https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

Deep Reinforcement Learning - Driven Optimization of Ice Cloud Bulk Scattering Properties in RTTOV-SCATT Using Arctic Weather Satellite Observations

Ke Chen[1, 2], Yuyang Yao[1], Fujia Meng[1]

[1]School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan, 430074, China
[2]Science and Technology on Multispectral Information Processing Laboratory, Wuhan, 430074, China

Correspondence to: Ke Chen (chenke@hust.edu.cn)

Abstract. The Arctic Weather Satellite (AWS), launched by the European Space Agency (ESA) in August 2024, has enabled the first-ever global ice cloud remote sensing using 325 GHz terahertz channels. Owing to its short wavelength, the 325 GHz frequency band is highly sensitive to scattering by the three frozen hydrometeor types present in ice clouds—cloud ice, snow, and graupel. Radiative transfer (RT) models are the core tools for satellite remote sensing retrieval and data assimilation, and their accuracy directly determines the capability to retrieve ice cloud parameters from satellite observations. Due to the complexity of cloud microphysical processes and insufficient observational constraints, the parameterization of the Bulk Scattering Properties (BSPs) of frozen hydrometeors in the current operational RT model RTTOV still harbors considerable uncertainty. The resulting large discrepancies between simulated brightness temperatures (TBs) from high-frequency channels and AWS-observed TBs severely limit the retrieval accuracy of ice cloud parameters.

This paper proposes a physically-informed deep reinforcement learning (DRL) framework for optimizing the computation of frozen hydrometeor BSPs. The framework employs the Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm to couple the RTTOV-SCATT v14.0 module with a differentiable surrogate neural network model, using AWS-observed ice cloud TBs as constraints to optimize the bulk extinction coefficient, scattering coefficient, and asymmetry factor of frozen hydrometeors. Training and validation experiments are conducted using approximately 176,000 AWS-observed TB data records collected since July 2025 (filtered to approximately 80,000 valid ice cloud samples). Results show that after optimization, the root-mean-square error (RMSE) between AWS-observed and simulated ice cloud TBs is reduced to 4.98–14.33 K for the 183 GHz channels and 9.54–13.32 K for the 325 GHz channels, representing an improvement of 56.3–69.7% over the standard RTTOV-SCATT module. The distribution of observation-minus-background (O–B) deviations shifts from markedly skewed to approximately Gaussian. The optimized model significantly improves the simulation accuracy of spaceborne terahertz-band ice cloud TBs while retaining the computational efficiency of RTTOV (approximately 0.2 ms per sample).

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

1 Introduction

Ice clouds are a major component of the Earth's atmosphere, composed of ice crystals of various shapes and sizes (primarily between 100 and 1000 μm) that are widely distributed across the globe in all seasons, particularly in the upper troposphere at 5–13 km in the midlatitudes, with a global coverage of approximately 20% (Liu et al., 2020). Ice clouds cool the Earth system by scattering and absorbing incoming solar shortwave radiation, while simultaneously absorbing longwave radiation emitted by the surface and atmosphere, leading to climate warming. They thus play a critical regulatory role in the Earth's radiative energy balance and the global water cycle (Sunil et al., 2024). Consequently, accurately detecting the mass distribution of ice clouds on a global scale is essential for understanding climate change, improving numerical weather prediction (NWP), and providing early warnings of extreme weather events.

The terahertz band (0.1–10 THz), by virtue of having wavelengths comparable to the sizes of typical ice cloud particles, is highly sensitive to the scattering properties of ice crystals. It can provide more refined information on ice cloud mass, particle size, and cloud altitude than microwave and infrared methods, making it an effective means for detecting the microphysical parameters of ice clouds (Evans and Stephens, 1995; Jiménez et al., 2007). The European Space Agency (ESA) has recently launched the Arctic Weather Satellite (AWS), which has realized the first operational spaceborne application of a 325 GHz terahertz-band channel, filling the gap left by conventional microwave instruments in detecting high-altitude thin ice clouds and frozen hydrometeors (Eriksson et al., 2025). However, the retrieval of ice cloud parameters depends critically on the accuracy of the simulated brightness temperatures (TBs) computed by radiative transfer (RT) models, which establish the link between satellite radiances and atmospheric state variables. In ice cloud remote sensing, the RT process is primarily affected by complex multiple-scattering effects induced by various frozen hydrometeor particles. Accurate computation requires assumptions about the particle size distribution, morphology, and orientation of three types of frozen hydrometeor particles (cloud ice, snow, and graupel), followed by integration over the particle size spectrum to generate Bulk Scattering Properties (BSPs). However, owing to the complexity of ice cloud composition and the consequent lack of accurate ice cloud microphysical parameterization schemes, the computation of ice cloud BSPs in current RT models remains highly uncertain (Baum et al., 2007; Geer et al., 2021; Xie et al., 2020).

Over the past decade, extensive research has been carried out to optimize the parameterization schemes for bulk-scattering properties of ice clouds in radiative transfer models. Sieron et al. (2018) introduced mass-conservation-scaled non-spherical particles, among which sector snowflakes yielded the best performance, to replace the original soft-sphere scattering properties in the CRTM model, which effectively corrected the biases in microwave brightness temperature simulations for Hurricane Karl. Ekelund et al. (2020) employed combined observations from tropical Pacific cloud radars and passive microwave satellites to examine the impacts of different ice-particle shape models on brightness temperature simulations. Mangla et al. (2021) tested 26 sets of scattering parameter schemes for snowflake particles to investigate the effects of snow, ice and other hydrometeor particle shapes on brightness temperature simulations, and quantified the uncertainties in the

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

radiative properties of hydrometeors. Huang et al. (2025) developed a set of physically more realistic scattering datasets for non-spherical ice clouds for the ARMS model, mitigating model errors in ice-cloud radiation simulations.
These studies have advanced RT computations for frozen hydrometeor particles from the assumption of spherical particles to the capability of accurately computing scattering for up to several tens of non-spherical particle types derived from real observations. These advances have been incorporated into the community operational RT model RTTOV through pre-computed non-spherical particle scattering lookup tables (Geer, 2021). Nevertheless, the ice cloud BSPs observed by the AWS satellite represent the integrated scattering contribution from all frozen hydrometeor particles within the observed region, which contains particles of different species, shapes, and sizes. Because the true shape distribution and particle size spectrum of frozen hydrometeors within real ice clouds are not obtainable, current RT models can only assume a single particle shape and select a particular particle size distribution function to compute the BSPs of ice clouds. This "one-shape-fits-all" assumption cannot capture the variability of frozen hydrometeor particle morphology in real ice clouds, leading to substantial discrepancies between the ice cloud TBs simulated by RTTOV and those observed by the AWS satellite in the 183 GHz and 325 GHz frequency bands. This indicates that current operational RT models still exhibit significant deficiencies in computing ice cloud TBs in the terahertz band.
Deep learning offers a data-driven approach to RT computation, with applications broadly falling into two categories: accelerating computation speed and improving computation accuracy. In terms of accelerating computation, Stegmann et al. (2022) proposed a deep learning method that replaces the regression coefficients in fast RT models with a shallow neural network. By training on transmittance profile data generated from line-by-line models, they achieved fast parameterization of atmospheric transmittance in the infrared and microwave bands. In terms of improving accuracy, Li et al. (2025) proposed PhySCAT-Net, a physically-informed deep learning framework that uses a deep neural network to parameterize the lookup tables in RTTOV, constrained by GMI 166.5 GHz dual-polarization observations, optimizing the O–B deviation to within ±1 K.
The aforementioned studies rely on scattering parameters pre-computed by RT models for supervised learning and are thus fundamentally constrained by fixed assumptions regarding particle shape and particle size spectrum. Because the true distribution of particle shapes in real ice clouds is complex and directly observable "ground truth" labels are unavailable, conventional supervised learning is difficult to apply directly. To address this challenge, this paper proposes RTTOV-ScattDRL, a deep reinforcement learning (DRL)-based optimization framework for the computation of frozen hydrometeor BSPs in RTTOV. The framework employs the Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm to couple the RTTOV v14 scattering module with a differentiable surrogate neural network model. Using AWS-observed ice cloud TBs as physical constraints, it enables the dynamic optimization of the extinction coefficient, scattering coefficient, and asymmetry factor of frozen hydrometeor particles without requiring explicit labels. This significantly improves the accuracy of RTTOV in simulating AWS terahertz-band ice cloud TBs and provides a superior forward operator for AWS ice cloud parameter retrieval and all-sky data assimilation.

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

The remainder of this paper is organized as follows. Section 2 introduces the RTTOV model, the atmospheric and ice cloud physical parameters used as RT model inputs, and the AWS satellite TB data. Section 3 describes the DRL framework, including dataset construction, pre-training of the differentiable scattering surrogate model, and the TD3 optimization algorithm. Section 4 presents a comprehensive validation of the proposed RTTOV-ScattDRL model through comparisons with AWS-observed ice cloud TBs at 183 GHz and 325 GHz. Finally, Section 5 summarizes the research findings.

2 Model and Data

The implementation of RTTOV-Scatt-DRL relies on three components: input atmospheric state variable data, the physical RT model RTTOV that calculates atmospheric RT processes, and AWS satellite observed TB data, which serve as constraints for the optimization. This section provides a detailed description of these models and datasets.

2.1 Atmospheric State Variables

The initial atmospheric fields used in this study are derived from ERA5 reanalysis data, provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). Vertically, the data consist of 37 layers, with a horizontal resolution of 0.25° × 0.25°, and include physical parameters required for RT calculations, such as atmospheric temperature, humidity, and pressure. ERA5 is ECMWF's latest reanalysis product, offering hourly atmospheric, land, and oceanic state variables since 1950, along with estimates of uncertainty.

Since the temporal resolution of ERA5 reanalysis data is only one hour, providing data at whole-hour intervals, and satellite observations are precise to the minute, this study uses forecasts from the Weather Research and Forecasting (WRF) model to generate atmospheric state variables that match the AWS satellite overpass times. The whole-hour ERA5 data from 6 hours before the satellite overpass are used as initial and boundary conditions for WRF simulations, producing continuous atmospheric evolution from the initial time to the satellite overpass. The WRF-forecasted atmospheric parameters are then interpolated to the locations of AWS observation pixels via latitude-longitude matching and input into the RTTOV fast RT model to compute TBs.

2.2 Radiative Transfer Model

This study employs RTTOV v14.0 as the baseline RT model. RTTOV is one of the most widely used fast RT models in operational meteorological applications and is capable of efficiently simulating satellite-observed TBs from atmospheric and surface state variables for a wide range of infrared and microwave sounding instruments (Saunders et al., 2018). Its core structure includes multiple modules, such as gas absorption parameterization, surface emission and reflection handling, scattering and absorption calculations for cloud and precipitation particles, and multi-scattering RT solutions. RTTOV requires inputs including layer-by-layer atmospheric profiles of temperature, humidity, and volume mixing ratios of key gases (e.g., ozone), surface temperature and emissivity parameters, as well as the mass content and particle size distribution

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

of hydrometeors such as cloud liquid water, cloud ice, rain, snow, and graupel. Surface emissivity is provided through internal empirical models and adjusted according to the satellite viewing zenith angle and surface type, while gas absorption coefficients are obtained by interpolating precomputed coefficient tables.

For cloud and precipitation scattering, RTTOV treats each hydrometeor type as an ensemble of equivalent particles with specific shapes and dielectric properties. The scattering and extinction coefficients, single-scattering albedo, and phase function parameters for each layer are obtained by referencing a predefined single-scattering property database. BSPs can be derived by numerically integrating the scattering properties of individual particles over the particle size distribution. Specifically, the bulk extinction coefficient β_e can be calculated by integrating the extinction cross-section of a single particle $\sigma_e(D_g)$ （m^2） according to the following expression:

$$\beta_e = \int_{D_{\min}}^{D_{\max}} \sigma_e\,(D_g)\,N(D_g)\,dD_g, \tag{1}$$

The integration range is from $D_{\min}$ to $D_{\max}$. $\mathrm{D_g}$ denotes the geometric diameter. $\sigma_e(D_g)$, the single-particle extinction cross-section, represents the extinction capability of a single particle with diameter D_g to incident radiation. $N(D_g)$ is the particle number density distribution function, indicating the number of particles with diameter D_g per unit volume. $D_{\min}$ and $D_{\max}$ are the lower and upper limit diameters of the integral. For fast RT models used in both active and passive remote sensing such as RTTOV-SCATT, the volume scattering coefficient β_s and asymmetry factor g are also required, which are likewise obtained by integrating the single-particle scattering cross-section $\sigma_s(D_g)$ and the dimensionless single-particle asymmetry factor $\mathrm{g_{single}}(D_g)$:

$$\beta_s = \int_{D_{\min}}^{D_{\max}} \sigma_s\,(D_g)\,N(D_g)\,dD_g, \tag{2}$$

$$\mathrm{g} = \frac{\int_{D_{\min}}^{D_{\max}} g_{\mathrm{single}}\,(D_g)\,\sigma_s(D_g)\,N(D_g)\,dD_g}{\beta_s}, \tag{3}$$

By repeatedly calculating the above integrals, a lookup table can be generated as a function of temperature, water content and channel. The lookup table is output in the form of a data file ("hydrotables"), with one file corresponding to each target instrument. It contains the following bulk optical properties: bulk extinction coefficient β_e、 dimensionless bulk single scattering albedo SSA, and dimensionless bulk asymmetry factor g (referred to as asym in the RTTOV14 User Manual). The single scattering albedo SSA is defined as the scattering coefficient divided by the extinction coefficient, with a value ranging from 0 to 1.

$$\mathrm{SSA} = \frac{\beta_s}{\beta_e}, \tag{4}$$

Subsequently, RTTOV numerically solves the RT equation using a simplified multi-stream approximation, propagating the radiation intensity layer by layer. Thanks to the aforementioned parameterizations and precomputed strategies, RTTOV greatly improves computational efficiency while maintaining a reasonable level of physical accuracy, making it well suited for simulations and assimilation applications in large-sample scenarios.

Under clear-sky conditions, the core formula used by RTTOV to compute radiative TB is:

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

$$TB^{Clear}(v,\theta) = \tau_s(v,\theta)\varepsilon_s(v,\theta)B(v,T_s) + \int_{\tau_s}^{1} B(v,T)d\tau + (1-\varepsilon_s(v,\theta))\tau_s^2(v,\theta)\int_{\tau_s}^{1}\frac{B(v,T)}{\tau^2}d\tau, \quad (5)$$

In the formula, TB represents brightness temperature, τ_s denotes the transmittance from the surface to the top of the atmosphere, ε_s stands for surface emissivity, v represents frequency, θ denotes the zenith angle, and $B(v,T)$ is the Planck function.

It can be seen that the calculation is divided into three components: the first corresponds to direct upward radiation from the surface, the second to direct upward radiation from the atmosphere, and the third to downward atmospheric radiation reflected by the surface and then propagating upward.

Surface emissivity is calculated using the FASTEM model. The computation of optical depth relies on the optical depth coefficient files provided in the official RTTOV design. Optical depth is computed layer by layer, with the optical depth from each layer to space calculated using the following formula:

$$d_{i,j} = d_{i,j-1} + \sum_{k=1}^{K} a_{i,j,k}X_{k,j}, \quad (6)$$

$d_{i,j}$ represents the optical thickness from level j to space at level i, $a_{i,j,k}$ denotes the coefficient in the coefficient file, and $X_{k,j}$ stands for the predictor calculated by RTTOV based on input variables such as temperature and humidity.

Under cloudy conditions, the formula used by RTTOV to compute TB is:

$$T_B^{\text{Total}} = (1-C)T_B^{\text{Clear}} + CT_B^{\text{Rainy}}, \quad (7)$$

In the formula, T_B^{Total} represents the total brightness temperature, T_B^{Clear} represents the clear-sky brightness temperature, T_B^{Rainy} represents the cloud and rain scattering brightness temperature, and C denotes the effective cloud fraction, which is calculated by RTTOV based on input parameters and coefficient files. It can be seen that the final TB is a linear weighting of the clear-sky TB and the cloud TB. The calculation formula for the T_B^{Rainy} scattering TB is as follows:

$$T_B^{Rainy} = \epsilon_{(\theta,v)}T_s\Gamma_{(\theta,v)} + (1-\epsilon_{(\theta,v)})T_{(\theta,v)}^{\downarrow}\Gamma_{(\theta,v)} + T_{(\theta,v)}^{\uparrow}, \quad (8)$$

T_s represents the surface temperature, $T_{(\theta,v)}^{\downarrow}$ denotes atmospheric downward radiation, and $T_{(\theta,v)}^{\uparrow}$ stands for atmospheric upward radiation. $\Gamma_{(\theta,v)}$ indicates cloud transmittance, $\epsilon_{(\theta,v)}$ represents surface emissivity, v denotes frequency, θ signifies zenith angle, and $B(v,T)$ refers to the Planck function.

The RT equation describes how the electromagnetic energy emitted by the surface and lower atmosphere is absorbed by atmospheric gas molecules and altered by scattering from cloud particles. By solving the atmospheric RT equation, the total upward radiative energy at the top of the atmosphere can be calculated, typically expressed in terms of radiance or brightness temperature. This represents the fundamental function of an atmospheric RT model. In this study, the latest version, RTTOV14, is employed, which integrates RTTOV-SCATT, the scattering module, directly into the RTTOV computation.

The RTTOV model parameterizes the calculation of atmospheric transmittance as well as the scattering properties of ice-phase and liquid-phase hydrometeors. In this study, the AWS satellite coefficient files rtcoef_aws_1_aws_o3.dat and rttov_hydrotable_aws_aws.dat (https://nwp-saf.eumetsat.int/site/software/rttov/download/coefficients/coefficient-download/),

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

provided on the official RTTOV website, were employed to simulate AWS TBs.The file rtcoef_aws_1_aws_o3.dat is a clear-sky (ozone-included) RT coefficient file that stores regression coefficients for gaseous absorption, optical depth, and transmittance associated with temperature, water vapour, and ozone for all AWS channels. The file rttov_hydrotable_aws_aws.dat is an AWS-specific lookup table containing the optical properties of hydrometeors. Together, these two coefficient files enable the rapid calculation of optical depths for different frequency channels at each atmospheric pressure level, thereby allowing efficient atmospheric RT simulations.

2.3 AWS Level-1B brightness temperature product

The Arctic Weather Satellite (AWS) is a new polar-orbiting meteorological operational satellite developed by the European Space Agency (ESA). It serves as a prototype demonstration satellite for the EUMETSAT Polar System – STERNA, a polar constellation of the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT). The satellite was lauched in August 2024, aiming to fill the observational gap in high-frequency measurements of temperature, humidity, clouds, and precipitation over polar regions. The AWS payload consists of a 19-channel cross-track microwave radiometer, covering four frequency bands: 54 GHz, 89 GHz, 174 GHz, and 325 GHz. A photograph of the payload is shown in Figure 1, and the configuration of each channel group is summarized in Table 1.

Figure 1 Arctic Weather Satellite instrument (Eriksson et al., 2025)

Table 1. Arctic Weather Satellite Instrument Characteristics

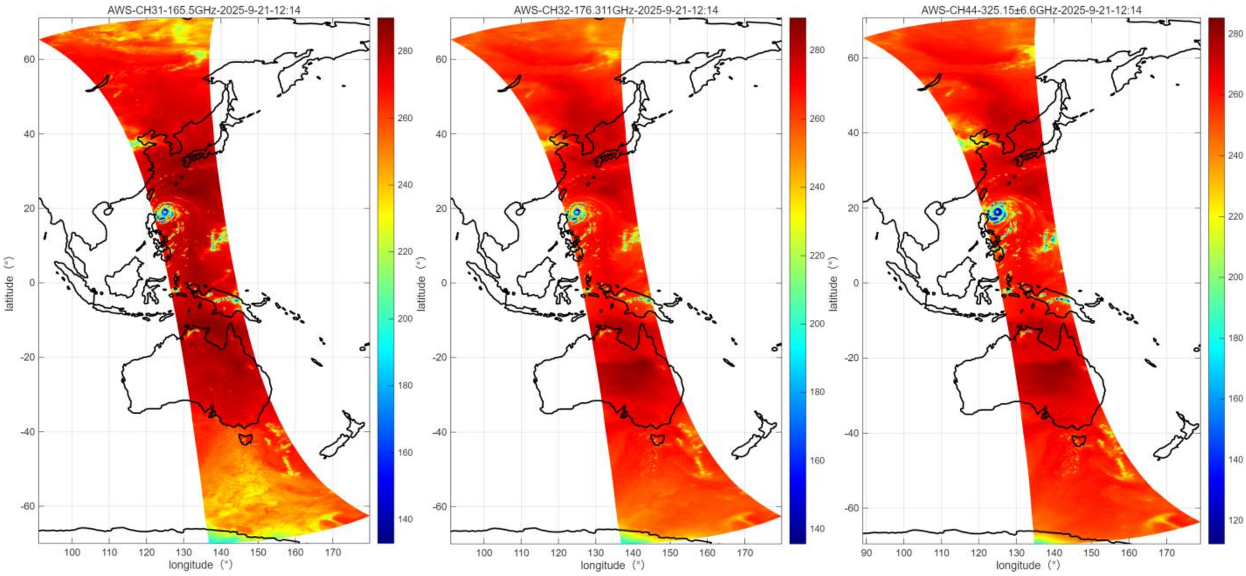
Channel	Frequency (GHz)	Bandwidth (MHz)	NEΔT (K)	Primary Use
11	50.3	180	<0.6	Temperature sounding
12	52.8	400	<0.4	Temperature sounding
13	53.246	300	<0.4	Temperature sounding
14	53.596	370	<0.4	Temperature sounding

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

15	54.4	400	<0.4	Temperature sounding
16	54.94	400	<0.4	Temperature sounding
17	55.5	330	<0.5	Temperature sounding
18	57.290344	330	<0.6	Temperature sounding
21	89	4000	<0.3	Window / Cloud detection
31	165.5	2800	<0.6	Window / Humidity sounding
32	176.311	2000	<0.7	Humidity sounding
33	178.811	2000	<0.7	Humidity sounding
34	180.311	1000	<1	Humidity sounding
35	181.511	1000	<1	Humidity sounding
36	182.311	500	<1.3	Humidity sounding
41	325.15±1.2	800	<1.7	Humidity sounding / Cloud detection
42	325.15±2.4	1200	<1.4	Humidity sounding / Cloud detection
43	325.15±4.1	1800	<1.2	Humidity sounding / Cloud detection
44	325.15±6.6	2800	<1	Humidity sounding / Cloud detection

Prior to launch, AWS underwent comprehensive ground calibration tests, and its performance was further validated during the in-orbit commissioning phase. The noise-equivalent temperature difference (NEΔT) of all channels meets design specifications with margin, and the in-orbit measured NEΔT values are consistent with ground test results, demonstrating stable receiver performance in orbit. The measured spectral response of each channel shows deviations from the target design within acceptable limits. Assimilation-based evaluations by ECMWF using numerical forecast models further confirmed the data quality of AWS (Duncan et al, 2025). Studies indicate that the bias and noise levels of AWS are comparable to those of the Advanced Microwave Sounding Unit-A (AMSU-A) and the Microwave Humidity Sounder (MHS), meeting the requirements for operational numerical weather prediction assimilation. An example observation (AWS1-MWR-1B-RAD_C_OHB__20250921135758_G_O_20250921121433_20250921135050_C_N.nc) is shown in Figure 2. The region affected by Super Typhoon Ragasa can be clearly observed. The in-orbit validation by ESA and ECMWF demonstrates that AWS data are reliable, providing a credible observational benchmark for the simulations in this study. In addition, AWS's unique 325 GHz channel is highly sensitive to ice crystals, offering, for the first time, a satellite-based reference for evaluating the capability of atmospheric RT models to simulate ice-cloud scattering in the terahertz frequency range.

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026



(a) **(b)** **(c)**

Figure 2. Observed TBs from the Arctic Weather Satellite payload: (a) Channel 31 (165.5 GHz); (b) Channel 32 (176.311 GHz); (c) Channel 44 (325.15 ± 6.6 GHz).

3 Deep Reinforcement Learning Framework

3.1 Deep Reinforcement Learning Theory

Reinforcement Learning (RL) is an important branch of machine learning. Its fundamental idea is that an agent interacts continuously with an environment and, through a trial-and-error process, gradually adjusts its actions to learn an optimal decision-making policy. Deep Reinforcement Learning (DRL) combines the feature extraction capability of deep learning with the decision-making optimization ability of reinforcement learning, approximating value functions or policy functions through neural networks. This overcomes the limitations of traditional RL in directly handling high-dimensional state and action spaces.

The basic framework of DRL consists of two components: the agent and the environment. Learning is achieved through a closed-loop feedback of "state – action – reward." The agent is responsible for decision-making and learning, and its structural design must be tailored to the specific problem. This includes the choice of neural network type (e.g., fully connected networks, convolutional neural networks), the configuration of input and output dimensions (inputs corresponding to state feature dimensions, outputs corresponding to action dimensions or value estimations), and the depth of the network and the number of neurons in each layer. These factors directly affect the algorithm's approximation ability and convergence characteristics. The environment provides the external conditions for the agent's interaction, and its proper abstraction and modeling are prerequisites for the effective operation of the algorithm.

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

In complex optimization problems with continuous action spaces, conventional DRL algorithms such as Deep Deterministic Policy Gradient (DDPG) often suffer from issues like overestimation of the value function and sensitivity of the training process to hyperparameters and noise, which can affect the accuracy and stability of policy optimization. To address these shortcomings, Fujimoto et al. (2018) proposed the Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm. Building upon the DDPG framework, TD3 introduces twin critic networks, target policy smoothing regularization, and delayed policy updates, which together help mitigate value overestimation and improve training stability and convergence accuracy in continuous control tasks.

For fine-tuning problems such as ice cloud scattering parameter optimization, which involve high-dimensional and continuous spaces, TD3 is theoretically and practically more suitable. Therefore, this study adopts TD3 as the core algorithm for optimizing ice cloud scattering characteristics. The main features of TD3 include the use of two critic (Q) networks, taking the minimum of the two as the target value during updates to reduce overestimation, target policy smoothing regularization where noise is added to the action of the next state to make value estimation smoother and more accurate, and delayed updates, where the actor network is updated only after multiple critic network updates to ensure more stable training of the actor network.

The core of TD3 lies in updating the Critic networks, and the improvement of this algorithm is the use of twin Critic networks. The action under state $s^{'}$ is calculated using the Target Actor network.

$$a^{'} = \mu^{'}\left(s^{'}|\theta^{\mu^{'}}\right), \tag{9}$$

Add noise to the target action $a^{'}$ based on target policy smoothing regularization:

$$a^{'} = a^{'} + \epsilon, \tag{10}$$

Based on the idea of the dual network, take the smaller one of the Q-values from the dual network to solve the problem of overestimating Q-values in traditional reinforcement learning algorithms, and calculate the target value:

$$y = r + \gamma min_{i=1,2} Q_i^{'}\left(s^{'}, a^{'}|\theta_i^{Q^{'}}\right), \tag{11}$$

The TD3 algorithm has achieved excellent performance across multiple continuous control tasks. Compared to the DDPG algorithm, TD3 is more stable and can better handle function approximation errors.

In this experiment, the algorithm's simulation performance is evaluated using metrics such as the root mean square error (RMSE), bias, and correlation coefficient. The definitions of RMSE and the correlation coefficient are as follows:

$$Bias = \frac{\sum_{i=1}^{n}(T_i - R_i)}{N}, \tag{12}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(T_i - R_i)2}{N}}, \tag{13}$$

$$Correlation = 1 - \frac{\sum_{i=1}^{n}(T_i - R_i)^2}{\sum_{i=1}^{n}(T_i - \overline{T})^2}, \tag{14}$$

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

Among them, T_i is the true value, R_i is the model simulated value, and $\overline{T}$ is the average of true values. The smaller the RMSE, the better the model fitting effect. The closer the correlation coefficient is to 1, the better the model simulation performance.

3.2 Dataset Construction

The construction of the dataset forms the foundation of this optimization scheme, directly influencing the model's ability to learn the scattering properties of real ice clouds. This study focuses on ice cloud TB simulations in the terahertz frequency range, using AWS satellite observations at 183 GHz and 325 GHz channels as the optimization targets. Multiple typical typhoon cases since July 2025 were selected, including Super Typhoon Ragasa, Typhoon Bualoi, and Severe Typhoon Matmo, encompassing typhoons, strong typhoons, and super typhoon events. These cases span different intensities, developmental stages, and tracks, ensuring that the dataset is both representative and diverse (including strong convective ice clouds, mixed-phase clouds, and high-altitude cirrus). AWS observed TB data are sourced from the L1B TB data products provided on the AWS data portal (https://data.eumetsat.int/data/map/EO:EUM:DAT:0905).
The atmospheric profiles were derived from the ERA5 reanalysis dataset, with a spatial resolution of 0.25°×0.25°, a temporal resolution of 1 hour, and 37 vertical levels. These data were used to drive the WRF model to generate high-resolution atmospheric and ice-cloud parameter profiles. The main variables include temperature, water vapor density, pressure, and graupel content. The graupel content was further optimized through inversion of AWS satellite-observed TBs using a UNet neural network, thereby improving the realism and consistency of the microphysical parameters. All profiles were interpolated onto 59 vertical levels to ensure compatibility with the RT model input requirements. The target outputs are the scattering property parameters of ice-cloud particles in RTTOV v14, including the extinction coefficient, scattering coefficient, and asymmetry factor. The extinction coefficient represents the total attenuation rate of electromagnetic wave energy caused jointly by particle absorption and scattering. The scattering coefficient characterizes the ability of particles to scatter incident electromagnetic wave energy into the surrounding space. It must satisfy the physical constraint that the ratio of the scattering coefficient to the extinction coefficient equals the single-scattering albedo, whose value ranges from 0 to 1, ensuring that the scattered energy does not exceed the total extinction. The asymmetry factor is used to quantify the directional preference of scattering (i.e., the relative strength of forward versus backward scattering), with a value range of [-1, 1], where 1 corresponds to complete forward scattering and -1 corresponds to complete backward scattering. This parameter represents the first angular moment of the scattering phase function and can be calculated by integrating the phase function.
The original dataset contains multiple typhoon scenarios, comprising a total of 175,740 training profiles with 59 vertical levels per profile. To focus on ice-cloud and mixed-phase cloud conditions and to avoid over-fitting caused by the dominance of clear-sky samples, a dynamic threshold scheme based on 183 GHz observed TBs, as proposed by Gang (2005), was adopted to screen cloud-region samples. After filtering, the dataset retained 80,741 valid training samples and 13,053 testing samples. To meet the requirements of neural network training, both input and output parameters were processed using

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

layer-wise logarithmic normalization. Specifically, independent normalizers were fitted for each vertical level (59 levels in total): the variables were first transformed into logarithmic space and then standardized using Z-score normalization. This approach preserves the vertical structural characteristics of the physical variables while enabling subsequent recovery of physical constraints through inverse normalization. The dataset covers multiple typhoon events from different sources, thereby ensuring strong generalization capability. In addition, the inclusion of effective radius information enhances the representation of particle microphysical properties.

3.3 Pre-trained Differentiable Scattering Surrogate Model

The purpose of pre-training a differentiable scattering surrogate model is twofold. First, it enables differentiability, facilitating gradient propagation during RL optimization. Second, directly inputting the generated scattering parameters into RTTOV would require invoking RTTOV for each sample in every training iteration. Although RTTOV itself is computationally efficient, RL optimization typically involves hundreds of thousands of samples and 300 – 1000 training iterations, resulting in substantial computational cost. Employing a pre-trained neural network surrogate can therefore accelerate computation and save training time.

The RTTOV-simulated TB consists of the clear-sky TB, which accounts only for atmospheric absorption, and the scattering TB computed by the RTTOV-SCATT module, which includes both liquid- and ice-phase hydrometeors. It is generally accepted that RTTOV provides high-accuracy atmospheric absorption calculations, even at terahertz frequencies such as 183 GHz and 325 GHz. This conclusion is further validated in Section 4.1 by comparing AWS observations in clear-sky regions with RTTOV-simulated TB. Therefore, the primary objective of this study is to optimize the BSPs of ice clouds computed by the RTTOV-SCATT module. We first pre-trained a neural network to map scattering property parameters to scattering TB based on RTTOV-SCATT outputs. Since RTTOV does not output the scattering TB separately, the scattering TB in the surrogate model was obtained as the difference between the actual ice-cloud TB and the RTTOV clear-sky TB. The surrogate model was trained with the objective of minimizing the mean squared error (MSE) between the surrogate's predicted scattering contribution and the actual scattering contribution.

Figure 3 illustrates the structure of the surrogate model, where the inputs are the optimized scattering parameters output by the Actor network, and the output is the scattering contribution, ΔTB. A non-negative constraint is applied through the activation function to ensure physical plausibility. Figure 4 shows that the loss on both the training and validation sets rapidly decreases during training, converging from 2250 to approximately 260.

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

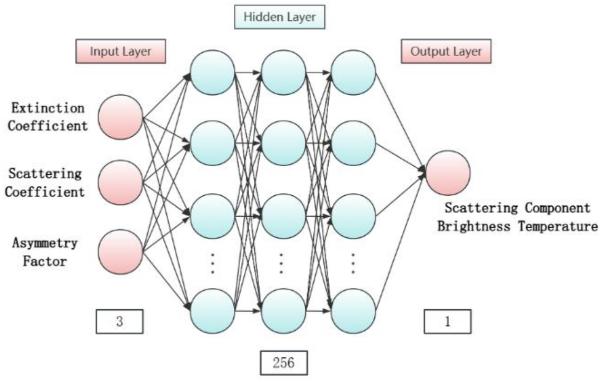


Figure 3. Scattering Surrogate Model

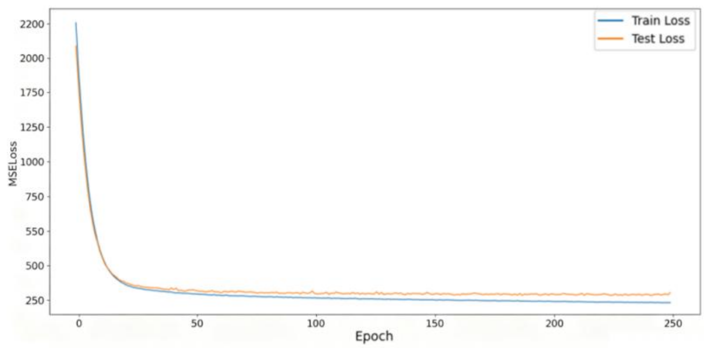


Figure 4. Training Loss Curve

3.4 Implementation of Deep Reinforcement Learning Algorithm

The simulation workflow of the standard RTTOV model is as follows: upon inputting the instrument configuration, atmospheric profiles, and viewing geometry parameters, the model first computes gas absorption to obtain the clear-sky optical depth and transmittance. Subsequently, surface emissivity is calculated based on the surface type, and hydrometeor scattering parameters (extinction coefficient, single-scattering albedo, and asymmetry factor) are loaded. Finally, a RT solver is executed to output the simulated TB. In this study, the pivotal enhancement lies in the integration of a DRL module. The Actor network takes atmospheric profiles (temperature, humidity, pressure, graupel content, and effective radius) as inputs and adaptively outputs optimized scattering parameters for ice-phase particles in real time. These parameters are directly injected into the RTTOV v14 scattering calculation via the esba interface, overriding the standard ice-phase scattering contributions to yield simulated TB that closely match actual AWS observations. This hybrid framework not only preserves the highly efficient physical computational capability of RTTOV but also achieves a data-driven optimization of the complex bulk scattering properties of ice clouds through DRL. A comparison between the proposed DRL-based RTTOV-SCATT-DRL architecture for ice cloud bulk scattering parameter optimization and the standard RTTOV-SCATT architecture is illustrated in Figure 5.

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

To identify the target for optimization, the sensitivity of TB at 183 GHz and 325 GHz to different species of hydrometeors was first analysed. Figure 6 shows the TB variations obtained by successively including the scattering effects of cloud liquid water, rain, ice, snow, and graupel particles. The results indicate that cloud liquid water and rain have a negligible impact on terahertz-band TB and can therefore be ignored, while ice and snow particles produce only small TB depressions. In contrast, graupel particles cause the largest TB depressions at both 183 GHz and 325 GHz and are the dominant hydrometeor type affecting ice-cloud scattering signals and the resulting TB variations. Therefore, considering that simultaneous optimization of the scattering parameters for multiple species of hydrometeors would substantially increase the dimensionality of the parameter space and the complexity of the optimization process, and given the dominant contribution of graupel to TB variations, this study focuses exclusively on optimizing the scattering parameters of graupel particles to effectively reduce ice-cloud TB simulation errors.

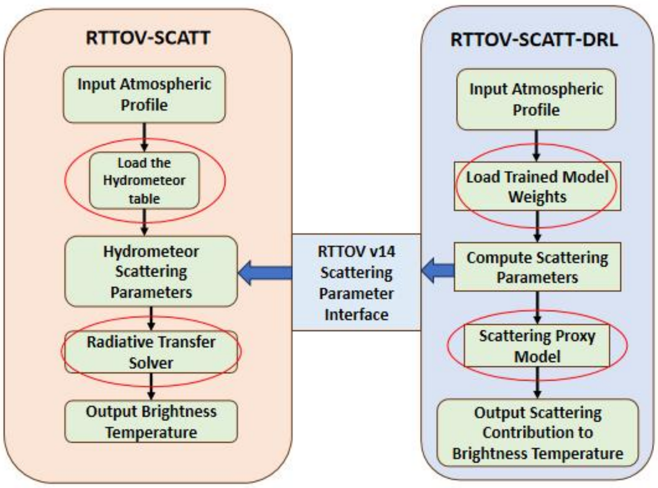


Figure 5. RTTOV Ice Cloud Scattering Parameter Optimization Architecture

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

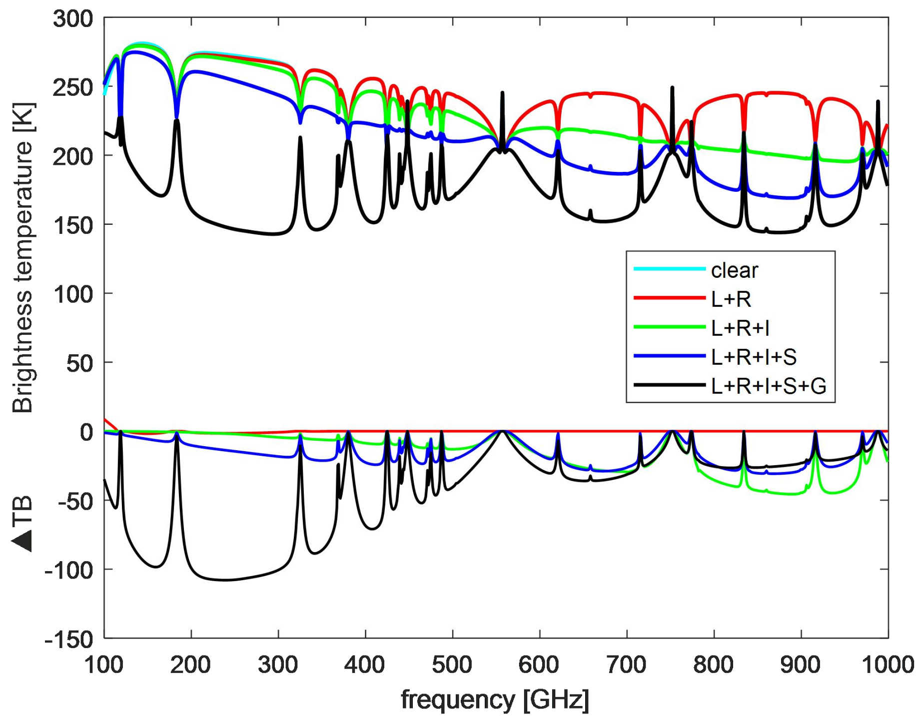


Figure 6 Simulated TBs for Different Hydrometeor Species from 100 to 1000 GHz

Within the DRL framework, this study models the problem as a collaborative decision-making process between an agent and an environment. The agent is responsible for selecting a set of scattering parameters to be optimized based on the current atmospheric state, while the environment consists of the RT computation process, which generates simulated TBs based on the selected parameters and provides feedback signals. The agent's structure and training details are designed to match the dimensionality of the scattering parameterization and the available observable information. Therefore, targeted designs were applied regarding network input/output dimensions, layer size, and constraint mechanisms. The algorithm implementation is divided into four components: environment modeling, network architecture, physical constraints, and update mechanism.

Figure 7 illustrates the environment modeling in the DRL algorithm. In the TD3 algorithm, the state (s) is characterized by the atmospheric profile and the effective radius, and the action (a) represents the adjustment to the scattering parameters, with the specific optimization targets being the normalized extinction (ext), scattering (sca), and asymmetry parameter (asym). The reward function is defined as a negative error related to observational consistency, specifically −RMSE. The simulated TB is obtained from "clear-sky + predicted scattering contribution" and compared to the observed TB to compute the RMSE. This design ensures that higher rewards correspond to parameters that more closely match the observations, directly aligning the optimization objective with the radiative consistency of interest in subsequent ice cloud retrieval application.

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

Figure 8a shows the Actor network constructed in this study. The network input is the environment state, and the output is the action. Figure 8b shows the Critic network, which takes both the state and the action from the Actor network as input and outputs the predicted reward. The Actor network maps states to the action space. It adopts a fully connected three-layer structure, with LayerNorm and ReLU applied after each layer. The output layer uses a Tanh activation to produce normalized values in the range [−1,1], which are then denormalized to physical scales usable by RTTOV. The Critic network uses a twin Critic structure to reduce overestimation bias. Each Critic employs 1D convolutional layers for feature extraction, with ReLU activation and Dropout to improve generalization. To reflect differences in physical sensitivity, the Critic evaluation only covers the vertical layers (approximately 200–450 hPa) where TB is most sensitive to scattering parameters, reducing ineffective gradient interference from less relevant layers and improving training efficiency and convergence stability.

To ensure physically feasible action outputs, explicit constraints are applied during the action execution phase. The scattering coefficient divided by the extinction coefficient is restricted to the range [0,1], and the asymmetry factor is limited to [0,1]. Considering the nonlinearity and multi-solution nature of the scattering parameter space, an adaptive noise mechanism is employed. An initial noise scale is set and gradually reduced with a decay rate of 0.99 as training progresses, allowing the algorithm to explore extensively during early stages and stabilize for fine optimization later.

Regarding the update mechanism, TD3 minimizes bias in the target Q values using twin Critic networks and updates the Actor network at a set frequency (once every two iterations), while applying soft updates to the target networks. During training, a batch size of 256 is used, with 800 epochs. The Adam optimizer is employed with a learning rate of 10^{-4}. Because the reward exhibits large fluctuations during early training but generally decreases, the RMSE on the test set is used as a monitoring metric, and the current optimal Actor and Critic models are saved during updates to ensure that the final output corresponds to the best generalization performance.

In summary, the TD3 framework achieves dynamic optimization of RTTOV scattering parameters through a combination of strategies: pre-training provides stable gradients, twin Critic networks enhance training robustness, explicit physical constraints ensure physically plausible outputs, and prioritizing sensitive layers reduces ineffective learning. This approach improves the consistency between simulated brightness temperatures and observations while maintaining physical feasibility. By integrating pre-training for stable gradients, TD3 for stable optimization, and physical constraints for output reliability, the framework enables dynamic optimization of RTTOV scattering parameters.

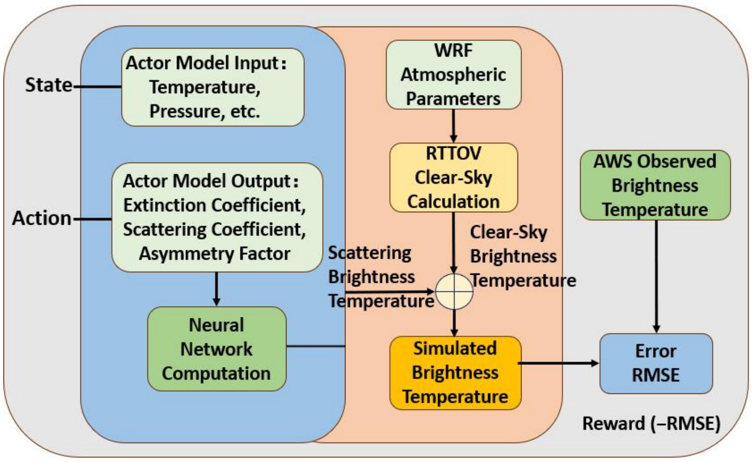
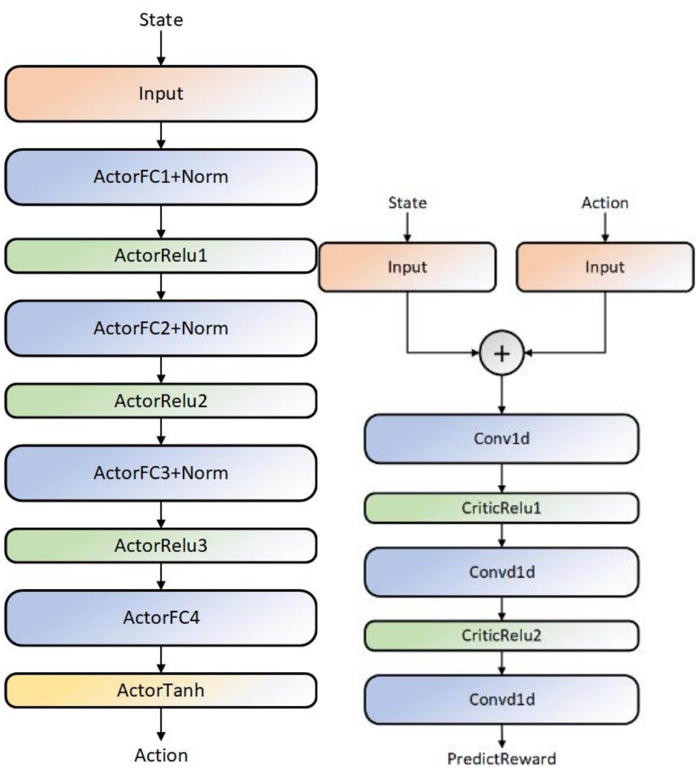


Figure 7 Environment Flowchart in the Deep Reinforcement Learning Algorithm



405 **(a)** **(b)**

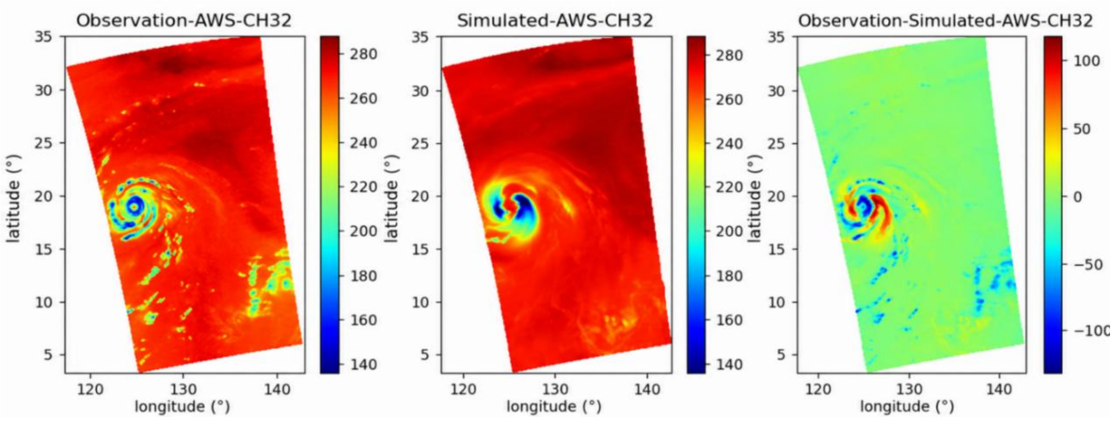
Figure 8. Actor and Critic Networks in the Reinforcement Learning Model of this Study: (a) Actor Network; (b) Critic Network

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

4 Validation Experiment

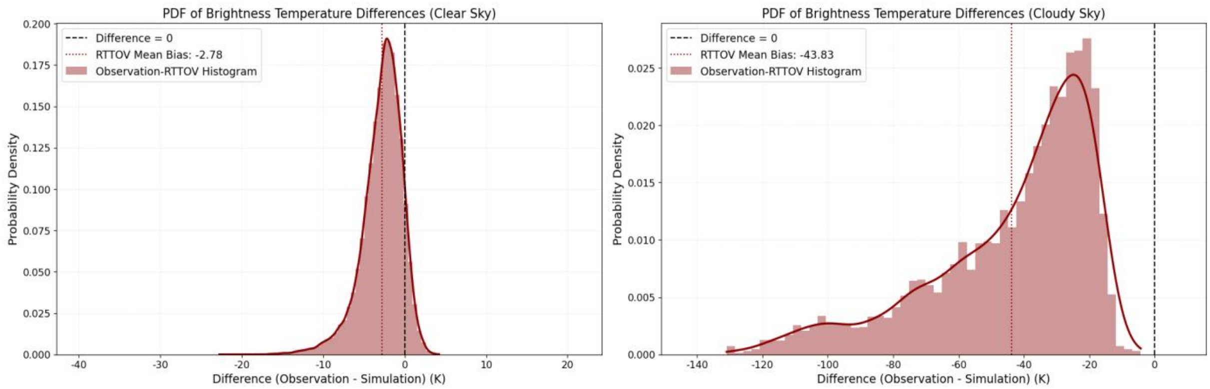
4.1 AWS Brightness Temperature Simulation Based on RTTOV

This section first evaluates the accuracy of RTTOV-simulated TBs against AWS observations in the region affected by Typhoon Haguibis, using the RTTOV configuration described in Section 2.2. All experiments employed the same WRF atmospheric profiles and surface parameters. The differences between simulated and observed TBs were analysed separately for clear-sky and cloudy samples. Cloud detection was performed using the dynamic threshold method based on 183 GHz observed TBs proposed by Gang (2005). Figures 9 and 10 present comparisons between RTTOV-simulated and observed TBs for the AWS CH32 channel (176.311 GHz) and CH44 channel (325.15 ± 6.6 GHz), respectively. The TB difference distributions shown in panel (c) indicate that the errors within typhoon ice-cloud regions are substantially larger than those under clear-sky conditions. As shown in panel (d), the Probability Density Function (PDF) of the differences between RTTOV-simulated and observed TBs in clear-sky regions for both the 183 GHz and 325 GHz channels generally follow Gaussian distributions, which is particularly important for data assimilation applications. The bias and RMSE values in clear-sky regions are also relatively small, demonstrating that RTTOV achieves good accuracy in calculating atmospheric absorption at these two frequency bands. However, in panel (e), both the bias and RMSE increase significantly in ice-cloud regions, and the TB difference distributions deviate markedly from Gaussian behaviour. This indicates that the current RTTOV simulations are still insufficiently accurate for AWS ice-cloud TBs, making them difficult to satisfy the requirements of retrieval and data assimilation applications. Table 2 lists the statistical metrics of the differences between simulated and observed TBs for the five AWS 183 GHz channels and four 325 GHz channels, separately for clear-sky and ice-cloud regions. The results are generally consistent with those shown in Figures 9 and Figures 10.



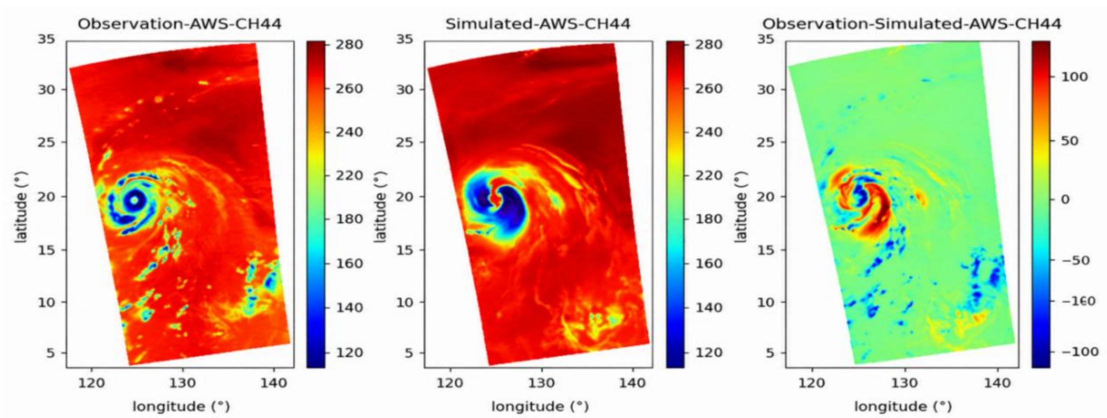
(a) (b) (c)

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

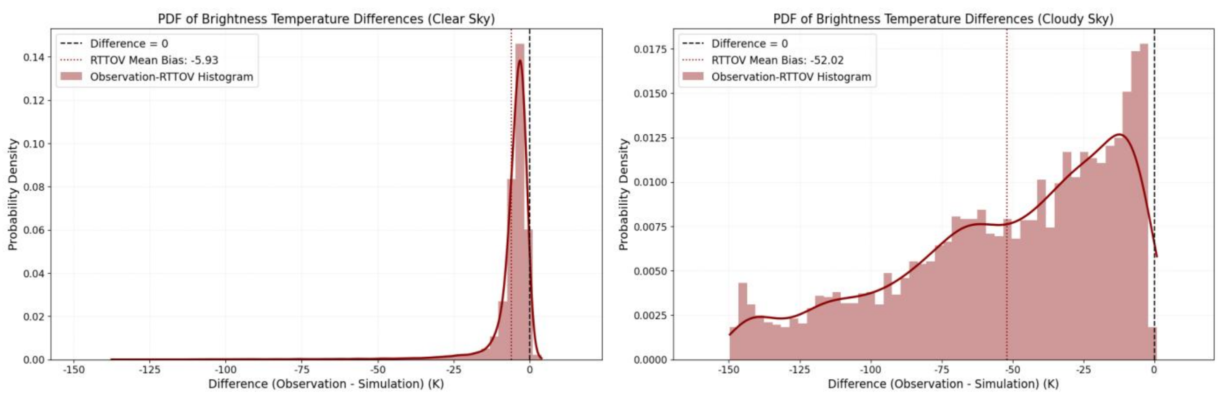


(d) (e)

Figure 9. RTTOV model simulation results for Channel 32 (176.311 GHz, RMSE: 15.36 K): (a) AWS-observed TBs, (b) simulated TBs, (c) differences between simulated and observed TBs, (d) PDF of TB differences under clear-sky conditions, (e) PDF of TB differences under cloudy-sky conditions.



(a) (b) (c)



(d) (e)

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

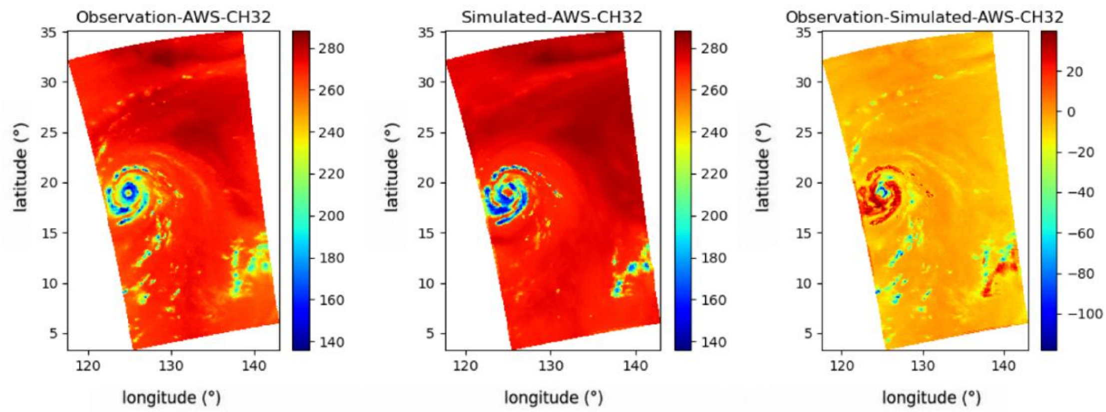
Figure 10. RTTOV model simulation results for Channel 44 (325.15±6.6 GHz, RMSE: 23.97 K): (a) AWS-observed TBs, (b) simulated TBs, (c) differences between simulated and observed TBs, (d) PDF of TB differences under clear-sky conditions, (e) PDF of TB differences under cloudy-sky conditions.

Table 2. RMSE for AWS the 183 GHz and 325 GHz channels

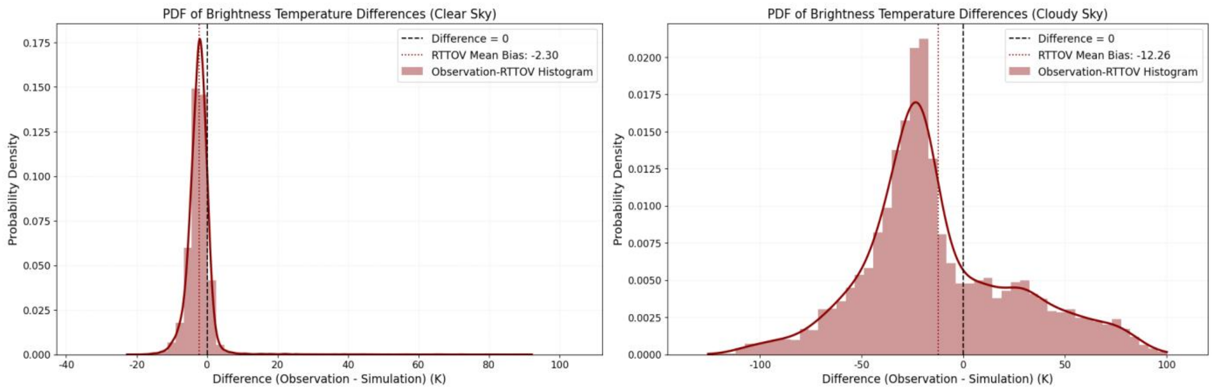
Channel	32	33	34	35	36	41	42	43	44
CLEAR SKY	4.99	4.76	4.45	3.81	4.46	5.98	7.44	8.43	10.80
CLOUD SKY	46.85	41.39	35.80	28.90	24.70	46.61	52.95	54.96	58.56

As shown in Figure 9 (c) and Figure 10 (c), one major source of the large TB simulation errors in ice-cloud regions is the spatial mismatch between the ice-cloud locations predicted by the WRF model and those observed by the AWS satellite. This mismatch introduces a "double penalty" effect in the differences between observed and simulated TBs ("double penalty" effect; Geer, 2021; Subich et al., 2025): first, the model fails to capture the observed ice-cloud features at the correct locations, and second, it generates spurious ice clouds elsewhere. To reduce the error amplification caused by the double-penalty effect in RTTOV-simulated TBs, it is necessary to reduce the biases in the spatial distribution and density representation of ice-phase particles in the WRF forecasts. However, the true distributions of all ice-phase particles are difficult to obtain directly. Since Section 3.4 has demonstrated that graupel particles dominate the TB variations at 183 GHz, our approach is to replace the graupel parameters in the WRF forecast fields with graupel parameters retrieved from AWS 183 GHz observed TBs. The graupel retrieval was performed using a U-Net network trained with ATMS observed TBs and graupel content profiles from the ATMS Level-2 products. The training dataset contains 1343 ice-cloud scenes over coastal regions of China. Figures 11 and 12 show comparisons between RTTOV-simulated and observed TBs for AWS Channel 32 (176.311 GHz) and Channel 44 (325.15 ± 6.6 GHz) after applying the retrieved graupel parameters. Table 3 further summarizes the statistical metrics of the differences between simulated and observed TBs for five AWS 183 GHz channels and four AWS 325 GHz channels using the retrieved graupel parameters. After replacing the WRF graupel parameters with the AWS-retrieved graupel parameters, the RMSEs for most channels decrease significantly by 20–50%. In particular, the RTTOV simulation RMSE for Channel 44 (325.15±6.6 GHz) decreases from approximately 58 K to 30 K. These results indicate that, under relatively reliable atmospheric profile and surface parameter conditions, errors in the content and spatial distribution of graupel particles have a substantial impact on TB simulations. The graupel parameters retrieved from 183 GHz observed TBs provide a more accurate representation of graupel spatial distribution and density than the WRF forecasts. Therefore, all subsequent optimization experiments adopt the retrieved graupel parameters.

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

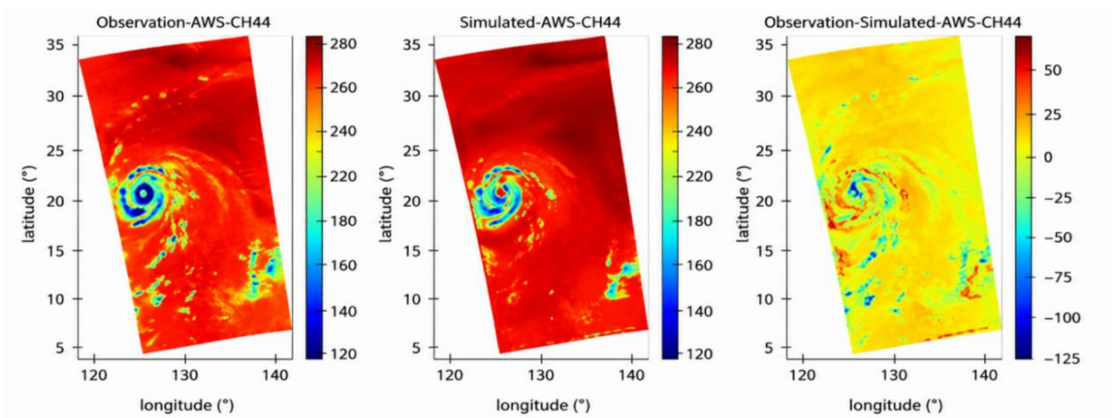


(a) **(b)** **(c)**



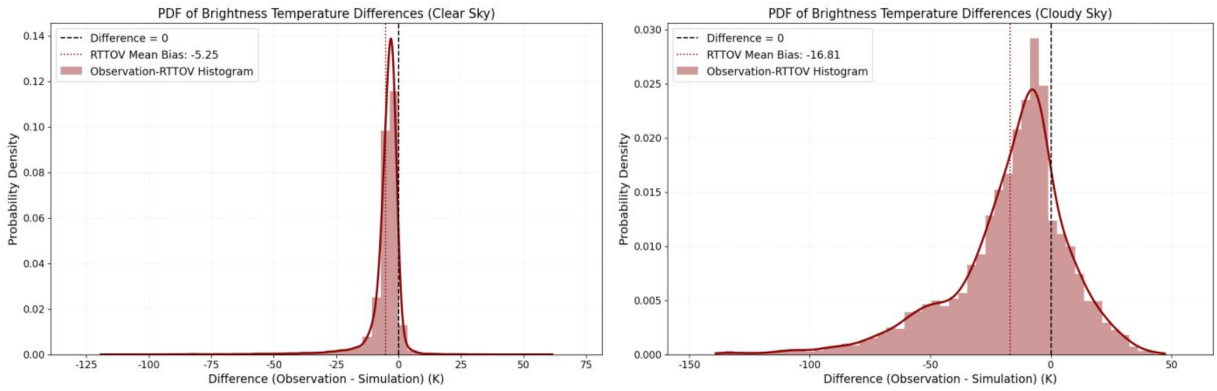
(d) **(e)**

Figure 11. RTTOV model simulation results for Channel 32 (176.311 GHz, RMSE: 9.33 K): (a) AWS-observed TBs, (b) simulated TBs, (c) differences between simulated and observed TBs, (d) PDF of TB differences under clear-sky conditions, (e) PDF of TB differences under cloudy-sky conditions.



(a) **(b)** **(c)**

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026



(d) **(e)**

Figure 12. RTTOV model simulation results for Channel 32 (325.15±6.6 GHz, RMSE: 14.96 K): (a) AWS-observed TBs, (b) simulated TBs, (c) differences between simulated and observed TBs, (d) PDF of TB differences under clear-sky conditions, (e) PDF of TB differences under cloudy-sky conditions.

Table 3. RMSE for the AWS 183 GHz and 325 GHz channels

Channel	32	33	34	35	36	41	42	43	44
CLEAR SKY	4.99	4.76	4.45	3.81	4.46	5.98	7.44	8.43	10.80
CLOUD SKY	27.25	21.73	18.49	16.99	13.52	7.29	27.41	28.80	30.03

4.2 Optimization Results of Bulk Scattering Parameters Using the Deep Reinforcement Learning Algorithm

As explained in Section 3.4, the RTTOV-SCATT-DRL model targeting ice-cloud TBs at 183 GHz and 325 GHz currently focuses exclusively on optimizing the scattering parameters of graupel particles. Accordingly, the input variables include the primary factors influencing graupel scattering characteristics: graupel water content (GWC), temperature (affecting the dielectric constant), atmospheric pressure (affecting particle distribution), and water vapor content (affecting particle growth processes). The asymmetry factor, a key parameter of the scattering phase function, is closely related to the particle size distribution, the forward-scattering intensity of large particles, the proportion of hydrometeor phases, particle shape, and operating frequency. Since atmospheric profile parameters do not explicitly contain particle size information, while RTTOV inherently assumes a fixed particle size distribution when calculating scattering parameters, RTTOV-SCATT-DRL additionally introduces the particle effective radius (r_{me}) as a variable associated with the particle size distribution.

In this study, graupel particles are represented using the RH size distribution model. The RH model assumes that the particle size distribution follows the following Gamma distribution:

$$\mathrm{n(D) = N_0 D^{\mu} exp(-\lambda D)} \quad (15)$$

Where: $N_0 = 4 \times 10^6 m^{-4}$ (intercept parameter); $\mu = 0$ (shape parameter for the simplified case); λ denotes the slope parameter (m^{-1}); D represents the equivalent diameter of particles (m). The effective radius of particles is defined as the average value of the ratio of the total particle volume to the total projected area, which is commonly used to characterize

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

scattering and radiation properties. For the Gamma distribution with $\mu = 0$, the expression for the effective radius r_e is simplified as:

$$r_e = \Gamma(4)/\big(\lambda\Gamma(3)\big) = 3/\lambda \tag{16}$$

Considering unit conversion (converting from meter to micrometer), the final formula for the effective radius is:

$$r_e = (3/\lambda) \times 10^6 \big(Unit\text{：} \ \mu m\big) \tag{17}$$

The slope parameter λ can be solved from the relationship between the mass mixing ratio of graupel particles and the slope parameter λ:

$$\lambda = \left(\frac{\pi\rho_p N_0 \Gamma(4+\mu)}{6 q_g \rho_a \Gamma(1+\mu)}\right)^{\frac{1}{3+\mu}} \tag{18}$$

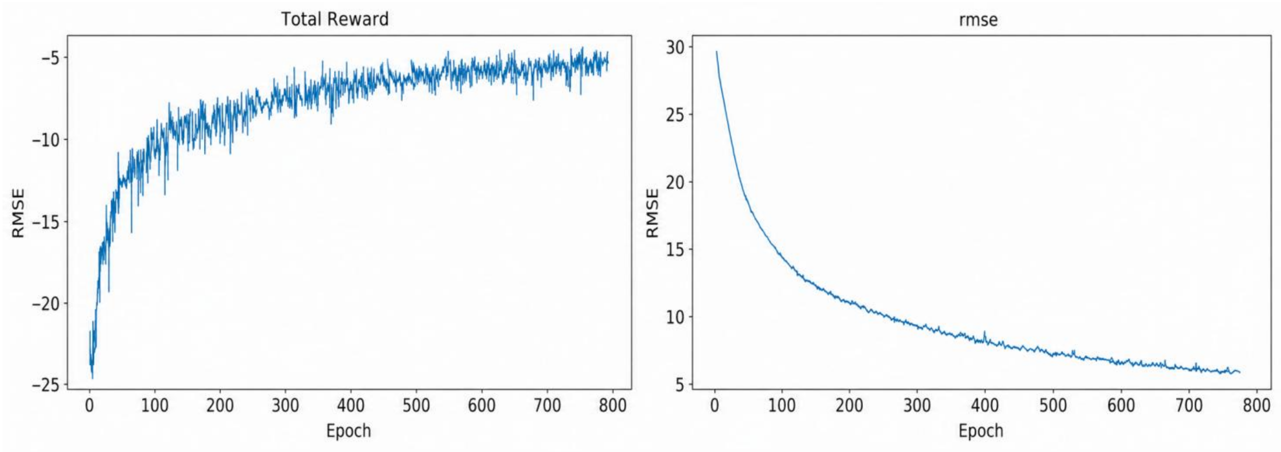
Where: $\rho_p = 400\ \mathrm{kg{\cdot}m^{-3}}$ (graupel particle density); $\rho_a = (p \times 100) / (R\_d * T)$ (air density, $\mathrm{kg{\cdot}m^{-3}}$), p denotes air pressure （hPa）, T denotes temperature （K）, and $R\_d = 287\ \mathrm{J{\cdot}kg^{-1}{\cdot}K^{-1}}$。

For the Gamma distribution with μ=0 , the slope parameter is determined by the following formula:

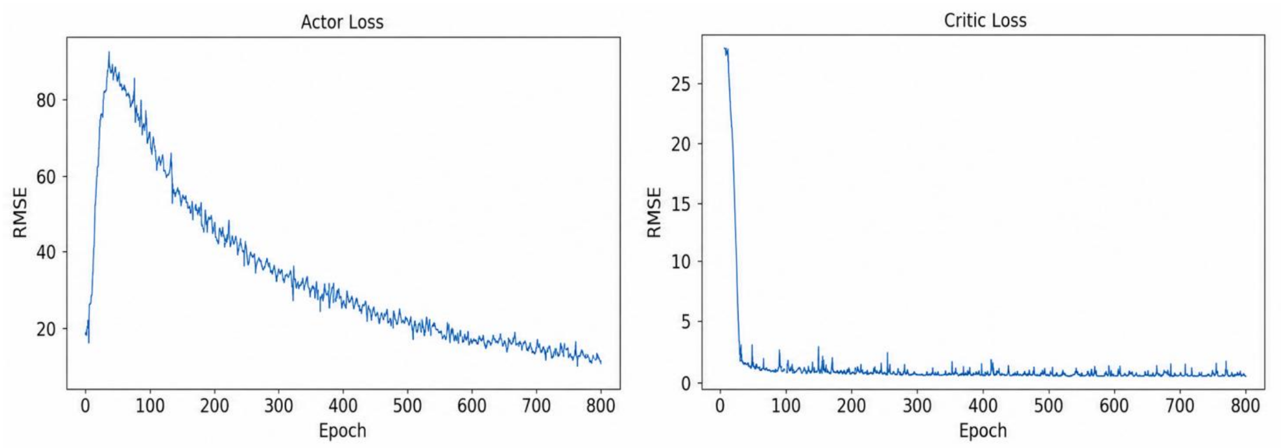
$$\lambda = \left(\frac{\pi\rho_p N_0}{q_g \rho_a}\right)^{\frac{1}{3}} \tag{19}$$

The optimization process of the RTTOV-SCATT-DRL model is illustrated using the AWS CH44 channel (325.15 ± 6.6 GHz) as an example. The training history, shown in Figure 13, presents the key metrics over 800 epochs for both the training and validation datasets, including training/validation RMSE, total reward, Actor Loss, and Critic Loss. Figure 14 further depicts the evolution of training and validation RMSE as a function of epoch. During the initial training stage (epochs 0–100), the RMSE decreased rapidly. The training RMSE dropped from approximately 30.5 K to below 15 K, while the validation RMSE decreased simultaneously from about 28.5 K to around 14 K, indicating that the model quickly learned the mapping between atmospheric states and scattering parameters. In the subsequent stage (epochs 100–300), the RMSE entered a stable convergence phase. The training RMSE was further reduced to approximately 8–10 K, while the validation RMSE stabilized within the range of 9–11 K with progressively smaller fluctuations, demonstrating the good generalization capability of the RTTOV-SCATT-DRL model. After epoch 300, the training RMSE continued to decrease gradually to approximately 5–7 K. Meanwhile, the validation RMSE exhibited only minor oscillations within the range of 7–9 K while maintaining an overall downward trend, with no evident signs of overfitting. The total reward increased steadily from an initial value of about −25 to approximately − 8 before reaching a plateau, indicating that the policy progressively approached the objective of minimizing the RMSE. The Actor Loss decreased rapidly from around 70 to below 30 and continued to improve gradually thereafter. The Critic Loss, although highly variable during the early training stage, converged rapidly to values close to zero. These results confirm the effectiveness of the TD3 framework, particularly the twin-Critic architecture and adaptive noise mechanism. Overall, the reinforcement learning framework achieved substantial error reduction within 800 epochs. The validation RMSE decreased from approximately 28 K at the beginning of training to about 8 K at convergence, corresponding to a reduction of

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026



(a) **(b)**



(c) **(d)**

Figure 13. Training process of the algorithm: (a) total reward; (b) RMSE; (c) Actor network loss; (d) Critic network loss.

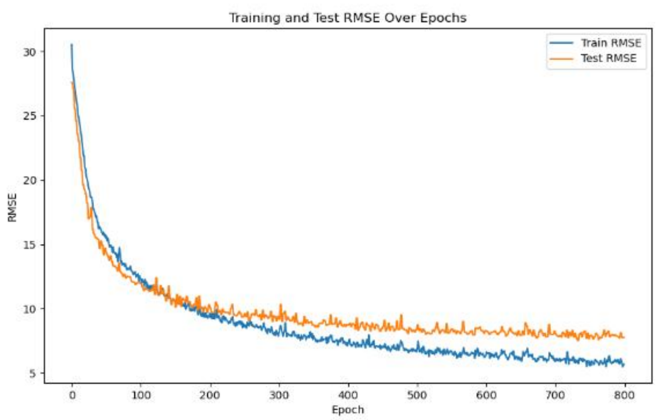
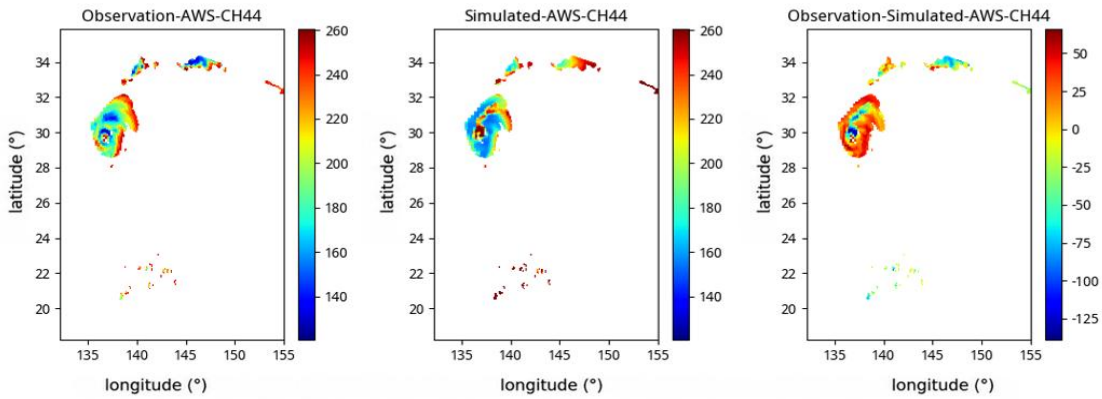


Figure 14. Curves of training and testing RMSE versus epochs.

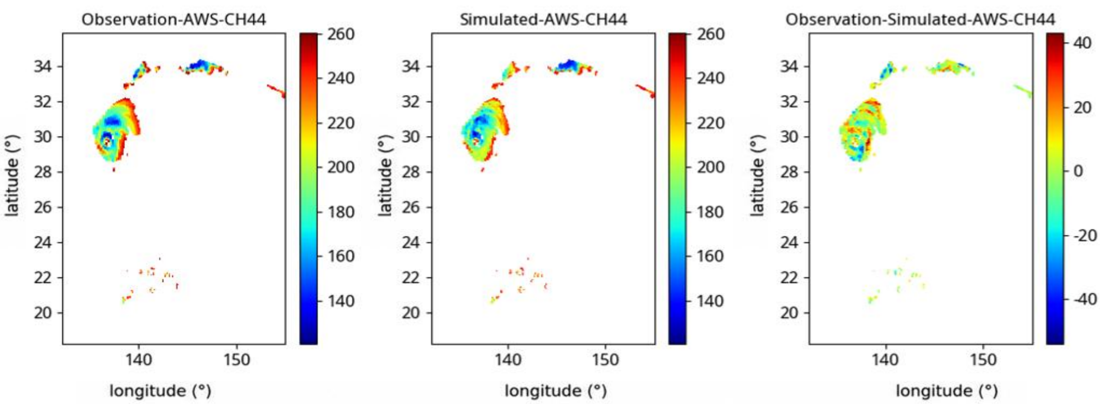
https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

more than 70%. Meanwhile, the reward approached saturation and both loss curves exhibited smooth and stable convergence behavior.

An independent test dataset is then used to evaluate the optimization performance of the trained RTTOV-SCATT-DRL model for the AWS Channel 44 (325.15 ± 6.6 GHz). The following results correspond to the Super Typhoon Halong case (Time: UTC 00:44, October 8, 2025, Area: Longitude 130-155, Latitude 18-36, AWS Data: W_NO-KSAT-Tromso,SAT,AWS1-MWR-1B-RAD_C_OHB__20251008023147_G_O_20251008004426_20251008022414_C_N____.nc). Figures 15−17 show that, after DRL optimization, the mean RTTOV-simulated TB is 199.1 K, which is very close to the observed mean value of 200.5 K, while the standard deviation (STD) further decreases to 31.1 K. The mean bias between observations and optimized simulations is substantially reduced to 1.33 K, and the STD decreases to 12.83 K. Compared with the standard RTTOV results (bias: −3.80 K; STD: 40.10 K), the RMSE decreases from approximately 40.28 K to 12.90 K, representing an improvement of more than 60%, while the correlation coefficient increases from 0.395 to 0.926.



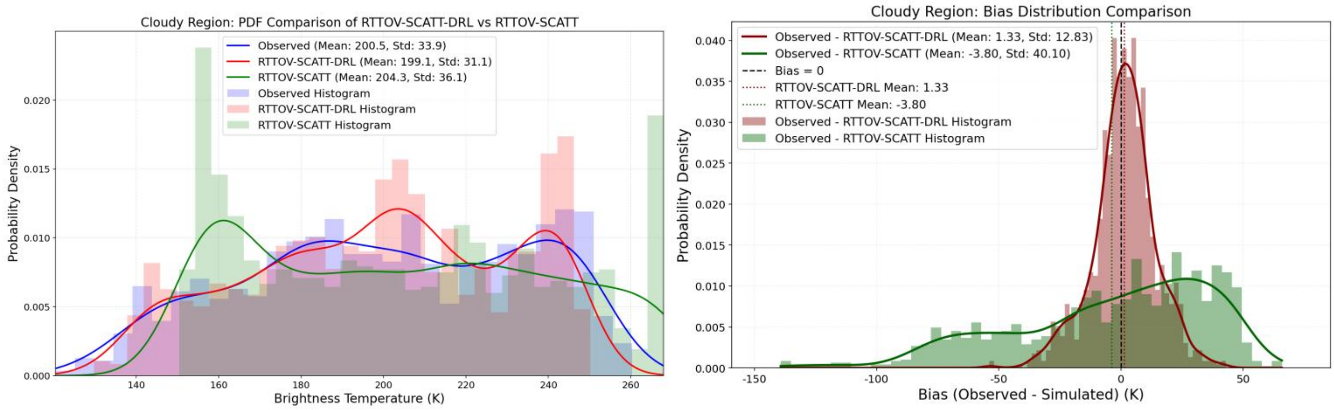
(a)



(b)

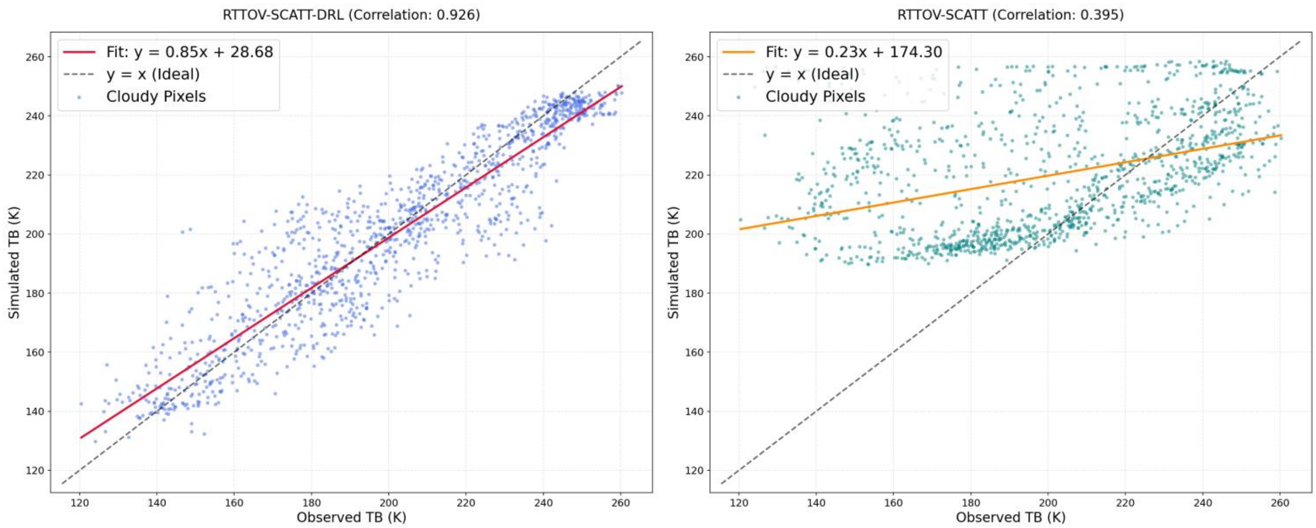
https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

Figure 15. Simulated cloud-region TBs for channel 44 (325.15 ± 6.6 GHz) before and after model optimization: (a) standard RTTOV simulation (RMSE: 40.28 K, bias: -3.80 K); (b) reinforcement learning-optimized RTTOV simulation (RMSE: 12.90 K, bias: 1.33 K). Left: AWS observed TB; middle: simulated TB; right: observation minus simulation (O-B).



(a) (b)

Figure 16. Comparison of cloud-region TBs for channel 44 (325.15 ± 6.6 GHz) before and after optimization: (a) TB distribution; (b) difference distribution.



(a) (b)

Figure 17. Scatter plots of channel 44 (325.15 ± 6.6 GHz) TBs: (a) reinforcement learning-optimized RTTOV; (b) standard RTTOV.

To quantitatively assess how closely the Observation minus Simulation (OMB) error distribution of the RTTOV-SCATT-DRL model approaches an ideal Gaussian distribution, we analyzed several statistical indicators, including skewness, excess kurtosis, the Kolmogorov-Smirnov (KS) statistic, and Kullback-Leibler (KL) divergence.

Skewness measures the symmetry of a probability distribution and is defined as:

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

$$S = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s}\right)^3 \quad (20)$$

where $\bar{x}$ is the sample mean, s is the sample STD, and n is the sample size. For an ideal Gaussian distribution, skewness should be close to zero.

Excess kurtosis quantifies the peakedness and tail thickness of a distribution, defined as:

$$K = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s}\right)^4 - \frac{3(n-1)^2}{(n-2)(n-3)} \quad (21)$$

For a Gaussian distribution, the excess kurtosis equals zero. Positive excess kurtosis indicate fatter tails and a sharper peak than the Gaussian distribution, and negative excess kurtosis indicates thinner tails and a flatter peak.

The KS statistic measures the maximum deviation between the empirical cumulative distribution function (CDF) and the theoretical Gaussian CDF:

$$D = \max_{x} |F_n(x) - F(x)| \quad (22)$$

where $F_n(x)$ is the empirical CDF and $F(x)$ is the Gaussian CDF. A smaller D indicates a distribution closer to Gaussian.

KL divergence measures the information distance between the actual probability distribution and the Gaussian distribution:

$$D_{KL}(P \parallel Q) = \int p(x) \ln \frac{p(x)}{q(x)} dx \quad (23)$$

where $p(x)$ is the actual OMB PDF and $q(x)$ is the corresponding Gaussian fit. KL values closer to zero indicate greater similarity.

Table 4 summarizes the statistical characteristics of the OMB distributions for the RTTOV-SCATT and RTTOV-SCATT-DRL models in cloudy regions of AWS CH44 channel samples.

Table 4. Statistical characteristics and Gaussianity metrics of OMB error distributions

Model	**Mean**	**Std**	**Skewness**	**Excess Kurtosis**	**KS D**	**KL Divergence**
RTTOV-SCATT	-3.8	40.1	-0.729	-0.149	0.091	0.155
RTTOV-SCATT-DRL	1.33	12.83	-0.286	1.016	0.059	0.048

As shown in Table 4, the RTTOV-SCATT-DRL model optimized with AWS observation exhibits clear improvements in error statistics. The mean of OMB errors decreases from -3.80 K to 1.33 K, and the STD reduces from 40.10 K to 12.83 K, indicating that both systematic and random errors are significantly suppressed. Skewness decreases from -0.729 to -0.286, demonstrating a more symmetric error distribution. The KS statistic drops from 0.091 to 0.059, and KL divergence decreases from 0.155 to 0.048, reduced by approximately 35% and 69%, respectively, indicating that the optimized OMB distribution is substantially closer to a Gaussian distribution. The post-optimization excess kurtosis is positive, showing a mild peakedness, with the majority of samples concentrated near zero bias and long-tail extreme errors effectively reduced. Consequently, the PDF of OMB errors exhibits a more pronounced Gaussian feature, which has important physical and practical implications. In variational data assimilation theory, system errors are typically based on the Gaussian distribution assumption, meaning both observation and background errors are assumed to follow zero-mean, symmetric Gaussian

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

distributions. Only when the OMB error distribution approximates a Gaussian can the assimilation system produce optimal analysis increments; otherwise, skewed distributions may lead to "double penalty" effects, increased analysis bias, and degraded forecast skill (Geer et al., 2014). Therefore, incorporating particle size distribution and shape information allows the model to more accurately capture the scattering effects of large particles, significantly improving high-frequency channel simulations, especially for the 325 GHz channel.

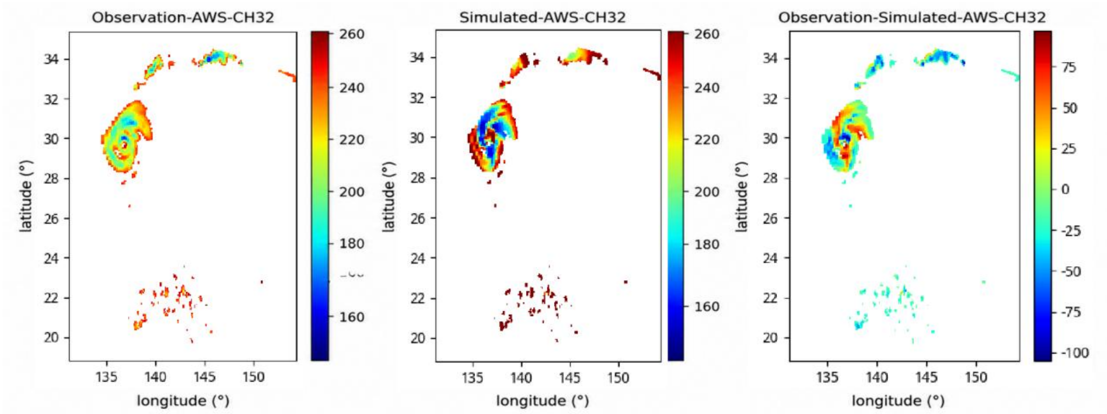
4.3 Evaluation of the RTTOV-SCATT-DRL Model for the AWS 183 GHz Channels

This section presents evaluation results of the trained RTTOV-SCATT-DRL model for simulating ice-cloud TBs across all five AWS 183 GHz channels (CH32-36). Channels 32–36 located near the 183 GHz water vapor absorption band with frequency ranges from 176.311 to 182.311 GHz. Owing to the strong water vapor absorption peak at 183.311 GHz, these channels are highly sensitive to ice-cloud scattering effects. Channels closer to the 183.311 GHz are dominated more strongly by atmospheric absorption, whereas channels farther from the center frequency exhibit enhanced sensitivity to ice-particle scattering.

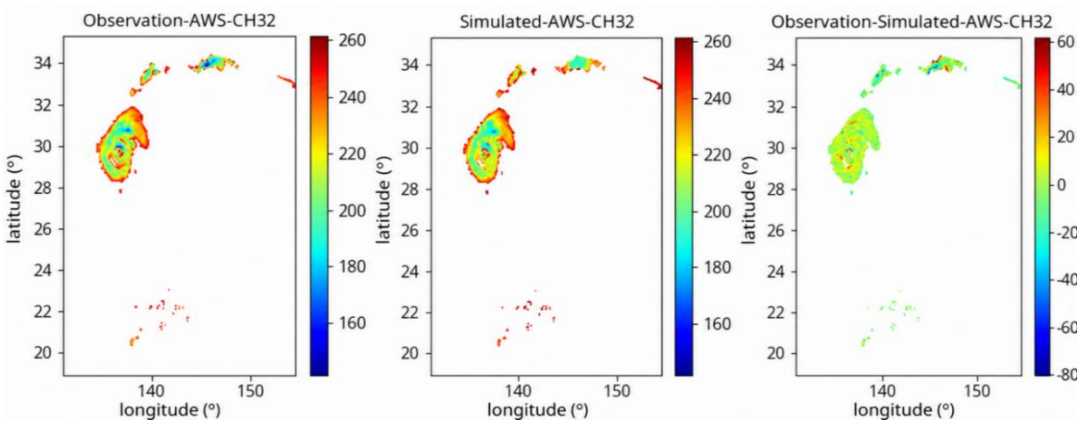
Figure 18 shows the observed and simulated TBs for Channel 32 within cloudy regions, together with the corresponding OMB differences. For RTTOV-SCATT, the RMSE reaches 36.78 K with a bias of −6.69 K. After optimization, the RTTOV-SCATT-DRL model reduces the RMSE to approximately 14.33 K and the bias to −1.68 K.

Figure 23 summarizes the validation experiment results at the AWS 183 GHz channels for the typhoon Halong (2025, No. 22). After optimization, the RMSE of each channel decreased significantly, with an average reduction of 45%. Among all channels, Channel 32 achieved the most remarkable reduction, dropping from 36.78 K to 14.33 K, representing a decrease of 61%. Meanwhile, Figure 24 presents the test results based on 2,153 atmospheric profile samples of the super typhoons Mitag (2025, No. 17). Under conditions with stronger scattering effects, the error of the standard RTTOV model increased further. By optimizing the scattering parameters via DRL, the error decreased substantially, with the reduction magnitude reaching approximately 60% for most channels and the average bias reduction exceeding 75%. The above results demonstrate that the reinforcement learning-based ice cloud bulk scattering model constructed in this study possesses high simulation accuracy and stability in the 183 GHz water vapor absorption band.

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

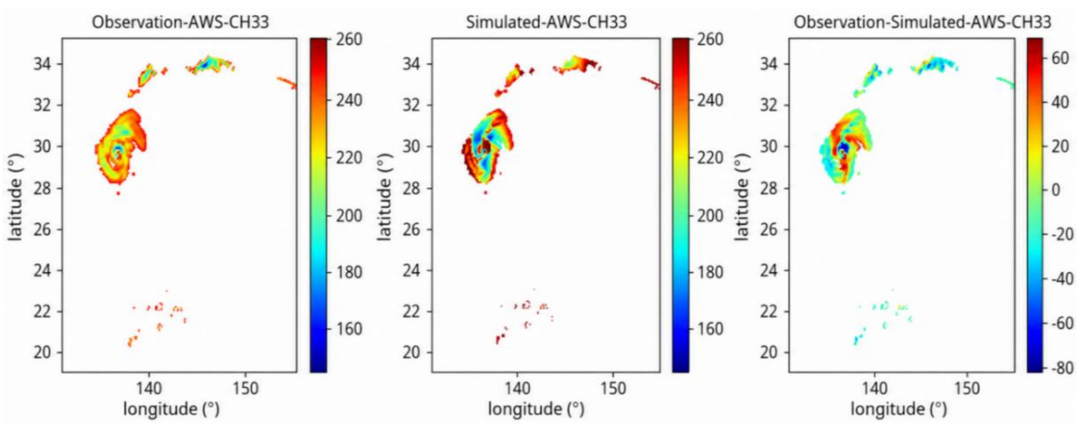


(a)



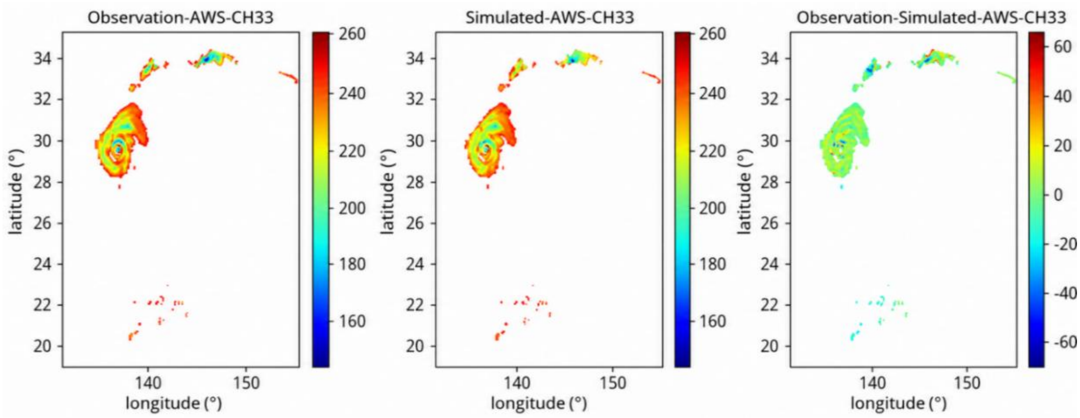
(b)

Figure 18. Simulated cloud-region TBs for channel 32 (176.311 GHz) before and after model optimization: (a) standard RTTOV simulation (RMSE: 36.78 K, bias: -6.69 K); (b) DRL optimized RTTOV simulation (RMSE: 14.33 K, Bias: -1.68 K). Left: AWS observed TB; middle: simulated TB; right: observation minus simulation (O-B).



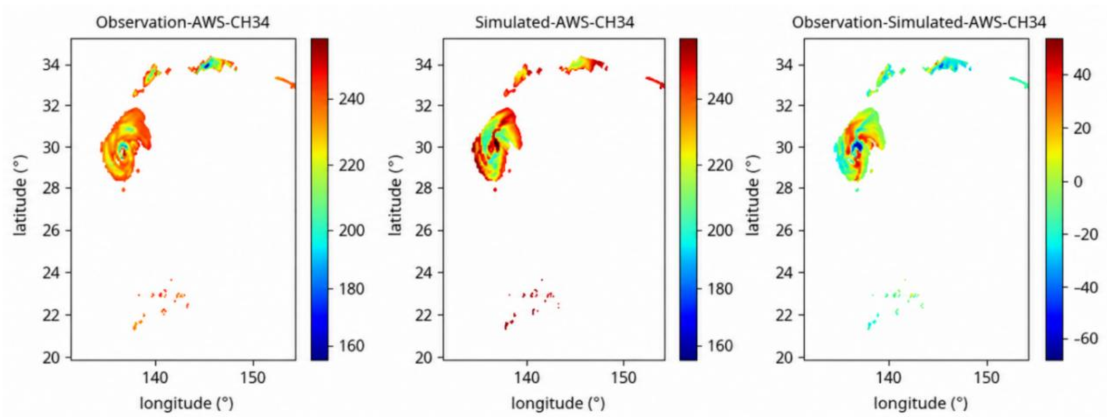
(a)

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

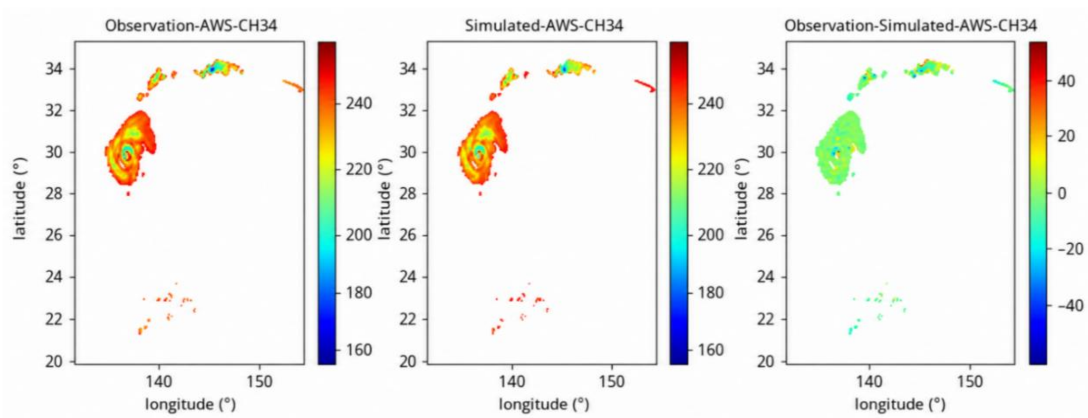


(b)

Figure 19. Simulated cloud-region TBs for channel 33 (178.811 GHz) before and after model optimization: (a) standard RTTOV simulation (RMSE: 28.48 K, Bias: -4.36 K); (b) DRL optimized RTTOV simulation (RMSE: 12.62 K, Bias: -0.35 K). Left: AWS observed TB; middle: simulated TB; right: observation minus simulation (O-B).



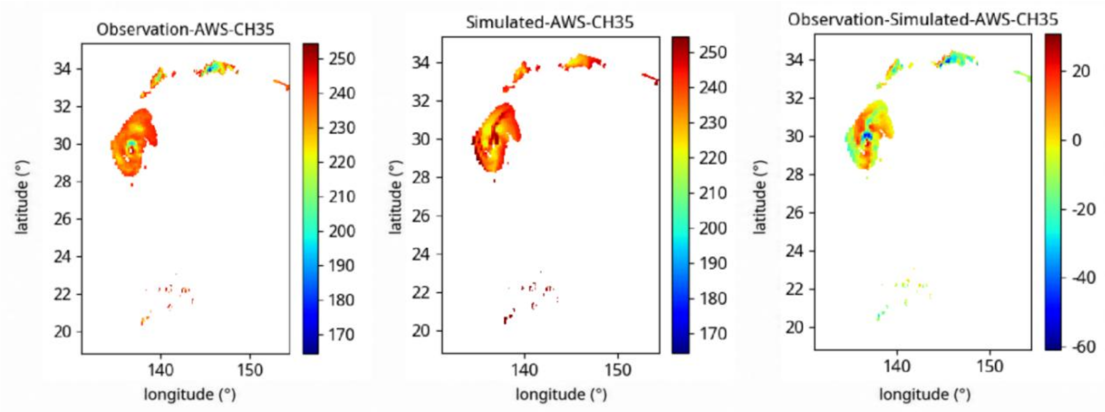
(a)



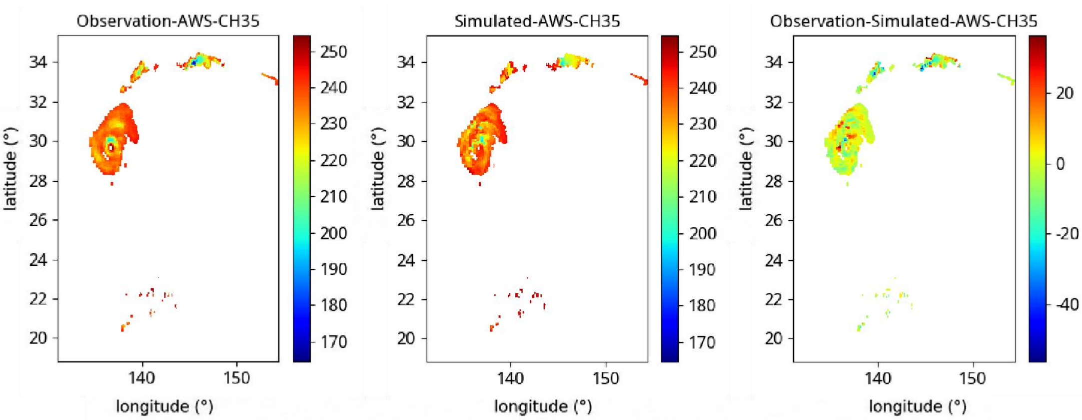
(b)

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

Figure 20. Simulated cloud-region TBs for channel 34 (180.311 GHz) before and after model optimization: (a) standard RTTOV simulation (RMSE: 20.97 K, Bias: -4.18 K); (b) DRL optimized RTTOV simulation (RMSE: 9.45 K, Bias: 0.03 K). Left: AWS observed TB; middle: simulated TB; right: observation minus simulation (O-B).



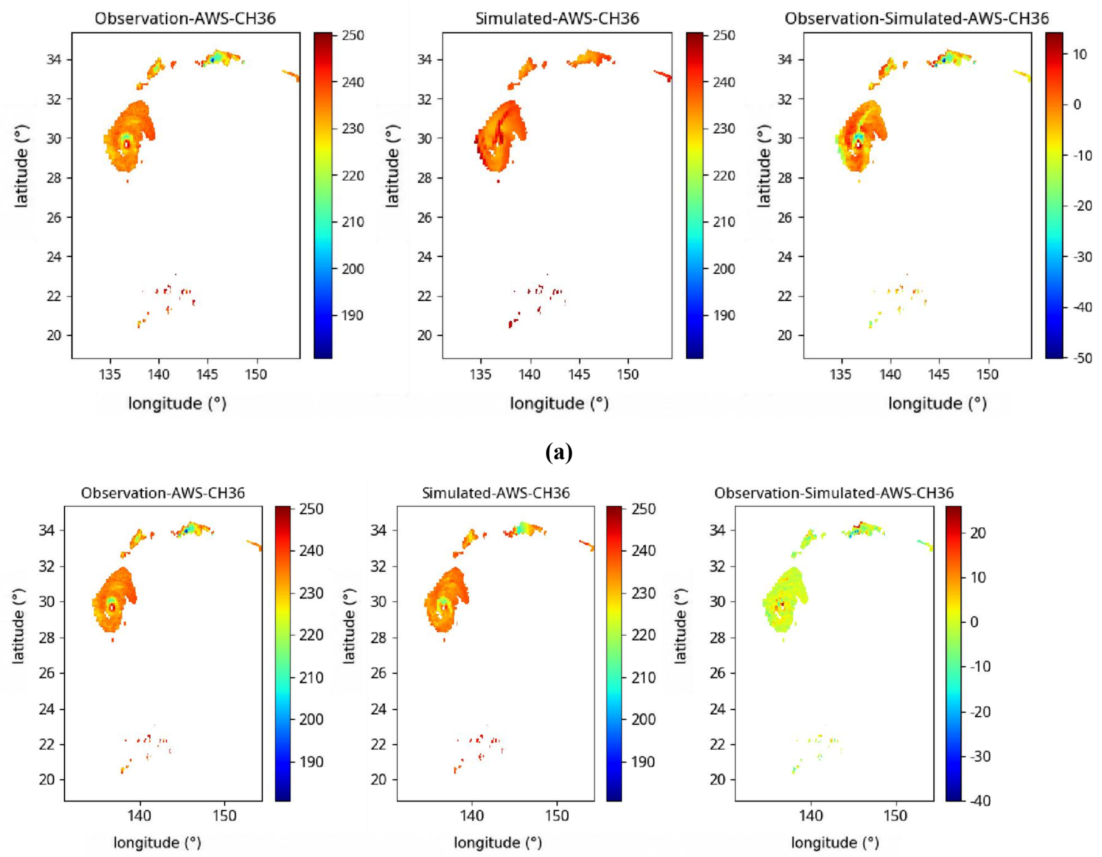
(a)



(b)

Figure 21. Simulated cloud-region TBs for channel 35 (181.511 GHz) before and after model optimization: (a) standard RTTOV simulation (RMSE: 13.95 K, Bias: -5.47 K); (b) DRL optimized RTTOV simulation (RMSE: 9.43 K, Bias: -1.57 K). Left: AWS observed TB; middle: simulated TB; right: observation minus simulation (O-B).

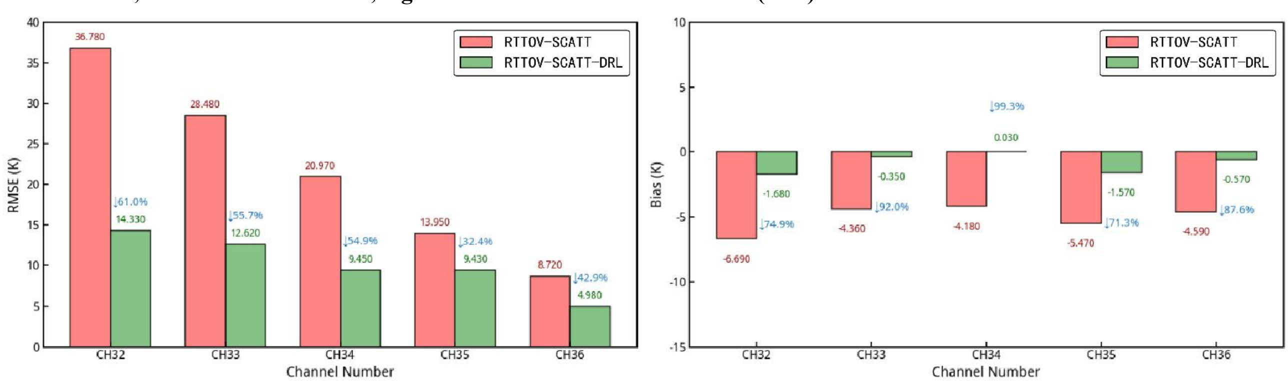
https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026



(a)

(b)

645 **Figure 22. Simulated cloud-region TBs for channel 36 (182.311 GHz) before and after model optimization: (a) standard RTTOV simulation (RMSE: 8.72 K, Bias: -4.59 K); (b) DRL optimized RTTOV simulation (RMSE: 4.98 K, Bias: -0.57 K). Left: AWS observed TB; middle: simulated TB; right: observation minus simulation (O-B).**

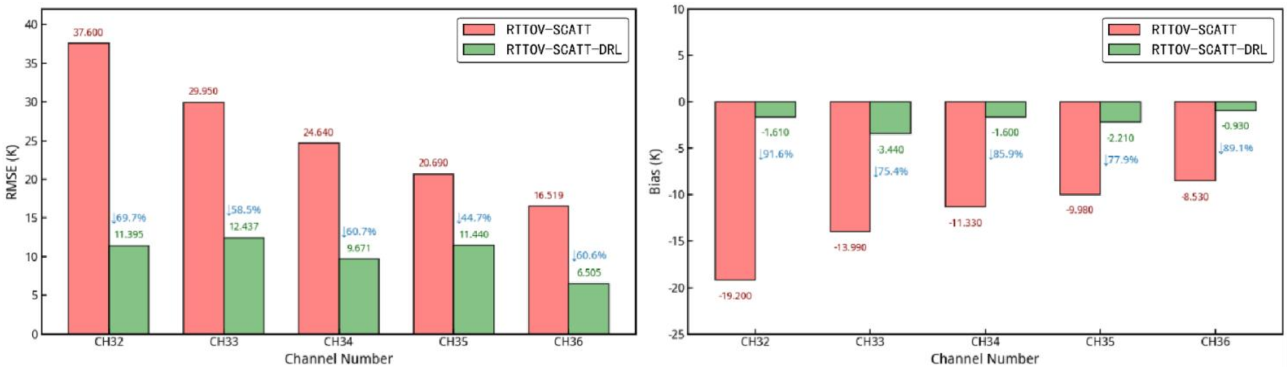


(a)

(b)

650 **Figure 23 Validation experiment results at the AWS 183 GHz channels for the typhoon Halong (Number of samples: 1151): (a) Comparison of RMSE before and after optimization; (b) Comparison of Bias before and after optimization**

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

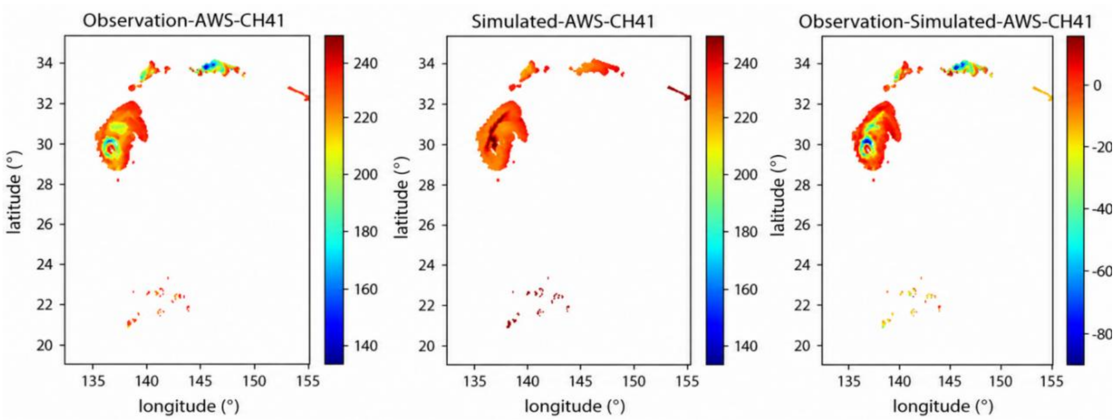


(a) (b)

Figure 24 Validation experiment results at the AWS 183 GHz channels for the super typhoon Mitag (Number of samples: 2153): (a) Comparison of RMSE before and after optimization; (b) Comparison of Bias before and after optimization

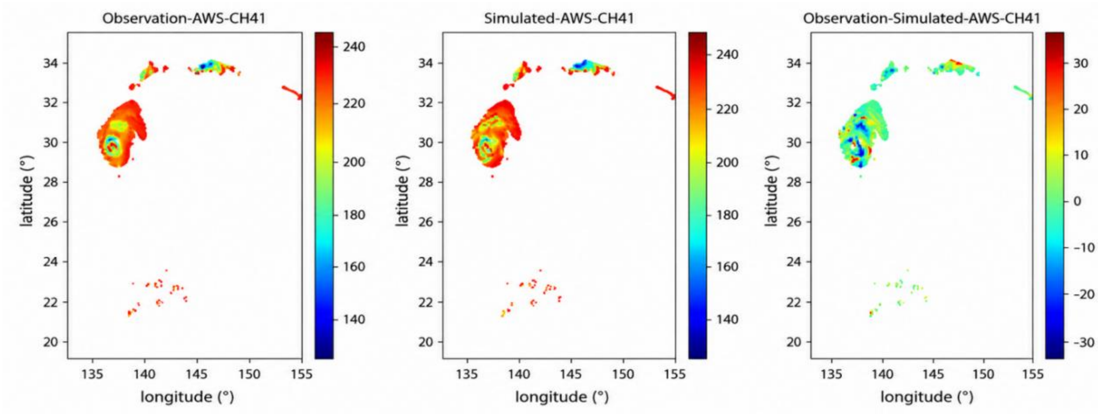
4.4 Evaluation of the RTTOV-SCATT-DRL Model for the AWS 325 GHz Channels

This section presents the evaluation results of the RTTOV-SCATT-DRL model for simulating ice-cloud TBs across all four AWS 325 GHz channels (CH41–CH43). Taking the super typhoon Halong as a case, we compare the RMSE and bias of the OMB in cloud areas between the simulated TB before and after optimization and actual AWS observations. The optimization results of Channel 41 are shown in Figure 25. For the standard RTTOV simulation, the RMSE is 22.81 K, the bias is −12.18 K. After optimization, the RMSE decreases to approximately 9.93 K, with a reduction of about 56.4%, and the bias is reduced to 0.57 K. In particular, the simulation performance is significantly improved in the central area of the typhoon.



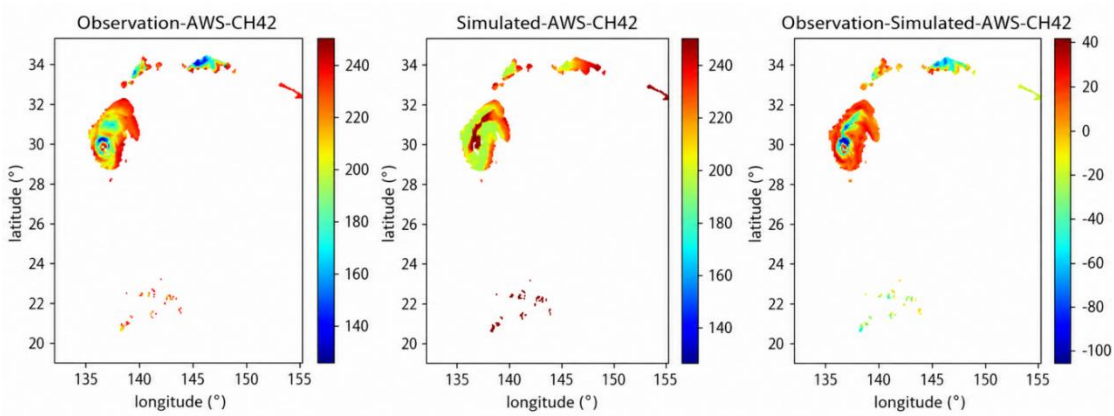
(a)

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

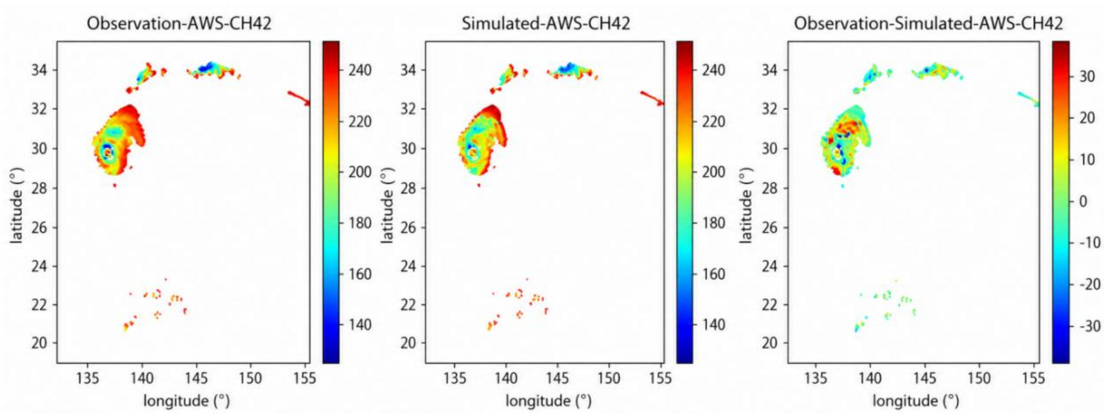


(b)

Figure 25 Simulated cloud-region TBs for channel 41 (325.15±1.2 GHz) before and after model optimization: (a) standard RTTOV simulation (RMSE: 22.81 K, Bias: −12.18 K); (b) DRL optimized RTTOV simulation (RMSE: 9.93 K, Bias: 0.57 K). Left: AWS observed TB; middle: simulated TB; right: observation minus simulation (O-B).



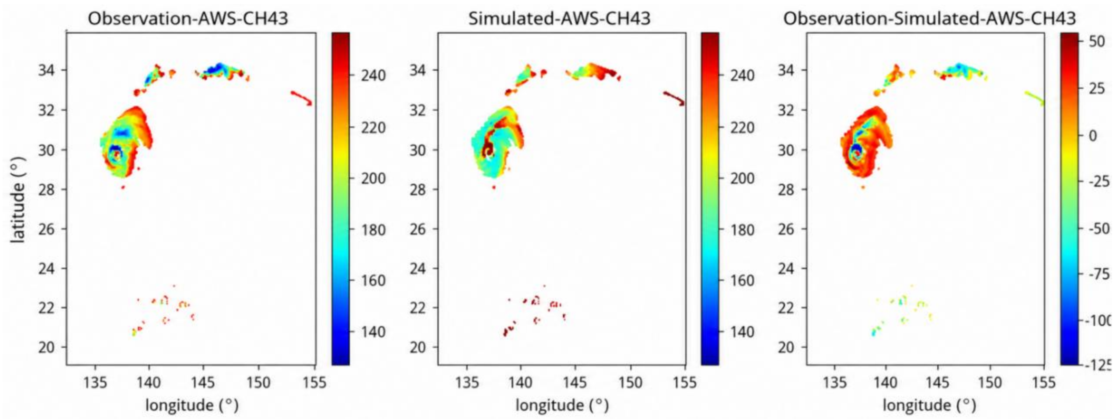
(a)



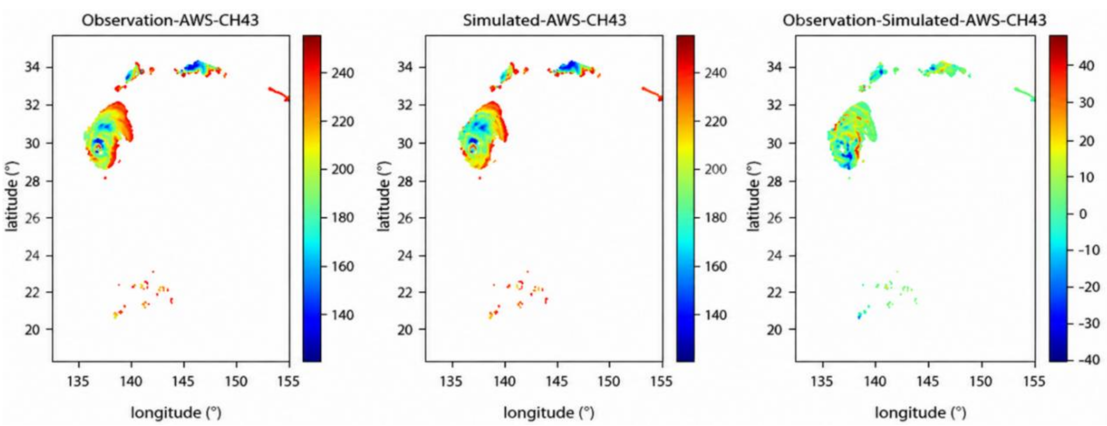
(b)

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

Figure 26. Simulated cloud-region TBs for channel 42 (325.15 ± 2.4 GHz) before and after model optimization: (a) standard RTTOV simulation (RMSE: 29.02 K, Bias: −7.86 K); (b) DRL optimized RTTOV simulation (RMSE: 12.67 K, Bias: 3.62 K). Left: AWS observed TB; middle: simulated TB; right: observation minus simulation (O-B).



(a)



(b)

Figure 27. Simulated cloud-region TBs for channel 43 (325.15 ± 4.1 GHz) before and after model optimization: (a) standard RTTOV simulation (RMSE: 36.01 K, Bias: −5.70 K); (b) DRL optimized RTTOV simulation (RMSE: 12.73 K, Bias: 3.82 K). Left: AWS observed TB; middle: simulated TB; right: observation minus simulation (O-B).

As a high-frequency channel in the terahertz band, 325 GHz is more sensitive to ice cloud scattering, and the standard RTTOV exhibits larger errors. Figure 28 summarizes the validation experiment results at the AWS 325 GHz channels for the typhoon Halong. After optimization, the RMSE of each channel is reduced on average by approximately 60%, with Channel 44 achieving a maximum reduction of 69.7%, resulting in an RMSE of 12.9 K. The bias converges from ±5–12 K to ±0.5–3.8 K.

For the super typhoon Mitag, the optimization effect at 325 GHz is even more pronounced. Figure 29 shows the error comparison for the four channels. Errors are larger at frequencies farther from the central frequency, and the optimization yields correspondingly greater improvements, with RMSE reduced by 68% to 13.32 K. These results indicate that the

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

reinforcement-learning-optimized RTTOV ice cloud scattering model significantly improves simulation accuracy for the AWS 325 GHz channels compared with the standard RTTOV.

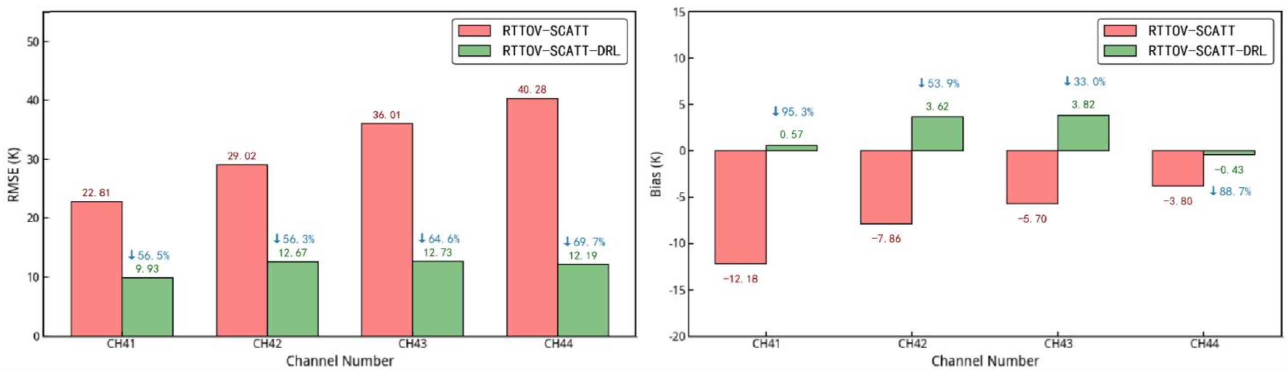


Figure 28. Validation experiment results at the AWS 325 GHz channels for the typhoon Halong (1,151 samples): (a) RMSE before and after optimization; (b) Bias before and after optimization.

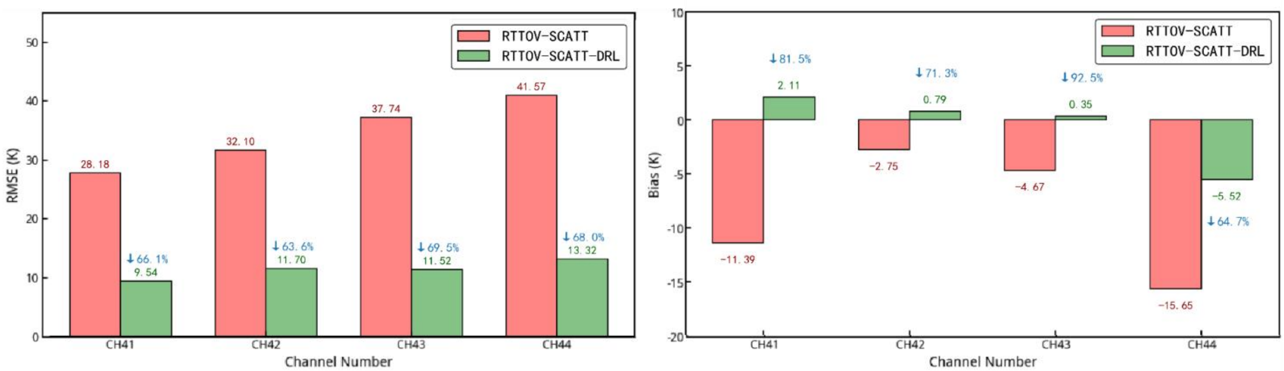


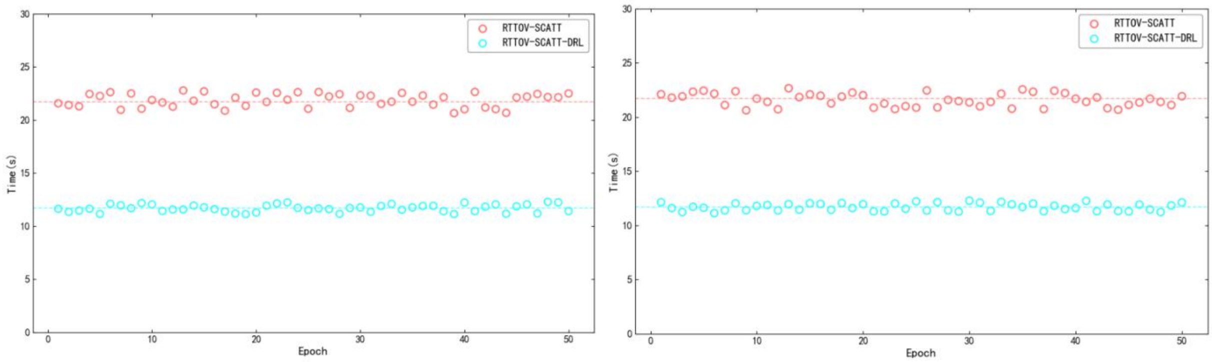
Figure 29. Validation experiment results at the AWS 325 GHz channels for the super typhoon Mitag (2,153 samples): (a) RMSE before and after optimization; (b) Bias before and after optimization.

4.5 Computational Efficiency Comparison between RTTOV and RTTOV-SCATT-DRL

To comprehensively evaluate the computational efficiency of different RT models in operational applications, this study assessed the average computation time for the AWS 183 GHz (five channels) and 325 GHz (four channels) frequency bands. To minimize fluctuations caused by computer performance and system-related factors, 50 repeated timing tests were conducted for each model. In each test, the sample size was fixed at 145×400, and all experiments were performed under identical testing conditions using the same hardware platform.

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

Figure 30 presents the distributions of computation times from the 50 repeated tests for both the RTTOV-SCATT model and the RTTOV-SCATT-DRL model at the 183 GHz and 325 GHz, respectively. The corresponding statistical results are summarized in Table 5.



(a) **(b)**

Figure 30. Comparison of average computation times (unit: s): (a) 183 GHz; (b) 325 GHz.

Table 5. Comparison of average computation times for the two frequency bands (unit: s).

Channel	RTTOV-SCATT	RTTOV-SCATT-DRL
183GHz	21.7	11.7
325GHz	22.2	11.9

At the 183 GHz frequency band, the average computation time of RTTOV-SCATT was 21.7 s, whereas the RTTOV-SCATT-DRL model required only 11.7 s on average. At the 325 GHz frequency band, the computation time of RTTOV-SCATT increased slightly to 22.2 s, while the RTTOV-SCATT-DRL model remained stable at approximately 11.9 s.

The theoretical analysis is consistent with the experimental results. As a parameterized fast RT model, RTTOV-SCATT exhibits little sensitivity of computational cost to frequency changes, resulting in only a small difference in runtime between the 183 GHz and 325 GHz bands. In contrast, the RTTOV-SCATT-DRL model is based on a fixed neural network architecture whose computational cost is determined solely by the forward-propagation process and is independent of input frequency. Consequently, its computation time remains nearly unchanged across the two frequency bands.

Overall, the RTTOV-SCATT-DRL model demonstrates a substantial advantage in computational efficiency. Its runtime is only about 54% of that of RTTOV. While maintaining high simulation accuracy, the proposed model significantly reduces computational cost, highlighting its strong potential for operational large-scale simulations, data assimilation applications, and real-time retrieval systems.

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

5 Conclusion

This study proposed a DRL (TD3)-based framework for optimizing ice-cloud BSPs. Using observed ice-cloud TBs from the AWS satellite at 183 GHz and 325 GHz as optimization targets and observational constraints, the framework effectively addresses the inaccurate representation of ice-phase particle BSPs in RTTOV-SCATT module under complex ice-cloud conditions. By combining a pre-trained differentiable scattering surrogate model with DRL optimizing, a new RTTOV-SCATT-DRL model was developed to replace the operational RTTOV-SCATT module and was tightly integrated into the standard RTTOV Ver. 14.0 framework. The proposed model enables dynamic optimization of the extinction coefficient, scattering coefficient, and asymmetry factor of graupel particles within ice clouds.

Approximately 176,000 typhoon atmospheric profiles publicly released by the AWS mission since July 2025 were collected for this study, from which about 80,000 valid ice-cloud samples were selected after quality control. Comprehensive training and performance evaluations of the RTTOV-SCATT-DRL model were conducted under both strong typhoon and super typhoon scenarios. The results demonstrate that the optimized model substantially improves simulation accuracy across all channels. For the 183 GHz channels (CH32 – CH36), the RMSE was reduced to 4.98 – 14.33 K, while for the 325 GHz channels (CH41–CH44), the RMSE decreased to 9.54–13.32 K, corresponding to improvements of 56.3–69.7% relative to the standard RTTOV model. In ice-cloud-dominated regions, the RMSE reduction reached up to 69%, with the error of the 325.15 ± 6.6 GHz channel (CH44) decreasing from approximately 58 K to 12.19 K. After introducing effective radius as a key input variable, the OMB error distribution evolved from a strongly skewed pattern toward an approximately Gaussian distribution, effectively eliminating the systematic cold bias, satisfying the Gaussian error assumption required by variational data assimilation systems. Meanwhile, the optimized model fully preserves the computational efficiency of RTTOV, with an average computational cost of only about 0.2 ms per sample, making it suitable for large-scale operational simulations and real-time retrieval applications.

The major innovations of this study can be summarized in three aspects. First, TD3 reinforcement learning was applied for the first time to the end-to-end optimization of ice-cloud bulk scattering parameters in RTTOV-SCATT, using real satellite-observed TBs as the reward signal and thereby avoiding the dependence of conventional supervised learning approaches on particle-shape labels. Second, a differentiable scattering surrogate model was developed, enabling a complete gradient propagation pathway from observed TBs to scattering parameters. Third, effective radius derived from the RH Gamma distribution was introduced as a physically meaningful input variable, significantly enhancing the model's capability to represent forward-scattering effects associated with large particles, particularly for the high-frequency 325 GHz channels. These innovations go beyond previous machine-learning approaches that primarily focused on computational acceleration or individual particle parameterizations, providing a practical and physically consistent solution for high-frequency ice-cloud scattering optimization.

Despite these advances, several limitations remain. First, the training and validation datasets were primarily derived from tropical and subtropical typhoon environments, and the generalization capability of the model to non-convective ice-cloud

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

regimes, such as midlatitude stratiform clouds and polar stratospheric clouds, requires further investigation. Second, the current framework mainly optimizes graupel-related scattering parameters and does not simultaneously constrain the scattering characteristics of ice crystals and snow particles at higher frequencies. Finally, although effective radius serves as an effective proxy for particle morphology and significantly improves simulation accuracy, it cannot fully represent detailed microphysical characteristics of natural ice particles, such as aspect ratio and particle orientation.

Future work will focus on incorporating the optimized RTTOV-SCATT-DRL scattering module into the AWS satellite data assimilation system and conducting cycling assimilation experiments to quantitatively assess the impact of improved ice-cloud scattering simulations on AWS data assimilation performance. Such efforts will provide more reliable technical support for the utilization of AWS observations and future satellite missions carrying terahertz ice-cloud sounding instruments.

In summary, this study demonstrates that the tight integration of deep reinforcement learning with a physically based RT model can effectively overcome the inherent limitations of traditional bulk scattering parameterizations. The proposed RTTOV-SCATT-DRL framework provides a highly accurate and computationally efficient forward operator for next-generation terahertz ice-cloud remote sensing, with significant scientific value and strong potential for operational applications.

Code and data availability

The source code and input datasets used in this study are publicly archived on Zenodo at https://doi.org/10.5281/zenodo.21290427 (Chen et al., 2026).

Author contributions

KC designed the research framework, determined the technical route and provided overall academic guidance for this work. YY completed the main body of the manuscript, conducted data processing and analysis, and implemented iterative optimization and function upgrading of the program codes. FM is the original developer of the core code, and participated in manuscript reviewing and proofreading.

Competing interests

The contact author has declared that neither they nor their co-authors have any competing interests.

Disclaimer

Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

785 **Acknowledgements**

We gratefully acknowledge EUMETSAT for providing AWS L1B observations and the RTTOV model, and ECMWF for the ERA5 reanalysis datasets. This work was supported by our university's high-performance computing platform. The authors also appreciate the reviewers' insightful comments.

Financial support

790 This study was jointly supported by National Natural Science Foundation of China (42275141, 62271218, and 42276183).

Review statement

The review statement will be added by Copernicus Publications listing the handling editor as well as all contributing referees according to their status anonymous or identified.

References

795 Baum, B. A., Yang, P., Nasiri, S., Heidinger, A. K., Heymsfield, A. J., and Li, J.: Bulk Scattering Properties for the Remote Sensing of Ice Clouds. Part III: High-Resolution Spectral Models from 100 to 3250 cm^{-1}, J. Appl. Meteorol. Climatol., 46, 423–434, https://doi.org/10.1175/JAM2473.1, 2007.

Blackwell, W. J., et al.: An overview of the TROPICS NASA earth venture mission, Q. J. R. Meteorol. Soc., 144, 16–26, https://doi.org/10.1002/qj.3210, 2018.

800 Chen, K., Yao, Y., and Meng, F.: Deep Reinforcement Learning - Driven Optimization of Ice Cloud Bulk Scattering Properties in RTTOV - SCATT Using Arctic Weather Satellite Observations, Zenodo [data set], https://doi.org/10.5281/zenodo.21290427, 2026.

Duncan, D. I., Bormann, N., Dahoui, M., et al.: Assessment of the Arctic Weather Satellite in NWP, ECMWF Tech. Memo., Reading, UK, https://www.ecmwf.int, 2025.

805 Ekelund, R., Eriksson, P., and Pfreundschuh, S.: Using passive and active observations at microwave and sub-millimetre wavelengths to constrain ice particle models, Atmos. Meas. Tech., 13, 501–520, https://doi.org/10.5194/amt-13-501-2020, 2020.

Eriksson, P., Ekelund, R., Mendrok, J., et al.: A general database of hydrometeor single scattering properties at microwave and sub-millimetre wavelengths, Earth Syst. Sci. Data, 10, 1301–1326, https://doi.org/10.5194/essd-10-1301-2018, 2018.

810 Eriksson, P., Emrich, A., Kempe, K., et al.: The Arctic Weather Satellite radiometer, Atmos. Meas. Tech., 18, 4709–4729, https://doi.org/10.5194/amt-18-4709-2025, 2025.

https://doi.org/10.5194/egusphere-2026-4021
Preprint. Discussion started: 24 August 2026

Evans, K. F., and Stephens, G. L.: Microwave radiative transfer through clouds composed of realistically shaped ice crystals. Part II: Remote sensing of ice clouds, J. Atmos. Sci., 52, 2058 – 2072, https://doi.org/10.1175/1520-0469(1995)052<2058:MRTTCC>2.0.CO;2, 1995.

Fujimoto, S., van Hoof, H., and Meger, D.: Addressing Function Approximation Error in Actor-Critic Methods, in: Proceedings of the 35th International Conference on Machine Learning, Stockholmsmässan, Stockholm, Sweden, 1587–1596, 2018.

Geer, A. J., and Baordo, F.: Improved scattering radiative transfer for frozen hydrometeors at microwave frequencies, Atmos. Meas. Tech., 7, 1839–1860, https://doi.org/10.5194/amt-7-1839-2014, 2014.

Geer, A. J., et al.: Bulk hydrometeor optical properties for microwave and sub-millimetre radiative transfer in RTTOV-SCATT v13.0, ECMWF Tech. Memo., Reading, UK, 2021.

Hong, G., Heygster, G., Miao, J., and Kunzi, K.: Detection of tropical deep convective clouds from AMSU-B water vapor channels measurements, J. Geophys. Res.-Atmos., 110, D13205, https://doi.org/10.1029/2004jd004949, 2005.

Huang, Z., Lu, H., Ma, Z., Shi, Y., Han, Y., Hu, H., and Yang, J.: A nonspherical cloud scattering database using aggregates of roughened bullet rosettes model for the advanced radiative transfer modeling System (ARMS), Adv. Atmos. Sci., 42, 1483–1498, https://doi.org/10.1007/s00376-024-4117-7, 2025.

Jiménez, C., et al.: Performance simulations for a submillimetre-wave satellite instrument to measure cloud ice, Q. J. R. Meteorol. Soc., 133, 129–149, https://doi.org/10.1002/qj.103, 2007.

Li, Z. T., Han, W., Xie, H. J., Xu, X. Z., Bi, L., Sun, X. Y., and Li, H.: PhySCAT-Net: A Physics-Informed Deep Learning Framework for Optimizing Hydrometeor Bulk Scattering Properties Using Satellite Observations, J. Adv. Model. Earth Syst., https://doi.org/10.1029/2025MS004218, 2025.

Lillicrap, T. P., Hunt, J. J., Pritzel, A., et al.: Continuous control with deep reinforcement learning, arXiv preprint, arXiv:1509.02971, 2015.

Liu, L., Weng, C. S., Li, S. L., Hu, S., Ye, J., Dou, F. L., and Shang, J.: Review of Terahertz Passive Remote Sensing of Ice Clouds, Adv. Earth Sci., 35, 1211–1221, 2020.

Mangla, R., J, I., Chambon, P., and Mahfouf, J.: First steps towards an all-sky assimilation framework for tropical cyclone event over Bay of Bengal region: Evaluation and assessment of GMI radiances, Atmos. Res., 257, 105564, https://doi.org/10.1016/j.atmosres.2021.105564, 2021.

Saunders, R., Hocking, J., Turner, E., et al.: An update on the RTTOV fast radiative transfer model, Geosci. Model Dev., 11, 2717–2737, https://doi.org/10.5194/gmd-11-2717-2018, 2018.

Sieron, S. B., Zhang, F., Clothiaux, E. E., Zhang, L. N., and Lu, Y.: Representing Precipitation Ice Species With Both Spherical and Nonspherical Particles for Radiative Transfer Modeling of Microphysics-Consistent Cloud Microwave Scattering Properties, J. Adv. Model. Earth Syst., 10, 1011–1028, https://doi.org/10.1002/2017MS001226, 2018.

Stegmann, P. G., et al.: A deep learning approach to fast radiative transfer, J. Quant. Spectrosc. Radiat. Transf., 280, 108082, https://doi.org/10.1016/j.jqsrt.2022.108082, 2022.