A Calibration Algorithm for Satellite Temperature Profile Products Based on Variation and Artificial Neural Network

Runze Zhao , Xiangde Xu * , Tian Xian , Wenyue Cai , Shengjun Zhang , Zhiying Cai , Lin Chen

Posted Date: 14 May 2026

doi: 10.20944/preprints202605.0954.v1

Keywords: temperature profile; satellite; machine learning; variation; data assimilation

 doi:10.20944/preprints202605.0954.v1

Article

A Calibration Algorithm for Satellite Temperature Profile Products Based on Variation and Artificial Neural Network

Runze Zhao [1,2], Xiangde Xu [3,*], Tian Xian [4], Wenyue Cai [3], Shengjun Zhang [3], Zhiying Cai [5] and Lin Chen [1,2]

[1] Key Laboratory of Radiometric Calibration and Validation for Environmental Satellites, National Satellite Meteorological Center (National Center for Space Weather), China Meteorological Administration, Beijing, China

[2] Innovation Center for FengYun Meteorological Satellite, Beijing, 100081

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

[4] National Key Laboratory of Space Target Awareness, School of Space Information, Space Engineering University, Beijing 101416, China

[5] Ningbo Meteorological Observatory of Zhejiang Province, Ningbo 315000, China

* Correspondence: xuxd@cma.gov.cn

Abstract

Accurate information on atmospheric temperature profiles is crucial for improving numerical weather forecasting and short-term numerical weather prediction (NWP). However; the harsh environment of the Tibetan Plateau (TP) limits the availability of station observations; which fails to meet the high spatial resolution required for NWP. In this study; we present a method to calibrate temperature profiles obtained from the Vertical Atmosphere Sounding System (VASS) using data from the polar-orbiting satellite FY-3C. The aim is to provide high-resolution atmospheric structure for NWP in the TP. The temperature profile in VASS exhibits temporal and spatial heterogeneity due to the significant impact of clouds on the radiative transfer mode (RTM). To address this; we employ a combination of variation and artificial neural network (Var-ANN) methods to calibrate the satellite product and improve its compatibility with the model. To confirm the feasibility of our method; we compare the calibrated results with the observed data from 121 radiosonde soundings and 2400 meteorological stations in China; both of which represent conditions closest to the real atmospheric states. The calibrated temperature shows improvements over the original temperature; with a root mean square error; bias; and agreement with radiosonde soundings of 2.11; -0.72; and 0.998; respectively. We also select two classical cases involving the eastward movement of the plateau vortex (PV) and the formation of precipitation to verify the applicability of the calibration in NWP. The results demonstrate that the performance of NWP improves after assimilating the calibrated data; with the Var-ANN data assimilation scheme achieving the highest threat score of 66.9 and 66.7 for case 1 and case 2; respectively. These findings suggest that the Var-ANN method is suitable for calibrating satellite temperature profiles; and the calibrated data holds potential for precipitation forecasting. Furthermore; the novel method can also be applied in global temperature profile correction and satellite cross-calibration.

Keywords: temperature profile; satellite; machine learning; variation; data assimilation

1. Introduction

The accuracy of numerical weather prediction (NWP) highly relies on initial atmosphere state, therefore, accurate initial fields is of great importance for NWP. It is especially true in the Tibetan

 doi:10.20944/preprints202605.0954.v1

Plateau (TP), where the weather and climate exert significant impacts on downstream regions, to which they are highly sensitive (Chen, Xu et al. 2012, Dong, Xu et al. 2019, Zhao, Xu et al. 2019). Many studies have revealed this process, that the convective system triggered over the southeastern TP can propagate eastward and intensify, eventually leading to precipitation in downstream areas. (Zhao, Xu et al. 2016, Zhao, Xu et al. 2019, Yang, Zhang et al. 2020). In contrast to plain areas, observational data on the TP feature relatively sparse spatial resolution. Meanwhile, the complexity of terrains and underlying surfaces in the region further degrades the spatial representativeness of observations stations. While sparse observation sites are sufficient for investigating large-scale and long-term climatic characteristics over the TP, they cannot accurately capture small-scale convective systems. Since the dynamic and thermodynamic structure of the atmosphere over the TP differs significantly from that over plain regions, such small-scale convective process occurring over the TP cannot be accurately simulated. (Raju, Parekh et al. 2015, Xu, Liu et al. 2019, Lv, Xu et al. 2020). Therefore, to better reproduce this process, assimilating more accurate initial information for simulation over the TP has become a key focus and challenge in current TP modeling research.

Recognized as best capturing the actual atmospheric conditions, radiosonde observations were directly assimilated into NWP models in early studies, leading to evident improvements in prediction skills. (Zhang, Xu et al. 2014, Li, Hu et al. 2018, Li, Xu et al. 2018). However, with the continuous advancement of high-resolution NWP, the insufficient station density of sounding data over the TP poses challenges for its assimilation in fine-scale numerical simulations. Therefore, some studies have attempted to assimilate satellite observations, which have broader spatial coverage, to improve the accuracy of high-resolution precipitation simulations (Lu, Yang et al. 2011, Maggioni, Reichle et al. 2013, Raju, Parekh et al. 2015, Eyre, English et al. 2020). However, it was found that discrepancies still exist between satellite retrievals and sounding observations (Lian, Zeng et al. 2017, Esmaili, Smith et al. 2020, Filei, Andreev et al. 2021), since satellites do not directly measure atmospheric profiles but indirectly retrieve them through atmospheric radiative transfer models. To reduce inaccuracies in satellite observations, several studies directly assimilated brightness temperatures from atmospheric radiative transfer models (Orlandi, Fierli et al. 2010, Shi, Li et al. 2018). This scheme imposes indirect constraints on the atmospheric structure but lacks validation from surface observations. Hence it is highly dependent on the accuracy and stability of the observation instrument. Another approach involves pre-correcting satellite retrievals with radiosonde data prior to assimilation into NWP models. It not only ensures the consistency between satellite and sounding data but also maintains the high-resolution characteristics of satellite observations. Variation method is one of the representative methods, which can minimum the value of Euler equation in error field by constructing functional function. This method can greatly reduce data errors and has been widely adopted for satellite data correction and model assimilation studies (Wang, Dickinson et al. 2012, Hu, Deng et al. 2016, Zhang, Zhao et al. 2018). Traditional variational approaches have limited effectiveness in correcting discrete observation fields, since error propagation is confined to small regions under high-resolution grid settings. Local errors caused by nonlinear variables, such as satellite elevation angles in the retrieval process, are also hard to compensate for via variational methods (Herman, Brunke et al. 2010, He, Zhang et al. 2012, Guo, An et al. 2016, Chen, Xu et al. 2025). Thus, this study develops a hybrid Var-ANN method integrating variational schemes and artificial neural networks to obtain a more precise high-resolution dataset, which benefits the research of the dynamic and thermal structure over the TP. The accuracy of the dataset is also confirmed by improved simulations of downstream precipitation. As numerical models continue to advance toward finer scales, higher requirements are placed on observational data, making high-accuracy, high-resolution satellite data essential to addressing this issue.

 doi:10.20944/preprints202605.0954.v1

2. Data and Method

2.1. Study Area

The study area covers Chinese mainland (70°E–140°E, 10°N–60°N). The FY-3C satellite passes over this region at 0000 and 1200 UTC, consistent with the observation times of conventional meteorological stations. This region features complex weather, climate, and topographic conditions, which pose considerable challenges to atmospheric temperature profile retrieval. The selection of study period is determined by considerations of sample size and practical significance. The research period covers July–August 2019, since summer clouds are more frequent and impose greater interference on temperature profile retrieval. As radiative transfer models (RTM) have become increasingly capable with enhanced retrieval and correction algorithms, the most recent 2019 dataset is employed.

2.2. Data

This study uses VASS temperature profiles and brightness temperatures (FY-3C) as the main background data for neural network training, supplemented by cloud amount and spatial location information. VASS employs both microwave and infrared measurements to retrieve atmospheric temperature profiles. FY-3C is equipped with three temperature-retrieval-related sensors: the Microwave Humidity Sounder (MWHS), the Microwave Thermometer Sounder (MWTS) and the Infrared Atmosphere Sounder (IRAS). Given the short service life of MWTS (only 2013–2015), VASS products are mainly produced using MWHS and IRAS. MWHS includes fifteen channels: five at 183.31 GHz (water vapor absorption), eight at 118.75 GHz (oxygen absorption), and two window channels at 89 and 150 GHz, respectively. IRAS measures infrared radiative emission from the surface and atmosphere across 26 channels spanning 669–6098 cm^{-1}. Given the impact of clouds, cloud cover products by Visible and Infrared Radiometer (VIRR) onboard FY-3C are adopted. VIRR has seven visible/near-infrared and three infrared bands with a 1.1 km ground resolution, and the cloud products are collocated to the 17 km VASS pixels (http://satellite.nsmc.org.cn). Quality control of the input data is essential to ensure a reliable training process and valid results. To remove outliers and prevent overfitting, we use the middle 50% of the data as the input to the Var-ANN, discarding the upper and lower 25% of extreme values. The dataset is then divided into three parts: 80% for model training, 10% for testing, and 10% for validation.

Temperature profile data from 121 radiosonde stations and 2,400 meteorological stations in China are employed to assess the calibration results. These observations are taken at 0000 and 1200 UTC, matching the overpass time of FY-3C. The vertical temperature profiles are interpolated into the RTM pressure levels, covering1013.25 hPa down to 0.1 hPa across a total of 43 levels.

 doi:10.20944/preprints202605.0954.v1

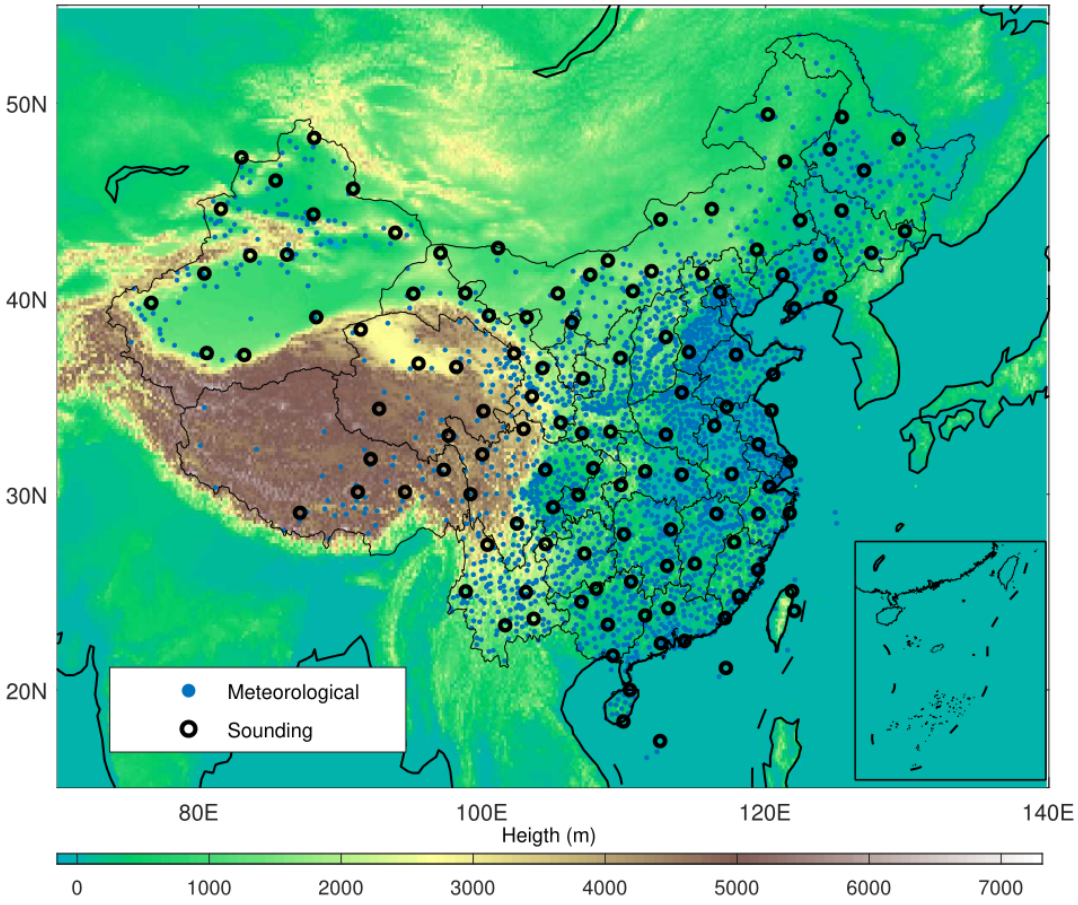


Figure 1. Distribution of sounding and meteorological stations. Black dots denote sounding stations, solid blue dots represent meteorological stations, and the colored background indicates geopotential height in meters.

2.3. Variation Method

The variation method is based on the principle that functional functions with multiple independent variables (1) must satisfy the Euler equation (2).

$$J[U(x,y)] = \iint_G F\,(x,y,U,\frac{\partial U}{\partial x},\frac{\partial U}{\partial y})\mathrm{d}x\mathrm{d}y \quad (1)$$

$$Fu - (\frac{\partial}{\partial x}Fu_x + \frac{\partial}{\partial y}Fu_y) = 0 \quad (2)$$

For this study, the satellite-derived temperature is set as $T_{sat}(x,y)$, the corresponding station observation field is $T_{obs}(I,J)$, and the difference between the two is the error field $\tilde{C}r(I,J)$.

$$\tilde{C}r(I,J) = T_{sat}(I,J) - T_{obs}(I,J) \quad (3)$$

Since station observations are discontinuous across the VASS product grid, the $\tilde{C}r(I,J)$ only covers grid points with available in-situ measurements. Thus, it is necessary to construct a more generalized error field function $Cr(x,y)$ and minimize this new function using the variational method.

$$J^* = \iint_D (Cr - \tilde{C}r)^2\,\mathrm{d}x\mathrm{d}y \to \min \quad (4)$$

For this variation problem, a functional function J^* can be constructed as:

$$J^* = \iint (Cr - \tilde{C}r)^2 + \lambda[(\frac{\partial Cr}{\partial x})^2 + (\frac{\partial Cr}{\partial y})^2]]\mathrm{d}x\mathrm{d}y \quad (5)$$

λ is the constraint coefficient, and the above formula can be rewritten into the following difference scheme：

$$\delta J = \sum\sum[(Cr - \tilde{C}r)^2 + \lambda[(\frac{\partial Cr}{\partial x})^2 + (\frac{\partial Cr}{\partial y})^2]] = 0 \quad (6)$$

The corresponding Euler equation is：

 doi:10.20944/preprints202605.0954.v1

$$(Cr - \tilde{Cr}) - \tilde{\lambda}(\frac{\partial^2 Cr}{\partial x^2} + \frac{\partial^2 Cr}{\partial y^2}) = 0 \quad (7)$$

The numerical solution of the above equation is derived through iterative expansion, yielding the variation-corrected precipitation field:

$$T(x, y) = T_{Sat} + Cr(x, y) \quad (8)$$

2.4. Back-Propagation Artificial Neural Network

Artificial neural network (ANN) is a nonlinear function to approach statistical relationship between input and output data. ANN typically consists of three components: neuronal nodes, an activation function, and the error back-propagation (BP) algorithm. The node values represent the neuron weights, which are continuously updated during training until the output is accurately fitted. The activation function maps neuronal inputs to outputs, enabling the Var-ANN to approximate nonlinear functions. As the core training component, the error back-propagation algorithm iteratively propagates output errors back to each layer's nodes until the error falls within a predefined threshold. In the Var-ANN model, the weight function adopts the dot-product weight function, the training function is the Levenberg-Marquardt algorithm, the transfer function is the hyperbolic tangent function, the learning function is the gradient descent algorithm, and the performance function is the mean square error. Details of this feedforward network configuration can be found in the MATLAB documentation.
(https://ww2.mathworks.cn/help/releases/R2019b/deeplearning/ref/feedforwardnet.html and https://ww2.mathworks.cn/help/releases/R2019b/deeplearning/ref/train.html).

2.5. Weather Research and Forecasting (WRF) Model

In this study, the WRF model (version 4.1.0) is employed to investigate the improvements of the calibrated dataset in simulating potential vorticity (PV) and predicting downstream precipitation. The WRF simulations are driven by the NCEP-GFS dataset. WRFDA-3DVAR is used to assimilate the calibrated data into the WRF initial fields. To ensure a controlled experimental setup, identical parameterization schemes are adopted across all comparative experiments, as detailed in Table 1.

Table 1. Parameterization scheme of Weather Research and Forecasting (WRF).

Physical process	**Parameterization scheme**	Reference
Microphysics	Purdue Lin	(Chen and Sun 2002)
Cumulus Parameterization	Kain-Fritsch	(Kain 2004)
Shortwave radiation	Dudhia	(Dudhia 1989)
Shortwave radiation	RRTM	(Mlawer, Taubman et al. 1997)

2.6. Evaluation Method

Three statistical indicators are employed to evaluation Var-ANN calibration results: the correlation coefficient (CC), root mean square error (RMSE), and bias. These indictors can be calculated as:

$$\mathrm{CC} = \frac{\sum_{i=1}^{n}(x-\bar{x})(y-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y-\bar{y})^2}} \quad (9)$$

$$\mathrm{RMSE} = \frac{\sum_{i=1}^{n}(x-\bar{y})^2}{n} \quad (10)$$

$$\mathrm{Bias} = \frac{\sum_{i=1}^{n}(x-\bar{y})}{n} \quad (11)$$

Where x denotes the satellite-derived temperature and calibrated temperature, and y represents the temperature observed by sounding stations.

T-score (TS), missing rate (MR), false alarm rate (FAR), and critical successful rate (CSI) are employed to further evaluate the WRF precipitation simulation. Listed in Table 2 are the calculation methods for these indicators, which are commonly applied in the simulation evaluation. n_{11} is called

 doi:10.20944/preprints202605.0954.v1

"Hit," indicating that precipitation is observed by the gauge and also predicted for the same grid points;

Table 2. Evaluation method for precipitation simulation.

Name	Short Name	Calculation	Ideal Value
Threat Score	TS	$\frac{n_{11}}{n_{11}+n_{10}+n_{01}+n_{00}}$	1
Missing Rate	MR	$\frac{n_{01}}{n_{01}+n_{00}}$	0
False Alarm Rate	FAR	$\frac{n_{10}}{n_{11}+n_{10}}$	0
Critical Success Index	CSI	$\frac{n_{11}}{n_{11}+n_{10}+n_{01}}$	1
Probability Of Detection	POD	$\frac{n_{11}}{n_{11}+n_{01}}$	1

n_{10} is called "False," denoting that precipitation is predicted but with no observed record; n_{01} is called "Miss," signifying that precipitation is observed by the gauge but not predicted; n_0 suggests that neither the gauge nor the prediction shows precipitation.

3. Results

3.1. Temperature Profile Correction Results

Many satellite products are retrieved using various artificial intelligence and deep learning methods. The core issue lies in how to construct the training dataset to capture the relationship between input and output data. The FY-3C VASS product combines microwave observations (MWTS and MWHS) and infrared observations (IRAS) to retrieve temperature profiles, with a spatial resolution of 17 km. Following the successive decommissioning of these sensors (MWTS in 2015 and IRAS in 2020), the VASS product ceased updating in 2020. There are two primary sources of error in the VASS product: observational limitations and retrieval algorithm uncertainties. For IRAS observations, infrared radiation cannot penetrate clouds; the measured radiation only originates from cloud tops and lacks full atmospheric information. Thus, IRAS observations are only valid under clear-sky conditions. To capture information within the cloud, window-channel microwave observations from MWHS and MWTS are incorporated into a multiple linear regression model. Constrained by the limitations of this model and the accuracy of RTM simulations under cloudy conditions, temperature profiles still cannot be retrieved accurately. As the first step of this study, the variational method is employed to construct a higher-quality training dataset. Widely applied in data assimilation and data merging, the variational method can adjust the satellite-derived error field toward the station-based observation field. In our study, radiosonde profiles are regarded as the most representative of real atmospheric conditions, and therefore the variational method is used to refine satellite-retrieved profiles toward these radiosonde observations.

As shown in Figure 2a, the bias at high temperatures (approximately 280–300 K) and low temperatures (around 190–230 K) is smaller than that at mid-level temperatures, indicating that temperature retrievals in the upper and lower atmosphere are more accurate than in the middle atmosphere (Figure 2 a). Scatter plots between satellite and radiosonde observations demonstrate that the calibrated dataset exhibits improved accuracy. Outliers deviating from the scatter distribution have been corrected, and the data points now cluster more closely around the linear relationship between satellite retrievals and station observations. Relative to radiosonde measurements, the

 doi:10.20944/preprints202605.0954.v1

calibrated satellite data yields an RMSE of 2.11, a bias of −0.72, and a correlation coefficient (CC) of 0.998, all of which represent significant improvements over the original satellite data (7.30, −4.77, and 0.984, respectively). The Var-ANN model is constructed to characterize the relationship between brightness temperature and atmospheric temperature. Vertical temperature retrieval relies on contributions from the vertical atmospheric structure, which exhibits stable climatic characteristics (Chen et al., 2020; Polyakov et al., 2018). The retrieval algorithm establishes regression relationships based on these stable features, which are encoded in the parameters of radiative transfer model equations. Accordingly, Var-ANN calibration not only improves low-quality temperature retrievals but also preserves the accuracy of high-quality data.

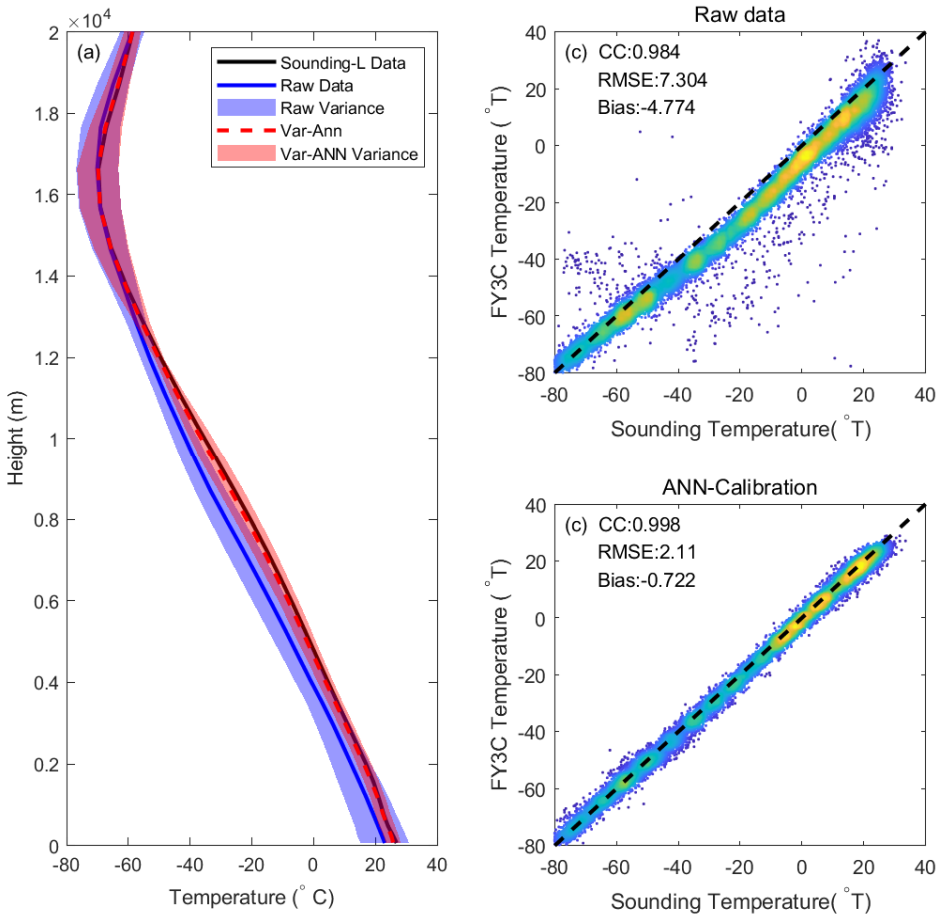


Figure 2. Comparison between satellite-derived temperature and radiosonde temperature. (a) Temperature profiles from radiosonde observations (black solid line), raw satellite retrievals, and Var-ANN calibrated retrievals. The blue area denotes the variance of the raw data and Var-ANN calibrated data; (b)–(c) Comparisons of raw and calibrated temperature retrievals against radiosonde observations.

The results elucidate that the VASS temperature profile is colder than station observations throughout the entire atmosphere, particularly in the middle and lower troposphere (0–12 km), a region closely linked to atmospheric dynamic and thermodynamic structures. Negative biases arise from multiple factors, including observational calibration uncertainties and the RTM performance (Jiang and Li 2009; Zhang et al. 2021), as well as aerosol effects (Feng and Wang 2021; Norris and Wild 2009). Although the bias is significantly reduced following Var-ANN calibration (from −4.77 in the raw data to −0.72), a residual bias remains unavoidable when training the network solely with VASS data. Notably, data quality at 0000 UTC (8 a.m. BJT) exhibits marked improvement compared

 doi:10.20944/preprints202605.0954.v1

to 1200 UTC (8 p.m. BJT). The increased fluctuations at 1200 UTC are likely attributed to orbital biases, a well-documented characteristic of the FY-3C Microwave Radiation Imager (Xie et al., 2019).

Besides the overall accuracy of temperature data, discontinuous regions appear at the edge of the satellite swath, which are clearly visible in the spatial distribution (Figure 3). Owing to the coarser resolution at the swath edge, we resample the data to 50 km resolution to highlight these discontinuities. The IRAS instrument records 56 pixels per scan line, and anomalous data typically occur periodically at approximately the 10th to 15th pixel (from left to right), persisting across about 10 consecutive scan lines. Such highly regular scan-line anomalies clearly do not represent real atmospheric signals and thus constitute a major source of error. Traditional methods for addressing this issue apply spatial smoothing over the entire domain (commonly 5-point or 9-point smoothing), at the cost of degrading high-resolution features. In contrast, after Var-ANN correction, the swath edges become significantly smoother while preserving most of the high-resolution information.

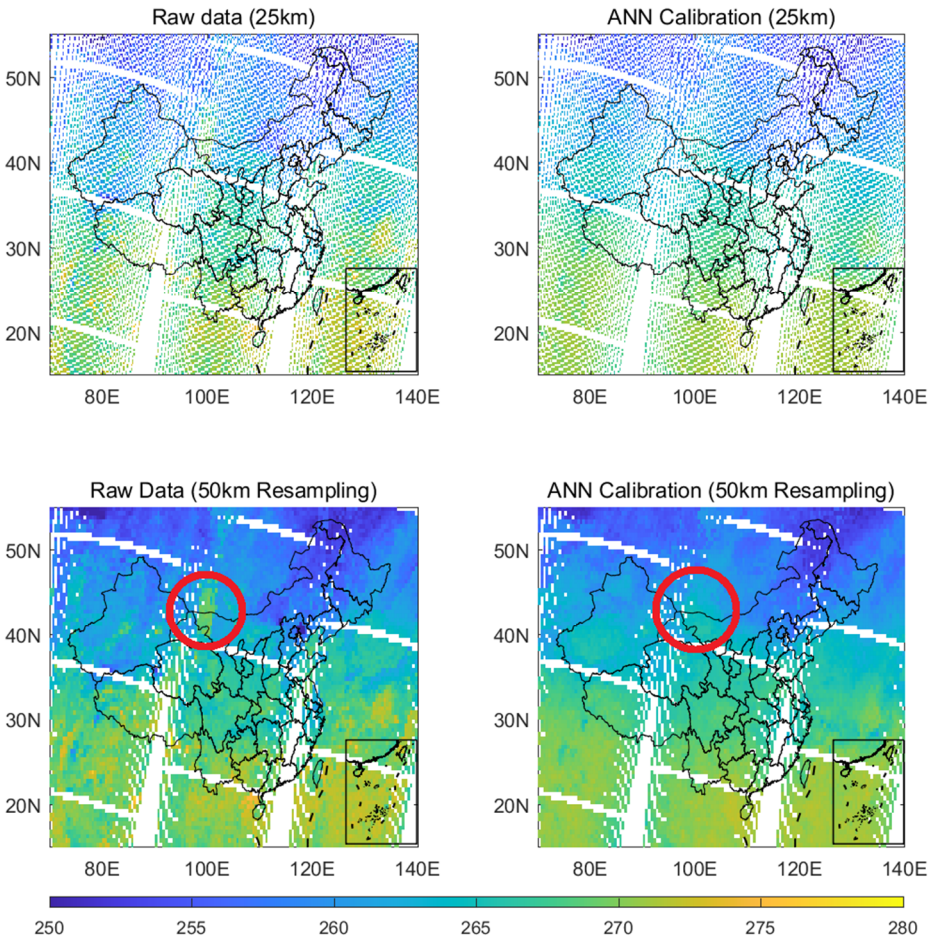


Figure 3. Spatial distribution of raw and calibrated temperature at 520 hPa. (a) and (b) show the original 25 km resolution, while (c) and (d) display the resampled products used to fill pixel gaps. The colorbar represents atmospheric temperature in Kelvin (K).

Temperatures at various isobaric levels are analyzed over the study period (Table 3). It is observed that upper-atmosphere temperatures exhibit a branch that is higher than radiosonde observations (Figure 2c), a feature that becomes more pronounced in the calibrated data. The results indicate that the Var-ANN model captures two types of relationships: one is strongly positively correlated with station temperatures (the diagonal dashed line in Figure 2), consistent with real atmospheric conditions; the other regression behavior is irregular: temperatures are underestimated in the middle and lower atmosphere but overestimated in the upper atmosphere. This may be

 doi:10.20944/preprints202605.0954.v1

attributed to MWHS channels located in water vapor absorption bands. Under cloudy conditions, higher water vapor content can cause the Var-ANN model to misjudge the vertical level of temperature signals, leading to underestimated retrievals in the middle and lower troposphere and overestimated values within the temperature inversion layer (220–240 K). Owing to the greater vertical variability in the temperature inversion layer, the upper-atmosphere branch is more distinct, as clearly reflected in the vertical temperature profiles (Figure 2). In addition to the regression branching, the negative bias is substantially reduced but still remains, with a value of −0.72 K. This result is reasonable, given that the calibration method relies solely on FY-3C satellite observations and is independent of radiosonde data. To further eliminate this bias, additional temperature profile data are required, which can be obtained through inter-calibration with other satellites.

Table 3. Standard deviation (Std.), root mean square error (RMSE) and bias relative to sounding observations at each atmospheric level.

Isobaric level		Raw data			Var-ANN		
No.	**Height(m)**	**Std.**	**Bias.**	RMSE	Std.	Bias.	RMSE
1	0	8	-4.5	11.22	2.3	-0.49	1.63
2	65.25	7.81	-3.9	9.92	2.49	-0.45	1.79
3	230.2	7.64	-3.66	9.23	2.5	-0.41	1.97
4	474.95	7.51	-4.02	8.9	2.51	-0.5	2.2
5	784.19	7.31	-4.66	8.91	2.48	-0.65	2.35
6	1146.51	6.94	-4.98	8.71	2.47	-0.73	2.38
7	1553.09	6.39	-5.16	8.24	2.4	-0.75	2.34
8	1997.17	5.87	-5.03	7.63	2.33	-0.75	2.14
9	2473.8	5.52	-4.95	7.2	2.34	-0.76	2.03
10	2979.43	5.33	-5.08	7.03	2.43	-0.78	1.96
11	3511.79	5.28	-5.5	7.21	2.61	-0.84	1.91
12	4069.59	5.41	-5.9	7.48	2.77	-0.91	1.92
13	4652.11	5.65	-6.25	7.79	2.94	-0.94	1.9
14	5259.54	5.75	-6.81	8.23	3.13	-1.02	1.9
15	5892.34	5.8	-7.27	8.54	3.41	-1.07	1.95
16	6551.03	5.68	-7.61	8.75	3.69	-1.14	2.05
17	7236.79	5.4	-7.89	8.89	3.91	-1.22	2.16
18	7950.11	5.07	-7.81	8.78	4.07	-1.25	2.25
19	8691.55	4.63	-7.06	8.12	4.02	-1.17	2.26
20	9461.6	4.01	-5.86	7.17	3.49	-1.04	2.31
21	10260.27	3.41	-4.77	6.22	2.45	-0.87	2.36
22	11087.37	3.12	-3.67	5.21	1.5	-0.68	2.36
23	11943.06	3.23	-2.35	4.04	2.45	-0.45	2.18
24	12829.77	4.35	-0.98	3.59	4.12	-0.19	1.97
25	13744.08	5.73	-0.33	3.99	5.62	-0.08	1.91
26	14684.76	6.77	-0.03	4.5	6.67	-0.06	1.9
27	15675.75	6.76	-0.36	4.37	6.62	-0.09	1.94
28	16643.1	6.31	-1.78	4.4	5.29	-0.29	1.79
29	17659	5.16	-1.95	3.94	3.82	-0.3	1.66
30	18701.29	4.04	-0.47	2.68	2.85	-0.06	1.52
31	19775.01	3.51	1.4	2.79	2.39	0.25	1.59
32	20884.27	3.66	2.62	3.91	2.02	0.41	1.73
33	22033.02	4.84	3.42	5.61	1.71	0.54	1.99
34	23227.87	5.42	4.05	6.5	1.91	0.59	2.35
35	24473.4	6	4.02	7.27	2.39	0.55	2.72
36	25773.19	5.94	0.93	8.58	1.93	0.2	3.16
37	27131.4	6.57	4.67	11.62	3.23	0.6	4.71

 doi:10.20944/preprints202605.0954.v1

38	28556.77	7.25	10.71	16.55	4.62	1.62	5.86
39	30049.14	7.77	17.16	18.61	5.29	4.36	6.27
40	31608.46	5.63	24.48	24.48	6.6	4.23	4.23
41	33236.83	2.67	13.07	13.07	7.05	1.9	1.9
42	34918.84	3.27	-6.11	6.1	9.95	-1.43	1.42
43	36639.07	8	-4.5	11.22	2.3	-0.49	1.63

Unit: Meter for Height, °C for Std., Bias and RMSE

3.2. Design and Description of Var-ANN

The structure of the Var-ANN model is often determined subjectively, lacking sufficient theoretical and physical support. In this study, we train and test several Var-ANN architectures with different configurations, aiming to achieve a more objective network design. We adopt a Var-ANN with two hidden layers, where the number of nodes in each layer is set up to 20. Here only several representative results are presented in Figure 5. The results indicate that a simple network structure cannot adequately characterize the relationship between brightness temperature and atmospheric temperature. As the network architecture becomes more complex, the ability of Var-ANN to represent this relationship improves, although at the cost of increased computational cost for training. Once the model capacity reaches its upper limit, further increasing the number of nodes no longer improves the model performance. The redundant nodes remain involved in training but contribute little to the final output. The number of training samples also requires careful tuning. Insufficient training samples prevent the model from adequately capturing the underlying relationship, while an excessively large sample size leads to overfitting of the temperature profiles. An overfitted Var-ANN tends to store individual sample details within each node, resulting in ineffective correction of the output temperature profiles. After considering the above factors, we finalize the Var-ANN structure (six nodes in the first layer and five nodes in the second layer) and select the training period (summer 2019).

Training data play an essential role, as they endow the Var-ANN with physical meaning and link it to atmospheric mechanisms. The output dataset consists of the 43-level temperature profiles retrieved by VASS, which represent the results of the RTM employed in the VASS system. The input data are composed of three parts, with the primary part being the brightness temperatures observed by fifteen MWHS channels. Although these channels are designed for water vapor sounding, their radiative frequency information can still be utilized for temperature calibration. Cloud cover retrieved by VIRR is also included as input to mitigate inaccuracies in the RTM under cloudy conditions. In addition to satellite radiance measurements, the longitude and latitude of each observed pixel are also incorporated into the input dataset as geographic information. Each input component carries distinct physical significance related to atmospheric temperature and RTM. Specifically, MWHS data contain integrated atmospheric information, which conventional RTMs rely on to retrieve temperature profiles. The inclusion of cloud cover enables the Var-ANN to characterize the radiation–temperature relationship under cloudy conditions. This relationship remains poorly constrained and cannot be accurately represented by RTMs (Mason 2002; Takeishi and Wang 2022; Yeo et al. 2022), which constitutes a major source of error in remote sensing observations. Many local factors also exert considerable influence on temperature profile retrieval, such as aerosols, atmospheric circulation, and land cover (Cheng et al. 2017; Zhao et al. 2021). These factors involve complex interaction mechanisms and are highly location-dependent. Therefore, longitude and latitude are also used as input data to approximately represent the influences of these local variables. Such auxiliary pixel information enables the Var-ANN model to consider the effects of local factors. S

The training process is illustrated in Figure 4, which summarizes the overall fitting procedure and performance. The error gradient decreases steadily throughout training, indicating that the Var-ANN solution is converging. The performance is evaluated using the mean squared error (MSE), which measures the discrepancy between the network-output temperatures and the original VASS

 doi:10.20944/preprints202605.0954.v1

temperatures. The MSE drops rapidly in the early stage of training. As the Var-ANN gradually approaches the temperature relationship, the rate of decrease slows and eventually stabilizes at epoch 29, with an optimal MSE of 5.45. The MU parameter is related to machine precision and iterative error. It increases whenever an iteration leads to a higher error. Validation testing uses 10% of the input data to verify the model output, with the MU increasing continuously each time the validation error rises. Training terminates if the MU exceeds 10^5 or the number of validation failures exceeds 6. These two parameters constrain the training process within reasonable boundaries, ensuring that the Var-ANN properly captures the underlying features.

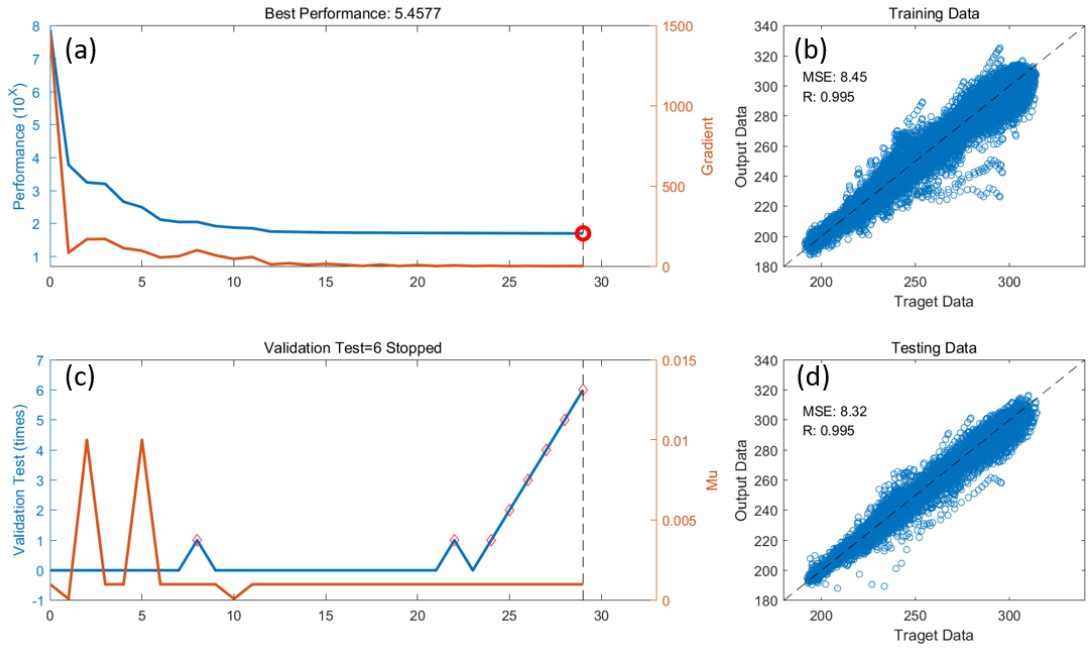


Figure 4. Training process of the Var-ANN. (a) The performance and gradient of the machine learning model. The blue axis and line represent the performance, defined as the mean squared error, while the orange axis and line denote the gradient; (c) the constraint boundaries during training. The blue axis and line correspond to the number of validation failures, and the orange axis and line represent the iterative parameter MU; (b) and (d) show comparisons for the training and testing datasets, respectively. The x-axis represents the radiosonde profiles used for training and testing, and the y-axis represents the corresponding model outputs.

Var-ANN encodes its input-to-output processing procedure through node values, which poses challenges for interpretation. However, the contribution of input to the output can be captured through correlation analysis, with the results presented in Figure 5. It is noteworthy that latitude is highly correlated with the temperature profile and exerts different effects on the stratosphere and troposphere. Specifically, the tropospheric temperature is heated by surface radiation, while the surface itself is heated by solar radiation. As higher latitudes receive less solar radiation, a negative correlation exists between latitude and tropospheric temperature. With the increase of altitude, the temperature is less affected by the surface and more influenced by temperature advection. The study area of this article is mainly located in the middle latitude region, which is controlled by the Ferrel Circulation. This circulation pattern transfers heat from high latitude to low latitude, resulting in a positive correlation in the stratosphere. The correlation with longitude is similar to that with latitude but not statistically significant. The negative correlation in the lower atmosphere is the result of the synergy between the heating effect of the TP and the westerlies: the TP heats the temperature of the lower atmosphere, while the westerlies transport this heat eastward. In contrast, the positive correlation in the stratosphere is caused by the westward component of the Ferrel Circulation. The research results demonstrate the necessity of additional geographical information, as it enables Var-

 doi:10.20944/preprints202605.0954.v1

ANN to extract surface-atmosphere coupling features and circulation characteristics, thereby further correcting the temperature profile in line with local conditions.

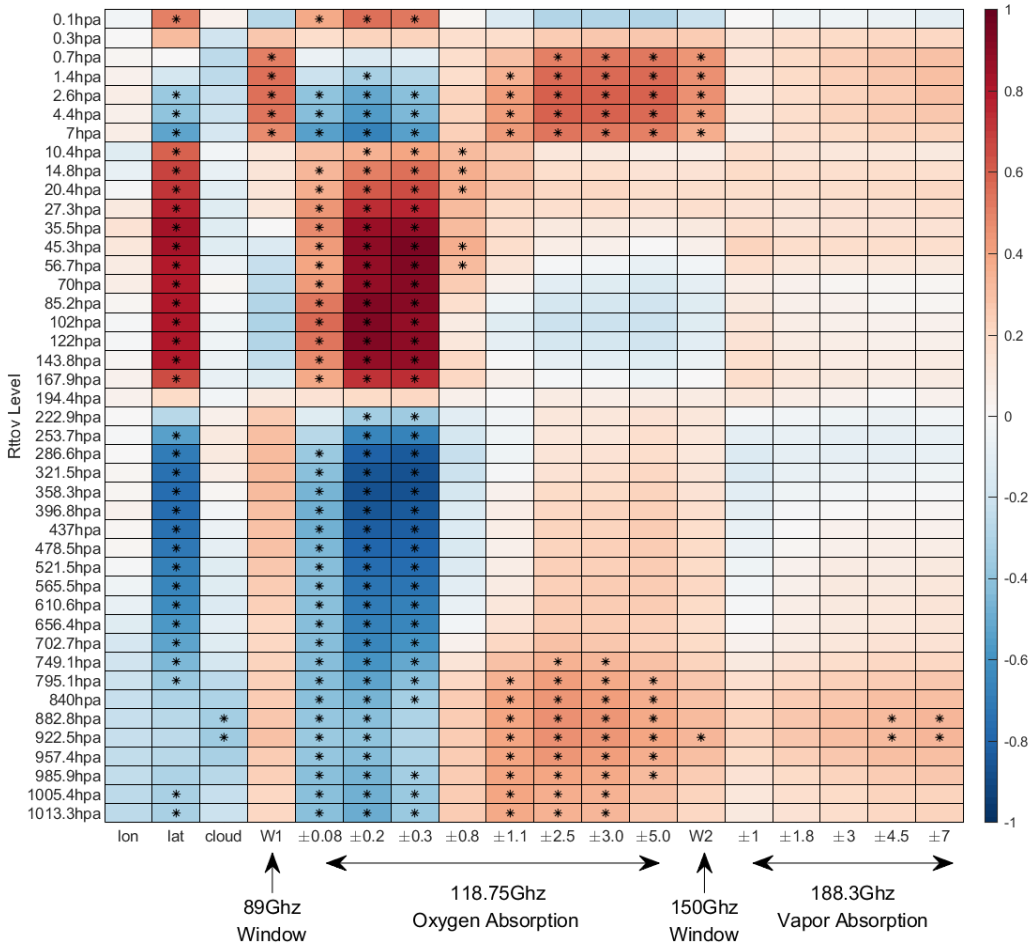


Figure 5. Correlation between input and output data in the machine learning model. The x-axis represents the input data, which includes geographical information, cloud parameters, and brightness temperature channels. The y-axis denotes the pressure levels of the raw data, which are determined by the radiative transfer model. The color gradient indicates the correlation coefficient between the input and output variables, while the markers signify that the corresponding correlation coefficients are statistically significant at the 95% confidence level.

Clouds show a negative correlation with the output across most levels, though this relationship is weak and non-significant. This indicates that Var-ANN fails to fully capture cloud-related features in the RTM. Although clouds do not absorb microwave radiation, refraction and scattering still occur, lowering observed brightness temperature and, in turn, retrieved temperature. Despite this weak linkage, the model still partially corrects cloud-induced error, especially in the lower atmosphere (882.8 hPa and 922.5 hPa). The unique predominance of negative cloud correlations across most levels suggests that cloud parameters could be an effective means of mitigating negative biases in station observations (Figure 2). To assist RTM, two window channels operating at 89 GHz and 150 GHz are designed to capture surface radiation parameters. The oxygen absorption channel (118.75 GHz) serves as the primary component for retrieving temperature profiles, with its narrow pass-bands (width ≤ 0.3 GHz) exhibiting a high correlation with almost all pressure levels. In contrast, the wide pass-bands (width > 0.3 GHz) are highly correlated with the temperatures of the upper and lower atmospheres, which supplements the levels where the narrow pass-bands show non-significant correlations. The water vapor absorption channel (183.75 GHz) is designed to observe humidity profiles. The positive features extracted by Var-ANN can be explained by the Clausius-Clapeyron (C-

 doi:10.20944/preprints202605.0954.v1

C) equation, which states that an increase in temperature leads to a corresponding increase in specific humidity. This finding demonstrates the functional efficacy of water vapor channels in correcting temperature profiles. The cloud and water vapor channels exert analogous roles within the Var-ANN framework, both displaying comparable correlation patterns with temperatures across nearly all pressure levels. This seemingly contradictory situation affords the Var-ANN two complementary perspectives for characterizing the influence of water on temperature—one rooted in radiative transfer processes and the other in atmospheric thermodynamics. Collectively, these findings demonstrate that the Var-ANN is capable of capturing the fundamental relational characteristics between atmospheric profiles and brightness temperature, mirroring the operational logic of the RTM. Thus, the trained model accurately retrieves temperature profiles, providing more reliable data input for follow-up research.

3.3. Calibrated Data Validation in Data Assimilation

To further validate the accuracy of the calibrated data and its applicability in NWP data assimilation, the Weather Research & Forecast (WRF) Model is employed to simulate precipitation downstream of the TP. Two cases, specifically 12 July 2019 and 14 June 2020, are selected for WRF simulations. In these instances, convective cells that formed over the southeastern TP propagated eastward and generated heavy rainfall over the downstream region to the east of the TP. The selection criteria focus on their typical development process: TP-based convective initiation, westerly-affected eastward movement, upscale growth into organized convective systems, and resultant precipitation. Large amounts of previous studies have investigated this process, and their results indicate that the atmospheric conditions in the southeastern TP are highly correlated with convection and precipitation in the downstream regions (Wang et al., 2011; Xu et al., 2012; Zhang et al., 2014). If the calibrated data can effectively simulate this process, it will not only verify the accuracy of the data but also facilitate a deeper understanding of this process. To confirm the reliability of our dataset and reveal this mechanism, three distinct simulation schemes are designed. Two typical summer cases are selected for WRF simulation, with the results presented in Figures 6 and 7. The “GFS” scheme uses the Global Forecast System (GFS) of the NCEP as the initial field, which is commonly employed in precipitation forecasting. The “FY3C” and “Sounding” schemes assimilate the original satellite and station data as the initial field, respectively, while the “Var-ANN” scheme assimilates the calibrated dataset developed in this study. The spatial distribution of gauge and simulated precipitation on June 10, 2019 is presented in Figure 6. It is found that the Sichuan Basin exhibits a pronounced precipitation center resulting from the aforementioned mechanism. Among the designed schemes, Var-ANN calibration alone accurately reproduces this case, whereas both the FY3C and Sounding schemes yield inaccurate results. This inaccuracy is attributed to the insufficient accuracy of the FY3C atmospheric profiles and the sparse distribution of sounding stations.

 doi:10.20944/preprints202605.0954.v1

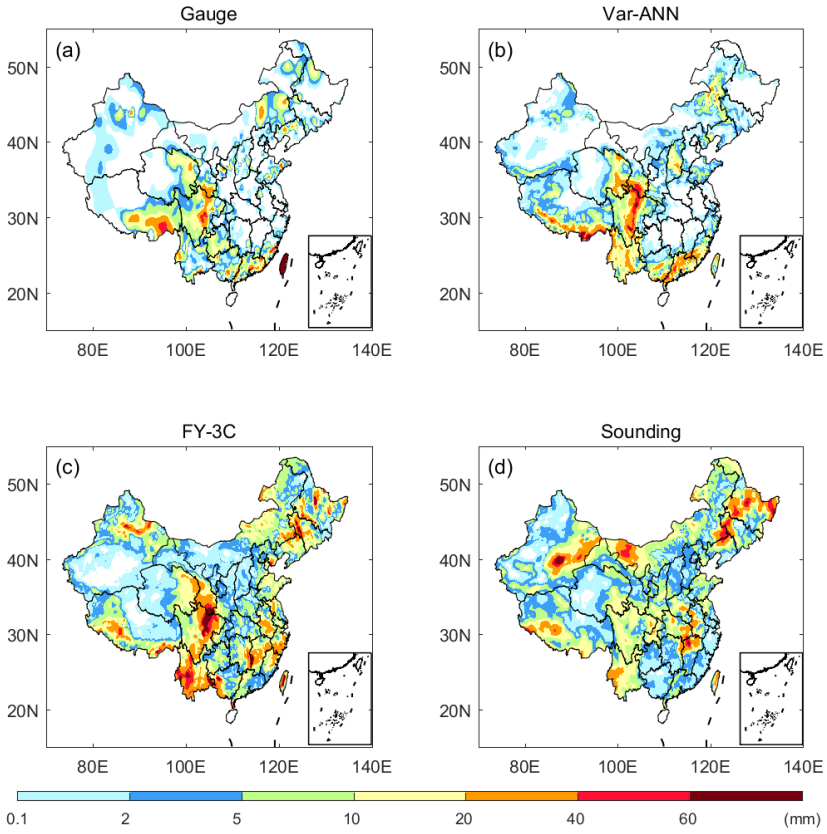
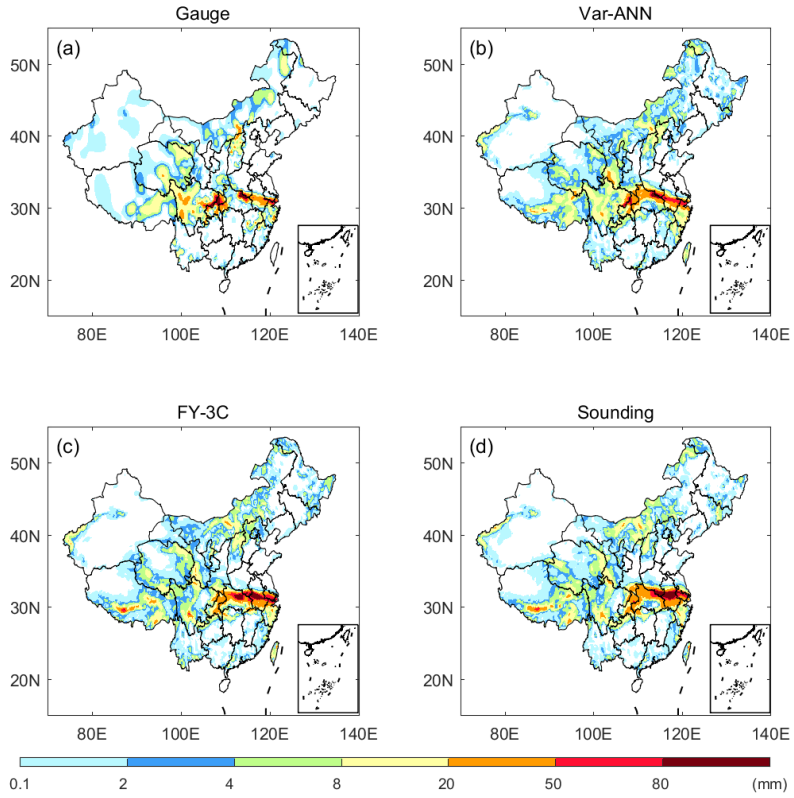


Figure 6. Daily precipitation on 10/June/2019. (a) gauge precipitation, (b) calibrated data simulation, (c) FY3D-VASS original data simulation, and (d) sounding profile simulation.



 doi:10.20944/preprints202605.0954.v1

Figure 7. Daily precipitation on 14/July/2020. (a) gauge precipitation, (b) calibrated data simulation, (c) FY3D-VASS original data simulation, and (d) sounding profile simulation.

To quantitatively analyze the errors associated with different data assimilation methods, several evaluation metrics for precipitation simulation results are presented in Table 4. Compared with the GFS scheme, the simulation accuracies of the other three schemes are improved, illustrating the role of data assimilation in simulation optimization. RMSE and bias reflect the overall errors across all samples; the FY-CLB scheme yields the lowest RMSE and bias, indicating its superior accuracy. TS, FAR, and CSI are commonly used for spatial verification of precipitation: FAR denotes the false alarm rate of precipitation, CSI represents the critical success index, and TS refers to the threat score, which combines both false alarms and successful detections. These metrics disregard precipitation intensity and instead focus on the spatial performance of simulated precipitation. For the FY-CLB scheme, the CSI value indicates that 60.76% of precipitation grids are correctly predicted, which is the highest among all schemes. The FAR indicates that 35.82% of no-raining grids are forecasted accurately, which outperforms the GFS scheme; however, the Sounding scheme achieves the best performance at 30.8%. By comprehensively considering both hit and false alarm events, the TS score further confirms that the FY-CLB scheme yields the best performance in simulating downstream precipitation. The research results demonstrate that the initial field incorporating both sounding and satellite data assimilation yields optimal performance in precipitation simulation.

Table 4. Two cases selected for daily evaluation: 12 July 2019 (Case 1) and 14 June 2020 (Case 2).

		GFS	FY3C	Sounding	FY-CLB
Case 1	RMSE	11.65	9.61	5.76	**5.6**
	Bias	7.94	3.33	-1.49	**1.21**
	TS	55.88	53.62	65.09	**66.93**
	FAR	44.16	44.19	**30.8**	35.82
	CSI	55.67	55.77	51.76	**60.76**
Case 2	RMSE	9.18	14.08	6.59	**6.46**
	Bias	5.07	2.37	1.62	**1.49**
	TS	57.7	55.75	63.31	**66.7**
	FAR	48.09	49.64	41.57	**38.79**
	CSI	50.25	48.31	55.44	**57.38**

Unit: mm for RMSE and Bias, % for TS, UR, FAR and CSI.

4. Conclusion and Discussion

To provide accurate, high-resolution initial atmospheric field for NWP simulation over the TP, this study integrates the variation method and artificial neural network (Var-ANN) to calibrate satellite-derived temperature profiles, and subsequently assimilates the calibrated data into the NWP model to simulate downstream precipitation. The background temperature field is derived from the Vertical Atmospheric Sounding System (VASS) product of the FY-3C satellite. As detailed in He et al. (2016), this product is retrieved from MWHS, MWTS, and IRAS measurements using a multiple linear regression algorithm, which derives regression coefficients from prior atmospheric samples and converts brightness temperature into atmospheric temperature profiles. While this method enables rapid temperature profile retrieval, it cannot adequately account for nonlinear factors such as satellite zenith angle. To address this issue, the FY-3C retrieval algorithm partitions atmospheric samples into multiple subsets. Although these subsets can approximate nonlinear variations, they often struggle to preserve spatial homogeneity. Figure 3 reveals a pronounced high-value region, typically occurring at large satellite zenith angles, which is attributed to the retrieval scheme. In addition, uncertainties of cloud radiation simulation within RTM remain a critical issue, which further degrades the quality of the retrieved atmospheric profiles. To cope with these issues in this study, we additionally incorporate satellite zenith angle, geographic information, and cloud

 doi:10.20944/preprints202605.0954.v1

parameters as input features in the Var-ANN framework to better capture the radiative signals associated with atmospheric temperature. The ANN method has been widely applied in atmospheric data calibration; most studies use satellite-observed brightness temperatures as input training data and radiosonde observations as the target output (Hu et al., 2023; Huang et al., 2021). A key advantage of the proposed Var-ANN configuration is that it treats radiosonde measurements as the most representative of real atmospheric conditions. Using variational analysis data as the training output not only significantly expands the training sample size but also enables the Var-ANN model to extract more accurate details of the atmospheric vertical structure.

Training results from the Var-ANN model demonstrate that it can effectively capture the dominant relationships between brightness temperature and atmospheric temperature profiles. As a result, the accuracy and stability of the corrected temperature profiles are notably improved. Correlation analysis between model inputs and outputs reveals that the Var-ANN framework successfully extracts key features associated with oxygen absorption and temperature profile structure. In addition, the supplementary input variables also show strong correlations with the output temperature field, particularly latitude. Latitude exhibits a negative correlation in the troposphere and a positive correlation in the stratosphere, which is consistent with large-scale circulation patterns. These findings confirm that the inclusion of supplementary input parameters is both necessary and beneficial for improving temperature profile correction. The effects of cloud amount and water vapor on the model output are opposite: cloud amount shows a negative correlation, while water vapor exhibits a positive correlation. Although most of these correlations are not statistically significant, they still suggest that the Var-ANN can partially capture signals related to latent heat absorption and release. In summary, the Var-ANN model is capable of extracting meaningful relationships between input variables and atmospheric temperature profiles, and the Var-ANN-based correction effectively improves the quality of the temperature profile product from VASS.

High-resolution satellite data can provide NWP models with more detailed information regarding the thermal structure of the atmosphere over the TP (Shi et al., 2008; Zhao et al., 2018). However, the accuracy of satellite retrievals relies heavily on both the precision of the retrieval algorithm and the accurate simulation of brightness temperature by RTM. To mitigate uncertainties introduced during the retrieval process, this study proposes a Var-ANN correction algorithm. This approach links discrete station observations with continuous satellite retrievals via variational processing, allowing the Var-ANN model to characterize the relationship between brightness temperature and atmospheric temperature. By integrating the complementary strengths of radiosonde and satellite observations, this method provides a more reliable initial condition for the numerical model. This improved initialization enables the model to better reproduce convective processes over the TP and its impact on downstream precipitation.

The dynamic and thermal effects of the TP exert a substantial influence on the development of small-scale convection and the occurrence of precipitation over the region. During summer, the TP is characterized by abundant water vapor, which readily forms precipitation under the influence of vortices and shear lines over the TP (Li et al., 2019; Shi et al., 2008). Local intense convection generated in this process propagates eastward under the steering of westerly winds, continuously incorporating additional water vapor and intensifying, ultimately leading to heavy precipitation events over the downstream middle and lower reaches (Peng et al., 2009; Xu et al., 2012; Zhao et al., 2019b). However, owing to the complex terrain of the TP, the spatial resolution of in-situ observations remains relatively sparse in comparison with the vast size of the region. Meanwhile, the complex terrain and underlying surface conditions are responsible for the poor spatial representativeness of the observation stations. While such sparse station coverage is sufficient for investigating large-scale and long-term climatic features of the TP, it fails to adequately resolve small-scale convective systems. Consequently, accurate short-term precipitation forecasting continues to pose a major challenge for NWP. To verify the improvements brought by this method to NWP simulations, this study selects two typical cases in which convective systems initiated over the TP and then moved eastward,

 doi:10.20944/preprints202605.0954.v1

leading to downstream precipitation. Relative to the control experiments, the Var-ANN data assimilation scheme yields noticeably improved precipitation forecasts. Root-mean-square error, bias, and correlation coefficient analyses all confirm that the calibrated temperature profiles agree better with radiosonde observations. Furthermore, the threat score, critical success index, and false alarm rate demonstrate that assimilating improved high-resolution atmospheric profiles over the TP significantly enhances WRF's quantitative precipitation forecast over downstream regions. These findings carry important implications for the development of TP-based reanalysis datasets and for advancing our understanding of atmospheric physical processes and systematic climate change over the TP.

Acknowledgments: This work was supported by the National Natural Science Foundation of China (Grant No. U2542206), National Key R&D Program of China (Grant No. 2025YFE0108100), National Natural Science Foundation of China (Grant No. 42505137) and National Natural Science Foundation of China (Grant No. 42475074).

Reference

Chen: B., X.-D. Xu, S. Yang and J.-C. Bian (2012). "On the characteristics of water vapor transport from atmosphere boundary layer to stratosphere over Tibetan Plateau regions in summer." Chinese Journal of Geophysics-Chinese Edition **55**(2): 406-414.

Chen, L., N. Xu, J. S. Wang, J. Shang, Y. X. Shou, B. Li, R. H. Xu, S. L. Wu, X. Wang, W. Zheng and S. Z. Jia (2025). "FengYun satellites： From observations to quantitative applications." National Remote Sensing Bulletin **29**(6): 1462-1479.

Chen, S. H. and W. Y. Sun (2002). "A one-dimensional time dependent cloud model." Journal of the Meteorological Society of Japan **80**(1): 99-118.

Dong, L., X. Xu, T. Zhao and H. Ren (2019). "Linkage between moisture transport over the Yangtze River Basin and a critical area of the Tibetan Plateau during the Meiyu." Climate Dynamics **53**(5-6): 2643-2662.

Dudhia, J. (1989). "Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model." Journal of the Atmospheric Sciences **46**(20).

Esmaili, R., N. Smith, M. Schoeberl and C. Barnet (2020). "Evaluating Satellite Sounding Temperature Observations for Cold Air Aloft Detection." Atmosphere **11**(12).

Eyre, J. R., S. J. English and M. Forsythe (2020). "Assimilation of satellite data in numerical weather prediction. Part I: The early years." Quarterly Journal of the Royal Meteorological Society **146**(726): 49-68.

Filei, A. A., A. I. Andreev and A. B. Uspensky (2021). "Using of a Neural Network Algorithm for Retrieval Temperature and Humidity Sounding of the Atmosphere from Satellite-Based Microwave Radiometer MTVZA-GY Measurements On-Board Meteor-M No. 2-2." Issledovanie Zemli iz Kosmosa(6): 83-95.

Guo, L., N. An and K. C. Wang (2016). "Reconciling the discrepancy in ground- and satellite-observed trends in the spring phenology of winter wheat in China from 1993 to 2008." Journal of Geophysical Research-Atmospheres **121**(3): 1027-1042.

He, J., S. Zhang and Z. Wang (2012). "The retrievals and analysis of clear-sky water vapor density in the Arctic regions from MWHS measurements on FY-3A satellite." Radio Science **47**.

Herman, B. M., M. A. Brunke, R. A. Pielke, Sr., J. R. Christy and R. T. McNider (2010). "Satellite Global and Hemispheric Lower Tropospheric Temperature Annual Temperature Cycle." Remote Sensing **2**(11): 2561-2570.

Hu, L., D. Deng, S. Gao and X. Xu (2016). "The seasonal variation of Tibetan Convective Systems: Satellite observation." Journal of Geophysical Research-Atmospheres **121**(10): 5512-5525.

Kain, J. S. (2004). "The Kain-Fritsch convective parameterization: An update." Journal of Applied Meteorology **43**(1): 170-181.

Li, H., Y. Hu, Z. Zhou, J. Peng and X. Xu (2018). "Characteristic Features of the Evolution of a Meiyu Frontal Rainstorm with Doppler Radar Data Assimilation." Advances in Meteorology **2018**.

Li, H., X. Xu, Y. Hu, Y. Xiao and Z. Wang (2018). "Assimilation of Doppler Radar Data and Its Impact on Prediction of a Heavy Meiyu Frontal Rainfall Event." Advances in Meteorology **2018**.

 doi:10.20944/preprints202605.0954.v1

Lian, X., Z. Zeng, Y. Yao, S. Peng, K. Wang and S. Piao (2017). "Spatiotemporal variations in the difference between satellite-observed daily maximum land surface temperature and station-based daily maximum near-surface air temperature." Journal of Geophysical Research-Atmospheres **122**(4): 2254-2268.

Lu, Q., X. Yang, C. Wu, J. Zheng, D. Qin, H. Yang, P. Zhang and A. Electromagnetics (2011). An Initial Study on Assimilating Satellite-derived Total Precipitable Water in a Variational Assimilation System. Progress In Electromagnetics Research Symposium, Suzhou, PEOPLES R CHINA.

Lv, M., Z. Xu and Z.-L. Yang (2020). "Cloud Resolving WRF Simulations of Precipitation and Soil Moisture Over the Central Tibetan Plateau: An Assessment of Various Physics Options." Earth and Space Science **7**(2).

Maggioni, V., R. H. Reichle and E. N. Anagnostou (2013). "The Efficiency of Assimilating Satellite Soil Moisture Retrievals in a Land Data Assimilation System Using Different Rainfall Error Models." Journal of Hydrometeorology **14**(1): 368-374.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono and S. A. Clough (1997). "Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave." Journal of Geophysical Research **102**: 16666-166682.

Orlandi, E., F. Fierli, S. Davolio, A. Buzzi and O. Drofa (2010). "A nudging scheme to assimilate satellite brightness temperature in a meteorological model: Impact on representation of African mesoscale convective systems." Quarterly Journal of the Royal Meteorological Society **136**(647): 462-474.

Raju, A., A. Parekh, P. Sreenivas, J. S. Chowdary and C. Gnanaseelan (2015). "Estimation of Improvement in Indian Summer Monsoon Circulation by Assimilation of Satellite Retrieved Temperature Profiles in WRF Model." Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing **8**(4): 1591-1600.

Shi, X., Y. Li, J. Liu, X. Xiang and L. Liu (2018). "Simulation of FY-2D infrared brightness temperature and sensitivity analysis to the errors of WRF simulated cloud variables." Science China-Earth Sciences **61**(7): 957-972.

Wang, K. C., R. E. Dickinson, M. Wild and S. Liang (2012). "Atmospheric impacts on climatic variability of surface incident solar radiation." Atmospheric Chemistry and Physics **12**(20): 9581-9592.

Xu, L., H. Liu, Q. Du and X. Xu (2019). "The assessment of the planetary boundary layer schemes in WRF over the central Tibetan Plateau." Atmospheric Research **230**.

Yang, S., W. Zhang, B. Chen, X. Xu and R. Zhao (2020). "Remote moisture sources for 6-hour summer precipitation over the Southeastern Tibetan Plateau and its effects on precipitation intensity." Atmospheric Research **236**.

Zhang, S., X. Xu, S. Peng, W. Yao and T. Koike (2014). "Three-Dimensional Variational Data Assimilation Experiments for a Heavy Rainfall Case in the Downstream Yangtze River Valley Using Automatic Weather Station and Global Positioning System Data in Southeastern Tibetan Plateau." Journal of the Meteorological Society of Japan **92**(5): 483-500.

Zhang, Y., L. Zhao, W. Wang, S. Tang and F. Huang (2018). Summer ozone variation derived from FY3/TOU satellite data and impacts of East Asian summer monsoon. Conference on Remote Sensing of the Atmosphere, Clouds, and Precipitation VII, Honolulu, HI.

Zhao, Y., X. Xu, B. Chen and Y. Wang (2016). "The upstream "strong signals" of the water vapor transport over the Tibetan Plateau during a heavy rainfall event in the Yangtze River Basin." Advances in Atmospheric Sciences **33**(12): 1343-1350.

Zhao, Y., X. Xu, T. Zhao and X. Yang (2019). "Effects of the Tibetan Plateau and its second staircase terrain on rainstorms over North China: From the perspective of water vapour transport." International Journal of Climatology **39**(7): 3121-3133.