All-Sky Temperature and Humidity Retrieval from the MWRI-RM Onboard the FY-3G Satellite

Minghua Liu[1], Wei Han[2, 3], Yunfan Yang[4, 5], Haofei Sun[4, 5], Ruoying Yin[2]

[1]School of Atmospheric Physics, Nanjing University of Information Science and Technology, Nanjing, 210044, China

[2]CMA Earth System Modeling and Prediction Centre (CEMC), China Meteorological Administration, Beijing, 100081, China

[3]State Key Laboratory of Severe Weather (LaSW), Chinese Academy of Meteorological Sciences, China Meteorological Administration, Beijing, 100081, China

[4]State Key Laboratory of Atmospheric Boundary Layer Physics and Atmospheric Chemistry (LAPC), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, 100029, China

[5]Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, 100029, China

Correspondence to: Wei Han (hanwei@cma.gov.cn)

Abstract. To investigate the application of deep learning in satellite remote sensing, this study employs brightness temperature observations from the remapped Micro-Wave Radiation Imager-Rainfall Mission (MWRI-RM) onboard the Fengyun-3G (FY-3G) satellite as input data, while temperature and relative humidity profiles (ranging from 1000 hPa to 100 hPa) obtained from ERA5 reanalysis data are used as label data. An Advanced Residual Convolutional Neural Network (AR-CNN) model was developed to retrieve atmospheric temperature and relative humidity profile data. The results show that: (1) The retrieval of temperature profiles achieves a root-mean-square error (RMSE) of approximately 1.24 K, and the RMSE for relative humidity profiles is 12.98%. (2) A comparison between predicted and labeled samples reveals consistent results for temperature retrieval but some discrepancies in extreme high and low humidity regions, suggesting the need for further refinement. (3) Gradient-based analyses and perturbation experiments confirm that 118 GHz oxygen channels are critical for mid-to-upper tropospheric temperature (500-200 hPa), indirectly impacting upper-level humidity (200-100 hPa) through thermal coupling, while 183 GHz water vapor channels dominate lower-to-mid tropospheric moisture retrievals (1000-500 hPa) and constrain temperature via moisture-radiation feedbacks. (4) Additional channel ablation experiments demonstrate that channels with smaller frequency offsets mainly affect upper atmospheric layers, whereas larger-offset channels have stronger impacts on lower layers, supporting the spectral contribution patterns identified in previous studies. These findings highlight the model's ability to capture temperature-humidity coupling and confirm the complementary roles of 118 GHz and 183 GHz channels in improving vertical profile retrievals.

1. Introduction

In atmospheric science, temperature, relative humidity, and cloud liquid water content profiles are crucial for characterizing atmospheric thermodynamics and dynamics. Real-time global monitoring of these parameters is essential for weather forecasting, climate research, and meteorological support (Liu et al., 2010; Ebell et al., 2013; He et al., 2017). With continuous advancements in data assimilation and numerical weather prediction technologies, the quality requirements for meteorological data, including temperature and humidity measurements, have become increasingly stringent (Xu et al., 2024). Traditional radiosonde observations from meteorological stations, while valuable, suffer from low spatiotemporal resolution due to geographical constraints, often failing to meet modern operational demands (Yao et al., 2022). This limitation is particularly pronounced in remote regions such as polar areas, oceans, and deserts.

Satellite remote sensing technology has emerged as a powerful solution for global atmospheric monitoring. Microwave remote sensing, with its unique ability to penetrate clouds and operate independently of solar illumination, provides enhanced spatial sampling density under all-sky conditions (Wang et al., 2010; Duan et al., 2017). This capability translates to improved spatiotemporal resolution for retrieving atmospheric variables like temperature and humidity profiles, overcoming the coverage gaps of conventional observations (Rosenkranz et al., 2001; Zhu et al., 2001). The development of atmospheric profile retrieval algorithms has evolved through decades of research. King's (1958) pioneering work on infrared-based temperature profiling laid the foundation for subsequent advancements. Current methodologies for atmospheric retrieval primarily fall into two categories: statistical regression and physical approaches. Statistical methods establish empirical radiance-parameter relationships using large datasets, offering computational efficiency but limited physical representation (Cao et al., 2021). For example, Wang et al. (2009) developed a log-linear algorithm for oceanic CWV retrieval using TRMM TMI microwave data, achieving bias < 0.7 kg/m² and RMSE 2.5 kg/m² against radiosondes with minimal ancillary data dependence. Physical methods, grounded in radiative transfer theory, provide rigorous solutions but face challenges in computational complexity and error sensitivity (Wu et al., 2005).

Recent advancements in artificial intelligence have transformed atmospheric retrieval, particularly through microwave-focused machine learning. For instance, Xia et al. (2023) used an AutoML

framework with 50M+ GPS samples to retrieve land PWV from AMSR2 data, achieving RMSE 3.1 mm across diverse surfaces. Du et al. (2015) developed an AMSR2-based statistical algorithm for PWV (RMSE 4.7 mm vs. GPS) and surface temperature (RMSE < 3.5 K), demonstrating all-weather utility. Cross-sensor synergies enhance microwave retrieval accuracy. Ji et al. (2017) combined AMSR-E microwave and MODIS data to retrieve TPW, optimizing emissivity and terrain effects for 0.05° resolution (RMSE 4.18 mm, r=0.95), suitable for cloudy conditions. These approaches highlight data-driven models for robust atmospheric parameter retrieval.

A significant advancement in observational capabilities came with China's April 2023 launch of the FY-3G satellite, equipped with the upgraded Micro-Wave Radiation Imager-Rainfall Mission (MWRI-RM). This instrument integrates 26 channels across oxygen (54/118 GHz) and water vapor (183 GHz) absorption bands, significantly enhancing atmospheric profiling capacity while maintaining all-weather operation (Gu et al., 2023; Zhang et al., 2023). Building on these technological developments, this study presents an Advanced Residual Convolutional Neural Network (AR-CNN) for retrieving temperature and relative humidity profiles from FY-3G/MWRI-RM observations. Our methodology employs brightness temperatures from all available channels, supplemented by standardized surface parameter preprocessing (Cadeddu et al., 2009; Cai et al., 2020). Key innovations include channel importance analysis for 118/183 GHz bands, model interpretability investigations. The relative humidity parameter selection is carefully justified through error propagation analysis considering its temperature dependence, with comparative validation against independent datasets ensuring methodological robustness.

2. Data

2.1 Instrument Overview

The primary purpose of the MWRI-RM is to collect microwave radiation data from both land and sea surfaces globally, serving the needs of global precipitation measurement. It employs a conical scanning system, featuring 1648 scan lines with 492 sampling points on each line. The MWRI-RM is equipped with 8 feed sources arranged in two rows. The first row consists of channels at 10 GHz, 18.7 GHz, 23.8 GHz, 36 GHz, and 89 GHz, where the 18.7 GHz and 23.8 GHz channels share a common feed source. The second row comprises channels at 54 GHz, 118 GHz, 166 GHz, and 183 GHz. The feed sources in

different rows have angular differences in their observation geometries. The MWRI-RM functions across 17 frequency points, encompassing 26 channels, with certain frequencies featuring dual-polarization capabilities. Detailed channel specifications are delineated in Table. 1.

Table 1: Channel information of FY-3G/MWRI-RM (Data source: FY-3G_L1 Data User's Guide_MWRI-RM_20230625.pdf (nsmc.org.cn), Gu et al., 2023)

Center Frequency (GHz)	Polarization Mode	Bandwidth (MHz)	Frequency Stability (MHZ)	IFOV (km×km)	Main Applications
10.65	V,H	180	10	21×35	Intense precipitation over the ocean, land surface products.
18.7	V,H	200	10	14×23	Precipitation over the ocean, land surface products.
23.8	V,H	400	15	13×21	Total water vapor over oceans.
36.5	V,H	900	20	9×15	Precipitation, land surface products.
50.30	V,H	400	25	7×11	Precipitation, drizzle, Snowfall, height and thickness of the melting layer.
52.61	V,H	400	25	7×11	
53.24	V,H	400	25	7×11	
53.75	V,H	400	25	7×11	
89.0	V,H	3000	25	5×8	Precipitation and snowfallover ocean and land areas, land surface products.
118.7503±3.2	V	2x500	25	4×7	
118.7503±2.1	V	2x400	25	4×7	Precipitation, drizzle, Snowfall, height and thickness of the melting layer.
118.7503±1.4	V	2x400	25	4×7	
118.7503±1.2	V	2x400	25	4×7	
165.5±0.75	V	2x1350	30	4×6	Precipitation and snowfall overland.
183.31±2.0	V	2x1500	30	4×7	
183.31±3.4	V	2x1500	30	4×7	Snowfall, cloud ice.
183.31±7	V	2x2000	30	4×7	

2.2 Sample Construction

This study utilizes MWRI-RM Level 1 products from October 23 to November 23, 2023 (UTC), containing brightness temperatures from all 26 channels. The dataset integrates sea surface observations with preprocessed static parameters obtained through decoding and quality checks. Oceanic data were exclusively selected to reduce computational complexity in the preprocessing stage. After remapping using an enhanced BGI algorithm that aligns spatial resolutions and footprint centers (Chen et al., 2024), all channels were resampled to the 23.8 GHz channel's resolution of 13 km×21 km. ERA5 reanalysis data (Hersbach et al., 2023), providing temperature profiles across 27 levels (1000-100 hPa) at 0.25° resolution, served as training labels. To address spatial resolution differences, ERA5 temperature profiles were spatially interpolated to satellite footprints using nearest-neighbor matching, with temporal alignment achieved through hourly binning of observation times.

The data were structured as 4D tensors (samples×height×width×channels): 13,100,161 samples each sized 5×5×26 channels. Corresponding ERA5 temperature/humidity profiles at sample centers were used as labels (Fig. 1). The dataset was partitioned chronologically into training (80%, 10,480,128 samples), testing (20%, 2,620,032), with 20% of training data (2,096,025 samples) allocated for validation.

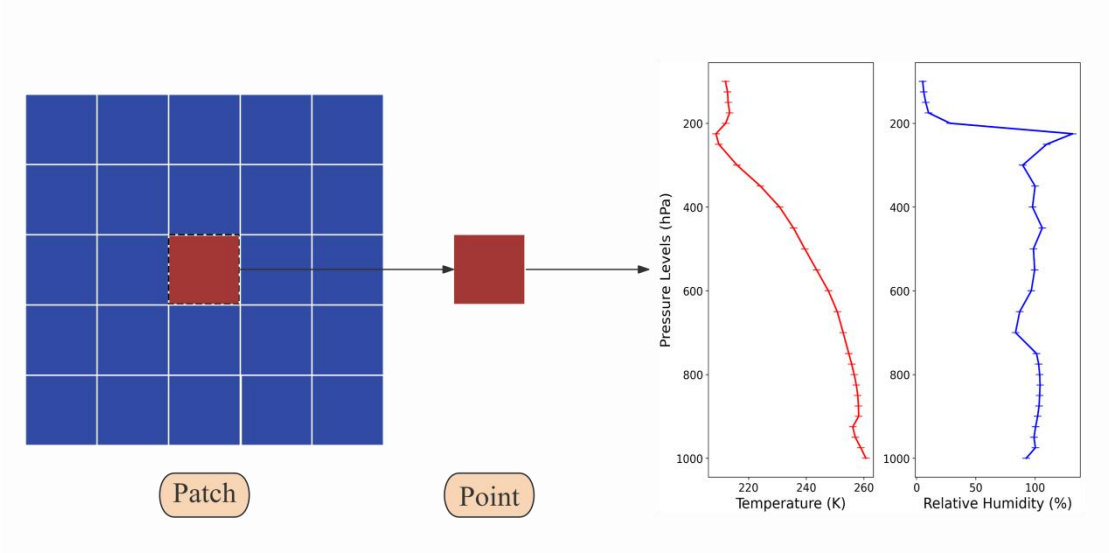

Figure 1: Illustration of retrieval from area to point

3 Retrieval Methods

3.1 AR-CNN Network Algorithm Principles

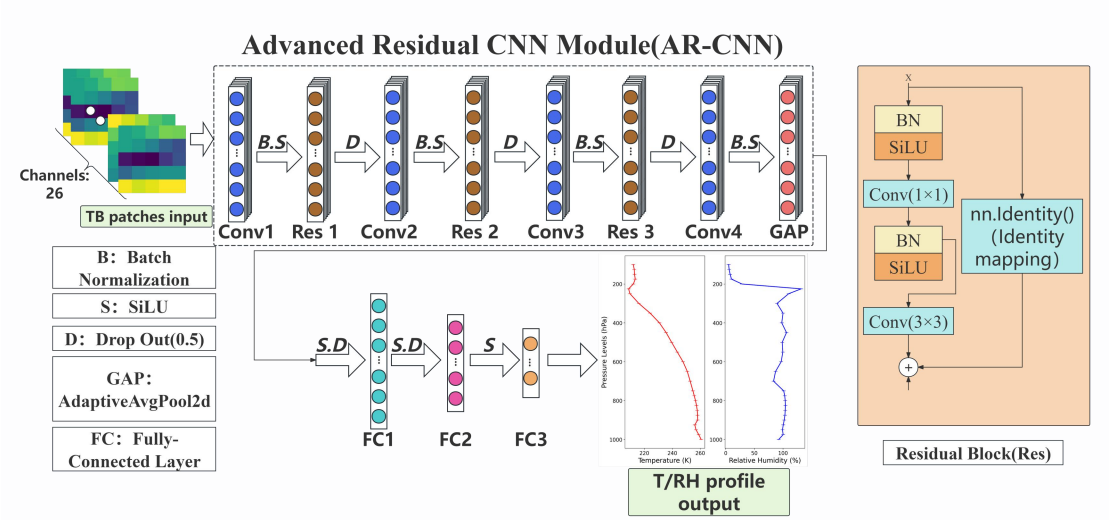

Figure 2: Network architecture for Advanced Residual CNN (AR-CNN)

In this study, an Advanced Residual Convolutional Neural Network (AR-CNN) was developed to retrieve temperature and humidity profiles (Fig. 2). The network architecture integrates convolutional layers, residual blocks, batch normalization, dropout, adaptive average pooling, and fully connected layers to efficiently extract features from 5×5 satellite brightness temperature images with 26 channels. SiLU activation, known for its smooth gradient properties and ability to avoid dead zones, is employed throughout the network (Elfwing et al., 2018). Residual blocks, equipped with skip connections, mitigate gradient vanishing in deep layers (He et al., 2016). A dropout layer with a rate of 0.25 is incorporated to prevent overfitting. The final convolutional layer generates a 512-channel feature map, which is normalized and compressed via adaptive average pooling to 1×1 per channel. The fully connected layers predict temperature and humidity profiles across 27 layers. An early stopping mechanism is implemented during training to enhance model generalization. The AR-CNN's design effectively captures multi-level features and improves training efficiency and generalization, making it well-suited for complex image classification tasks.

3.2 Retrieval Algorithm Validation

In this paper, bias and root-mean-squared error (RMSE) are chosen as evaluation metrics for the accuracy of atmospheric temperature and relative humidity profile retrieval. The specific formulas are as follows:

$$\text{Bias} = \frac{1}{N}\sum_{i=1}^{Ns}(x_i - y_i) \quad (1)$$

Where is the x_i retrieved value, y_i is the target value, and N is the number of samples.

Bias evaluates the systematic error of the model, representing the average difference between the predicted and true values. A bias of zero indicates no systematic deviation, while a positive or negative bias indicates that the model tends to overestimate or underestimate the target variable, respectively.

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{Ns}(x_i - y_i)^2} \quad (2)$$

Where x_i is the retrieved value, y_i is the target value, and N is the number of samples.

RMSE assesses the magnitude of errors between predicted and true values. It is sensitive to large errors due to the squaring of the differences. A lower RMSE indicates better model performance, with smaller deviations between the retrieved and target values.

4. Results

4.1 Temperature Retrieval

The retrieved temperature profiles exhibit height-independent consistency, with vertically uniform bias (0.11 K) and RMSE (1.24 K) distributions throughout the atmospheric column (Fig. 3a-b). This stability suggests minimal height-dependent variability in retrieval performance. Nevertheless, The observed vertical RMSE structure (0.97-1.98 K) primarily reflects the instrument's inherent sensitivity characteristics, as evidenced by the Jacobian weighting functions in Fig. 5(b) showing peak sensitivity at 500-700 hPa with diminishing response above 200 hPa. This vertical sensitivity pattern fundamentally constrains retrieval accuracy in the upper troposphere. Notably, the RMSE is relatively more pronounced in the lower troposphere around 850 hPa. This is likely attributable to the substantial surface influences, especially those stemming from the variability of sea surface temperature (SST). Such variability engenders complexity in the retrieval of temperature near the Earth's surface (Susskind et al., 2003; Stähli et al., 2013). The non-linear interactions between the surface and the atmosphere, in conjunction with the sensor's sensitivity to the lower atmospheric layers, might not have been adequately learned by the model.

With the progressive increase in altitude from 850 hPa to 500 hPa, the RMSE gradually dwindles, reaching a minimum value of approximately 1 K at 500 hPa. This decrease can be ascribed to the

relatively stable and homogeneous temperature field within the mid-troposphere, where the vertical mixing is less conspicuous compared with that in the lower and upper troposphere. Within this region, the model derives advantages from more consistent radiometric data, resulting in more accurate predictions.

Above 500 hPa, the RMSE commences to increase once more, reaching its zenith around 100 hPa. This increase is ascribed to the complex dynamics in the upper troposphere and the lower stratosphere, involving intense vertical motion, convective processes, along with the phase transitions of water vapor, clouds, and ice particles (Chen et al., 2003; Susskind et al., 2003). Furthermore, the sensitivity of the instrument declines at higher altitudes on the grounds of the reduced radiative energy emission and the limited sensitivity of the sensor in these regions (Sahoo et al., 2015), giving rise to lower signal - to - noise ratios and less informative data for temperature inversion.

Fig. 3(c)-(h) further validate these findings. The histogram presented in Fig. 4, which compares the counts of retrieved samples with those of labeled samples, discloses the difficulties in accurately retrieving extremely high and low temperatures. This challenge is prevalent in satellite remote sensing, especially in the retrieval of microwave channels, and might originate from data imbalance, in which extreme temperatures are inadequately represented in training datasets. Consequently, the model may exhibit a proclivity towards more common temperature ranges, thereby undermining its performance on extreme values. Moreover, complex meteorological interactions, such as those related to moisture content, pressure systems, and radiative transfer processes, might not be comprehensively captured by the input features of the model (Wang et al., 2023). The inherent variability and non-linearity of atmospheric processes, particularly during extreme weather events, introduce further complexity, thereby exacerbating the retrieval inaccuracies in the context of microwave channel data (Wang et al., 2021). To address these issues, we could employ advanced data preprocessing techniques, such as data augmentation or transfer learning, to enhance the representation of extreme temperatures in training datasets. Additionally, incorporating more comprehensive input features that account for complex meteorological interactions might improve the model's ability to handle extreme temperature retrieval.

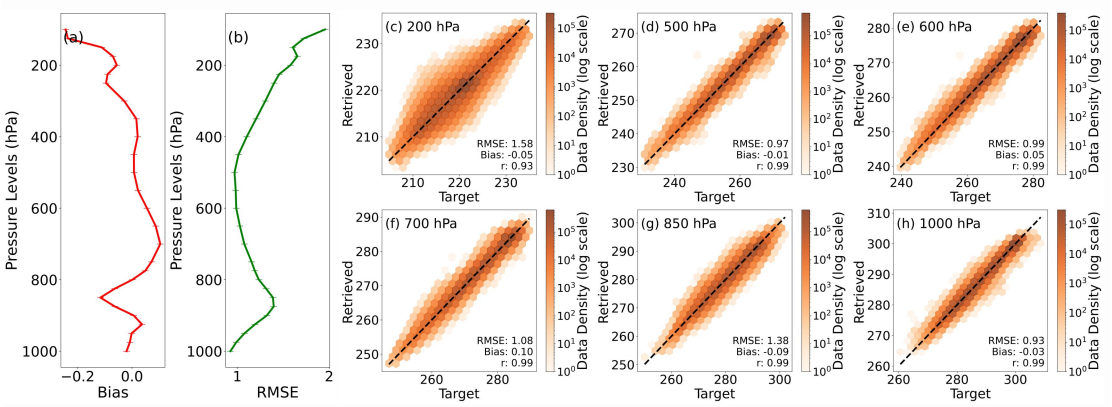

Figure 3: (a)-(b) Bias and RMSE of the retrieved temperature profile; (c)-(h) Two-dimensional histogram of retrieved temperature at specific pressure levels, where color intensity represents data point density (logarithmic binning). Darker colors indicate higher data counts in each bin, while lighter colors denote sparser distributions. A dashed black line (y=x) denotes the "1:1 consistency line" for perfect agreement between retrieved and target values. Pearson correlation coefficient (***r***) is shown for statistical assessment of linear agreement. (c) 200 hPa; (d) 500 hPa; (e) 600 hPa; (f) 700 hPa; (g) 850 hPa; (h) 1000 hPa.

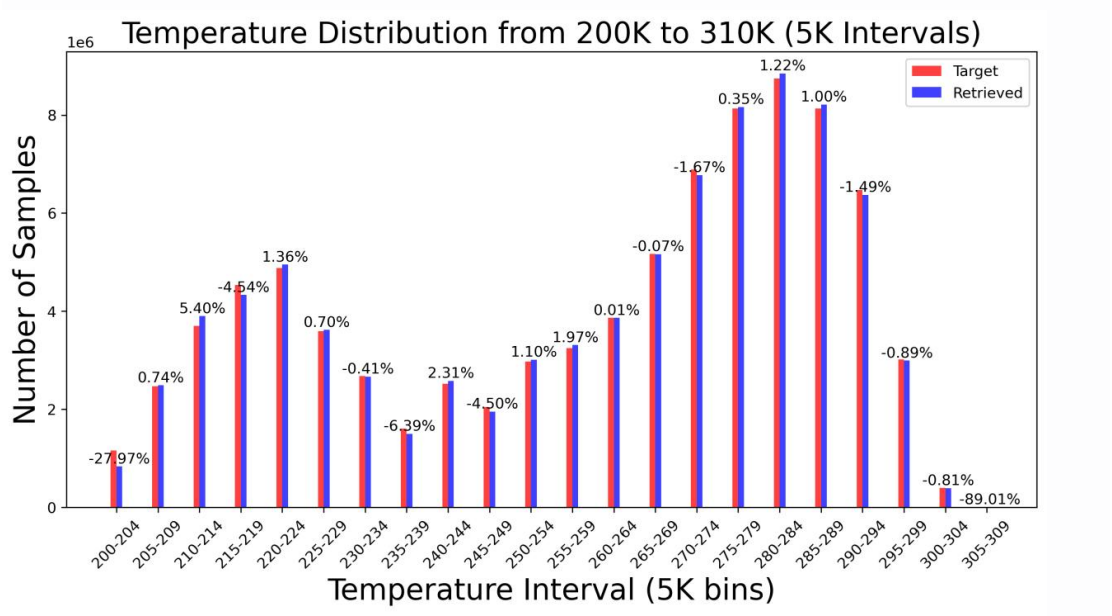

Figure 4: Statistics of the number of retrieved and labeled samples in the test set (5K bins)

The analysis of the Jacobians for temperature, derived using the Radiative Transfer for TOVS (RTTOV) model customized for the MWRI-RM sensor, was conducted via gradient backpropagation to evaluate individual channel contributions. Results (extracting weights from the first convolutional layer parameters (model.conv1), summing absolute weights across output channels and spatial dimensions, and normalizing for comparison) indicate that the highest weight in temperature profile retrieval resides within the 118.75±1.2 GHz frequency band (Fig. 5(a)), aligning with the peak sensitivity of temperature Jacobians (Fig. 5(b)). This highlights the model's utilization of the oxygen absorption line

near 118 GHz, a spectral feature with significant sensitivity to mid-to-upper tropospheric temperature gradients (500-200 hPa) and boundary layer thermal structure (1000-850 hPa). Oxygen rotational transitions at 118.75±1.2 GHz dominate radiative transfer in the mid-tropospheric lapse rate region (500-300 hPa) and lower stratosphere-upper troposphere transition (200-100 hPa), which are critical for atmospheric dynamics (Gasiewski et al., 1989, 1993; Xiao et al., 2023; Li et al., 2025).

Moreover, the 183.31±7 GHz channel, located within the water vapor absorption band, contributes secondarily to temperature profile refinement through moisture-radiation interaction mechanisms. This channel's sensitivity to water vapor variations in the boundary layer and lower troposphere (1000-700 hPa) influences radiance measurements via absorption/emission processes, providing cross-constraints for temperature retrievals in regions with high moisture content (Xiao et al., 2023; Li et al., 2025). Water vapor's impact on radiative budgets in these layers enhances the model's ability to resolve temperature-humidity coupling, particularly in tropical and maritime environments.

To assess model stability and physical interpretability, we perturbed brightness temperatures of the 118.75±1.2 GHz and 183.31±7 GHz channels (Fig. 6). Results show the 118 GHz channel primarily affects temperature retrievals in the mid-to-upper troposphere (500-100 hPa), with discernible but weaker signals extending to the boundary layer (850-1000 hPa). The 183 GHz channel demonstrates significant influence on lower tropospheric temperatures (1000-700 hPa), consistent with its dominant role in boundary layer moisture constraint. These findings align with Jacobian profiles: the 118 GHz channel's oxygen sensitivity peaks in the mid-troposphere, while the 183 GHz channel's water vapor absorption introduces localized temperature adjustments via humidity-radiation interactions.

Moreover, the channel ablation experiments (Fig. 7) further corroborate these sensitivities. When channels 21-26 are removed individually, the resulting bias and RMSE patterns reveal a clear vertical hierarchy: channels with smaller frequency offsets (particularly around the 118 GHz band) exert a more pronounced influence on higher atmospheric layers, whereas those with larger offsets predominantly affect lower layers. Specifically, the 118.75±1.2 GHz channel, with its relatively small offset, impacts the mid-to-upper troposphere significantly, while the 183.31±7 GHz channel, characterized by a larger offset, mainly modulates the lower troposphere. These distinct vertical sensitivities align well with prior findings by Xiao et al. (2023) and Li et al. (2025).

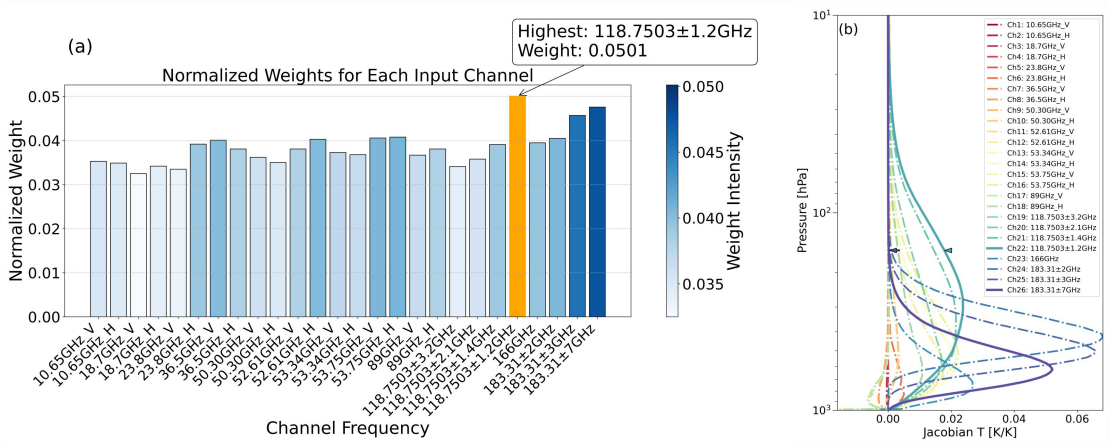

Figure 5: (a) Weighting map for temperature profile retrieval channels; (b) Jacobian for temperature, obtained using the RTTOV model tailored for the MWRI-RM sensor. Solid lines denote Channel 22 (118.75±1.2 GHz) and 26 (183.31±7 GHz), while dash-dot lines represent other channels.

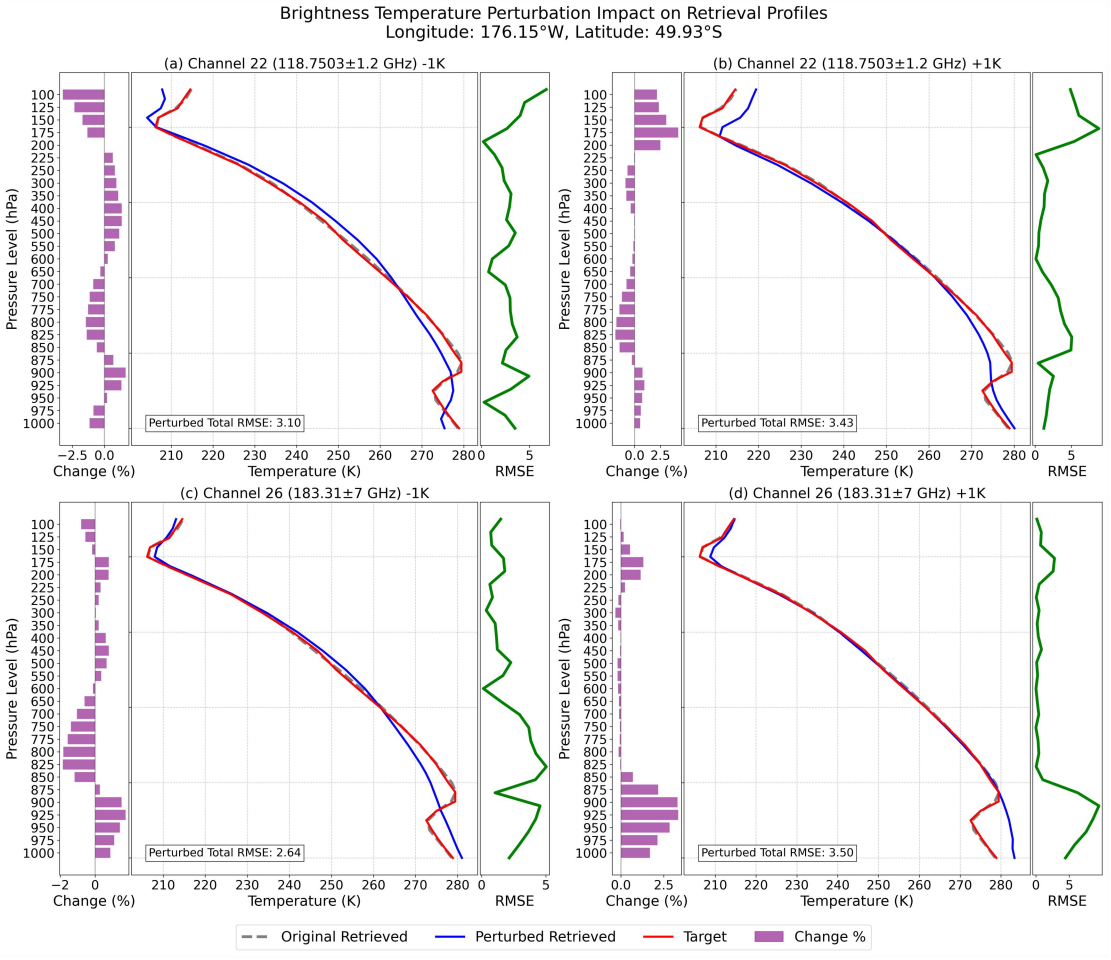

Figure 6: Single temperature profile perturbation experiment (±1K perturbations are introduced to the brightness temperatures of channels 22 and 26). Each subplot adopts a triple-axis composite layout: the left-hand purple bar chart reveals the vertical relative differences of the perturbed retrieval profile from the original profile (calculated as [(perturbed - original)/original]×100%); the central area presents the temperature profiles comparison (gray dashed line denotes the original retrieval result, blue solid line represents the perturbed retrieval result, and red solid line is the target profile); the right-hand green curve signifies the vertical RMSE between the perturbed retrieval profile and the target profile.

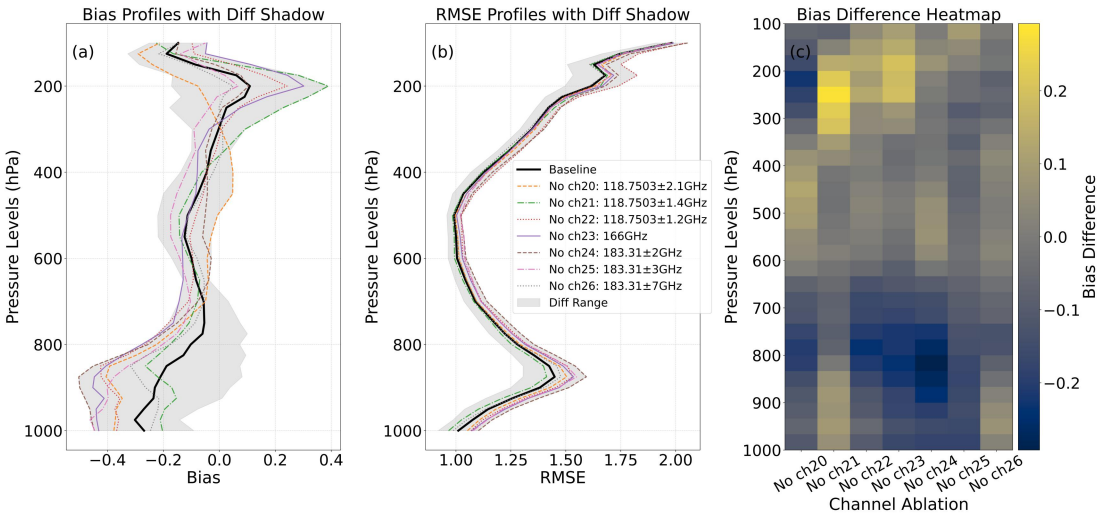

Figure 7: Results of channel ablation experiments illustrating the impact of removing individual channels (channels 21-26) on temperature profile retrieval accuracy. In (a), vertical bias profiles are shown relative to the baseline (black solid line). In (b), the corresponding RMSE profiles are presented. In (c), the bias difference heatmap displays the magnitude of deviations relative to the baseline.

4.2 Relative Humidity(RH) Retrieval

Across various atmospheric layers, the model demonstrated a relatively consistent performance, presenting a bias of 0.87% and a RMSE of 12.98%, as graphically depicted in Fig. 7(a)-(b). The consistent performance indicates the model's general stability in handling different atmospheric conditions, yet the non-zero bias and RMSE values also suggest room for improvement in its accuracy.

The analysis identified notable RMSE variations, particularly in the lower troposphere around 900 hPa. As height increases from 800 hPa upward, RMSE increases, peaking at around 200 hPa. This trend implies increasing complexity in humidity retrieval at higher altitudes, likely due to enhanced dynamic and turbulent influences in the mid-to-upper troposphere. In this region, convective transport of heat and moisture, along with cloud processes such as particle formation and phase changes, introduce significant uncertainty and alter local humidity distributions. These processes, especially in areas with deep convection or liquid water droplets and ice crystals, obscure the direct link between radiometric observations and humidity, challenging the model's ability to accurately retrieve humidity values solely from brightness temperatures (Burns et al., 1995, 1997; Blackwell et al., 2005; Zhang et al., 2021; Wang et al., 2010; Tan et al., 2015). Fig. 8(c)-(h) further validate these findings. Fig. 9 illustrates the suboptimal retrieval performance in low-humidity regions at low latitudes and high-humidity regions at high latitudes. This graphical representation clearly highlights the difficulties encountered in accurately retrieving extremely high and low relative humidity values.

Similar to the challenges faced in temperature profile retrieval, the issues in humidity retrieval can be traced back to data-related problems, such as data imbalance in the training dataset, and the complex non-linear relationships between the input variables and the humidity values in extreme conditions. To enhance retrieval accuracy, we could refine the deep learning model by incorporating high-resolution SST data and advanced surface flux models. Additionally, leveraging data augmentation techniques to expand the diversity of training samples, especially those representing extreme relative humidity conditions, may improve the model's generalization ability. Exploring the potential of hybrid neural network architectures that better capture non-linear ocean-atmosphere interactions could also be a promising avenue for future research.For maritime areas, these variations are primarily linked to SST and surface flux dynamics.

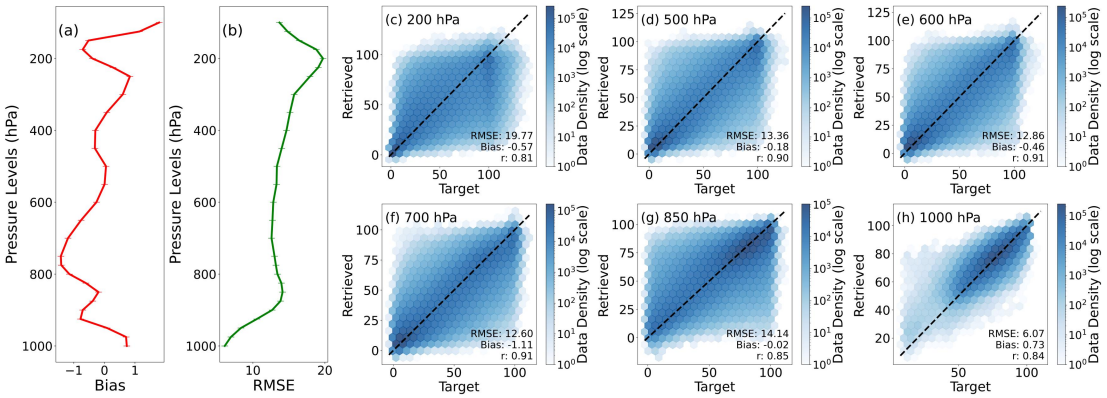

Figure 8: (a)-(b) Bias and RMSE of the retrieved RH profile; (c)-(h) Two-dimensional histogram of retrieved RH at specific pressure levels, where color intensity represents data point density (logarithmic binning). Darker colors indicate higher data counts in each bin, while lighter colors denote sparser distributions. A dashed black line (y=x) denotes the "1:1 consistency line" for perfect agreement between retrieved and target values. Pearson correlation coefficient (***r***) is shown for statistical assessment of linear agreement. (c) 200 hPa; (d) 500 hPa; (e) 600 hPa; (f) 700 hPa; (g) 850 hPa; (h) 1000 hPa.

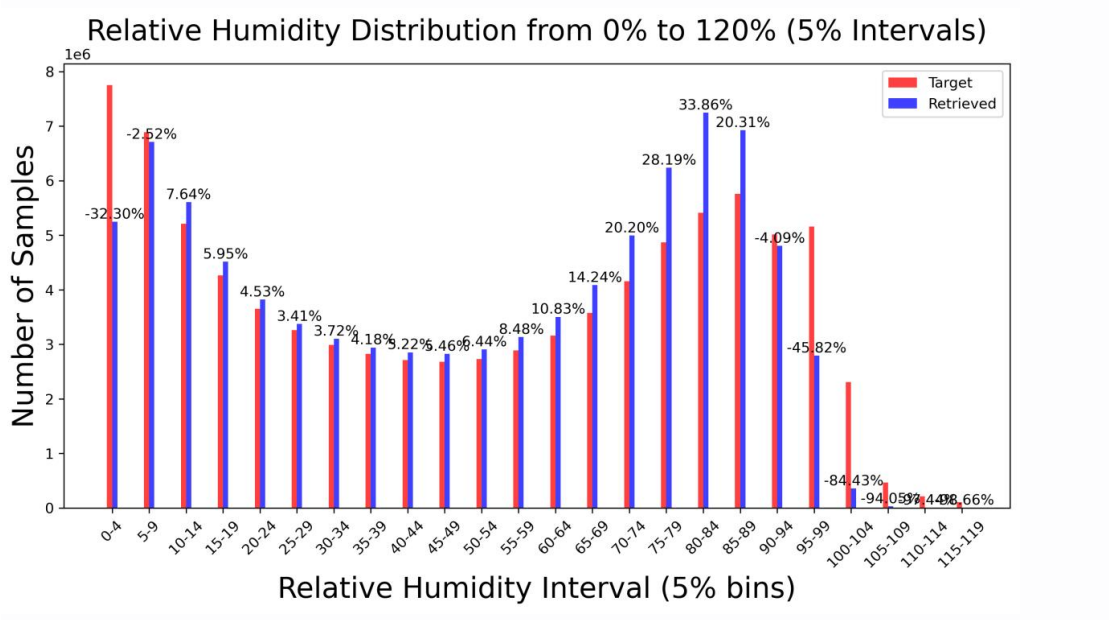

Figure 9: Statistics of the number of retrieved and labeled samples in the test set (5% bins)

Through gradient-based backpropagation analysis, we first characterize the physical mechanisms of channel contributions to humidity retrieval. The 183.31±7 GHz channel dominates atmospheric humidity profile retrieval, with peak sensitivity concentrated in the boundary layer and lower troposphere (1000-700 hPa) (Fig. 10(a,b)). This aligns with the direct absorption and emission of water vapor at this frequency, enabling precise resolution of moisture gradients in convective and maritime environments (He et al., 2017; Yao et al., 2022; Xiao et al., 2023; Li et al., 2025). In contrast, the 118.75±1.2 GHz channel is primarily sensitive to mid-to-upper tropospheric temperature gradients (500-200 hPa) via oxygen absorption lines (Xiao et al., 2023; Li et al., 2025), with negligible direct humidity sensitivity (Fig. 10(b)). However, perturbation experiments (Fig. 11) reveal unexpected impacts of this channel on upper tropospheric humidity (200-100 hPa), alongside its dominant temperature adjustments. This apparent discrepancy can be attributed to temperature-humidity dynamical coupling, whereby upper tropospheric temperature retrievals from the 118 GHz channel influence moisture distribution through processes such as vertical advection and stability modifications. Additionally, weight diffusion in the convolutional layer architecture (e.g., model.conv1) may propagate temperature-sensitive features to adjacent humidity retrievals, representing a secondary effect of spatial feature mixing inherent to deep neural networks.

Moreover, the channel ablation experiments (Fig. 12) further corroborate and refine these sensitivity patterns. When removing channels 21-26 individually, we observe that the 183.31±7 GHz channel

(marked by a larger frequency offset) dominates lower-tropospheric humidity constraints. Notably, for channels in the vicinity of 183 GHz, those with smaller frequency offsets (although their impact is not strikingly pronounced), suggest a tentative tendency to influence higher tropospheric layers. In contrast, the 118.75±1.2 GHz channel exerts a more notable influence on upper-level humidity fields, presumably via indirect dynamical pathways. These findings align with prior results by Xiao et al. (2023) and Li et al. (2025), underscoring the distinct vertical contributions of different channels to humidity retrieval. However, our current model's lack of explicit physical constraints leads to certain deviations from traditional interpretations, as highlighted in earlier studies. Future work will focus on integrating physics-informed network designs to disentangle temperature-humidity coupling and enhance the interpretability of subtle vertical signals.

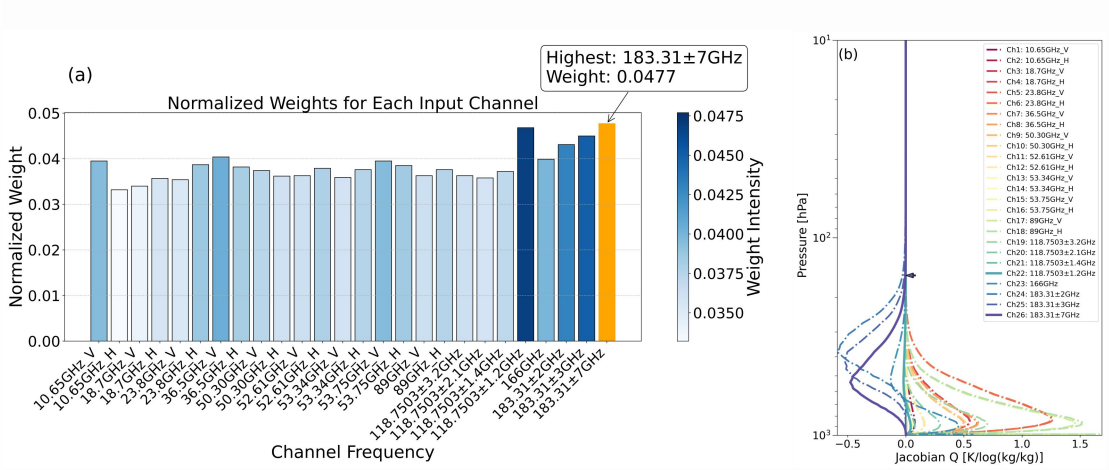

Figure 10: (a) Weighting map for RH profile retrieval channels; (b) Jacobian for relative humidity, obtained using the RTTOV model tailored for the MWRI-RM sensor. Solid lines denote Channel 22 (118.75±1.2 GHz) and 26 (183.31±7 GHz), while dash-dot lines represent other channels.

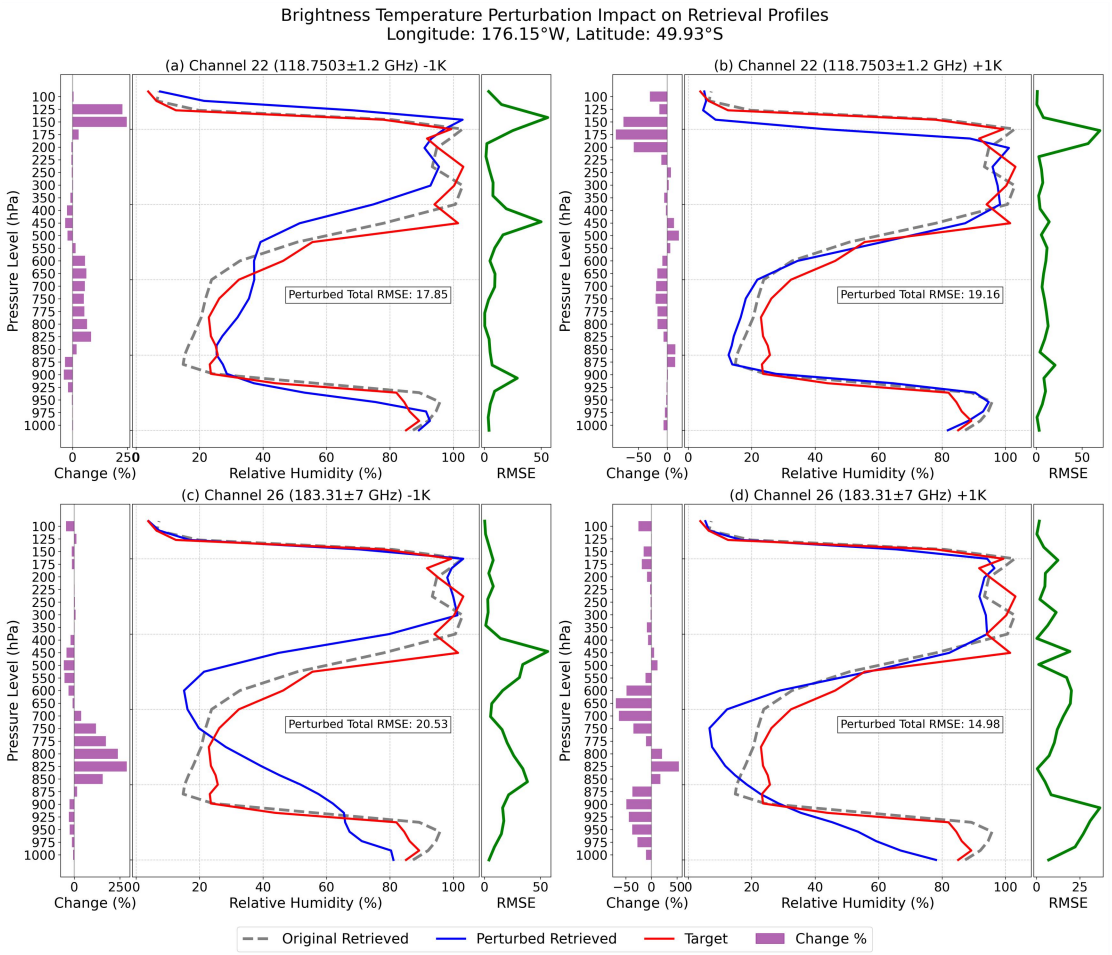

Figure 11: Single relative humidity profile perturbation experiment (±1K perturbations are introduced to the brightness temperatures of channels 22 and 26). Each subplot adopts a triple-axis composite layout: the left-hand purple bar chart reveals the vertical relative differences of the perturbed retrieval profile from the original profile (calculated as [(perturbed - original)/original] × 100%); the central area presents the relative humidity profiles comparison (gray dashed line denotes the original retrieval result, blue solid line represents the perturbed retrieval result, and red solid line is the target profile); the right-hand green curve signifies the vertical RMSE between the perturbed retrieval profile and the target profile.

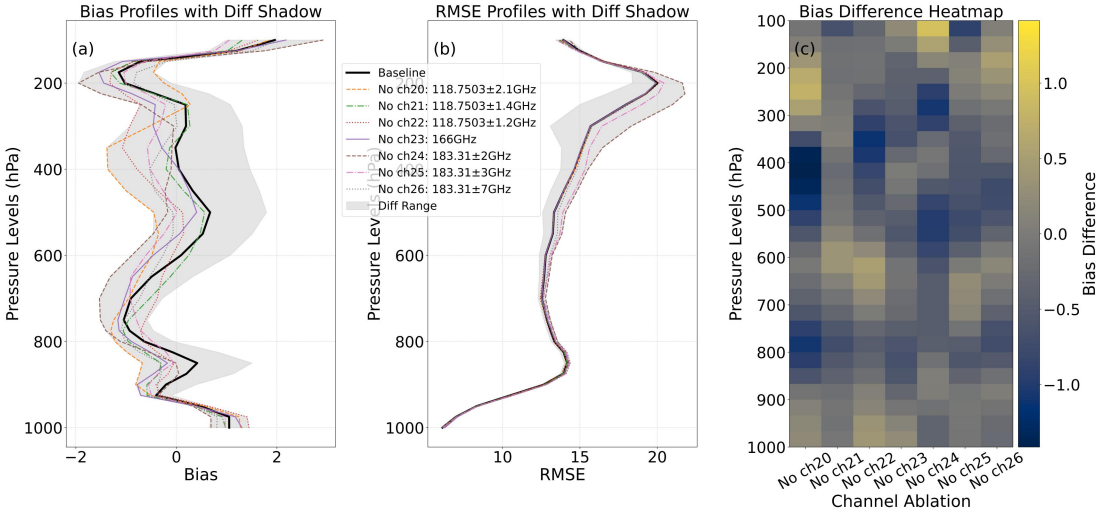

Figure 12: Results of channel ablation experiments illustrating the impact of removing individual channels

(channels 21-26) on relative humidity profile retrieval accuracy. In (a), vertical bias profiles are shown relative to the baseline (black solid line). In (b), the corresponding RMSE profiles are presented. In (c), the bias difference heatmap displays the magnitude of deviations relative to the baseline.

5. Conclusions

This study employs remapped MWRI-RM sea-surface-observed brightness temperature data, with ERA5 reanalysis temperature and relative humidity profile sets functioning as label data, to retrieve temperature and relative humidity profiles by means of an Advanced Residual Convolutional Neural Network (AR-CNN). The findings can be summarized as follows:

The retrieval of temperature profiles exhibits an RMSE of approximately 1.24 K, while the RMSE for relative humidity profile retrieval is 12.98%. These values demonstrate a high level of accuracy in both temperature and humidity retrieval, indicating the effective performance of the model in capturing atmospheric variables.

Statistical comparisons between the retrieved and label samples reveal that, for temperature profiles, the counts exhibit substantial consistency, demonstrating the model's robust capability to generalize across diverse temperature conditions. However, discrepancies were observed in regions with extreme humidity (both extremely high and low), implying that performance in such environments may be less accurate, likely due to the challenges of resolving water vapor content under extreme conditions. Consequently, further refinement of the model and the adoption of data augmentation techniques may be necessary to improve performance in these scenarios.

A comprehensive analysis integrating channel weights and Jacobian matrices validated the model's physical consistency in leveraging spectral information, showing that the 118 GHz band excels in mid-to-upper tropospheric temperature retrieval (500-200 hPa) with indirect upper-tropospheric (200-100 hPa) humidity adjustments via temperature-humidity dynamical coupling, while the 183 GHz band dominates boundary layer to mid-tropospheric moisture profiling (1000-500 hPa) through direct water vapor absorption/emission and provides humidity-driven cross-constraints for temperature estimates. Perturbation experiments revealed that the 118 GHz channel influences upper-tropospheric temperature-humidity covariance, likely due to both physical temperature-driven moisture redistribution and model weight diffusion in deep learning layers, while the 183 GHz channel's impact on lower tropospheric humidity (850-700 hPa) aligns with its strong water vapor sensitivity. Moreover,

channel ablation experiments further confirmed these vertical sensitivities, demonstrating that channels with smaller frequency offsets primarily affect upper atmospheric layers, while those with larger offsets exert stronger impacts on lower layers. These results highlight the importance of spectral synergy in improving vertical profile retrievals. Future work will focus on introducing explicit physics-based constraints into the network architecture to further enhance layer-specific interpretability and retrieval accuracy.

Appendix A.

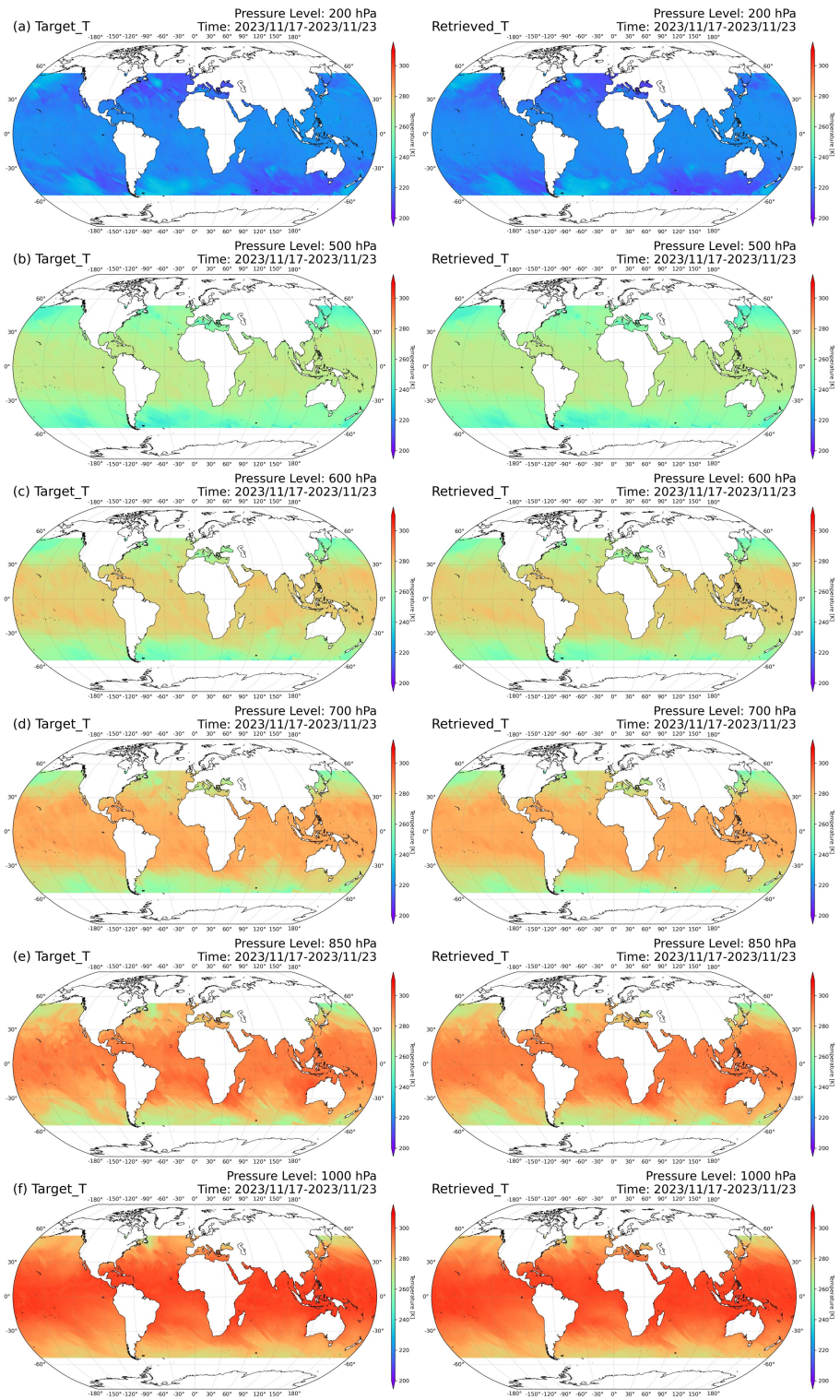

Fig. A1: Global Comparison Map of the Retrieved Temperature at Specific Pressure Levels (a)200hPa; (b)500hPa; (c)600Pa; (d)700hPa; (e)850hPa; (f)1000hPa

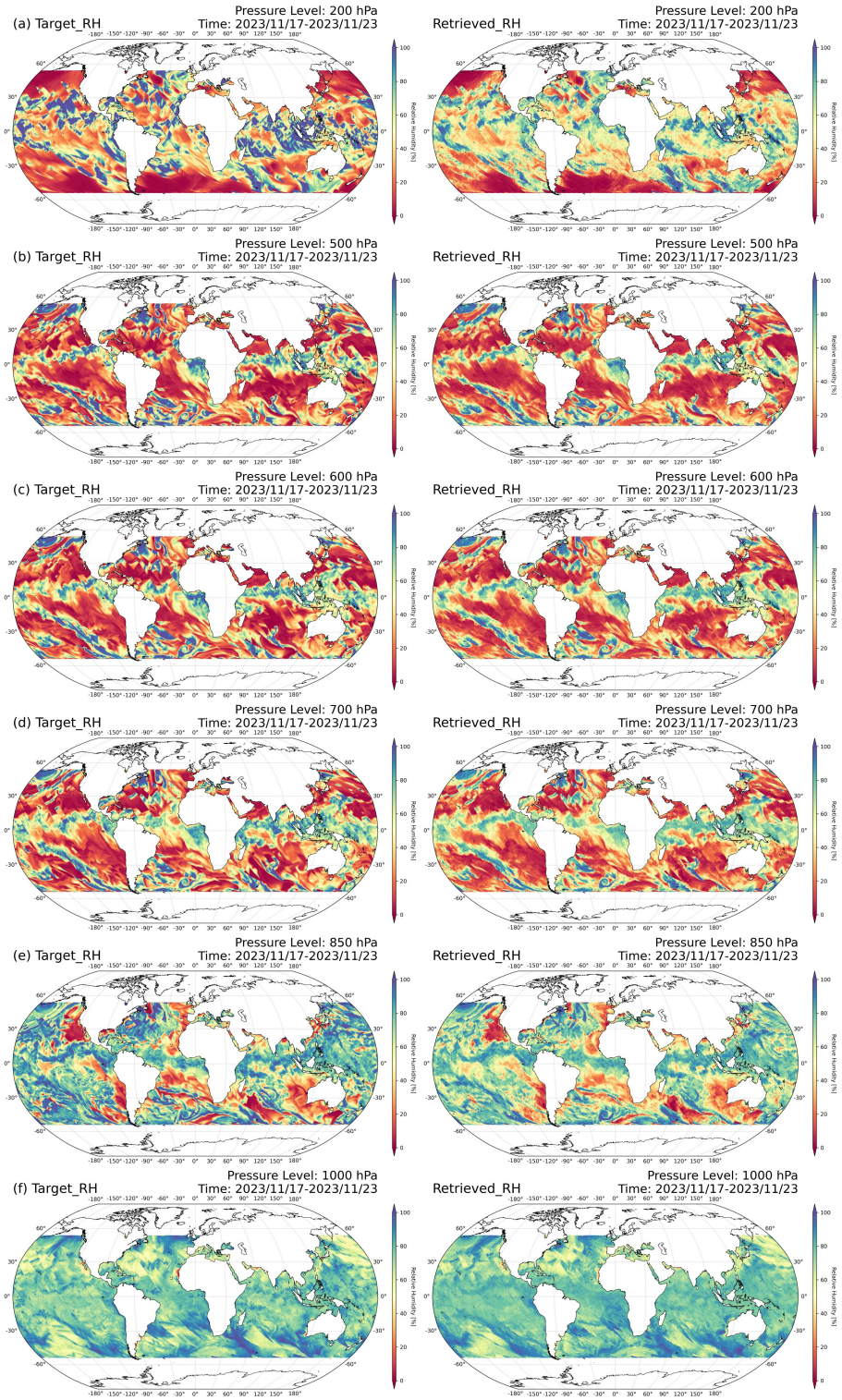

Fig. A2: Global Comparison Map of the Retrieved Relative Humidity at Specific Pressure Levels (a)200hPa; (b)500hPa; (c)600Pa; (d)700hPa; (e)850hPa; (f)1000hPa

Model	Train_RMSE	Val_RMSE	Test_RMSE	Epochs

MLP	T：1.52 K RH：15.33 %	T：1.61 K RH：16.87 %	T：1.77 K RH：17.46 %	105
CNN	T：1.38 K RH：14.14 %	T：1.46 K RH：14.95 %	T：1.53 K RH：16.32 %	135
ResNet	T：1.27 K RH：13.55 %	T：1.28 K RH：14.12 %	T：1.33 K RH：14.73 %	153
AR-CNN	T：1.16m/s RH：11.76m/s	T：1.21m/s RH：12.39m/s	**T：1.24 K** **RH：12.98 %**	181

Table. A1: Comparison of temperature and humidity profile inversion accuracy for different models, including root mean square error (RMSE) on training, validation, and test sets, and number of training epochs.

Author contributions. In this research, each author played a vital role. WH and ML designed the experiments, setting up the framework, procedures, and variables. ML carried out all experiments, handling equipment and data accurately. HS optimized the experimental code for better efficiency. RY verified the results through rigorous analysis. YY revised and edited the manuscript to meet the journal's high standards. Their combined efforts led to the success of this research.

Data availability. The ERA5 data are from the Copernicus Climate Change Service (C3S) Climate Data Store (CDS) (https://doi.org/10.24381/cds.bd0915c6, Hersbach et al., 2023). The training datasets used for this study can be requested from the authors.

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

Acknowledgments. This research was supported by many individuals. We gratefully acknowledge the faculty and senior researchers at the CMA Earth System Modeling and Prediction Centre, especially Bowen Cai for satellite data remapping and Shuting Zhou for RTTOV simulations. We also thank Professors Hao Li and Xiuyu Sun from the Shanghai Academy of AI for Science for guidance on algorithm optimization and model application, which advanced our work in intelligent science and

meteorological forecasting. Finally, we are grateful to all mentors who supported this research and its completion.

Financial support. This work was supported by the National Natural Science Foundation of China under Grants（U2442219 and 42205158)

Reference

Aires, F., Boucher, E., and Pellet, V.: Convolutional neural networks for satellite remote sensing at coarse resolution. Application for the SST retrieval using IASI, Remote Sens. Environ., 263, 112553, https://doi.org/10.1016/j.rse.2021.112553, 2021.

Blackwell, W. J.: A neural-network technique for the retrieval of atmospheric temperature and moisture profiles from high spectral resolution sounding data, IEEE Trans. Geosci. Remote Sens., 43, 2535–2546, https://doi.org/10.1109/tgrs.2005.855071, 2005.

Burns, B. A., Wu, X., and Diak, G. R.: Model-derived brightness temperature in AMSU moisture channels for various precipitation structures: Comparison of two radiative transfer formulations, in: 1995 International Geoscience and Remote Sensing Symposium, IGARSS '95: Quantitative Remote Sensing for Science and Applications, vol. 2, Firenze, Italy, 876–878, https://doi.org/10.1109/IGARSS.1995.521084, 1995.

Burns, B. A., Wu, X., and Diak, G. R.: Effects of precipitation and cloud ice on brightness temperatures in AMSU moisture channels, IEEE Trans. Geosci. Remote Sens., 35, 1429–1437, https://doi.org/10.1109/36.649797, 1997.

Cabrera-Mercader, C. R., and Staelin, D. H.: Passive microwave humidity profile retrievals using neural networks, in: Proceedings of IGARSS '94 – 1994 IEEE International Geoscience and Remote Sensing Symposium, vol. 4, Pasadena, CA, USA, 2057–2059, https://doi.org/10.1109/IGARSS.1994.399653, 1994.

Cadeddu, M. P., Turner, D. D., and Liljegren, J. C.: A Neural Network for Real-Time Retrievals of PWV and LWP From Arctic Millimeter-Wave Ground-Based Observations, IEEE Trans. Geosci. Remote Sens., 47, 1887–1900, https://doi.org/10.1109/tgrs.2009.2013205, 2009.

Cai, X., Bao, Y., Petropoulos, G. P., Lu, F., Lu, Q., Zhu, L., and Wu, Y.: Temperature and Humidity Profile Retrieval from FY4-GIIRS Hyperspectral Data Using Artificial Neural Networks, Remote Sens., 12, 1872, https://doi.org/10.3390/rs12111872, 2020.

Cao, X. F., Li, X. Y., Luo, Q., Liu, S. H., Li, P., and Liu, X.: Review of temperature profile inversion of satellite-borne infrared hyperspectral sensors, Natl. Remote Sens. Bull., 25, 577–598, https://doi.org/10.11834/jrs.20210009, 2021.

Chen, K., Cai, B., Han, W., and Suo, Z.: Matching of Observation Footprints in the FY-3G MWRI-RM Using BGI, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 17, 1–12, https://doi.org/10.1109/JSTARS.2024.3468437, 2024.

Chen, H., and Lin, L.: Numerical simulation of temperature profile retrievals from the brightness temperatures in 6 channels near 118.75 GHz, Chin. J. Atmos. Sci., 27, 894–900, https://doi.org/10.3878/j.issn.1006-9895.2003.05.10, 2003.

Du, J., Kimball, J. S., and Jones, L. A.: Satellite microwave retrieval of total precipitable water vapor and surface air temperature over land from AMSR2, IEEE Trans. Geosci. Remote Sens., 53, 2520–2531, https://doi.org/10.1109/TGRS.2014.2361344, 2015.

Duan, S.-B., Li, Z.-L., and Leng, P.: A framework for the retrieval of all-weather land surface temperature at a high spatial resolution from polar-orbiting thermal infrared and passive microwave data, Remote Sens. Environ., 195, 107–117, https://doi.org/10.1016/j.rse.2017.04.008, 2017.

Ebell, K., Orlandi, E., Hünerbein, A., Löhnert, U., and Crewell, S.: Combining ground-based with satellite-based measurements in the atmospheric state retrieval: Assessment of the information content, J. Geophys. Res.-Atmos., 118, 6940–6956, https://doi.org/10.1002/jgrd.50548, 2013.

Elfwing, S., Uchibe, E., and Doya, K.: Sigmoid-weighted linear units for neural network function approximation in reinforcement learning, Neural Netw., 108, 319–331, https://doi.org/10.1016/j.neunet.2017.12.012, 2018.

Gasiewski, A. J., and Staelin, D. H.: Statistical precipitation cell parameter estimation using passive 118-GHz O2 observations, J. Geophys. Res., 94, 18367–18378, https://doi.org/10.1029/JD094iD15p18367, 1989.

Gasiewski, A. J., and Johnson, J. T.: Statistical temperature profile retrievals in clear-air using passive 118-GHz O_2 observations, IEEE Trans. Geosci. Remote Sens., 31, 106–115, https://doi.org/10.1109/36.210450, 1993.

Gu, S., Zhang, P., Chen, L., Shang, J., Zhang, H., Lin, M., Zhu, A., Jia, S., Yin, H., Sun, F., Xu, H., Wang, H., Li, L., Wu, Q., Guo, Y., Dou, F., and Wu, S.: Overview and prospect of the detection capability of China's first precipitation measurement satellite FY-3G, Torrential Rain Disasters, 42, 489–498, https://doi.org/10.12406/byzh.2023-106, 2023.

He, K., Zhang, X., Ren, S., and Sun, J.: Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 770–778, https://doi.org/10.1109/CVPR.2016.90, 2016.

He, Q. R., Wang, Z. Z., and He, J. Y.: Retrieval of clear sky temperature and humidity profiles over land using measurements of FY-3C/MWHTS, J. Remote Sens., 21, 27–39, https://doi.org/10.11834/jrs.20176006, 2017.

Hersbach, H., Bell, B., Berrisford, P., et al.: ERA5 hourly data on pressure levels from 1940 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), 2023, https://doi.org/10.24381/cds.bd0915c6.

Ioffe, S. and Szegedy, C.: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, arXiv preprint, arXiv:1502.03167, 2015.

Ji, D., Shi, J., Xiong, C., Wang, T., and Zhang, Y.: A total precipitable water retrieval method over land using the combination of passive microwave and optical remote sensing, Remote Sens. Environ., 191, 313–327, https://doi.org/10.1016/j.rse.2017.01.028, 2017.

King, J. I.: The Radiative Heat Transfer of Planet Earth, 1958.

Liu, Y., Mao, J., Liu, J., and Li, F.: Study on BP neural network retrieval method of atmospheric profile using ground-based microwave radiometer, Plateau Meteorology, 29, 1514–1523, 2010.

Li, Z., and Han, W.: Impact of Implementing All-Sky Radiance Assimilation for FY-3E MWHS-2 in the CMA-GFS, Mon. Wea. Rev., https://doi.org/10.1175/MWR-D-24-0093.1, 2025 (in press).

Malmgren-Hansen, D., Laparra, V., Nielsen, A. A., and Camps-Valls, G.: Statistical retrieval of atmospheric profiles with deep convolutional neural networks, ISPRS J. Photogramm. Remote Sens., 158, 231–240, 2019, https://doi.org/10.1016/j.isprsjprs.2019.10.002.

Renju, R., Raju, C. S., Swathi, R., and Milan, V. G.: Retrieval of atmospheric temperature and humidity profiles over a tropical coastal station from ground-based Microwave Radiometer using deep learning technique, Journal of Atmospheric and Solar-Terrestrial Physics, 249, 106094, https://doi.org/10.1016/j.jastp.2023.106094, 2023.

Rosenkranz, P. W.: Retrieval of temperature and moisture profiles from AMSU-A and AMSU-B measurements, IEEE Trans. Geosci. Remote Sens., 39, 2429–2435, 2001, https://doi.org/10.1109/36.964979.

Sahoo, S., Bosch-Lluis, X., Reising, S. C., and Vivekanandan, J.: Radiometric Information Content for Water Vapor and Temperature Profiling in Clear Skies Between 10 and 200 GHz, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 8, 859–871, 2015, https://doi.org/10.1109/JSTARS.2014.2364394.

Smith, W. L.: Iterative solution of the radiative transfer equation for the temperature and absorbing gas profile of an atmosphere, Appl. Opt., 9, 1993–1999, 1970.

Smith, W. L. and Woolf, H. M.: The use of eigenvectors of statistical covariance matrices for interpreting satellite sounding radiometer observations, J. Atmos. Sci., 33, 1127–1140, 1976.

Solheim, F., Godwin, J. R., Westwater, E. R., Han, Y., Keihm, S. J., Marsh, K., and Ware, R.: Radiometric profiling of temperature, water vapor and cloud liquid water using various inversion methods, Radio Sci., 33, 393–404, 1998, https://doi.org/10.1029/97RS03656.

Stähli, O., Murk, A., Kämpfer, N., Mätzler, C., and Eriksson, P.: Microwave radiometer to retrieve temperature profiles from the surface to the stratopause, Atmos. Meas. Tech., 6, 2477–2494, https://doi.org/10.5194/amt-6-2477-2013, 2013.

Susskind, J., Barnet, C. D., and Blaisdell, J. M.: Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds, IEEE Trans. Geosci. Remote Sens., 41, 390–409, https://doi.org/10.1109/TGRS.2002.808236, 2003.

Tan, Q., Yao, Z., Zhao, Z., Han, Z., and Sun, X.: Performance analysis of atmospheric temperature and humidity profile retrieval using multiband microwave radiometer, Remote Sens. Technol. Appl., 30, 170–177, 2015.

Wang, D., Tong, L., Gong, X., Guan, X., Wang, P., and Gao, B.: Retrieval of atmospheric temperature profiles from hyperspectral microwave radiative data based on the neural network, in: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Brussels, Belgium, 7095–7098, https://doi.org/10.1109/IGARSS47720.2021.9554282, 2021.

Wang, H., Liu, D., Xia, Y., Xie, W., and Wang, Y.: Retrieval of atmospheric temperature profile from historical data and ground-based observations by using a machine learning algorithm, Remote Sens., 15, 2717, https://doi.org/10.3390/rs15112717, 2023.

Wang, X., Song, G., Yao, Z., and Li, W.: Study on the retrieval of atmospheric humidity profiles over the northwestern Pacific Ocean using AMSU data, J. Peking Univ. (Nat. Sci. Ed.), 46, 69–78, https://doi.org/10.13209/j.0479-8023.2010.011, 2010.

Wang, Y., Fu, Y., Liu, G., Liu, Q., and Sun, L.: A new water vapor algorithm for TRMM Microwave Imager (TMI) measurements based on a log linear relationship, J. Geophys. Res., 114, D21304, https://doi.org/10.1029/2008JD011057, 2009.

Weinman, J. A.: The effect of cirrus clouds on 118-GHz brightness temperatures, J. Geophys. Res., 93, 11059–11062, https://doi.org/10.1029/JD093iD09p11059, 1988.

Wu, X., Li, J., Zhang, W., and Wang, F.: Atmospheric profile retrieval with AIRS data and validation at the ARM CART site, Adv. Atmos. Sci., 22, 647–654, https://doi.org/10.1007/BF02918708, 2005.

Xia, X., Fu, D., Shao, W., Jiang, R., Wu, S., Zhang, P., et al.: Retrieving precipitable water vapor over land from satellite passive microwave radiometer measurements using automated machine learning, Geophys. Res. Lett., 50, e2023GL105197, https://doi.org/10.1029/2023GL105197, 2023.

Xiao, H., Han, W., Zhang, P., and Bai, Y.: Assimilation of data from the MWHS-II onboard the first early morning satellite FY-3E into the CMA global 4D-Var system, Meteorol. Appl., 30(3), e2133, https://doi.org/10.1002/met.2133, 2023.

Xu, X., Han, W., Gao, Z., Li, J., and Yin, R.: Retrieval of atmospheric temperature profiles from FY-4A/GIIRS hyperspectral data based on TPE-MLP: analysis of retrieval accuracy and influencing factors, Remote Sens., 16, 1976, https://doi.org/10.3390/rs16111976, 2024.

Yao, S. H., and Guan, L.: Retrieval of atmospheric temperature and humidity profiles using a machine learning algorithm based on satellite-borne infrared hyperspectral observations, Infrared Laser Eng., 51, 461–472, 2022.

Zhang, J., Liu, P., Zhang, F., and Song, Q.: CloudNet: ground-based cloud classification with deep convolutional neural network, Geophys. Res. Lett., 45, 8665–8672, https://doi.org/10.1029/2018GL077787, 2018.

Zhang, P., Gu, S., Shang, J., Zhang, H., Chen, L., Lin, M., Zhu, A., Jia, S., Wu, S., and Xian, D.: Overview of the mission of China's first precipitation satellite—FY-3G, Int. Space, 6, 17–21, 2023.

Zhu, X., Yee, J., and Talaat, E. R.: Diagnosis of dynamics and energy balance in the mesosphere and lower thermosphere, J. Atmos. Sci., 58, 2441–2454, https://doi.org/10.1175/1520-0469(2001)058<2441:DODAEB>2.0.CO;2, 2001.