Applied Energy 423 (2026) 128373

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apen

Real-time estimation of cloud properties and surface downwelling longwave radiation from satellite imagery: A two-stage surrogate modeling approach ☆

Yuying Xie, Nan Deng, Tao Jing, Mengying Li *

Department of Mechanical Engineering & Research Centre for Data Science and Artificial Intelligence, The Hong Kong Polytechnic University, Hung Hom, Hong Kong Special Administrative Region of China

HIGHLIGHTS

- Propose a novel two-stage surrogate modeling framework that outperforms direct end-to-end modeling.
- Facilitate real-time, regional-scale retrievals of cloud optical properties and surface downwelling longwave radiation.
- Accelerate complex radiative modeling while preserving accuracy through advanced deep learning techniques.
- Enhance accessibility of radiative modeling for non-experts and supporting broader operational applications.

ARTICLE INFO

Keywords:
Sky radiative cooling
Surface downwelling longwave radiation
Cloud optical properties
Radiative transfer model
Remote sensing
Deep learning

ABSTRACT

Passive sky radiative cooling systems, which utilize the universe as a natural heat sink, offer promising low-carbon solutions for urban cooling. Surface downwelling longwave radiation (SDLR), originating from the atmosphere, significantly influences the cooling potential of such systems. However, spatiotemporal continuous SDLR data remain scarce, primarily due to the complex dynamics of clouds and their non-linear interactions with radiation. To address this challenge, we propose a novel two-stage surrogate modeling framework that combines high-resolution (5-min, 2-km) geostationary satellite imagery with deep learning techniques. When validated against one year (2019) of SDLR measurements from the Surface Radiation Budget Network (SURFRAD) across diverse climatic regions in the contiguous United States, the proposed model significantly outperforms direct end-to-end satellite-to-SDLR mapping, with MBE values ranging from 0.10 W/m^2 to 17.62 W/m^2 and RMSE ranging from 20.78 W/m^2 to 30.54 W/m^2. Benchmark comparisons underscore its reliability over empirical, physical, reanalysis, and satellite-based methods, which demonstrate competitive performance with notable bias reduction. Furthermore, this model enables robust and efficient retrieval of spatiotemporal cloud properties and SDLR, while exhibiting strong transferability and minimal reliance on ground-based data, thereby supporting real-time, regional-scale operational applications.

1. Introduction

Rapid urbanization necessitates innovative technologies to regulate the urban thermal environment, particularly to mitigate urban heat island effects. Over the past decade, passive sky radiative cooling systems have gained significant attention due to their ability to dissipate heat in the form of thermal radiation through the atmospheric window (8 to 13 μm) to the cold universe. These systems have emerged as vital low-carbon technologies for urban cooling [1,2]. As a technology whose performance heavily depends on atmospheric radiative transfer processes, especially surface downwelling longwave radiation (SDLR), real-time SDLR estimation is crucial for the design and performance evaluation of these systems [3]. However, high-quality, spatiotemporal continuous SDLR estimation methods and corresponding datasets remain scarce or lack-accuracy, primarily due to the high cost of sensors and the difficulty of characterizing cloud fields–the major modulators of atmospheric radiative transfer processes.

Previous studies have demonstrated remarkable accuracy in estimating clear-sky SDLR, primarily using physical models [4] and

☆ The short version of the paper was presented at CUE2025, Kitakyushu, Japan, July 18–22, 2025. This paper is a substantial extension of the short version of the conference paper.

* Corresponding author.
Email address: mengying.li@polyu.edu.hk (M. Li).

https://doi.org/10.1016/j.apenergy.2026.128373
Received 30 January 2026; Received in revised form 17 June 2026; Accepted 4 July 2026
Available online 14 July 2026

Nomenclature

Abbreviations

ABI Advanced Baseline Imager
AFGL Air Force Geophysical Laboratory
BON Bondville
CF Cloud Fraction
CMF Cloud Modification Factor
COD Cloud Optical Depth
DL Deep Learning
DNN Deep Neural Network
DRA Desert Rock
FPK Fort Peck
GHI Global Horizontal Irradiance
GOES Geostationary Operational Environmental Satellites
GWN Goodwin Creek
IR Infrared
KNN K-Nearest Neighbor
LBL Line-by-Line
LBLRTM Line-by-Line Radiative Transfer Model
MBE Mean Bias Error
MLP Multilayer Perceptron
PSU Pennsylvania State University
RH Relative Humidity
RMSE Root Mean Squared Error
RTM Radiative Transfer Model
SDLR Surface Downwelling Longwave Radiation
SURFRAD Surface Radiation Budget Network
SXF Sioux Falls
TBL Table Mountain

Notations

$\hat{F}_b^{\uparrow}$ Measured Upwelling Flux
$\hat{F}_0^{\downarrow}$ Surface Downwelling Flux
$\hat{I}_{v,b}^{\uparrow}$ Satellite Spectral Intensity
ϕ Relative Humidity
τ Cloud Optical Depth
θ_z Solar Zenith Angle
$F_b^{\uparrow}$ Modeled Upwelling Flux
$F_0^{\downarrow}$ Modeled Downwelling Flux
$I_{v,b}^{\uparrow}$ Modeled Spectral Intensity
$k_b^{\uparrow}$ Clear-sky Index
R_b Spectral Response Function
T_a Ambient Temperature
z_t Cloud-top Height

Subscripts

v Wavenumber
b Band Number

Superscripts

$*$ Normalized Values
$\downarrow$ Downwelling Flux
$\uparrow$ Upwelling Flux

parameterized algorithms [5] that correlate SDLR (or sky emissivity as an alternative form) with atmospheric water vapor content. However, under cloudy-sky conditions, SDLR estimation remains challenging due to the inherent variability and dynamic nature of cloud properties [6]. Accurate, real-time characterization of cloud fields is therefore essential for improving SDLR estimates. While some efforts have been made to quantify the impact of clouds on SDLR using parameterized schemes based on cloud characteristics [7–9], these approaches are constrained by simplified representations of cloud optical properties.

To enable real-time atmospheric characterization, geostationary weather satellites, such as the Geostationary Operational Environmental Satellite (GOES) [10] and Himawari [11], have emerged as valuable sources of open-access data with high spatiotemporal resolution for top-of-atmosphere upwelling radiance (Level 1 product). However, significant gaps remain in accurately retrieving cloud properties and estimating downwelling radiative fluxes, primarily due to the complex dynamics of clouds and their non-linear interactions with radiation. Currently, retrieval methods predominantly rely on well-established radiative transfer models (RTMs) to infer cloud properties and radiative fluxes, as these models can simulate the intricate interactions between radiation and atmospheric constituents [12]. For example, our previous work [13] integrated a high-fidelity line-by-line radiative transfer model (LBLRTM) [14] with GOES observations in the longwave spectrum ($\lambda > 4$ μm) to estimate cloud optical properties in real time using a constrained optimization approach. Exclusive reliance on the longwave spectrum offers three distinct advantages over the use of solar shortwave measurements ($\lambda < 4$ μm). First, longwave radiance is more sensitive to cloud vertical structure, including cloud-top height and thickness. Second, the angular distribution of longwave radiation in the atmosphere is nearly isotropic, making it largely insensitive to solar geometry and associated angular effects [12,15]. Finally, longwave radiance is available during both daytime and nighttime, enabling continuous 24/7 retrieval of cloud properties. However, despite these advantages, RTM-based approaches remain challenging to implement and computationally expensive, which limits their practicality for real-time or near-real-time applications [16].

The need for accurate, efficient, and easy to implement radiative transfer calculations has motivated the exploration of alternative approaches [17]. In recent years, deep learning (DL) [18] has demonstrated significant potential as surrogate modeling approach to accelerate atmospheric retrievals and radiative transfer simulations. These methods reduce computation time from hours to minutes or less while maintaining much of the fidelity of physics-based methods [17,19,20]. For example, Ukkonen et al. [21] employed recurrent neural networks to accelerate RTMs, while feed-forward neural networks were devised to replace shortwave radiation schemes. Li et al. [22] proposed a machine learning algorithm that integrates RTM simulations to facilitate multilayer cloud property retrievals from Himawari-8. Likewise, Lagerquist et al. [23] developed DL models to emulate RTMs, targeting key radiative transfer variables, such as surface downwelling flux and top-of-atmosphere upwelling flux. Despite these advancements, surrogate models generally suffer from reduced accuracy, as computational speedup is achieved through significant approximations [24]. This accuracy reduction is particularly pronounced in end-to-end DL models, which often fail to fully capture the physical complexities of atmospheric radiative transfer. These challenges lead to an increased dependency on large, high-quality datasets, as additional knowledge-guided inputs are required to enhance performance [25,26]. A key research gap remains in establishing reliable surrogate frameworks that balance computational efficiency and accuracy while minimizing data dependence for operational SDLR estimation.

To address these limitations, this work introduces a novel two-stage surrogate modeling framework designed to efficiently retrieve high-resolution cloud optical properties and estimate SDLR from satellite data in the longwave spectrum. The proposed framework leverages DL techniques to emulate RTM processes in two stages. In Stage I, the cloud optical properties are retrieved from multi-band upwelling radiance data. In Stage II, the cloud optical properties, combined with ancillary meteorological data (e.g., temperature and humidity), are mapped to surface downwelling longwave fluxes. The main contributions of this work are:

- An efficient surrogate modeling framework is developed to replace complex radiative models using advanced DL techniques, thereby addressing the computational bottlenecks of physical RTMs for real-time, regional-scale applications.
- A novel two-stage modeling framework is developed in which cloud optical properties serve as effective intermediate physical descriptors. This two-stage architecture demonstrates a significant improvement in accuracy compared to direct end-to-end satellite-to-SDLR mapping. Furthermore, the incorporation of physical descriptors has substantially enhanced the interpretability of the model.
- The proposed surrogate model enables both cloud property retrieval and SDLR estimations from remote sensing data with high spatiotemporal resolution. This ensures consistent performance evaluations of outdoor energy applications, such as sky radiative cooling devices, with minimal reliance on ground-sensing data. Furthermore, this framework enhances the accessibility of radiative transfer modeling for non-experts and supports a broader range of operational applications.

The rest of this paper is structured as follows: Section 2 describes the data acquisition and pre-processing procedures for model development and evaluation; Section 3 presents the methodology of the surrogate framework. The results are evaluated and discussed both quantitatively and qualitatively in Section 4. Finally, Section 5 summarizes the key findings.

2. Data

In this work, two publicly available datasets from 2019 are used: (1) remote sensing data from the latest generation of the Geostationary Operational Environmental Satellite (GOES) series (specifically, GOES-16 launched in 2016), and (2) ground-level meteorological measurements (air temperature, relative humidity, SDLR) from the Surface Radiation Budget Network (SURFRAD) sites.

2.1. Remote sensing data

The GOES series consists of four satellites equipped with nadir-pointing, solar-pointing, and in situ instruments [27]. This work uses multi-band data from the Advanced Baseline Imager (ABI), covering the Contiguous United States. To enable SDLR estimation during both daytime and nighttime, only data from the longwave bands are considered. These bands offer data with a temporal resolution of 5-min, and a spatial resolution ranging from 0.5 to 2.0 km, as detailed in Table 1. Spectral radiance data from the pixel directly above the ground stations are considered.

Table 1
Summary of data from longwave spectral bands of GOES-16. IR denotes infrared.

Band No.	Type	Name	Center Wavelength (Spectral Range), µm
7	IR	Shortwave Window	3.90 (3.80 – 4.00)
8	IR	Upper-level Water Vapor	6.15 (5.77 – 6.60)
9	IR	Mid-level Water Vapor	7.00 (6.75 – 7.15)
10	IR	Lower-level Water Vapor	7.40 (7.24 – 7.44)
11	IR	Cloud-top Phase	8.50 (8.3 – 8.70)
12	IR	Ozone	9.70 (9.4 – 9.80)
13	IR	Clean IR Longwave Window	10.30 (10.10 – 10.60)
14	IR	IR Longwave Window	11.20 (10.80 – 11.60)
15	IR	Dirty Longwave Window	12.30 (11.80 – 12.80)
16	IR	CO_2 Longwave Infrared	13.30 (13.00 – 13.60)

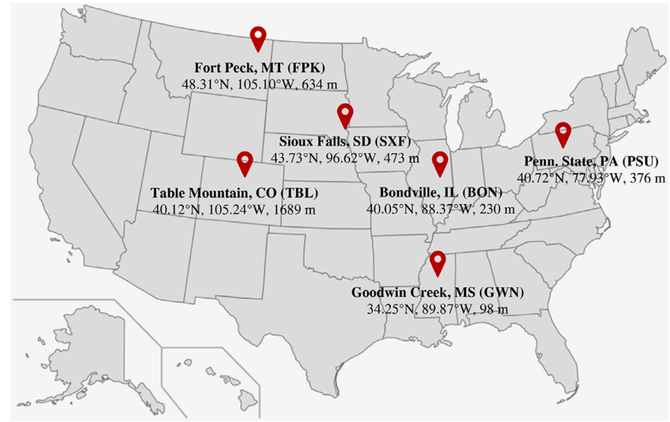

Fig. 1. Geographical distribution of the SURFRAD stations with annotated geographic coordinates (latitude, longitude) and elevation.

2.2. Ground-based measurements

Ground measurements are collected from six SURFRAD stations, namely Bondville, Illinois (BON); Fort Peck, Montana (FPK); Goodwin Creek, Mississippi (GWN); Pennsylvania State University, Pennsylvania (PSU); Sioux Falls, South Dakota (SXF); and Table Mountain, Boulder, Colorado (TBL), as shown in Fig. 1. Measurements from Desert Rock, Nevada (DRA) are excluded from this study due to previously identified issues related to erroneous SDLR values [28]. To ensure consistency with satellite data, all ground-based measurements are converted to standard international units and resampled into 5-min backward-averaged values. The data from each station are classified into daytime and nighttime based on the solar zenith angle θ_z, where daytime is defined as $\theta_z \leq 85°$ and nighttime is defined otherwise. Following the sky classification methodology proposed by [29], daytime data are further divided into clear and cloudy conditions. Overall, the data from each station are classified into three categories for further analysis: (1) daytime cloudy, (2) daytime all-sky (clear and cloudy), and (3) nighttime all-sky.

3. Methodology

This section presents the detailed working principles of the proposed two-stage surrogate modeling framework for deriving cloud properties and SDLR from satellite data. Two deep learning (DL) models are trained using data from a physical radiative transfer model (RTM), thereby minimizing reliance on ground-sensing data and improving model generalizability. In the following, the RTM used in this framework is introduced in Section 3.1. The DL model architecture is described in Section 3.2. The overall surrogate modeling framework is detailed in Section 3.3, and the performance assessment metrics are presented in Section 3.4.

3.1. Line-by-line radiative transfer model

Line-by-line RTMs (LBLRTM) are widely regarded as the most accurate tools for capturing the physical processes of atmospheric radiative transfer and have served as the foundation for satellite retrieval algorithms [16]. In this work, the LBLRTM model we previously developed is adopted [14]. Within this model, the 120 km-thick atmosphere is represented by a series of plane-parallel layers. Vertical profiles of temperature, pressure, and atmospheric composition are prescribed using the Air Force Geophysical Laboratory (AFGL) mid-latitude summer standard atmosphere. Seven atmospheric gases are included: water vapor, carbon dioxide, ozone, methane, nitrous oxide, oxygen, and nitrogen. Spectral absorption coefficients for these gases are obtained from the HITRAN database, and continuum absorption coefficients are computed using the MT_CKD continuum model [14]. To capture real-time variability in atmospheric conditions, a linear calibration is applied to the tropospheric temperature profile over 0–14 km, together with a nighttime temperature inversion at 1 km [13], as defined in Eq. (1). The vertical water vapor profile is calibrated using ground-level temperature and humidity measurements [13], while the profiles of the remaining gases are adjusted using the most recent global average measurements, as described in Eqs. (2) and (3).

$$T(z) = \begin{cases} T_0 + \frac{z}{z_{\text{trop}}}\left[T_{\text{AFGL}}(z_{\text{trop}}) - T_0\right], & \text{day, } 0 \le z < z_{\text{trop}} \\ T_0 + \frac{z}{z_{\text{inv}}}\left[T_{\text{off}} + \frac{z_{\text{inv}}}{z_{\text{trop}}}\left[T_{\text{trop}} - (T_0 + T_{\text{off}})\right]\right], & \text{night, } 0 < z < z_{\text{inv}} \\ T_0 + T_{\text{off}} + \frac{z}{z_{\text{trop}}}\left[T_{\text{AFGL}}(z_{\text{trop}}) - (T_0 + T_{\text{off}})\right], & \text{night, } z_{\text{inv}} \le z < z_{\text{trop}} \end{cases} \tag{1}$$

$$w_{\text{gas}}(z) = w_{\text{gas}}(0)\frac{w_{\text{AFGL}}(z)}{w_{\text{AFGL}}(0)}, \qquad \text{for z ranges from 0 to 120 km} \tag{2}$$

$$w_{\text{H}_2\text{O}}(0) = \frac{\phi_0 P_s(T_0)}{P_{\text{atm}}}, \quad \text{where} \quad P_s(T_0) = P_{\text{ARM}} \exp\left(\frac{c_{\text{ARM}}(T_0 - 273.15)}{T_0 - 30.11}\right), \tag{3}$$

Here, T_0 and ϕ_0 denote the ground-level air temperature (ranging from 234 K to 324 K) and relative humidity (ranging from 0% to 100%), respectively. z_{trop} denotes the altitude of the tropopause, which is 14 km. z_{inv} and T_{off} denote the altitude and temperature offset for nighttime temperature inversion, which occurs at 1 km with an offset of 12 K. $w_{\text{gas}}(0)$ denotes the most recent global average surface volumetric mixing ratio of a given gas (e.g., 430 ppm for CO_2) [14]. For water vapor, $w(0)$ is a function of surface temperature and relative humidity, as given in Eq. (3). $P_s(T_0)$ denotes the saturated water vapor pressure at temperature T_0, calculated using the August–Roche–Magnus (ARM) expression, with $P_{\text{ARM}} = 610.94$ Pa and $c_{\text{ARM}} = 17.625$.

The vertical aerosol profile is derived from Cloud-Aerosol LIDAR and Infrared Pathfinder Satellite Observations (CALIPSO) measurements over North America. Aerosol particles are assumed to be spherical, with size distributions following a standard bimodal lognormal distribution,

$$\frac{dN}{d\ln r} = r\frac{dN}{dr} = rn(r) = \sum_{i=1}^{I} \frac{N_i}{\sqrt{2\pi}\ln\sigma_i} \exp\left[-\frac{1}{2}\left(\frac{\ln(r/r_{m,i})}{\ln\sigma_i}\right)^2\right], \tag{4}$$

where I denotes the total number of modes. For each mode i, $r_{m,i}$ is the modal radius, σ_i is the geometric standard deviation, and N_i is the mode amplitude. To account for interactions with water vapor, both the aerosol size distribution and refractive index are allowed to vary with the ambient relative humidity. The aerosol microphysical radiative properties, including absorption and scattering coefficients as well as the asymmetry factor, are evaluated using Mie theory, as shown in Fig. 2. Additional details are provided in [14].

Water clouds are modeled using a similar approach to that used for aerosols [30]. Specifically, each cloud droplet is assumed to be spherical, with a size distribution following a Gamma distribution,

$$n(r) = r^{1/\sigma_e - 3} \exp\left(-\frac{r}{r_e\sigma_e}\right), \tag{5}$$

where the effective radius r_e is set to 10 μm and the effective variance σ_e is fixed at 0.1 [30].

The cloud microphysical radiative properties, including scattering, absorption, and extinction coefficients, as well as the asymmetry factor, are calculated using Mie theory and are shown in Fig. 2. In the LBLRTM, clouds are assumed to be vertically uniform within each atmospheric layer, although they may span multiple layers to represent realistic clouds with varying thicknesses and altitudes. In this work, the term cloud optical properties refers to the macroscopic vertical cloud structure, including cloud thickness, cloud-top height, cloud-base height, and cloud optical depth (COD).

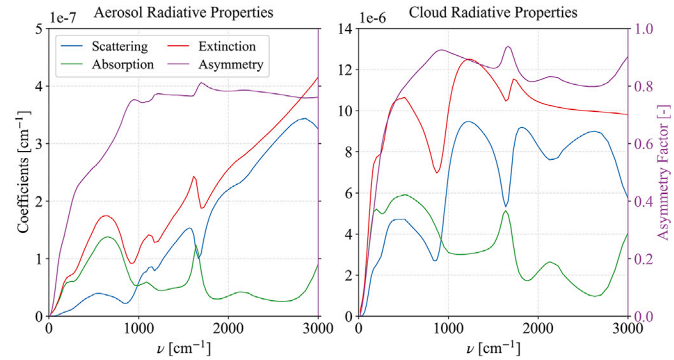

Fig. 2. Spectral microphysical radiative properties for aerosol (left) and cloud (right), including scattering, absorption, and extinction coefficients, as well as the asymmetry factor, across the longwave range (0-3000 cm^{-1}). The aerosol properties shown correspond to an average altitude of 2.5 km under ambient relative humidity conditions of 100%, accounting for the presence of clouds within this atmospheric layer.

Table 2
Altitudinal information of three typical cloud types in polar, temperate, and tropical regions [31].

Type	Polar Region	Temperate Region	Tropical Region
High Clouds	3–8 km	5–14 km	6–18 km
Mid Clouds	2–4 km	2–7 km	2–8 km
Low Clouds	0–2 km	0–2 km	0–2 km

To systematically represent vertical cloud distributions, clouds are categorized into three altitudinal types—low, mid, and high clouds [31], as presented in Table 2. For each cloud type, all possible vertical distributions are considered, including both single-layer and multilayer structures. In total, 225 distribution patterns are included for each set of COD at each set of ground level conditions (i.e., ambient temperature and relative humidity), which is computationally intensive. For each case, the cloud properties are represented by an N-dimensional cloud vector, corresponding to the N-layer atmospheric configuration. In this study, a 32-layer atmosphere is adopted according to the grid convergence test in [13]. Within this vector, a value of zero indicates the absence of clouds (clear-sky) in this layer, while positive values denote the presence of clouds (cloudy layer), with the specific value representing the COD (τ).

To improve computational efficiency, similar cloud distributions are eliminated by quantifying the similarity between the modeled upwelling spectral fluxes at TOA using Euclidean distance (ED) [32]. The ED between two modeled flux vectors corresponding to different cloud distributions is expressed as Eq. (6). A threshold of 0.30 is applied to identify representative cloud distributions, as illustrated in Fig. 3. Specifically, only one case is preserved in instances where ED < 0.30, ensuring a balance between accuracy and computational efficiency.

$$\|\mathbf{F}^{*\uparrow}(x_i), \mathbf{F}^{*\uparrow}(x_j)\| = \sqrt{\sum_{b=1}^{n}\left(F_b^{*\uparrow}(x_i) - F_b^{*\uparrow}(x_j)\right)^2} \tag{6}$$

Here, $F_b^{*\uparrow}$ denotes the normalized modeled upwelling fluxes for band b, with surface conditions of $T_a = 294$ K and $\phi = 50\,\%$. x_i and x_j denote two different cloud vectors.

To account for the anisotropic scattering of aerosols and clouds, the δ-M approximation is used to correct the extinction coefficient and single-scattering albedo in each atmospheric layer. The longwave radiative transfer is then treated as diffuse, allowing the use of a two-flux approach. Specifically, the spectral radiosity (J) and irradiation (G) in each atmospheric layer are first obtained by simultaneously solving the

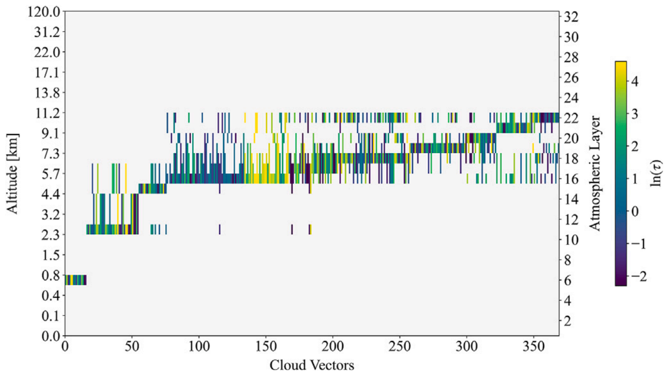

Fig. 3. Demonstration of representative cloud vectors sorted by the cloud vertical distribution and COD. The left axis shows atmospheric altitude while the right axis indicates the atmospheric layers. Each column represents a cloud vector within the 32-layer atmospheric configuration, where non-zero values indicate the logarithmically transformed CODs ($\ln(\tau)$) for cloudy layer.

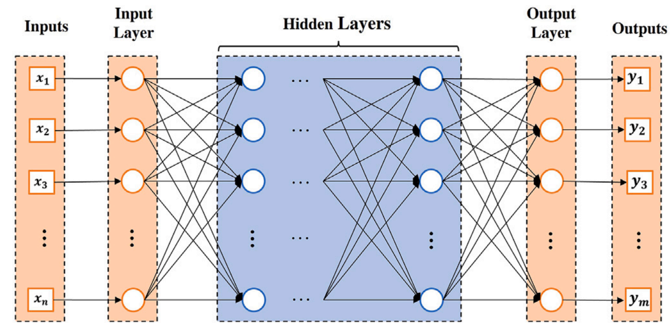

Fig. 4. The fundamental configuration of the MLP architecture, with the fully connected structure comprising one input layer, one or more hidden layers, and one output layer.

radiosity–irradiation balance equations for all layers. The spectral upwelling and downwelling fluxes across each layer boundary are then calculated from the corresponding radiosity and irradiation using appropriate transfer factors. Additional details of solving for the fluxes are available in [14].

The measured (from satellites) and estimated (from LBLRTM) upwelling fluxes at the top of the atmosphere (TOA), together with the estimated (from LBLRTM) downwelling flux at the surface (SDLR), as measured by ground-based pyrgeometers, are calculated as follows:

$$F_b^{\uparrow} = \pi I_{\nu,b}^{\uparrow} \int R_b(\nu) d\nu \qquad \text{and} \qquad \hat{F}_b^{\uparrow} = \pi \int_b \hat{I}_\nu^{\uparrow} R_b(\nu) d\nu \tag{7}$$

$$\hat{F}_0^{\downarrow} = \pi \int_{\nu_1}^{\nu_2} \hat{I}_\nu^{\downarrow} d\nu \tag{8}$$

Here, ν denotes the wavenumber, $F_b^{\uparrow}$ represents the upwelling flux in band b of the satellite data, calculated by multiplying the satellite-measured spectral intensity $I_{\nu,b}^{\uparrow}$ with the integration of the spectral response function (SRF) R_b of the ABI over the wavenumber range of band b. Similarly, $\hat{F}_b^{\uparrow}$ represents the estimated upwelling flux in band b, calculated using the spectral intensity output from the LBLRTM, weighted by the SRF in band b. Finally, $\hat{F}_0^{\downarrow}$ denotes the estimated SDLR, derived by integrating the modeled surface downwelling spectral intensity over the spectral response range of the pyrgeometers.

3.2. The chosen deep learning models

Deep neural networks (DNNs) have evolved into compelling and effective tools in various climate and earth science related research [25]. Among the various DNN models, the multilayer perceptron (MLP) is, by far, the most popular architecture due to its structural flexibility, excellent representational capabilities, and the availability of a large number of training algorithms [33]. In this study, two MLPs are incorporated to construct the two-stage surrogate modeling framework: an MLP classifier for cloud retrieval and an MLP predictor (hereafter referred to as a DNN regressor) for SDLR estimations.

The MLP is an advanced feedforward neural network designed to map input vectors to the corresponding output vectors [34]. Its architecture consists of multiple layers, including the input layer, hidden layers, and the output layer. This structure forms a fully connected network, where each node in a given layer is connected to all nodes in the adjacent preceding and succeeding layers. Typically, MLPs include one or more hidden layers, with each neuron employing a nonlinear activation function. The primary training method for MLPs is the backpropagation algorithm, a widely used supervised learning technique that optimizes network weights by minimizing the error between the predicted and actual outputs through gradient descent. This iterative process enables the MLP to generalize effectively and learn from labeled datasets, making it a powerful tool for solving both classification and regression problems [35]. The fundamental architecture of the MLP is illustrated in Fig. 4.

3.3. The proposed two-stage surrogate model

In this work, a novel two-stage surrogate modeling framework is proposed, which emulates high-fidelity LBLRTM by leveraging DL techniques with powerful learning capabilities. The overall working principle is illustrated in Fig. 5.

During the training stage, the spectral TOA upwelling fluxes ($F_b^{\uparrow}$) are simulated using the LBLRTM for various cloud optical properties, which are characterized by cloud vertical distribution and COD, under a range of ground-level conditions with different air temperatures and relative humidity levels. Based on these LBLRTM simulations, a two-stage surrogate modeling framework is constructed, consisting of two DL models. In Stage I, the normalized TOA upwelling fluxes in all longwave bands from the LBLRTM serve as input features for an MLP classifier, with the corresponding cloud vectors (Fig. 3) as the prediction targets. In Stage II, another DNN regressor maps the N-dimensional cloud vector, along with normalized ancillary meteorological variables (e.g., temperature and humidity), to the modeled SDLR ($F_0^{\downarrow}$) from LBLRTM.

Prior to model training, all input features are pre-processed for improved model performance. To eliminate the effects of other atmospheric constituents, the clear-sky index of upwelling flux is used, as defined by Eq. (9). The cloud vectors are constructed using a natural logarithmic scale (see Eq. (10)), following established practices in satellite cloud retrieval and atmospheric science research [15,36]. Ancillary meteorological variables are normalized using the Min-Max normalization technique (see Eq. (11)), which has been widely employed in DL to scale data within certain ranges [37].

$$k_b^{\uparrow} = \frac{\hat{F}_b^{\uparrow}}{\hat{F}_{b,clc}^{\uparrow}} \tag{9}$$

$$\tau^* = \begin{cases} \ln(\tau), & \text{if cloudy layer, where } \tau \in [0.1, 100]. \\ \tau, & \text{if clear-sky layer, where } \tau = 0. \end{cases} \tag{10}$$

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \qquad \text{with } x^* \in [0, 1]. \tag{11}$$

where $\hat{F}_b^{\uparrow}$ denotes the modeled TOA upwelling flux for given cloud optical properties in the spectral band b, and $F_{b,clc}^{\uparrow}$ denotes the corresponding values under clear-sky conditions. τ^* denotes the logarithmically transformed CODs. x^* and x denote the normalized and original ancillary

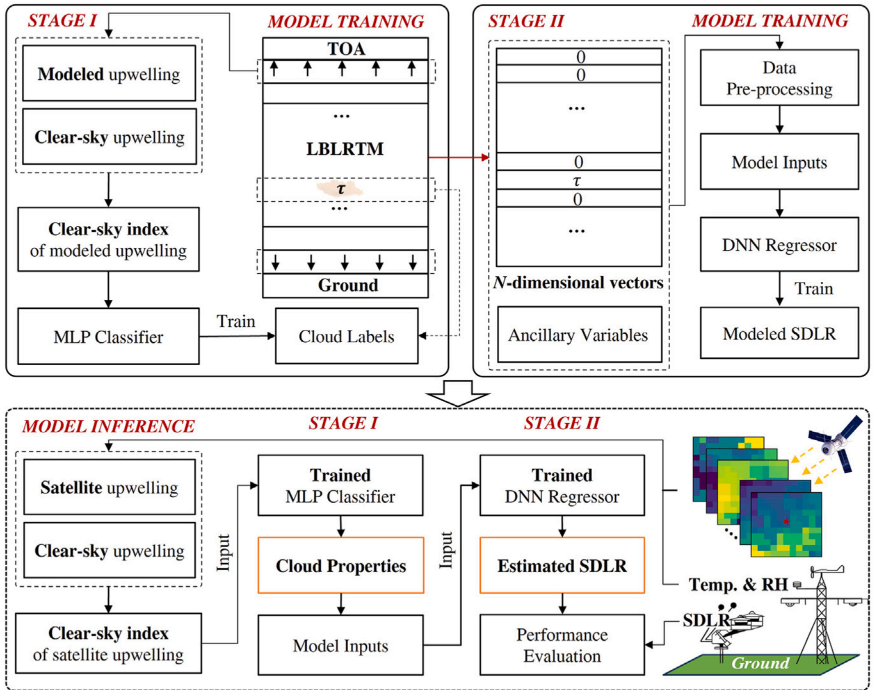

Fig. 5. Schematic overview of the proposed two-stage surrogate modeling framework.

Table 3
Optimized hyperparameter setting for the MLP classifier and the DNN regressor, as determined by Keras Tuner [38].

Sky Conditions	MLP Classifier		DNN Regressor	
	Parameters	Values	Parameters	Values
Daytime	Hidden unit 1	64	Input unit	256
	Hidden unit 2	40	Hidden units	64-32-16
	Learning rate	0.01	Learning rate	0.0001
Nighttime	Hidden unit 1	56	Input unit	512
	Hidden unit 2	16	Hidden units	64-32-32
	Learning rate	0.001	Learning rate	0.0001

meteorological variables (e.g., T_a, ϕ), $x_{\min}$ and $x_{\max}$ represent the minimum and maximum values.

In this work, the MLP classifier comprises two hidden layers, while the DNN regressor consists of five fully connected layers: one input layer, three hidden layers, and one output layer. Both architectures are optimized using Keras Tuner [38], and the optimized hyperparameters are summarized in Table 3. For the MLP classifier, the number of units in each hidden layer is tuned within the range of 16 to 64, in increments of 8. For the DNN regressor, the number of units in each layer is tuned independently as 2^x, where $x \in [3, 9]$. The learning rates for both architectures are tuned as 10^{-x}, over the range 10^{-5} to 10^{-1}.

During the inference stage, the TOA upwelling spectral flux in all longwave bands, $\hat{F}_b^{\uparrow}$, measured by the satellite at time t, is first converted to the clear-sky index. The resulting clear-sky indices are then fed into the well-trained MLP classifier (Stage I) to produce the retrieved N-dimensional cloud vectors. These cloud vectors, along with normalized ancillary meteorological variables, are input into the well-trained DNN regressor (Stage II) to estimate the SDLR. Finally, the SDLR estimates are compared against ground-based measurements to evaluate the model's performance.

3.4. Performance evaluation

The accuracy of SDLR estimation is quantified using the root mean squared error (RMSE) and the mean bias error (MBE) [39] when compared with ground measurements, which are defined as,

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(f_i - o_i)^2} \tag{12}$$

$$\mathrm{MBE} = \frac{1}{N}\sum_{i=1}^{N}(f_i - o_i) \tag{13}$$

where f_i and o_i represent the estimated and measured SDLR, respectively, and N denotes the total number of data points.

4. Results and discussion

4.1. Sensitivity analysis

To examine the impact of atmospheric and cloud parameters on SDLR and ensure physical interpretability, a comprehensive sensitivity analysis was conducted based on outputs from both the physical model (LBLRTM) and the proposed surrogate model. The sensitivities of four key parameters are investigated: cloud-top height (z_t), COD (τ), ambient temperature (T_a), and relative humidity (ϕ).

4.1.1. Sensitivity analysis based on the physical model (LBLRTM)

To investigate the sensitivity of the physical behavior captured by LBLRTM, we first examine the spectral top-of-atmosphere upwelling flux, $F_{\mathrm{TOA}}^{\uparrow}$, and the surface downwelling flux, $F_0^{\downarrow}$, under a range of atmospheric conditions using LBLRTM simulations.

As shown in Fig. 6, increasing cloud top height z_t reduces both $F_{\mathrm{TOA}}^{\uparrow}$ and $F_0^{\downarrow}$, as higher cloud tops emit less longwave radiation due to their colder temperatures. However, the influence on $F_0^{\downarrow}$ is smaller, as SDLR is primarily dominated by emission from the lower atmosphere. Cloud optical properties have an evident impact on $F_0^{\downarrow}$, while their effect on $F_{\mathrm{TOA}}^{\uparrow}$ is comparatively weaker. Increasing τ slightly decreases $F_{\mathrm{TOA}}^{\uparrow}$ in the atmospheric window region (approximately 800–1200 cm^{-1}), which includes ABI bands 11, 13, 14, and 15. This occurs because optically thick clouds replace the warmer surface emission with emissions from colder cloud tops. In contrast, ABI bands 8–10, which correspond to water vapor bands, are largely insensitive to changes in τ because they

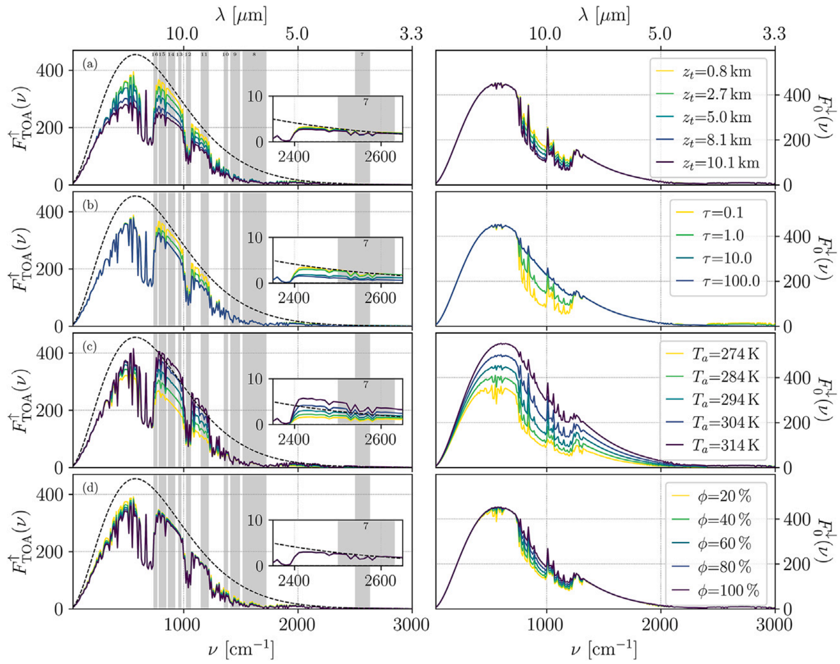

Fig. 6. Sensitivity analysis of TOA upwelling spectral flux $F^{\uparrow}_{TOA}$ (mW m^{-2} cm) and surface downwelling spectral flux $F^{\downarrow}_{0}$ (mW m^{-2} cm), as a function of (a) cloud-top height (z_t) when $\tau = 1.0, \phi = 50\%, T_a = 294$ K; (b) COD (τ) when $z_t = 2.7$ km, $\phi = 50\%, T_a = 294$ K; (c) ambient temperature (T_a), when $z_t = 2.7$ km, $\tau = 1.0, \phi = 50\%$; and (d) ambient relative humidity (ϕ) when $z_t = 2.7$ km, $\tau = 1.0, T_a = 294$ K. Single-layer clouds within the 32-layer atmospheric configuration are assumed throughout the analysis. For example, $z_t = 2.7$ km corresponds to a single-layer cloud with a base height of 2.3 km and cloud average temperature of 282 K. The ten longwave ABI bands (7–16) for GOES-16 are shaded with gray color. The black dashed line in each subplot shows the blackbody spectral emissive power ($E_{b,\nu}$) at $T_a = 294$ K.

are primarily dominated by atmospheric moisture. Conversely, $F^{\downarrow}_{0}$ increases with τ but gradually approaches a plateau, as optically thicker clouds behave more like blackbody emitters and enhance SDLR up to the blackbody emissive power.

Surface-level temperature and humidity also play critical roles in radiative fluxes, especially in SDLR. Increasing T_a results in larger $F^{\uparrow}_{TOA}$ because a warmer atmosphere and clouds emit more thermal radiation. The increase in $F^{\downarrow}_{0}$ with T_a is even more substantial, as warmer near-surface atmospheric layers emit more radiation toward the surface. Higher ϕ increases $F^{\downarrow}_{0}$ by enhancing longwave emission from the moist lower atmosphere, particularly in spectral bands where water vapor is the primary radiative contributor. In contrast, the effect of ϕ on $F^{\uparrow}_{TOA}$ is much weaker under cloudy conditions because the TOA primarily receives emission from clouds.

Overall, the sensitivity analysis highlights the strong linkage between cloud optical properties and TOA upwelling flux, as well captured by the LBLRTM, which lays the theoretical foundation for retrieving cloud properties from satellite observations. Furthermore, in addition to cloud properties, ambient atmospheric conditions play a dominant role in SDLR estimation. These results underscore the physical relationships among cloud properties, atmospheric conditions, and radiative fluxes, providing theoretical guidance for the surrogate modeling framework.

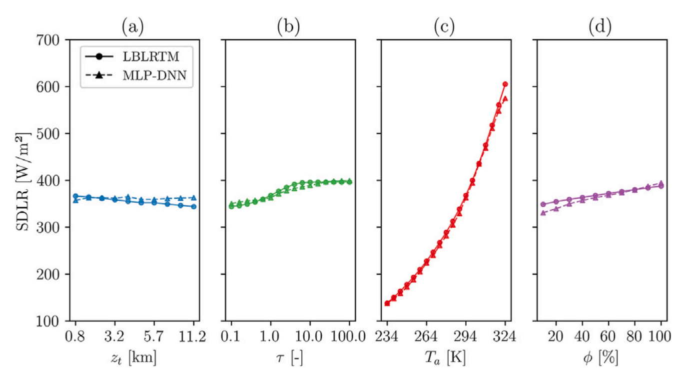

Fig. 7. Sensitivity analysis of SDLR estimations from the surrogate model (dashed lines) compared to the LBLTRM simulations (solid lines), in response to variations in (a) cloud-top height (z_t) when $\tau = 0.6, \phi = 50\%, T_a = 294$ K, (b) COD (τ) when $z_t = 2.7$ km, $\tau = 1.0, \phi = 50\%$, (c) ambient temperature (T_a) when $z_t = 2.7$ km, $\tau = 1.0, \phi = 50\%$, and (d) relative humidity (ϕ) when $z_t = 2.7$ km, $\tau = 1.0, T_a = 294$ K. Single-layer clouds are assumed throughout the analysis, consistent with Fig. 6.

4.1.2. Sensitivity analysis based on the surrogate model

To further examine how the input parameters affect the surrogate model outputs, a sensitivity analysis was conducted for SDLR estimation, as shown in Fig. 7.

Among the investigated variables, cloud-top height exhibits a limited impact, indicating that SDLR is less sensitive to cloud-top emission than to the radiative contributions from the lower atmosphere. The surrogate model reproduces this behavior well, showing a relatively stable trend with only slight deviations from the LBLRTM results at higher z_t. For COD (τ), the surrogate model excellently captures the initial increase and subsequent plateau in SDLR, reflecting the effect as clouds approach blackbody-like emission. Ambient temperature demonstrates the most pronounced nonlinear positive correlation with SDLR, consistent with enhanced longwave radiation from warmer near-surface atmospheric layers toward the surface. The surrogate model follows this behavior closely, with minor underestimation at extremely high temperatures (e.g., above 320 K). Relative humidity also shows steady positive correlations with SDLR. Although the surrogate model slightly underestimates SDLR at low humidity levels (e.g., below 20%), it converges closely with the LBLRTM results at higher values, indicating that it captures the radiative contribution of atmospheric water vapor well. Overall, these

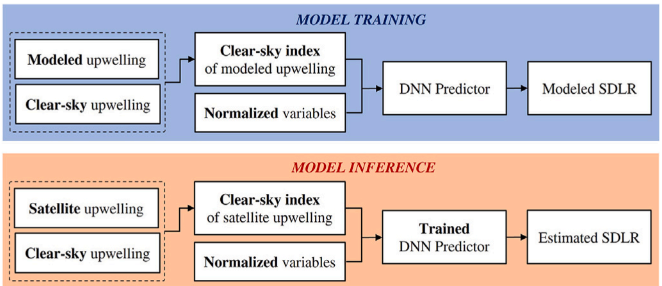

Fig. 8. Flowchart of the end-to-end satellite-to-SDLR surrogate approach, in which a DNN predictor is employed to directly estimate SDLR from satellite data.

results demonstrate that the proposed surrogate model accurately captures the complex, nonlinear dependencies inherent in radiative transfer processes, confirming its ability to represent the underlying physical mechanisms.

4.2. Comparison with end-to-end surrogate approach

The performance of the proposed two-stage model is compared against a direct satellite-to-SDLR mapping approach. The end-to-end benchmark model employs a DNN predictor, using the same architecture as the DNN regressor in Stage II of the proposed method, to estimate SDLR directly from satellite-measured TOA upwelling radiation and ancillary meteorological variables (e.g., temperature and humidity). This bypasses the intermediate step of estimating the cloud vector, as illustrated in Fig. 8. To ensure a fair and robust comparison between the two approaches, identical pre-processing procedures are applied. In particular, the clear-sky index of upwelling flux (Eq. 9) is used to emphasize the effects of clouds, while ancillary variables (T_a and ϕ) are normalized using the Min-Max normalization technique (see Eq. (11)). Additionally, the DNN architecture is optimized using the Keras Tuner [38] to ensure optimal performance.

The joint and marginal distributions of measured and estimated SDLR across the SURFRAD stations are shown in Fig. 9. The results underscore the superior performance of the proposed two-stage surrogate approach compared to the direct end-to-end approach (DNN). Specifically, the direct approach exhibits noticeable discrepancies in both joint and marginal distributions, with more outlier values. Quantitatively, the direct model exhibits higher prediction errors and bias, with MBE values ranging from $-13.44\,\mathrm{W/m^2}$ to $37.78\,\mathrm{W/m^2}$, and RMSE ranging from $23.63\,\mathrm{W/m^2}$ to $46.53\,\mathrm{W/m^2}$. Moreover, its performance deteriorates more during nighttime than during daytime, primarily due to its inability to account for radiative effects associated with atmospheric temperature inversion at night. These results indicate that a direct regression relationship between TOA upwelling radiation and SDLR is

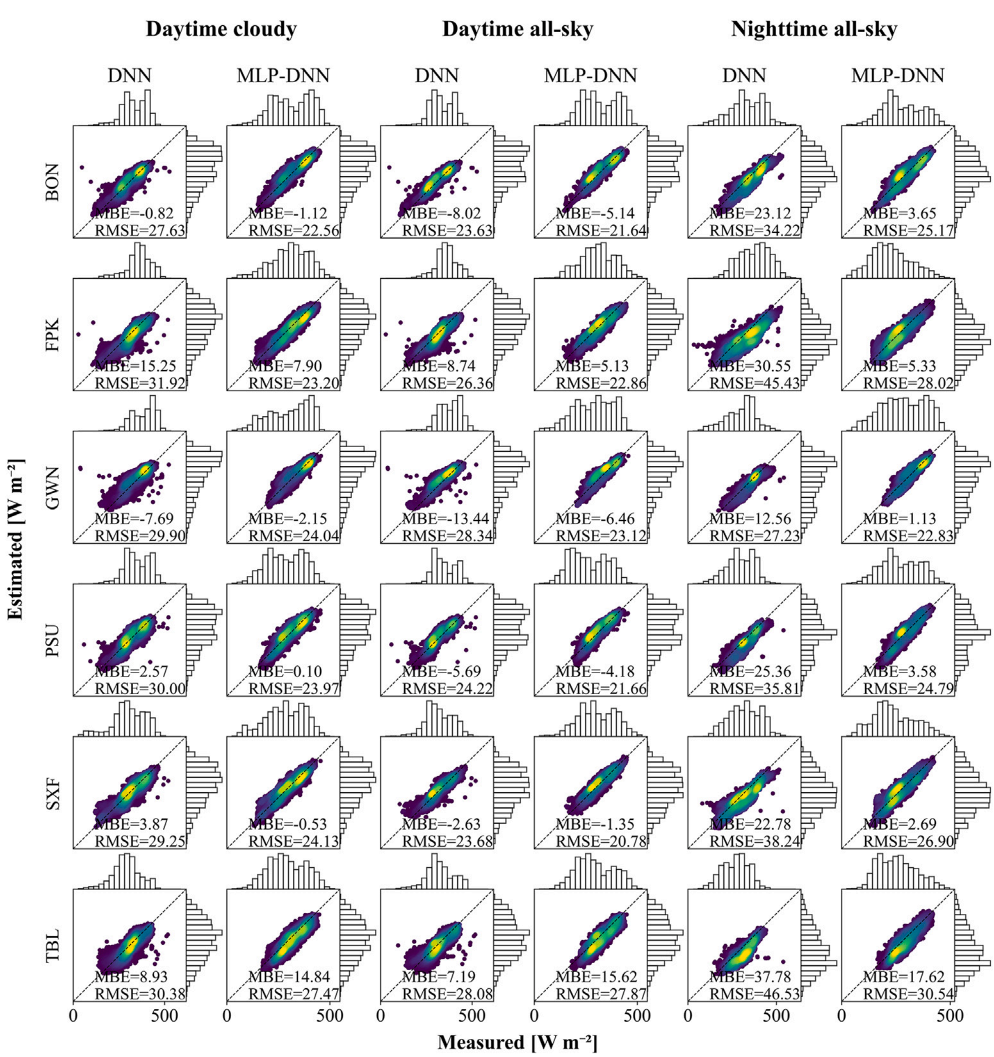

Fig. 9. Joint and marginal distributions for SDLR measurements and estimations across the SURFRAD stations. The joint distributions illustrate the alignment of estimated SDLR with measurements, while the marginal distributions demonstrate the distribution of the SDLR.

inadequate to capture the complexities of atmospheric radiative transfer processes.

In contrast, the two-stage model exhibits strong alignment between estimated and measured SDLR, along with improved agreement in the SDLR marginal distributions. Furthermore, notable improvements in error metrics are also achieved, with MBE values ranging from 0.10 W/m^2 to 17.62 W/m^2 and RMSE values ranging from 20.78 W/m^2 to 30.54 W/m^2, consistently outperforming the direct DNN approach. These findings emphasize the effectiveness of incorporating cloud optical properties as intermediate physical descriptors, enhancing the ability to capture radiative transfer complexities while improving physical interpretability.

4.3. Comparison with benchmark methods

To comprehensively evaluate the performance of the proposed surrogate model, two previously developed benchmark methods are included for comparison: one empirical method [1] and one physical method [13]. The empirical method [1] incorporates measured cloud information, such as cloud fraction (CF) or cloud modification factor (CMF, available only during daytime) to estimate SDLR, as described in Eq. (14). The CMF data are calculated from global horizontal irradiance (GHI) and clear-sky GHI as defined in Eq. (15), with both values obtained from SURFRAD. The CF data were collected from the closest automated surface observing system stations corresponding to the SURFRAD stations. Further details can be found in [1].

$$\mathrm{SDLR} = \begin{cases} \mathrm{SDLR_{clear}}(1 - c_1 \mathrm{CMF}^{c_2}) + c_3 \sigma T_a^4 \mathrm{CMF}^{c_4} \phi^{c_5}, & \text{Daytime} \\ \mathrm{SDLR_{clear}}(1 - c_1 \mathrm{CF}^{c_2}) + c_3 \sigma T_a^4 \mathrm{CF}^{c_4} \phi^{c_5}, & \text{Daytime or Nightime} \end{cases} \tag{14}$$

$$\mathrm{CMF} = 1 - \frac{\mathrm{GHI}}{\mathrm{GHI_{clear}}} \tag{15}$$

where $\mathrm{SDLR_{clear}}$ denotes the clear-sky values calculated from the SDLR clear-sky model. T_a and ϕ denote the measured air temperature and relative humidity, and σ is the Stefan-Boltzmann constant. The coefficients c_1 through c_5 are empirical parameters regressed from observational SDLR data [1]. The physical method, referred to as the SCOPE method, utilizes a pre-computed look-up table to retrieve cloud properties from satellite data. These cloud properties are then used to compute SDLR through the LBLRTM model [13]. A summary of all evaluated methods is provided in Table 4.

Fig. 10 presents the RMSE and MBE of SDLR estimations across SURFRAD stations under different sky conditions. Among the evaluated methods, the empirical methods generally demonstrate the smallest errors. Specifically, the EMP (CMF) method performs best during daytime scenarios (both cloudy and all-sky). However, this approach relies on ground-level GHI data and is unavailable during nighttime, limiting its ability to provide consistent day-and-night operation. The EMP (CF) method exhibits competitive performance but shows poorer results at the SXF station, with MBE ranging from 16 W/m^2 to 23 W/m^2 and RMSE ranging from 25 W/m^2 to 30 W/m^2. This indicates a strong dependence on the applicability of regressed coefficients and ground-based cloud information specific to the target location. Additionally, its 60-min temporal resolution hinders compatibility with continuously operating applications that require high-resolution data for performance evaluation. Moreover, compared to the proposed surrogate method, the empirical approaches focus solely on surface-level products and lack real-time cloud optical properties, which are essential for accurately characterizing cloud fields for outdoor energy applications.

The SCOPE method demonstrates superior performance compared to the proposed surrogate during nighttime, achieving a relatively low RMSE (20-26 W/m^2). However, it suffers from higher bias, with MBE ranging from 5.0 W/m^2 to approximately 17.5 W/m^2. In contrast, the proposed surrogate model consistently outperforms SCOPE during the daytime (both cloudy and all-sky scenarios), with the exception of the TBL. Significant improvements are observed with the surrogate model, including reduced bias (MBE ranging from 0.1 W/m^2 to 17.5 W/m^2) and improved accuracy (RMSE ranging from 20.5 W/m^2 to 30.5 W/m^2). These results underscore the superiority of the surrogate model over the physics-based approach (e.g., SCOPE). Furthermore, compared with SCOPE, the proposed method eliminates the need to run complex RTMs, thereby improving the accessibility of radiative transfer modeling for non-experts and supporting a broader range of operational applications.

The distributions of RMSE and MBE across all stations are further evaluated, as illustrated in Fig. 11. Compared to the benchmarks, the proposed surrogate method achieves notable bias reductions, consistently demonstrating the lowest or near-zero median MBE across all stations. This improvement is particularly pronounced at night, where its median MBE is 3.6 W/m^2, compared to 4.6 W/m^2 and 6.8 W/m^2 for the empirical and SCOPE methods, respectively. In terms of RMSE, the proposed method performs competitively during the daytime, although an increase in RMSE is observed at night. During daytime cloudy periods, the median RMSE for MLP-DNN is 24.0 W/m^2, which is lower than that of SCOPE (26.0 W/m^2). Comparable patterns are observed in daytime all-sky scenarios, where its median RMSE is 22.3 W/m^2, compared to benchmark values ranging from 18.2 W/m^2 to 25.5 W/m^2.

The mean and standard deviation of MBE and RMSE across geographic regions further highlight the superior performance of the proposed approach. As shown in Fig. 11, the proposed surrogate model shows low variability across different stations, with a standard deviation typically below 6.5 W/m^2 for MBE and 2.5 W/m^2 for RMSE. Notably, it achieves the lowest MBE mean in daytime all-sky conditions (0.6 W/m^2), outperforming EMP(CMF) (5.6 W/m^2) and SCOPE (-4.4 W/m^2), while maintaining a competitive RMSE mean (23.0 W/m^2). In nighttime all-sky conditions, the proposed surrogate model achieves a lower MBE mean (5.7 W/m^2) compared to SCOPE (8.5 W/m^2), with a slightly higher RMSE standard deviation. In contrast, EMP(CF) exhibits high MBE variability, with standard deviation values reaching up to 9.1 W/m^2, signifying its reliance on location-specific coefficients and cloud information. While EMP(CMF) performs well in daytime cloudy conditions, with an RMSE mean of 20.6 W/m^2 and a standard deviation of 4.2 W/m^2, its unavailability during nighttime limits its applicability. These findings indicate the robust performance of the proposed approach across diverse geographic regions, making it a reliable method for continuous, high-resolution cloud properties and SDLR retrieval.

Table 4
Summary of all evaluated methods. EMP (CMF) and EMP (CF) are empirical method based on CMF and CF, while SCOPE and MLP-DNN represent physical and the proposed surrogate modeling approaches, respectively.

Methods	Inputs	Temporal Resolution	Availability	Outputs
EMP (CMF)	GHI, $\mathrm{GHI_{clear}}$, T_a, ϕ	5-min	Daytime	SDLR
EMP (CF)	CF, T_a, ϕ	60-min	All day	SDLR
SCOPE	$\hat{F}_b^{\uparrow}$, T_a, ϕ	5-min	All day	SDLR, cloud properties
MLP-DNN	$\hat{F}_b^{\uparrow}$, T_a, ϕ	5-min	All day	SDLR, cloud properties

4.4. Comparison with existing datasets

Reanalysis and satellite-based datasets are widely used in atmospheric and climate studies due to their global coverage and the availability of comprehensive meteorological variables. In this work, three benchmark datasets with SDLR products are adopted as benchmarks, as summarized below:

- ECMWF Reanalysis, Version 5 (ERA5): ERA5 [40] provides hourly SDLR at 0.25° (~31 km) resolution based on the European Center for Medium-Range Weather Forecasts (ECMWF) reanalysis. It is

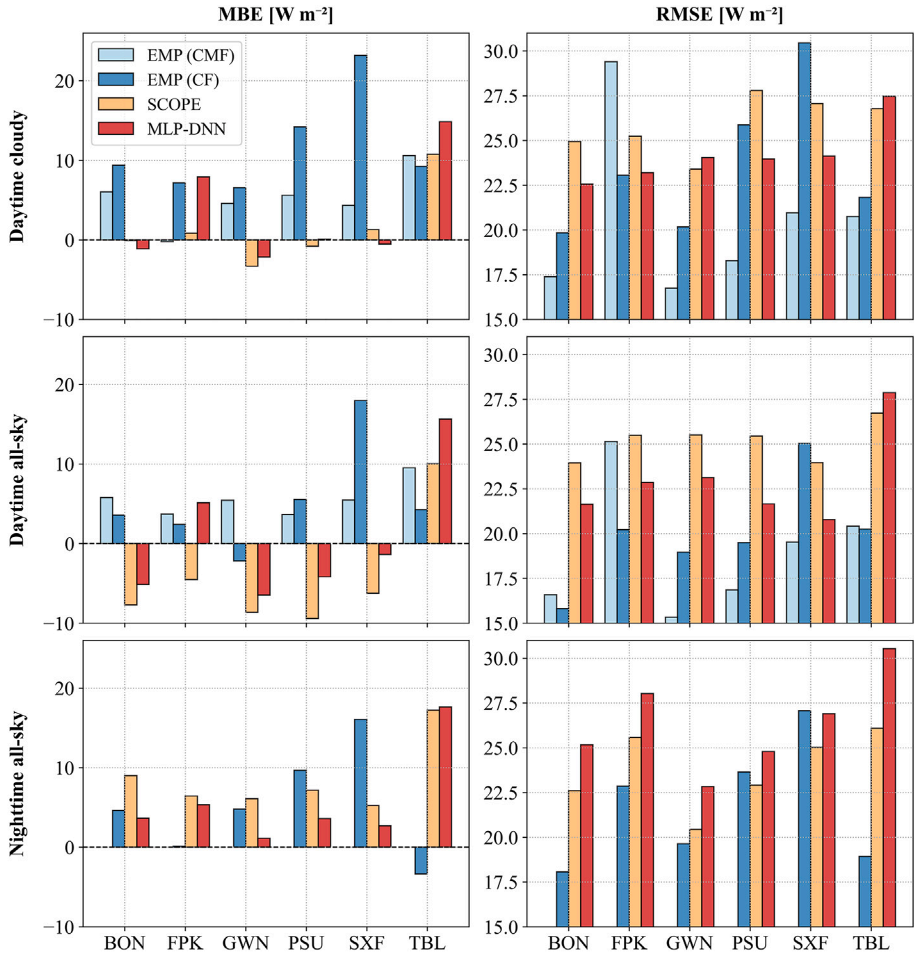

Fig. 10. Comparison of RMSE and MBE of SDLR estimations between the proposed surrogate model and benchmark methods across the SURFRAD stations.

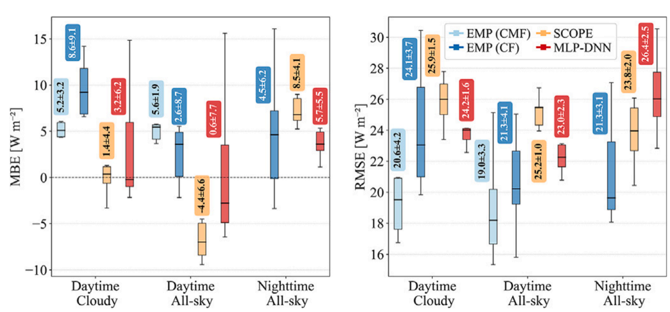

Fig. 11. Comparison of the distributions of MBE and RMSE for the proposed two-stage surrogate model (MLP-DNN) compared to benchmark methods. The black horizontal line within each box represents “median”, while the annotated text indicates “mean ± standard deviation”.

updated daily with a latency of about 5 days, while the final consolidated product is typically released 2—3 months later.

- Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2): MERRA-2 [41] provides hourly output at 0.5° × 0.625° resolution (∼ 62 km) with the Goddard Earth Observing System Model (GEOS). The dataset is generally available about 3 weeks after the end of each month.
- Clouds and the Earth’s Radiant Energy System (CERES): CERES [42] delivers hourly longwave radiative flux estimates at a spatial resolution of approximately 1° (∼124 km). These fluxes are derived by combining CERES observations with cloud properties retrieved from MODIS and geostationary satellites, together with radiative transfer calculations from the Langley Fu–Liou model [43]. The currently released official record spans 2000 to 2025, with additional data becoming available as forward processing continues.

To ensure fair comparisons with the reanalysis and satellite-based datasets, SDLR outputs from the proposed surrogate model and ground-based SDLR measurements are resampled to an hourly resolution.

As shown in Table 5, the proposed surrogate model (MLP-DNN) exhibits markedly improved bias characteristics relative to the benchmark datasets. ERA5 and, more notably, MERRA-2 exhibit systematic negative biases, indicating persistent underestimation of SDLR. CERES shows more heterogeneous behavior, with MBEs spanning $-20.88\,\mathrm{W/m^2}$ to $8.00\,\mathrm{W/m^2}$. In contrast, the surrogate model achieves MBEs ranging from $-4.27\,\mathrm{W/m^2}$ to $8.17\,\mathrm{W/m^2}$ in most scenarios, indicating substantially reduced bias and more balanced SDLR estimates across diverse geographical locations.

In terms of RMSE, ERA5 and CERES remain the strongest benchmarks overall, generally yielding the lowest errors across stations and sky conditions. Nevertheless, the proposed surrogate model remains within competitive ranges of 21.15–28.91 $\mathrm{W/m^2}$ and demonstrates clear superiority over MERRA-2, whose RMSE values range from $25.00\,\mathrm{W/m^2}$ to $35.80\,\mathrm{W/m^2}$. The advantage of the proposed model becomes even more evident under complex geographical and meteorological conditions. Specifically, at the high-elevation TBL station, the proposed model yields significantly smaller RMSE values, especially when compared to

Table 5
Comparison of reanalysis data, satellite-based retrievals, and SDLR estimations from the proposed surrogate model at the SURFRAD stations under different sky conditions.

Station	Method	Daytime Cloudy		Daytime All-sky		Nighttime All-sky	
		MBE	RMSE	MBE	RMSE	MBE	RMSE
BON	ERA5	−2.32	15.88	−5.22	16.34	−2.70	18.47
	MERRA-2	−19.00	28.39	−23.22	31.14	−12.60	27.57
	CERES	0.19	17.18	−3.31	16.83	−7.73	20.47
	MLP-DNN	−0.57	21.86	−3.03	21.28	3.72	24.16
FPK	ERA5	−6.51	20.13	−10.03	21.55	−7.62	22.16
	MERRA-2	−21.63	32.25	−26.85	35.80	−16.05	31.67
	CERES	8.00	21.39	4.81	20.15	0.44	21.54
	MLP-DNN	8.17	22.11	7.98	23.20	4.95	26.65
GWN	ERA5	−5.17	15.87	−7.23	16.76	−1.26	16.29
	MERRA-2	−19.59	27.60	−22.88	30.01	−11.30	25.00
	CERES	−2.47	15.08	−4.49	15.44	−8.13	17.52
	MLP-DNN	−1.48	23.31	−4.27	22.56	1.02	22.51
PSU	ERA5	−7.04	18.68	−11.21	20.13	−10.07	21.12
	MERRA-2	−19.04	29.13	−23.92	32.59	−15.22	29.41
	CERES	6.97	18.58	3.49	17.07	−4.60	18.50
	MLP-DNN	0.63	22.71	−2.08	21.35	2.88	23.76
SXF	ERA5	−5.84	19.15	−7.99	19.24	−6.98	21.21
	MERRA-2	−16.39	28.63	−19.92	29.19	−9.73	27.74
	CERES	7.31	18.01	4.83	16.90	−2.95	18.53
	MLP-DNN	0.32	23.31	0.66	21.15	2.25	25.92
TBL	ERA5	−25.06	31.39	−27.23	33.27	−27.15	34.26
	MERRA-2	−24.90	33.15	−27.32	34.64	−20.56	30.29
	CERES	−10.99	26.00	−11.70	25.88	−20.88	31.30
	MLP-DNN	15.46	26.40	17.64	28.10	16.50	28.91

reanalysis datasets, highlighting its robustness in handling challenging environments.

More importantly, compared with benchmark products that depend on computationally intensive data assimilation systems or multisensor retrieval frameworks and are released with non-negligible latency, the proposed surrogate offers clear advantages for near-real-time applications. Furthermore, it delivers superior spatiotemporal resolution (e.g., 2 km–5-min), compared with the coarser hourly benchmark datasets, whose spatial resolutions are ~31 km for ERA5, ~62 km for MERRA-2, and ~124 km for CERES.

Overall, these findings demonstrate the reliability of the proposed surrogate model relative to reanalysis and satellite-based datasets, with improved bias characteristics, competitive accuracy, and robust performance. Furthermore, the proposed surrogate model provides a practical alternative for real-time scenarios and operational applications, owing to its evident advantages over existing datasets in spatiotemporal resolution, data latency, and implementation efficiency.

4.5. Retrieval results over a large region

To demonstrate the applicability and transferability of the proposed two-stage surrogate modeling framework, retrieval results over a large spatial domain are presented, as shown in Fig. 12. The two-stage surrogate model is applied to an 11 × 11 pixel satellite image centered at the BON station, supplemented with ground-based measurements of ambient temperature and relative humidity from the central BON station.

The retrieved SDLR fields exhibit pronounced spatial variability that closely follows the cloud distribution patterns, consistent with well-established physical principles. Regions with reduced SDLR are associated with high-altitude cloud cover, which is typically colder and emits less longwave radiation toward the surface compared to lower, warmer clouds [44]. In contrast, areas dominated by low clouds or clear-sky conditions exhibit higher SDLR values. Optical thickness also plays a critical role, as evidenced by the spatial correspondence between low-altitude, optically thick [45] cloud cover and elevated SDLR.

These findings highlight the reliability of the proposed surrogate model, which aligns well with established physical knowledge and provides spatiotemporal SDLR estimates. By leveraging open-source remote

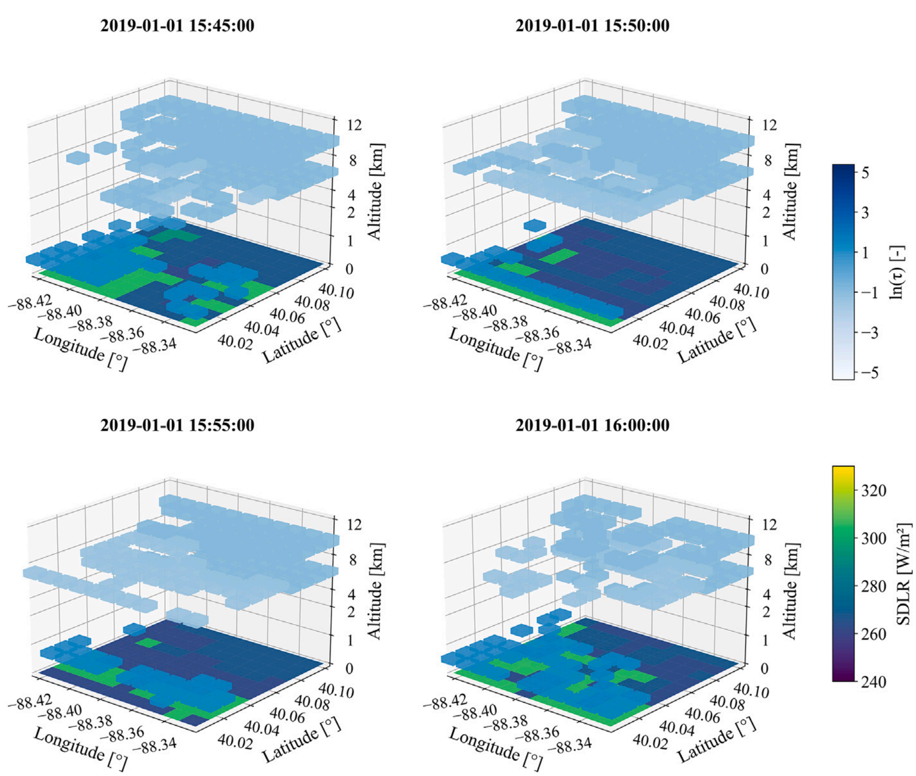

Fig. 12. Spatiotemporal retrieval results of SDLR and corresponding 3D cloud fields over a regional domain. Local time is used for illustration.

sensing data (e.g., GOES [27], Himawari [11]), the surrogate model can produce transferable SDLR estimates with minimal reliance on ground-based data. Furthermore, the retrieval results demonstrate the model's ability to generalize effectively to nearby locations without direct ground-based observations, using only ambient temperature and relative humidity data from the nearest station. This highlights the potential to extend the proposed approach to other regions with sparse observational networks, thereby enabling a broader range of operational applications under diverse meteorological conditions.

5. Conclusions

This study presents a novel two-stage surrogate modeling framework for the efficient retrieval of cloud properties and SDLR from satellite data. The two-stage architecture incorporates cloud optical properties as effective intermediate physical descriptors, resulting in significant improvements over direct end-to-end satellite-to-SDLR mapping. The results reveal substantial reductions in prediction error and bias, with MBE values ranging from 0.10 W/m^2 to 17.62 W/m^2, and RMSE ranging from 20.78 W/m^2 to 30.54 W/m^2. Benchmark comparisons underscore the reliability and robustness of the proposed method compared to empirical, physical, reanalysis, and satellite-based approaches, which demonstrate competitive accuracy with notable bias reduction. Retrieval results over large regions further demonstrate its applicability and transferability for real-time, regional-scale applications, with minimal reliance on ground-sensed observations. This work provides an efficient and reliable framework for estimating cloud properties and SDLR, ensuring consistent retrievals with minimal errors for applications in operational outdoor energy systems.

CRediT authorship contribution statement

Yuying Xie: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation. **Nan Deng:** Writing – original draft, Methodology, Investigation, Conceptualization. **Tao Jing:** Writing – original draft, Validation, Methodology, Investigation, Conceptualization. **Mengying Li:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used GPT-4 to improve language and readability. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the content of the publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors gratefully acknowledge the substantial support from the Research Grants Council of Hong Kong (Grant numbers 25213022 and C6003-22Y), and the Innovation and Technology Commission of Hong Kong (Grant number ITS/188/23).

Data availability

Data will be made available on request.

References

[1] Li M, Jiang Y, Coimbra CF. On the determination of atmospheric longwave irradiance under all-sky conditions. Sol Energy 2017;144:40–8.
[2] Shang W, Lv Z. Low carbon technology for carbon neutrality in sustainable cities: a survey. Sustain Cities Soc 2023;92:104489.
[3] Coimbra CF. Energy meteorology for the evaluation of solar farm thermal impacts on desert habitats. Adv Atmos Sci 2025;42(2):313–26.
[4] Guo Y, Cheng J, Liang S. Comprehensive assessment of parameterization methods for estimating clear-sky surface downward longwave radiation. Theor Appl Climatol 2019;135(3):1045–58.
[5] Yu S, Li L, Cao B, Zhang H, Zhu L, Xin X, Liu Q. Surface downward longwave radiation estimation from new generation geostationary satellite data. Atmos Res 2022;276:106255.
[6] Paletta Q, Hu A, Arbod G, Lasenby J. ECLIPSE: envisioning cloud induced perturbations in solar energy. Appl Energy 2022;326:119924.
[7] Zhou Y, Cess RD. Algorithm development strategies for retrieving the downwelling longwave flux at the earth's surface. J Geophys Res Atmos 2001;106(D12):12477–88.
[8] Wang T, Shi J, Ma Y, Letu H, Li X. All-sky longwave downward radiation from satellite measurements: general parameterizations based on LST, column water vapor and cloud top temperature. ISPRS J Photogramm Remote Sens 2020;161:52–60.
[9] Diak GR, Bland WL, Mecikalski JR, Anderson MC. Satellite-based estimates of longwave radiation for agricultural applications. Agric For Meteorol 2000;103(4):349–55.
[10] McCorkel J, Efremova B, Hair J, Andrade M, Holben B. GOES-16 ABI solar reflective channel validation for earth science application. Remote Sens Environ 2020;237:111438.
[11] Bessho K, Date K, Hayashi M, Ikeda A, Imai T, Inoue H, Kumagai Y, Miyakawa T, Murata H, Ohno T, et al. An introduction to Himawari-8/9—Japan's new-generation geostationary meteorological satellites. J Meteorol Soc Jpn II 2016;94(2):151–83.
[12] Clough SA, Shephard MW, Mlawer EJ, Delamere JS, Iacono MJ, Cady-Pereira K, Boukabara S, Brown PD. Atmospheric radiative transfer modeling: a summary of the AER codes. J Quant Spectrosc Radiat Transf 2005;91(2):233–44.
[13] Larson DP, Li M, Coimbra CF. SCOPE: spectral cloud optical property estimation using real-time GOES-R longwave imagery. J Renew Sustain Energy 2020;12(2).
[14] Li M, Liao Z, Coimbra CF. Spectral model for clear sky atmospheric longwave radiation. J Quant Spectrosc Radiat Transf 2018;209:196–211.
[15] Liou K. An introduction to atmospheric radiation, Vol. 84, Elsevier; 2002.
[16] Alvarado M, Payne V, Mlawer E, Uymin G, Shephard M, Cady-Pereira K, Delamere J, Moncet J. Performance of the line-by-line radiative transfer model (LBLRTM) for temperature, water vapor, and trace gas retrievals: recent updates evaluated with IASI case studies. Atmos Chem Phys 2013;13(14):6687–711.
[17] Pal A, Mahajan S, Norman MR. Using deep neural networks as cost-effective surrogate models for super-parameterized E3SM radiative transfer. Geophys Res Lett 2019;46(11):6069–79.
[18] Goodfellow I. Deep learning; 2016.
[19] Su P, Liu Y, Tarkoma S, Rebeiro-Hargrave A, Petäjä T, Kulmala M, Pellikka P. Retrieval of multiple atmospheric environmental parameters from images with deep learning. IEEE Geosci Remote Sens Lett 2022;19:1–5.
[20] Brence J, Tanevski J, Adams J, Malina E, Džeroski S. Surrogate models of radiative transfer codes for atmospheric trace gas retrievals from satellite observations. Mach Learn 2023;112(4):1337–63.
[21] Ukkonen P. Exploring pathways to more accurate machine learning emulation of atmospheric radiative transfer. J Adv Model Earth Syst 2022;14(4):e2021MS002875.
[22] Li W, Zhang F, Guo B, Fu H, Letu H. Physics-driven machine learning algorithm facilitates multilayer cloud property retrievals from geostationary passive imager measurements. IEEE Trans Geosci Remote Sens 2024;62:1–18.
[23] Lagerquist R, Turner D, Ebert-Uphoff I, Stewart J, Hagerty V. Using deep learning to emulate and accelerate a radiative transfer model. J Atmos Ocean Technol 2021;38(10):1673–96.
[24] Himes MD, Harrington J, Cobb AD, Baydin AG, Soboczenski F, O'Beirne MD, Zorzan S, Wright DC, Scheffer Z, Domagal-Goldman SD, et al. Accurate machine-learning atmospheric retrieval via a neural-network surrogate model for radiative transfer. Planet Sci J 2022;3(4):91.
[25] Gentine P, Pritchard M, Rasp S, Reinaudi G, Yacalis G. Could machine learning break the convection parameterization deadlock? Geophys Res Lett 2018;45(11):5742–51.
[26] Dueben PD, Schultz MG, Chantry M, Gagne DJ, Hall DM, McGovern A. Challenges and benchmark datasets for machine learning in the atmospheric sciences: definition, status, and outlook. Artif Intell Earth Syst 2022;1(3).
[27] Schmit TJ, Gunshor MM. ABI imagery from the GOES-R series. In: The GOES-R series. Elsevier; 2020. p. 23–34.
[28] Matsunobu L, Coimbra CFM. On effective spectral wideband models for clear sky atmospheric emissivity and transmissivity. J Geophys Res Atmos 2024;129(7).
[29] Inman RH, Edson JG, Coimbra CF. Impact of local broadband turbidity estimation on forecasting of clear sky direct normal irradiance. Sol Energy 2015;117:125–38.
[30] Li M, Peterson HB, Coimbra CFM. Radiative cooling resource maps for the contiguous United States. J Renew Sustain Energy 2019;11(3).
[31] Howard L. Essay on the modifications of clouds. Cambridge University Press; 2011.
[32] Natarajan P, Savarimuthu C, Al-Zadjali NSH, Venusamy K. Peak-shadow-distance method for solar energy mapping of Al Hajar mountains using data mining techniques. Mater Today Proc 2023;80:717–23.
[33] Han H, Qiao J. A structure optimisation algorithm for feedforward neural network construction. Neurocomputing 2013;99:347–57.
[34] Windeatt T. Accuracy/diversity and ensemble MLP classifier design. IEEE Trans Neural Netw 2006;17(5):1194–211.
[35] Thomas P, Suhner M-C. A new multilayer perceptron pruning algorithm for classification and regression applications. Neural Process Lett 2015;42(2):437–58.
[36] Zelinka MD, Klein SA, Hartmann DL. Computing and partitioning cloud feedbacks using cloud property histograms. Part II: Attribution to changes in cloud amount, altitude, and optical depth. J Clim 2012;25(11):3736–54.

[37] Cabello-Solorzano K, Ortigosa de Araujo I, Peña M, Correia L, Tallón-Ballesteros AJ. The impact of data normalization on the accuracy of machine learning algorithms: a comparative analysis. In: International conference on soft computing models in industrial and environmental applications. Springer; 2023. p. 344–53.
[38] Joshi S, Owens JA, Shah S, Munasinghe T. Analysis of preprocessing techniques, keras tuner, and transfer learning on cloud street image data. In: 2021 IEEE international conference on big data (big data). IEEE; 2021. p. 4165–8.
[39] Chakraborty D, Elzarka H. Performance testing of energy models: are we using the right statistical metrics? J Build Perform Simul 2018;11(4):433–48.
[40] Store CCD, et al. ERA5 hourly data on single levels from 1940 to present. Datos recuperados entre noviembre. 2024. https://doi.org/10.24381/cds.adbb2d47
[41] Pawson S, et al. MERRA-2 tavg1_2d_rad_Nx: 2D, 1-hourly, time-averaged, single-level, assimilation, radiation diagnostics v5. 12.4. NASA Goddard Earth Sciences Data and Information Services Center (DAAC) data set 2015.
[42] NASA/LARC/SD/ASDC. CERES and GEO-enhanced TOA, within-atmosphere and surface fluxes, clouds and aerosols 1-hourly Terra-Aqua-NOAA20 Edition4B; September 2017. https://doi.org/10.5067/TERRA-AQUA-NOAA20/CERES/SYN1DEG-1HOUR_L3.004B.
[43] Fu Q, Liou K. Lafuli: NASA Langley fu-liou radiative transfer code. Astrophysics Source Code Library (ascl–2104 2021.
[44] Stephens GL. Radiative properties of cirrus clouds in the infrared region. J Atmos Sci 1980;37(2):435–46.
[45] Yang Y, Marshak A, Mao J, Lyapustin A, Herman J. A method of retrieving cloud top height and cloud geometrical thickness with oxygen a and B bands for the deep space climate observatory (DSCOVR) mission: radiative transfer simulations. J Quant Spectrosc Radiat Transf 2013;122:141–9.