https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

Physics-Constrained Transfer Learning with a Spectral-Fidelity-Preserving Model for Satellite Remote Sensing Applications

Min Min [1], Qiang Yu [1], Jun Li [2*], Han Lin[3], Yunheng Xue [4]

Xinran Xia [1], Di Di [5], Bo Li [2], Peng Zhang [6]

1. School of Atmospheric Sciences and Guangdong Province Key Laboratory for Climate Change and Natural Disaster Studies, Sun Yat-sen University and Southern Laboratory of Ocean Science and Engineering, Zhuhai 519082, China
2. National Satellite Meteorological Center (National Centre for Space Weather), Innovation Center for FengYun Meteorological Satellite (FYSIC), and Key Laboratory of Radiometric Calibration and Validation for Environmental Satellites, China Meteorological Administration (CMA), Beijing 100081, China
3. Key Laboratory of Spatial Data Mining and Information sharing of Ministry of Education, National & Local Joint Engineering Research Center of Satellite Geospatial Information Technology, and Academy of Digital China (Fujian), Fuzhou University, Fuzhou 350108, China
4. Key Laboratory of Ecosystem Carbon Source and Sink, China Meteorological Administration (ECSS-CMA), Wuxi University, 214063, Wuxi, China
5. School of Atmospheric Physics, Nanjing University of Information Science and Technology, Nanjing 210044, China
6. China Meteorological Administration (CMA) Meteorological Observation Center, Beijing 100081, China

Corresponding Author: Jun Li (junli@cma.gov.cn)

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

Highlight 1: Physics-constrained transfer learning framework enables rapid transfer of retrieval models across different satellite sensors while preserving physical consistency.

Highlight 2: Both the calibration uncertainty and the similarity of spectral response function of satellite channel jointly affect the performance of the transfer learning framework.

Highlight 3: The effectiveness of the transfer learning tool is demonstrated through its application to cloud amount profile and precipitation retrieval models based on the satellite observations.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

Abstract

Accurate spectral transformation across satellite sensors with similar but different spectral response functions (SRFs) are essential for applying the same retrieval algorithms. A novel physics-constrained transfer learning (TL) framework is developed for transferring satellite radiance observations across different sensors while preserving physical consistency. It integrates a core Spectral-Fidelity-Preserving (SFP) model based on extensive radiative transfer simulations, allowing broad adaptability for radiance transformation under diverse satellite observational conditions. Sensitivity experiments demonstrate the robustness of the TL framework relating to radiometric calibration uncertainties, particularly in infrared (IR) channels, and further highlight the critical role of SRF similarity between sensors. Specifically, the scaling factor between the SRFs of the target and reference channels should be constrained within the range of 0.5 – 1.5. Meanwhile, the shift in central wavenumber should remain below 200 cm^{-1} for visible or near IR channels, and more strictly below 20 cm^{-1} for infrared window channels (e.g., 10.80 μm). Applying to radiance observations from Fengyun-4A/B (FY-4A/B) geostationary (GEO) satellites explicitly indicates that the TL approach improves retrieval accuracy for key geophysical parameters such as cloud amount profile and quantitative precipitation estimation, when compared those without applying TL. Thus, the TL approach enhances cross-satellite data consistency and provides a practical tool for operational satellite data applications (e.g., adopt algorithms of F-4A to FY-4B without operational interruption).

Keywords: Radiative transfer, transfer learning, Physics-Constrained

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

1. Introduction

For operational continuity, machine learning (ML) enables rapid spectral transformation between similar spaceborne sensors with slightly different response functions, allowing algorithms trained on one sensor to be applied seamlessly to another without interruption. The history of transfer learning (TL) with ML dates back to the 1990s (Caruana, 1997), which was introduced to reuse knowledge across tasks or domains when labeled data or training resources are limited. After 2006, with the rise of artificial intelligence (AI), TL entered a new era of rapid development (Hinton et al., 2006; Lecun et al., 2015). Current TL techniques can be broadly categorized into five types: fine-tuning-based transfer learning (FTL), multi-task learning (MTL), few-shot learning (FL), unsupervised domain adaptation (UDA), and self-supervised learning (SSL) (Ma et al., 2024). Each approach has emerged in response to rapid technological advances and evolving application scenarios. For example, the emergence of pre-trained models (e.g., ImageNet) (Krizhevsky et al., 2012) particularly enabled models to adapt to various downstream tasks through fine-tuning, establishing it as a core paradigm in modern TL (Pan and Yang, 2010). Besides, in recent years, large language models (LLM, e.g., the GPT (Generative Pre-trained Transformer) series) have further advanced the widespread application of TL, enabling cross-domain knowledge transfer from image processing to natural language processing (Jablonka et al., 2024; Joseph et al., 2024). Notably, the applications of TL in remote sensing have also grown explosively since 2017, including land-cover mapping, soil property estimation, vegetation monitoring, natural disaster management, etc. (Ma et al., 2024).

In remote sensing, TL is most widely applied to land-cover mapping and vegetation monitoring, accounting for approximately 48% and 17% of reported applications, respectively (Ma et al., 2024; Li et al., 2025). In land-cover mapping, models are commonly trained on high-accuracy, high-resolution land-cover reference data or obtained by fine-tuning pre-trained networks and are then transferred to other unmanned aerial vehicle (UAV) or satellite imagery for broader application. For example, Nowakowski et al. (2021) similarly fine-tuned VGG-16 (Visual Geometry Group 16-layer network) and GoogLeNet on high-resolution imagery for crop mapping in Africa, achieving accuracies of approximately 83%–90% (Nowakowski et al., 2021). More recently, Zermatten et al. (2025) propose an open-vocabulary

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

land-cover segmentation model that aligns pixel-level visual features with language embeddings using a pre-trained text encoder and a text augmentation strategy, enabling more flexible transfer across land-cover datasets with different label sets (Zermatten et al., 2025). The key concept of method is similar to that used in land-cover mapping, but vegetation monitoring primarily focuses on vegetation health, biomass estimation, and productivity. For instance, Bhadra et al. (2024) proposed a TL dual-stream neural network of PROSAIL-Net that leverages prior knowledge from numerical simulations to improve estimates of corn leaf chlorophyll concentration and average leaf angle under limited ground-truth data (Bhadra et al., 2024). Zhou et al. (2023) presented a deep TL framework that pretrains a Bi-LSTM (Long Short-Term Memory) model on MODIS (Moderate Resolution Imaging Spectroradiometer) reflectance and LAI (leaf area index) to learn temporally informed reflectance–LAI relationships, and then applies the transferred model to generate long-term Landsat LAI maps, outperforming traditional retrieval methods in both accuracy and robustness (Zhou et al., 2023).

With the exception of land-cover mapping and vegetation monitoring, natural disaster management ranks as the third most prevalent remote sensing application of TL (Ma et al., 2024), accounting for approximately 14% of studies. In this domain, low-resolution meteorological satellite observations (Xia et al., 2024; Zhou et al., 2024a) consistently leverage knowledge from historical disaster or weather records to facilitate faster and more effective response mechanisms. Generally, dominant TL applications of meteorological satellites remote sensing primarily follow a self-supervised learning or unsupervised domain adaptation paradigm, often combined with fine-tuning-based TL. For example, several studies have applied data-driven TL by fine-tuning ImageNet-pre-trained models or multi-task feature TL to improve tropical cyclone (TC) center localization using geostationary satellite measurements (Zhou et al., 2024b; Wang and Li, 2023; Lee et al., 2025). For sparsely observed regions such as the Tibetan Plateau, a TL framework (UDA mode) combining deep neural networks (DNNs) and a U-Net (U-shaped convolutional network) was developed to improve MODIS-based estimates of near-surface air temperature, achieving strong agreement with independent stations (R^2 (coefficient of determination) = 0.92, RMSE (root-mean-square error) = 2.29 °C) and reducing RMSE by at least 7% overall (Wang et al., 2025). Similar TL-based U-Net models

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

have been developed and applied to Himawari-8 geostationary (GEO) satellite observations, substantially enhancing the retrieval accuracy of cloud-top height, cloud optical thickness, and cloud effective radius (Li et al., 2023). As another example of combination of FTL and SSL modes, Mateo-García et al. (2020) developed a deep learning model for cloud masking using Landsat-8 imagery and then transferred it for PROBA-V data using spectrally similar channels (Mateo-García et al., 2020). Besides, Wang et al. (2021) also pre-trained a precipitation retrieval model using the U.S. Geostationary Operational Environmental Satellite (GOES) data and then transferred it to generate China Fengyun (FY) GEO meteorological satellite precipitation dataset (Wang et al., 2021).

As discussed above, AI-based application models for meteorological satellites (Min et al., 2020; Wang et al., 2024; KüHnlein et al., 2014; Li et al., 2024) are typically trained on data from a specific time window, often in a SSL mode (e.g., using a full year of data from 2025), and then generalized to subsequent periods (e.g., deployed from 2026 onward). These models have been widely used for a range of retrieval tasks, including cloud, aerosol, precipitation, wind, and convective system (Ma et al., 2021; Miller et al., 2024; Mcgovern et al., 2023; Min et al., 2019; Min et al., 2020; Fu et al., 2024). Particularly in applications like tropical cyclone monitoring (Zhang et al., 2024), where building a long-term labeled training dataset is particularly slow, TL is crucial to maintaining operational continuity and performance consistency. However, in practice, temporal generalization and applicability of TL are frequently coupled with cross-sensor adaptation, where a pre-trained model is transferred to similar satellite instruments through a combination of UDA and FTL.

Accordingly, transfer learning in these AI-based remote sensing models still faces two major challenges: (1) within-series transfer, where models trained on earlier sensors' data must be migrated to newer satellites in the same series (e.g., transferring a model from China FY-4A to the newly launched FY-4B/C satellites); and (2) cross-platform transfer, where algorithms developed for one satellite sensor must be adapted to comparable sensors on other satellite systems (e.g., transferring approaches developed for the U.S. MODIS sensor to similar sensor onboard China Fengyun meteorological satellites). Hence, without an effective TL model, algorithmic consistency and retrieval quality can be severely compromised, which could interrupt

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

operational applications of satellite observations from sensors onboard the same series or different platforms.

It turns out that, Bayes theorem has explicitly indicated that the label marginal distributions differ between the source (subscript: 0) and target (subscript: 1) domains, i.e., $P_0(X) \neq P_1(X)$, which in turn leads to slight differences in the conditional distributions $P_0(X|Y) \approx P_1(X|Y)$ (Ma et al., 2024). In other words, this represents a typical prior shift, where the observation variables (e.g., reflectance (Ref) or brightness temperature (BT) observed by satellite imager) used to pre-train an application model on the source satellite may differ from those available from the target satellite for transfer learning (Figure 1). Figure 1 shows the different radiative signatures of cloud and ocean targets observed by nominally similar channels across satellites, which may degrade TL performance. According to physical theory, differences in $P(X|Y)$ across satellite observation periods are primarily caused by time-varying changes in an instrument measurement and calibration performance (Min et al., 2022; Doelling et al., 2013). In cross-satellite-instrument settings, however, beyond calibration effects, disparities in spectral response characteristics are the most critical driver of such shifts in $P(X|Y)$.

In addition, Figure 2 compares the typical spectral response functions (SRFs) from the visible (VIS) to infrared (IR) channels across China FY-4A/B GEO imagers, the U.S. GOES-16 GEO imager, China FY-3D polar-orbiting imager, and the U.S. Aqua/MODIS (Yang et al., 2017; Platnick et al., 2017; Schmit et al., 2017). As shown in Figure 2, differences in SRFs across satellites lead to substantial discrepancies in how nominally similar VIS and IR channels respond to atmospheric absorption and scattering effects (Min et al., 2017a; Hocking et al., 2021; Pearlman et al., 2013; Wang et al., 2024). Particularly, it shows that the SRFs at 0.86 μm and 2.25 μm channels on these satellite imagers differ in both spectral coverage and shape. Besides, FY-3D/MERSI-II SRFs are roughly 1.5 times wider than those of the corresponding MODIS bands, particularly for the longwave IR channels at 11 and 12 μm. These SRF discrepancies imply that TL for meteorological satellite applications can be highly sensitive to those discrepancies, potentially inducing observation biases (i.e., shifts in $P(X|Y)$ as illustrated in Figure 1) and retrieval errors that are difficult to quantify.

Therefore, it is essential to enable the reliable transfer of observed radiance across similar satellite channels, facilitating the inheritance of generalized retrieval

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

algorithms for scientific products (Wang et al., 2021; Liu et al., 2024), and the propagation of radiometric calibration reference (Yu et al., 2026). For such purpose, a physics-constrained TL framework is developed for satellite remote sensing applications, which bridges the physical measurement (radiance) discrepancies between nominally similar channels across different satellite imaging sensors. Specifically, the core Spectral-Fidelity-Preserving (SFP) model, built on the radiative transfer simulations is developed, its accuracy and robustness are evaluated through numerical sensitivity analyses and retrieval experiments using real satellite SRFs and observations. Theoretically, this new TL framework for satellite remote sensing is grounded in classical atmospheric radiative transfer theories and calculations, thus differs from the conventional machine-learning-based approaches (Ma et al., 2024; Li et al., 2025). Superior transferability and generalization capabilities of the new TL model are achieved by offering stronger physical interpretability.

The remainder of this paper is structured as follows. Section 2 presents the proposed physics-constrained transfer learning framework. Section 3 examines the sensitivity of the core Spectral-Fidelity-Preserving model to satellite spectral response functions and radiometric calibration performance across the visible to infrared channels. Section 4 demonstrates the application of this TL framework to satellite retrieval experiments, highlighting its positive impact on satellite-based applications. Section 5 concludes the paper with a summary of the main findings, along with a discussion of limitations and directions for future work.

2. Physics-constrained Transfer Learning Framework

This section introduces the construction of the core Spectral-Fidelity-Preserving model, along with the overall framework and workflow of the proposed physics-constrained TL approach. In the following sections, we use the primary observation channels of the Advanced Geosynchronous Radiation Imager (AGRI) onboard the FY-4A and FY-4B satellites (Min et al., 2017b) as a test case to evaluate the performance of the SFP model. Because the two instruments have different numbers of channels, we only select 14 matched channels spanning 0.47–13.5 μm for the analysis (Figure 2).

2.1 Spectral-Fidelity-Preserving Model

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

Based on atmospheric radiative transfer theory, the satellite-observed top-of-atmosphere (TOA) monochromatic radiance L^{TOA} in the visible/near-infrared (VIS/NIR) and thermal infrared (TIR) channels can be expressed in the following simplified forms (Kotchenova et al., 2006):

$$L_{VIS/NIR}^{TOA} = L_{scat}T^{\uparrow} + \frac{(1-\varepsilon)}{\pi}E_0\mu_0T^{\downarrow}T^{\uparrow} \quad (1)$$

$$L_{TIR}^{TOA} = \varepsilon B(T_s)T^{\uparrow} + \int_0^{\infty} B(T(z))W(z)dz \quad (2)$$

where L_{scat} denotes the combined contribution from both single and multiple scattering by particles (e.g., clouds and aerosols) as well as gas molecules. ε is the surface emissivity. E_0 and μ_0 represent the incoming solar irradiance and the cosine of the solar zenith angle, respectively. T with arrows indicates the atmospheric transmittance along the downward and upward paths. It is primarily controlled by the total extinction optical depth along the line of sight, which depends on gaseous absorption (e.g., H_2O, CO_2, O_3) and scattering by clouds and aerosols, as well as the atmospheric temperature–pressure structure and viewing geometry (path length/zenith angle). In Eq. (2), B is the Planck function, which is used to compute the emitted radiance associated with the surface temperature T_s. The integral term accounts for the radiance emitted from each atmospheric layer at altitude z, and W(z) is the corresponding weighting function. W(z) quantifies the relative contribution of emission from atmospheric layer (z) to the TOA radiance; it is essentially the product of the layer's absorptivity (or emissivity) and the transmittance from that layer to the sensor, thereby indicating the altitude range to which a given channel is most sensitive (Matricardi, 2010).

From the physical interpretation of the above two equations, the primary factors affecting satellite-observed radiance include the atmospheric thermodynamic structure (temperature and humidity profiles), satellite viewing geometry, surface reflective and emissive properties, and extinction by atmospheric gas composition, clouds, and aerosols. A key scientific challenge for the TL model is therefore to develop a generalizable framework that can accurately propagate these physical effects across satellites for spectrally similar channels. Because these channels are spectrally similar (e.g., the examples shown in Figure 2), we aim to exploit their underlying physical relationships to enable robust transfer of spectral radiance features across satellite observations.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

Building on the rationale outlined above, we develop a physics-constrained TL model using simulated, satellite-sensor-consistent high-resolution spectral radiances. Specifically, we employ the classical MODTRAN (Version 4.2) model (Berk and Hawes, 2017) as the forward radiative transfer model (RTM) to generate TOA radiances (Unit = W/(m^2·sr·μm)) over 0.40–13.5 μm at a spectral resolution of 1 cm^{-1}. For computational convenience, the simulations are conducted over two subranges: 0.3 – 2.4 μm (VIS and NIR) and 3.3 – 13.5 μm (TIR). To ensure diversity and global representativeness, we compile 83 representative atmospheric profiles from the European Centre for Medium-Range Weather Forecasts (ECMWF) archive spanning multiple climate regimes. In addition to the 83 independent atmospheric profiles, we further enhance the diversity of the simulations by specifying a comprehensive set of cloud, aerosol, surface, and viewing-geometry parameters, as summarized in Table 1. This shows that cloud optical depth at 550 nm ranges from 0 (clear sky) to 216 (cumulus), with other four different cloud types. Surface reflectivity (ρ) is prescribed by land-cover type, varying from 0.003 – 0.979 (snow) and 0.08 – 0.83 (forest) to 0.01 – 0.05 (ocean) and 0.019 – 0.137 (urban). The observation geometry is configured with a set of discrete angles, which together cover the major satellite observation modes. Overall, these combinations generate 50,400 simulation cases. The resulting SFP dataset constitutes a large-scale database of simulated satellite hyperspectral radiances, primarily spanning two spectral ranges as noted above: 0.3 – 2.4 μm (VIS/NIR) and 3.3 – 13.5 μm (TIR).

This SFP model was originally developed for the cross-calibration of the reflective solar bands of FY-3D/MERSI-II. By employing TL, it enables the transfer of the radiometric reference from the MODIS sensor while preserving the inherent radiometric observation characteristics of MERSI-II (Yu et al., 2026). Application of this model to the FY-3D/MERSI-II solar reflective channels over a period of more than seven years resulted in a mean relative bias of approximately 1%, with a 2%–3% improvement compared with the operational calibration coefficients.

2.2 Application of New Physics-constrained Transfer Learning Framework

Radiance TL between the target and reference instruments (using FY-4A and FY-4B AGRI as an example) is achieved through an adaptive polynomial framework. In this approach, transfer coefficients are derived by fitting the simulated radiances of

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

both the reference and target sensors, with the theoretical radiances generated from MODTRAN simulations mentioned in section 2.1. The underlying polynomial model is expressed as follows:

$$y = \sum_{i=0}^{n} a_i x^i \tag{3}$$

where y (i.e., the radiance observed by the target sensor) denotes the fitted value of the dependent variable corresponding to the input variable x (i.e., the radiance observed by the reference sensor), n is the polynomial order, a_i represents the polynomial coefficients to be estimated, and i is the index of the polynomial term.

The new physics-constrained TL framework systematically optimizes the fitting procedure. It begins with a linear transformation; if the correlation coefficient is below 0.95, the polynomial order is adaptively increased until the threshold is satisfied. To balance practical applicability and computational efficiency, the maximum polynomial degree is limited to four (n≤4). If the threshold is still not achieved at the fourth order, the model with the highest correlation coefficient is selected. Before fitting, the MODTRAN-simulated hyperspectral TOA radiances are convolved with the SRFs of the target and reference sensors to obtain the corresponding broadband channel radiances. The channel-averaged radiance $L_{Averaged}^{TOA}$ measured by a sensor channel is calculated by:

$$L_{Averaged}^{TOA} = \frac{\int_{\lambda_2}^{\lambda_1} L^{TOA}(\lambda) * SRF(\lambda) d\lambda}{\int_{\lambda_2}^{\lambda_1} SRF(\lambda) d\lambda} \tag{4}$$

where λ_1 and λ_2 denote the starting and ending wavelengths of the normalized spectral response function (SRF), respectively. Overall, Eqs. (3) and (4) are designed to provide physically consistently observed radiance for bridging the target and reference channels. Furthermore, for practical remote sensing applications, the observed radiances of the VIS/NIR channels and IR channels need to be converted into reflectance (ρ, Eq. 5) and brightness temperature (BT, Eq. 6), respectively, which are written as follows:

$$\rho = \frac{\pi L_{Averaged}^{TOA}}{cos(\theta_0) F_0} \tag{5}$$

$$BT = Planck(L_{Averaged}^{TOA}, \lambda_{emw}) \tag{6}$$

where θ_0 is the solar zenith angle, F_0 represents the channel SRF-convolved solar exo-atmospheric irradiance. Additionally, *Planck* denotes the Planck function, and λ_{emw} corresponds to the effective middle wavelength weighted by the channel SRF.

Note that, the IR channels require additional correction because they are not ideal narrowband filters. As a result, a slight difference exists between the nominal central wavelength and the effective central wavelength used in BT calculations based on the Planck function (Chen et al., 2012). The fitting coefficients and Planck-weighted correction coefficients (for IR channels only) for the FY-4A/B AGRI channels used in this study are provided in Tables S1 and S2 of the Supplementary Documentation.

Figure 3 illustrates the TL models used to convert four representative FY-4B/AGRI channels (Channels 1 (0.47 μm), 5 (1.61 μm), 13 (10.80 μm), and 15 (13.30 μm)) to their corresponding FY-4A/AGRI channels within the new physics-constrained framework. This conversion spans VIS, NIR, and TIR spectral bands, and includes a case employing higher-order polynomial fitting. As shown in this figure, the fitting performance is excellent in the solar reflective bands, such as Channels 1 and 5, with only very small fitting residuals. In contrast, the longer-wavelength Channel 15 exhibits not only larger residuals but also a need for a quadratic term in the fitting. Besides the slight differences in the central wavelengths of this channel (see the bottom panel in Figure 2), the larger errors in the TIR bands are primarily caused by the more complex gaseous absorption characteristics at these wavelengths, especially carbon dioxide (CO_2) absorption (Teixeira et al., 2024).

Differences in central wavelengths or SRFs, combined with strong path absorption and emission, introduce substantial nonlinearities into the radiative transfer process. As a result, lower-order linear fitting becomes insufficient, and higher-order terms must be incorporated. Therefore, the next section 3 investigates the sensitivity of the physics-constrained TL model to those discrepancies in the SRFs of satellite sensors and their radiometric calibration uncertainties, using the AGRIs aboard the FY-4A and FY-4B satellites as examples.

3. Transfer Learning Sensitivity to Changes in Spectral Response Functions and Calibration Status

3.1 Simulated SRFs and their Discrepancies

We simulate discrepancies in the SRFs of spaceborne imager by scaling the spectral width and shifting the central wavelength. The spectral width scaling factor ranges from 0.1 to 2.0 in increments of 0.1. The central wavelength is shifted from −100 cm^{-1} (“−” means a leftward shift) to +100 cm^{-1} (“+” means a rightward shift) in

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

increments of 10 cm^{-1} for IR channels. For VIS/NIR channels, due to the relatively large wavenumber, the central wavelength is shifted from −1000 cm^{-1} to +100 cm^{-1} in increments of 100 cm^{-1}. Figure 4 shows examples of these simulated (or adjusted) SRFs for FY-4B/AGRI Channel 2 (0.66 μm) and Channel 13 (10.80 μm), including cases with narrower and wider spectral widths as well as shifts in central wavelength. As shown in Figure 4, the discrepancies between the adjusted SRFs and the original SRFs are still substantial. Whether excessively large SRF differences between the target and reference sensors may affect the performance of this new TL approach is a question that warrants careful evaluation and consideration.

In this investigation, we employ three classical metrics, namely Jensen–Shannon divergence (JSD), Wasserstein distance, and Kolmogorov–Smirnov distance (KS), to evaluate the similarity between the adjusted SRFs and the original SRFs. For the calculation of the Jensen–Shannon divergence (Endres and Schindelin, 2006), the two curves (or SRFs) are first discretized into histograms or grids and represented as probability distributions, which is defined as:

$$P = (p_1, p_2, \dots, p_n),\ \ Q = (q_1, q_2, \dots, q_n) \tag{7}$$

where $\sum p_i = \sum q_i = 1$ and $M = \frac{1}{2}(P+Q)$. JSD is a symmetric and smoothed variant of the Kullback–Leibler divergence (KL), which is defined as follows:

$$JSD(P||Q) = \frac{1}{2}KL(P||M) + \frac{1}{2}KL(Q||M) \tag{8}$$

$$KL(P||M) = \sum_i p_i \log\left(\frac{p_i}{m_i}\right) \tag{9}$$

The JSD is dimensionless and ranges from 0 to 1, with smaller values indicating greater similarity between the two curves.

The second metric is the Wasserstein distance (Mémoli, 2011), also referred to as the Earth Mover's Distance, which measures the minimum transportation cost required to transform one probability distribution into another. For one-dimensional distributions, the first-order Wasserstein distance can be expressed as:

$$W(P,Q) = \int_{-\infty}^{+\infty} |F_P(x) - F_Q(x)|dx \tag{10}$$

where $F_P(x)$ and $F_Q(x)$ are the cumulative distribution functions (CDFs) of P and Q curves, respectively. This metric reflects not only distributional overlap but also the geometric displacement between them.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

The third metric is the Kolmogorov–Smirnov distance (Massey, 1951), which is defined as the maximum absolute difference between the two CDFs and can be written as:

$$D_{KS} = \sup_x |F_P(x) - F_Q(x)| \quad (11)$$

It characterizes the largest pointwise discrepancy between two curve distributions. Together, these three metrics provide complementary perspectives on distributional differences: JSD emphasizes probabilistic similarity, Wasserstein distance captures transportation-based discrepancy, and KS measures the maximum deviation between cumulative distributions.

Figures 5 and 6 show the distributions of Jensen–Shannon divergence, Wasserstein distance, and Kolmogorov–Smirnov distance as functions of the scaling factor (from 0.1 to 2.0) for FY-4B/AGRI Channel 2 (0.66 μm, with shifted wavenumber from −1000 cm^{-1} to +1000 cm^{-1}) and Channel 13 (10.80 μm, with shifted wavenumber from −100 cm^{-1} to +100 cm^{-1}), respectively. From these two figures, the shortwave band, Channel 2 (0.66 μm), exhibits relatively low sensitivity to wavenumber shifts within ±1000 cm^{-1} because of its comparatively broad spectral width (about 3000 cm^{-1}), whereas the narrower longwave band, Channel 13 (10.80 μm), is considerably more sensitive. However, when the wavenumber shift reaches more than 1000 cm^{-1}, even the shortwave band is likely to be affected significantly. Considering the limitation of the manuscript length and the similar distribution characteristics, the results of other FY-4B/AGRI channels are presented in Figures S1 to S13 in the Supplementary Documentation.

Moreover, both the VIS and IR channels of FY-4B/AGRI exhibit much greater sensitivity to variations in SRF width than to shifts in central wavenumber. When the SRF becomes narrower, all three similarity metrics mentioned above increase markedly (smaller values indicate greater similarity); for example, the JSD approaches 1.0. These results indirectly indicate that excessively large discrepancies in central wavelength and SRF width between the target and reference channels could seriously impair the effectiveness of the proposed transfer learning method.

3.2 Evaluation of Fitting Errors in Transfer Learning using Simulated SRFs

To more quantitatively evaluate the impact of SRF discrepancies, fitting experiments are conducted using both simulated SRFs and the original SRFs of FY-4B/AGRI mentioned before. The fitting errors in physics-constrained TL between

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

reference and target channel radiance observations are calculated using simulated FY-4B/AGRI SRFs, with the following configurations: spectral width scaling factor from 0.1 to 2.0 (step 0.1); central wavelength shift from −100 cm⁻¹ to +100 cm⁻¹ (step 10 cm⁻¹) for IR channels, and from −1000 cm⁻¹ to +1000 cm⁻¹ (step 100 cm⁻¹) for VIS/NIR channels.

During the fitting process, the coefficient of determination (R^2) is further used to determine whether a first-order linear fit or a second-order polynomial fit is more appropriate, and it is given by:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{12}$$

where y, $\hat{y}$, and $\bar{y}$ denote the observed value, the fitted value, and the mean of the observed values, respectively. R^2 is commonly used to measure the extent to which the fitted model explains the variability in the observed data, with values closer to 1.0 indicating better agreement between fitted and observed values and stronger explanatory power of the model. In this study, a second-order polynomial fit is adopted when the R^2 value of the initial linear fit is below 0.95.

In addition, three metrics are used for assessing the fitting performance of the TL model, including the RMSE (root-mean-square-error), the correlation coefficient (R), and the number of outliers (N_{outlier}). They can be written as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{13}$$

$$R = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})^2}} \tag{14}$$

where $\bar{\hat{y}}$ denotes the mean of the residuals. For the final metric, N_{outlier}, outliers are identified using the upper and lower fences. Any value lying beyond these two fences is regarded as an abnormal value, defined as:

$$N_{outlier} = \begin{cases} N_{outlier}^{Upper\ Fence} = Q_{25} - 1.5 \times IQR \\ N_{outlier}^{Lower\ Fence} = Q_{75} + 1.5 \times IQR \end{cases} \tag{15}$$

$$IQR = Q_{75} - Q_{25} \tag{16}$$

where Q_{25} and Q_{75} represent the 25th and 75th percentiles of the residuals, respectively. A higher outlier count (N_{outlier}) corresponds to an increased presence of extreme errors (or abnormal value).

Figure 7 shows the distributions of RMSE, R, and N_{outlier} as functions of scaling factor (0.1 – 2.0) and shifted wavenumber (−1000 cm⁻¹ to +1000 cm⁻¹) for

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

FY-4B/AGRI Channel 2 (0.66 μm) and Channel 5 (1.61 μm). The results in the left panel for Channel 2 are generally consistent with those in Figure 5: as the scaling factor increases, the discrepancy between the target and reference SRFs becomes larger, resulting in higher RMSE and N_{outlier} and a lower R. When the scaling factor exceeds 1.8, most cases require second-order polynomial fitting, suggesting that the relationship no longer follows a simple linear conversion. Notably, although scaling factors of about 0.1 – 0.2 correspond to large similarity differences in Figure 5, they do not substantially degrade the TL performance, and the fitting remains predominantly linear. This may be because the narrower target channel still preserves particle scattering characteristics similar to those of the reference channel. By contrast, due to relatively low wavenumber, Channel 5 (1.61 μm) is much more sensitive to wavenumber shifts. Once the shift exceeds about 300 cm^{-1}, the RMSE increases significantly, the correlation coefficient decreases markedly, and both the fitting type and N_{outlier} exhibit abrupt changes.

Figure 8 presents the results for FY-4B/AGRI Channel 9 (6.25 μm) and Channel 13 (10.80 μm). Channel 9 is a typical upper-tropospheric water vapor absorption channel, with its weighting function peaking around 300–400 hPa, corresponding to an altitude of approximately 8 – 9 km (Li et al., 2022). In terms of the correlation coefficient R, Channel 9 appears to be more sensitive to consistency in SRF shape. However, the RMSE results shown in the upper panels indicate that Channel 9 yields smaller fitting errors, suggesting that it is more suitable for observation radiance transfer by using the physics-constrained TL model. It is also noteworthy that, compared with the VIS/NIR channels discussed above (see Figure 7), the IR channels produce N_{outlier} values about one order of magnitude smaller, indicating that particle scattering in the VIS/NIR bands introduces larger deviations into the TL performance.

The results of Figures 7 and 8 indicate that the performance of the physics-constrained TL model is strongly affected by SRF discrepancies. Larger differences in scaling factor or shifted wavenumber generally lead to higher fitting errors, more outliers, and lower correlation, while IR channels show better transfer performance than VIS/NIR channels, likely because particle scattering introduces larger deviations in the latter. As a conservative guideline, for VIS/NIR channels, the difference in central wavenumber should be kept below 200 cm^{-1}, while the scaling factor should be constrained to the range of 0.5 - 1.5. In comparison, IR channels,

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

particularly the water-vapor bands, appear to be somewhat more tolerant. However, for atmospheric window channels around 11 μm, the shifted wavenumber should ideally remain below 20 cm^{-1}, with the scaling factor likewise limited to 0.5 - 1.5. If the similarity between the target and reference channels is too low, reliable conversion becomes infeasible; the resulting conversion errors will then propagate into the retrieval model and further degrade retrieval performance.

3.3 Influence of Radiometric Calibration Uncertainties in Transfer Learning

Beyond SRF differences, radiometric calibration uncertainty in each target channel can also influence the stability and reliability of the physics-constrained TL model. According to Eq. (3) and Fig. 3, when the TL model is linear, i.e., $y=ax+b$, where a and b are fitted coefficients, and x and y denote the radiances in the target (FY-4B/AGRI) and reference (FY-4A/AGRI) channels, respectively, the associated error (σ, or uncertainty) propagation is obtained by differentiating the linear relationship, as follows:

$$\sigma_y = |a| \times \sigma_x \tag{17}$$

However, as shown in Fig. 3d (FY-4B/AGRI Channel 15 centered at 13.30 μm), when the TL model is nonlinear and represented by a quadratic function, i.e., $y=ax^2+bx+c$, the corresponding error propagation can be expressed as follows (first-order approximation):

$$\sigma_y = |2ax + b| \times \sigma_x \tag{18}$$

This equation suggests that the error in the reference-channel radiance (y) is strongly dependent on the magnitude of the radiance observed in the target channel (x).

According to Eq. (5), the propagation of uncertainties or errors in the reflectance observed by the VIS/NIR channels, which are typically on the order of 2–5% (Xiong et al., 2019), can be readily expressed as follows:

$$\sigma_x = |\frac{F_{0,t}}{\pi}| \times \sigma_{\rho,t} \tag{19}$$

By combining Eqs. (5), (17), (18), and (19), the reflectance errors ($\sigma_{\rho,r}$) in the reference channel with different fitting formula (Eqs. (17) and (18)) can be expressed as follows:

$$\sigma_{\rho,r} = |a| \times |\frac{F_{0,t}}{F_{0,r}}| \times \sigma_{\rho,t} \tag{20}$$

$$\sigma_{\rho,r} = |2a\frac{F_{0,t}}{\pi}\rho_t + b| \times |\frac{F_{0,t}}{F_{0,r}}| \times \sigma_{\rho,t} \tag{21}$$

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

where the footnotes of ρ, t, and r respectively signify the reflectance, target, and reference. It should be noted that the formulation of ρ_t does not explicitly include the term of $\cos(\theta_0)$. According to Eq. (20), the propagation of reflectance error in ρ_t is controlled solely by the slope of the TL model and the solar irradiance terms of the two different channels. By comparison, the nonlinear form in Eq. (21) additionally depends on the reflectance observed by the target channel.

Compared with the VIS/NIR channels, the propagation of uncertainties (0.1 ~ 1.0 K) (He et al., 2021) in the BTs measured in the target IR channels is nonlinear and more complex, owing to the additional Planck-function conversion and Planck-weighted correction calculation (Chen et al., 2012). In this study, the final equations for uncertainty propagation in the IR channels are given as Eqs. (A20) and (A21), with the full derivation provided in the Appendix.

To address the inherent calibration-drift errors among different spectral channels, chain-rule-based error-propagation sensitivity experiments were conducted according to Eqs. (20), (21), (A20), and (A21). Figure 9 presents the heatmaps derived from the corresponding error-sensitivity analysis for the FY-4/AGRI sensors based on the physics-constrained TL models. Note that the target and reference designations in Figure 9 are opposite to those in Figure 3. In Figure 9, the A→B case treats FY-4A as the target and FY-4B as the reference, allowing for a symmetric assessment of transfer errors in both directions. For the solar reflective channels, propagated errors are generally attenuated because the error-transfer relationship is predominantly linear, with absolute slopes smaller than unity. However, Channel 4 at 1.38 μm exhibits notably unstable behavior and pronounced error amplification. This channel is primarily sensitive to upper-level ice crystals and is widely used for cirrus cloud detection (Gao et al., 1993). In addition, as shown in Fig. 2, the SRF of FY-4B/AGRI Channel 4 (1.38 μm) covers a substantially narrower range than that of FY-4A/AGRI, which may directly result in a relatively large slope in the fitted TL model and, consequently, a larger difference between the two F_0 terms in Eq. (20). Regardless of the specific mechanism, Fig. 4 further indicates that when the radiometric calibration bias becomes large (e.g., reaching 6%), the resulting transfer error also increases approximately linearly. In other words, channels subject to substantial radiometric calibration bias are not well suited for transfer-learning-based retrieval applications.

In contrast, for most IR channels, small perturbations tend to be amplified, whereas larger perturbations remain close to the magnitude of the original input. A notable exception is FY-4A/AGRI Channel 14 (13.30 μm), which is modeled using a polynomial relationship (see Fig. 3d). The propagated errors for this channel are substantially amplified, owing to pronounced SRF discrepancies and its high sensitivity to atmospheric variability. In addition, the 7.10 and 10.80 μm channels exhibit relatively larger transfer errors because these bands underwent notable upgrades from FY-4A to FY-4B, whereas the remaining channels are generally more stable. Overall, although the IR channels involve a more complicated Planck-function-based conversion (see Eqs. (A20) and (A21)), they are comparatively more stable and thus more suitable for TL applications. This is largely because onboard blackbody calibration constrains the radiometric biases of most IR channels to within about 0.5 K (Min et al., 2022; He et al., 2021), leading to perturbations of a similar magnitude. Relative to typical observed BTs of around 270 K, this corresponds to a relative error of only about 0.2%, which is much smaller than that in the solar reflective channels (about 2% - 5%).

4. Transfer-Learning-Based Satellite Retrieval Experiments

4.1 A Sensitivity Experiment on Retrieving Cloud Amount Profile

Lin et al. (2025) developed CLANN (Cloud Amount Neural Network), a ML framework designed to retrieve three-dimensional (3D) cloud amount profiles from observations acquired by the FY-4A geostationary meteorological satellite passive imager, with important applications in cloud climatology and weather analysis (Lin et al., 2025). The model integrates multi-channel observations from FY-4A/AGRI, including VIS/NIR (0.47, 0.66, and 0.86, 1.38, 1.61, and 2.25 μm), and IR channels (3.75, 6.25, 7.10, 8.60, 10.80, 12.00 μm), together with reanalysis-based environmental predictors. CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations) / CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization) cloud products from 1 March 2018 to 31 December 2020 were used as the reference for model training and independent evaluation. In essence, CLANN learns the vertical distribution of cloud amount from collocated passive-sensor radiances and environmental fields, thereby extending 3D cloud information beyond the narrow swath of active sensors. Validation results showed a correlation coefficient of 0.73 for

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

cloud amount profile retrieval. In addition, the cloud-amount-weighted height derived from CLANN agreed well with CALIOP, with an RMSE of approximately 1.88 km and a correlation coefficient of 0.92. Feature-importance analysis further indicated that AGRI water-vapor-band observation BTs and upper-tropospheric temperature are among the most influential predictors for improving CLANN retrieval skill (Lin et al., 2025).

FY-4B did not enter operational service until December 2022 (Xia et al., 2025). By the time reliable FY-4B/AGRI observations became routinely available, CALIPSO observations had largely ceased in the second half of 2023, leaving only a limited period of effective overlap between FY-4B/AGRI and CALIPSO (approximately June 2022 to March 2023). Consequently, previous work did not attempt to develop an FY-4B/AGRI-based CLANN retrieval model. Although limited, this overlap period is sufficient to conduct a targeted validation experiment for the newly developed physics-constrained TL model. Consequently, previous work did not attempt to develop an FY-4B/AGRI-based CLANN retrieval model. This relatively short overlap period, however, provides a valuable opportunity to test and extend the newly developed physics-constrained TL model in this study. In the first control experiment, FY-4B/AGRI observations from channels corresponding to those used in CLANN training were directly ingested into the FY-4A/AGRI-based CLANN model mentioned above (Lin et al., 2025) to retrieve cloud amount profiles. In the second sensitivity experiment, the newly developed TL models (Figure 3 and Table S1 of the Supplementary Documentation) were applied to convert each FY-4B/AGRI channel observation, in either reflectance (VIS/NIR) or BT (IR) form (target), into its FY-4A/AGRI-equivalent value (reference), which was then used as input to the CLANN model for estimating cloud amount profile. Finally, independent CALIPSO observations from June 1, 2022, to March 31, 2023, were used to comprehensively and accurately evaluate the effectiveness of the TL model.

Figure 10 presents the weighted mean cloud-top height comparisons between CLANN (trained using FY-4A data) and CALIPSO based on FY-4B/AGRI observations, with and without the physics-constrained transfer learning model. As shown in Figures 10b and 10c, the incorporation of the TL model reduces the RMSE by 0.08 km and improves the correlation coefficient from 0.841 to 0.851. Although the R value remains lower than that achieved using FY-4A/AGRI data (R = 0.92) (Lin

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

et al., 2025), the physics-constrained TL model still yields a measurable improvement. In addition, cross-sectional cloud amount profiles for two representative cases (05:35 UTC on June 8, 2022, and 07:15 UTC on June 21, 2022) are presented in the middle and bottom panels, respectively. These typical cases explicitly demonstrate that the physics-constrained TL model significantly improves the estimation of cloud amount profiles derived from FY-4B observations. For the case at 05:35 UTC on June 8, 2022, the TL model more effectively captures the vertical distribution characteristics of cloud amount in Figure 10e, resulting in an increased correlation coefficient and a reduced RMSE.

4.2 A Sensitivity Experiment on Estimating Quantitative Precipitation

We also evaluate the impact of the physics-constrained TL model on quantitative precipitation estimation (QPE) using a Multi-Task UNet model (Jaegle et al., 2021) trained on four FY-4A/AGRI IR channels (C09, C10, C12, and C13 at 6.25, 7.10, 10.80, 12.00 μm), with GPM (Global Precipitation Measurement) IMERG (Integrated Multi-satellitE Retrievals for GPM) products as the reference (Skofronick-Jackson et al., 2017). The FY-4A-based model was trained and validated using observations from 2019, with data from May to August used as the primary training period and randomly selected days from each of the remaining months included to ensure seasonal representativeness across both the validation and inference stages. The study domain covered 90° – 130°E and 0° – 40°N. To assess cross-satellite generalizability, three independent FY-4B/AGRI observation cases were selected from the extrapolation period spanning June 2023 to July 2024, which was entirely outside the FY-4A training window. These cases include Case 1 on 11 June 2024 at 09:15 UTC, Case 2 on 10 August 2024 at 00:00 UTC, and Case 3 on 12 July 2024.

Figures 11a–c and 11d–f show the spatial distributions of QPE results for Cases 1 and 2 mentioned above, respectively, including GPM IMERG rain rates, the raw FY-4B model output, and the TL corrected FY-4B output. Without TL correction, the raw FY-4B model systematically underestimates precipitation in both cases, producing substantially lower rain rates than GPM IMERG. The corresponding CSI (Critical Success Index) values are 0.392 and 0.412 for Cases 1 and 2, respectively. After applying the physics-constrained TL, the predicted precipitation fields become much more consistent with GPM IMERG, with the CSI increasing to 0.498 in both

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

cases. This improvement indicates a robust and reproducible enhancement in precipitation detection skill.

Figures 11g–i further evaluate model performance using scatter plots of predicted rain rates against GPM IMERG estimates. The FY-4A in-domain baseline (Figure 11g; 15 May 2019), selected from the primary training period, achieves a CSI of 0.537, confirming the expected performance of the retrieval model under the original FY-4A observation domain. For the raw FY-4B configuration (Figure 11h), the model yields R (correlation coefficient) = 0.462, MAE (Mean Absolute Error) = 1.678 mm h^{-1}, and bias = −0.959 mm h^{-1}, indicating a systematic underestimation when the FY-4A-trained model is directly applied to FY-4B observations without correction. The underestimation of precipitation, especially for extreme rainfall intensities, is a common issue in AI-based QPE models (Kühnlein et al., 2014). After TL correction (Figure 11i), R increases to 0.475, MAE decreases to 1.564 mm h^{-1}, and bias improves to −0.450 mm h^{-1}, demonstrating a reduction in systematic error.

The latitudinal rain-rate profiles in Figures 11j–l further confirm the improvement across all three cases. The raw FY-4B model consistently underestimates the precipitation structure represented by GPM IMERG, whereas the transfer-learning-corrected FY-4B output better captures the overall latitudinal variability. The consistent improvement across Cases 1–3, which were selected from the FY-4B extrapolation period and span different months and meteorological conditions, demonstrates that the physics-constrained TL model can effectively recover the retrieval capability of the FY-4A-trained model when applied to FY-4B observations.

5. Discussions and Conclusions

This paper presents a recently developed physics-constrained TL framework for satellite remote sensing applications, which differs fundamentally from conventional AI-based TL approaches, offering strong physical interpretability and high generalization capability. The core of this physics-constrained TL framework is a Spectral-Fidelity-Preserving model built on extensive radiative transfer computations, enabling the transformation of radiance observations across different satellites and similar spectral channels spanning 0.47 – 13.5 μm. The TL is implemented using either linear or quadratic regression models, and the approach can be well applied to

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

satellite radiometric calibration transfer, cross-satellite retrieval model adaptation, and related applications.

We further examine the quantitative effects of satellite channel SRF differences, including spectral width variations and central wavelength shifts, as well as radiometric uncertainties, on the accuracy of this physics-constrained transfer learning framework. To validate this approach, we performed some sensitivity experiments using a neural-network-based cloud amount profile retrieval model and a quantitative precipitation estimation model with transferred radiometric observations between FY-4A and FY-4B GEO satellite imagers. The results clearly demonstrate that the proposed physics-constrained TL technique improves the retrieval accuracy of key satellite-derived parameters.

However, from both theoretical and practical perspectives, several key issues regarding the applicability of this physics-constrained transfer learning framework need to be further investigated.

(1) According to Table 1 and the description in Section 2, the core Spectral-Fidelity-Preserving model is constructed based on 50,400 sets of simulated results under diverse conditions. From a sample diversity perspective, there is currently no need to further increase the observation geometries or atmospheric conditions to enhance representativeness. The wide distribution of observed radiances shown in Figure 2 also directly confirms the reliability of the training dataset. Compared with mainstream AI-based TL approaches (Ma et al., 2024), the proposed physics-constrained TL model exhibits distinct advantages. Only a single large-scale radiative transfer computation is required, after which the model can be applied conveniently. Users need only provide the SRFs of the target and reference satellite channels to rapidly fit the observation radiance transformation model of training quality, without undergoing the extensive retraining and testing process required by AI-based models. Consequently, the model demonstrates excellent generalization capability, eliminating the need for substantial computational resources for repeated modeling. As long as the SRFs of the satellite channels to be transferred fall within the 0.47 – 13.5 μm spectral range, the transformation model can be directly fitted. Notably, the core SFP model within this TL framework preserves the simulated information from satellite-based hyperspectral observations, encompassing physical

processes such as atmospheric reflection and absorption, thereby effectively accounting for differences in fitting performance across channels.

(2) As shown in Section 3, the proposed TL framework is most sensitive to differences in the SRFs among satellite channels. Figures 7 and 8 illustrate that IR channels achieve superior fitting performance, with the scaling factor required to remain within the 0.5 – 1.5 range and the central wavelength shift below 200 cm^{-1}. Notably, channels in the IR atmospheric window are highly sensitive to central wavelength shifts, which should be controlled within 20 cm^{-1}. Prior to applying this transfer learning model, it is recommended to evaluate the SRF similarity between target and reference channels using metrics such as Jensen-Shannon divergence, Wasserstein distance, and Kolmogorov-Smirnov distance, as described in Section 3.2. The empirical analysis in Figures 7 – 8 shows that when JSD $\leqslant$ 0.3, the RMSE remains low and $R^2 > 0.95$, while exceeding this threshold leads to a marked increase in fitting errors. Therefore, JSD $\leqslant$ 0.3 is recommended as an empirical threshold for channel transferability.

In addition to SRF differences, radiometric calibration uncertainties of individual channels also affect TL model performance. Overall, the physics-constrained TL model is relatively insensitive to calibration uncertainties in IR channels (errors < 0.2%). In contrast, errors in VIS/NIR channels propagate proportionally to their actual calibration uncertainties. State-of-the-art satellite imagers exhibit calibration errors of approximately 1.5% in the VIS/NIR channels (Choi et al., 2024), an order of magnitude higher than for IR channels. This indicates that the use of this TL model requires particular caution when large calibration uncertainties are present in VIS/NIR channels.

(3) Early radiometric calibration studies (Yu et al., 2026) have demonstrated that this TL model can propagate high-quality radiometric calibration standards from one satellite (reference) to lower-quality calibration channels on other satellites, primarily for VIS/NIR channels. In contrast, IR channels have traditionally relied on high-spectral-resolution IR imagers for synchronous nadir overpass (SNO) (Hewison et al., 2013) and error assessment. However, differences in horizontal spatial resolution between imaging instruments (0.25 – 1.0 km) and high-spectral infrared sounders (10 – 20 km) still introduce significant spatial uncertainties and noise when applying SNO techniques to polar-orbiting satellite imagers. The newly developed TL

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

model presented in this study offers a novel approach for calibrating and evaluating IR channels on polar-orbiting satellites.

A key application scenario is the transfer of core geophysical parameter retrieval algorithms across satellites of the same generation or different generations. This scenario has been demonstrated in Section 4 with experiments on cloud amount profile and quantitative precipitation. Currently, the physics-constrained TL model is used to satellites with relatively low spatial resolution, and its extension to higher-resolution Earth-observing satellites (e.g., Sentinel-2 (Drusch et al., 2012)) requires caution. At higher spectral resolutions, atmospheric radiative transfer calculations based on the plane-parallel approximation can introduce distortions, representing a primary limitation and bottleneck of the TL model.

Finally, the physics-constrained TL framework developed in this study is fully open-source and freely available for satellite channels spanning the VIS to IR spectral range (from 0.47 to 13.5 μm). Due to the extremely large size of the radiative transfer simulation dataset, it cannot be uploaded to online repositories. Researchers interested in using this physics-constrained TL tool can contact the first (Min Min, minm5@mail.sysu.edu.cn) or corresponding authors to obtain assistance in generating the relevant TL models.

Data availability. The authors would like to thank the GPM and CALIPSO team for providing the GPM/CALIPSO Level-2 products and acknowledge the PyTorch open-source library (https://gpm1.gesdisc.eosdis.nasa.gov/opendap/GPM_L2/contents.html and https://asdc.larc.nasa.gov/data/CALIPSO/LID_L2_05kmCPro-Standard-V4-21/).

Author contributions. MM proposed the essential research idea. QY, JL, MM, HL, YH, DD, and XX performed the analysis and drafted the manuscript. BL and PZ provided useful comments. All the authors contributed to the interpretation and discussion of results and the revision of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors would like to thank the GPM and CALIPSO team for providing the GPM/CALIPSO Level-2 products and acknowledge the PyTorch open-source library. We also acknowledge the high-performance computing support from School of Atmospheric Science of Sun Yat-sen University. This research was supported by the National Key Research and Development Program of China (Grant No. 2024YFC3711703), the Innovation Group Project of Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai) (No.311024009), the Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai) (No. SML2024SP011), the National Natural Science Foundation of China (Grant No. U2342201), and the Science and Technology Planning Project of Guangdong Province (Grant No. 2023B1212060019).

Appendix:

The Planck-weighted correction method (Chen et al., 2012) likewise employs a linear form, $T_{b,t}=S_t \cdot T_{b,0}+Q_t$, in which $T_{b,t}$ and $T_{b,0}$ respectively represent the brightness temperature (BT) observed by the satellite target IR channel and the corrected BT for the calculation using the Planck function, respectively. By contrast, for the reference IR channel, the observed BT is also corrected using a linear relationship, $T_{b,r}=S_r \cdot T_{b,1}+Q_r$. As in Eq. (17), the corresponding error propagation can be formulated as follows:

$$\sigma_{T_{b,0}} = |\frac{1}{S_t}| \times \sigma_{T_{b,t}} \tag{A1}$$

$$\sigma_{T_{b,r}} = |S_r| \times \sigma_{T_{b,1}} \tag{A2}$$

For a given wavelength λ, the Planck function for radiance is expressed as:

$$L_\lambda(T_{b,0}) = \frac{c_1}{\lambda^5}\frac{1}{\left[exp\left(\frac{c_2}{\lambda T_{b,0}}\right)-1\right]} \tag{A3}$$

where c_1 (=1.1910439×10^{-16} (W·m^2·sr^{-1})) and c_2 (=1.438769×10^{-2} (m·K)) are the constants in the Planck function. To obtain $\frac{dL_\lambda}{dT_{b,0}}$, we first define:

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

$$u = \frac{c_2}{\lambda T_{b,0}} \tag{A4}$$

$$\frac{du}{dT_{b,0}} = -\frac{c_2}{\lambda T_{b,0}^2} \tag{A5}$$

Then, the Eq. (A2) and $\frac{dL_\lambda}{dT_{b,0}}$ are transformed into

$$L_\lambda = \frac{c_1}{\lambda^5}(\mathrm{e}^{\mathrm{u}} - 1)^{-1} \tag{A6}$$

$$\frac{dL_\lambda}{dT_{b,0}} = \frac{c_1}{\lambda^5}\left[-(e^u - 1)^{-2} e^u \frac{du}{dT_{b,0}}\right] \tag{A7}$$

By substituting Eq. (A3) into Eq. (A6), we obtain:

$$\frac{dL_\lambda}{dT_{b,0}} = \frac{c_1 c_2}{\lambda^6 T_{b,0}^2} \frac{exp\left(\frac{c_2}{\lambda T_{b,0}}\right)}{\left[exp\left(\frac{c_2}{\lambda T_{b,0}}\right) - 1\right]^2} \tag{A8}$$

Therefore, according to the first-order error propagation formula, the propagation of observed-radiance errors ($\sigma_{L_{\lambda,t}}$) (Here, we use $L_{\lambda,t}$ to replace L_λ in Eq. (8)) from $T_{b,t}$ and $T_{b,0}$ can be expressed as follows:

$$\sigma_{L_{\lambda,t}} = \left|\frac{c_1 c_2}{\lambda^6 T_{b,0}^2} \frac{exp\left(\frac{c_2}{\lambda T_{b,0}}\right)}{\left[exp\left(\frac{c_2}{\lambda T_{b,0}}\right) - 1\right]^2}\right| \times \sigma_{T_{b,0}} =$$

$$\left|\frac{c_1 c_2}{\lambda^6 [(T_{b,t} - Q_t)/S_t]^2} \frac{exp\left(\frac{c_2}{\lambda[(T_{b,t} - Q_t)/S_t]}\right)}{\left[exp\left(\frac{c_2}{\lambda[(T_{b,t} - Q_t)/S_t]}\right) - 1\right]^2}\right| \times \left|\frac{1}{S_t}\right| \times \sigma_{T_{b,t}} \tag{A9}$$

In turn, the propagation of radiance error (L_λ) into the calculated BT ($T_{b,1}$) can be expressed in terms of $\frac{dT_{b,1}}{dL_\lambda}$. The inverse of the Planck function in Eq. (A3) can be written as:

$$T_{b,1}(L_\lambda) = \frac{c_2}{\lambda \cdot \ln\left(1 + \frac{c_1}{\lambda^5 L_\lambda}\right)} \tag{A10}$$

Then, we first define:

$$A = \frac{c_1}{\lambda^5} \tag{A11}$$

$$u = \ln\left(1 + \frac{A}{L_\lambda}\right) \tag{A12}$$

$$\frac{du}{dL_\lambda} = \frac{1}{1 + A/L_\lambda} \cdot \frac{d(A/L_\lambda)}{dL_\lambda} = \frac{1}{1 + A/L_\lambda}\left(-\frac{A}{L_\lambda^2}\right) \tag{A13}$$

By substituting Eqs. (A11) and (A12) into Eq. (A10), we obtain:

$$T_{b,1} = \frac{c_2}{\lambda \cdot u} \tag{A14}$$

$$\frac{dT_{b,1}}{dL_\lambda} = -\frac{c_2}{\lambda \cdot u^2} \frac{du}{dL_\lambda} \tag{A15}$$

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

Then, we substitute Eq. (A13) into Eq. (A15), we obtain:

$$\frac{dT_{b,1}}{dL_\lambda}=\frac{c_2}{\lambda u^2}\frac{A}{L_\lambda^2(1+A/L_\lambda)} \tag{A16}$$

Finally, by rearranging Eq. (A16), we obtain:

$$\frac{dT_{b,1}}{dL_\lambda}=\frac{c_1c_2/\lambda^6}{L_\lambda\left(L_\lambda+\frac{c_1}{\lambda^5}\right)\left[ln\left(1+\frac{c_1}{\lambda^5 L_\lambda}\right)\right]^2} \tag{A17}$$

At this stage, the TL model established in this study is applied to $L_{\lambda,t}$ to derive the transformed radiance of the reference IR channel, $L_{\lambda,r}$. As described in Section 3.3, for both the linear and quadratic nonlinear TL models, the corresponding error propagation based on Eq. (A17) can be expressed as follows:

$$\sigma_{T_{b,r}}=|S_r|\times\sigma_{T_{b,1}}=|S_r|\times\left|\frac{c_1c_2a/\lambda^6}{(ax+b)\left(ax+b+\frac{c_1}{\lambda^5}\right)\left[ln\left(1+\frac{c_1}{\lambda^5(ax+b)}\right)\right]^2}\right|\times\sigma_x \tag{A18}$$

$$\sigma_{T_{b,r}}=|S_r|\times\sigma_{T_{b,1}}=|S_r|\times\left|\frac{c_1c_2(2ax+b)/\lambda^6}{(ax^2+bx+c)\left(ax^2+bx+c+\frac{c_1}{\lambda^5}\right)\left[ln\left(1+\frac{c_1}{\lambda^5(ax^2+bx+c)}\right)\right]^2}\right|\times\sigma_x \tag{A19}$$

where x is the radiance ($L_{\lambda,t}$) observed by target IR channel in Eq. (A9).

According to the equations in Section 3.3, by combining Eqs. (A9), (A18), and (A19), the error propagation from the initially observed BT ($T_{b,t}$) in the target IR channel to the BT ($T_{b,r}$) observed by the final reference IR channel used as model input can be expressed as follows:

$$\sigma_{T_{b,r}}=\left|\frac{S_r}{S_t}\right|\times\left|\frac{c_1c_2a/\lambda^6}{(ax+b)\left(ax+b+\frac{c_1}{\lambda^5}\right)\left[ln\left(1+\frac{c_1}{\lambda^5(ax+b)}\right)\right]^2}\right|\times$$
$$\left|\frac{c_1c_2}{\lambda^6[(T_{b,t}-Q_t)/S_t]^2}\frac{exp\left(\frac{c_2}{\lambda[(T_{b,t}-Q_t)/S_t]}\right)}{\left[exp\left(\frac{c_2}{\lambda[(T_{b,t}-Q_t)/S_t]}\right)-1\right]^2}\right|\times\sigma_{T_{b,t}} \tag{A20}$$

$$\sigma_{T_{b,r}}=\left|\frac{S_r}{S_t}\right|\times\left|\frac{c_1c_2(2ax+b)/\lambda^6}{(ax^2+bx+c)\left(ax^2+bx+c+\frac{c_1}{\lambda^5}\right)\left[ln\left(1+\frac{c_1}{\lambda^5(ax^2+bx+c)}\right)\right]^2}\right|\times$$
$$\left|\frac{c_1c_2}{\lambda^6[(T_{b,t}-Q_t)/S_t]^2}\frac{exp\left(\frac{c_2}{\lambda[(T_{b,t}-Q_t)/S_t]}\right)}{\left[exp\left(\frac{c_2}{\lambda[(T_{b,t}-Q_t)/S_t]}\right)-1\right]^2}\right|\times\sigma_{T_{b,t}} \tag{A21}$$

where x is calculated from the Planck function in Eq. (A3). a, b, and c represent the fitted coefficients of the IR channel TL model introduced above. We will also use the derived error-propagation Eqs. (A20) and (A21) to conduct sensitivity experiments in Section 3.3.

References:

Berk, A. and Hawes, F.: Validation of MODTRAN®6 and its line-by-line algorithm, Journal of Quantitative Spectroscopy and Radiative Transfer, 203, 542-556, https://doi.org/10.1016/j.jqsrt.2017.03.004, 2017.

Bhadra, S., Sagan, V., Sarkar, S., Braud, M., Mockler, T. C., and Eveland, A. L.: PROSAIL-Net: A transfer learning-based dual stream neural network to estimate leaf chlorophyll and leaf angle of crops from UAV hyperspectral images, ISPRS Journal of Photogrammetry and Remote Sensing, 210, 1-24, https://doi.org/10.1016/j.isprsjprs.2024.02.020, 2024.

Caruana, R.: Multitask Learning, Machine Learning, 28, 41-75, 10.1023/A:1007379606734, 1997.

Chen, Y., Weng, F., Han, Y., and Liu, Q.: Planck-Weighted Transmittance and Correction of Solar Reflection for Broadband Infrared Satellite Channels, Journal of Atmospheric and Oceanic Technology, 29, 382-396, https://doi.org/10.1175/JTECH-D-11-00102.1, 2012.

Choi, T., Cao, C., Blonski, S., Shao, X., and Wang, W.: NOAA-20 VIIRS On-Orbit Reflective Solar Band Radiometric Calibration Five-Year Update, IEEE Transactions on Geoscience and Remote Sensing, 62, 1-10, 10.1109/TGRS.2024.3363661, 2024.

Doelling, D. R., Daniel, M., Benjamin, R. S., Rajendra, B., and Arun, G.: The characterization of deep convective clouds as an invariant calibration target and as a

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

visible calibration technique, IEEE Transactions on Geoscience and Remote Sensing, 51, 1147-1159, 2013.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., and Bargellini, P.: Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services, Remote Sensing of Environment, 120, 25-36, https://doi.org/10.1016/j.rse.2011.11.026, 2012.

Endres, D. M. and Schindelin, J. E.: A new metric for probability distributions, IEEE Trans. Inf. Theor., 49, 1858–1860, 10.1109/tit.2003.813506, 2006.

Fu, D., Shi, H., Gueymard, C. A., Yang, D., Zheng, Y., Che, H., Fan, X., Han, X., Gao, L., Bian, J., Duan, M., and Xia, X.: A Deep-Learning and Transfer-Learning Hybrid Aerosol Retrieval Algorithm for FY4-AGRI: Development and Verification over Asia, Engineering, 38, 186-198, 2024.

Gao, B.-C., Goetz, A. F. H., and Wiscombe, W. J.: Cirrus cloud detection from Airborne Imaging Spectrometer data using the 1.38 µm water vapor band, Geophys Res Lett, 20, 301-304, https://doi.org/10.1029/93GL00106, 1993.

He, X., Xu, N., Feng, X., Hu, X., Xu, H., and Peng, Y.: Assessing radiometric calibration of FY- 4A/AGRI thermal infrared channels using CrIS and IASI, IEEE Transactions on Geoscience and Remote Sensing, 10.1109/TGRS.2021.3111975, 2021.

Hewison, T., Wu, X., Yu, F., and Hu, X.: GSICS inter-calibration of infrared channels of geostationary imagers using Metop/IASI, IEEE Transactions on Geoscience and Remote Sensing, 51, 1160–1170, 2013.

Hinton, G. E., Osindero, S., and Teh, Y.-W.: A Fast Learning Algorithm for Deep Belief Nets, Neural Computation, 18, 1527-1554, 10.1162/neco.2006.18.7.1527, 2006.

Hocking, J., Vidot, J., Brunel, P., Roquet, P., Silveira, B., Turner, E., and Lupu, C.: A new gas absorption optical depth parameterisation for RTTOV version 13, Geosci. Model Dev., 14, 2899-2915, 10.5194/gmd-14-2899-2021, 2021.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

Jablonka, K. M., Schwaller, P., Ortega-Guerrero, A., and Smit, B.: Leveraging large language models for predictive chemistry, Nature Machine Intelligence, 6, 161-169, 10.1038/s42256-023-00788-1, 2024.

Jaegle, A., Gimeno, F., Brock, A., Vinyals, O., Zisserman, A., and Carreira, J.: Perceiver: General Perception with Iterative Attention, Proceedings of the 38th International Conference on Machine Learning, Proceedings of Machine Learning Research, 2021/07/18, 4651-4664, 10.48550/arXiv.2103.03206,

Joseph, S., Parthi, A. G., Maruthavanan, D., Jayaram, V., Veerapaneni, P. K., and Parlapalli, V.: Transfer Learning in Natural Language Processing, 2024 7th International Conference on Information and Communications Technology (ICOIACT), 20-21 Nov. 2024, 30-36, 10.1109/ICOIACT64819.2024.10912895,

Kotchenova, S. Y., Vermote, E. F., Matarrese, R., and Klemm, J. F. J.: Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part I: Path radiance, Applied Optics, 45, 6762-6774, 10.1364/AO.45.006762, 2006.

Krizhevsky, A., Sutskever, I., and Hinton, G. E.: ImageNet Classification with Deep Convolutional Neural Networks, Advances in Neural Information Processing Systems, 1106-1114,

Kühnlein, M., Appelhans, T., Thies, B., and Nauß, T.: Precipitation estimates from MSG SEVIRI daytime, nighttime, and twilight data with random forests, Journal of Applied Meteorology and Climatology, 53, 2457-2480, 10.1175/JAMC-D-14-0082.1, 2014.

Kühnlein, M., Appelhans, T., Thies, B., and Nauss, T.: Improving the accuracy of rainfall rates from optical satellite sensors with machine learning — A random forests-based approach applied to MSG SEVIRI, Remote Sensing of Environment, 141, 129-143, 10.1016/j.rse.2013.10.026, 2014.

LeCun, Y., Bengio, Y., and Hinton, G.: Deep learning, Nature, 521, 436-444, 10.1038/nature14539, 2015.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

Lee, J., Moon, I.-J., Im, J., Kim, D.-H., and Jung, H.: Multi-task feature transfer deep learning-based tropical cyclone center estimation (MFT–TC) using geostationary satellite observations, Weather and Climate Extremes, 49, 100796, https://doi.org/10.1016/j.wace.2025.100796, 2025.

Li, J., Menzel, W. P., Schmit, T. J., and Schmetz, J.: Applications of geostationary hyperspectral infrared sounder observations – progress, challenges, and future perspectives, Bulletin of the American Meteorological Society, 10.1175/BAMS-D-21-0328.1, 2022.

Li, J., Zhang, F., Li, W., Tong, X., Pan, B., Li, J., Lin, H., Letu, H., and Mustafa, F.: Transfer-Learning-Based Approach to Retrieve the Cloud Properties Using Diverse Remote Sensing Datasets, IEEE Transactions on Geoscience and Remote Sensing, 61, 1-10, 10.1109/TGRS.2023.3318374, 2023.

Li, J., Zheng, J., Li, B., Min, M., Liu, Y., Liu, C.-Y., Li, Z., Menzel, W. P., Schmit, T. J., Cintineo, J. L., Lindstrom, S., Bachmeier, S., Xue, Y., Ma, Y., Di, D., and Lin, H.: Quantitative applications of weather satellite data in nowcasting: progress and challenges, Journal of Meteorological Research, 38, 399-413, 10.1007/s13351-024-3138-6 2024.

Li, L., Han, L., Ye, Y., Xiang, Y., and Zhang, T.: Deep learning in remote sensing image matching: A survey, ISPRS Journal of Photogrammetry and Remote Sensing, 225, 88-112, https://doi.org/10.1016/j.isprsjprs.2025.04.001, 2025.

Lin, H., Li, J., Min, M., Zhang, F., Wang, K., and Wu, Q.: CLANN: Cloud amount neural network for estimating 3D cloud from geostationary satellite imager, Remote Sensing of Environment, 318, 114600, 10.2139/ssrn.4914490, 2025.

Liu, J., Yu, J., Lin, C., He, M., Liu, H., Wang, W., and Min, M.: Near real-time atmospheric and oceanic science products of Himawari-8 and Himawari-9 geostationary satellites over the South China Sea, Earth System Science Data, 16, 4949–4969, 10.5194/essd-16-4949-2024, 2024.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

Ma, Y., Chen, S., Ermon, S., and Lobell, D. B.: Transfer learning in environmental remote sensing, Remote Sensing of Environment, 301, 113924, https://doi.org/10.1016/j.rse.2023.113924, 2024.

Ma, Z., Li, J., Han, W., Li, Z., Zeng, Q., Menzel, W. P., Schmit, T. J., Di, D., and Liu, C.-Y.: Four-dimensional wind fields from geostationary hyperspectral infrared sounder radiance measurements with high temporal resolution, Geophys Res Lett, 48, e2021GL093794, 10.1029/2021GL093794, 2021.

Massey, F. J., Jr.: The Kolmogorov-Smirnov test for goodness of fit, Journal of the American Statistical Association, 46, 68-78, 10.2307/2280095, 1951.

Mateo-García, G., Laparra, V., López-Puigdollers, D., and Gómez-Chova, L.: Transferring deep learning models for cloud detection between Landsat-8 and Proba-V, ISPRS Journal of Photogrammetry and Remote Sensing, 160, 1-17, https://doi.org/10.1016/j.isprsjprs.2019.11.024, 2020.

Matricardi, M.: A principal component based version of the RTTOV fast radiative transfer model, Quarterly Journal of the Royal Meteorological Society, 136, 1823-1835, 10.1002/qj.680, 2010.

McGovern, A., Chase, R. J., Flora, M., Gagne, D. J., Lagerquist, R., Potvin, C. K., Snook, N., and Loken, E.: A Review of Machine Learning for Convective Weather, Artificial Intelligence for the Earth Systems, 2, e220077, https://doi.org/10.1175/AIES-D-22-0077.1, 2023.

Mémoli, F.: Gromov–Wasserstein Distances and the Metric Approach to Object Matching, Foundations of Computational Mathematics, 11, 417-487, 10.1007/s10208-011-9093-5, 2011.

Miller, L., Pelletier, C., and Webb, G. I.: Deep Learning for Satellite Image Time Series Analysis: A Review, arXiv preprint, 10.48550/arXiv.2404.03936, 2024.

Min, M., Li, J., Wang, F., Liu, Z., and Menzel, W. P.: Retrieval of cloud top properties from advanced geostationary satellite imager measurements based on machine learning algorithms, Remote Sensing of Environment, 239, 111616, 10.1016/j.rse.2019.111616 2020.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

Min, M., Chen, B., Xu, N., He, X., Wei, X., and Wang, M.: Nonnegligible diurnal and long-term variation characteristics of the calibration biases in Fengyun-4A/AGRI infrared channels based on the oceanic drifter data, IEEE Transactions on Geoscience and Remote Sensing, 60, 1-15, 10.1109/TGRS.2022.3160450, 2022.

Min, M., Deng, J., Liu, C., Lu, N., Hu, X., Chen, L., Guo, J., Zhang, P., Lu, Q., and Wang, L.: An investigation of the implications of lunar illumination spectral changes for Day/Night Band based cloud property retrieval due to lunar phase transition, Journal of Geophysical Research: Atmospheres, 122, 9233-9244, 10.1002/2017JD027117, 2017a.

Min, M., Bai, C., Guo, J., Sun, F., Liu, C., Wang, F., Xu, H., Tang, S., Li, B., Di, D., Dong, L., and Li, J.: Estimating summertime precipitation from Himawari-8 and global forecast system based on machine learning, IEEE Transactions on Geoscience and Remote Sensing, 57, 2557-2570, 10.1109/TGRS.2018.2874950, 2019.

Min, M., Wu, C., Li, C., Liu, H., Xu, N., Wu, X., Chen, L., Wang, F., Sun, F., Qin, D., Wang, X., Li, B., Zheng, Z., Cao, G., and Dong, L.: Developing the science product algorithm testbed for Chinese next-generation geostationary meteorological satellites: FengYun-4 series, Journal of Meteorological Research, 31, 708-719, 10.1007/s13351-017-6161-z, 2017b.

Nowakowski, A., Mrziglod, J., Spiller, D., Bonifacio, R., Ferrari, I., Mathieu, P. P., Garcia-Herranz, M., and Kim, D.-H.: Crop type mapping by using transfer learning, International Journal of Applied Earth Observation and Geoinformation, 98, 102313, https://doi.org/10.1016/j.jag.2021.102313, 2021.

Pan, S. J. and Yang, Q.: A Survey on Transfer Learning, IEEE Transactions on Knowledge and Data Engineering, 22, 1345-1359, 10.1109/TKDE.2009.191, 2010.

Pearlman, A., Pogorzala, D., and Cao, C.: GOES-R Advanced Baseline Imager: spectral response functions and radiometric biases with the NPP Visible Infrared Imaging Radiometer Suite evaluated for desert calibration sites, Applied Optics, 52, 7660-7668, 10.1364/AO.52.007660, 2013.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

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.

Schmit, T. J., Griffith, P., Gunshor, M. M., Daniels, J. M., Goodman, S. J., and Lebair, W. J.: A closer look at the ABI on the GOES-R Series Bulletin of the American Meteorological Society, 98, 681-698 10.1175/BAMS-D-15-00230.1, 2017.

Skofronick-Jackson, G., Petersen, W. A., Berg, W., Kidd, C., Stocker, E., Kirschbaum, D. B., and Huffman, G. J.: The Global Precipitation Measurement (GPM) Mission for Science and Society, Bulletin of the American Meteorological Society, 98, 1679-1695, 2017.

Teixeira, J., Wilson, R. C., and Thrastarson, H. T.: Direct observational evidence from space of the effect of CO2 increase on longwave spectral radiances: the unique role of high-spectral-resolution measurements, Atmos. Chem. Phys., 24, 6375-6383, 10.5194/acp-24-6375-2024, 2024.

Wang, C. and Li, X.: Developing a Data-Driven Transfer Learning Model to Locate Tropical Cyclone Centers on Satellite Infrared Imagery, Journal of Atmospheric and Oceanic Technology, 40, 1605-1618, https://doi.org/10.1175/JTECH-D-23-0026.1, 2023.

Wang, C., Tang, G., Xiong, W., Ma, Z., and Zhu, S.: Infrared precipitation estimation using convolutional neural network for FengYun satellites, Journal of Hydrology, 603, 127113, https://doi.org/10.1016/j.jhydrol.2021.127113, 2021.

Wang, M., Min, M., Li, J., Lin, H., Liang, Y., Chen, B., Yao, Z., Xu, N., and Zhang, M.: Technical note: Applicability of physics-based and machine-learning-based algorithms of geostationary satellite in retrieving the diurnal cycle of cloud base height, Atmospheric Chemistry and Physics, 24, 14239–14256, 10.5194/acp-24-14239-2024, 2024.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

Wang, W., Brönnimann, S., Zhou, J., Li, S., and Wang, Z.: Near-surface air temperature estimation for areas with sparse observations based on transfer learning, ISPRS Journal of Photogrammetry and Remote Sensing, 220, 712-727, https://doi.org/10.1016/j.isprsjprs.2025.01.021, 2025.

Xia, P., Meng*, D., Min*, M., Li, J., Liu, Y.-A., Yang, X., and Xu, N.: Impacts of Fengyun-4B satellite rapid imaging-derived atmospheric motion vectors on typhoon forecasts, Journal of Geophysical Research: Atmospheres, 130, e2025JD044269, 10.1029/2025JD044269, 2025.

Xia, P., Zhang, L., Min, M., Li, J., Wang, Y., Yu, Y., and Jia, S.: Accurate nowcasting of cloud cover at solar photovoltaic plants using geostationary satellite images, Nature Communications, 15, 1-10, 10.1038/s41467-023-44666-1, 2024.

Xiong, X., Angal, A., Twedt, K. A., Chen, H., Link, D., Geng, X., Aldoretta, E., and Mu, Q.: MODIS Reflective Solar Bands On-Orbit Calibration and Performance, IEEE Transactions on Geoscience and Remote Sensing, 57, 6355-6371, 10.1109/TGRS.2019.2905792, 2019.

Yang, J., Zhang, Z., Wei, C., Lu, F., and Guo, Q.: Introducing the new generation of Chinese geostationary weather satellites, FengYun-4, Bulletin of the American Meteorological Society, 98, 1637-1658, 10.1175/BAMS-D-16-0065.1, 2017.

Yu, Q., Min, M., Di, D., Zhao, Y., Li, G., Li, J., Wang, L., Hu, X., and Zhang, P.: Spectral-fidelity-preserving recalibration of reflective solar bands for MERSI-II onboard Fengyun-3D satellite, IEEE Transactions on Geoscience and Remote Sensing, 64, 4104116, 10.1109/TGRS.2026.3671371, 2026.

Zermatten, V., Castillo-Navarro, J., Marcos, D., and Tuia, D.: Learning transferable land cover semantics for open vocabulary interactions with remote sensing images, ISPRS Journal of Photogrammetry and Remote Sensing, 220, 621-636, https://doi.org/10.1016/j.isprsjprs.2025.01.006, 2025.

Zhang, C.-J., Wang, Y., Ma, L.-M., and Lu, X.-Q.: TCIC-SC-RNet: A Hybrid Model for TC Intensity Estimation Based on Transfer Learning, Weather and Forecasting, 39, 1469-1485, https://doi.org/10.1175/WAF-D-23-0080.1, 2024.

Zhou, J., Yang, Q., Liu, L., Kang, Y., Jia, X., Chen, M., Ghosh, R., Xu, S., Jiang, C., Guan, K., Kumar, V., and Jin, Z.: A deep transfer learning framework for mapping high spatiotemporal resolution LAI, ISPRS Journal of Photogrammetry and Remote Sensing, 206, 30-48, https://doi.org/10.1016/j.isprsjprs.2023.10.017, 2023.

Zhou, R., Pan, X., Xiaohu, Z., Na, X., and Min, M.: Research progress and prospects of atmospheric motion vector based on meteorological satelliteimages, Reviews of Geophysics and Planetary Physics (In Chinese), 55, 184-194, 10.19975/j.dqyxx.2022-077, 2024a.

Zhou, Y., Min, M., Li, J., Cao, Z., and Gao, L.: Enhanced typhoon center localization using geostationary satellite imagery, Geophysical Research Letters, 51, e2024GL112252, 10.1029/2024GL112252, 2024b.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

Table 1. Input variables and the corresponding values for RTM simulations.

Parameter	Classification	Values
Aerosol	Ocean aerosol	$\tau = 0.17$
	Desert aerosol	$\tau = 0.78$
Cloud	Clear-sky (no cloud)	$\tau = 0.0$
	Cumulus cloud	$\tau = 216$
	Stratus cloud	$\tau = 38$
	Altostratus cloud	$\tau = 76$
	Nimbostratus cloud	$\tau = 46$
	Cirrus cloud	$\tau = 1.5$
Surface	Snow	ρ= 0.003–0.979
	Forest	P = 0.08–0.83
	Farm	ρ = 0.09–0.90
	Desert	ρ= 0.10–0.71
	Urban	ρ= 0.019–0.137
	Ocean	ρ= 0.01–0.05
	Grass	ρ= 0.00606–0.0325
	Tundra	ρ= 0.032–0.273
	Wind speed	ws= 6.9 m/s
View Geometry	Solar zenith angle	*Ang* = 0°, 30°, 45°, 55°, 60°
	Solar azimuth angle	*Ang* = 0°
	Satellite view zenith angle	*Ang* = 0°, 30°, 45°, 55°, 60°
	Satellite view azimuth angle	*Ang* = 0°, 45°, 90°, 135°, 180°

Note: τ = Optical depth at 550 nm. ρ = Reflectivity. *Ang* = Angle (degree). ws= wind speed.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

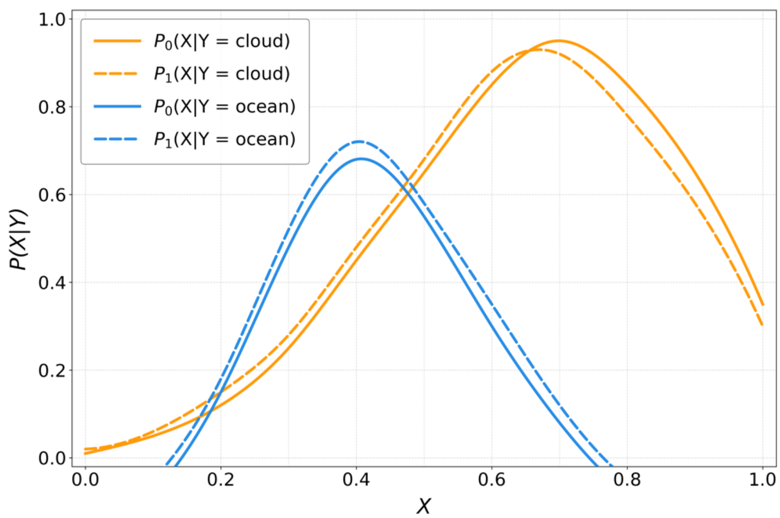

Figure 1. Prior shift in the conditional distributions $P(X|Y)$ of satellite-observed reflectance (Ref) or brightness temperature (BT) for assumed cloud (orange) and ocean (blue) targets. This shift (with subscripts 0 (solid line) and 1 (dashed line)) can occur when data are collected either by different satellite sensors or by the same sensor but over different time periods (Example inspired by (Ma et al., 2024)).

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

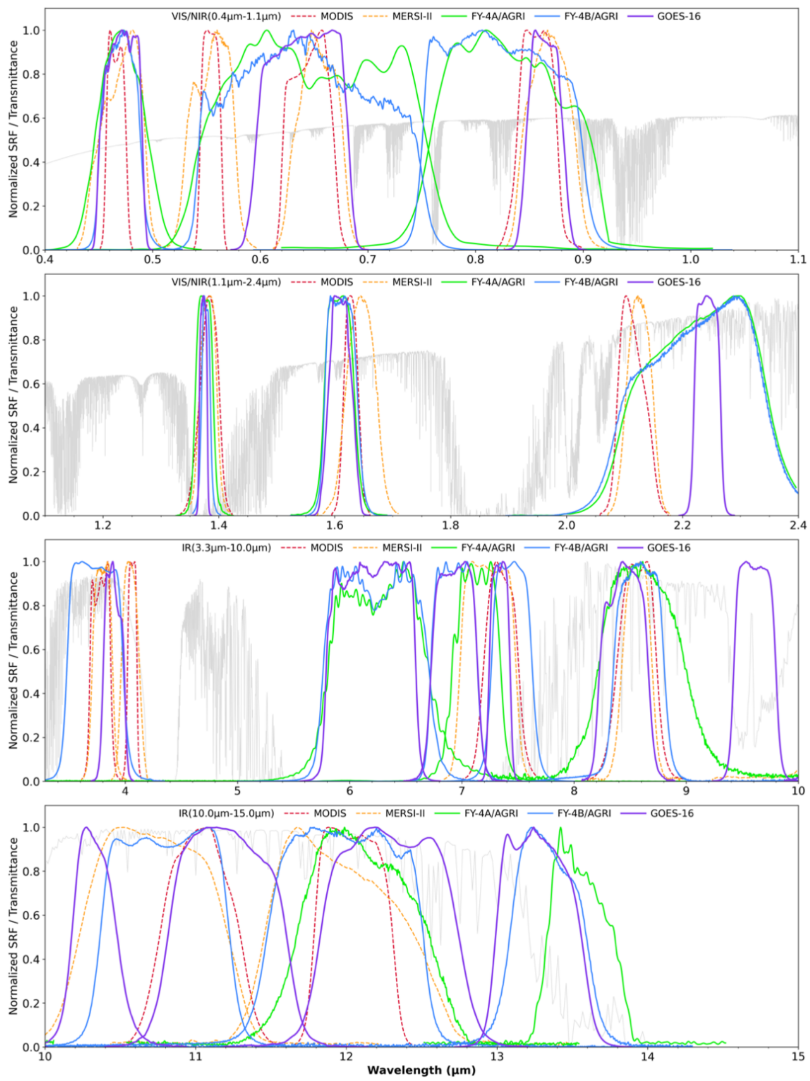

Figure 2. Normalized Spectral response functions from the visible to infrared channels across China FY-4A/B GEO imagers (AGRI, Advanced Geosynchronous Radiation Imager, green/blue solid line) , the U.S. GOES-16 GEO imager (Advanced Baseline Imager, ABI, purple solid line), China FY-3D polar-orbiting imager (MERSI-II, Medium Resolution Spectral Imager II, orange dashed line), and the U.S. Aqua/MODIS polar-orbiting sensor (cardinal red dashed line). The light gray solid curve in each subfigure denotes the simulated transmittance using the 1976 U.S. Standard Atmosphere profile.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

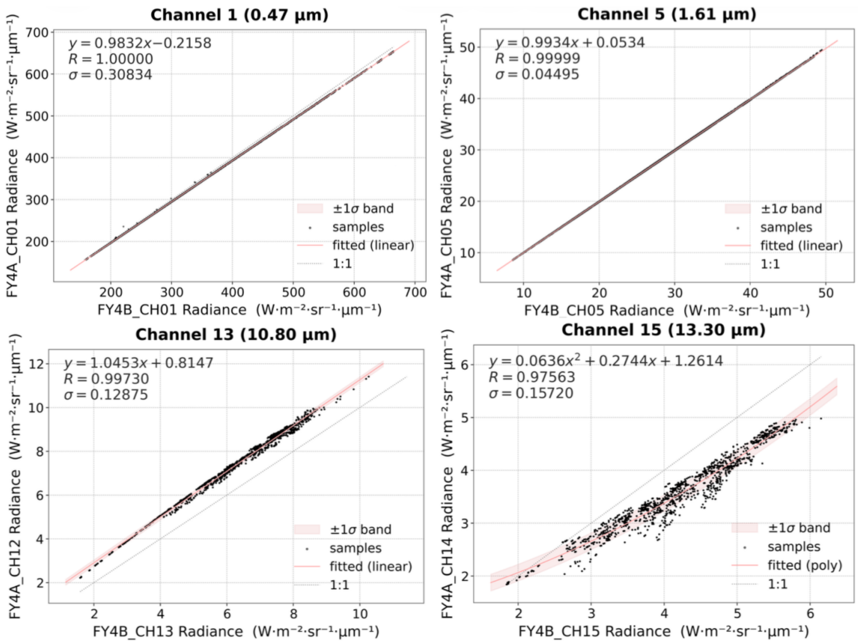

Figure 3. Application of the new physics-constrained transfer learning framework from FY-4B to FY-4A AGRI radiance observations. From left to right, the panels show Channels 1 (0.47μm), 5 (1.61μm), 13 (10.80μm), and 15 (13.30μm) (following the FY-4B channel designations). They demonstrate the framework's linear fitting performance in the visible, near-infrared, and thermal infrared bands, as well as an example of higher-order polynomial fitting for FY-4B/AGRI Channel 15 (13.30μm). R and σ denote the correlation coefficient and fitting residuals, respectively.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

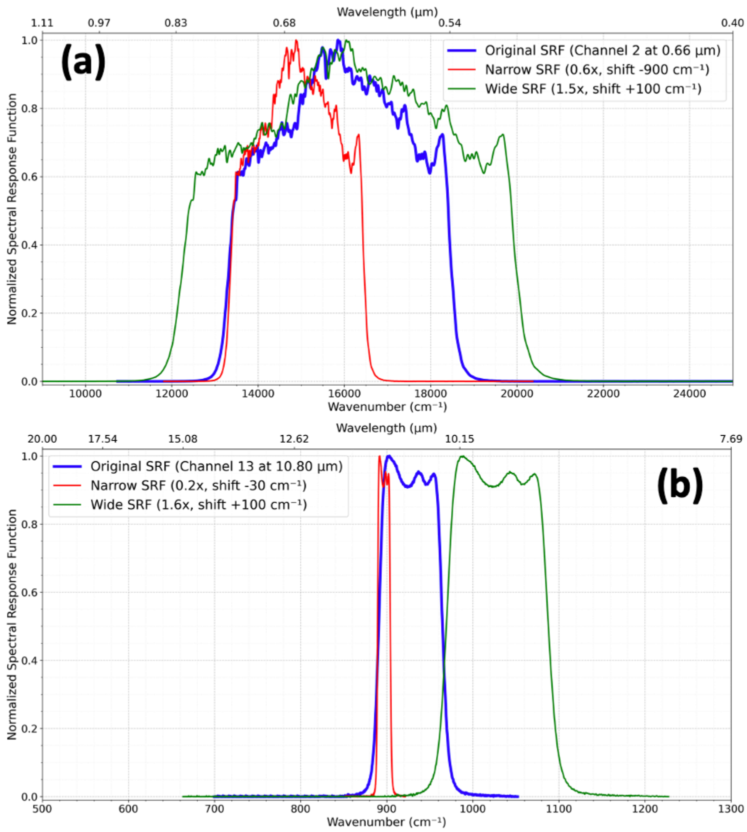

Figure 4. Original SRFs (blue solid lines) and two simulated SRFs, including a narrower SRF (red solid lines) and a wider SRF (green solid lines), for FY-4B/AGRI: (a) Channel 2 at 0.66 μm and (b) Channel 13 at 10.80 μm. Here, 0.2/0.6/1.5/1.6× indicates 0.2/0.6/1.5/1.6 times (or scaling factor) the width of the original SRF, shift −30/−900 cm^{-1} denotes a leftward shift of 30/900 cm^{-1}, and shift +100 cm^{-1} denotes a rightward shift of 100 cm^{-1}.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

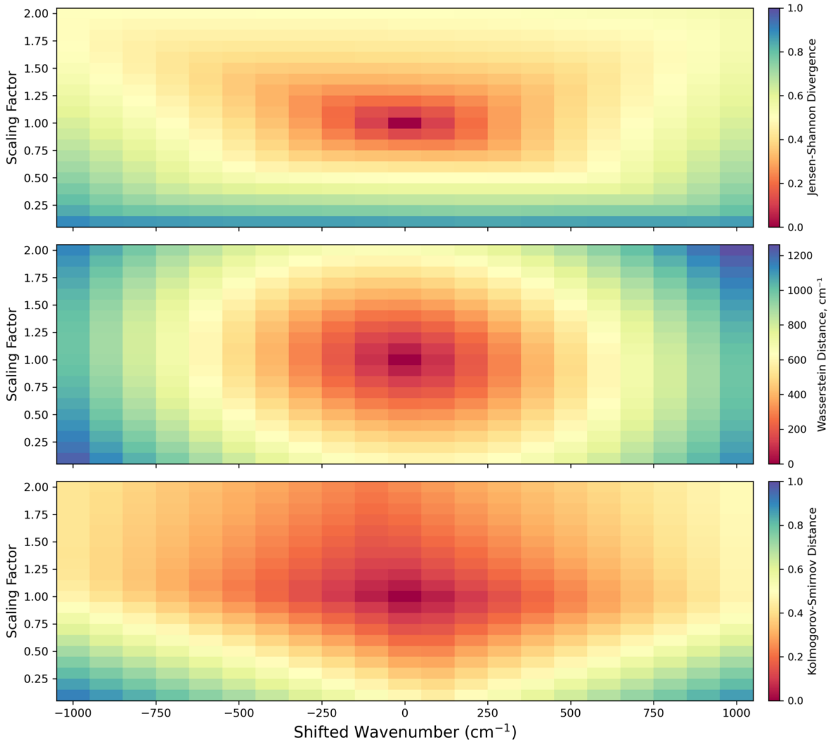

Figure 5. Distributions of Jensen–Shannon divergence (top panel), Wasserstein distance (middle panel), and Kolmogorov–Smirnov distance (bottom panel) as functions of scaling factor (from 0.1 to 2.0) and shifted wavenumber (from -1000 cm^{-1} to +1000 cm^{-1}) for FY-4B/AGRI Channel 2 at 0.66 μm.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

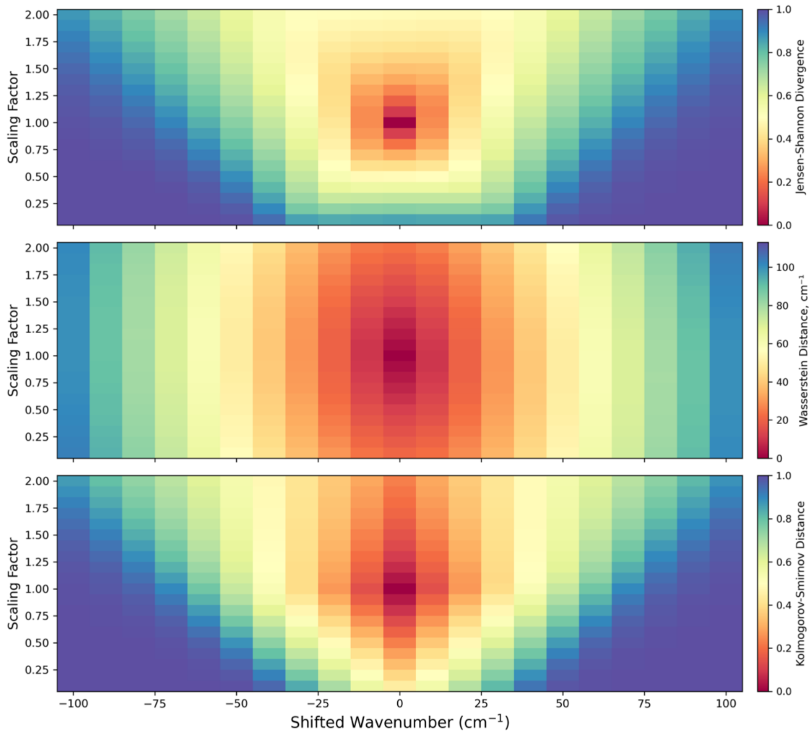

Figure 6. Same as Figure 5, but for the FY-4B/AGRI Channel 13 at 10.80 µm with shifted wavenumber (from -100 cm⁻¹ to +100 cm⁻¹).

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

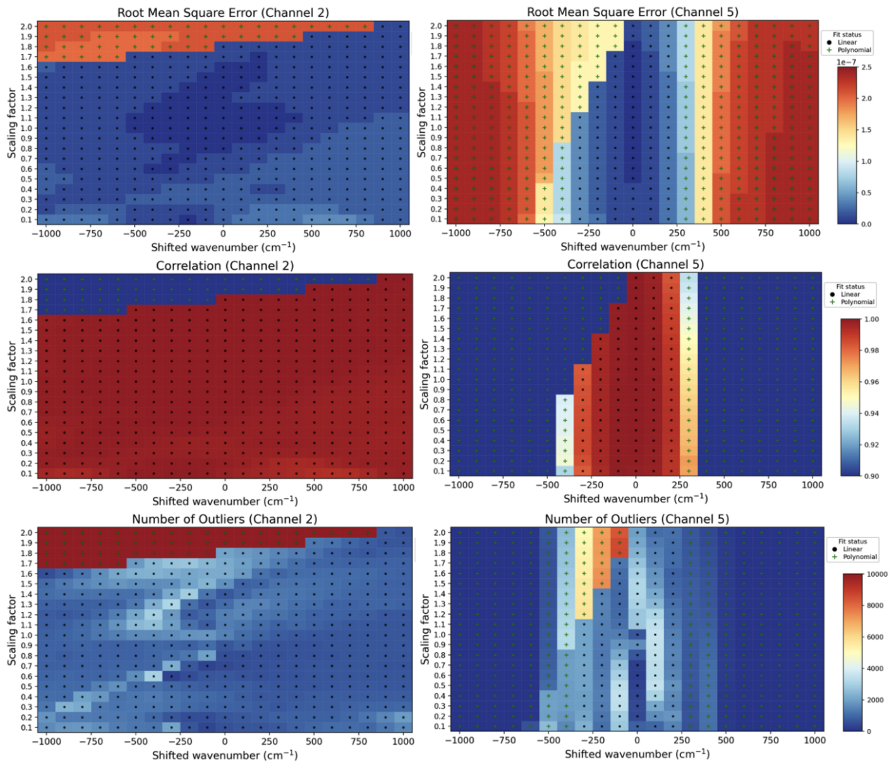

Figure 7. Distributions of RMSE (top panel), R (middle panel), and $N_{outlier}$ (bottom panel) as functions of scaling factor (from 0.1 to 2.0) and shifted wavenumber (from -1000 cm^{-1} to +1000 cm^{-1}) for FY-4B/AGRI Channel 2 at 0.66 μm (left panel) and Channel 5 at 1.61 μm (right panel). Black solid dots denote linear fits, and green plus signs denote polynomial fits, respectively, for each regression-based transfer learning model.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

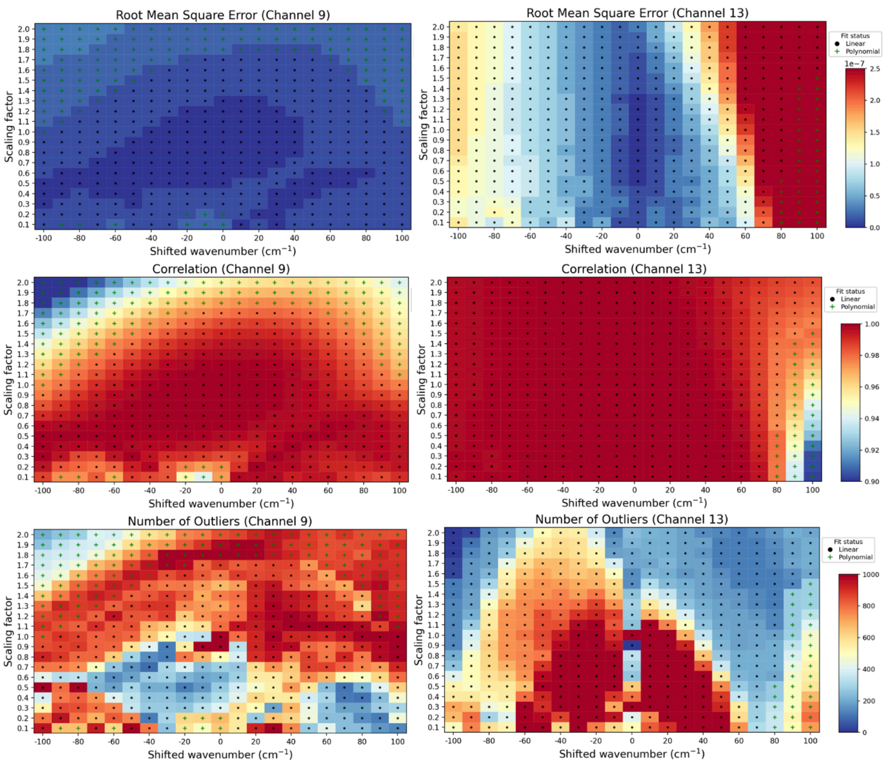

Figure 8. Same as Figure 7, but for the FY-4B/AGRI Channel 9 at 6.25 μm (left panel) and Channel 13 at 10.80 μm (right panel).

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

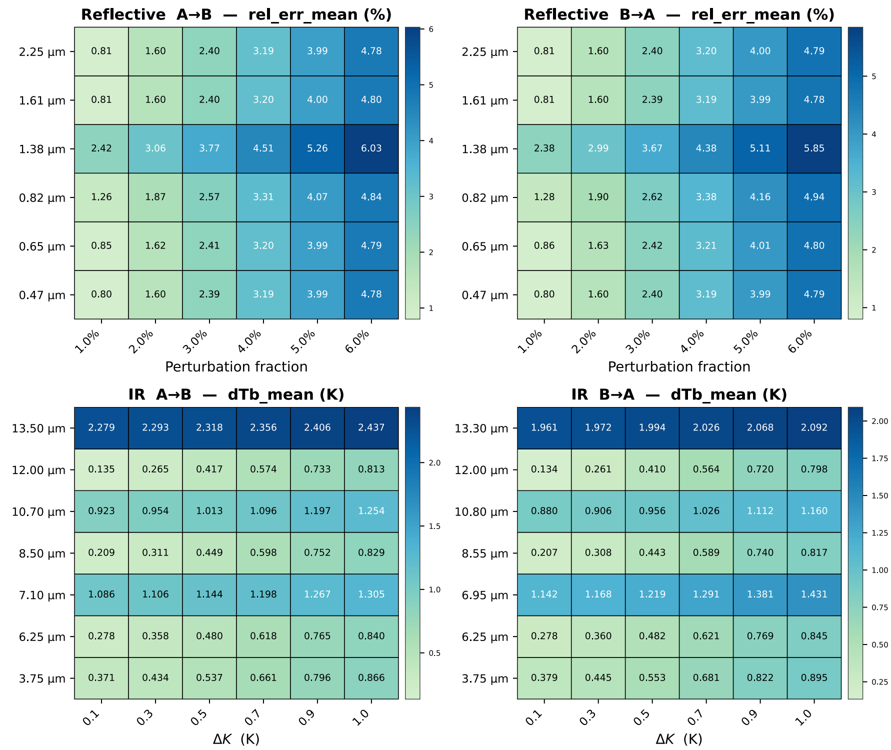

Figure 9. Heatmaps derived from the error-sensitivity analysis based on the physics-constrained TL models show error perturbations ranging from 1% to 6% in reflectance for the solar reflective channels and from 0.1 to 1 K within an interval of 0.2 K in brightness temperature for the thermal infrared emission channels. The A→B case (left panel) denotes the transformation in which FY-4A/AGRI channels are treated as the target channels and FY-4B/AGRI channels as the reference channels, whereas the B→A case (right panel) represents the exact reverse process.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

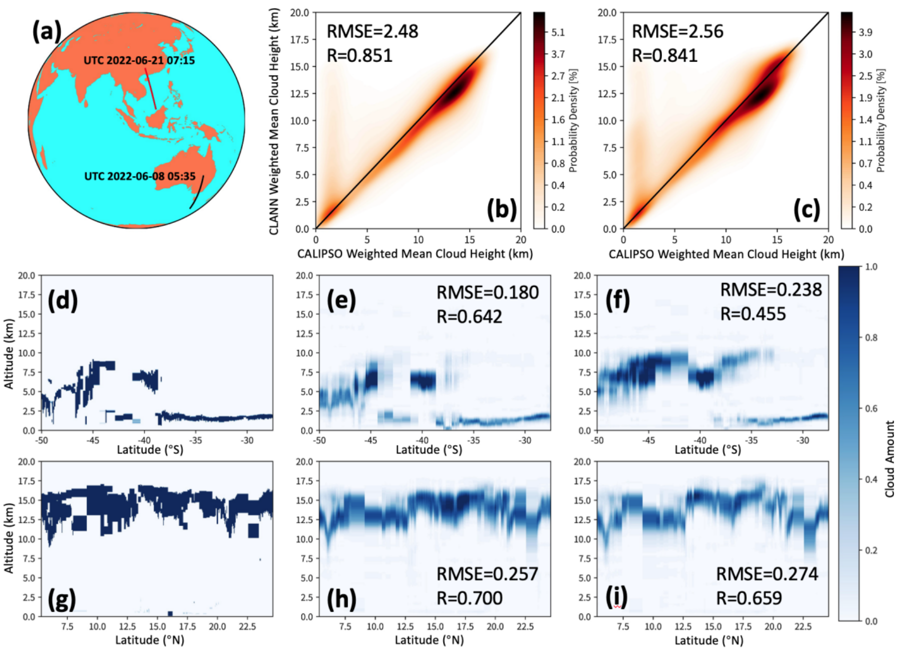

Figure 10. Map projection of CALIPSO tracks for two cases is shown in (a). Panels (b–c) present weighted mean cloud-top height comparisons between CLANN and CALIPSO using FY-4B/AGRI data, with and without TL model, respectively. Panels (d–i) show cross-sectional cloud amount profiles, including CALIPSO observations (d, g), CLANN with TL model (e, h), and without TL model (f, i). Cases 1 and 2 at 05:35 UTC on June 8, 2022 (black solid line) and 07:15 UTC on June 21, 2022 (red solid line), respectively. RMSE and correlation coefficients (R) are indicated in each panel.

https://doi.org/10.5194/egusphere-2026-3959
Preprint. Discussion started: 14 July 2026

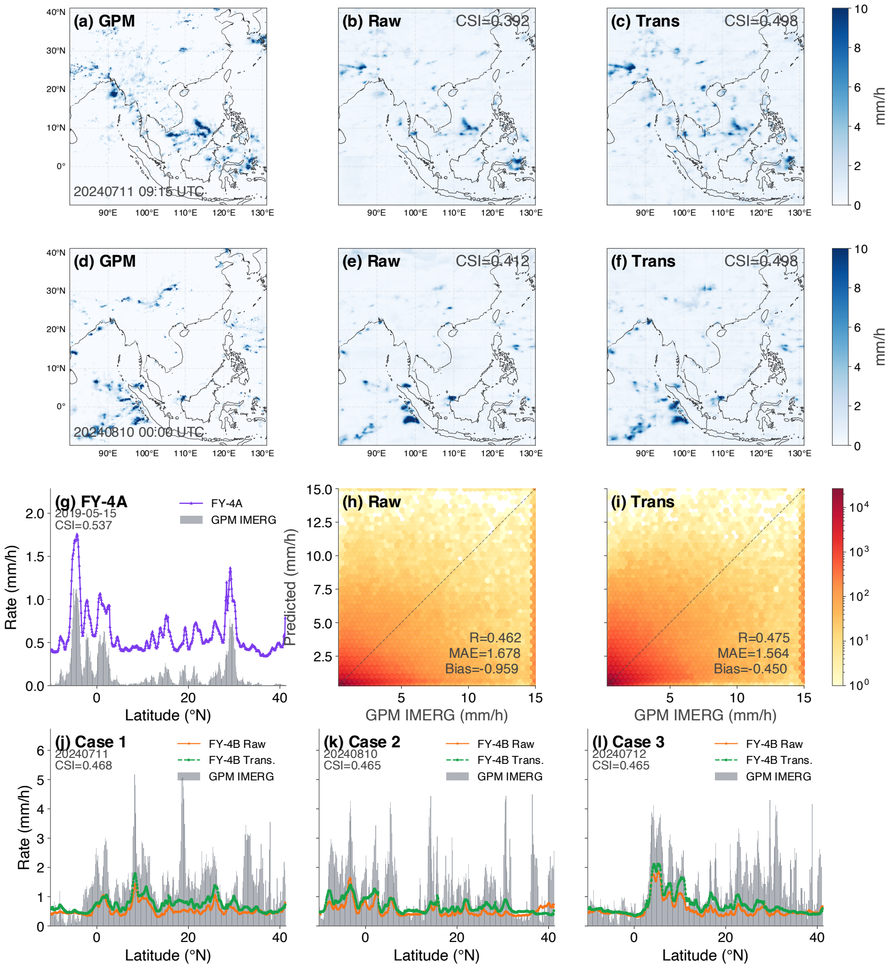

Figure 11. Evaluation of the physics-constrained transfer learning framework for QPE using a Multi-Task UNet model trained with 2019 FY-4A/AGRI observations. The training samples were mainly drawn from May–August and supplemented with five randomly selected days from each remaining month. (a–c) Spatial distributions of rain rate (mm h⁻¹) from GPM IMERG, the raw FY-4B output, and the transfer-learning-corrected FY-4B output for Case 1 on 11 July 2024 at 09:15 UTC. CSI values are annotated for the raw and corrected outputs. (d–f) Same as (a–c), but for Case 2 on 10 August 2024 at 00:00 UTC. (g–i) Scatter-density comparisons between model-estimated and GPM IMERG rain rates for the FY-4A in-domain baseline on 15 May 2019, the raw FY-4B output, and the TL-corrected FY-4B output, respectively. R, MAE, and bias are annotated in each panel. (j–l) Latitudinal rain-rate profiles for Cases 1–3, corresponding to 11 July 2024, 10 August 2024, and 12 July 2024, respectively, comparing the raw FY-4B (orange solid line) output, TL-corrected FY-4B (green solid line) output , and GPM IMERG reference (gray). CSI values for the corrected output are annotated in each panel. All independent FY-4B cases were selected from the extrapolation period spanning June 2023 to August 2024, which is entirely outside the FY-4A training window.