, we can directly calculate the Cloud Radiative Effect (CRE). CRE is defined as the difference between all-sky radiative flux ($F_{all\text{-}sky}$) and clear-sky radiative flux ($F_{clear\text{-}sky}$). Using the all-sky fluxes observed by CERES and the clear-sky fluxes inferred by CSRFormer, the calculation formulas are as follows:

$$CRE_{SW} = F^{SW}_{clear-sky} - F^{SW}_{all-sky} \quad (3)$$

$$CRE_{LW} = F^{LW}_{clear-sky} - F^{LW}_{all-sky} \quad (4)$$

Through this application, we are able to generate an Aqua-overpass-time CRE estimate for analyzing the contributions of different cloud systems to the Earth's energy balance. To deeply investigate the differentiated contributions of cloud systems with different vertical structures to global radiative forcing, this study used a vertical stratified classification system containing six cloud types based on instantaneous Cloud Top Pressure (CTP) and Cloud Base Pressure (CBP) data retrieved by MODIS within the CERES SSF product (Platnick et al., 2017). This system uses 680 hPa and 420 hPa as boundaries to divide the atmosphere into three layers: low, middle, and high, and defines the following categories based on the vertical extent of the cloud layers, as shown in Table 2 (Dietel et al., 2024).

Table 2. Definition of different cloud types.

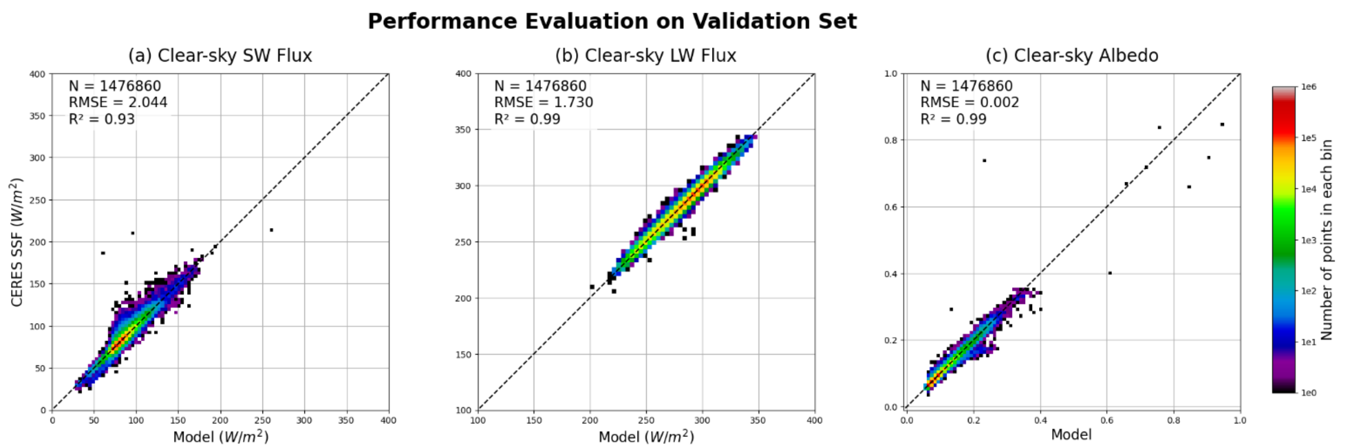
https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

Type	Cloud Top Pressure	Cloud Base Pressure
Low-level Clouds	CTP > 680 hPa	CBP > 680 hPa
Mid-Low Clouds	420 hPa < CTP ≤ 680 hPa	CBP > 680 hPa
Deep Convective Clouds	CTP ≤ 420 hPa	CBP > 680 hPa
Mid-level Clouds	CTP > 420 hPa	CBP ≤ 680 hPa
High-Mid Clouds	CTP ≤ 420 hPa	CBP ≤ 680 hPa
High-level Clouds	CTP ≤ 420 hPa	CBP ≤ 420 hPa

4 Results

4.1 Accuracy assessment

We evaluated CSRFormer by comparing predicted clear-sky radiative fluxes with CERES SSF clear-sky reference values at the pixel level for the validation and independent test sets. Figures 2 and 3 display the scatter density distributions and error statistics of the model on the Validation Set and Test Set, respectively. Across approximately 1.4 million samples, predictions for both SW and LW fluxes clustered closely around the 1:1 line. For clear-sky shortwave radiation in validation set (Fig. 2a), the model predictions were highly consistent with observations, with a R^2 as high as 0.93 and a RMSE of only 2.044 W/m^2, showing no obvious overestimation or underestimation. For clear-sky longwave radiation (Fig. 2b), the model also performed well, with R^2 reaching 0.99 and RMSE further reducing to 1.730 W/m^2. For the calculated clear-sky albedo (Fig. 2c), R^2 reached 0.99 with an RMSE of 0.002. This result indicates that the current model performance has exceeded that of the prediction of clear-sky albedo achieved based on an MLP model in Zheng et al. (2025) (R^2 of 0.96, RMSE of 0.004).



https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

Figure 2. Probability density distributions for the validation set. The x-axis shows CSRFormer predictions and the y-axis shows CERES SSF clear-sky reference values. Greater proximity to the 1:1 line indicates better agreement. (a) Clear-sky shortwave flux; (b) clear-sky longwave flux; (c) clear-sky albedo.

In the independent test set not involved in training (Fig. 3a), SW performance decreased modestly (R^2 = 0.81, RMSE = 3.207 W m^{-2}), likely reflecting the higher sensitivity of shortwave fluxes to aerosol and surface conditions. Additionally, 3D cloud radiative effects—specifically, 'brightening' from lateral scattering by adjacent clouds—contribute to shortwave prediction biases. While CERES faithfully records these non-local enhancements (Loeb and Davies, 1996; Várnai and Marshak, 2002; Tijhuis et al., 2024), CSRFormer relies on single-point atmospheric column inputs lacking surrounding 3D cloud geometry. This observation-input asymmetry limits the model's ability to capture cloud adjacency effects. In contrast, the generalization performance for longwave radiation (Fig. 3b) appeared excellent and stable. R^2 on the test set remained at the high level of 0.99, and RMSE only increased marginally to 1.970 W/m^2, indicating stable performance across the held-out samples. Clear-sky albedo also remained in close agreement with the reference values (R^2 = 0.97, RMSE = 0.004).

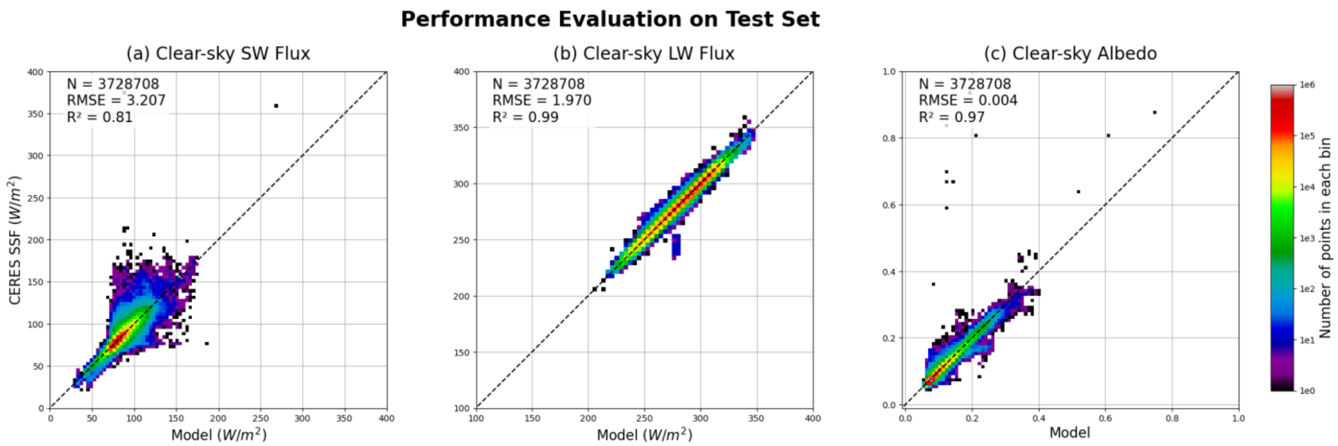


Figure 3. Probability density distributions for the independent test set. The x-axis shows CSRFormer predictions and the y-axis shows CERES SSF clear-sky reference values. Greater proximity to the 1:1 line indicates better agreement. (a) Clear-sky shortwave flux; (b) clear-sky longwave flux; (c) clear-sky albedo.

Due to satellite resolution limits, pixels marked as 100% clear may still contain sub-pixel or thin clouds (cloud contamination) (Sun et al., 2011; Eytan et al., 2025). These residual clouds strongly reflect solar radiation, causing anomalously high observed clear-sky SW fluxes (Zheng et al., 2025). Fig. A1 and A2 confirm this: clouds remain visible in MODIS imagery within regions CERES identifies as clear, demonstrating that the CERES clear fraction alone cannot completely remove contamination. To address this data noise, we introduced the MODIS 1km cloud mask to stringently filter samples (Platnick et al., 2017; Frey et al., 2008) and CSRFormer's reanalysis inputs are inherently unaffected by instantaneous cloud reflection. As shown in Fig. A3, predictions relying solely on CERES labels exhibit high-value outliers due to contamination, however, applying strict

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

MODIS screening largely eliminates them (Fig. 3). In addition, by learning from massive datasets, CSRFormer is robust enough to ignore sporadic noise from residual contamination (Zheng et al., 2025; Rolnick et al., 2017; Rahaman et al., 2019), reducing the influence of residual cloud-contaminated labels in the screened training samples.

4.2 Indirect evaluation and physical interpretability

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 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 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 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.

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

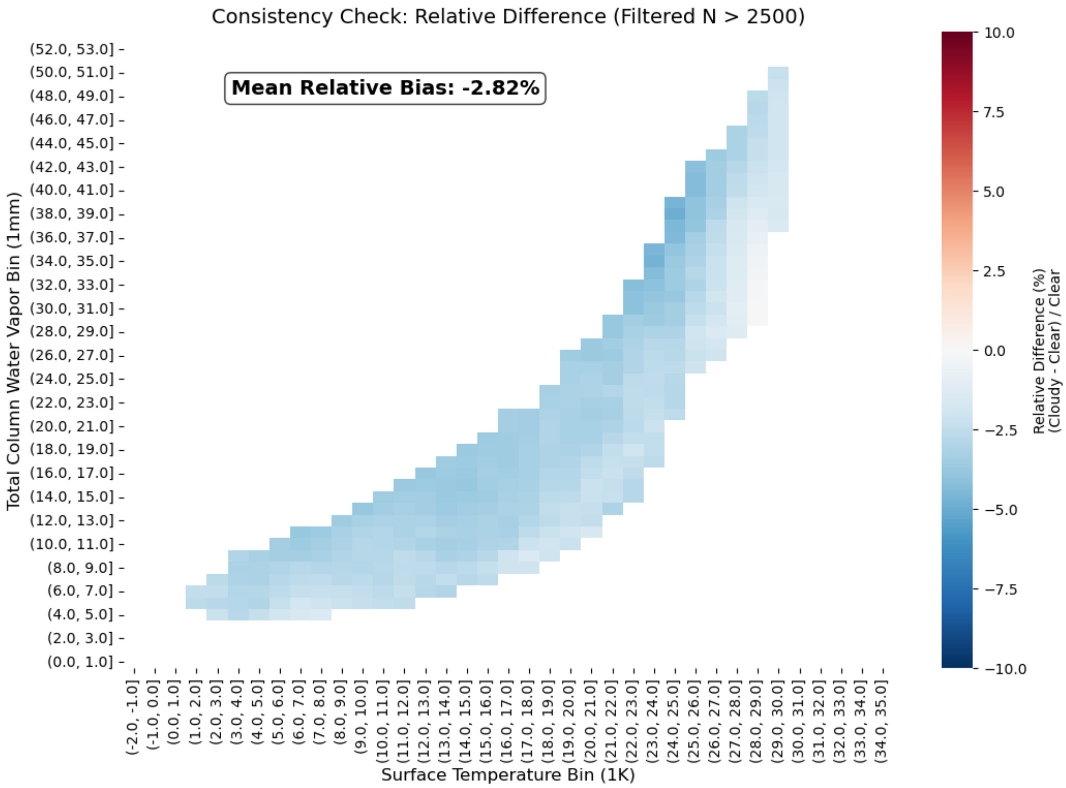


Figure 4. Distribution of relative differences in CSRFormer-predicted clear-sky longwave radiative fluxes between cloudy and clear-sky pixels across varying meteorological conditions. The X-axis represents surface skin temperature (NOAA SST), and the y-axis represents TCWV. The color shading indicates the relative difference (calculated as the cloudy-pixel prediction minus the clear-pixel prediction, normalized by the clear-pixel prediction) within the same meteorological bins. Note: To ensure statistical robustness, only grid cells containing more than 2,500 samples are calculated and displayed; regions with insufficient sampling are masked in white.

We next used permutation feature importance to assess which inputs most strongly influence the predictions (Fig. 5). 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).

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

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

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). 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.

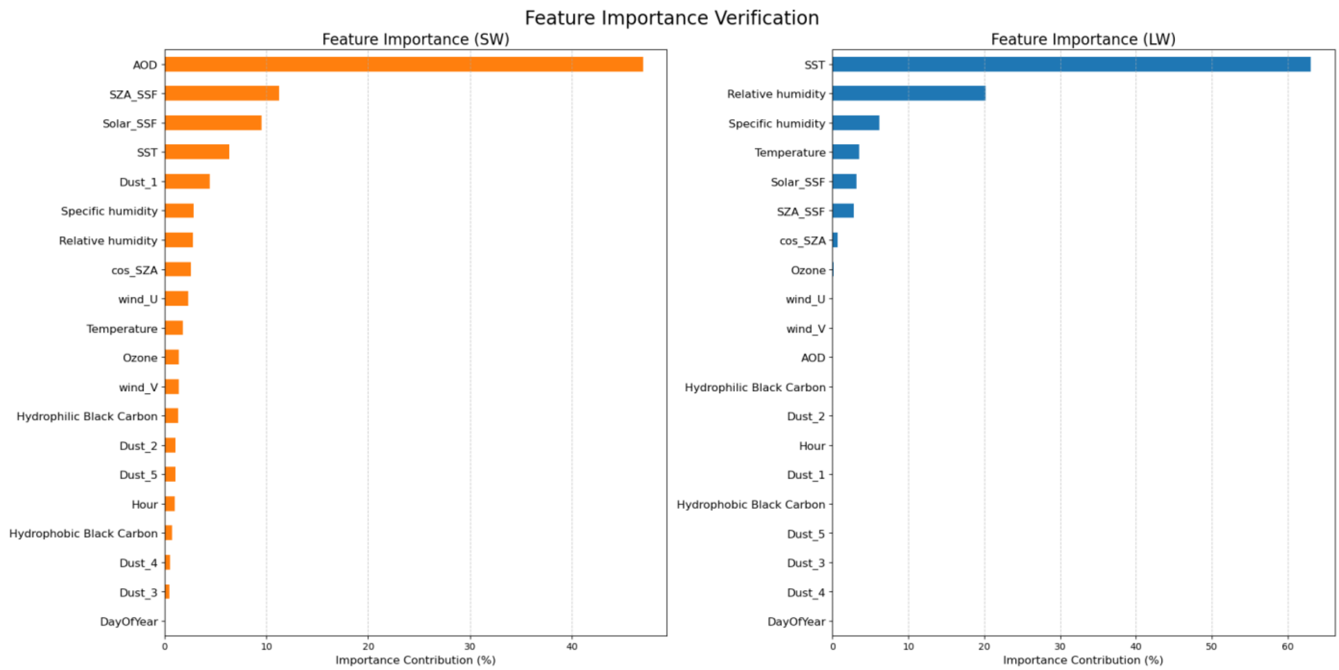


Figure 5. Feature importance verification for CSRFormer input variables. Importance is quantified by the increase in RMSE when a specific feature is randomly permuted while others remain unchanged (Permutation Feature Importance). Features are sorted in descending order of importance; a larger increase 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%.

4.3 Comparison with existing radiation products

To place CSRFormer in the context of existing radiation products, we compared it with CERES CRS, CERES SYN1deg, ERA5, and MERRA-2. CRS was compared at the footprint level, whereas SYN1deg, ERA5, and MERRA-2 were compared on a common 1° ocean grid. The comparisons were designed to assess not only the similarity of the reconstructed clear-sky

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

fluxes to CERES SSF in clear-sky regions, but also the systematic differences among products arising from their different algorithms, input fields, and temporal representations.

4.3.1 Comparison with CERES CRS

In clear-sky regions, the joint comparison of CERES SSF, CERES CRS, and CSRFormer (Fig. 6) 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 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.

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

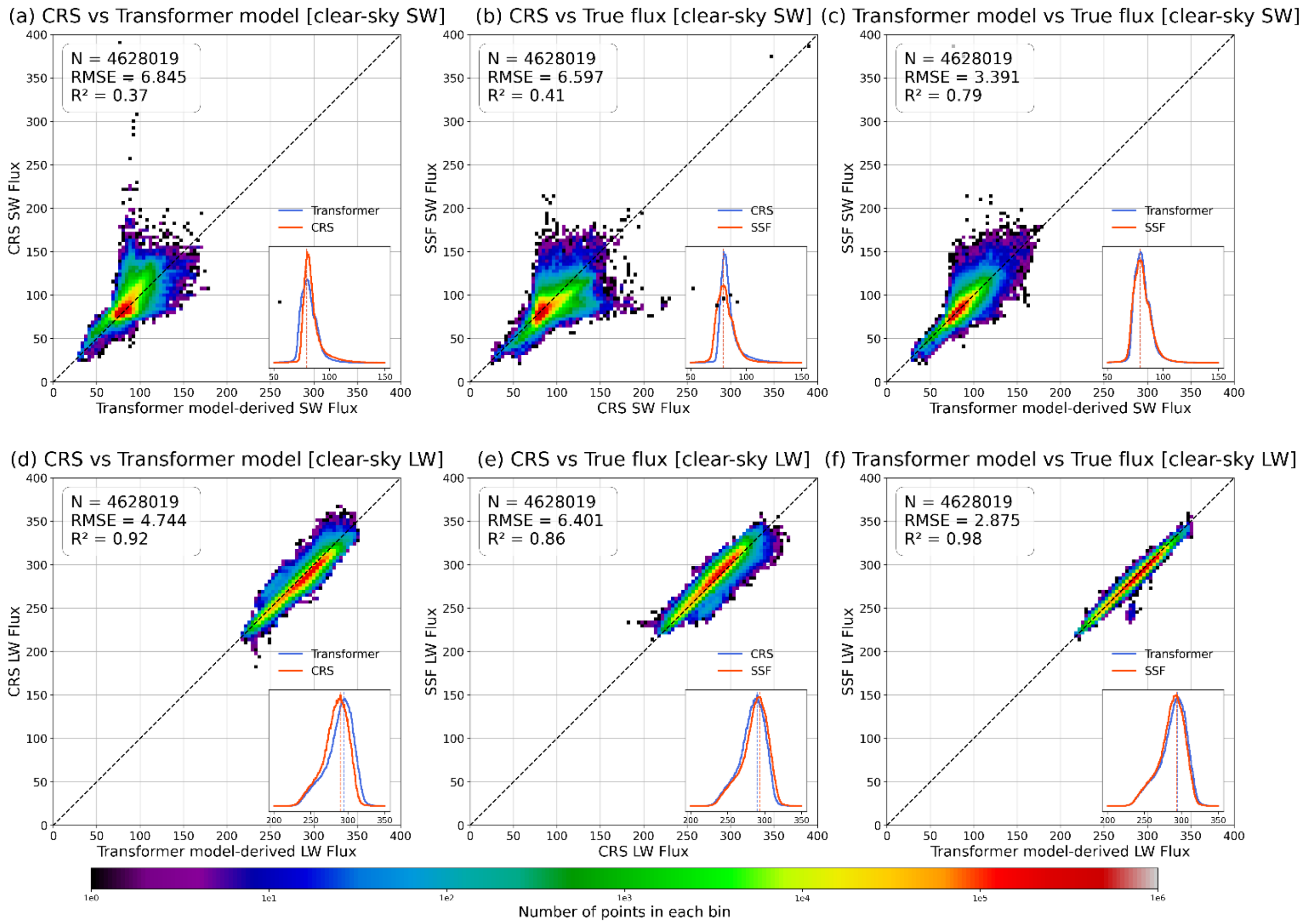


Figure 6. Probability density distribution plot of the comparisons of the CERES SSF, CERES CRS and CSRFormer in clear-sky areas. The closer to the diagonal line represents the more similar the two products are. (a) Comparison of clear-sky shortwave fluxes between CRS and CSRFormer; (b) Comparison of clear-sky shortwave fluxes between CRS and SSF; (c) Comparison of clear-sky shortwave fluxes between SSF and CSRFormer; (d) Comparison of clear-sky longwave fluxes between CRS and CSRFormer; (e) Comparison of clear-sky longwave fluxes between CRS and SSF; (f) Comparison of clear-sky longwave fluxes between SSF and CSRFormer.

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. 7). 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

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

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.

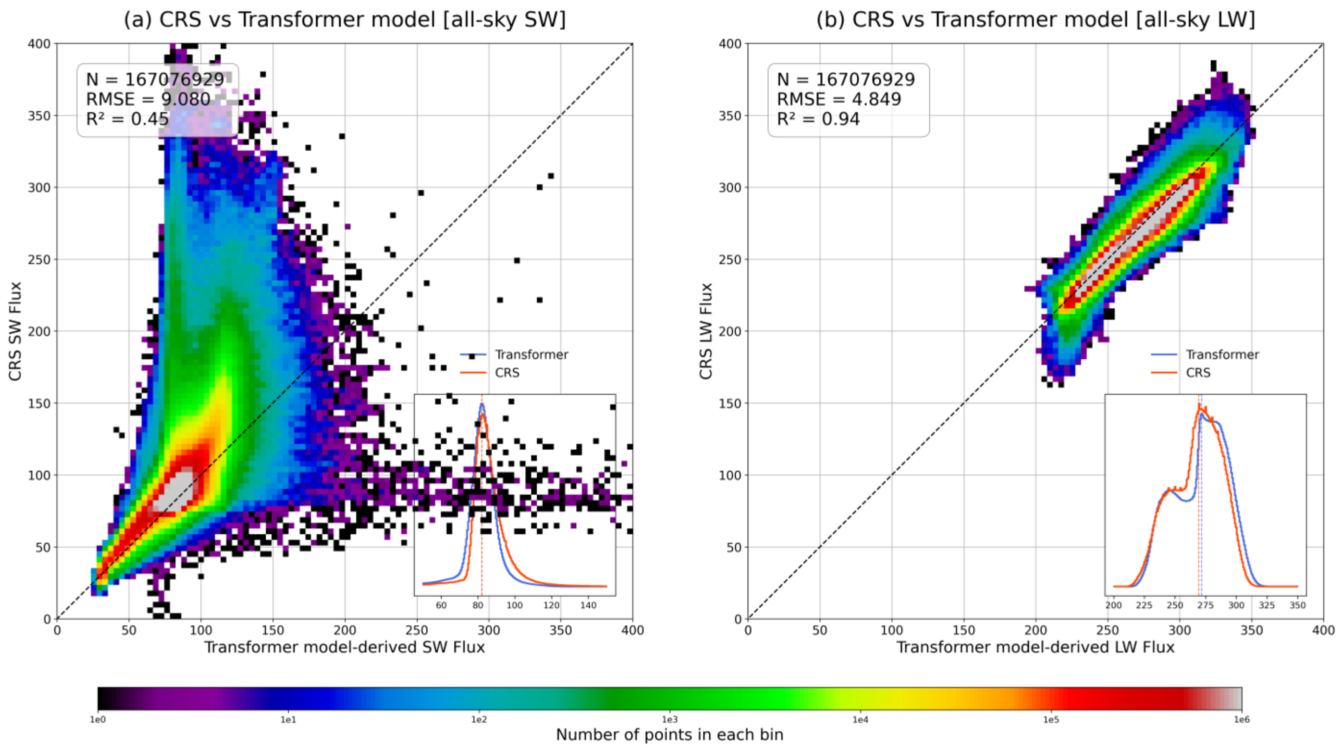


Figure 7. Probability density distribution plot of the comparisons of the CERES CRS and CSRFormer in all-sky areas. The closer to the diagonal line represents the more similar the two products are. (a) Comparison of clear-sky shortwave fluxes between CRS and CSRFormer; (b) Comparison of clear-sky longwave fluxes between CRS and CSRFormer.

The single-day spatial comparison in Fig. 8 supports this interpretation. In the shortwave band, CRS shows higher reflected clear-sky fluxes than CSRFormer in several high-aerosol regions, including areas near North Africa, India, and eastern China,

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

Open Access Earth System Science Data Discussions

which is consistent with the MATCH–MERRA-2 aerosol difference discussed above (Fillmore et al., 2022). CRS also exhibits striping artifacts in some tropical and subtropical sunglint-related regions, although the exact source of these artifacts remains unclear. In the longwave band, the differences are spatially broader and more systematic, suggesting that they are more closely linked to differences in radiative-transfer assumptions and meteorological driving fields than to localized input effects alone (Scott et al., 2022).

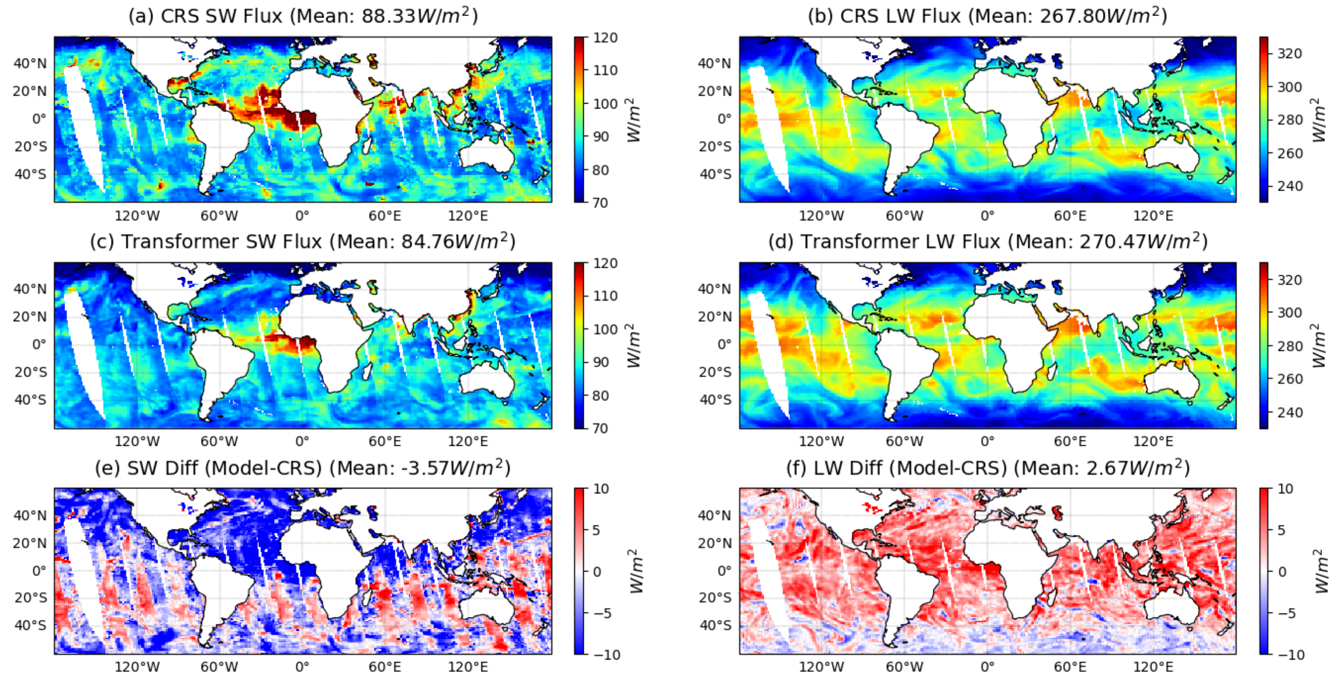


Figure 8. Spatial distributions and differences of clear-sky radiative fluxes from the CERES CRS product and the CSRFormer model in January 1st, 2018. (a)–(b) Clear-sky shortwave and longwave radiative fluxes from the CERES CRS product, respectively. (c)–(d) Clear-sky shortwave and longwave radiative fluxes predicted by CSRFormer, respectively. (e)–(f) Differences in shortwave and longwave fluxes between CSRFormer and CERES CRS (calculated as CSRFormer minus CRS). The global mean value is annotated in each panel (Unit: W/m²).

4.3.2 Comparison with CERES SYN1deg

Comparing CSRFormer with the CERES SYN1deg product (Fig. 9) 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

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

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°×1°) inevitably introduces smoothing errors.

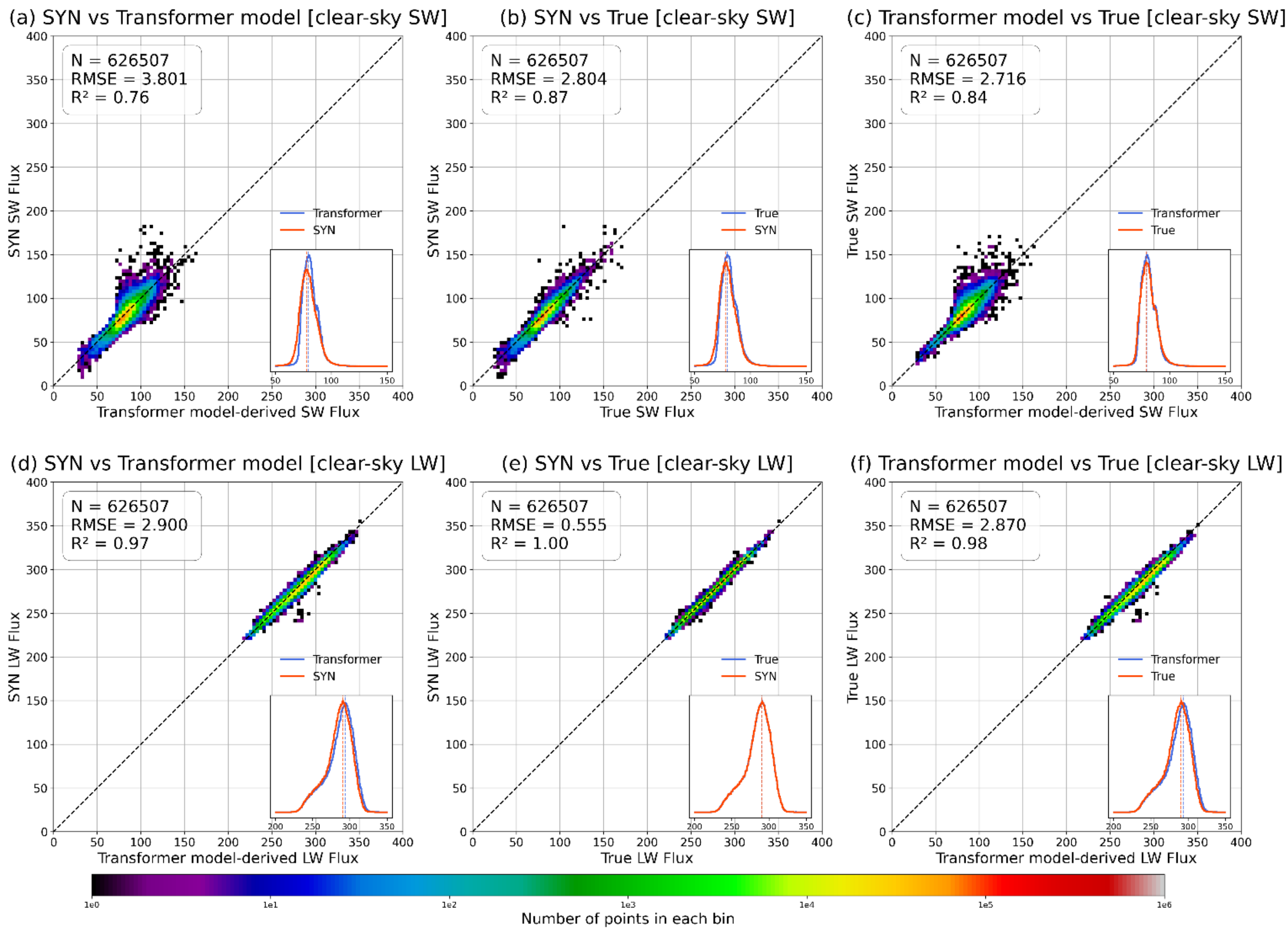


Figure 9. Probability density distribution plot of the comparisons of the CERES SSF, CERES SYN and CSRFormer in clear-sky areas. The closer to the diagonal line represents the more similar the two products are. (a) Comparison of clear-sky shortwave fluxes between SYN and CSRFormer; (b) Comparison of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

Extending the analysis to all-sky areas (i.e., estimating clear-sky fluxes under clouds), the deviation between CSRFormer and the CERES SYN product expands significantly (Fig. 10). For shortwave fluxes, this discrepancy primarily stems from temporal

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

misalignment and differing aerosol treatments. SYN estimates hourly fluxes based on hourly-averaged SZA, causing systematic biases compared to the instantaneous SSF overpass (~13:30 LT) anchored by CSRFormer (Doelling et al., 2013; Zheng et al., 2025). Furthermore, because SYN relies on MODIS, it lacks sub-cloud AOD retrievals and is expected to represent aerosol effects differently in cloudy regions (Doelling et al., 2013; Zheng et al., 2025). To address this, CSRFormer incorporates MERRA-2 AOD, effectively accounting for sub-cloud aerosol effects (Zheng et al., 2025). Additionally, SYN exhibits extreme high-value outliers that severely deviate from normal physical ranges, likely caused by residual cloud contamination or complex sunglint interference. For longwave fluxes, the discrepancies highlight the limitations of SYN's linear interpolation between adjacent clear-sky observations (Doelling et al., 2016). While effective for low-frequency variations, this interpolation may not fully capture non-linear fluctuations in surface temperatures and atmospheric profiles during persistent cloud cover. Conversely, CSRFormer leverages all-weather ERA5 profiles and NOAA SST to provide physically consistent estimates, avoiding the direct use of temporal interpolation in the CSRFormer.

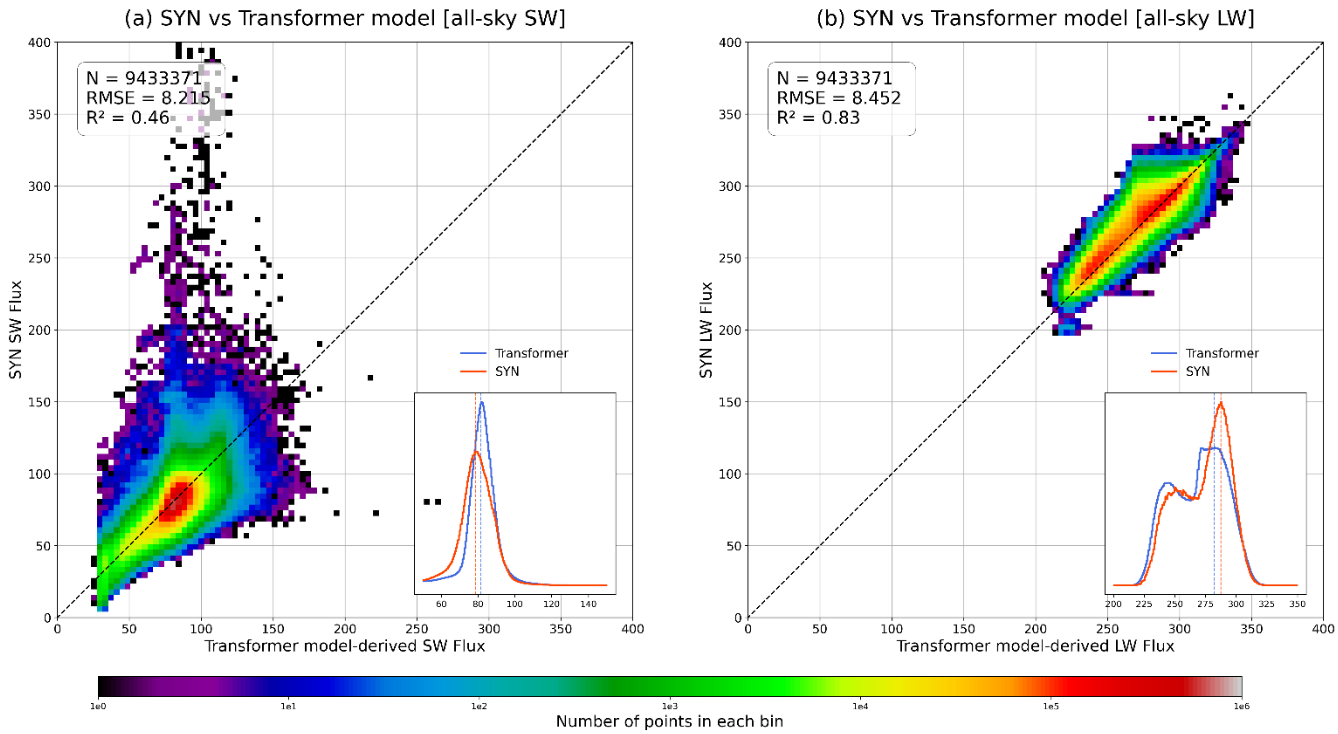


Figure 10. Probability density distribution plot of the comparisons of the CERES SYN and CSRFormer in all-sky areas. The closer to the diagonal line represents the more similar the two products are. (a) Comparison of clear-sky shortwave fluxes between SYN and CSRFormer; (b) Comparison of clear-sky longwave fluxes between SYN and CSRFormer.

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

A single-day spatial distribution analysis (Fig. 11) intuitively highlights these differences. For longwave fluxes, SYN exhibits horizontal discontinuity, noise-like jumps, and misses certain weather system impacts. This stems from SYN's reliance on linear interpolation between adjacent clear-sky observations (Doelling et al., 2016). During persistent cloud cover, this interpolation fails to reconstruct smooth thermodynamic gradients, causing splicing artifacts and omitting high-frequency fluctuations. For shortwave fluxes, the differences expose divergent aerosol handling. Because SYN relies on MODIS AOD, it cannot retrieve sub-cloud data, often ignoring aerosol scattering or reverting to climatological backgrounds in cloudy regions (Doelling et al., 2013; Zheng et al., 2025). Conversely, CSRFormer utilizes MERRA-2, ensuring continuous sub-cloud aerosol fields. This is prominent during Saharan dust transport: CSRFormer captures a continuous high-value belt reflecting the dust plume's extinction effect, whereas SYN may lead to weaker aerosol-related signals in these regions due to missing potential sub-cloud AOD (Zheng et al., 2025).

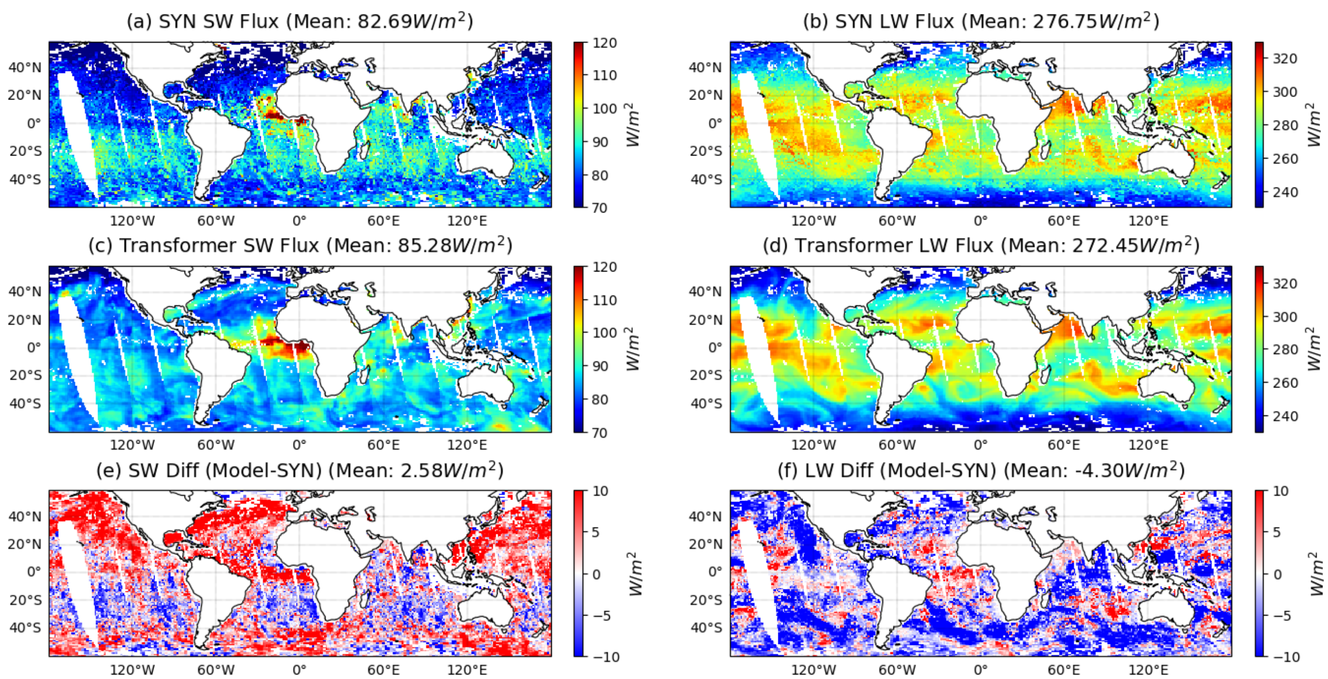


Figure 11. Spatial distributions and differences of clear-sky radiative fluxes from the CERES SYN product and the CSRFormer model in January 1st, 2018. (a)–(b) Clear-sky shortwave and longwave radiative fluxes from the CERES SYN product, respectively. (c)–(d) Clear-sky shortwave and longwave radiative fluxes predicted by CSRFormer, respectively. (e)–(f) Differences in shortwave and longwave fluxes between CSRFormer and CERES SYN (calculated as CSRFormer minus SYN). The global mean value is annotated in each panel (Unit: W/m^2).

4.3.3 Comparison with reanalysis products (ERA5 and MERRA-2)

We finally compared CSRFormer with ERA5 and MERRA-2 as representatives of reanalysis-based radiation products. Both products differ systematically from CSRFormer and CERES SSF, which is consistent with their reliance on model-based radiation schemes, coarser spatiotemporal representation, and the lack of strict synchronization with the instantaneous CERES observation time.

For ERA5, the contrast between the longwave and shortwave bands is particularly clear. As shown in Fig. A6 and A7, ERA5 longwave clear-sky fluxes remain close to CSRFormer and CERES SSF, which is expected because CSRFormer uses ERA5 temperature, humidity, and ozone profiles as key inputs, so both products are anchored to a similar atmospheric thermodynamic state. In the shortwave band, however, substantial differences persist. These are likely related to ERA5's idealized treatment of ocean-surface albedo and its use of climatological aerosol inputs (Tegen et al., 1997; Taylor et al., 2006; Hersbach et al., 2020), whereas CSRFormer uses dynamic SST, wind speed, and aerosol information to reconstruct the instantaneous shortwave response more flexibly.

MERRA-2 shows a different behavior. Compared with ERA5, the shortwave discrepancy is smaller (Fig. A8 and A9), which is consistent with the stronger role of aerosol forcing in shortwave radiation and with the fact that CSRFormer directly uses MERRA-2 AOD and absorbing aerosol variables as inputs. However, the MERRA-2 radiation product still exhibits an overall overestimation tendency relative to CERES SSF and CSRFormer. This remaining difference likely reflects both biases in the native MERRA-2 radiation scheme and the mismatch between hourly-mean reanalysis fluxes and instantaneous CERES observations (Gelaro et al., 2017; Randles et al., 2017; Su et al., 2015b). Extending the comparison to all-sky regions slightly improves the overall fit statistics because of the larger sample size, but the overestimation tendency remains.

4.3.4 Summary of product intercomparison

Taken together, the product comparisons show that CSRFormer agrees closely with CERES SSF in clear-sky regions and differs in systematic ways from CRS, SYN, ERA5, and MERRA-2. Many of these differences can be understood from the underlying product designs, input fields, and temporal treatments, although some specific attributions should still be regarded as tentative unless directly verified. A practical advantage of CSRFormer is that it produces predictions at the CERES observation time, which avoids the temporal mismatch inherent in hourly mean products and may be useful when combining the fluxes with MODIS cloud-property retrievals on Aqua.

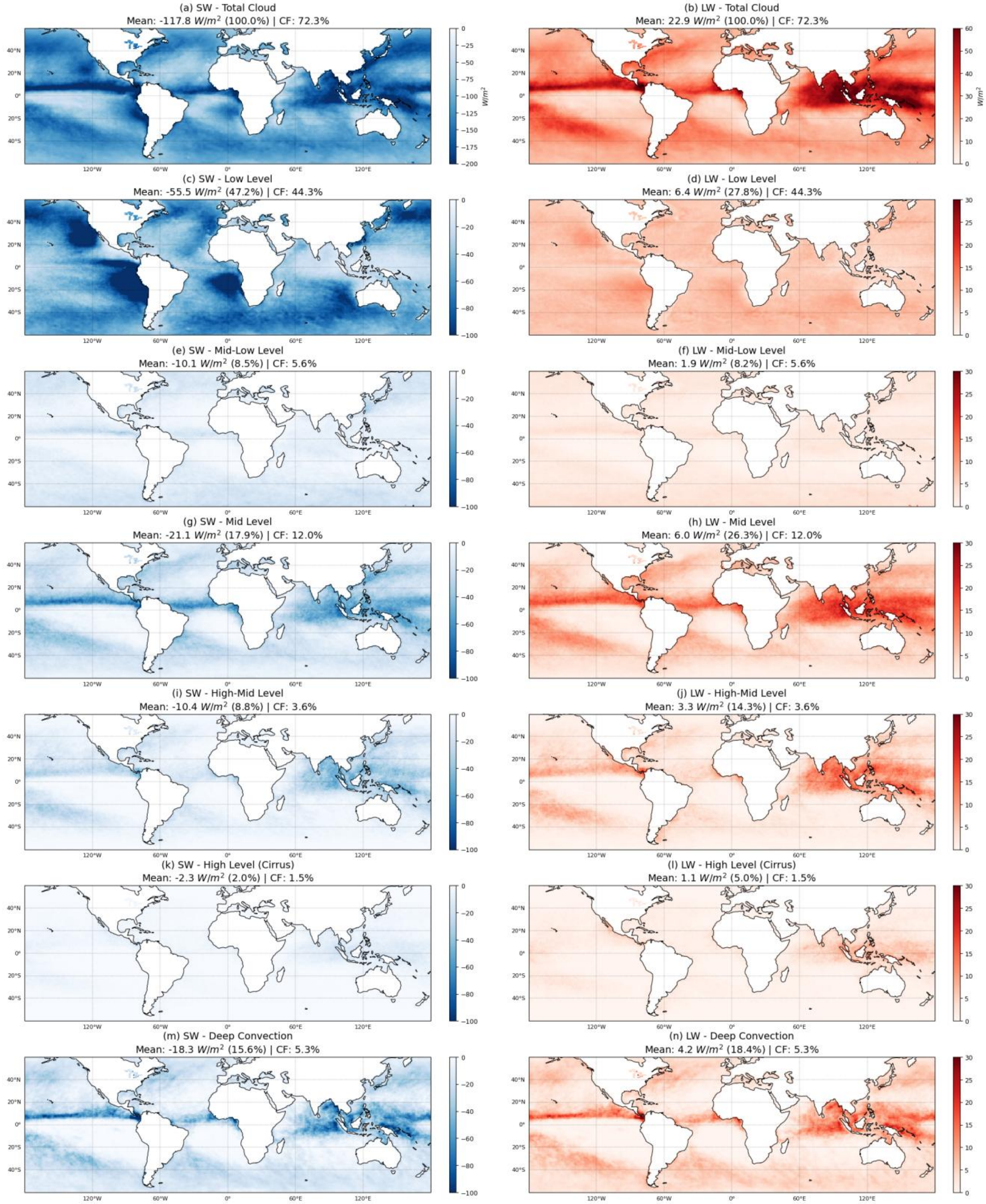
https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

4.4 Application of clear-sky flux

4.4.1 Estimation of CRE for different clouds

Using CSRFormer clear-sky fluxes together with CERES SSF all-sky TOA fluxes, we estimated instantaneous CRE over the global ocean. The latitude-longitude weighted means are -117.8 W m^{-2} for shortwave CRE and 22.9 W m^{-2} for longwave CRE. These values represent the Aqua overpass time rather than a full diurnal mean, so they should be interpreted accordingly. Spatially, the strongest CRE signals occur over the ITCZ, the western Pacific warm pool, and the midlatitude storm tracks, consistent with the distribution of deep convective and frontal cloud systems (Allan, 2011; Matus and L'ecuyer, 2017; L’ecuyer et al., 2019; Zheng et al., 2025).

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026



520

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

Figure 12. Spatial distribution of annual mean CRE for clouds grouped by cloud-top pressure class. (a) All-cloud SWCRE; (b) all-cloud LWCRE; (c) low-cloud SWCRE; (d) low-cloud LWCRE; (e) mid-low-cloud SWCRE; (f) mid-low-cloud LWCRE; (g) mid-cloud SWCRE; (h) mid-cloud LWCRE; (i) high-mid-cloud SWCRE; (j) high-mid-cloud LWCRE; (k) high-cloud SWCRE; (l) high-cloud LWCRE; (m) deep-convective-cloud SWCRE; (n) deep-convective-cloud LWCRE. Parenthetical values in the subtitles denote the mean cloud albedo.

By integrating MODIS cloud properties with CERES SSF data, we quantified the specific contributions of different vertical cloud structures to the SW and LW CRE. First, low-level clouds dominate global SW cooling, particularly regulating planetary albedo in stratocumulus regions like the Eastern Pacific, Southeast Atlantic, and Southeast Indian Ocean (Zheng et al., 2025; Wood, 2012; Matus and L'ecuyer, 2017; Cao et al., 2024). Although their warm cloud-top temperatures yield weak LW forcing per unit area, their massive global coverage results in a non-negligible LWCRE (Wood, 2012; L'ecuyer et al., 2019). Second, deep convective and mid-level clouds contribute through distinct mechanisms. Despite low coverage, deep convective clouds possess large optical thickness and high cloud tops, producing strong dual SW and LW CRE per unit area (L'ecuyer et al., 2019; Matus and L'ecuyer, 2017). Conversely, mid-level clouds exert weaker individual radiative effects but have higher coverage, making them the second most important SW regulator. Finally, high-level (cirrus) clouds show weaker radiative effects in this assessment, which may partly reflect the sampling characteristics of the Aqua overpass time and previously reported diurnal variations in cirrus occurrence (Noel et al., 2018). Overall, this cloud-type decomposition illustrates the utility of CSRFormer for linking instantaneous CRE to collocated cloud-property retrievals and for diagnosing how different cloud regimes shape the radiative budget at the satellite overpass time.

4.4.2 Comparison of CRE between CERES CRS and CSRFormer

We further compared CRE estimated using CSRFormer and CERES CRS clear-sky fluxes. CRS was selected because it is a CERES Level-2 radiative-transfer-based product and, like SSF, is matched to the satellite sampling time, allowing a controlled comparison without the additional temporal smoothing present in hourly products. To isolate the impact of the clear-sky benchmark, both CRE calculations used the same CERES SSF all-sky flux, while only the clear-sky flux was varied. The resulting CRE difference therefore directly reflects the effect of the clear-sky reference.

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

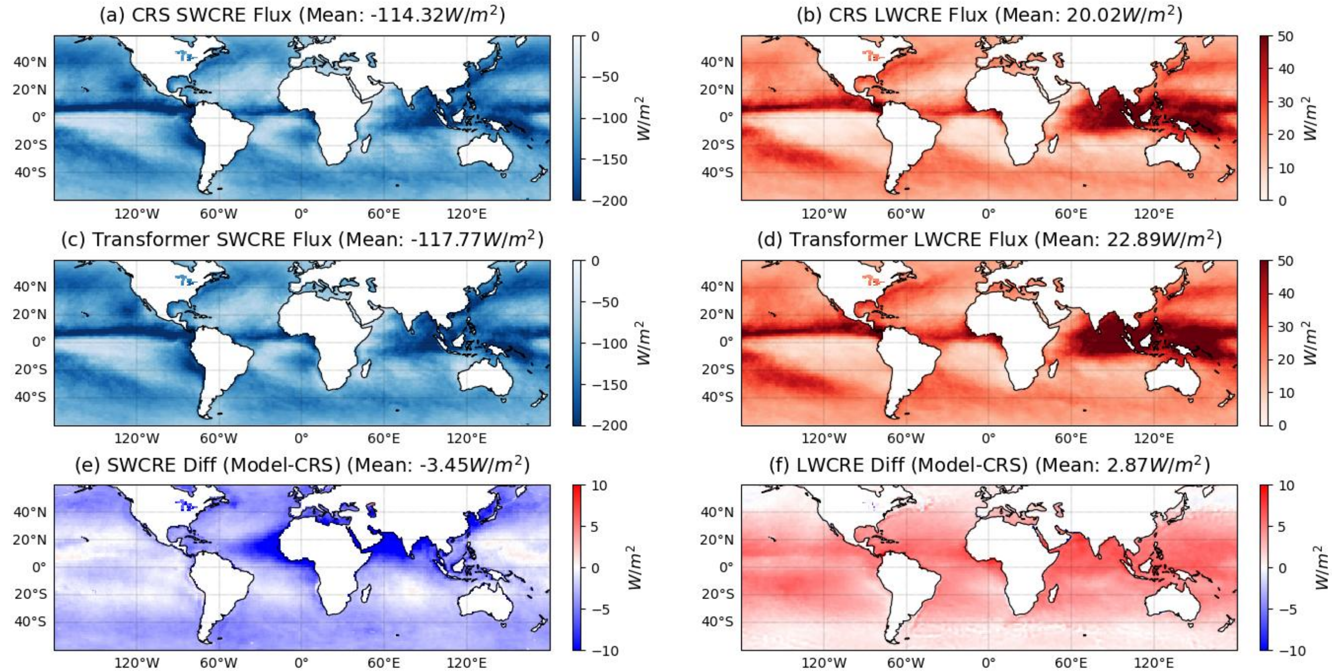


Figure 13. Spatial distribution of annual mean CRE. (a) SWCRE estimated using CERES CRS clear-sky fluxes; (b) LWCRE estimated using CERES CRS clear-sky fluxes; (c) SWCRE estimated using CSRFormer clear-sky fluxes; (d) LWCRE estimated using CSRFormer clear-sky fluxes; (e) CSRFormer SWCRE minus CERES CRS SWCRE; (f) CSRFormer LWCRE minus CERES CRS LWCRE.

The spatial distribution of annual mean CRE differences is shown in Fig. 13. 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).

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

Overall, this comparison shows that CRE estimates are sensitive to the choice of the clear-sky benchmark. Even when the same all-sky observations are used, systematic differences in the reconstructed clear-sky state can produce regionally coherent changes in both SWCRE and LWCRE. This further highlights the importance of constructing clear-sky fluxes that are consistent with the satellite observation time and with the intended instantaneous CRE application.

4.5 Dataset specifications

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.

5 Conclusions and discussion

This study developed CSRFormer, a physically constrained Transformer model for reconstructing instantaneous clear-sky shortwave and longwave radiative fluxes over the global ocean. The model shows strong agreement with CERES SSF in clear-sky conditions and provides estimates matched to the CERES overpass time. This temporal consistency is a practical advantage for instantaneous CRE analysis and collocated studies with satellite cloud-property retrievals. The primary purpose of constructing this model is to provide an estimate of the theoretical clear-sky radiation fluxes under clouds, in order to estimate the CRE, rather than the actual observed clear-sky flux.

Taken together, the comparisons with CRS, SYN, ERA5, and MERRA-2 show that CSRFormer remains closest to CERES SSF in clear-sky regions, while differing systematically from the other products. These differences are not random, but are broadly consistent with the distinct formulations, input fields, and temporal representations used by each dataset. CRS mainly differs from CSRFormer in ways that are consistent with radiative-transfer assumptions and with differences in aerosol and surface treatment, especially in cloudy scenes (Scott et al., 2022; Fillmore et al., 2022). SYN is expected to remain very close to CERES SSF because of its observation-based construction, but it still depends on temporal interpolation and Level-3 gridding (Doelling et al., 2013; Doelling et al., 2016). ERA5 and MERRA-2 represent reanalysis-based approaches; compared with ERA5, MERRA-2 is more consistent with CSRFormer in shortwave flux because of its aerosol inputs, whereas ERA5 is more limited by climatological aerosol assumptions (Tegen et al., 1997; Randles et al., 2017; Hersbach et al., 2020). However, reanalysis radiation products still remain affected by model-based radiation schemes and by the mismatch between hourly fields and instantaneous CERES observations. In this context, the main practical value of CSRFormer is that it combines dynamic environmental inputs with strict matching to the CERES observation time, making it particularly useful for instantaneous CRE analyses and for collocated studies with Aqua/MODIS cloud-property retrievals.

In addition to the product evaluation, CSRFormer was applied together with CERES SSF all-sky fluxes to quantify instantaneous CRE over the global ocean. The resulting latitude-longitude weighted means are -117.8 W m^{-2} for SWCRE and 22.9 W m^{-2} for LWCRE, and these values should be interpreted as Aqua-overpass-time estimates rather than diurnal means. Further combining the fluxes with collocated MODIS cloud-property information allows CRE to be decomposed by cloud type, showing that low clouds dominate shortwave cooling, while deep convective and mid-level clouds contribute in distinct ways to the total radiative effect (Zheng et al., 2025; Matus and L'ecuyer, 2017; L’ecuyer et al., 2019; Cao et al., 2024). This application illustrates the practical value of CSRFormer for linking observation-time-matched clear-sky fluxes with cloud-property retrievals in instantaneous CRE analyses.

The quality of the input data also limits the quality of the predictions. NOAA OISST, ERA5 temperature and humidity profiles, and MERRA-2 aerosol fields are all widely used products, but they are not error free (Buchard et al., 2017; Randles et al., 2017; Gueymard and Yang, 2020; Su et al., 2023; Castro et al., 2016; Gupta and Sil, 2024). In particular, the choice of aerosol input can influence the shortwave reconstruction substantially. MERRA-2 offers the practical advantage of spatially complete fields, including cloudy regions, but differences from other aerosol products should not be interpreted automatically as proof that one product is physically correct in every situation. Future work could assess this sensitivity more explicitly by repeating the analysis with alternative aerosol inputs.

A data-driven model can partially compensate for persistent input biases when those biases are represented systematically in the training data (Krasnopolsky, 2007; Rasp and Lerch, 2018; Reichstein et al., 2019). This is one practical advantage of machine-learning-based reconstruction relative to purely prescribed parameterizations. At the same time, such compensation remains empirical and depends on the stability of the bias structure represented in the training set. The model may therefore be less robust for rare events, regime shifts, or sporadic input errors, and these uncertainties should be quantified more systematically in future work.

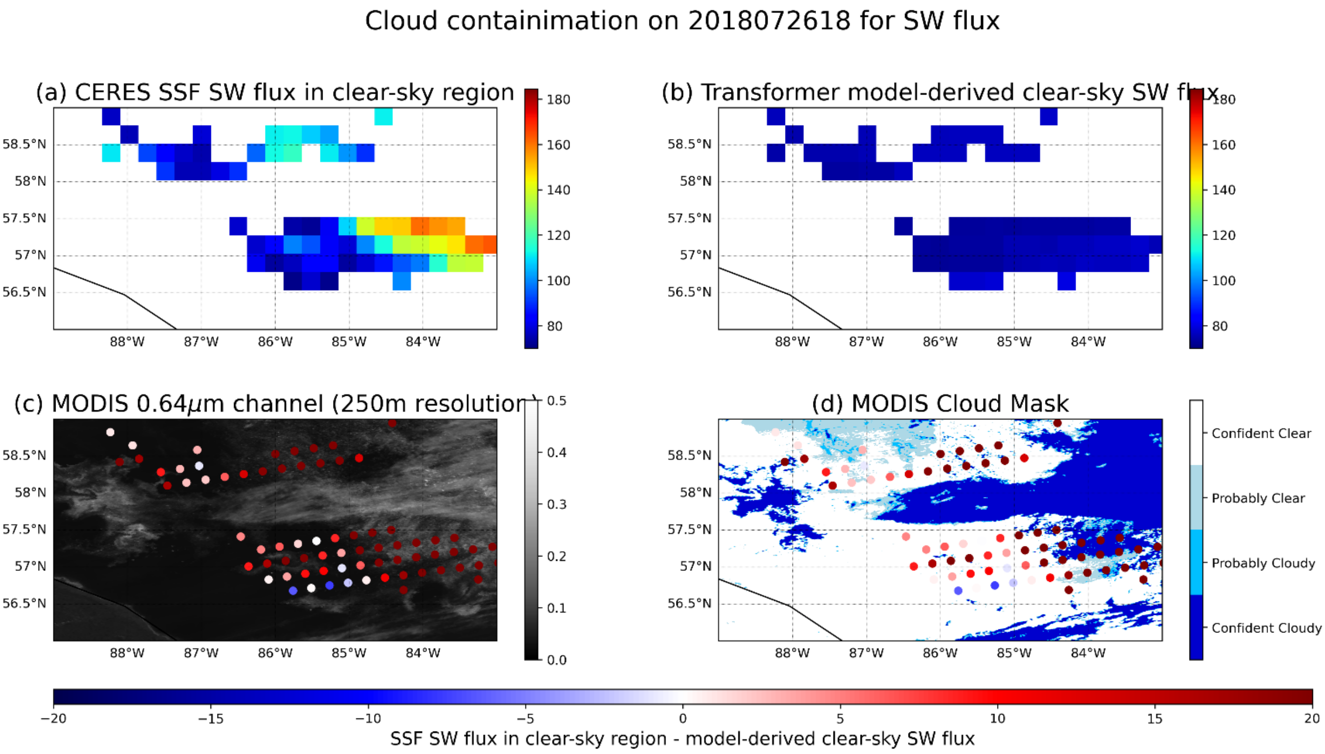
A further point to note is that CSRFormer uses column-based inputs and therefore does not explicitly represent 3D radiative effects such as cloud adjacency (Loeb and Davies, 1996; Várnai and Marshak, 2002; Tijhuis et al., 2024). This may affect comparisons near cloud edges, where observations can still contain lateral scattering signals associated with nearby clouds. However, the target quantity in this study is the theoretical cloud-free radiative flux used as the clear-sky reference for CRE, and such 3D cloud effects are not part of that reference state by definition. In this sense, their omission should be viewed

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

primarily as a consequence of the target definition and modeling framework, rather than as a fundamental deficiency in the intended clear-sky reconstruction.

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 cloud-clear transition zones, and it may also be extendable to Terra in future work.

Appendix A: Additional figures



https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

Open Access Earth System Science Data Discussions

Figure A1. Observed on July 26, 2018. The data in the figure is filtered for clear-sky pixels only using the CERES clear fraction. (a) CERES SSF product albedo in the region of CERES clear fraction >99.9%. (b) CSRFormer model predicted clear-sky SW fluxes in the region of CERES clear fraction >99.9%. (c) MODIS channel observations at 250m resolution. (d) MODIS cloud mask product. Where the upper scatter in (c) and (d) is the difference between the albedo of the CERES SSF product and CSRFormer model predicted clear-sky SW fluxes in regions with CERES clear fraction >99.9%.

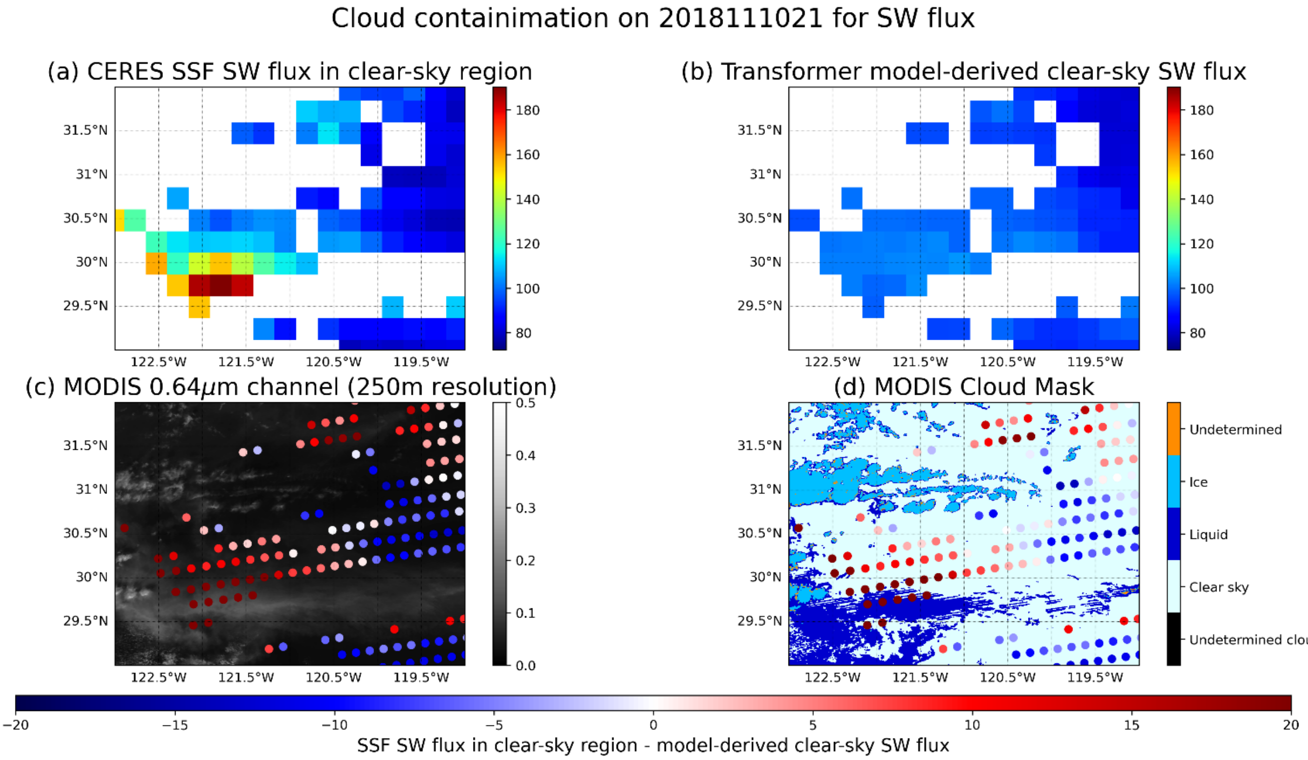


Figure A2. Observed on November 10, 2018. The data in the figure is filtered for clear-sky pixels only using the CERES clear fraction. (a) CERES SSF product albedo in the region of CERES clear fraction >99.9%. (b) CSRFormer model predicted clear-sky SW fluxes in the region of CERES clear fraction >99.9%. (c) MODIS channel observations at 250m resolution. (d) MODIS cloud mask product. Where the upper scatter in (c) and (d) is the difference between the albedo of the CERES SSF product and CSRFormer model predicted clear-sky SW fluxes in regions with CERES clear fraction >99.9%.

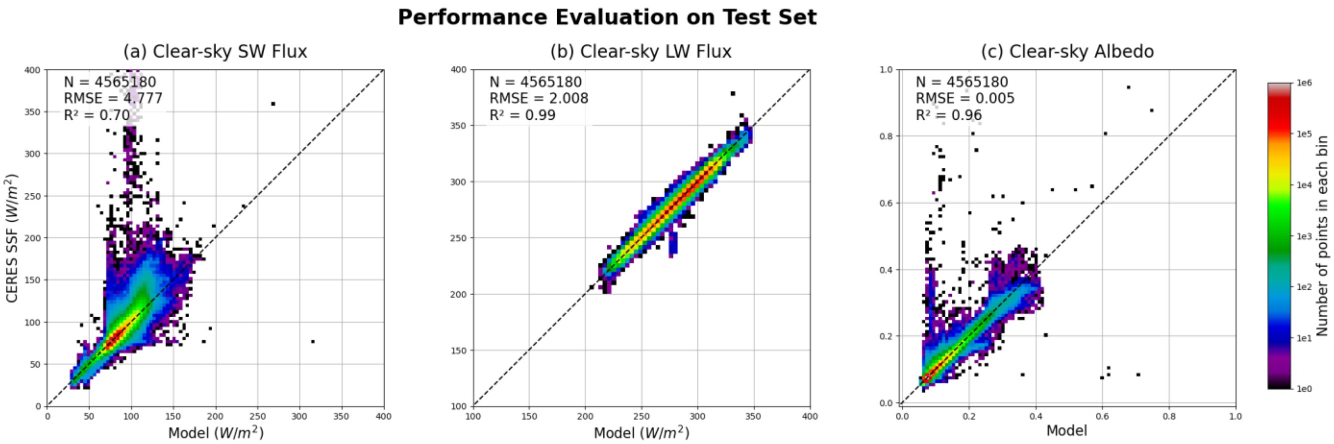


Figure A3. Probability density distributions for the independent test set defined using only the CERES SSF Clear Fraction, without the MODIS 1 km cloud mask. The x-axis shows CSRFormer predictions and the y-axis shows CERES SSF clear-sky reference values. Greater proximity to the 1:1 line indicates better agreement. (a) Clear-sky shortwave flux; (b) clear-sky longwave flux; (c) clear-sky albedo.

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

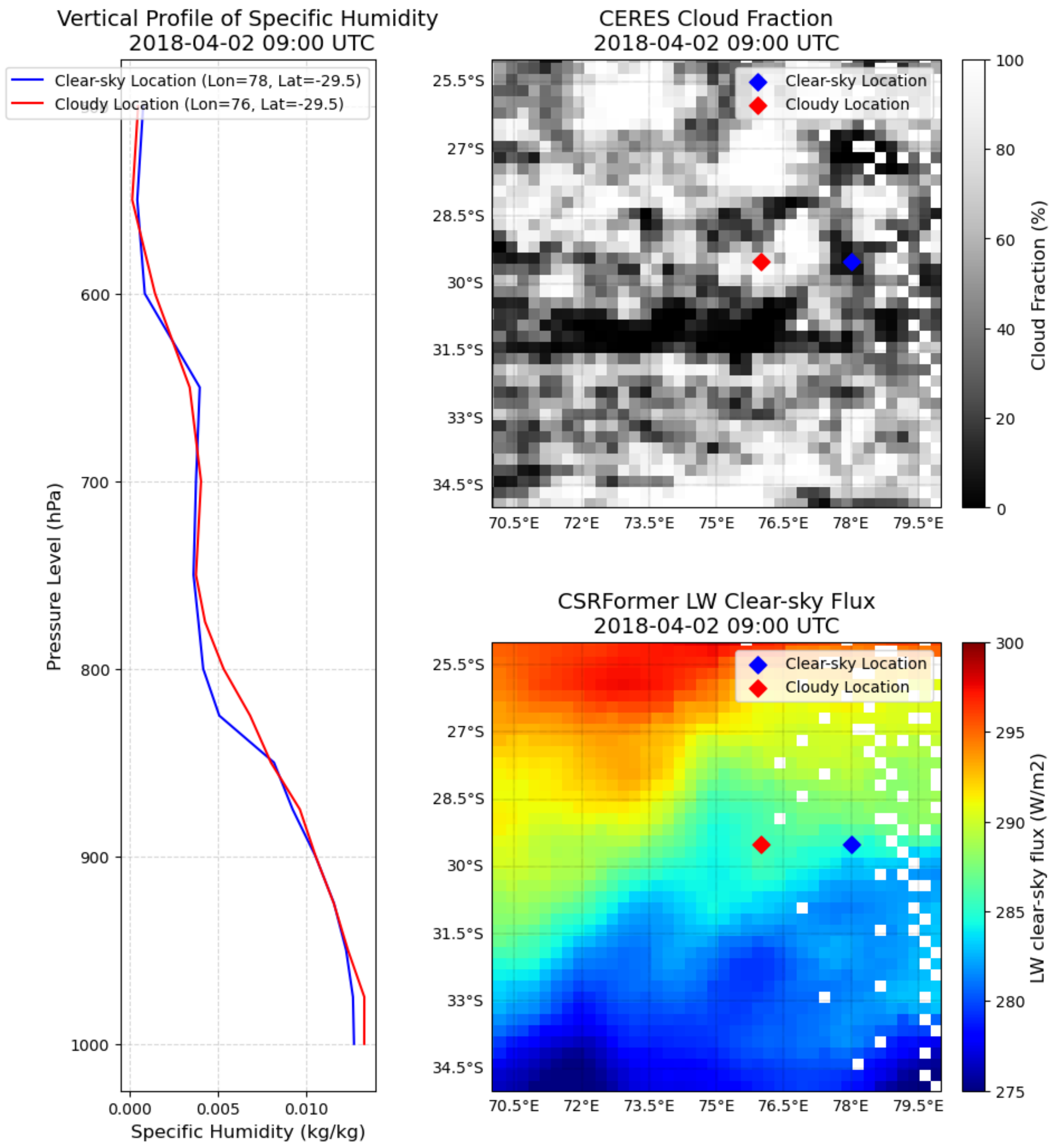
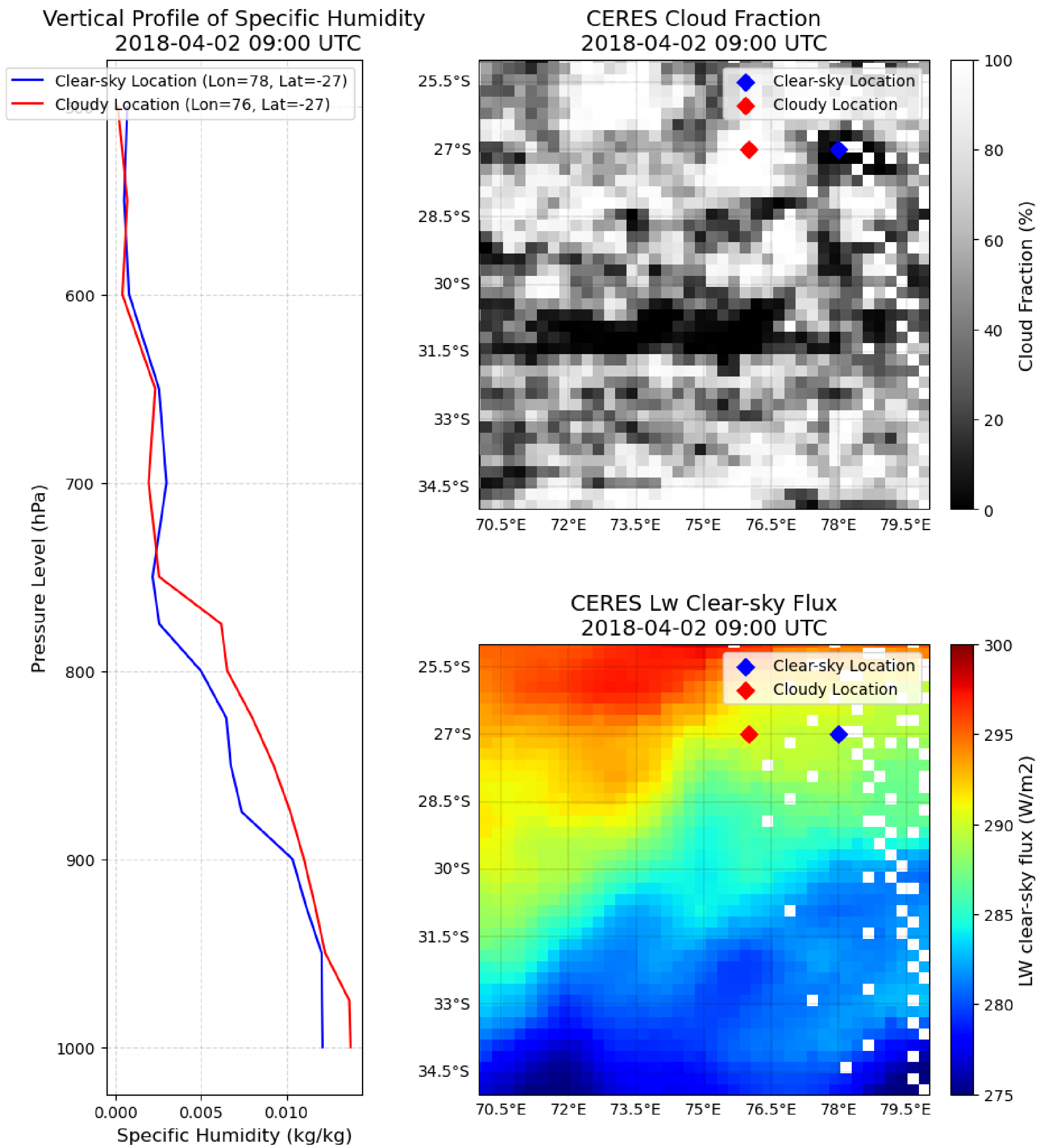


Figure A4. Case1: Vertical profiles of atmospheric humidity and spatial distribution of radiative fluxes for selected pixels. (a) Vertical profiles of specific humidity for the selected clear-sky and cloudy pixels. (b) Spatial distribution of

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

cloud fraction. (c) Spatial distribution of clear-sky longwave radiative fluxes predicted by CSRFormer. The diamond markers in (b) and (c) indicate the geographic locations of the pixels analyzed in (a).

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026



https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

Figure A5. Case2: Vertical profiles of atmospheric humidity and spatial distribution of radiative fluxes for selected pixels. (a) Vertical profiles of specific humidity for the selected clear-sky and cloudy pixels. (b) Spatial distribution of cloud fraction. (c) Spatial distribution of clear-sky longwave radiative fluxes predicted by CSRFormer. The diamond markers in (b) and (c) indicate the geographic locations of the pixels analyzed in (a).

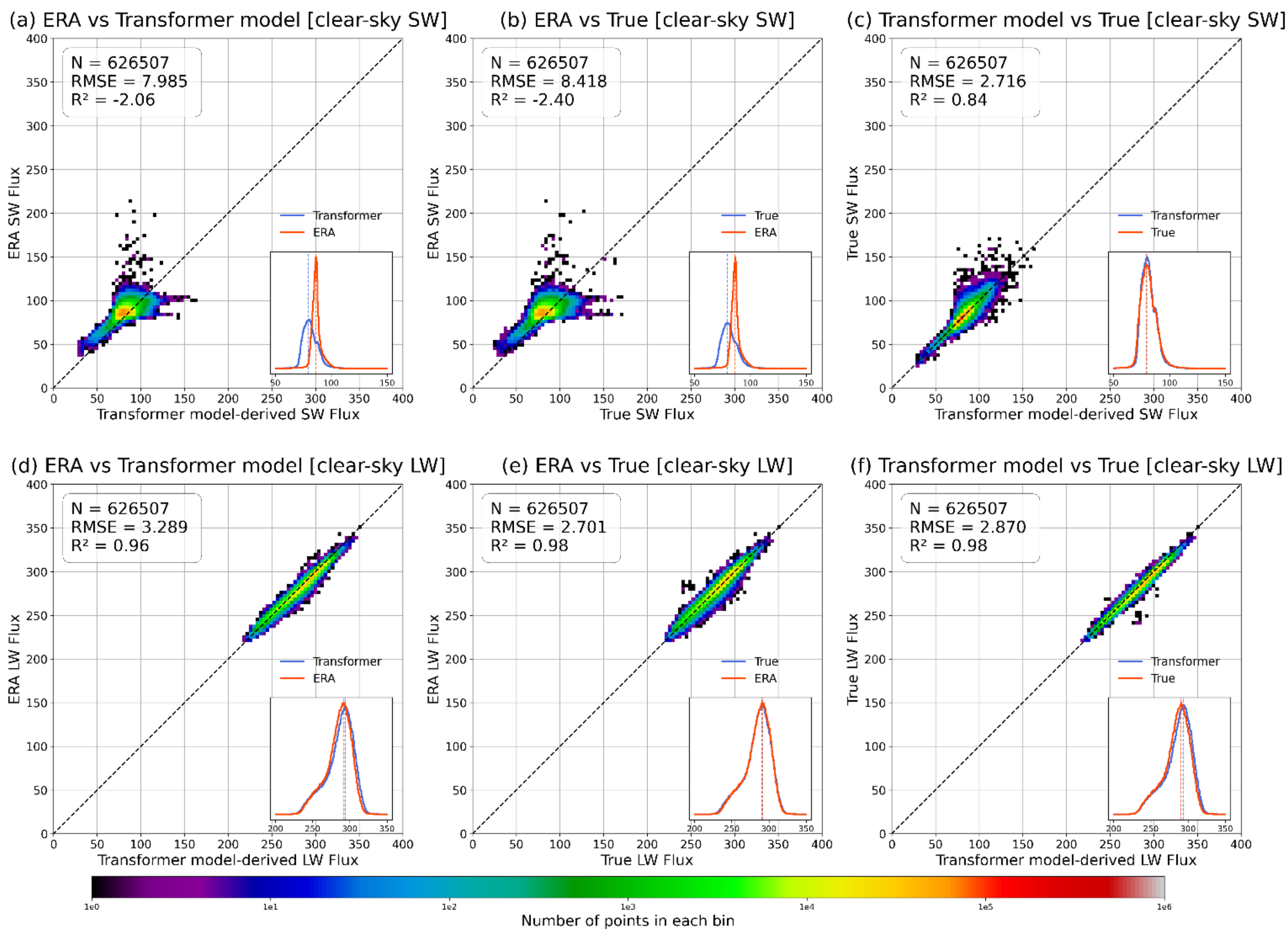


Figure A6. Probability density distribution plot of the comparisons of the CERES SSF, ERA5 and CSRFormer in clear-sky areas. The closer to the diagonal line represents the more similar the two products are. (a) Comparison of clear-sky shortwave fluxes between ERA5 and CSRFormer; (b) Comparison of clear-sky shortwave fluxes between ERA5 and SSF; (c) Comparison of clear-sky shortwave fluxes between SSF and CSRFormer; (d) Comparison of clear-sky longwave fluxes between ERA5 and CSRFormer; (e) Comparison of clear-sky longwave fluxes between ERA5 and SSF; (f) Comparison of clear-sky longwave fluxes between SSF and CSRFormer.

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

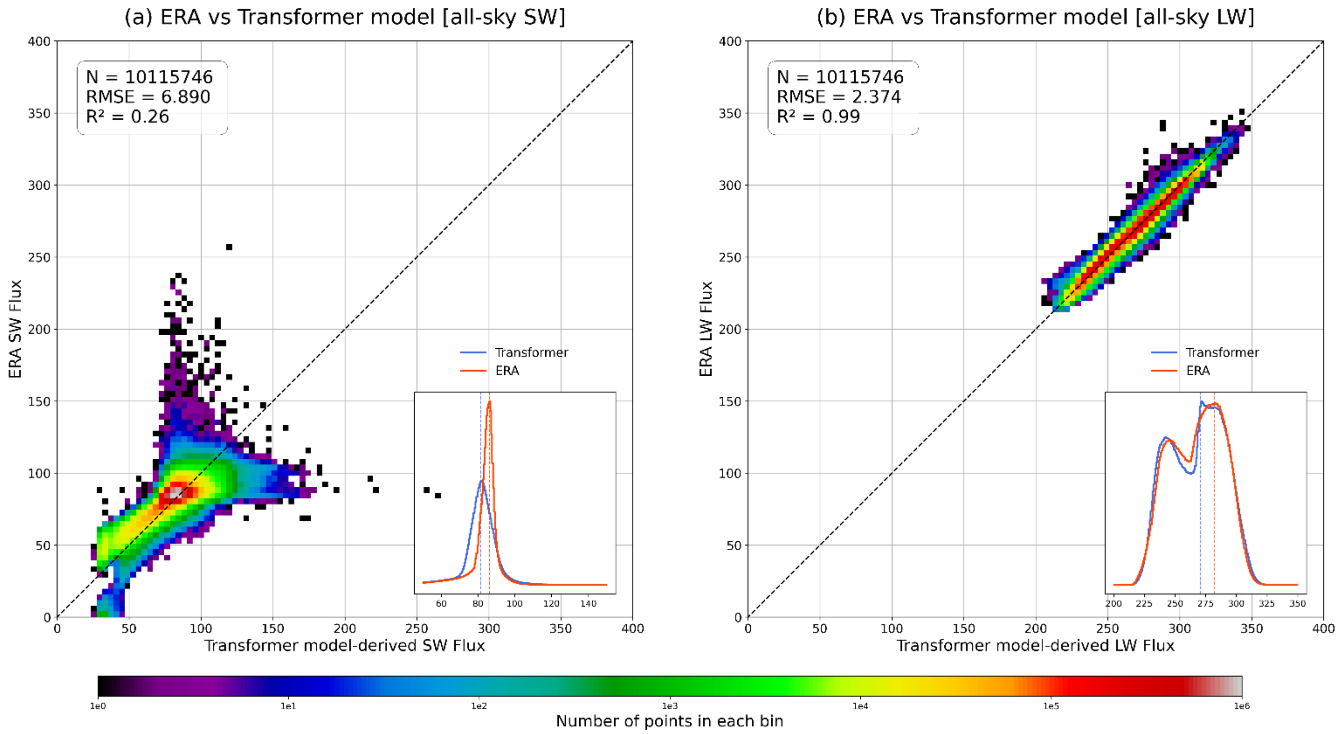


Figure A7. Probability density distribution plot of the comparisons of the ERA5 and CSRFormer in all-sky areas. The closer to the diagonal line represents the more similar the two products are. (a) Comparison of clear-sky shortwave fluxes between ERA5 and CSRFormer; (b) Comparison of clear-sky longwave fluxes between ERA5 and CSRFormer.

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

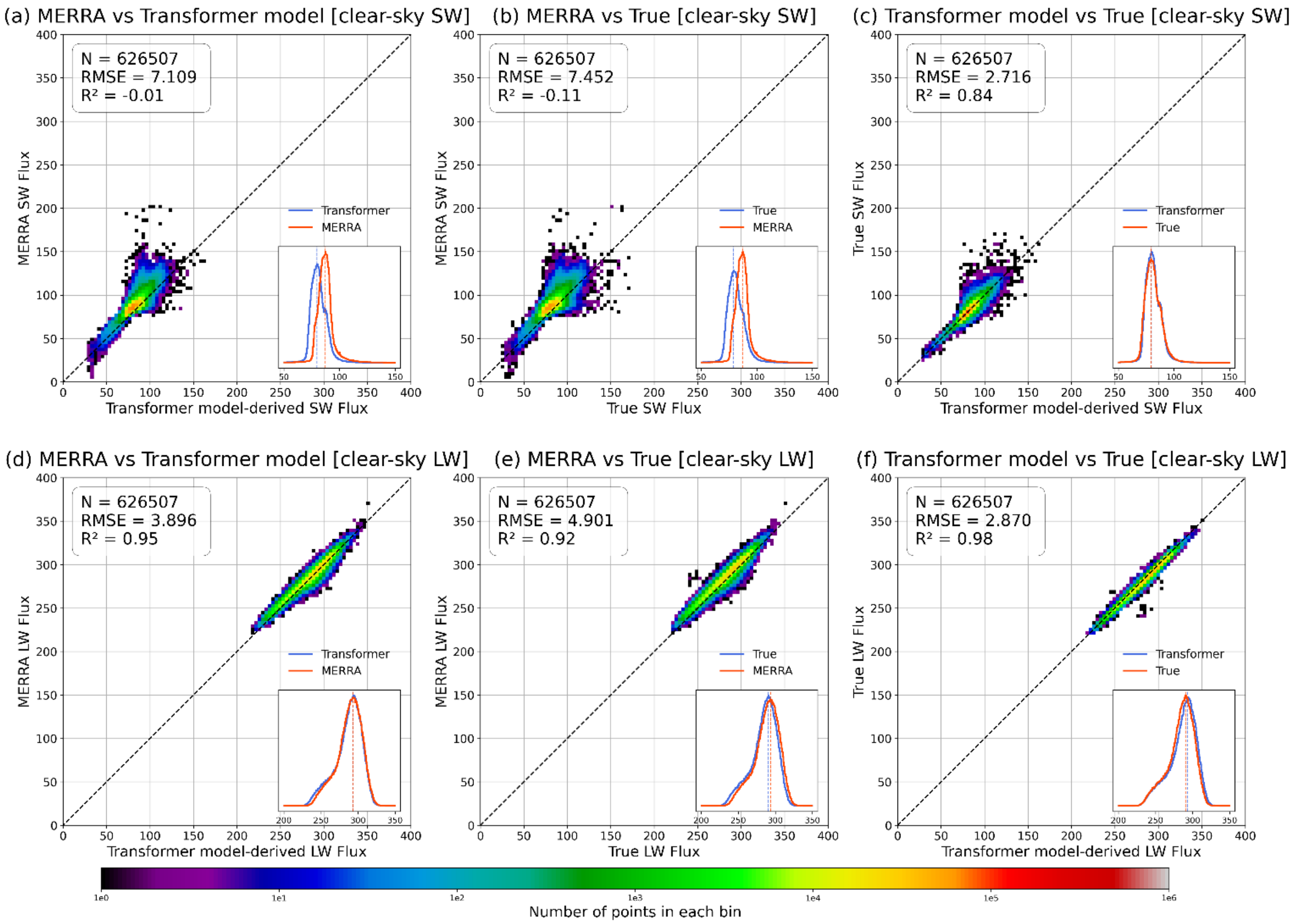


Figure A8. Probability density distribution plot of the comparisons of the CERES SSF, MERRA-2 and CSRFormer in clear-sky areas. The closer to the diagonal line represents the more similar the two products are. (a) Comparison of clear-sky shortwave fluxes between MERRA-2 and CSRFormer; (b) Comparison of clear-sky shortwave fluxes between MERRA-2 and SSF; (c) Comparison of clear-sky shortwave fluxes between SSF and CSRFormer; (d) Comparison of clear-sky longwave fluxes between MERRA-2 and CSRFormer; (e) Comparison of clear-sky longwave fluxes between MERRA-2 and SSF; (f) Comparison of clear-sky longwave fluxes between SSF and CSRFormer.

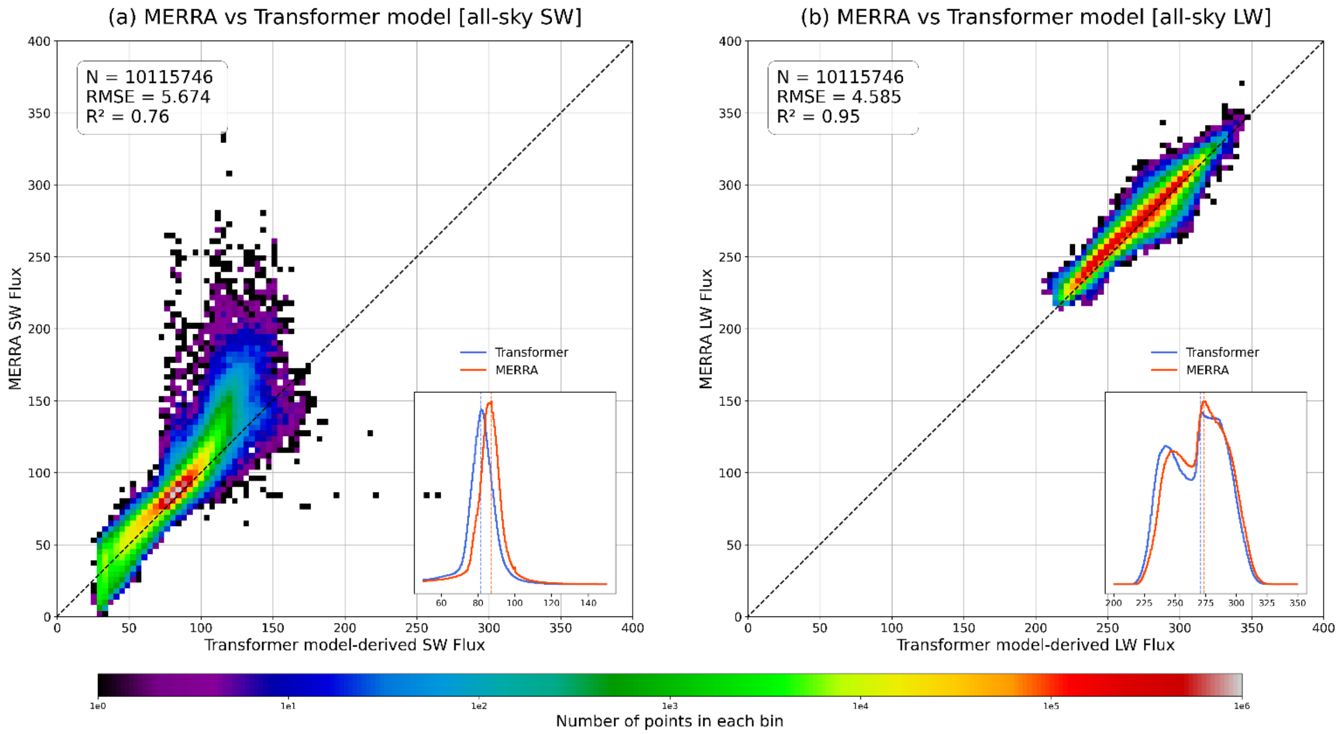


Figure A9. Probability density distribution plot of the comparisons of the MERRA-2 and CSRFormer in all-sky areas. The closer to the diagonal line represents the more similar the two products are. (a) Comparison of clear-sky shortwave fluxes between MERRA-2 and CSRFormer; (b) Comparison of clear-sky longwave fluxes between MERRA-2 and CSRFormer.

Code and data availability

The CSRFormer clear-sky radiative flux dataset generated in this study is publicly available at Zenodo with DOI: 10.5281/zenodo.20046058 (Zheng et al., 2026). The source code used for data preprocessing, model training, inference, and figure generation is available at https://github.com/Boya928861918/CSRFormer. The third-party input and benchmark datasets used in this study are publicly available from their original providers, including CERES SSF/CRS/SYN1deg, ERA5, MERRA-2, NOAA OISST, and MODIS MYD06, and are cited in the main text and reference list.

The CERES Aqua SSF Edition4A products were obtained from https://ceres-tool.larc.nasa.gov/ord-tool/products?CERESProducts=SSFlevel2_Ed4 (Su et al., 2015b). The CERES SYN1deg Ed4A products were obtained from https://ceres-tool.larc.nasa.gov/ord-tool/jsp/SYN1degEd42Selection.jsp (Doelling et al., 2013; Doelling et al., 2016). The CERES CRS products were obtained from https://ceres-tool.larc.nasa.gov/ord-tool/products?CERESProducts=CRS (Scott et al., 2022). ERA5 hourly data is available as part of the "ERA5 hourly data on single levels from 1940 to present" was extracted from https://cds.climate.copernicus.eu (accessed on 13 November 2023) (Hersbach et al., 2020). The NOAA SST reanalysis data used in this study were obtained from: https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.highres.html (Schlax et al., 2007). The MERRA-2 2d,3-Hourly,Instantaneous,Single-Level,Assimilation,Aerosol Optical Depth Analysis V5.12.4 products were obtained from https://disc.gsfc.nasa.gov/datasets/M2I3NXGAS_5.12.4/summary and the MERRA-2 inst3_3d_aer_Nv: 3d,3-Hourly,Instantaneous,Model-Level,Assimilation,Aerosol Mixing Ratio were obtained from https://disc.gsfc.nasa.gov/datasets/M2I3NVAER_5.12.4/summary (Gelaro et al., 2017). The MYD06 data were obtained from https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/MYD06_L2/ (Platnick, 2015).

Author contributions

Boyang Zheng: Data curation, Formal analysis, Investigation, Methodology, Validation, Writing – original draft. Yannian Zhu: Conceptualization, Funding acquisition, Supervision, Project administration. Minghuai Wang, Daniel Rosenfeld, and Jihu Liu: Supervision, Project administration. Yang Cao: Conceptualization, Methodology, Writing – review & editing. Kang-En Huang and Yichuan Wang: Methodology, Formal analysis. Jun Shi and Yicheng Wei: Formal analysis.

Competing interests

The contact author has declared that none of the authors has any competing interests.

Disclaimer

Copernicus Publications adds a standard disclaimer: "Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher."

Acknowledgements

We acknowledge the teams of CERES, ERA5, MERRA-2, NOAA OISST, and MODIS for providing the datasets used in this study.

Financial support

This research is supported by Natural Science Foundation of China (U2342223, 42475089, and 91744208), the Ministry of Science and Technology of China (Grant 2024YFF0809403), and National Key R&D Program of China (Grant 2023YFC3007600), also supported by the Collaborative Innovation Center of Climate Change, Jiangsu Province.

References

Allan, R. P.: Combining satellite data and models to estimate cloud radiative effect at the surface and in the atmosphere, Meteorological Applications, 18, 324-333, 10.1002/met.285, 2011.

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.

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., Boucher, O., Carslaw, K. S., Christensen, M., Daniau, A. L., Dufresne, J. L., Feingold, G., Fiedler, S., Forster, P., Gettelman, A., Haywood, J. M., Lohmann, U., Malavelle, F., Mauritsen, T., McCoy, D. T., Myhre, G., Mulmenstadt, J., Neubauer, D., Possner, A., Rugenstein, M., Sato, Y., Schulz, M., Schwartz, S. E., Sourdeval, O., Storelvmo, T., Toll, V., Winker, D., and Stevens, B.: Bounding Global Aerosol Radiative Forcing of Climate Change, Rev Geophys, 58, e2019RG000660, 10.1029/2019RG000660, 2020.

Beucler, T., Pritchard, M., Rasp, S., Ott, J., Baldi, P., and Gentine, P.: Enforcing analytic constraints in neural networks emulating physical systems, Physical review letters, 126, 098302, 2021.

Bond, T. C., Doherty, S. J., Fahey, D. W., Forster, P. M., Berntsen, 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.

Bony, S., Stevens, B., Frierson, D. M. W., Jakob, C., Kageyama, M., Pincus, R., Shepherd, T. G., Sherwood, S. C., Siebesma, A. P., Sobel, A. H., Watanabe, M., and Webb, M. J.: Clouds, circulation and climate sensitivity, Nature Geoscience, 8, 261-268, 10.1038/ngeo2398, 2015.

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.

Castro, S. L., Wick, G. A., and Steele, M.: Validation of satellite sea surface temperature analyses in the Beaufort Sea using UpTempO buoys, Remote Sensing of Environment, 187, 458-475, 10.1016/j.rse.2016.10.035, 2016.

Dietel, B., Sourdeval, O., and Hoose, C.: Characterisation of low-base and mid-base clouds and their thermodynamic phase over the Southern Ocean and Arctic marine regions, Atmospheric Chemistry and Physics, 24, 7359-7383, 10.5194/acp-24-7359-2024, 2024.

Doelling, D. R., Sun, M., Nguyen, L. T., Nordeen, M. L., Haney, C. O., Keyes, D. F., and Mlynczak, 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.

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.

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

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.

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.

Forster, P., T. Storelvmo, K. Armour, W. Collins, J.-L. Dufresne, D. Frame, D.J. Lunt, T. Mauritsen, M.D. Palmer, M. Watanabe, M. Wild, and H. Zhang: The earth's energy budget, climate feedbacks, and climate sensitivity. In Climate Change 2021: The physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change., in: Climate Change 2021 – The Physical Science Basis, edited by: Intergovernmental Panel on Climate, C., Cambridge University Press, Cambridge, 923-1054, 10.1017/9781009157896.009, 2021.

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.

Gelaro, R., McCarty, W., Suarez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A., Gu, W., Kim, G. K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), J Clim, Volume 30, 5419-5454, 10.1175/JCLI-D-16-0758.1, 2017.

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.

Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G., and Yacalis, G.: Could Machine Learning Break the Convection Parameterization Deadlock?, Geophysical Research Letters, 45, 5742-5751, 10.1029/2018gl078202, 2018.

Grundner, A., Beucler, T., Gentine, P., Iglesias-Suarez, F., Giorgetta, M. A., and Eyring, V.: Deep Learning Based Cloud Cover Parameterization for ICON, Journal of Advances in Modeling Earth Systems, 14, 10.1029/2021ms002959, 2022.

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.

Gupta, H. and Sil, S.: Assessment of GHRSST and OISST Datasets in Identification of Marine Heat-Waves and Heat-Spikes, IEEE Geoscience and Remote Sensing Letters, 21, 1-5, 10.1109/lgrs.2024.3362474, 2024.

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.

Henderson, D. S., L'Ecuyer, T., Stephens, G., Partain, P., and Sekiguchi, M.: A Multisensor Perspective on the Radiative Impacts of Clouds and Aerosols, Journal of Applied Meteorology and Climatology, 52, 853-871, 10.1175/jamc-d-12-025.1, 2013.

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.

Huang, B., Liu, C., Banzon, V., Freeman, E., Graham, G., Hankins, B., Smith, T., and Zhang, H.-M.: Improvements of the Daily Optimum Interpolation Sea Surface Temperature (DOISST) Version 2.1, Journal of Climate, 34, 2923-2939, 10.1175/jcli-d-20-0166.1, 2021.

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.

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., and Yang, L.: Physics-informed machine learning, Nature Reviews Physics, 3, 422-440, 10.1038/s42254-021-00314-5, 2021.

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

Open Access Earth System Science Data Discussions

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.

Konsta, D., Dufresne, J. L., Chepfer, H., Vial, J., Koshiro, T., Kawai, H., Bodas-Salcedo, A., Roehrig, R., Watanabe, M., and Ogura, T.: Low-Level Marine Tropical Clouds in Six CMIP6 Models Are Too Few, Too Bright but Also Too Compact and Too Homogeneous, Geophysical Research Letters, 49, 10.1029/2021gl097593, 2022.

Krasnopolsky, V. M.: Neural network emulations for complex multidimensional geophysical mappings: Applications of neural network techniques to atmospheric and oceanic satellite retrievals and numerical modeling, Reviews of Geophysics, 45, 10.1029/2006rg000200, 2007.

L'Ecuyer, T. S., Hang, Y., Matus, A. V., and Wang, Z.: Reassessing the Effect of Cloud Type on Earth's Energy Balance in the Age of Active Spaceborne Observations. Part I: Top of Atmosphere and Surface, Journal of Climate, 32, 6197-6217, 10.1175/jcli-d-18-0753.1, 2019.

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.

Matus, A. V. and L'Ecuyer, T. S.: The role of cloud phase in Earth's radiation budget, Journal of Geophysical Research: Atmospheres, 122, 2559-2578, 10.1002/2016jd025951, 2017.

Nam, C., Bony, S., Dufresne, J. L., and Chepfer, H.: The 'too few, too bright' tropical low-cloud problem in CMIP5 models, Geophysical Research Letters, 39, n/a-n/a, 10.1029/2012gl053421, 2012.

Noel, V., Chepfer, H., Chiriaco, M., and Yorks, J.: The diurnal cycle of cloud profiles over land and ocean between 51° S and 51° N, seen by the CATS spaceborne lidar from the International Space Station, Atmospheric Chemistry and Physics, 18, 9457-9473, 10.5194/acp-18-9457-2018, 2018.

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, 1972.

Pincus, R. and Evans, K. F.: Computational Cost and Accuracy in Calculating Three-Dimensional Radiative Transfer: Results for New Implementations of Monte Carlo and SHDOM, Journal of the Atmospheric Sciences, 66, 3131-3146, 10.1175/2009jas3137.1, 2009.

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.

Platnick, S., Ackerman, S., King, M., Meyer K., Menzel W., Holz R., Baum B., and Yang P.: MODIS Atmosphere L2 Cloud Product (06_L2) [dataset], 10.5067/MODIS/MYD06_L2.061, 2015.

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.

Rahaman, N., Baratin, A., Arpit, D., Draxler, F., Lin, M., Hamprecht, F., Bengio, Y., and Courville, A.: On the spectral bias of neural networks, International conference on machine learning, 5301-5310,

Ramanathan, V., Cess, R., Harrison, E., Minnis, P., Barkstrom, B., Ahmad, E., and Hartmann, D.: Cloud-radiative forcing and climate: Results from the Earth Radiation Budget Experiment, Science, 243, 57-63, 1989.

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.

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

Rasp, S. and Lerch, S.: Neural Networks for Postprocessing Ensemble Weather Forecasts, Monthly Weather Review, 146, 3885-3900, 10.1175/mwr-d-18-0187.1, 2018.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, Nature, 566, 195-204, 10.1038/s41586-019-0912-1, 2019.

Rolnick, D., Veit, A., Belongie, S., and Shavit, N.: Deep learning is robust to massive label noise, arXiv preprint arXiv:1705.10694, 2017.

Samset, B. H., Myhre, G., Schulz, M., Balkanski, Y., Bauer, S., Berntsen, 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.

Schlax, M. G., Casey, K. S., Chelton, D. B., Liu, C., Smith, T. M., and Reynolds, R. W.: Daily High-Resolution-Blended Analyses for Sea Surface Temperature, Journal of Climate, 20, 5473-5496, 10.1175/2007jcli1824.1, 2007.

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.

Sun, W., Videen, G., Kato, S., Lin, B., Lukashin, C., and Hu, Y.: A study of subvisual clouds and their radiation effect with a synergy of CERES, MODIS, CALIPSO, and AIRS data, Journal of Geophysical Research: Atmospheres, 116, n/a-n/a, 10.1029/2011jd016422, 2011.

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.

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.

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.

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, 2002.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I.: Attention is all you need, Advances in neural information processing systems, 30, 2017.

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.

https://doi.org/10.5194/essd-2026-360
Preprint. Discussion started: 11 August 2026

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, 1996.

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.

Wood, R.: Stratocumulus Clouds, Monthly Weather Review, 140, 2373-2423, 10.1175/mwr-d-11-00121.1, 2012.

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.

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.