



Article

A Study on the Retrieval of Temperature and Humidity Profiles Based on FY-3D/HIRAS Infrared Hyperspectral Data

Chunming Zhang ^{1,2,3} , Mingjian Gu ^{1,3,4,*}, Yong Hu ^{1,3,4}, Pengyu Huang ^{1,2,3}, Tianhang Yang ^{1,3}, Shuo Huang ^{1,3}, Chunlei Yang ⁴ and Chunyuan Shao ^{1,3}

- ¹ Key Laboratory of Infrared System Detection and Imaging Technologies, Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China; zhangchunming@mail.sitp.ac.cn (C.Z.); huyong@mail.sitp.ac.cn (Y.H.); huangpengyu18@mailsucas.ac.cn (P.H.); yangtianhang@mail.sitp.ac.cn (T.Y.); huangshuo@mail.sitp.ac.cn (S.H.); shaochunyu@mail.sitp.ac.cn (C.S.)
- ² University of Chinese Academy of Sciences, Beijing 100049, China
- ³ Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China
- ⁴ Suzhou Academy, Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Suzhou 215000, China; yangcl@sitpfy.com
- * Correspondence: gumingjian@mail.sitp.ac.cn

Abstract: Satellite infrared hyperspectral instruments can obtain a wealth of atmospheric spectrum information. In order to obtain high-precision atmospheric temperature and humidity profiles, we used the traditional One-Dimensional Variational (1D-Var) retrieval algorithm, combined with the information capacity-weight function coverage method to select the spectrum channel. In addition, an Artificial Neural Network (ANN) algorithm was introduced to correct the satellite observation data error and compare it with the conventional error correction method. Finally, to perform the temperature and humidity profile retrieval calculation, we used the FY-3D satellite HIRAS (Hyperspectral Infrared Atmospheric Sounder) infrared hyperspectral data and combined the RTTOV (Radiative Transfer for TOVS) radiative transfer model to build an atmospheric temperature and humidity profile retrieval system. We used data on the European region from July to August 2020 to carry out the training and testing of the retrieval system, respectively, and used the balloon-retrieved sounding data of temperature and humidity published by the University of Wyoming as standard truth values to evaluate the retrieval accuracy. Our preliminary research results show that, compared with the retrieval results of conventional deviation correction, the introduction of ANN algorithm error correction can improve the retrieval accuracy of the retrieval system effectively and the RMSE (Root-Mean-Square Error) of the temperature and humidity has a maximum accuracy of improvement of about 0.5 K (The K represents the thermodynamic temperature unit) and 5%, respectively. The temperature and humidity results obtained by the retrieval system were compared with Global Forecast System (GFS) forecast data. The retrieved temperature RMSE was less than 1.5 K on average, which was better than that for the GFS; the humidity RMSE was less than 15% as a whole, and better than the forecast profile between 100 hpa (1 hpa is 100 pa, the pa represents the air pressure unit) and 600 hpa. Compared with AIRS (Atmospheric Infrared Sounder) products, the result of the retrieval system also had a higher accuracy. The main improvement of the temperature was at 200 hpa and 800 hpa, with maximum accuracy improvements of 2 K and 1.5 K, respectively. The RMSE of the humidity retrieved by the system was also better than the AIRS humidity products at most pressure levels, and the error of maximum difference could reach 15%. After combining the two algorithms, the FY-3D/HIRAS infrared hyperspectral retrieval system could obtain higher-precision temperature and humidity profiles, and relevant results could provide a reference for improving the accuracy of business products.

Keywords: temperature and humidity profiles; FY-3D/HIRAS; retrieval system; infrared hyperspectral; remote sensing



Citation: Zhang, C.; Gu, M.; Hu, Y.; Huang, P.; Yang, T.; Huang, S.; Yang, C.; Shao, C. A Study on the Retrieval of Temperature and Humidity Profiles Based on FY-3D/HIRAS Infrared Hyperspectral Data. *Remote Sens.* **2021**, *13*, 2157. <https://doi.org/10.3390/rs13112157>

Academic Editors: Filomena Romano and Elisabetta Ricciardelli

Received: 29 April 2021
Accepted: 29 May 2021
Published: 31 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Temperature and humidity profiles are important atmospheric parameters. High-accuracy temperature and humidity profiles improve the accuracy of numerical weather prediction. In order to make up for the low spatial resolution and uneven distribution of measured sites, many countries have successively developed infrared hyperspectral vertical sounders on satellites; the detected spectral information can fill the gaps in areas where relevant data are scarce [1–6].

The FY-3D satellite was successfully launched on 15 November 2017. The infrared hyperspectral atmospheric sounder (HIRAS) on board, which was developed by the Shanghai Institute of Technical Physics, Chinese Academy of Sciences, has 2287 channels [7]. FY-3D/HIRAS is the first infrared hyperspectral detector of polar orbiting meteorological satellites in China. Reports from multiple numerical forecast centers around the world show that the infrared hyperspectral detection instrument load on the polar orbiting satellite has the greatest impact on the assimilation forecast effect in terms of the contribution rate of a single instrument to the model [8,9]. The use of satellite remote sensing data has accounted for more than 90% of the available observational data of the ECMWF (European Centre for Medium-Range Weather Forecasts) numerical forecast operational system, and it is already widely used in global numerical weather forecasting, climate change forecasting, environmental monitoring, and other fields, producing great social benefits [9].

Based on satellite observation data, combined with the infrared radiative transfer model and retrieval algorithm, the radiation value received by satellites can be used to retrieve temperature and humidity profiles [10]. At present, the main retrieval methods include Feature Vector Regression algorithms, Artificial Neural Network (ANN) algorithms, and physical retrieval algorithms. The feature vector regression algorithm usually uses relatively stable atmospheric profiles, such as reanalysis profiles or sounding profiles, to establish a relevant statistical relationship with the simulated radiation value, and uses specific channels (such as carbon dioxide or water vapor) to retrieve temperature and humidity profiles. However, the physical process of the atmospheric radiative transfer model is not considered. Generally, the retrieval profile is used as the initial guess of the physical retrieval algorithm [11].

Cai et al. [12], based on the artificial neural network algorithm, used FY-4A/GIIRS (Geosynchronous Interferometric Infrared Sounder) to retrieve temperature and humidity profiles, and used AIRS products to verify the result of the retrieval. The results showed that the temperature RMSE was less than 1 K in the troposphere, and the humidity was less than 10% below 200 hpa. Huang et al. [13] used ANN algorithms combined with an improved one-dimensional variational (1D-Var) algorithm to retrieve temperature profiles at different pressure levels using FY-4A/GIIRS data, and they used the public sounding data of the University of Wyoming as ground-truth values to evaluate the retrieval accuracy. The results showed that the RMSE error of temperature profiles was less than 2 K, and the retrieval accuracy was better than the temperature products released by GIIRS.

Zhu et al. [14] used FY-3D/HIRAS to retrieve atmospheric temperature and humidity profiles based on a 1D-Var retrieval algorithm. In their experiment, the error average E between the observed brightness temperature and the model simulated brightness temperature was defined, and the error average E was used to correct the observation data. The corrected observation data were used for retrieval. The results showed that the accuracy of temperature and humidity profiles was better than the background profile. The RMSE error of the temperature profiles was about 0.5 K–2.6 K, and the RMSE of the humidity profiles was less than 20%. The 1D-Var retrieval algorithm takes into account the atmospheric radiative transfer path and has high retrieval accuracy, but the error between the observed brightness temperature and the model simulation's brightness temperature greatly limits the improvement of the retrieval accuracy in the algorithm. Guan et al. [15] used IASI infrared hyperspectral data to retrieve temperature and humidity profiles based on the 1D-Var retrieval algorithm, and corrected the deviation of the IASI observation data

using the error mean E. The retrieval results show that the RMSE error of the results after the deviation correction was smaller than before the deviation correction [16].

Through the above analysis, in the conventional 1D-Var retrieval algorithm, we focus on introducing a Neural Network algorithm to correct the observation data error, using the information capacity-weight function coverage method to select temperature and humidity channels, and combine the RTTOV radiative transfer model to calculate the background error covariance and the observation error covariance matrix. Finally, we constructed an FY-3D/HIRAS atmospheric temperature and humidity retrieval system. We used the infrared hyperspectral data of Western Europe in late July 2020 to train the system model, and used the same region data in early August 2020 to verify and test the system. In the actual retrieval experiment of atmospheric parameters, the reanalysis data of the day cannot be imported into the retrieval system in time; that is, the background information is more difficult to obtain. Therefore, the background information used by our retrieval system was collected from the training data set, which fits the actual business operation effectively. We mainly used the sounding data published by the University of Wyoming as the standard true values, and compared the retrieval results with the retrieval results produced by the conventional deviation correction method, so as to justify the neural network error correction introduced in this experiment. We also compared the retrieval results with background forecast field data and AIRS temperature and humidity products to analyze and evaluate the accuracy and retrieval capabilities of the FY-3D/HIRAS retrieval system.

2. Materials and Methods

2.1. Datasets and the RTTOV

2.1.1. FY-3D/HIRAS Data

In this study, we used the FY-3D/HIRAS L1 satellite data (<http://satellite.nsmc.org.cn/portalsite/default.aspx>, accessed on 15 May 2021). The selected area for this experiment is a part of Western Europe, and the data were recorded from July 15th to 31st, 00:00–01:00 UTC, and from 1–5 August 2020, 00:00–01:30 UTC. The relevant performance and parameters of HIRAS are shown in the following Table 1:

Table 1. FY-3D/HIRAS spectral characteristics and detection indicators.

FY-3D/HIRAS			
Performance and Parameters	>Wavenumber (cm ⁻¹)	Spectral Resolution (cm ⁻¹)	Number of Channels
Spectral Characteristics	Long Wave: 648.75~1134	0.625	781
	Medium Wave1: 1208.5~1749.375	0.625	869
	Medium Wave2: 2153.75~2549.375	0.625	637
Detection Indicators	Scan cycle		10s
	Field of view		1.1°
	Pixel/scan line		58
	Maximum scanning angle		±50.4°
	Radiative calibration accuracy		0.7 K
	Spectral calibration accuracy		7 ppm

As shown in Table 1, HIRAS has three wavebands: Long Wave, Medium Wave1, and Medium Wave2, with a total of 2275 spectral channels. It has 4 probe array detectors in each waveband to observe the target area simultaneously. Each probe element has an opening angle of 1.1°, and the instantaneous field of view of each probe element corresponding to the nadir is about 16 km [7,17,18].

2.1.2. GFS Forecast Data

In this experiment, we used GFS forecast data as a priori background information. During the retrieval, the GFS forecast data were used as initial profiles to input the retrieval system. GFS forecast data are provided by the National Center for Environmental Prediction (NCEP) (<https://www.ncei.noaa.gov/has/HAS.FileAppRouter?datasetname=GFSGRB24&subqueryby=STATION&applname=&outdest=FILE>, accessed on 15 May 2021), the forecast time limit is 6 h, the spatial resolution is 0.5° , and the GFS data were recorded from 15 July to 15 August 2020, at 00:00UTC.

2.1.3. ERA5 Reanalysis Data

The reanalysis data we selected were mainly the ERA5 daily average datasets from the ECMWF (<https://apps.ecmwf.int/datasets/>, accessed on 15 May 2021), dated from 5 July to 5 August, 00:00–01:00UTC, and includes specific humidity, temperature profiles, 2-m temperature, 2-m humidity, surface temperature, the 10-m wind U vector, the 10 m wind V vector, surface pressure, and other surface information and pressure-level information.

2.1.4. Sounding Data

The sounding data of the temperature and humidity profile in the experiment were provided by the University of Wyoming (<http://weather.uwyo.edu/upperair/sounding.html>, accessed on 15 May 2021). We used the sounding data as standard true values to verify the accuracy of the retrieval results. The sounding balloons are launched at midnight and noon every day. We mainly selected the sounding data from Western Europe at 00:00 UTC in late July and early August 2020 in our experiment.

2.1.5. RTTOV Radiative Transfer Model

RTTOV was developed from the TOVS (TIROS Operational Vertical Sounder) fast radiative transfer model (<https://nwp-saf.eumetsat.int/site/software/rttov/>, accessed on 15 May 2021). RTTOV is a fast radiative transfer model that can be used to simulate satellite observation data in visible light, infrared, microwave, and other wavebands [19,20]. This model is based on a given temperature and humidity profile, various gas concentrations, clouds, aerosols, and surface characteristics, combined with the satellite sensor channel response function, which can simulate the radiation value received by satellite sensors [21].

In this study, we used RTTOV for an FY-3D/HIRAS retrieval system for the related calculations. The FY-3D/HIRAS sensor coefficient file was downloaded from the official RTTOV website (<https://nwpsaf.eumetsat.int/site/software/rttov/download/coefficients/>, accessed on 15 May 2021). The RTTOV surface emissivity data were downloaded from the official website (https://nwpsaf.eumetsat.int/site/software/rttov/download/#Emissivity_BRF_atlas_data, accessed on 15 May 2021).

2.2. Main Principles and Methods of the System

2.2.1. Information Capacity-Weight Function Coverage Channel Selection

The HIRAS infrared hyperspectral sounder has 2275 spectral channels. Using all channels for retrieval cannot improve the retrieval accuracy. On the contrary, too many channels will increase the retrieval error [22]. This is because some channels have similar feature information, there is a large correlation between channels, and some channels will be affected by multiple atmospheric parameters. Therefore, it is particularly important to select a suitable channel for atmospheric parameter retrieval. The channel used for retrieval in this study was selected based on the information entropy capacity analysis and the peak of the weight function; that is, we selected the channel containing more atmospheric parameter information (temperature and humidity), and combined the peak position, peak size, and curve steepness of the channel weight function.

Shannon entropy can be used to express the uncertainty of information [22]. The difference of information entropy before and after the retrieval is defined as D , which represents the information capacity observed by the satellite. The background error covariance

is S_{ap} in the following Formula (1), and \hat{S} is the error covariance matrix of the retrieval parameters:

$$D = -\frac{1}{2} \log(\hat{S} S_{ap}^{-1}) \quad (1)$$

For the satellite retrieval process, the atmospheric state column vector constitutes the vector x , and the estimated error covariance of the parameters is estimated as Formula (2) and Formula (3), where K is the Jacobian matrix:

$$\hat{S} = S_{ap} - S_{ap} K^T \cdot (K S_{ap} K^T + S_{\epsilon})^{-1} \cdot K S_{ap} \quad (2)$$

$$\hat{S}^{-1} = (K S_{\epsilon} K^T + S_{ap})^{-1} \quad (3)$$

Some theories and parameter settings in our experiments in this section mainly refer to the article [23]. We set the initial error covariance matrix \hat{S} to S_{ap} , and the iterative process of the information entropy distribution is as follows Figure 1:

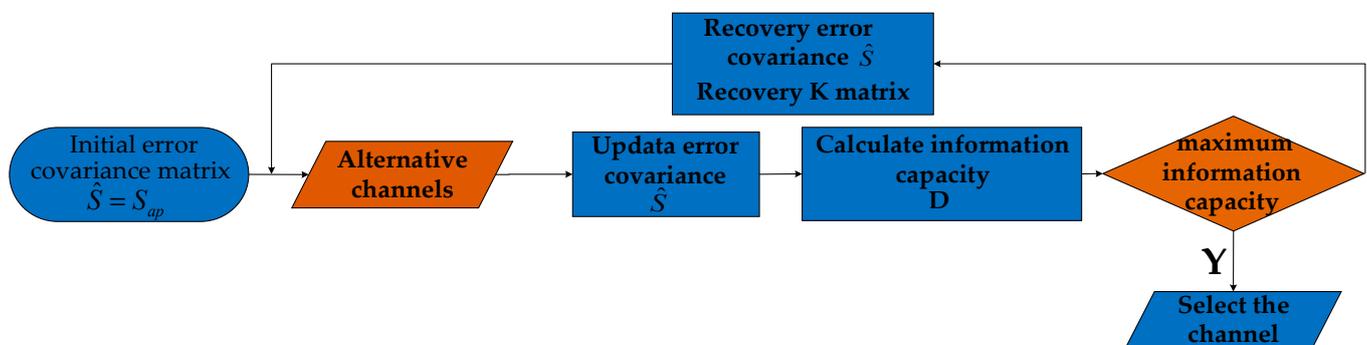


Figure 1. Iterative process of information entropy capacity.

As shown in Figure 1, in each iteration, we select a channel from the alternative channels and update the error covariance matrix and the information capacity D separately according to Formulas (1)–(3). The related calculation of error covariance is introduced in Section 3.3. We compare the information capacity of all alternative channels after a loop iteration and select the channel with the largest information capacity as the selected channel in this iteration. After multiple iterations, the channels are arranged in the order of iteration. In the end, 99 temperature channels and 50 humidity channels were selected for this experiment.

The weight function curve is the first derivative of the transmittance with the vertical height. The peak value of the weight function indicates the strongest absorption position on the radiation transmission path; that is, the peak position is the position at which the specific pressure level can be detected by the corresponding spectral channel [22]. This study introduces the peak height of the weight function as the channel selection criterion. As shown in Figure 2, the Jacobian matrix was normalized according to the maximum value to obtain the normalized weight function. It can be seen that the peak value of the weight function of temperature and humidity almost covers most of the pressure levels, which could be because the channels we selected were combinations after optimization. The position of the selected channel of temperature and humidity in the spectrogram is shown in the Figure 3:

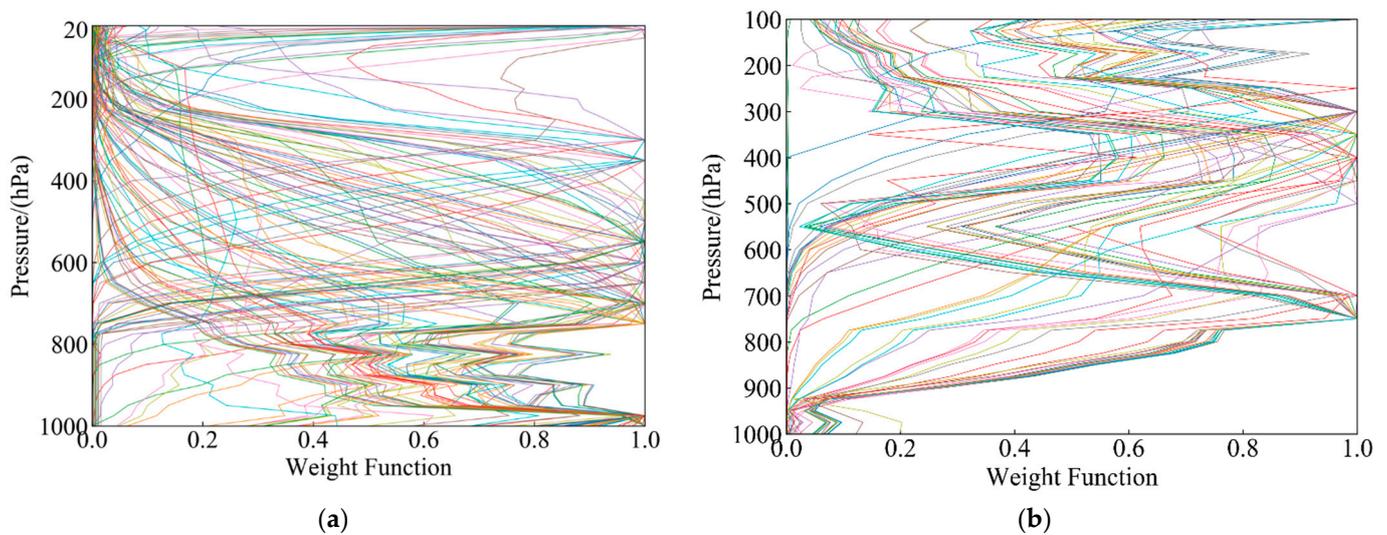


Figure 2. Normalized weight function. (a) Temperature; (b) humidity.

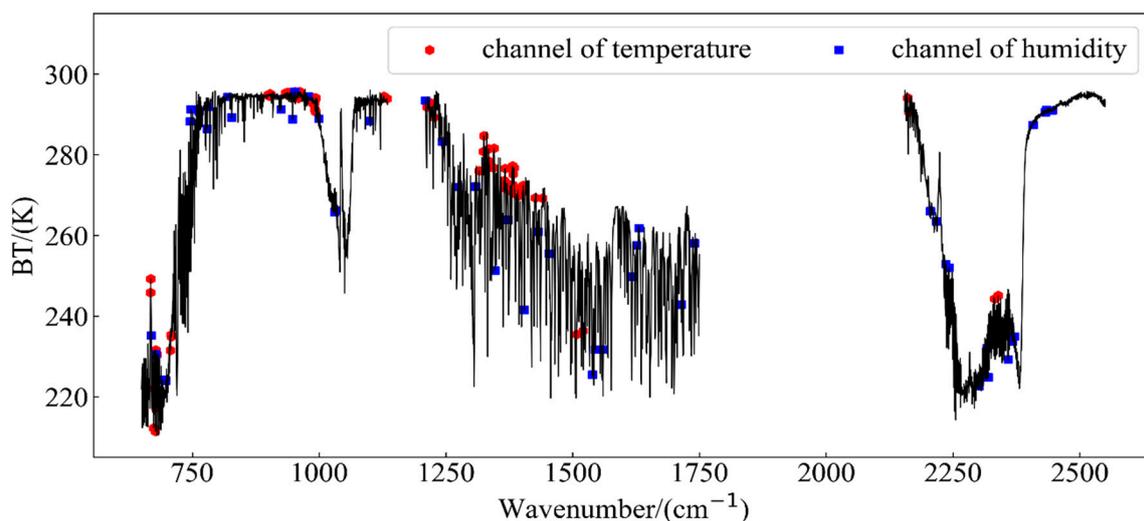


Figure 3. The position of channel selection about temperature and humidity in the spectrum.

As shown in Figure 3, the channels of temperature are mostly concentrated in the CO₂-sensitive channels of Long Wave and Medium Wave1, while the channels of water vapor are distributed in the entire wave band.

2.2.2. Neural Network Observation Error Correction

The 1D-Var retrieval algorithm and related variational assimilation theories assume that the observation field is unbiased. However, due to errors in the radiative transfer model and errors in the satellite observation data, the observation field data used in the actual retrieval often have systematic deviations, which limits the full use of observation data and affects the accuracy of the retrieval results. Therefore, before the observational data are used in a retrieval experiment, they need to be corrected for their deviation [24–26]. The observation correction algorithm in the conventional 1D-Var retrieval algorithm is shown in the following formula [14]:

$$E = \frac{\sum(Y - F(x))}{n} \quad (4)$$

where $F(x)$ is the forward simulated brightness temperature, Y is the satellite observation, x is the atmospheric state parameter (temperature and humidity profiles), and n is the number of samples. The conventional observation correction algorithm is used to calculate the average error E between the simulated brightness temperature and the observed brightness temperature. When the observation data are biased, the average error E needs to be subtracted from the observed brightness temperature to achieve the purpose of correcting the observed brightness temperature. The calculation of the observation error covariance is also calculated based on the observed brightness temperature corrected by the average error E .

The neural network algorithm has strong nonlinear processing capabilities and good fault tolerance [27]. In this experiment, the neural network algorithm was used to correct the observation error. The specific principle of the algorithm is shown in the Figure 4:

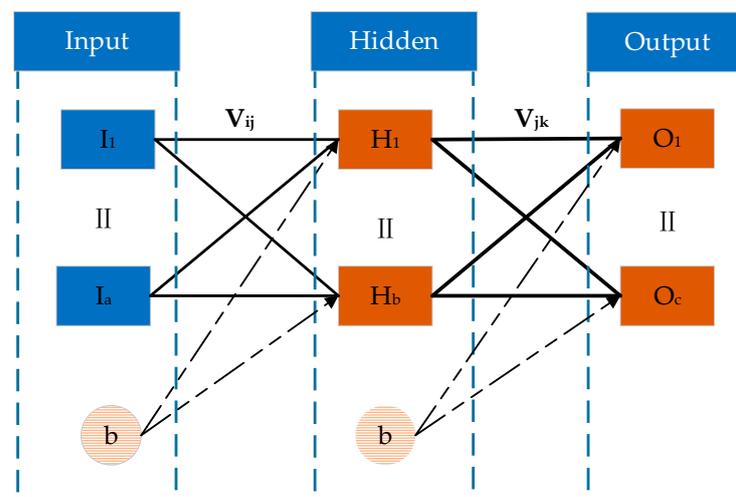


Figure 4. Schematic diagram of the neural network algorithm principle.

In Figure 4, I is the neural network input data (Input), H is the output data of the neural network training hidden layer (Hidden), O is the output data of the neural network training output layer (Output). In this experiment, a hidden layer was set. The neurons in the adjacent layers were connected by weight V , intercept b , and an activation function. The activation function of the hidden layer is the sigmoid function [13]:

$$y = \frac{1}{1 + e^{-x}} \quad (5)$$

where x is the input, and y is the output. In the training process, we calculated the $RMSE_{ANN}$ between the output value and the actual output (Equation (6)). The weight vector and intercept vector were continuously updated and fed back by using Formulas (7) and (8) of the gradient descent algorithm, so that the actual output became increasingly closer to the target data [28]. The relevant calculations are as follows:

$$RMSE_{ANN} = \sum_{k=1}^c (T_k - O_k)^2 \quad (6)$$

$$V_{jk} = V_{jk} - \eta \cdot \frac{\partial RMSE_{ANN}}{\partial V_{jk}} = V_{jk} + \eta (T_k - O_k) \cdot O_k \cdot (1 - H_j) \quad (7)$$

$$V_{ij} = V_{ij} - \eta \cdot \frac{\partial RMSE_{ANN}}{\partial V_{ij}} = V_{ij} + \eta \left\{ \sum_{k=1}^c [(T_k - O_k) \cdot O_k \cdot (1 - H_j) \cdot V_{jk}] \right\} \cdot H_j \cdot (1 - H_j) \cdot I_i \quad (8)$$

where i represents the i -th neuron of the input, j represents the j -th neuron of the hidden layer, and k represents the k -th neuron of the output layer. In Formulas (7) and (8), I , O , and H are consistent as described in Figure 4 above, T is the target output data, η is the step size, V_{ij} and V_{jk} , respectively, represent the weights between the input layers and the hidden layers and the weights between the hidden layers and the output layers after the gradient descent algorithm. In the iterative process of ANN algorithm, the threshold of the gradient was set to 1×10^{-6} . If the value of gradient was less than 1×10^{-6} , the iteration was terminated. After multiple iterations, the final weight function was obtained, and the network model was saved. The output data were the radiation spectrum data after error correction.

2.2.3. D-Var Retrieval Algorithm

In the FY-3D/HIRAS retrieval system, the retrieval of the atmospheric temperature and humidity profiles was mainly based on the basic theory of the atmospheric radiation transfer equation. However, this equation is the first type of nonlinear Fredholm integral equation [29,30]. Because the weight function of the instrument has a certain width range, in the instrument noise range, there are multiple solutions that can satisfy the equation, which is a typical non-well-posed problem. The way to solve this kind of problem is to put deterministic or random restrictions on the parameters to be solved, which is to introduce prior information and expand the solvable domain. Non-well-posed problems are then transformed into well-posed problems, and we can finally focus on solving non-linear problems [31,32].

The 1D-Var retrieval algorithm takes into account the atmospheric radiation transfer process, and it combines the observation value, the a priori background value, and the error information. Non-linear equations are transformed into optimization problems. The 1D-Var algorithm mainly consists of two parts: (1) construct the objective function (cost function) and (2) determine the optimization strategy.

(1) Build the objective function

Based on the above analysis, in our study, we used a 1D-Var retrieval algorithm, took the prior background profile x_b as the starting value (first guess profile), and adjusted the atmospheric state column vector X (that is, the atmospheric temperature and humidity profiles) continuously to minimize the objective function [33–36]:

$$J(x) = \frac{1}{2} \left[(Y^m - F(x))^T \cdot S_\varepsilon^{-1} \cdot (Y^m - F(x)) \right] + \frac{1}{2} \left[(x - x_b)^T \cdot \gamma S_{ap}^{-1} \cdot (x - x_b) \right] \quad (9)$$

where x_b is the GFS forecast data, γ is the smoothing factor, S_{ap} is the background error covariance, S_ε is the observation error covariance, $F(x)$ is the simulated observation, Y^m is the FY-3D/HIRAS observation, and the number of preferred channels participating in the retrieval is m .

(2) Optimization theory—Newton's nonlinear iteration

Newton's nonlinear iterative method was used to minimize the above objective function, and the temperature and humidity profiles were retrieved by step-by-step iteration. The general form of Newton's nonlinear iterative equation is as follows [37]:

$$x_{n+1} = x_n - \left[\nabla^2 J(x_n) \right]^{-1} \cdot \nabla J(x_n) \quad (10)$$

where $\nabla J(x_n)$ and $\nabla^2 J(x_n)$ represent the first-order derivative and the second-order derivative, respectively, which are specifically expressed as

$$\nabla J(x_n) = -K^T S_\varepsilon^{-1} \cdot [Y^m - F(x)] + \gamma S_{ap}^{-1} \cdot (x - x_b) \quad (11)$$

$$\nabla^2 J(x_n) = \gamma S_{ap}^{-1} + K^T \cdot S_\varepsilon^{-1} \cdot K - \frac{d^2 K^T}{dx} \cdot S_\varepsilon^{-1} \cdot [Y^m - F(x)] \quad (12)$$

where K is the Jacobian matrix calculated by the radiative transfer model. In the actual calculation, the HIRAS radiation value data Y^m and the forward simulation observation radiation $F(x)$ were constantly closed in the iteration. In order to reduce the calculation amount, the second-order partial derivative part of the objective function was usually ignored [38], and the final iterative equation is as follows [39]:

$$x_{n+1} = x^b + \left(\gamma S_{ap}^{-1} + K_n^T S_\varepsilon^{-1} K_n \right)^{-1} K_n^T S_\varepsilon^{-1} \left\{ [Y^m - F(x_n)] + K_n (x_n - x^b) \right\} \quad (13)$$

where n is the number of iterations, x_{n+1} and x_n are the profile results of the $(n+1)$ -th and the n -th iteration, respectively. In general, when $J(x_n) - J(x_{n+1})$ is less than a certain threshold in the iteration process, it is regarded as convergence. When the background field is not normally distributed or only a local normal distribution, it is easy to cause instability in the solution. The function of smoothing factor γ was used to adjust the influence of observation information and prior background information on the solution, so as to achieve the purpose of accelerating convergence [40].

2.2.4. Error Evaluation Method

In this experiment, we used the calculation of Mean Error (ME) and RMSE as the accuracy evaluation standards. It is known that the smaller the RMSE of the retrieval temperature and humidity profiles and the true value of the temperature and humidity profiles, the higher the accuracy of the retrieval algorithm [14,40]. The specific calculation formula is as follows:

$$ME = \frac{\sum_{i=1}^n (x_i - x'_i)}{n}, \quad (14)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - x'_i)^2}{n}}, \quad (15)$$

where x_i is the retrieval value; x'_i is the true value, and n is the number of samples.

Since the humidity profiles calculated by the RTTOV radiative transfer model in this experiment were specific humidity profiles, the unit is $\text{kg} \cdot \text{kg}^{-1}$. We need to convert specific humidity to relative humidity (%) to be applicable to the above error evaluation method. To avoid confusion, we use relative humidity to calculate humidity in all of the following sections (except for the calculation of the background error covariance matrix), and the conversion formula is as follows [14]:

$$e = p \frac{q}{0.622 + q} \quad (16)$$

$$e_s = 6.1078 \exp \left[\frac{17.2693882(T - 273.16)}{T - 38} \right] \quad (17)$$

$$RH = \frac{e}{e_s} \times 100\% \quad (18)$$

where e represents water vapor pressure, q represents humidity (specific humidity), p is the pressure layer, and the unit is hpa. e_s is the water vapor saturation pressure, T is the temperature, and RH is the relative humidity to be calculated.

2.3. The Specific Process of the Experiment

In our experiment, the FY-3D/HIRAS retrieval system for temperature and humidity profiles was built, which mainly includes system training data and test data preprocessing and retrieval system construction and verification. First, the FY-3D/HIRAS L1 data, GFS forecast data, and ERA5 reanalysis data were used to perform strict time and space matching by linear interpolation, and we then performed data preprocessing for HIRAS

observations, including clear-sky pixel selection and spectral apodization. Secondly, we used the RTTOV radiative transfer model to calculate the forward simulated brightness temperature and the Jacobian matrix, and we selected the channel based on the information entropy capacity-weight function coverage algorithm and performed neural network error correction on the observed brightness temperature after channel optimization. At this point, the training of the retrieval system model was complete. Finally, we used test data from 1–5 August 2020, as input into the FY-3D/HIRAS retrieval system to perform the retrieval experiment of the temperature and humidity profiles. We used the University of Wyoming sounding data as the standard truth values and compared the retrieval results based on conventional error correction methods with the retrieval results of our system. The relevant results can provide a foundation for introducing the neural network algorithm to the retrieval system. In order to test the reliability and stability of the retrieval system, we compared the retrieval results with GFS and AIRS temperature and humidity products for verification. A flow chart of the retrieval system is shown in Figure 5 below:

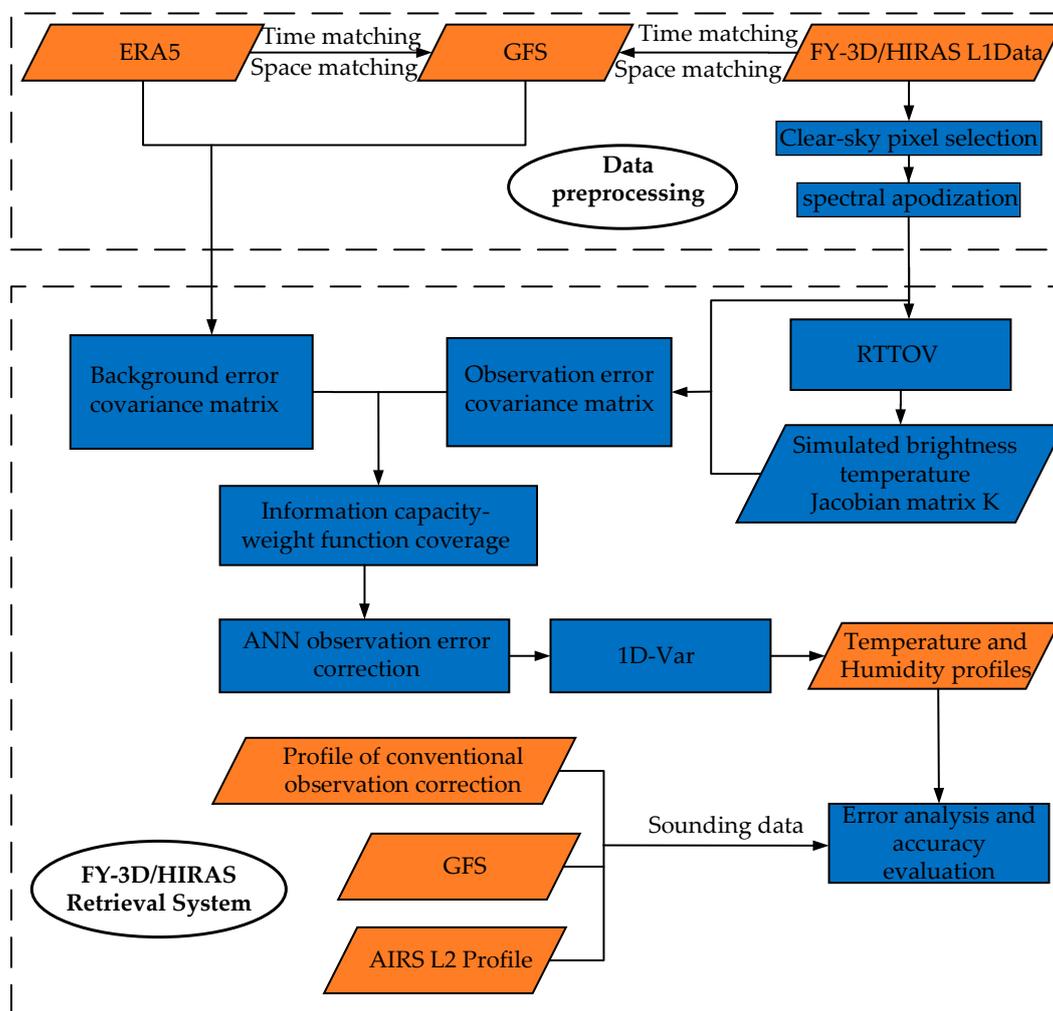


Figure 5. A flow diagram of the FY-3D/HIRAS retrieval system.

2.3.1. Data Preprocessing

The geographical location selected in this experiment has a longitude of 15° – 30° E and a latitude of 35° – 45° N. We used data from 15–31 July 2020, to train the retrieval system, and used data from 1–5 August 2020, to test and verify the system. In order to improve the quality of data usage, we needed to preprocess the data input to the retrieval system, which included the selection of clear-sky pixels of satellite data, spectral apodized operations, and

the matching of profile data and satellite data. HIRAS L1 business data are a high-precision nonapodized radiation value product produced after multi-step processing and spectral radiation calibration processing [41,42]. Therefore, before HIRAS L1 participates in the retrieval, we should convert the radiation values into brightness temperature values.

(1). Selection of clear-sky pixels

In the process of using satellite data in the retrieval of temperature and humidity profile, the presence of clouds in the observation field of view will cause a large change in brightness temperature, which will affect the accuracy of the retrieval. In order to reduce the influence of clouds in our experiment, we used the cloud detection product of FY-3D/MERSI-II (Moderate resolution spectral imager) on the same platform as HIRAS to select training set samples and test set samples of clear-sky pixels in the satellite data. The product is a judgment mark of cloud and clear sky obtained through a multi-feature threshold method [43–45]. The product needs to convert bytes into bit operations for reading. The classification of the product is 0 (cloud), 1 (possible cloud), 2 (possible clear sky), and 3 (clear sky). The location of the clear-sky point selected by the test set data of the retrieval system is shown in the Figure 6:

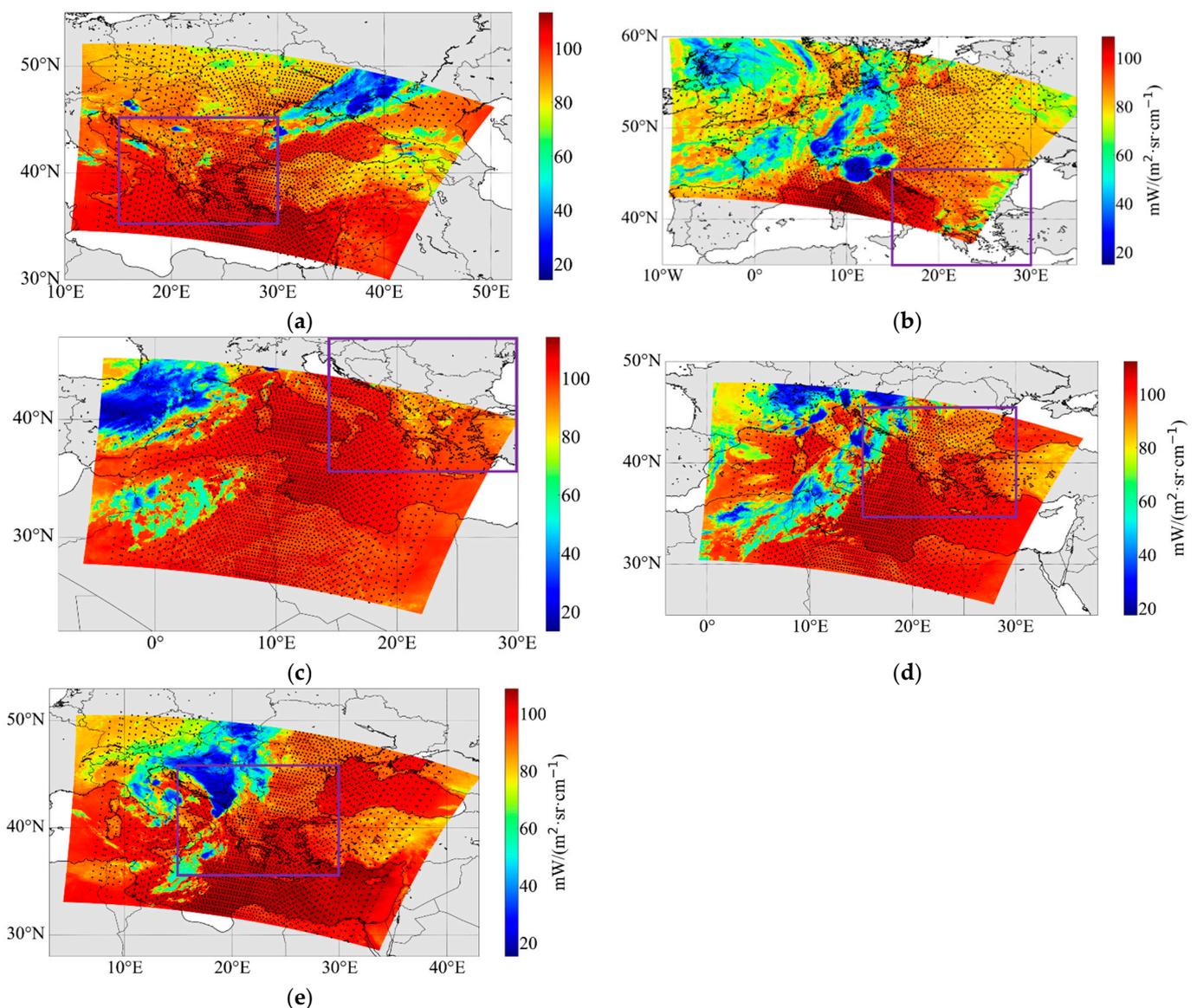


Figure 6. Clear-sky pixel selection. The satellite observation test data are as follows: (a) 1 August 2020, 00:05 UTC; (b) 2 August 2020, 01:25 UTC; (c) 3 August 2020, 01:10 UTC; (d) 4 August 2020, 00:50 UTC; (e) 5 August 2020, 00:30 UTC.

In Figure 6, the bottom image was the radiation data of the channel selected by MERSI-II. The black dots represented the selected clear-sky pixels, and the black dots in the purple frame are the selected HIRAS observation sample points for this experiment. As can be seen in Figure 6, most of the pixels in the cloud area with low brightness temperature were not selected, and the distribution of the pixels selected by the clear-sky selection was more consistent with the change in MERSI-II clear-sky brightness temperature.

In this experiment, we determined the final observation training set samples and test set samples of the retrieval system after the clear-sky selection. The GFS forecast data and ERA5 reanalysis data were linearly interpolated according to time, longitude, and latitude of HIRAS sample points. In terms of atmospheric pressure levels, we interpolated the forecast data and sounding data to the corresponding atmospheric pressure levels according to the reanalysis data of ERA5. Sounding data lack temperature and humidity information in the higher atmospheric pressure levels, and the water vapor content in the atmosphere will rapidly decrease with a pressure drop, the water vapor is mainly concentrated below 200 hpa and the water vapor content in the upper stratosphere is almost zero [9]. Therefore, in this experiment, we set the pressure levels of the temperature from 1000 hpa to 20 hpa, a total of 33 pressure levels, and the pressure levels of the humidity were from 1000 hpa to 100 hpa, a total of 27 pressure levels.

(2) Spectral data apodization

As mentioned above, HIRAS L1 level data have not undergone apodization. In actual business applications, in order to reduce the side lobe effect, it is necessary to perform apodization calculations on spectral radiation data. Referring to HIRAS L1 level business data [46], we selected the Hamming function as a function of apodization in this experiment. We apodized all the observed sample data, and a comparison of the spectral brightness temperature before and after processing is shown in the Figure 7:

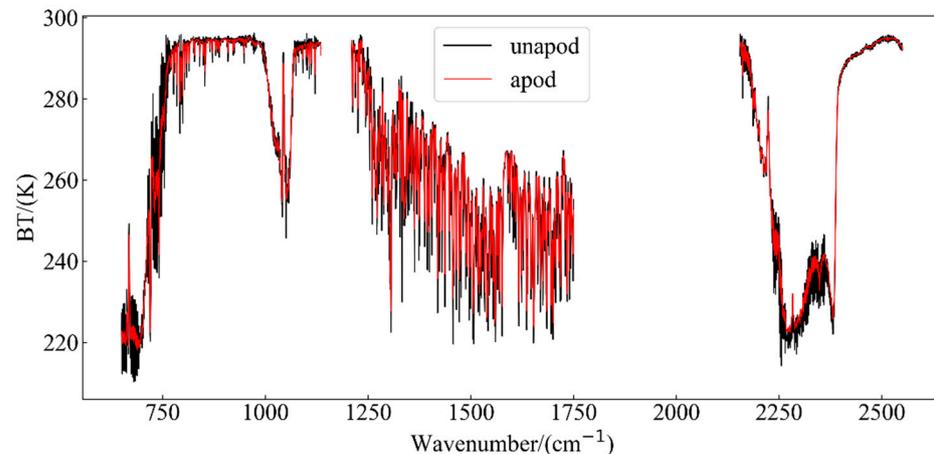


Figure 7. Spectral brightness temperature apodization before and after comparison.

In Figure 7 above, the black solid line represents the spectral brightness temperature before apodization, and the red solid line represents the spectral brightness temperature after apodization. It can be seen that the burr phenomenon is reduced when the HIRAS spectrum passes through apodization, and the spectrum is smoother.

2.3.2. Observation Error Correction

The ultimate goal of the 1D-Var retrieval algorithm is to find a set of atmospheric conditions that can minimize the difference between the satellite observation and the simulation brightness temperature [32]. Observation error correction based on a neural network algorithm before the retrieval experiment is performed brings the satellite observation closer to the simulation value, which can greatly reduce the error impact by satellite observations. We can then make full use of the observational information of the satellite's infrared hyperspectral data.

Since sounding sites are relatively sparse, there are relatively few acquired sounding data, so the data are not suitable as input data for the RTTOV model for obtaining the accurate simulated brightness temperature. The ERA5 reanalysis data are used as a standard to test the authenticity of the profiles in most of the literature [12–14,28]. It has a large amount of sample data and high temporal and spatial resolution, so the reanalyzed data were regarded as the reference profiles in this model training. In the matched 11,677 sets of training set and 4272 sets of test set samples, we used the temperature and humidity profile in the reanalysis data, the surface data in the reanalysis data, and the angle information of satellite observations to substitute into the RTTOV to obtain the simulated brightness temperature. In the training of the neural network, the target output was the simulated brightness temperature value of the training samples, and the corresponding satellite observed brightness temperature (the channel after the information capacity optimization) was used as the input. After the training was completed, we saved the network and used it for later testing.

The coefficient of determination R^2 was set as the evaluation index for regression model to evaluate the degree of agreement between the predicted value and the actual value. The coefficient of determination of the training set was 0.9680, and the coefficient of determination R^2 of the test set was 0.7792. Based on the experimental results and Figure 8, the training network was ideal and can be used for the error correction of observation data. We substituted 4272 sets of satellite observation test data from 1–5 August into the trained neural network to obtain the corrected observation data. The comparison before and after correction of the temperature and humidity channels is shown in the Figure 9.

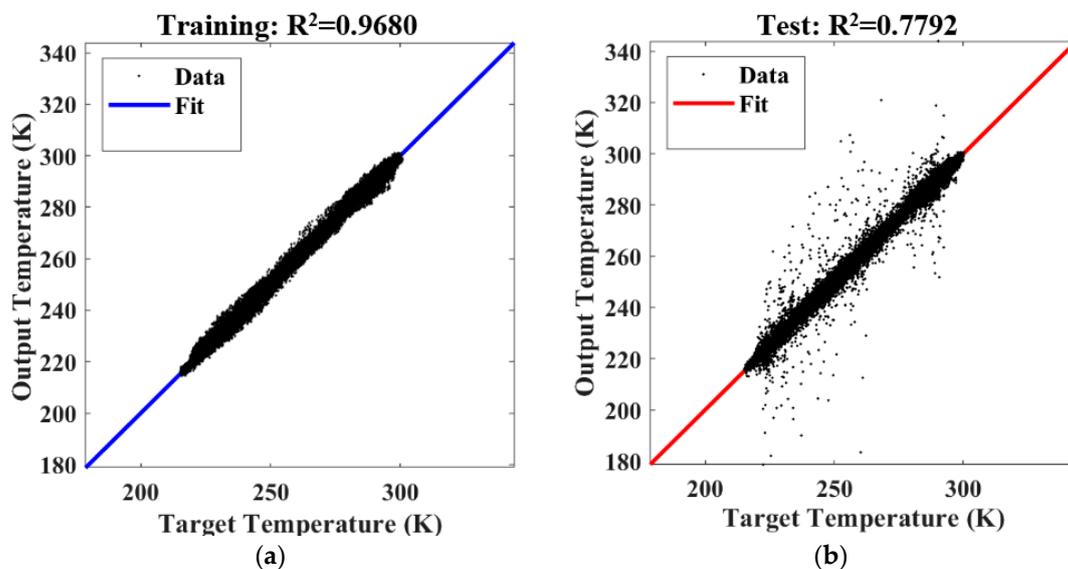


Figure 8. Result of training, Fit represents fitting results. (a) Correlation coefficient of the training data; (b) correlation coefficient of the test data.

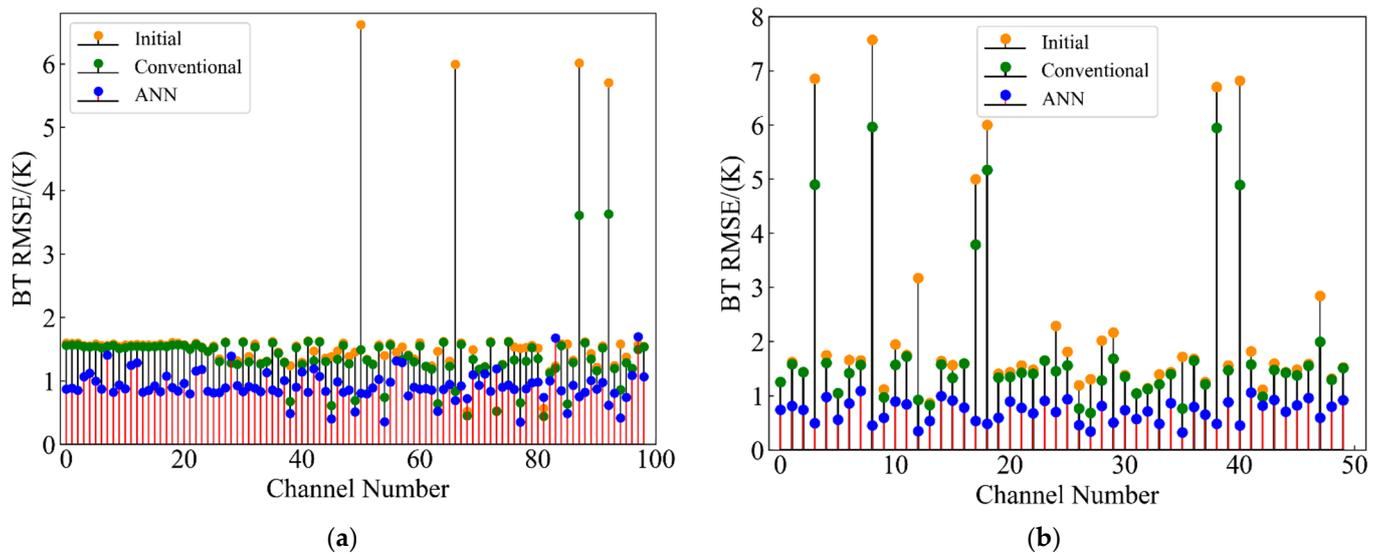


Figure 9. Channel observation error correction before and after comparison. (a) Temperature; (b) humidity.

We calculated the RMSE between the simulated brightness temperature and the observed brightness temperature. As shown in Figure 9 of the cotton swab, the yellow dots represent the uncorrected initial observation error distribution (Initial), the blue dots are the observation error distribution after the neural network correction (ANN), and the green dots represent the results of conventional deviation correction (Conventional). The RMSE of the brightness temperature of the temperature channel corrected by the neural network is less than about 2 K as a whole, and the RMSE of the brightness temperature of the humidity channel is less than 1.5 K as a whole. Compared with the conventional deviation correction method, neural network observation error correction has a certain extent of accuracy improvement in both temperature and humidity channels. The maximum RMSE of the brightness temperature difference between the two methods reached 3.5 K in the temperature channels, and the maximum RMSE of the difference was 5.5 K in the humidity channels. Figure 9 clearly shows that the neural network algorithm correction effect was stronger. Therefore, using a neural network method to correct the observation data of the selected channel can more effectively reduce the error between the observation data and the forward model simulation values.

2.3.3. Error Covariance Calculation

In the actual business temperature and humidity retrieval process, the reanalysis data of the day are difficult to obtain in time; therefore, in this experiment, we did not use the reanalysis data for retrieval in the system, and the reanalysis data were only used in the model training of the system. In this experiment, from the perspective of actual business operations, the system's training data set was used to calculate the error covariance.

(1) Observation error covariance calculation

Since the observation errors are regarded as independent of each other among the channels, the observation error covariance matrix is a diagonal matrix, and its diagonal elements are

$$\varepsilon = \frac{\sum(Y^m - F(x))^2}{n} \quad (19)$$

where Y^m is the FY-3D/HIRAS observed brightness temperature after neural network error correction, $F(x)$ is the simulated brightness temperature, and n is the number of samples. The error covariance of the satellite observation data on each channel is shown in the Figure 10:

As shown in Figure 10, Using the method introduced in Section 3.1, we selected 99 channels and 50 channels for temperature and humidity, respectively.

(2) Background error covariance calculation

The background error covariance matrix describes the error and correlation between the forecast value and the true value of each pressure level of the atmospheric state vector [47]. In this experiment, we used the GFS forecast data and ERA5 reanalysis data in the training data of the retrieval system model to calculate the deviation. The specific calculation method is as follows:

$$s_{ij} = \text{cov}(x^i, x^j) = \frac{1}{n} \sum_{k=1}^n [(x_k^i - E(x^i)) \cdot (x_k^j - E(x^j))] \quad (20)$$

where x represents the background error; x_k^i is the k -th sample data of the i -th level; $E(x^i)$ is the average error of the forecast value of the i -th level; n is the total number of samples.

Figure 11 shows that the background error covariance matrix of temperature and humidity has a larger value in the diagonal, which indicates that the farther the pressure levels are, the smaller the correlation between them. The temperature error covariance is less than 3 K^2 as a whole, and the overall humidity error covariance is better than $3.35 \times 10^{-6} (\text{kg}/\text{kg})^2$.

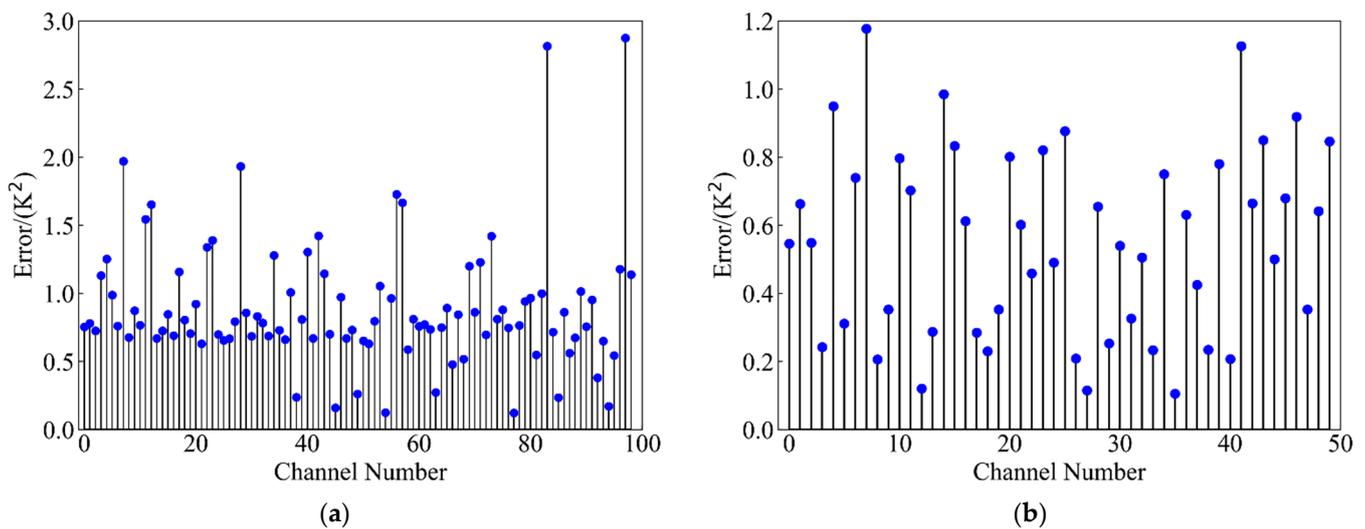


Figure 10. Observation error covariance. (a) Temperature; (b) humidity.

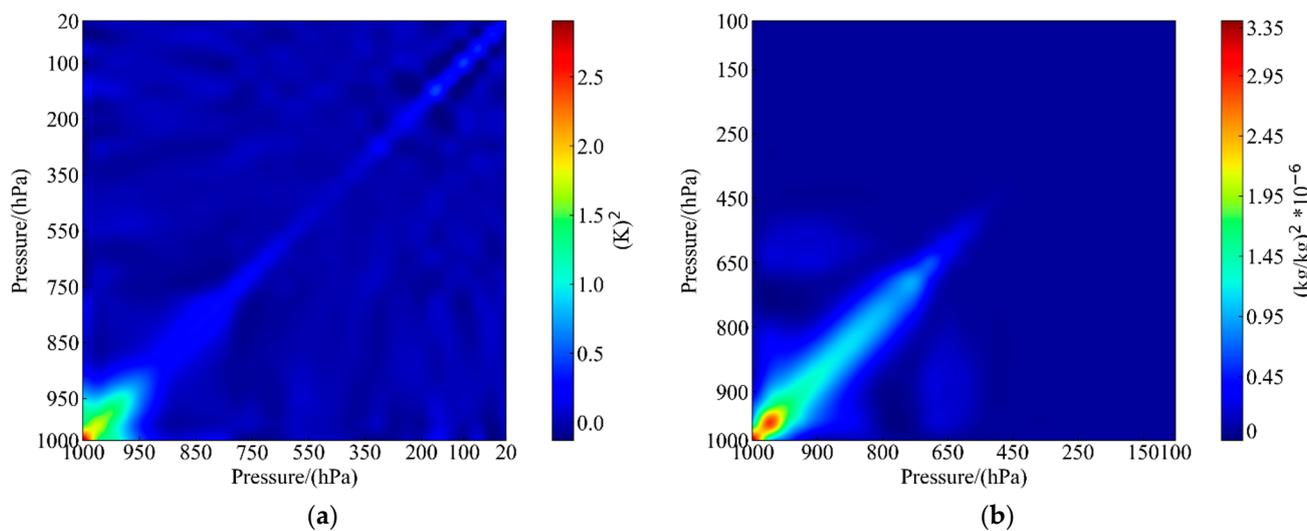


Figure 11. Background error covariance. (a) Temperature; (b) humidity.

3. Results

We used the training data set from July 15th to 31st, 2020, to train the retrieval system and obtained the relevant retrieval parameters. The test data set includes satellite data, GFS forecast data, and sounding data as standard truth values from 1–5 August 2020. In this experiment, the distribution of the sounding stations we selected is shown in Figure 12.

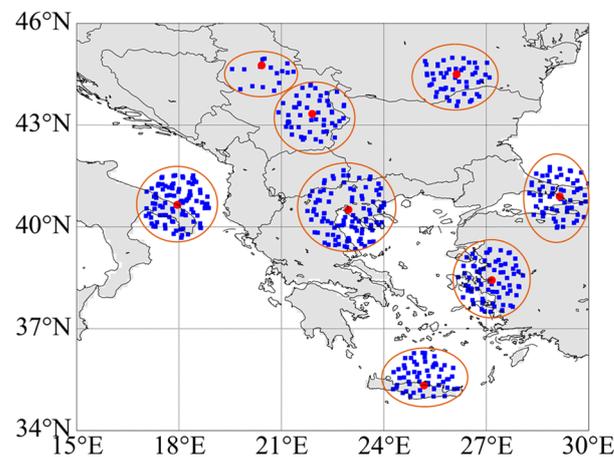


Figure 12. Schematic diagram of sounding sites and matching satellite observations.

The red dots are the sounding sites. We looked for HIRAS observation data points that satisfy Formula (21) and that have been preprocessed by the retrieval system around the sounding station. A sounding station should be matched by at least 4 observation sample points [13] and cannot be concentrated on one side. We selected 497 groups of HIRAS observation sample points (blue dots in the figure). Among these groups, we matched the time, longitude, and latitude of the GFS forecast data with the selected HIRAS observation samples through linear interpolation, and we finally obtained 497 sets of GFS profile data.

$$|lat_i - lat_j| + |lon_i - lon_j| < 1^\circ \quad (21)$$

where lat_i and lat_j , respectively, are the latitude and longitude of the observation data, and lat_j and lon_j , respectively, are the latitude and longitude of the sounding data.

3.1. Retrieval Results of Temperature and Humidity Profiles

The mean value of the retrieval results of temperature and humidity from 1–5 August 2020, is shown in the Figure 13. The red line represents the GFS forecast profile, the blue line is the temperature and humidity profiles retrieved by the system, and the green line is the sounding data.

As shown in Figure 13, in this experiment, our retrieval results of temperature and humidity profiles are consistent with the sounding data and forecast field data as a whole, and the trend of changes in temperature and humidity with altitude is the same as the trend of the sounding data and the forecast field data. The specific manifestations of temperature and humidity changes with height are as follows: the temperature gradually decreases as the height rises, and the pressure rises to around 100 hpa and then gradually increases. The relative humidity reaches its maximum at 1000 hpa near the surface. The humidity distribution in the remaining pressure levels decreases with the increase in altitude, and the water vapor content is almost zero above 100 hpa.

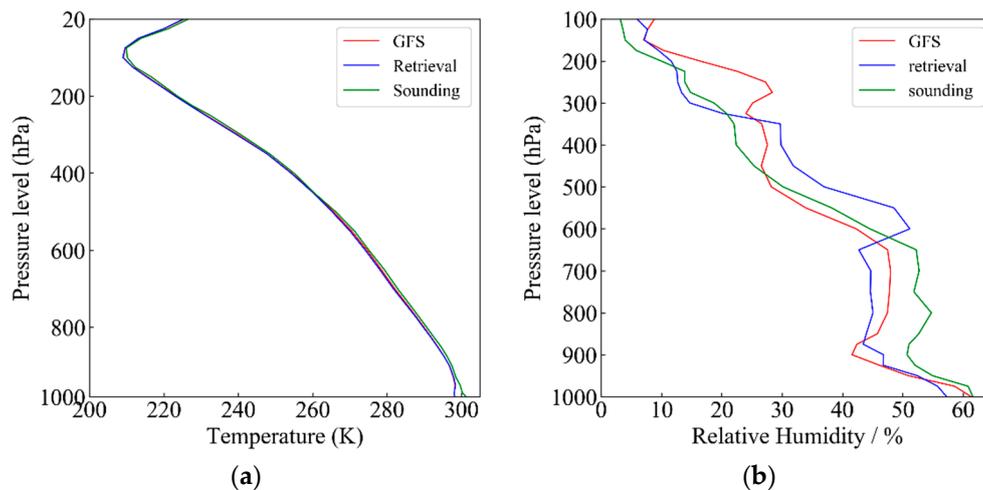


Figure 13. Comparison of retrieval results with sounding data (ground truth) and forecast data. (a) Temperature; (b) humidity.

3.2. Comparison of Retrieval Results with GFS and Conventional Deviation Correction Methods

At present, the international requirements for the retrieval accuracy of the temperature and humidity profiles are that the RMSE of temperature retrieval is better than 2 K as a whole (part of the pressure level is required to be less than 1 K), and the RMSE of humidity is less than 20% as a whole (part of the pressure level is required to be better than 10%) [12]. We used the retrieval results obtained from the satellite test set data matched with the sounding sites on 1 August 2020, to calculate the RMSE and ME. Secondly, we also calculated the RMSE and ME by using the retrieval results obtained by the conventional error correction method (see Formula (4), hereinafter referred to as the conventional method) and the forecast profiles with the sounding true values, respectively. The accuracy comparison results of the temperature and humidity among the three are shown in the Figure 14 below. The red is the error of forecast profiles, the blue is the error of our retrieval profiles, and the green is the error of the conventional method.

As shown in Figure 14a, the accuracy of the temperature profile results obtained by our retrieval system is higher than that of the forecast data and conventional methods as a whole and is closer to the sounding true value. The RMSE error is between 0.5 and 1.85 K, and the mean error is less than 1.5 K. The improvement of retrieval accuracy is the most obvious, near 1000 hpa and 350 hpa compared to the forecast profiles, and the improvement range is 1.8 K and 0.35 K, respectively. Compared with the conventional method, our overall accuracy is improved by 0.5 K, and the accuracy of the retrieval results meets the current temperature retrieval accuracy requirements. Figure 14c shows the humidity RMSE error curve. Similar to the temperature RMSE, there is a large error in humidity in the middle and lower troposphere from around 800 hpa to 1000 hpa, which may be due to the drastic change in water vapor in the lower atmosphere, and there is a large difference in a short period of time. Compared with forecast profiles and the conventional method, our retrieval humidity profiles show an accuracy improvement from 100 hpa to 600 hpa. The improvement extent is greater than 5% around 500 hpa. The RMSE of retrieval humidity is also better than the conventional methods as a whole. Our overall retrieval result is better than 15%, which also meets the current humidity retrieval requirements.

The closer the ME deviation value is to $x = 0$, the higher the accuracy is. Figure 14b,d show that the trends of our retrieved temperature and humidity profile ME curves are similar to those of the GFS forecast profile and conventional method. The ME curves of temperature and humidity are closer to $x = 0$ than the ME curve of the GFS forecast and the conventional method. Based on the above analysis, we can infer that the FY-3D/HIRAS retrieval system has a higher retrieval accuracy, and the temperature and humidity RMSE

error curve after neural network error correction is better than the conventional method at most pressure levels, which indicates that our method is more effective.

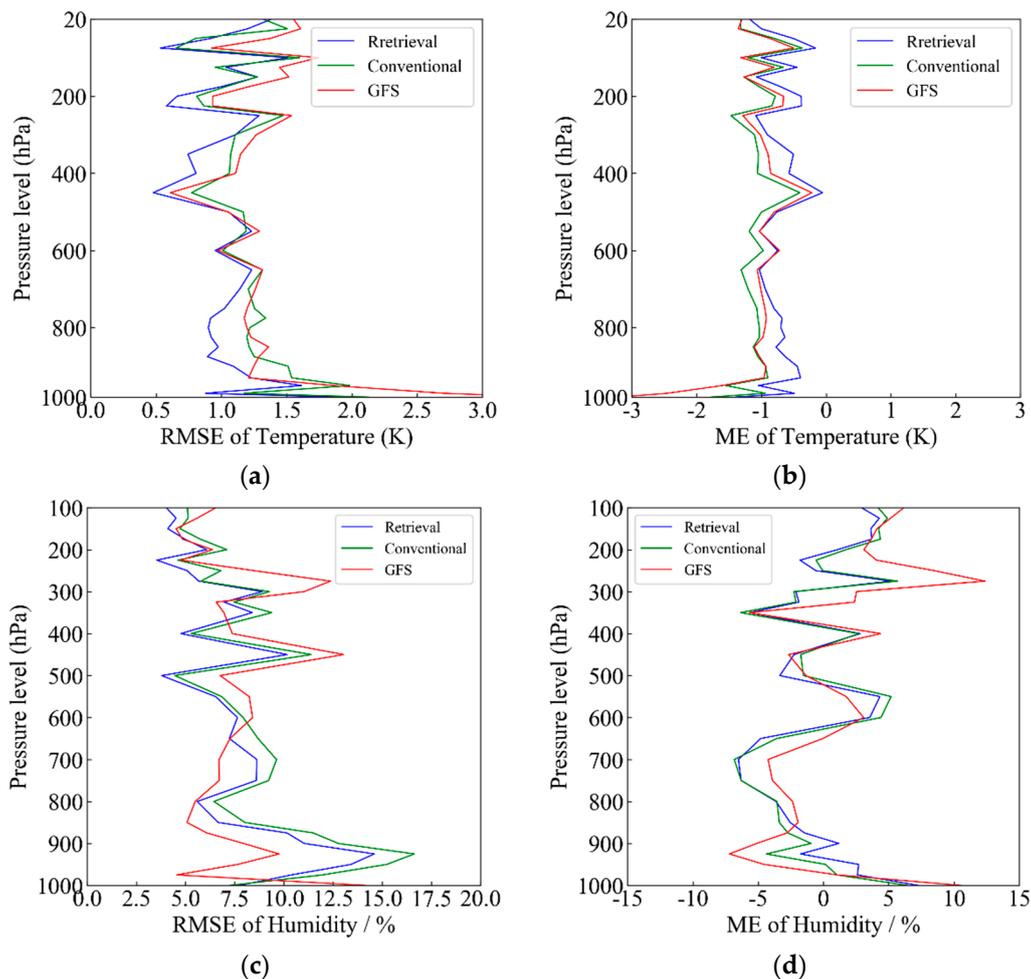


Figure 14. Comparison between the retrieval profile and GFS and retrieval profile by the conventional method. (a) RMSE of temperature; (b) ME of temperature; (c) RMSE of humidity; (d) ME of humidity.

3.3. Comparison of Retrieval Results with Similar, International Instruments

In order to improve the stability and reliability of the retrieval system, we also used the test sets matched with the sounding sites from 2–4 August 2020, to perform the temperature and humidity profile retrieval experiment, and we compared the retrieval results with the L2 products of similar, international instruments. The AIRS detector's L2 temperature and humidity profile products have been released to the V7 version [48] (<https://earthdata.nasa.gov/earth-observation-data/near-real-time/download-nrt-data/airs-nrt>, accessed on 15 May 2021). In our experiment, we selected AIRS L2 products from 2–5 August 2020, and mainly used the same area as HIRAS, with a time difference of ± 1.5 h. The AIRS temperature and relative humidity were both interpolated to the 33 temperature levels and 27 relative humidity levels using linear interpolation methods. We matched the time, latitude, and longitude of the AIRS products with 345 sets of HIRAS sample data using the same linear interpolation method and obtained 345 sets of AIRS temperature and humidity product data. In the end, we calculated the RMSE and ME by the retrieval results and the AIRS product profile with the sounding true values, and the calculation results are shown in the Figure 15:

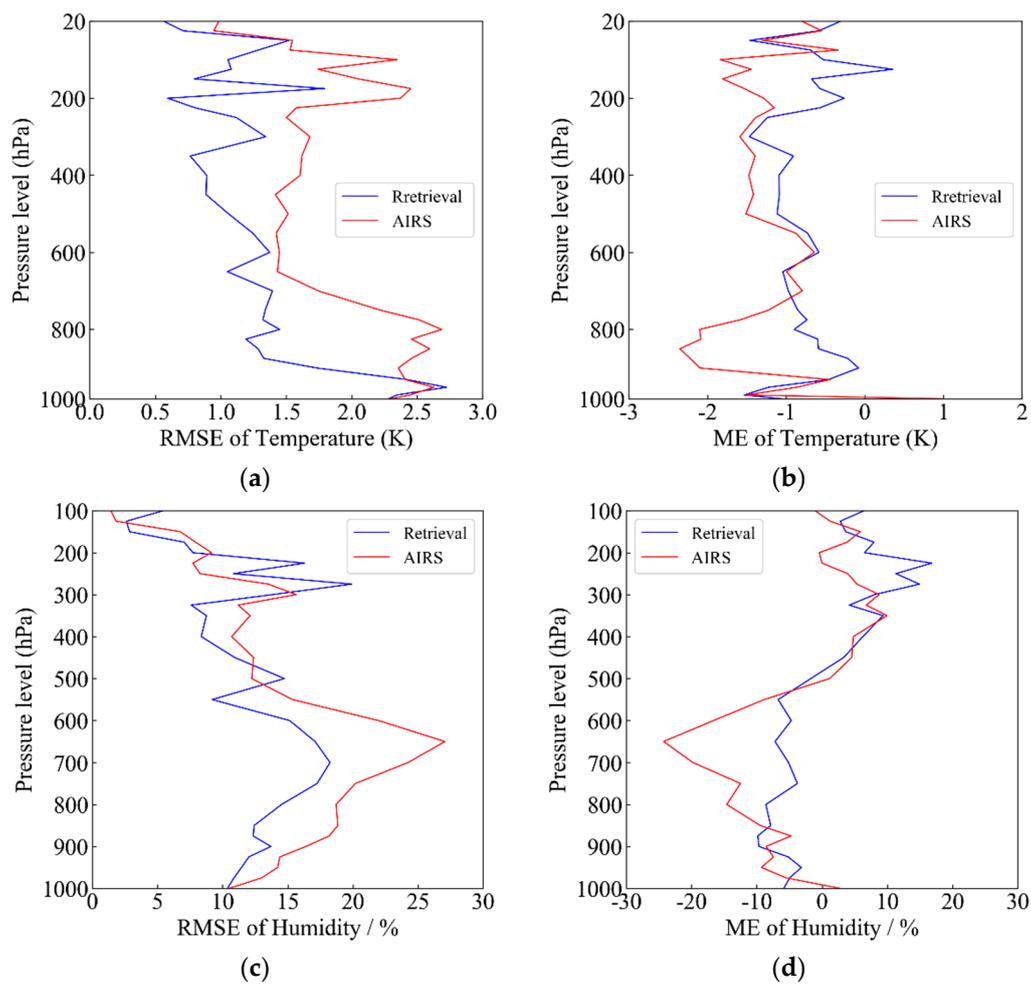


Figure 15. Comparison of the retrieval profile with the AIRS product. (a) Temperature RMSE, (b) temperature ME, (c) humidity RMSE, and (d) humidity ME.

As shown in the figure above, the average RMSE error of the AIRS temperature and humidity products is better than 2 K and 20%, respectively. The RMSE and ME error curves of the temperature and humidity profiles obtained by our retrieval system are similar to those of the AIRS temperature and humidity profile products. As shown in Figure 15a, the temperature RMSE obtained by our retrieval system shows a greater accuracy compared with the AIRS temperature product. The retrieval temperature values are closer to the sounding values as a whole, and it is a greater improvement near 200 hpa and 800 hpa. The maximum difference between the two can reach 1.5 K. The temperature ME retrieved by the system is closer to $x = 0$ than the AIRS product, and the value is stable between -1.5 K and 1 K.

The humidity RMSE error curve is shown in Figure 15c. The RMSE error curve of the humidity retrieved by our system is better than that of the AIRS products near 500–1000 hpa, and the accuracy is improved by up to 10%. It also has a good accuracy performance at 300–400 hpa, which shows that the humidity profile retrieved by our system is closer to the sounding true value. As shown in Figure 15d, the ME error curve of the AIRS products is farther away from $x = 0$ than the retrieval humidity, with a maximum deviation of 25% at about 650 hpa. It can be seen that the humidity profile retrieved by the FY-3D/HIRAS system has a higher accuracy than that of similar, international instruments.

4. Discussion

In this retrieval system construction experiment, we integrated the information entropy capacity-weight function coverage algorithm for channel selection and focused on

introducing the neural network error correction algorithm to correct the deviation of the observation data. We used sounding data as the standard truth value for evaluating the accuracy of the results. We showed through experiments that the average RMSE error of the temperature and humidity profiles obtained by our retrieval system was less than the GFS forecast profile, the retrieval result obtained by the conventional method, and the AIRS temperature and humidity products. Our results reached the international retrieval accuracy requirements of the temperature and humidity profiles.

However, we found that the retrieval RMSE value of humidity was higher than that of the GFS data in the humidity RMSE error graph (Figure 14c), around 800 hpa to 1000 hpa in the middle and lower troposphere, but the humidity RMSE of our overall retrieval was better than the accuracy requirement of 20%. There are two possible reasons for the unobvious improvement: The response peak of the weight function may have fewer channels in the middle and lower levels of the troposphere in our selected channels, which means that the accuracy of the bottom levels is lower than that of others in this experiment (see Figure 2b). Another possible reason is that the water vapor in the lower atmosphere changed drastically, the water vapor content undergoes great changes in a short time, and the time resolution of polar orbiting satellites is not high.

In addition, we believe that the error correction of the observation data must be considered in the retrieval process [19,49]. The neural network error correction proposed in this experiment has a higher accuracy than the conventional deviation correction method. From the perspective of the error correction effect, compared with the conventional method, the ANN has a maximum accuracy improvement of 3 K and 3.5 K in the temperature and humidity channels, respectively, (see Figure 9a,b). Based on the retrieval results, compared with the conventional method, the ANN error correction has an accuracy improvement of a maximum of 0.5 K and 5% in temperature and humidity retrieval, respectively (see Figure 14a–d). We believe that, in the process of the correction of observational data deviation, the neural network algorithm can find a law between the systemic deviation of the instrument itself and the error of the radiative transfer model. It can learn and master this part of the law, and subsequently perform error correction, so that the deviation between observation and simulation was reduced, and the retrieval can achieve good results.

In the end, there are still problems in the correction of neural network observation errors based on polar orbiting satellite data. Due to its lower time resolution than that of geostationary orbiting satellites, there are fewer training samples in the same transit time. At the same time, the sample diversity is poor, which means that the amount of information contained in the observations may be relatively small. There are also factors such as satellite angles and the surface temperature differences of day and night that need to be considered, but these factors are more difficult to calibrate, and the existence of these factors may affect the accuracy of the model. However, we have showed through this experiment that our method improves in accuracy in the retrieval of polar orbiting satellites, and the effect is better after this method is used. In future experiments, the number of samples can be increased through the joint observation of multiple satellites. The algorithm can be improved by the comprehensive consideration of various factors that affect the experimental results, and the accuracy of the model can be improved.

5. Conclusions

We built the FY-3D/HIRAS retrieval system in this study. We used the data of the European region from 15–31 July 2020, to train the relevant parameters of the system. We used test set data from 1–5 August 2020, to test the retrieval system. Channel selection was based on the information entropy capacity-weight function coverage algorithm, and 99 sets of temperature channels and 50 sets of humidity channels were finally selected. The peaks of the weight function of temperature and humidity cover most of the pressure levels, which proves that it is effective after channel optimization. At the same time, the retrieval system used a neural network algorithm to correct the observation error. The corrected temperature channel's RMSE and the humidity channel's RMSE of brightness temperature

were, respectively, less than 2 K and 1.5 K. Taking the sounding data as the true values, we compared the profile obtained by our retrieval system with the profiles obtained by the conventional method and found that the temperature and humidity profile's RMSE and ME obtained by our retrieval were overall better than the conventional method's results. We proposed that an ANN algorithm was used to correct the observation error, which effectively reduced the error between the radiative transfer model and the observation and improved the retrieval accuracy of our system. Comparing the temperature and humidity profiles retrieved by our system with the GFS forecast profiles, it was found that the temperature and humidity profiles retrieved by our system can meet the current retrieval accuracy requirements. It is closer to the sounding value than the GFS is, and most of the atmospheric pressure levels are better than that of the GFS, which proves that the retrieval results of the system have a high accuracy. In order to improve the stability and reliability of the retrieval system, we selected a similar, international instrument, AIRS L2 temperature and humidity products, for comparison and analysis. Our retrieval results have good accuracy at different pressure levels. The preliminary results show that our system can reach or exceed the product accuracy of similar, international instruments.

We highlighted through experiments that the two methods of retrieval system fusion have high credibility and stability. Since our retrieval system does not rely on reanalysis data that are difficult to obtain in real time, it can be applied to business operations and can act as a reference for the production of related products. It also provides a basis for the development of subsequent instruments and the improvement of related retrieval algorithms. Therefore, the FY-3D/HIRAS temperature and humidity profile retrieval system proposed in this study has broad application prospects. In the second-generation infrared hyperspectral instrument on board the later Morning ham orbit satellite, the time resolution and spectral resolution of the satellite data have been improved, and the training data that can be used for neural network error correction and the data that can be used for retrieval will increase. Therefore, the method proposed in this study will lead to greater accuracy improvements in the future.

Author Contributions: Conceptualization, C.Z., P.H., and M.G.; methodology, C.Z.; software, C.Z. and P.H.; validation, C.Z., P.H., T.Y., S.H., and C.Y.; formal analysis, C.Z. and P.H.; investigation, C.Z., P.H., and T.Y.; resources, C.Z.; data curation, M.G. and Y.H.; writing—original draft preparation, C.Z. and P.H.; writing—review and editing, C.Z., P.H., and Y.H.; visualization, M.G.; supervision, M.G. and Y.H.; project administration, M.G.; funding acquisition, C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Key R&D Program of China (2018YFB0504800 (2018YFB0504802)).

Data Availability Statement: All data generated or analyzed during this study are included in this article.

Acknowledgments: We would like to thank the NSMC (National Satellite Meteorological Center) for sharing FY3D/HIRAS data. Meanwhile, we would also like to thank the University of Wyoming, NCEP, and ECMWF for their openly accessible data and software. We would also like to thank Zhaojun Zheng and Na Xu from NSMC, CMA (China Meteorological Administration), for giving us advice on how to use the cloud-mask products of MERIS-II. Finally, we are very grateful to Suzhou Academy, Shanghai Institute of Technical Physics, Chinese Academy of Sciences, for our equipment resources.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Collard, A.D.; McNally, A.P. The assimilation of infrared atmospheric sounding interferometer radiances at ECMWF. *Q. J. R. Meteorol. Soc.* **2009**, *135*, 1044–1058. [[CrossRef](#)]
2. Guidard, V.; Fourrié, N.; Brousseau, P.; Rabier, F. Impact of IASI assimilation at global and convective scales and challenges for the assimilation of cloudy scenes. *Q. J. R. Meteorol. Soc.* **2011**, *137*, 1975–1987. [[CrossRef](#)]

3. Joo, S.; Eyre, J.; Marriott, R. The impact of Metop and other satellite data within the Met Office global NWP system using an adjoint-based sensitivity method. *Mon. Weather Rev.* **2013**, *141*, 3331–3342. [[CrossRef](#)]
4. Chen, Y.; Han, Y.; van Delst, P.; Weng, F. Assessment of Shortwave Infrared Sea Surface Reflection and Nonlocal Thermodynamic Equilibrium Effects in the Community Radiative Transfer Model Using IASI Data. *J. Atmos. Ocean. Technol.* **2013**, *30*, 2152–2160. [[CrossRef](#)]
5. Sun, B.; Reale, A.; Tilley, F.H.; Petty, M.E.; Nalli, N.R.; Barnet, C.D. Assessment of NUCAPS S-NPP CrIS/ATMS Sounding Products Using Reference and Conventional Radiosonde Observations. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 2499–2509. [[CrossRef](#)]
6. Gambacorta, A.; Barnet, C.D. Methodology and information content of the NOAA NESDIS operational channel selection for the Cross-Track Infrared Sounder (CrIS). *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 3207–3216. [[CrossRef](#)]
7. Qi, C.L.; Gu, M.J.; Hu, X.Q.; Wu, C.Q. FY-3 Satellite infrared high spectral sounding technique and potential application. *Adv. Met. ST* **2016**, *6*, 88–93.
8. Li, J.; Fang, Z.Y. The development of satellite meteorology-challenges and opportunities. *Meteorol. Mon.* **2012**, *38*, 129–146.
9. Luo, S.; Di, D.; Cui, L.L. Study on FY-4A/GIIRS infrared spectrum detection capability based on information content. *J. Infrared Millim. Waves* **2019**, *38*, 765–776.
10. Ren, J. Study on the Atmospheric Temperature and Humidity Profiles of Satellite Remote Sensing Based on One-Dimensional Variational Algorithm. Master's Thesis, Nanjing University of Information Science and Technology, Nanjing, China, 2018.
11. Zhou, A.M. Atmospheric Temperature and Humidity Profiles Retrieval from Hyperspectral Infrared Simulation Data Based on FY-4. Master's Thesis, Nanjing University of Information Science and Technology, Nanjing, China, 2017.
12. Cai, X.; Bao, Y.S.; Petropoulos, G.P.; Lu, F.; Zhu, L.H.; Wu, Y. Temperature and Humidity Profile Retrieval from FY4-GIIRS Hyperspectral Data Using Artificial Neural Networks. *J. Remote Sens.* **2020**, *12*, 1872. [[CrossRef](#)]
13. Huang, P.Y.; Guo, Q.; Han, C.P.; Zhang, C.M.; Yang, T.H.; Huang, S. An Improved Method Combining ANN and 1D-Var for the Retrieval of Atmospheric Temperature Profiles from FY-4A/GIIRS Hyperspectral Data. *J. Remote Sens.* **2021**, *13*, 481. [[CrossRef](#)]
14. Zhu, L.H.; Bao, Y.S.; Petropoulos, G.P.; Zhang, P.; Lu, F.; Lu, Q.F.; Wu, Y.; Xu, D. Temperature and Humidity Profiles Retrieval in a Plain Area from Fengyun-3D/HIRAS Sensor Using a 1D-VAR Assimilation Scheme. *J. Remote Sens.* **2020**, *12*, 435. [[CrossRef](#)]
15. Guan, Y.H.; Ren, J.; Bao, Y.S.; Lu, Q.F.; Liu, H.; Xiao, X.J. Research of the infrared high spectral (IASI) satellite remote sensing atmospheric temperature and humidity profiles based on the one-dimensional variational algorithm. *Trans. Atmos. Sci.* **2019**, *42*, 602–611.
16. Ma, P.F.; Chen, L.F.; Tao, J.H.; Su, L.; Tao, M.H.; Wang, Z.F.; Zou, M.Y.; Zhang, Y. Study on Simulation of infrared hyperspectral CrIS data retrieval of atmospheric temperature and humidity profiles. *Spectrosc. Spectr. Anal.* **2014**, *34*, 1894–1897.
17. Yang, T.H.; Hu, X.Q.; Xu, H.L.; Wu, C.Q.; Qi, C.L.; Gu, M.J. Radiation Calibration Accuracy Assessment of FY-3D Hyperspectral Infrared Atmospheric Sounder Based on Inter-Comparison. *Acta Opt. Sin.* **2019**, *39*, 377–387.
18. Qi, C.L.; Zhou, F.; Wu, C.Q.; Hu, X.Q.; Gu, M.J. Spectral Calibration of Fengyun-3 high-spectral resolution infrared sounder. *Opt. Precis. Eng.* **2019**, *27*, 747–755.
19. Di, D. Data Assimilation Research for Geosynchronous Interferometric Infrared Sounder onboard FengYun-4 Satellite. Ph.D. Thesis, University of Chinese Academy of Sciences, Beijing, China, 2019.
20. Saunders, R.; Hocking, J.; Turner, E.; Rayer, P.; Rundle, D.; Brunel, P.; Vidot, J.; Roquet, P.; Matricardi, M.; Geer, A.; et al. An update on the RTTOV fast radiative transfer model (currently at version 12). *Geosci. Model Dev.* **2018**, *11*, 2717–2737. [[CrossRef](#)]
21. Di, D.; Li, J.; Han, W.; Bai, W.G.; Wu, C.Q.; Menzel, W.P. Enhancing the fast radiative transfer model for FengYun-4 GIIRS by using local training profiles. *J. Geophys. Res. Atmos.* **2018**, *123*, 12583–12596. [[CrossRef](#)]
22. Rodgers, C.D. Information content and optimization of high spectral resolution remote measurements. *Adv. Space Res.* **1998**, *21*, 361–367. [[CrossRef](#)]
23. Yang, Y.H.; Yin, Q.; Shu, J. Channel selection of atmosphere vertical sounder (GIIRS) onboard the FY-4A geostationary satellite. *J. Infrared Millim. Waves* **2018**, *37*, 545–552.
24. Yu, Y. Research on Key Technologies of Data Assimilation of High-Resolution Satellite Observations. Ph.D. Thesis, National University of Defense Technology, Changsha, China, 2017.
25. Dee, D.P. Bias and data assimilation. In Proceedings of the 4th WMO International Symposium on Assimilation of Observations in Meteorology and Oceanography, Prague, Czech Republic, 18–22 April 2005.
26. Kanamitsu, M.; Saha, S. Systematic tendency error in budget calculations. *Mon. Weather Rev.* **1996**, *124*, 1145–1160. [[CrossRef](#)]
27. Yang, Y.; Shen, F.; Yang, Z.C.; Feng, X.S. Prediction of Solar Wind Speed at 1 AU Using an Artificial Neural Network. *Space Weather* **2018**, *16*, 1227–1244. [[CrossRef](#)]
28. Zhang, K.; Wu, C.Q.; Li, J. Retrieval of Atmospheric Temperature and Moisture Vertical Profiles from Satellite Advanced Infrared Sounder Radiances with a New Regularization Parameter Selecting Method. *J. Meteor. Res.* **2016**, *30*, 356–370. [[CrossRef](#)]
29. Chahine, M.T. Inverse Problems in Radiative Transfer: Determination of Atmospheric Parameters. *J. Atmos. Sci.* **1970**, *27*, 960–967. [[CrossRef](#)]
30. Fleming, H.E.; Goldberg, M.D.; Crosby, D.S. Minimum variance simultaneous retrieval of temperature and water vapor from satellite measurements. In Proceedings of the Second Conference on Satellite Meteorology/Remote Sensing and Applications, Williamsburg, VA, USA, 13–16 May 1986.

31. Bai, W.G.; Zhang, P.; Zhang, W.J.; Li, J. An efficient method for hyper-spectral infrared atmospheric radiation transfer calculation. *J. Infrared Millim. Waves* **2016**, *35*, 99–108.
32. Dong, C.H.; Li, J.; Zhang, P. *The Principle and Application of Satellite Hyperspectral Infrared Atmospheric Remote Sensing*, 1st ed.; Science Press: Beijing, China, 2013.
33. Boukabara, S.A.; Garrett, K.; Chen, W.C.; Iturbide-Sanchez, F.; Grassotti, C.; Kongoli, C.; Chen, R.Y.; Liu, Q.H.; Yan, B.H.; Weng, F.Z.; et al. MiRS: An All-Weather 1DVAR Satellite Data Assimilation and Retrieval System. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 3249–3272. [[CrossRef](#)]
34. Yang, J.; Min, Q. Retrieval of Atmospheric Profiles in the New York State Mesonet Using One-Dimensional Variational Algorithm. *J. Geophys. Res. Atmos.* **2018**, *123*, 7563–7575. [[CrossRef](#)]
35. Che, Y.; Ma, S.; Xing, F.; Li, S.; Dai, Y. An improvement of the retrieval of temperature and relative humidity profiles from a combination of active and passive remote sensing. *Meteorol. Atmos. Phys.* **2019**, *131*, 681–695. [[CrossRef](#)]
36. Jang, H.S.; Sohn, B.J.; Chun, H.W.; Li, J.; Weisz, E. Improved AIRS temperature and moisture soundings with local a priori information for the 1DVAR method. *J. Atmos. Ocean. Technol.* **2017**, *34*, 1083–1095. [[CrossRef](#)]
37. Duncan, D.I.; Kummerow, C.D. A 1DVAR retrieval applied to GMI: Algorithm description, validation, and sensitivities. *J. Geophys. Res. Atmos.* **2016**, *121*, 7415–7429. [[CrossRef](#)]
38. Ma, X.L.; Schmit, T.J.; Smith, W.L. A nonlinear physical retrieval algorithm-Its application to the GOES-8/9 sounder. *J. Appl. Meteorol.* **1999**, *38*, 501–513. [[CrossRef](#)]
39. Martinet, P.; Dabas, A.; Donier, J.M.; Douffet, T.; Guillot, O.G.R. 1D-Var temperature retrievals from microwave radiometer and convective scale model. *Tellus A Dyn. Meteorol. Oceanogr.* **2015**, *67*, 27925. [[CrossRef](#)]
40. Yang, T.H. Tropospheric Wind Field Measurement Based on Infrared Hyperspectral Observations. Ph.D. Thesis, University of Chinese Academy of Sciences, Beijing, China, 2020.
41. Carmi, F.; Xiao, X.J.; Lu, Q.F.; Atkinson, N.; Hocking, J. Assessment of the Hyperspectral Infrared Atmospheric Sounder (HIRAS). *J. Remote Sens.* **2019**, *11*, 2950. [[CrossRef](#)]
42. Wu, C.Q.; Qi, C.L.; Hu, X.Q.; Gu, M.J.; Yang, T.H.; Xu, H.L.; Li, L.; Yang, Z.D.; Zhang, P. FY-3D HIRAS radiometric calibration and accuracy assessment. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 3965–3976. [[CrossRef](#)]
43. Xu, N.; Chen, L.; Hu, X.Q.; Zhang, L.Y.; Zhang, P. Assessment and Correction of on-Orbit Radiometric Calibration for FY-3 VIRR Thermal Infrared Channels. *J. Remote Sens.* **2014**, *6*, 2884–2897. [[CrossRef](#)]
44. Yan, L.; Hu, Y.H.; Zhang, Y.; Li, X.M.; Dou, C.Y.; Li, J.; Si, Y.D. Radiometric Calibration Evaluation for FY3D MERSI-II Thermal Infrared Channels at Lake Qinghai. *J. Remote Sens.* **2021**, *13*, 466. [[CrossRef](#)]
45. Chen, S.L.; Zheng, X.B.; Li, X.; Wei, W.; Du, S.D.; Guo, F.X. Vicarious Radiometric Calibration of Ocean Color Bands for FY-3D/MERSI-II at Lake Qinghai, China. *Sensors* **2021**, *21*, 139. [[CrossRef](#)]
46. Qi, C.L.; Wu, C.Q.; Hu, X.Q.; Xu, H.L.; Li, L.; Zhou, F.; Gu, M.J.; Yang, T.H.; Shao, C.Y.; Yang, Z.D.; et al. High Spectral Infrared Atmospheric Sounder (HIRAS): System Overview and On-Orbit Performance Assessment. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 4335–4351. [[CrossRef](#)]
47. Hewison, T.J. 1D-VAR retrieval of temperature and humidity profiles from a ground-based microwave radiometer. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 2163–2168. [[CrossRef](#)]
48. Ishimoto, H.; Okamoto, K.; Okamoto, H.; Sato, K. One-dimensional variational (1D-Var) retrieval of middle to upper tropospheric humidity using AIRS radiance data. *J. Geophys. Res. Atmos.* **2014**, *119*, 7633–7654. [[CrossRef](#)]
49. Susskind, J.; Rosenfield, J.; Reuter, D. An accurate radiative transfer model for use in the direct physical inversion of HIRS-2 and MSU temperature sounding data. *J. Geophys. Res.* **1983**, *88*, 8550–8568. [[CrossRef](#)]