（a） The point-by-point response to the reviews

Dear Referee #2,

We thank you for your review of our manuscript and your detailed remarks. We would like to improve the article here in an immediate reply. Please, find our answers/comments on your notes below:

(1) p. 5, l. 108: ...from the thousands of channels' observations... please rephrase, but I have no suggestions

Yes, you are right.

“It is important to properly … from the thousands of channels’ observations ….” has been modified to “In order to improve the calculation efficiency and retrieval quality, it is very important to properly select a set of channels that can provide as much information as possible.” (L 104-L 107)

Thanks.

(2) p. 8, l. 155: Channel selection mostly uses the information content and delievers the largest amount of information for the selected channel combination during the retreival.

Yes, you are right.

“Currently, information content is often employed in channel selection. During retrieval, this method delivers the largest amount of information for the selected channel combination (Rodgers, 1996; Du et al., 2008; He et al., 2012; Richardson et al., 2018).” has been modified to “Channel selection mostly uses the information content and delivers the largest amount of information for the selected channel combination during the retrieval (Rodgers, 1996; Du et al., 2008; He et al., 2012; Richardson et al., 2018)." (L 153-L 156)

Thanks.

(3) p. 9, l. 194: after the retrieval

Yes, you are right.

"…after retrieval" has been modified to "…after the retrieval" (L 193)

Thanks.

(4) p. 11, l. 236: (2) The number looks like an equation number, so it is a bit confusing.

Sorry for the confusion.

The step number has been modified to (I), (II), (III) and (IV). (L 222; L 234; L 255; L 270; L 272)

Thanks.

(5) p. 18, l. 389: The footprint size is 13.5 km

Yes, you are right.

This has been modified. (L 288)

Thanks.

(6) p. 19, l. 401/402: in Fig. 1. The measurement error is not below 0.2K for all the instrument channels. There are...

Yes, you are right.

"The root mean square error of an AIRS infrared channel is shown in Fig. 1, with black spots, indicating that not all the instrument channels possess a measurement error of less than 0.2 K." has been modified to "The root mean square error of an AIRS infrared channel is shown in Fig. 1. The measurement error is not below 0.2K for all the instrument channels." (L 400-L 401)

Thanks.

(7) p. 21, l. 434: ranking -> in the range

Yes, you are right.

This has been modified. (L 432)

Thanks.

(8) p. 27, l. 532: temperature, a better observation can be obtained for higher temperatures

Yes, you are right.

"…temperature, the higher temperature is, the better observation can be obtained;" has been modified to "…temperature, a better observation can be obtained for higher temperatures" (L 524-L 525)

Thanks.

(9) p. 28, l. 547: to the -> at

Yes, you are right.

This has been modified. (L 538)

Thanks.

(10) p. 30, l. 567: reach to -> reaches

Yes, you are right.

This has been modified. (L 560)

Thanks.

(11) p. 36, l. 650: humidity and 10m wind speed (remove dot after speed)

Sorry for my carelessness.

This has been modified. (L 643)

Thanks.

(12) p. 43, l. 778: to the -> at

Yes, you are right.

This has been modified. (L 771; L 828)

Thanks again for your careful review.

Dear Referee #4,

We thank you for your review of our manuscript and your detailed remarks. We would like to improve the article here in an immediate reply. Please, find our answers/comments on your notes below:

(1) p1, l16-20: Suggest to replace the first two sentences of the abstract by: "This study introduces an effective channel selection method for hyperspectral infrared sounders. The method is illustrated for the Atmospheric InfraRed Sounder (AIRS) instrument."

Yes, we agree with you.

This has been modified. (L 16-L 18)

Thanks.

(2) p1, l21: "improved method" -> "improved channel selection (ICS) method"

Yes, you are right.

This has been modified. (L 19-L 21)

Thanks.

(3) p2, l24: suggest to delete "and closer to the actual atmosphere"

Yes, you are right.

This has been deleted. (L 22)

Thanks.

(4) p2, l30: suggest to replace "In general, ..." by "Also at lower heights, ..."

Yes, you are right.

This has been modified. (L 28)

Thanks.

(5) p2, l31-32: delete "(Improved Channel Selection)"

Yes, you are right.

This has been deleted. (L 29)

Thanks.

(6) p2, l32-36: suggest to rephrase this as "Statistical inversion comparison experiments for four different regions illustrate latitudinal

and seasonal variations and better performance of ICS compared to the NWP Channel Selection (NCS) and Primary Channel Selection (PCS) methods. The ICS method shows potential for future applications."

Yes, we agree with you.

“Statistical inversion comparison experiments in four typical regions indicate that ICS in this paper is significantly better than NCS (NWP Channel Selection) and PCS (Primary Channel Selection) in different regions and shows latitudinal variations, which shows potential for future applications.” has been modified to “Statistical inversion comparison experiments for four different regions illustrate latitudinal and seasonal variations and better performance of ICS compared to the NWP Channel Selection (NCS) and Primary Channel Selection (PCS) methods. The ICS method shows potential for future applications.” (L 30-L 34)

Thanks.

(7) p4, l70: please fix: "Infrared Atmospheric Sounding Interferometer"

Yes, you are right.

This has been modified. (L 68-L 69)

Thanks.

(8) p5, l110: replace "the channel selection algorithm" by "channel selection algorithms"

Yes, you are right.

This has been modified. (L 108)

Thanks.

(9) p11, l236: replace "p matrix" by "for calculating the p value"?

Yes, you are right.

This has been modified. (L 234)

Thanks.

(10) p11, l237: replace "matrixes, it" by "matrices, they"

Yes, you are right.

This has been modified. (L 235)

Thanks.

(11) p12, l257: replace "According to" by "Using"

Yes, you are right.

This has been modified. (L 257)

Thanks.

(12) p18, l383: The reference to Hoffmann and Alexander (200) can be deleted here, I think. Maybe move reference to Susskind et al. (2003) from line 389 to this place.

Yes, we agree with you.

"…(Aumann et al., 2003; Hoffmann and Alexander, 2009)." has been modified to "…(Aumann et al., 2003; Susskind et al., 2003).". (L 382-L383)

"Susskind et al. (2003)" has been deleted. (L 388)

Thanks.

(13) p27, l515-534: Suggest to delete all the sentences relating the wavelengths to the wavenumbers, as this is trivial. For instance, replace "When the wavenumber approaches 1000... Near this band..." by "Near the 10 um band..."

Yes, you are right.

"(1) When the wavenumber approaches 1000, the wavelength is 10 μm (1/1000 cm-1). Near this band, fewer channels are selected by PCS because the retrieval of ground temperature is considered by NCS; (2) When the wavenumber is near 1200, the wavelength is 9 μm (1/1200 cm-1). Near this band, no channels are selected by PCS because the retrieval of O3 is not considered in this paper; (3) When the wavenumber approaches 1500, the wavelength is 6.7 μm (1/1500 cm-1). As is known, the spectral range from 6 μm to 7 μm corresponds to water vapor absorption bands, but fewer channels are selected by NCS; (4) When the wavenumber is close to 2000, it derives a wavelength of 5 μm (1/2000 cm-1), which includes 4.2 μm for N2O and 4.3 μm for CO2 absorption bands." has been modified to "(1) Near 10 μm band, fewer channels are selected by PCS because the retrieval of ground temperature is considered by NCS; (2) Near 9 μm band, no channels are selected by PCS because the retrieval of O3 is not considered in this paper; (3) As is known, the spectral range from 6 μm to 7 μm corresponds to water vapor absorption bands, but fewer channels are selected by NCS; (4) Near 5 μm band, it includes 4.2 µm for N_2O and 4.3 µm for CO_2 absorption bands." (L 514-L 521)

"(5) In the near infrared area, the wavenumber exceeds 2200, deriving a wavelength of less than 4 µm (1/2000 cm-1). A small number of channels is selected by NCS, but no channels are selected by PCS." has been modified to "(5) Near 4 µm band, a small number of channels is selected by NCS, but no channels are selected by PCS." (L 526-L 527)

Thanks.

(14) p27, l529-533: It is still not clear what is meant by "high temperature zone", I think. Does this refer to temperatures at high altitudes?

Sorry for the confusion.

"high temperature zone" is not clear for readers which we want to say "under the higher temperature conditions" But it can be deleted in this sentence, because we have explained it in the following sentences.

"PCS is favorable for atmospheric temperature observation in the high temperature zone. Because 4.2 µm and 4.3 µm bands are sensitive to high temperature, a better observation can be obtained for higher temperatures;" has been modified to "PCS is favorable for atmospheric temperature observation. Because 4.2 µm and 4.3 µm bands are sensitive to high temperature, a better observation can be obtained for higher temperatures;" (L 522-L 525)

"Due to the method selected in this paper, there are more channels at 4.2

μm for N2O and 4.3 μm for CO2 absorption bands, and the channel combination of PCS is superior to that of NCS for atmospheric temperature observation in the high temperature zone.” has been modified to “Due to the method selected in this paper, there are more channels at 4.2 μm for N2O and 4.3 μm for CO2 absorption bands, and the channel combination of PCS is superior to that of NCS for atmospheric temperature observation.” (L 689-L 693)

Thanks.

(15) p28, l537: replaced "used in this paper" by "considered in this study"

Yes, you are right.

This has been modified. (L 528)

Thanks.

(16) p29, Fig. 5: I would like to suggest to combine the curves of plot a) and b) into a single plot.

Yes, you are right.

Fig. 5 has been modified to “The relationship between the number of iterations and ARI. Blue line represents the result of ICS. Red dotted line stands for the result of PCS.” (L 547: Fig. 5)

Thanks again for your careful review.

(b) The list of all relevant changes made in the manuscript

A channel selection method for hyperspectral atmospheric infrared sounders based on layering

Shujie Chang[1, 2,3], Zheng Sheng[1,2], Huadong Du[1,2], Wei Ge[1,2] and Wei Zhang[1,2]

[1] College of Meteorology and Oceanography, National University of Defense Technology, Nanjing, China

[2] Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science and Technology, Nanjing, China

[3] South China Sea Institute for Marine Meteorology, Guangdong Ocean University, Zhanjiang, China

Correspondence: Zheng Sheng (19994035@sina.com)

Abstract. ~~Because a satellite channel's ability to resolve hyperspectral data varies with height,~~This study introduces an ~~improved channel selection method is proposed based on information content. An~~ effective channel selection ~~scheme~~ method for ~~a~~ hyperspectral ~~atmospheric~~ infrared ~~sounder using~~ sounders. The method is illustrated for the Atmospheric InfraRed Sounder

(AIRS~~ data based on layering is proposed.~~) instrument. The results are as follows: (1) Using the improved ~~method,~~channel selection (ICS), the atmospheric retrievable index is more stable, the value reaching 0.54. The coverage of the weighting functions is more evenly distributed over height with this method~~ and closer to the actual atmosphere~~; (2) Statistical inversion comparison experiments show that the accuracy of the retrieval temperature, using the improved channel selection method in this paper, is consistent with that of 1Dvar channel selection. In the stratosphere and mesosphere especially, from 10 hPa to 0.02 hPa, the accuracy of the retrieval temperature of our improved channel selection method is improved by about 1 K. ~~In general~~Also at lower heights, the accuracy of the retrieval temperature of ICS ~~(Improved Channel Selection)~~ is improved; (3) Statistical inversion comparison experiments ~~in~~for four ~~typical regions indicate that ICS in this paper is significantly~~ different regions illustrate latitudinal and seasonal variations and better ~~than NCS (~~performance of ICS compared to the NWP Channel Selection (NCS) and ~~PCS (~~Primary Channel Selection~~) in different regions and shows latitudinal variations, which~~ (PCS) methods. The ICS method shows potential for future applications.

1 Introduction

Since the successful launch of the first meteorological satellite, TIROS in the 1960s, satellite observation technology has developed rapidly. Meteorological satellites observe the Earth’s atmosphere from space and are able to record data from regions which are otherwise difficult to observe. Satellite data greatly enrich the content and range of meteorological observations, and consequently, atmospheric exploration technology and meteorological observations have taken us to a new stage in our understanding of weather systems and related phenomena (Fang, 2014). From the perspective of vertical atmospheric observation, satellite instruments are developing rapidly. In their infancy, the traditional infrared measurement instruments for detecting atmospheric temperature and moisture profiles, such as TOVS (Smith et al., 1991) or HIRS in ATOVS (Chahine, 1972; Li et al., 2000; Liu, 2007), usually employed filter spectrometry. Even though such instruments have played an important role in improving weather prediction, it is difficult to continue to build upon improvements in terms of observation accuracy and vertical resolution due to the limitation of low spectral resolution. By using this kind of filter-based spectroscopic measurement instrument, therefore, it is difficult to meet today’s needs in numerical weather prediction (Eyre et al., 1993; Prunet et al., 2010; Menzel et al., 2018). To meet this challenge, a series of plans for the creation of high-spectral resolution atmospheric measurement instruments has been executed in the United States and in Europe in recent years: One example is the AIRS (Atmospheric InfraRed Sounder) on the Earth Observation System, “Aqua”, launched on May 4, 2002 from the United States. AIRS has 2378 spectral channels providing sensitivity from the ground to up to about 65 km of altitude (Aumann et al., 2003; Hoffmann and Alexander, 2009; Gong et al., 2011). The United States and Europe, in 2010 and in 2007, also installed the CRIS (Cross-track Infrared Sounder) and the IASI (~~Inter-Attractive~~Infrared Atmospheric Sounding Interferometer) on polar-orbiting satellites.

China also devotes great importance to the development of such advanced sounding technologies. In the early 1990s, the National Satellite Meteorological Center began to investigate the principles and techniques of hyperspectral resolution atmospheric observations. China’s development of interferometric atmospheric vertical detectors eventually led to the launch of Fengyun No. 3, on May 27, 2008, and Fengyun No. 4 on December 11, 2016, both of which were equipped with infrared atmospheric instruments. How best to use the hyperspectral resolution observation data obtained from these instruments, to obtain reliable atmospheric temperature and humidity profiles, is an active area of study in atmospheric inversion theory.

Due to technical limitations, only a limited number of channels could at first be built into the typical satellite instruments. In this case, channel selection generally involved controlling the channel weighting function by utilizing the spectral response characteristics of the channel (such as center frequency and bandwidth). With the development of measurement technology, increasing numbers of hyperspectral detectors were carried on meteorological satellites. Due to the large number of channels and data supported by such instruments today (such as AIRS with 2378 channels and IASI with 8461 channels), it has proven extremely cumbersome to store, transmit, and process such data. Moreover, there is often a close correlation between the channel, causing an ill-posedness of the inversion, potentially compromising accuracy of the retrieval product based on hyperspectral resolution data.

However, hyperspectral detectors have many channels and provide real-time mode prediction systems with vast quantities of data, which can significantly improve prediction accuracy. But, if all the channels are used to retrieve data, the retrieval time considerably increases. Even more problematic are the glut of information produced, and the unsuitability of the calculations for real-time forecasting. Concurrently, the computer processing power must be large enough to meet the demands of simulating all the channels simultaneously within the forecast time. ~~It~~In order to improve the calculation efficiency and retrieval quality, it is very important to properly select a ~~group~~set of channels that can provide as much information as possible ~~from the thousands of channels' observations to improve the calculation efficiency and retrieval quality~~.

Many researchers have studied ~~the~~ channel selection ~~algorithm.~~algorithms. Menke (1984) first chose channels using a data precision matrix method. Aires et al. (1999) made the selection using the Jacobian matrix, which has been widely used since then (Aires et al., 2002; Rabier et al., 2010). Rodgers (2000) indicated that there are two useful quantities in measuring the information provided by the observation data: Shannon information content and degrees of freedom. The concept of information capacity then became widely used in satellite channel selection. In 2007, Xu (2007) compared the Shannon information content with the relative entropy, analyzing the information loss and information redundancy. In 2008, Du et al. (2008) introduced the concept of the atmospheric retrievable index (ARI) as a criterion for channel selection, and in 2010, Wakita et al. (2010) produced a scheme for calculating the information content of the various atmospheric parameters in remote sensing using

Bayesian estimation theory. Kuai et al. (2010) analyzed both the Shannon information content and degrees of freedom in channel selection when retrieving CO_2 concentrations using thermal infrared remote sensing and indicated that 40 channels could contain 75% of the information from the total channels. Cyril et al. (2003) proposed the optimal sensitivity profile method based on the sensitivity of different atmospheric components. Lupu et al. (2012) used degrees of freedom for signals (DFS) to estimate the amount of information contained in observations in the context of observing system experiments. In addition, the singular value decomposition method has also been widely used for channel selection (Prunet et al., 2010; Zhang et al., 2011; Wang et al., 2014). In 2017, Chang et al. (2017) selected a new set of Infrared Atmospheric Sounding Interferometer (IASI) channels using the channel score index (CSI). Richardson et al. (2018) selected 75 from 853 channels based on the high spectral-resolution oxygen A-band instrument on NASA's Orbiting Carbon Observatory-2 (OCO-2), using information content analysis to retrieve the cloud optical depth, cloud properties, and position.

Today's main methods for channel selection use only the weighting function to study appropriate numerical methods, such as the data precision matrix method (Menke, 1984), singular value decomposition method (Prunet et al., 2010; Zhang et al., 2011; Wang et al., 2014), and the Jacobi method (Aires et al., 1999; Rabier et al., 2010). The use of the methods allows sensitive channels to be selected. The above-mentioned studies also take into account the sensitivity of each channel to atmospheric parameters during channel selection, while ignoring some factors that impact retrieval results. The accuracy of retrieval results depends not only on the channel weighting function but also on the channel noise, background field, and the retrieval algorithm.

~~Currently,~~ Channel selection mostly uses the information content ~~is often employed in channel selection. During retrieval, this method~~and delivers the largest amount of information for the selected channel combination during the retrieval (Rodgers, 1996; Du et al., 2008; He et al., 2012; Richardson et al., 2018).

This method has made great breakthroughs in both theory and practice, and the concept of information content itself does consider all the height dependencies of the kernel matrix K (Rodgers, 2000). However, earlier works have neglected the height dependencies of K for simplicity. This paper uses the atmospheric retrievable index (ARI) as the index, which is based on information content (Du et al., 2008; Richardson et al. 2018). Channel selection is made at different heights, and an effective channel selection scheme is proposed which fully considers various factors, including the influence of different channels on the retrieval results at different heights. This ensures the best accuracy of the retrieval product when using the selected channel. In addition, statistical inversion comparison experiments are used to verify the effectiveness of the method.

2 Channel selection indicator, scheme and method

2.1 Channel selection indicator

According to the concept of information content, the information content contained in a selected channel of a hyperspectral instrument can be described as H (Rodgers, 1996; Rabier et al., 2010). The final expression of H is:

$$\mathrm{H} = -\frac{1}{2} ln\left|\hat{S} S_a^{-1}\right|$$

$$= -\frac{1}{2} ln|(S_a - S_a K^T (K S_a K^T + S_\varepsilon)^{-1} K S_a) S_a^{-1}|, \qquad (1)$$

where S_a is the error covariance matrix of the background or the estimated value of atmospheric profile, S_ε represents the observation error covariance matrix of each hyperspectral detector channel, $\hat{S} = (S_a - S_a K^T (K S_a K^T + S_\varepsilon)^{-1} K S_a)$ denotes the covariance matrix after retrieval, K is the weighting function matrix.

In order to describe the accuracy of the retrieval results visually and quantitatively, the atmospheric retrievable index (ARI), p, (Du et al., 2008) is defined as follows:

$$p = 1 - \exp\left(\frac{1}{2n} ln\left|\hat{S}S_a^{-1}\right|\right), \tag{2}$$

Assuming that before and after the retrieval, the ratio of the root mean square error of each element in the atmospheric state vector is 1-p, then $\left|\hat{S}S_a^{-1}\right| = (1 - p)^{2n}$ is derived. By inverting the equation, the ARI that is p can be obtained in Eq. (2), which indicates the relative portion of the error that is eliminated by retrieval. In fact, before and after retrieval, the ratio of the root mean square error of each element cannot be 1-p. Therefore, p defined by Eq. (1) is actually an overall evaluation of the retrieval result.

2.2 Channel selection scheme

The principle of channel selection is to find the optimum channel combination after numbering the channels. This combination makes the information content, H, or the ARI defined in this paper as large as possible, in order to maintain the highest possible accuracy in the retrieval results.

Let there be M layers in the vertical direction of the atmosphere and N satellite channels. Selecting n from N channels, there will be

C_N^n combinations in each layer, leading C_N^n calculations to get C_N^n kinds of p results. Furthermore, there are M layers in the vertical direction of the atmosphere. Therefore, the entire atmosphere must be calculated $M \cdot C_N^n$ times. However, the calculation $M \cdot C_N^n$ times will be particularly large, which makes this approach impractical in calculating p for all possible combinations. Therefore, it is necessary to design an effective calculation scheme, and such a scheme, i.e., a channel selection method, using iteration is proposed, called the "sequential absorption method" (Dudhia et al., 2002; Du et al., 2008). The method's main function is to select ("absorb") channels one by one, taking the channel with the maximum value of p. Through n iterations, n channels can be selected as the final channel combination. The steps are as follows:

(~~1~~I) The expression of information content in a single channel:

First, we use only one channel for retrieval. A row vector, k, in the weighting function matrix, K, is a weighting function corresponding to the channel. After observation in this channel, the error covariance matrix is:

$$\hat{S} = S_a - S_a k^T (s_\varepsilon + k S_a k^T)^{-1} k S_a. \tag{3}$$

It should be noted that $(s_\varepsilon + k S_a k^T)$ is a scalar value in Eq. (3), so Eq. (3) can be converted to:

$$\hat{S} = \left(I - \frac{S_a k^T k}{(s_\varepsilon + k S_a k^T)}\right) S_a = \left(I - \frac{(k S_a)^T k}{(s_\varepsilon + k (k S_a)^T)}\right) S_a. \tag{4}$$

Substituting Eq. (4) into Eq. (2) gives:

$$\mathrm{p} = 1 - \exp\left(\frac{1}{2n} ln\left(\left|I - \frac{(kS_a)^T k}{(s_\varepsilon + k(kS_a)^T)}\right|\right)\right). \qquad (5)$$

(~~2~~II) Simplification of Eq. (5) for calculating the p ~~matrix~~value:

Since S_a and S_ε are positive definite symmetric ~~matrixes, it~~matrices, they can be decomposed into $S_a = (S_a^{1/2})^T(S_a^{1/2})$ and $S_\varepsilon = (S_\varepsilon^{1/2})^T(S_\varepsilon^{1/2})$.

Define $\mathrm{R} = S_\varepsilon^{1/2} K S_a^{1/2}$. (6)

The matrix R can then be regarded as a weighting function matrix, normalized by the observed error and a priori uncertainty. A row vector of R, $\mathrm{r} = s_\varepsilon^{-1/2} k S_a^{1/2}$, represents the normalized weighting function matrix of a single channel. Substituting r into Eq. (5) gives:

$$\mathrm{p} = 1 - \exp\left(\frac{1}{2n} ln\left(\left|I - \frac{rr^T}{1 + r^T r}\right|\right)\right). \qquad (7)$$

For arbitrary row vectors, a and b, using the matrix property $\det(\mathrm{I} + a^T \mathrm{b}) = 1 + \mathrm{b}a^T$, the new expression for p is:

$$\mathrm{p} = 1 - \exp\left(\frac{1}{2n} ln\left(1 - \frac{r^T r}{1 + r^T r}\right)\right)$$

$$= 1 - \exp\left(\frac{1}{2n} ln\left(\frac{1}{1+r^T r}\right)\right)$$

$$= 1 - \exp\left(-\frac{1}{2n} ln(1 + r^T r)\right). \quad (8)$$

(~~3~~III) Iteration in a single layer:

First, the iteration in a single layer requires the calculation of R. ~~According to~~Using S_a, S_ε , K and Eq. (6), R can be calculated. Second, using Eq. (8), p of each candidate channel can be calculated. Moreover, the channel corresponding to maximum p is the selected channel for this iteration. After a channel has been selected, according to Eq. (3) we can use $\hat{S}$ to get S_a for the next iteration. Finally, channels which are not selected during this iteration are used as the candidate channels for the next iteration.

When selecting n from N channels, it is necessary to calculate (N-n/2)n≈Nn p values, which is much smaller than C_N^n. In addition to high computational efficiency by using this method, another advantage is that all channels can be recorded in the order in which they are selected. In the actual application, if n' channels are needed, and $n' < n$, we will not need to select the channel again, but record the selected channel only.

(~~4~~IV) Iteration for different altitudes:

Because satellite channel sensitivity varies with height, repeating the iterative process of step (~~3~~III), selects the optimum channels at different heights. Assuming there are M layers in the atmosphere and selecting n from N channels, it is necessary to calculate $\mathrm{M} \cdot (\mathrm{N} - \mathrm{n}/2)\mathrm{n} \approx \mathrm{M} \cdot \mathrm{Nn}$ p values, a much smaller number than $\mathrm{M} \cdot \mathrm{C}_N^n$. In this way, different channel sets can be used to evaluate corresponding height in the retrieved profiles.

2.3 Statistical inversion method

The inversion methods for the atmospheric temperature profiles can be summarized in two categories: statistical inversion and physical inversion. Statistical inversion is essentially a linear regression model which uses a large number of satellite measurements and atmospheric parameters to match samples and calculate their correlation coefficient. Then, based on the correlation coefficient, the required parameters of the independent measurements obtained by the satellite are retrieved. Because the method does not directly solve the radiation transfer equation, it has the advantages of fast calculation speed. In addition, the solution is numerically stable, which makes it one of the highest precision methods (Chedin et al., 1985). Therefore, the statistical inversion method will be used for our channel selection experiment and a regression equation will be established.

According to an empirical orthogonal function, the atmospheric temperature (or humidity), T, and the brightness temperature, T_b, are expanded as:

$$\mathrm{T} = T^* \cdot A, \tag{9}$$

$$T_b = T_b^* \cdot A, \tag{10}$$

where T^* and T_b^* are the eigenvectors of the covariance matrix of temperature (or humidity) and brightness temperature, respectively. A and B stand for the corresponding expansion coefficient vectors of temperature (humidity) and brightness temperature.

Using the least squares method and the orthogonal property, the coefficient conversion matrix, V, is introduced:

$$\mathrm{A} = \mathrm{V} \cdot B, \tag{11}$$

where $V = AB^T(BB^T)^{-1}$. (12)

Using the orthogonality, we get:

$$\mathrm{B} = (T_b^*)^T T_b, \tag{13}$$

$$A = (T^*)^T T. \tag{14}$$

For convenience, the anomalies of the state vector (atmospheric temperature), T, and the observation vector (brightness temperature), T_b, are taken:

$$\hat{T} = \overline{T} + \hat{T}' = \overline{T} + GT_b' = \overline{T} + G(T_b - \overline{T_b}), \tag{15}$$

where $\hat{T}$ stands for the retrieval atmospheric temperature. $\overline{T}$ and $\overline{T_b}$ are the corresponding average values of the elements, respectively. $\hat{T}'$ and T_b' represent the corresponding anomalies of the elements, respectively.

Assuming there are k sets of observations, a sample anomaly matrix with k vectors can be constructed:

$$T' = (t_1', t_2', \cdots, t_k'), \tag{16}$$

$$T_b' = (t_{b1}', t_{b2}', \cdots, t_{bk}'). \tag{17}$$

Define the inversion error matrix as：

$$\delta = \overline{T} - \widehat{T} = \widehat{T}' - T' . \tag{18}$$

The retrieval error covariance matrix is:

$$\begin{aligned} S_\delta &= \frac{1}{k-n-1}\delta\delta^T \\ &= \frac{1}{k-n-1}(T' - GT_b')(T' - GT_b')^T \\ &= \frac{k-1}{k-n-1}(S_e - G^T S_{xy} - S_{xy}G^T + GS_yG^T), \end{aligned} \tag{19}$$

where

$$\begin{aligned} S_e &= \frac{1}{k-1}T' T'^T, \\ S_y &= \frac{1}{k-1}T_b' T_b'^T, \\ S_{xy} &= \frac{1}{k-1}T' T_b'^T. \end{aligned} \tag{20}$$

S_e stands for the sample covariance matrix of T, S_y denotes the sample covariance matrix of T_b, and S_{xy} represents the covariance matrix of T and T_b. The elements on the diagonal of the error covariance matrix, S_δ, represent the retrieval error variance of T. The matrix G that minimizes the overall error variance is the least squares coefficient matrix of the regression equation (15), which meets the criteria:

$$\delta^2 = \mathrm{tr}(S_\delta) = \min. \tag{21}$$

Taking a derivative of Eq. (21) with respect to G, $\frac{\partial}{\partial G}\mathrm{tr}(S_\delta) = 0 = (-2S_{xy} + 2GS_y)$, which means that:

$$G = S_{xy}S_y^{-1}. \tag{22}$$

Substituting Eq. (22) into Eq. (15) finally gives the least squares solution as:

$$\hat{T} = \overline{T} + S_{xy}S_y^{-1}(T_b - \overline{T_b}). \tag{23}$$

It should be noted that the least squares solution obtained here aims to minimize the sum of the error variance for each element in the atmospheric state vector after retrieval for several times. At present, statistical multiple regression is widely used in the retrieval of atmospheric profiles based on atmospheric remote sensing data. As long as there are enough data, S_{xy} and S_y can be determined.

3. Channel selection experiment

3.1 Data and model

The Atmospheric Infrared Sounder (AIRS) is primarily designed to measure the Earth's atmospheric water vapor and temperature profiles on a global scale (Aumann et al., 2003; ~~Hoffmann and Alexander, 2009).~~Susskind et al., 2003). AIRS is a continuously operating cross-track scanning sounder, consisting of a telescope that feeds an echelle spectrometer. The AIRS infrared spectrometer acquires 2378 spectral samples at a resolution $\lambda/\Delta\lambda$, ranging from 1086 to 1570, in three bands: 3.74 μm to 4.61 μm, 6.20 μm to 8.22 μm, and 8.8 μm to 15.4 μm. The footprint size is 13.5 km~~ at nadir (Susskind et al., 2003).~~. The spectral range includes 4.3 μm and 15.5 μm for important temperature observation and CO_2, 6.3 μm for water vapor, and 9.6 μm for ozone absorption bands (Menzel et al., 2018). The root mean square error (RMSE) of the measured radiation is better than 0.2 K (Susskind et al., 2003). Moreover, global atmospheric profiles can be detected every day. Due to radiometer noise and faults, there are currently only 2047 effective channels. However, compared with previous infrared detectors, AIRS boasts a significant improvement in both the number of channels and spectral resolution (Aumann, 1994; Huang et al., 2005; Li et al., 2005).

The root mean square error of an AIRS infrared channel is shown in Fig. 1~~, with black spots, indicating that~~. The measurement error is not below 0.2K for all the instrument channels ~~possess a measurement error of less than 0.2 K.~~. There are a few channels with extremely large measurement errors, which reduce the accuracy of prediction to some extent. Among them, some extremely large measurement errors reduce the accuracy of prediction to some extent (Susskind et al., 2003). At present, more than 300 channels have not been used because their errors exceed 1 K. If data from these channels were to be used for retrieval, the accuracy of the retrieval could be reduced. Therefore, it is necessary to select a group of channels to improve the calculation efficiency and retrieval quality. In this paper we study channel selection for temperature profile retrieval by AIRS.

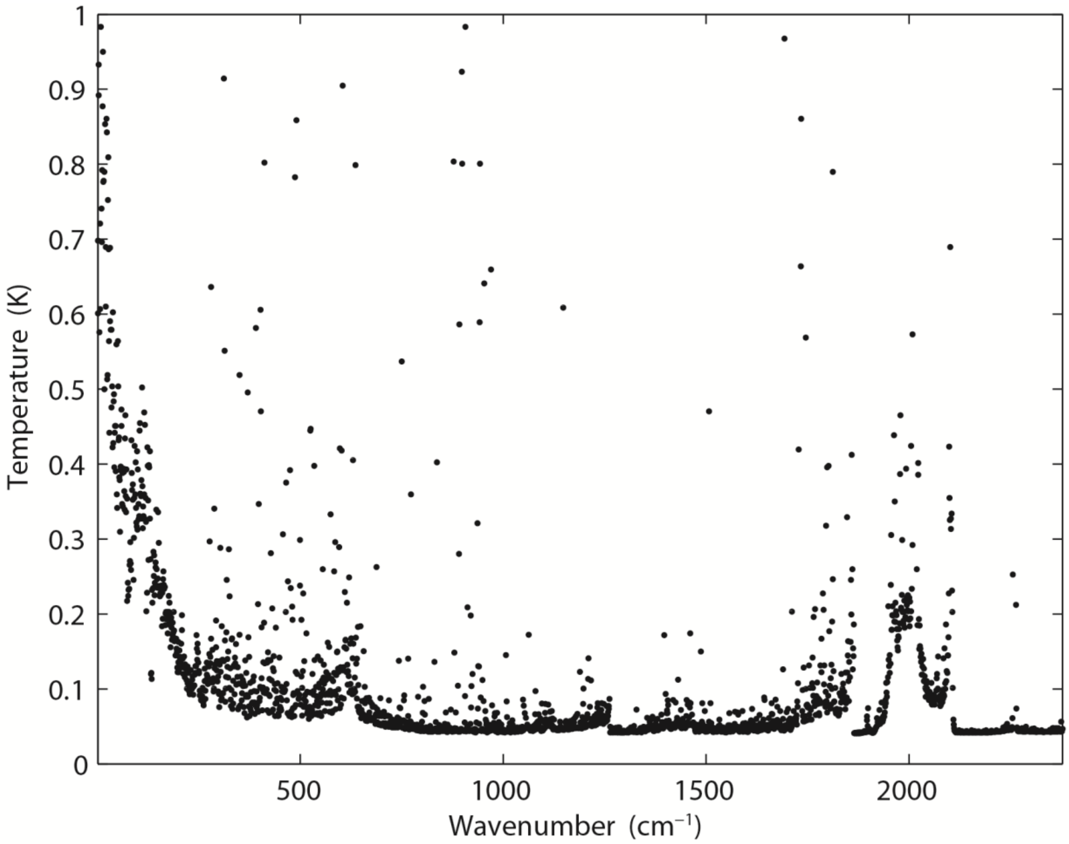


Figure 1. Root mean square error of AIRS infrared channel (black spots).

For the calculation of radiative transfer and the weighting function matrix, K, the RTTOV (Radiative Transfer for TOVS) v12 fast radiative transfer model is used. Although initially developed for the TOVS (TIROS Operational Vertical Sounder) radiometers, RTTOV can now simulate around 90 different satellite sensors measuring in the MW (microwave), IR (infrared) and VIS (visible) regions of the spectrum (Saunders et al., 2018). The model allows rapid simulations (1 ms for 40 channel ATOVS (Advanced TOVS) on a desktop PC) of radiances for satellite visible, infrared, or microwave nadir scanning radiometers given atmospheric profiles of temperature and trace gas concentrations, and cloud and surface properties. The only mandatory gas included as a variable for RTTOV v12 is water vapor. Optionally, ozone, carbon dioxide, nitrous oxide, methane, carbon monoxide, and sulfur dioxide can be included, with all other constituents assumed to be constant. RTTOV can accept input profiles on any defined set of pressure levels. The majority of RTTOV coefficient files are based on the 54 levels (see Table A1 in Appendix A), ~~ranking~~in the range from 1050 hPa to 0.01 hPa, though coefficients for some hyperspectral sounders are also available on 101 levels.

In order to correspond to the selected profiles, the atmosphere is divided into 137 layers, each of which contains corresponding atmospheric characteristics, such as temperature, pressure, and the humidity distribution. Each element in the weighting function matrix can be written as $\partial yi/\partial xj$. The subscript i is used to identify the satellite channel, and the subscript j is used to identify the atmospheric variable. Therefore, $\partial yi/\partial xj$ indicates the variation in brightness temperature in a given satellite channel, when a given atmospheric variable in a given layer changes. We are thus able to establish which layer of the satellite channel is particularly sensitive to which atmospheric characteristic (temperature, various gas contents) in the vertical atmosphere. The RTTOV_K (the K mode), is used to calculate the matrix H(X0) (Eq. (1)) for a given atmospheric profile characteristic.

3.2 Channel selection comparison experiment and results

In order to verify the effectiveness of the method, three sets of comparison experiments were conducted. First, 324 channels used by the EUMETSAT Satellite Application Facility on Numerical Weather Prediction (NWP SAF) were selected. NCS is short for NWP channel selection in this paper. NCS were released by the

NWPSAF 1DVar (one-dimensional variational analysis) scheme, in accordance with the requirements of the NWPSAF (Saunders et al., 2018). Second, 324 channels were selected using the information capacity method. This method was adopted by Du et al. (2008) without the consideration of layering. PCS is short for primary channel selection in this paper.

Third, 324×M channels were selected using the information capacity method for the M layer atmosphere. ICS is short for improved channel selection in this paper. In order to verify the retrieval effectiveness after channel selection, statistical inversion comparison experiments were performed using 5000 temperature profiles provided by the ECMWF dataset, which will be introduced in Sect. 4.

The observation error covariance matrix, S_ε, in the experiment is provided by NWP SAF 1Dvar. In general, it can be converted to a diagonal matrix, the elements of which are the observation error standard deviation of each hyperspectral detector channel, which is the square of the root mean square error for each channel. The root mean square error of the AIRS channels is shown in Fig. 1. The error covariance matrix of the background, S_a, is calculated using 5000 samples of the IFS-137 data provided by the ECMWF dataset (The detailed information will be introduced in Sect. 4). The last access date is April 26th, 2019 (download address: https://www.nwpsaf.eu/site/update-137-level-nwp-profile-dataset/, 2019). The covariance matrix of temperature is shown in Fig. 2. The results are consistent with the previous study by Du et al. (2008).

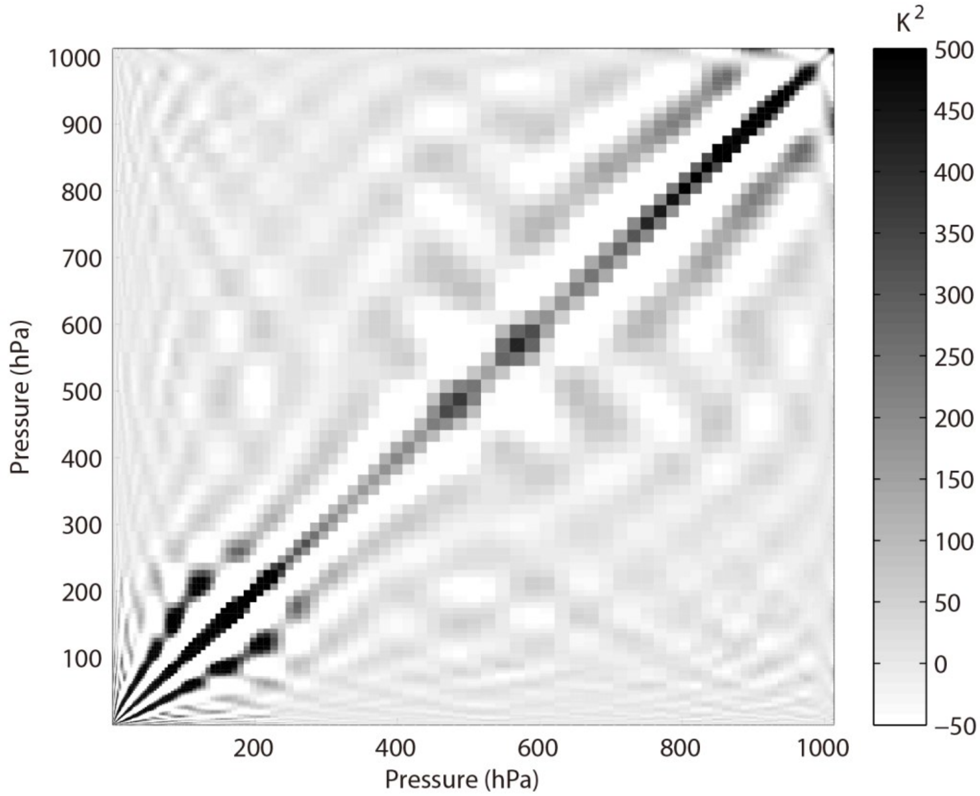


Figure 2. Error covariance matrix of temperature (shaded).

The reference atmospheric profiles are from the IFS-137 database, and the temperature weighting function matrix is calculated using the RTTOV_K mode, as shown in Fig. 3; the results are consistent with those of the previous study by Du et al. (2008). For the air-based passive atmospheric remote sensing studied in this paper, when the same channel detects the atmosphere from different observation angles, the value of the weighting function matrix K changes due to the limb effect. The goal of this section is focusing on the selection methods of selecting channels; therefore the biases produced from different observation angles can be ignored.

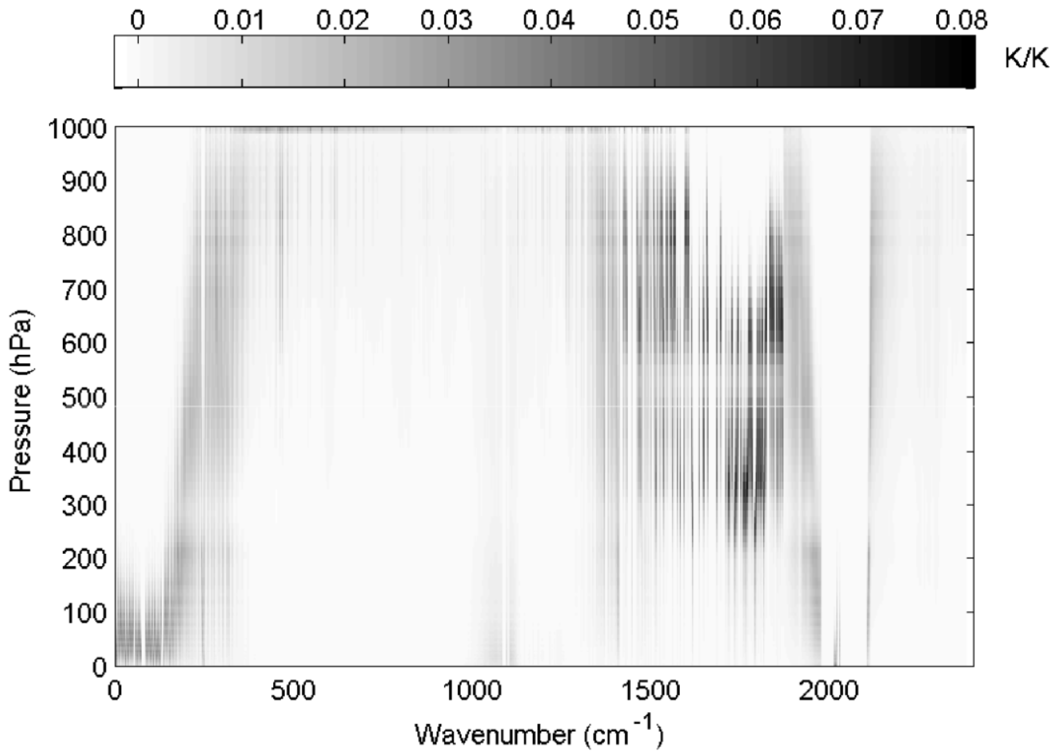


Figure 3. Temperature weighting function matrix (shaded).

In order to verify the effectiveness, the distribution of 324 channels, without considering layering, in the AIRS brightness temperature spectrum is indicated in Fig. 4. The background brightness temperature is the simulated AIRS observation brightness temperature, which is from the atmospheric profile in RTTOV put into the model. Figure 4(a) shows the 324 channels selected by PCS, while Fig. 4(b) shows the 324 channels selected by NCS.

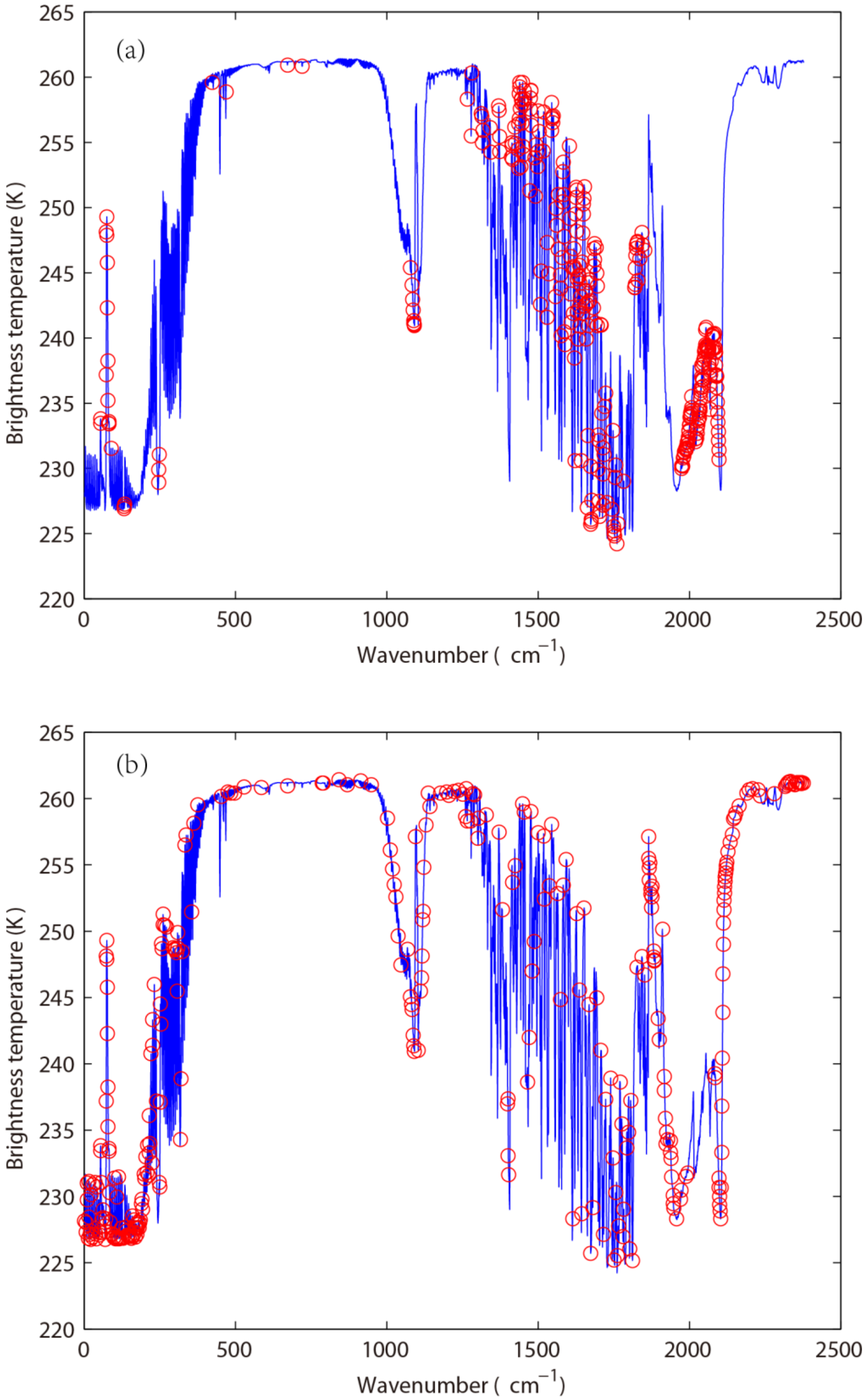


Figure 4. The distribution of different channel selection methods without considering layering in the AIRS brightness temperature

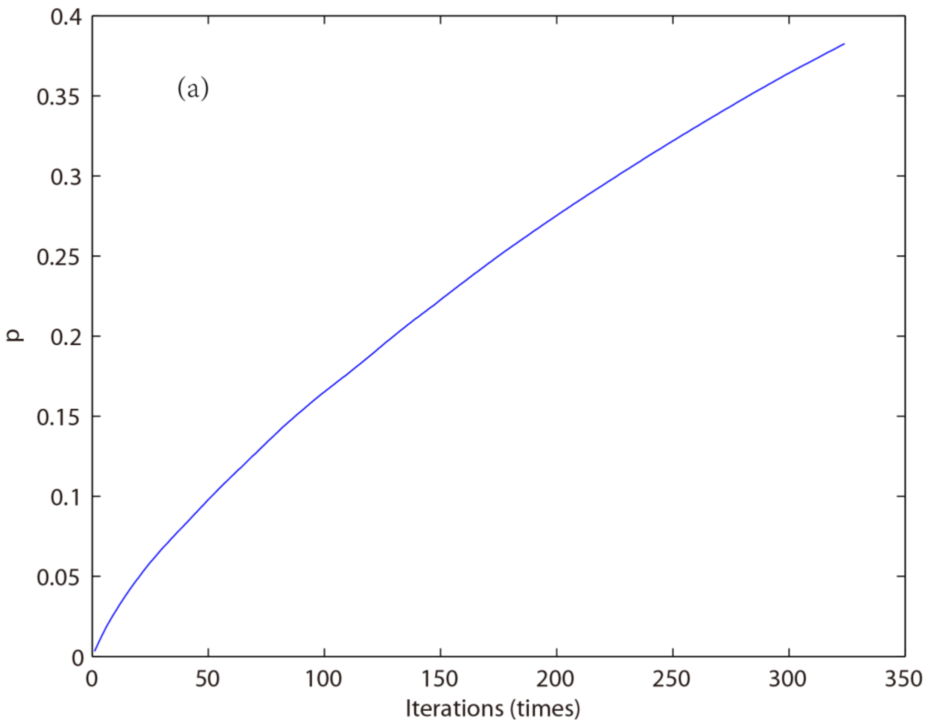
spectrum (blue line). (a) 324 channels selected by PCS (red circles). (b) 324 channels selected by NCS (red circles).

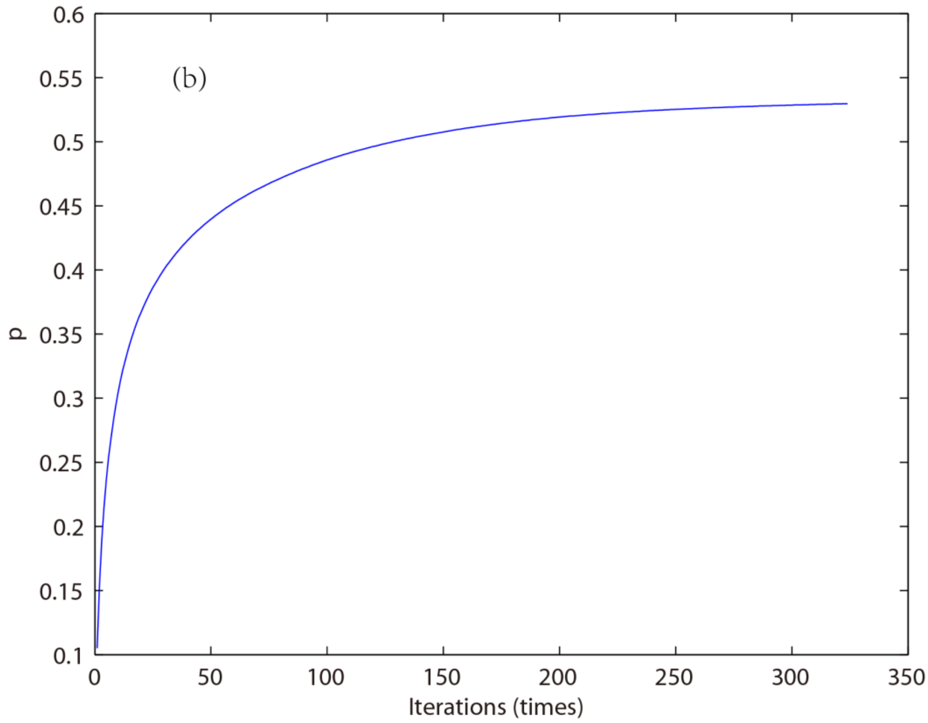
Without considering layering, the main differences between the 324 channels selected by PCS and NCS are as follows: (1) ~~When the wavenumber approaches 1000, the wavelength is 10 μm (1/1000 cm^{-1}). Near this~~Near 10 μm band, fewer channels are selected by PCS because the retrieval of ground temperature is considered by NCS; (2) ~~When the wavenumber is near 1200, the wavelength is 9 μm (1/1200 cm^{-1}). Near this~~Near 9 μm band, no channels are selected by PCS because the retrieval of O3 is not considered in this paper; (3) ~~When the wavenumber approaches 1500, the wavelength is 6.7 μm (1/1500 cm^{-1}).~~ As is known, the spectral range from 6 μm to 7 μm corresponds to water vapor absorption bands, but fewer channels are selected by NCS; (4) ~~When the wavenumber is close to 2000, it derives a wavelength of 5 μm (1/2000 cm^{-1}), which~~Near 5 μm band, it includes 4.2 μm for N_2O and 4.3 μm for CO_2 absorption bands. As is shown in Fig. 4, fewer channels are selected by PCS in those bands. PCS is favorable for atmospheric temperature observation~~ in the high temperature zone.~~. Because 4.2 μm and 4.3 μm bands are sensitive to high temperature, ~~the higher temperature is, the~~a better observation can be obtained for higher temperatures; (5) ~~In the near infrared area, the wavenumber exceeds 2200, deriving a~~

~~wavelength of less than~~Near 4 μm ~~(1/2000 cm^{-1}). A~~band, a small number of channels is selected by NCS, but no channels are selected by PCS.

Above all, the information content ~~used~~considered in this ~~paper~~study only takes the temperature profile retrieval into consideration, so the channel combination of PCS is inferior to that of NCS for the retrieval of surface temperature and the O_3 profile. The advantages of the channel selection method based on information content in this paper are mainly reflected in: (1) Stratosphere and mesosphere is less affected by the ground surface, so the retrieval result of PCS is better than that of NCS. (2) Due to the method selected in this paper there are more channels at 4.2 μm for N_2O and 4.3 μm for CO_2 absorption bands; the channel combination of PCS is better than that of NCS for atmospheric temperature observation ~~to the~~at higher temperature.

By comparing channel selection without considering layering, we note the general advantages and disadvantages of PCS and NCS for the retrieval of temperature and can improve the channel selection scheme. First, the retrieval of the temperature profile for 324 channels selected by PCS is obtained. The relationship between the number of iterations and the ARI is shown in Fig. 5.


(a)
p
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0 50 100 150 200 250 300 350
Iterations (times)


(b)
p
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0 50 100 150 200 250 300 350
Iterations (times)

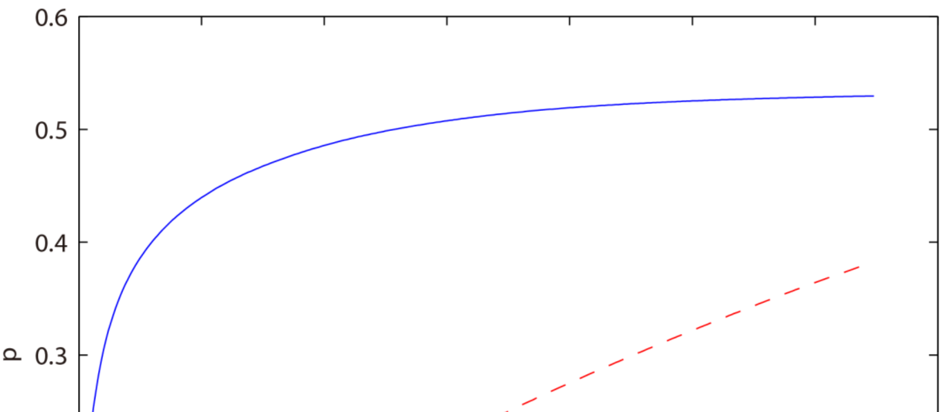

p
0.6
0.5
0.4
0.3

Figure 5. The relationship between the number of iterations and ARI. ~~(a)~~Blue line represents the result of ICS. Red dotted line stands for the result of PCS. ~~(b) ICS.~~

The ARI for PCS tends to be 0.38 and is not convergent, so the PCS method needs to be improved. In this paper, the atmosphere is divided into 137 layers, and based on the information content and iteration, 324 channels are selected for each layer. Then, the temperature profile of each layer can be retrieved based on statistical inversion (see at Sect. 4). The relationship between the number of iterations and the ARI for ICS is shown in Fig. 5b. When the number of iterations approaches 100, the ARI of ICS tends to be stable, and ~~reach to~~reaches 0.54. Thus, in terms of the ARI and convergence, the ICS method is better than that of PCS.

Furthermore, because an iterative method is used to select channels, the order of each selected channel is determined by the contribution from the ARI. The weighting function matrix of the top 324 selected channels, according to channel order, is shown in Fig. 6.

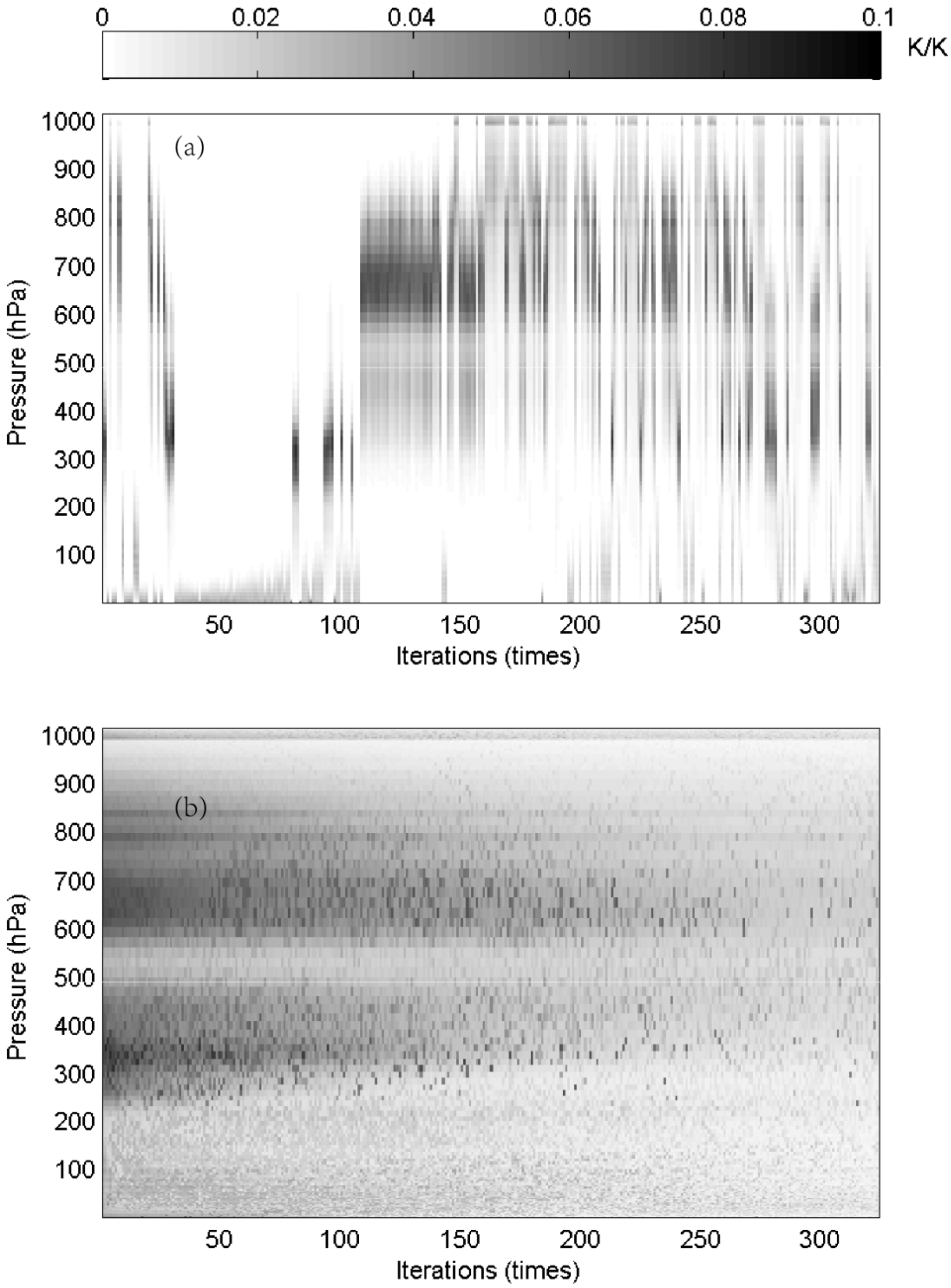


Figure 6. The relationship between the number of iterations and the weighting function of the top 324 selected channels (shaded). (a) ICS. (b) PCS.

As illustrated in Fig. 6, in the first 100 iterations, the distribution of the temperature weighting function for PCS is relatively scattered; it does not reflect continuity between the adjacent layers of the atmosphere. Besides, the ICS result is better than that of PCS, showing that: (1) the distribution of the temperature weighting function is more continuous and reflects the continuity between adjacent layers of the atmosphere; (2) regardless of the number of iterations, the maximum value of the weighting function is stable near 300–400 hPa and 600–700 hPa, without scattering, which is closer to the situation in real atmosphere.

4. Statistical multiple regression experiment

4.1 Temperature profile database

A new database including a representative collection of 25,000 atmospheric profiles from the European Centre for Medium-range Weather Forecasts (ECMWF) was used for the statistical inversion experiments. The profiles were given in a 137-level vertical grid extending from the surface up to 0.01 hPa. The database was divided into five subsets focusing on diverse sampling characteristics such as temperature, specific humidity, ozone mixing ratio, cloud condensates, and precipitation. In contrast with earlier releases of the

ECMWF diverse profile database, the 137-level database places greater emphasis on preserving the statistical properties of sampled distributions produced by the Integrated Forecasting System (IFS) (Eresmaa and McNally, 2014; Brath et al., 2018). IFS-137 spans the period from September 1, 2013 to August 31, 2014. There are two operational analyses each day (at 00z and 12z), and approximately 13 000 atmospheric profiles over the ocean. The pressure levels adopted for IFS-137 are shown in Table A2 (see Table A2 in Appendix A).

The locations of selected profiles of temperature, specific humidity, and cloud condensate subsets of the IFS-91 and IFS-137 databases are plotted on the map in Fig. 7. In the IFS-91 database, the sampling is fully determined by the selection algorithm, which makes the geographical distributions very inhomogeneous. Selected profiles represent those regions where gradients of the sampled variable are the strongest: in the case of temperature, mid- and high-latitudes dominate, while humidity and cloud condensate subsets concentrate at low latitudes. However, the IFS-137 database shows a much more homogeneous spatial distribution in all the sampling subsets, which is a consequence of the randomized selection.

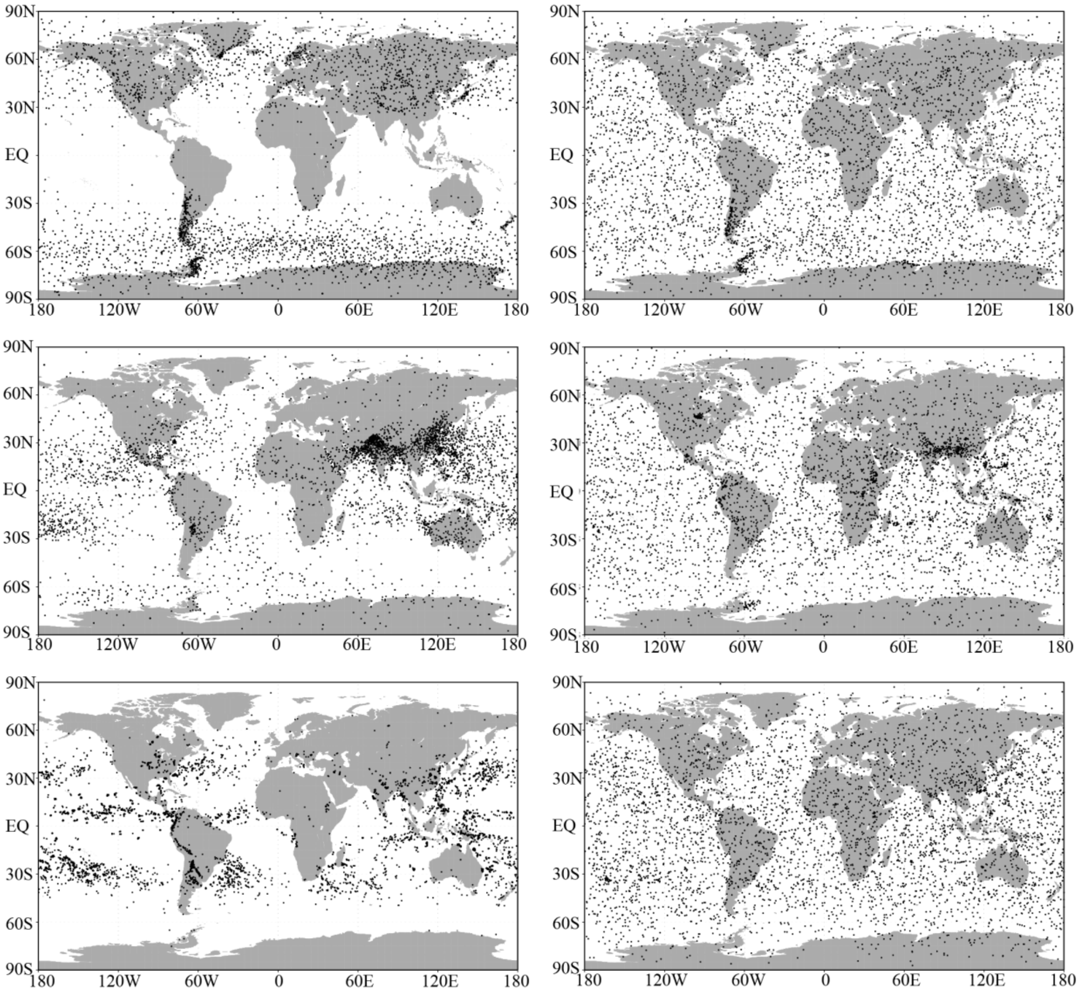


Figure 7. Locations of selected profiles in the temperature (top), specific humidity (middle), and cloud condensate (bottom), sampled subsets of the IFS-91 (left) and IFS-137 (right) databases (from https://www.nwpsaf.eu/site/update-137-level-nwp-profile-dataset/ , 2019).

The temporal distribution of the selected profiles is illustrated in Fig. 8. The coverage of the IFS-137 data set is more homogeneous than the IFS-91 data set. Moreover, the IFS-137 database supports the mode with input parameters, such as detection angle, 2 m temperature, and cloud information. Therefore, it is feasible to use the selected samples in a statistical multiple regression experiment.

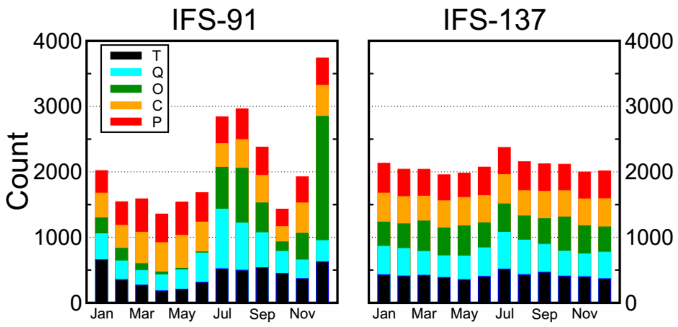


Figure 8. Distribution of profiles within the calendar months in IFS-91 (left) and IFS-137 (right) databases. Different subsets are shown in different colors.Black parts stand for temperature. Blue parts represent specific humidity. Green parts indicate ozone subset. Orange parts stand for cloud condensate. Red parts represent precipitation. The last access date is April 26th, 2019. (from https://www.nwpsaf.eu/site/update-137-level-nwp-profile-dataset/ , 2019).

4.2 Experimental scheme

In order to verify the retrieval effectiveness of ICS, 5000 temperature profiles provided by the IFS-137 were used for statistical inversion comparison experiments. The steps are as follows:

(1) 5000 profiles and their corresponding surface factors, including surface air pressure, surface temperature, 2 m temperature, 2 m specific humidity, 10 m wind speed~~,~~ are put into the RTTOV mode. Then, the simulated AIRS spectra are obtained.

(2) The retrieval of temperature is carried out in accordance with Eq. (23). The 5000 profiles are divided into two groups. The first group of 2500 profiles is used to obtain the regression coefficient, and the second group of 2500 is used to test the result.

(3) Verification of the results. The test is carried out based on the standard deviation between the retrieval value and the true value.

4.3 Results and Discussion

For the statistical inversion comparison experiments, the standard deviation of temperature retrieval is shown in Fig. 9. First, because PCS does not take channel sensitivity as a function of height into consideration, the retrieval result of PCS is inferior to that of ICS. Second, by comparing the results of ICS and NCS we found that below 100 hPa, since the method used in this paper considers near ground to be less of an influencing factor, the channel combination of ICS is slightly inferior to that of NCS, but the difference is small.

From 100 hPa to10 hPa, the retrieval temperature of ICS in this paper is consistent with that of NCS, slightly better than the channel selected for NCS. From 10 hPa to 0.02 hPa, near the space layer, the retrieval temperature of ICS is better than that of NCS. In terms of the standard deviation, the channel combination of ICS is slightly better than that of PCS from 100 hPa to 10 hPa. From 10 hPa to 0.02 hPa, the standard deviation of ICS is lower than that of NCS at about 1 K, meaning that the retrieval result of ICS is better than that of NCS.

In order to further illustrate the effectiveness of ICS, the mean improvement value of the ICS and its percentages compared with the PCS and NCS at different heights are shown in Table 1. Because PCS does not take channel sensitivity as a function of height into consideration, the retrieval result of PCS is inferior to that of ICS. In general, the accuracy of the retrieval temperature of ICS is improved. Especially, from 100 hPa to 0.01 hPa, the mean value of ICS is evidently improved by more than 0.5 K which means the accuracy can be improved by more than 11%. By comparing the results of ICS and NCS we found that below 100 hPa, since the method used in this paper considers near ground to be less of an influencing factor, the channel combination of ICS is slightly inferior to that of NCS, but the difference is small. From 100 hPa to 0.01 hPa, the mean value of ICS is improved by more than 0.36 K which means the accuracy can be improved by more than 9.6%.

Table 1. The mean improvement value of the ICS and its percentages compared with the PCS and NCS at different heights.

Pressure	Improved mean value /Percentage compared with PCS	Improved value /Percentage compared with NCS
hPa	K/%	K/%
surface-100hPa	0.24/10.77%	-0.04/-3.27%
100hPa-10hPa	0.15/5.08%	0.06/2.4%
10hPa-1hPa	0.04/0.64%	0.17/2.99%
1hPa-0.01hPa	0.52/11.92%	0.36/9.57%

This is because, as shown in Fig. 4: (1) Stratosphere and mesosphere is less affected by the ground surface, so the retrieval result of PCS is better than that of NCS. (2) Due to the method selected in this paper, there are more channels at 4.2 μm for N_2O and 4.3 μm for CO_2 absorption bands, and the channel combination of PCS is superior to that of NCS for atmospheric temperature observation in the high temperature zone. Moreover, ICS takes channel sensitivity as a function of height into consideration, so its retrieval result is improved.

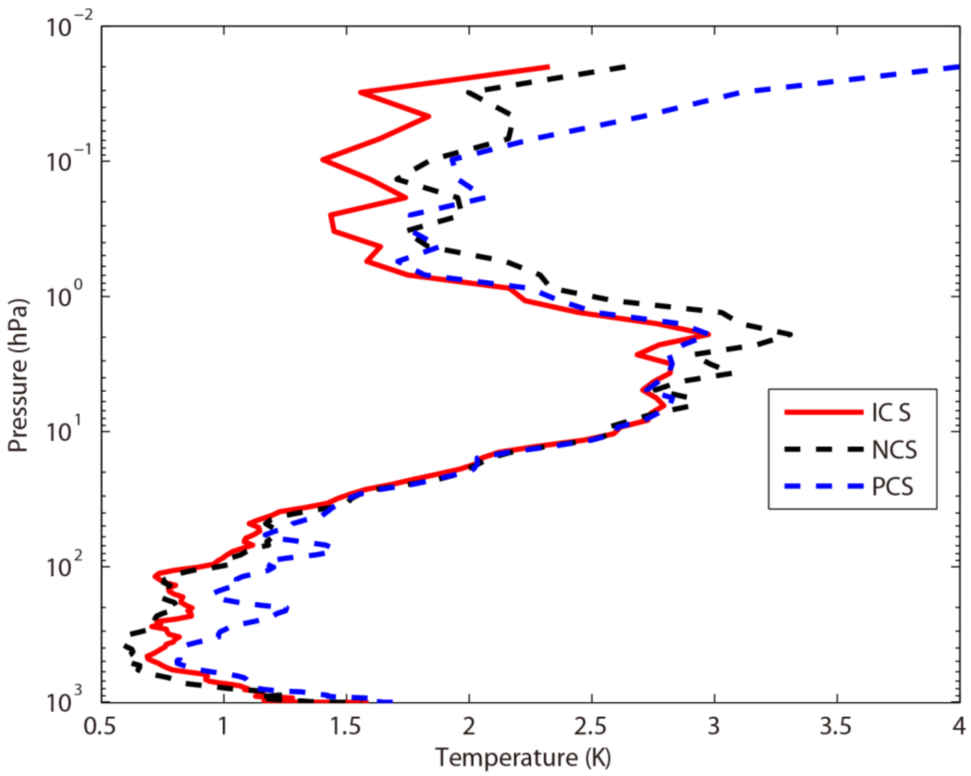


Figure 9. The temperature profile standard deviation of statistical inversion comparison experiments. Red line indicates the result of ICS. Black dotted line stands for the result of NCS. Blue dotted line represents the result of PCS.

5 Statistical inversion comparison experiments in four typical regions

The accuracy of the retrieval temperature varies from place to place and changes with atmospheric conditions. Therefore, in order to further compare the inversion accuracy under different atmospheric conditions, this paper has divided the atmospheric profile from the IFS-137 database introduced in Sect. 4 into four regions: equatorial zone, subtropical region, mid-latitude region and Arctic. The average temperature profiles in these four regions are shown in Fig. 10. The retrieval temperature varies from place to place and changes with atmospheric conditions. In order to further compare the regional differences of inversion accuracy, the temperature standard deviations of ICS in four typical regions are compared in Sect. 5.2.

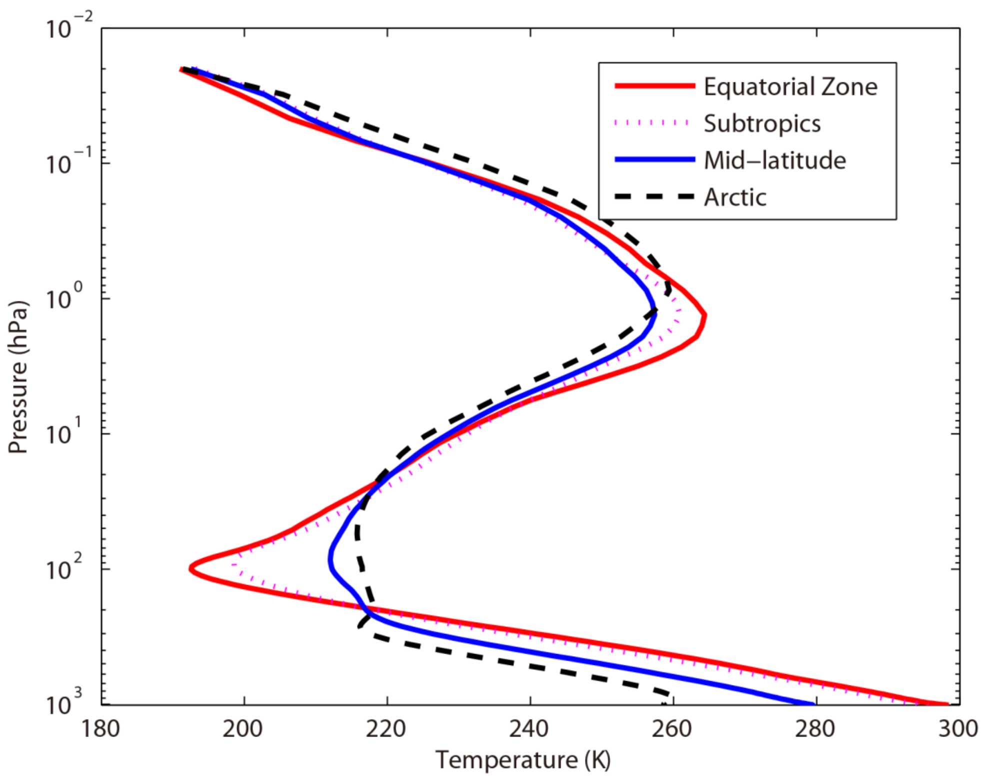


Figure 10. The average temperature profiles in four typical regions. Red line indicates the equatorial zone. Pink dotted line stands for the subtropics. Blue dotted line represents the mid-latitude region. Black dotted line stands for the Arctic.

5.1 Experimental scheme

In order to further illustrate the different accuracy of the retrieval temperature using our improved channel selection method under different atmospheric conditions, the profiles in four typical regions were used for statistical inversion comparison experiments. The experimental steps are as follows:

(1) 2500 profiles in Sect. 4 are used to work out the regression coefficient.

(2) The atmospheric profiles of the four typical regions: equatorial zone, subtropical region, mid-latitude region and Arctic are used for statistical inversion comparison experiments and test the result.(3) Verification of the results. The test is carried out based on the standard deviation between the retrieval value and the true value.

5.2 Results and Discussion

Using statistical inversion comparison experiments in four typical regions, the standard deviation of temperature retrieval is shown in Fig. 11. Generally, the retrieval temperature by ICS is better than that of NCS and PCS. In particular, above 1 hPa (the stratosphere and mesosphere), the standard deviation of atmospheric temperature can be improved by 1 K with PCS and NCS. Thus, ICS shows a great improvement. The results were consistent with Sect. 4.

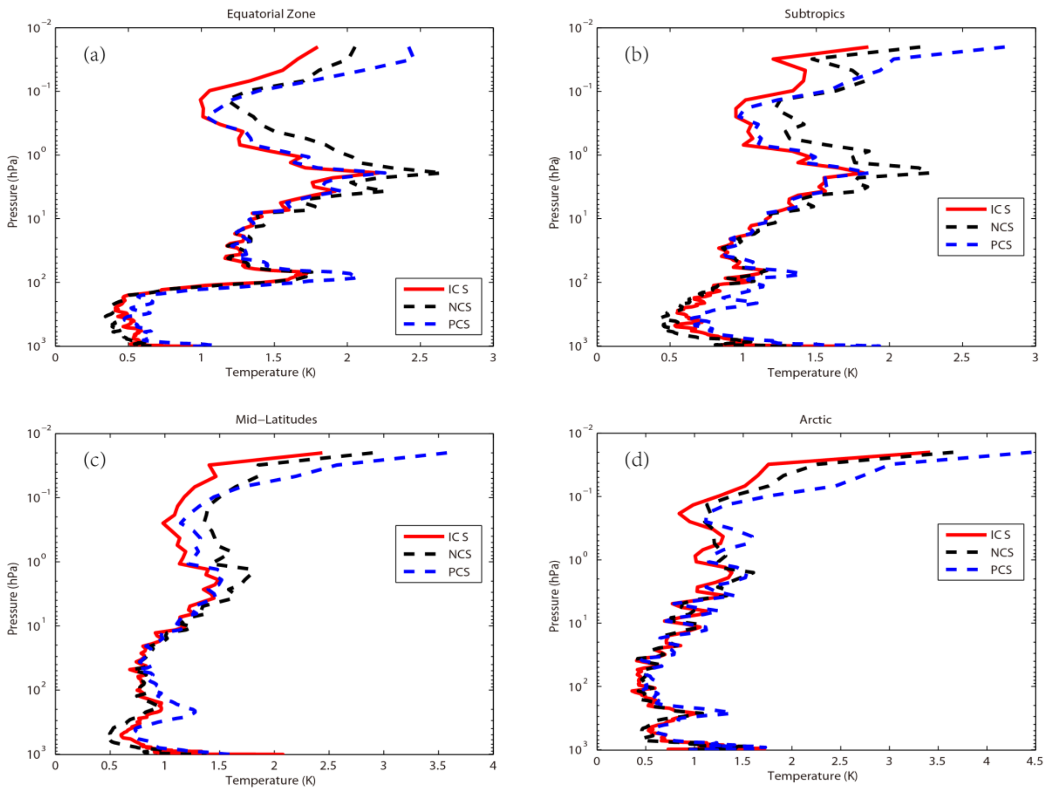


Figure 11. The temperature profile standard deviation of statistical inversion comparison experiments in four typical regions. Red line indicates the result of ICS. Black dotted line stands for the result of NCS. Blue dotted line represents the result of PCS. (a) Equatorial zone. (b) Subtropics. (c) Mid-latitudes. (d) Arctic.

In order to further compare the regional differences of inversion accuracy, the temperature standard deviation of ICS in four typical regions are compared in Fig. 12.

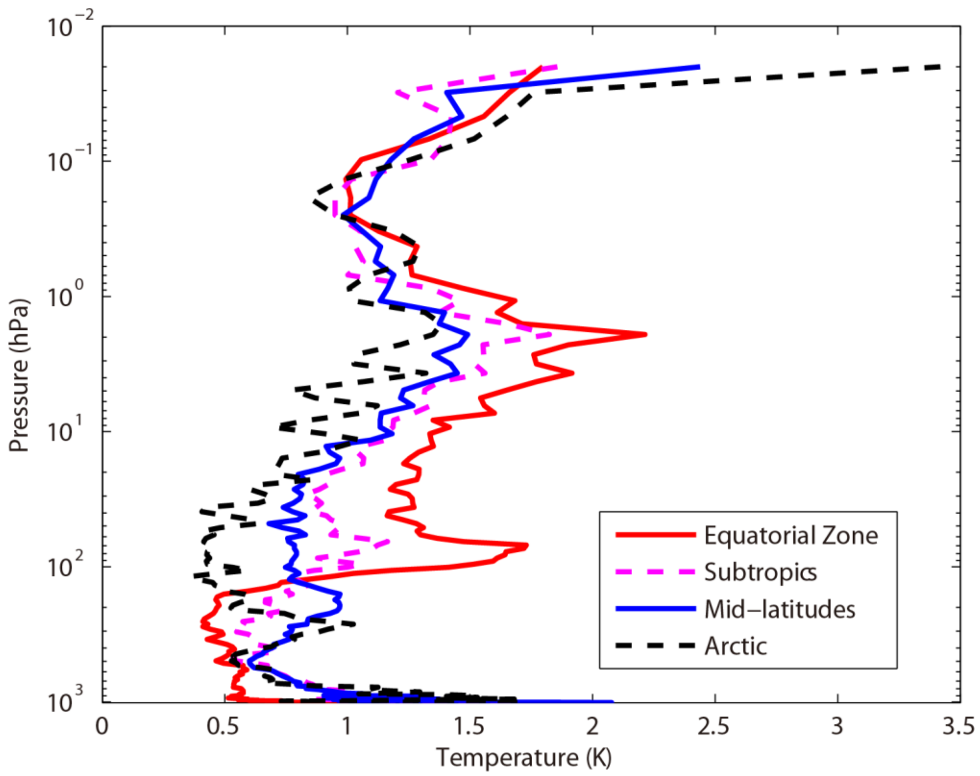


Figure 12. The temperature standard deviation of ICS in four typical regions. Red line indicates the result of equatorial zone. Pink dotted line represents the result of Subtropics. Blue line represents the result of Mid-latitudes. Black dotted line stands for the result of Arctic.

The temperature standard deviations of the ICS in the four typical regions are large (Fig. 12). Below100 hPa, due to the high temperature in the equatorial zone, the channel combination of ICS is better than that of PCS and NCS for atmospheric temperature observation ~~to the~~at higher temperature. The standard deviation is 0.5K. Due to the method selected in this paper there are more channels at 4.2 μm for N_2O and 4.3 μm for CO_2 absorption bands which has been previously described in Sect. 3. Near the tropopause, the standard deviation of the equatorial zone increases sharply. It is also due to the sharp drops in temperature. However, the standard deviation of the Arctic is still around 0.5K. From 100hPa to 1hPa, the standard deviation of ICS is 0.5 K to 2K. With the increase of latitude, the effectiveness considerably increases. According to Fig. 11, ICS takes channel sensitivity as a function of height into consideration, so its retrieval result is better.

Although the improvements of ICS in the four typical regions are different, in general, the accuracy of the retrieval temperature of ICS is improved. Because PCS does not take channel sensitivity as a function of height into consideration, the retrieval result of PCS is inferior to that of ICS. In general, the accuracy of the retrieval temperature of ICS is improved.

7 Conclusions

In recent years, the atmospheric layer in the altitude range of about 20–100 km has been named "the near space layer" by aeronautical and astronautical communities. It is between the space-based satellite platform and the aerospace vehicle platform, which is the transition zone between aviation and aerospace. Its unique resource has attracted a lot of attention from many countries. Research and exploration, therefore, on and of the near space layer are of great importance. A new channel selection scheme and method for hyperspectral atmospheric infrared sounder AIRS data based on layering is proposed. The retrieval results of ICS concerning the near space atmosphere are particularly good. Thus, ICS aims to provide a new and an effective channel selection method for the study of the near space atmosphere using the hyperspectral atmospheric infrared sounder.

An improved channel selection method is proposed, based on information content in this paper. A robust channel selection scheme and method are proposed, and a series of channel selection comparison experiments are conducted. The results are as follows:

(1) Since ICS takes channel sensitivity as a function of height into consideration, the ARI of PCS only tends to be 0.38 and is not convergent. However, as the 100th iteration is approached, the ARI of ICS tends to be stable, reaching 0.54, while the distribution of the temperature weighting function is more continuous and closer to that of the actual atmosphere. Thus, in terms of the ARI, convergence, and the distribution of the temperature weighting function, ICS is better than PCS.

(2) Statistical inversion comparison experiments show that the

retrieval temperature of ICS in this paper is consistent with that of NCS. In particular, from 10 hPa to 0.02 hPa (the stratosphere and mesosphere), the retrieval temperature of ICS is obviously better than that of NCS at about 1 K. In general, the accuracy of the retrieval temperature of ICS is improved. Especially, from 100 hPa to 0.01 hPa, the accuracy of ICS can be improved by more than 11%. The reason is that stratosphere and mesosphere are less affected by the ground surface, so the retrieval result of ICS is better than that of NCS. Additionally, due to the method selected in this paper there are more channels at 4.2 μm for the N_2O and at 4.3 μm for the CO_2 absorption bands; the channel combination of ICS is better than that of NCS for atmospheric temperature observation ~~to the~~at higher temperature.

(3) Statistical inversion comparison experiments in four typical regions indicate that ICS in this paper is significantly better than NCS and PCS in different regions and shows latitudinal variations, which shows potential for future applications.

Data availability. The data used in this paper are available from the corresponding author upon request.

Appendices

Appendix A

Table A1. Pressure levels adopted for RTTOV v12 54 pressure level coefficients and profile limits within which the transmittance calculations are valid. Note that the gas units here are ppmv. (From https://www.nwpsaf.eu/site/software/rttov/, RTTOV Users guide, 2019).

Level	Pressure	Tmax	Tmin	Qmax	Qmin	Q_2max	Q_2min	Q_2Ref
Number	hPa	K	K	ppmv*	ppmv*	ppmv*	ppmv*	ppmv*
1	0.01	245.95	143.66	5.24	0.91	1.404	0.014	0.296
2	0.01	252.13	154.19	6.03	1.08	1.410	0.069	0.321
3	0.03	263.71	168.42	7.42	1.35	1.496	0.108	0.361
4	0.03	280.12	180.18	8.10	1.58	1.670	0.171	0.527
5	0.13	299.05	194.48	8.44	1.80	2.064	0.228	0.769
6	0.23	318.64	206.21	8.59	1.99	2.365	0.355	1.074
7	0.41	336.24	205.66	8.58	2.49	2.718	0.553	1.471
8	0.67	342.08	197.17	8.34	3.01	3.565	0.731	1.991
9	1.08	340.84	189.50	8.07	3.30	5.333	0.716	2.787
10	1.67	334.68	179.27	7.89	3.20	7.314	0.643	3.756
11	2.50	322.5	17627	7.75	2.92	9.191	0.504	4.864
12	3.65	312.51	175.04	7.69	2.83	10.447	0.745	5.953
13	5.19	303.89	173.07	7.58	2.70	12.336	1.586	6.763
14	7.22	295.48	168.38	7.53	2.54	12.936	1.879	7.109
15	9.84	293.33	166.30	7.36	2.46	12.744	1.322	7.060
16	13.17	287.05	16347	7.20	2.42	11.960	0.719	6.574
17	17.33	283.36	161.49	6.96	2.20	11.105	0.428	5.687
18	22.46	280.93	161.47	6.75	1.71	9.796	0.278	4.705
19	28.69	282.67	162.09	6.46	1.52	8.736	0.164	3.870
20	36.17	27993	162.49	6.14	1.31	7.374	0.107	3.111

21	45.04	27315	164.66	5.90	1.36	6.799	0.055	2.478
22	55.44	265.93	166.19	6.21	1.30	5.710	0.048	1.907
23	67.51	264.7	167.42	9.17	1.16	4.786	0.043	1.440
24	81.37	261.95	159.98	17.89	0.36	4.390	0.038	1.020
25	97.15	262.43	163.95	20.30	0.01	3.619	0.016	0.733
26	114.94	259.57	168.59	33.56	0.01	2.977	0.016	0.604
27	134.83	259.26	169.71	102.24	0.01	2.665	0.016	0.489
28	156.88	260.13	169.42	285.00	0.01	2.351	0.013	0.388
29	181.14	262.27	17063	714.60	0.01	1.973	0.010	0.284
30	207.61	264.45	174.11	1464.00	0.01	1.481	0.013	0.196
31	236.28	270.09	177.12	2475.60	0.01	1.075	0.016	0.145
32	267.10	277.93	181.98	4381.20	0.01	0.774	0.015	0.110
33	300.00	285.18	184.76	6631.20	0.01	0.628	0.015	0.086
34	334.86	293.68	187.69	9450.00	1.29	0.550	0.016	0.073
35	371.55	300.12	190.34	12432.00	1.52	0.447	0.015	0.063
36	409.89	302.63	194.40	15468.00	2.12	0.361	0.015	0.057
37	449.67	304.43	198.46	18564.00	2.36	0.284	0.015	0.054
38	490.&5	307.2	201.53	21684.00	2.91	0.247	0.015	0.052
39	532.56	31217	202.74	24696.00	3.67	0.199	0.015	0.050
40	572.15	31556	201.61	27480.00	3.81	0.191	0.012	0.050
41	618.07	318.26	189.95	30288.00	6.82	0.171	0.010	0.049
42	661.00	321.71	189.95	32796.00	6.07	0.128	0.009	0.048
43	703.59	327.95	189.95	55328.00	6.73	0.124	0.009	0.047
44	745.48	333.77	189.95	37692.00	8.71	0.117	0.009	0.046
45	786.33	336.46	189.95	39984.00	8.26	0.115	0.008	0.045
46	825.75	338.54	189.95	42192.00	7.87	0.113	0.008	0.043
47	863.40	342.55	189.95	44220.00	7.53	0.111	0.007	0.041
48	898.93	346.23	189.95	46272.00	7.23	0.108	0.006	0.040
49	931.99	34924	189.95	47736.00	6.97	0.102	0.006	0.038
50	962.26	349.92	189.95	51264.00	6.75	0.099	0.006	0.034

51	989.45	350.09	189.95	49716.00	6.57	0.099	0.006	0.030
52	1013.29	360.09	189.95	47208.00	6.41	0.094	0.006	0.028
53	1033.54	350.09	189.95	47806.00	6.29	0.094	0.006	0.027
54	1050.00	350.09	189.95	47640.00	6.19	0.094	0.006	0.027

Table A2. Pressure levels adopted for IFS-137 137 pressure levels (in hPa).

Level number	pressure hPa	Level number	pressure hPa	Level number	pressure hPa	Level number	pressure hPa	Level number	pressure hPa
1	0.02	31	12.8561	61	106.4153	91	424.019	121	934.7666
2	0.031	32	14.2377	62	112.0681	92	441.5395	122	943.1399
3	0.0467	33	15.7162	63	117.9714	93	459.6321	123	950.9082
4	0.0683	34	17.2945	64	124.1337	94	478.3096	124	958.1037
5	0.0975	35	18.9752	65	130.5637	95	497.5845	125	964.7584
6	0.1361	36	20.761	66	137.2703	96	517.4198	126	970.9046
7	0.1861	37	22.6543	67	144.2624	97	537.7195	127	976.5737
8	0.2499	38	24.6577	68	151.5493	98	558.343	128	981.7968
9	0.3299	39	26.7735	69	159.1403	99	579.1926	129	986.6036
10	0.4288	40	29.0039	70	167.045	100	600.1668	130	991.023
11	0.5496	41	31.3512	71	175.2731	101	621.1624	131	995.0824
12	0.6952	42	33.8174	72	183.8344	102	642.0764	132	998.8081
13	0.869	43	36.4047	73	192.7389	103	662.8084	133	1002.225
14	1.0742	44	39.1149	74	201.9969	104	683.262	134	1005.356
15	1.3143	45	41.9493	75	211.6186	105	703.3467	135	1008.224
16	1.5928	46	44.9082	76	221.6146	106	722.9795	136	1010.849
17	1.9134	47	47.9915	77	231.9954	107	742.0855	137	1013.25
18	2.2797	48	51.199	78	242.7719	108	760.5996		
19	2.6954	49	54.5299	79	253.9549	109	778.4661		
20	3.1642	50	57.9834	80	265.5556	110	795.6396		
21	3.6898	51	61.5607	81	277.5852	111	812.0847		
22	4.2759	52	65.2695	82	290.0548	112	827.7756		
23	4.9262	53	69.1187	83	302.9762	113	842.6959		
24	5.6441	54	73.1187	84	316.3607	114	856.8376		
25	6.4334	55	77.281	85	330.2202	115	870.2004		
26	7.2974	56	81.6182	86	344.5663	116	882.791		
27	8.2397	57	86.145	87	359.4111	117	894.6222		
28	9.2634	58	90.8774	88	374.7666	118	905.7116		
29	10.372	59	95.828	89	390.645	119	916.0815		

30	11.5685	60	101.0047	90	407.0583	120	925.7571

Author contributions. ZS contributed the central idea. SC, ZS and HD conceived the method, developed the retrieval algorithm and discussed the results. SC analyzed the data, prepared the figures and wrote the paper. WG contributed to refining the ideas, carrying out additional analyses. All co-authors reviewed the paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The study was supported by the National Key Research Program of China: Development of high-resolution data assimilation technology and atmospheric reanalysis data set in East Asia (Research on remote sensing telemetry data assimilation technology, Grant no. 2017YFC1501802). The study was also supported by the National Natural Science Foundation of China (Grant no. 41875045) and Hunan Provincial Innovation Foundation for Postgraduate (Grant no. CX2018B033 and no. CX2018B034).

References

Aires, F., Schmitt, M., Chedin, A., and Scott, N.: The “weighting smoothing” regularization of MLP for Jacobian stabilization, IEEE. T. Neural. Networks., 10, 1502-1510, https://doi.org/10.1109/72.809096, 1999.

Aires, F., Chédin, Alain., Scott, N. A., and Rossow, W. B.: A regularized neural net approach for retrieval of atmospheric and surface temperatures with the IASI instrument, J. Appl. Meteorol., 41,144-159, https://doi.org/10.1175/1520-0450(2002)041<0144:ARNNAF>2.0.CO;2, 2002.

Aumann, H. H.: Atmospheric infrared sounder on the earth observing system, Optl. Engr., 33, 776-784, https://doi.org/10.1117/12.159325, 1994.

Aumann, H. H., Chahine, M. T., Gautier, C., and Goldberg, M.: AIRS/AMSU/HSB on the Aqua mission: design, science objective, data products, and processing systems, IEEE. Trans. GRS., 41,253-264, http://dx.doi.org/10.1109/TGRS.2002.808356, 2003.

Brath, M., Fox, S., Eriksson, P., Harlow, R. C., Burgdorf, M., and Buehler, S. A.: Retrieval of an ice water path over the ocean from ISMAR and MARSS millimeter and submillimeter brightness temperatures, Atmos. Meas. Tech., 11, 611–632, https://doi.org/10.5194/amt-11-611-2018, 2018.

Chahine, M. I.: A general relaxation method for inverse solution of the full radiative transfer equation, J. Atmos. Sci., 29, 741-747, https://doi.org/10.1175/1520-0469(1972)029<0741:AGRMFI>2.0.CO;2, 1972.

Chang, K. W, L'Ecuyer, T. S., Kahn, B. H., and Natraj, V.: Information content of visible and midinfrared radiances for retrieving tropical ice cloud properties, J. Geophys. Res., 122, https://doi.org/10.1002/2016JD026357, 2017.

Chedin, A., Scott, N. A., Wahiche, C., and Moulinier, P.: The improved initialization inversion method: a high resolution physical method for temperature retrievals from satellites of the tiros-n series, J. Appl. Meteor., 24, 128-143, https://doi.org/10.1175/1520-0450(1985)024<0128:TIIIMA>2.0.CO;2, 1985.

Cyril, C., Alain, C., and Scott, N. A.: Airs channel selection for CO_2 and other trace-gas retrievals, Q. J. Roy. Meteor. Soc., 129, 2719-2740, https://doi.org/10.1256/qj.02.180, 2003.

Du, H. D., Huang, S. X., and Shi, H. Q.: Method and experiment of channel selection for high spectral resolution data, Acta. Physica. Sinica., 57, 7685-7692, 2008 .

Dudhia, A., Jay, V. L., and Rodgers, C. D.: Microwindow selection for high-spectral-resolution sounders, Appl. Opt. 41, 3665-3673, https://doi.org/10.1364/AO.41.003665, 2002.

Eresmaa, R. and McNally, A. P.: Diverse profile datasets from the

ECMWF 137-level short-range forecasts, Tech. rep., ECMWF, 2014.
Eyre, J. R., Andersson E., and McNally, A. P.: Direct use of satellite sounding radiances in numerical weather prediction, High Spectral Resolution Infrared Remote Sensing for Earth's Weather and Climate Studies, Springer, Berlin, Heidelberg, https://doi.org/10.1007/978-3-642-84599-4_25, 1993.
Fang, Z. Y.: The evolution of meteorological satellites and the insight from it, Adv. Meteorol. Sci. Technol., 4, 27-34, https://doi.org/ 10.3969/j.issn.2095-1973.2014.06.003, 2014.
Gong, J., Wu, D. L., and Eckermann, S. D.: Gravity wave variances and propagation derived from AIRS radiances, Atmos. Chem. Phys., 11, 11691-11738, https://doi.org/10.5194/acp-12-1701-2012, 2011.
He, M. Y., Du, H. D., Long, Z. Y., and Huang, S. X.: Selection of regularization parameters using an atmospheric retrievable index in a retrieval of atmospheric profile, Acta. Physica Sinica., 61, 024205-160, 2012.
Hoffmann, L. and Alexander, M. J.: Retrieval of stratospheric temperatures from atmospheric infrared sounder radiance measurements for gravity wave studies, J. Geophys. Res. Atm., 114, https://doi.org/10.1029/2008JD011241, 2009.

Huang, H. L., Li, J., Baggett, K., Smith, W. L., and Guan, L.: Evaluation of cloud-cleared radiances for numerical weather prediction and cloud-contaminated sounding applications, Atmospheric and Environmental Remote Sensing Data Processing and Utilization: Numerical Atmospheric Prediction and Environmental Monitoring, I. S. O. Photonics., https://doi.org/10.1117/12.613027, 2005.

Kuai, L., Natraj, V., Shia, R. L., Miller, C., and Yung, Y. L.: Channel selection using information content analysis: a case study of CO_2 retrieval from near infrared measurements. J. Q. S. Radiative. Transfer., 111, 1296-1304, https://doi.org/10.1016/j.jqsrt.2010.02.011, 2010.

Li, J., Wolf, W. W., Menzel, W. P., Paul, Menzel. W., Zhang, W. J., Huang, H. L., and Achtor, T. H.: Global soundings of the atmosphere from ATOVS measurements: the algorithm and validation, J. Appl. Meteor., 39, 1248-1268, https://doi.org/10.1175/1520-0450(2000)039<1248:GSOTAF>2.0.CO;2, 2000.

Li, J., Liu, C. Y., Huang, H. L., Schmit, T. J., Wu, X., Menzel, W. P., and Gurka, J. J.: Optimal cloud-clearing for AIRS radiances using MODIS, IEEE. Trans. GRS. , 43, 1266-1278, http://dx.doi.org/10.1109/tgrs.2005.847795, 2005.

Liu, Z. Q.: A regional ATOVS radiance-bias correction scheme for rediance assimilation, Acta. Meteorologica. Sinica., 65, 113-123, 2007.

Lupu, C., Gauthier, P., and Laroche, Stéphane.: Assessment of the impact of observations on analyses derived from observing system experiments, Mon. Weather. Rev., 140, 245-257, https://doi.org/10.1175/MWR-D-10-05010.1, 2012.

Menke, W.: Geophysical Data Analysis: Discrete Inverse Theory, Acad. Press., Columbia University, New York, https://doi.org/10.1016/B978-0-12-397160-9.00019-9, 1984.

Menzel, W. P., Schmit, T. J., Zhang, P. and Li, J.: Satellite-based atmospheric infrared sounder development and applications, Bull. Amer. Meteor. Soc., 99, 583–603, https://doi.org/10.1175/BAMS-D-16-0293.1, 2018.

Prunet, P., Thépaut J. N., and Cass, V.: The information content of clear sky IASI radiances and their potential for numerical weather prediction, Q. J. Roy. Meteor. Soc., 124, 211-241, https://doi.org/10.1002/qj.49712454510, 2010.

Xu, Q.: Measuring information content from observations for data assimilation: relative entropy versus shannon entropy difference, Tellus. A., 59, 198-209, https://doi.org/10.1111/j.1600-0870.2006.00222.x, 2007.

Rabier, F., Fourrié, N., and Chafäi, D.: Channel selection methods for infrared atmospheric sounding interferometer radiances, Q. J. Roy. Meteor. Soc., 128, 1011-1027, https://doi.org/10.1256/0035900021643638, 2010.

Richardson, M. and Stephens, G. L.: Information content of oco-2 oxygen a-band channels for retrieving marine liquid cloud properties, Atmospheric Measurement Techniques, 11, 1-19, https://doi.org/10.5194/amt-11-1515-2018, 2018.

Rodgers, C. D.: Information content and optimisation of high spectral resolution remote measurements, Adv. Spa. Research, 21, 136-147, https://doi.org/10.1016/S0273-1177(97)00915-0, 1996.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding, Inverse methods for atmospheric sounding, World Scientific, https://doi.org/10.1142/3171, 2000.

Saunders, R., Hocking, J., Turner, E., Rayer, P., Rundle, D., Brunel, P., Vidot, J., Roquet, P., Matricardi, M., Geer, A., Bormann, N., and Lupu, C.: An update on the RTTOV fast radiative transfer model (currently at version 12), Geosci. Model Dev., 11, 2717-2737, https://doi.org/10.5194/gmd-11-2717-2018, 2018.

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

Smith, W. L., Woolf, H. M., and Revercomb, H. E.: Linear simultaneous solution for temperature and absorbing constituent profiles from radiance spectra, Appl. Optics., 30, 1117, https://doi.org/10.1364/AO.30.001117, 1991.

Wakita, H., Tokura, Y., Furukawa, F., and Takigawa, M.: Study of the information content contained in remote sensing data of atmosphere, Acta. Physica. Sinica., 59, 683-691, 2010.

Wang, G., Lu, Q. F., Zhang, J. W., and Wang, H. Y.,.: Study on method and experiment of hyper-spectral atmospheric infrared sounder channel selection, Remote Sensing Technology and Application., 29, 795-802 , 2014.

Zhang, J. W., Wang, G., Zhang, H., Huang J., Chen J., and Wu, L. L.: Experiment on hyper-spectral atmospheric infrared sounder channel selection based on the cumulative effect coefficient of principal component, Journal of Nanjing Institute of meteorology, 1, 36-42, http://dx.doi.org/10.3969/j.issn.1674-7097.2011.01.005, 2011.

（c） The marked-up manuscript version

A channel selection method for hyperspectral atmospheric infrared sounders based on layering

Shujie Chang[1, 2,3], Zheng Sheng[1,2], Huadong Du[1,2], Wei Ge[1,2] and Wei Zhang[1,2]

[1] College of Meteorology and Oceanography, National University of Defense Technology, Nanjing, China

[2] Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science and Technology, Nanjing, China

[3] South China Sea Institute for Marine Meteorology, Guangdong Ocean University, Zhanjiang, China

Correspondence: Zheng Sheng (19994035@sina.com)

Abstract. This study introduces an effective channel selection method for hyperspectral infrared sounders. The method is illustrated for the Atmospheric InfraRed Sounder (AIRS) instrument. The results are as follows: (1) Using the improved channel selection (ICS), the atmospheric retrievable index is more stable, the value reaching 0.54. The coverage of the weighting functions is more

evenly distributed over height with this method; (2) Statistical inversion comparison experiments show that the accuracy of the retrieval temperature, using the improved channel selection method in this paper, is consistent with that of 1Dvar channel selection. In the stratosphere and mesosphere especially, from 10 hPa to 0.02 hPa, the accuracy of the retrieval temperature of our improved channel selection method is improved by about 1 K. Also at lower heights, the accuracy of the retrieval temperature of ICS is improved; (3) Statistical inversion comparison experiments for four different regions illustrate latitudinal and seasonal variations and better performance of ICS compared to the NWP Channel Selection (NCS) and Primary Channel Selection (PCS) methods. The ICS method shows potential for future applications.

1 Introduction

Since the successful launch of the first meteorological satellite, TIROS in the 1960s, satellite observation technology has developed rapidly. Meteorological satellites observe the Earth's atmosphere from space and are able to record data from regions which are otherwise difficult to observe. Satellite data greatly enrich the content and range of meteorological observations, and consequently, atmospheric exploration technology and meteorological observations have taken us to a new stage in our understanding of weather systems and related phenomena (Fang, 2014). From the perspective of vertical atmospheric observation, satellite instruments are developing rapidly. In their infancy, the traditional infrared measurement instruments for detecting atmospheric temperature and moisture profiles, such as TOVS (Smith et al., 1991) or HIRS in ATOVS (Chahine, 1972; Li et al., 2000; Liu, 2007), usually employed filter spectrometry. Even though such instruments have played an important role in improving weather prediction, it is difficult to continue to build upon improvements in terms of observation accuracy and vertical resolution due to the limitation of low spectral resolution. By using this kind of filter-based spectroscopic measurement instrument, therefore, it is difficult to meet today’s needs in numerical weather prediction (Eyre et al., 1993; Prunet et al., 2010; Menzel et al., 2018). To meet this challenge, a series of plans for the creation of high-spectral resolution atmospheric measurement instruments has been executed in the United States and in Europe in recent years: One example is the AIRS (Atmospheric InfraRed Sounder) on the Earth Observation System, “Aqua”, launched on May 4, 2002 from the United States. AIRS has 2378 spectral channels providing sensitivity from the ground to up to about 65 km of altitude (Aumann et al., 2003;

Hoffmann and Alexander, 2009; Gong et al., 2011). The United States and Europe, in 2010 and in 2007, also installed the CRIS (Cross-track Infrared Sounder) and the IASI (Infrared Atmospheric Sounding Interferometer) on polar-orbiting satellites.

China also devotes great importance to the development of such advanced sounding technologies. In the early 1990s, the National Satellite Meteorological Center began to investigate the principles and techniques of hyperspectral resolution atmospheric observations. China's development of interferometric atmospheric vertical detectors eventually led to the launch of Fengyun No. 3, on May 27, 2008, and Fengyun No. 4 on December 11, 2016, both of which were equipped with infrared atmospheric instruments. How best to use the hyperspectral resolution observation data obtained from these instruments, to obtain reliable atmospheric temperature and humidity profiles, is an active area of study in atmospheric inversion theory.

Due to technical limitations, only a limited number of channels could at first be built into the typical satellite instruments. In this case, channel selection generally involved controlling the channel weighting function by utilizing the spectral response characteristics of the channel (such as center frequency and bandwidth). With the development of measurement technology, increasing numbers of hyperspectral detectors were carried on meteorological satellites. Due to the large number of channels and data supported by such instruments today (such as AIRS with 2378 channels and IASI with 8461 channels), it has proven extremely cumbersome to store, transmit, and process such data. Moreover, there is often a close correlation between the channel, causing an ill-posedness of the inversion, potentially compromising accuracy of the retrieval product based on hyperspectral resolution data.

However, hyperspectral detectors have many channels and provide real-time mode prediction systems with vast quantities of data, which can significantly improve prediction accuracy. But, if all the channels are used to retrieve data, the retrieval time considerably increases. Even more problematic are the glut of information produced, and the unsuitability of the calculations for real-time forecasting. Concurrently, the computer processing power must be large enough to meet the demands of simulating all the channels simultaneously within the forecast time. In order to improve the calculation efficiency and retrieval quality, it is very important to properly select a set of channels that can provide as much information as possible.

Many researchers have studied channel selection algorithms. Menke (1984) first chose channels using a data precision matrix method.

Aires et al. (1999) made the selection using the Jacobian matrix, which has been widely used since then (Aires et al., 2002; Rabier et al., 2010). Rodgers (2000) indicated that there are two useful quantities in measuring the information provided by the observation data: Shannon information content and degrees of freedom. The concept of information capacity then became widely used in satellite channel selection. In 2007, Xu (2007) compared the Shannon information content with the relative entropy, analyzing the information loss and information redundancy. In 2008, Du et al. (2008) introduced the concept of the atmospheric retrievable index (ARI) as a criterion for channel selection, and in 2010, Wakita et al. (2010) produced a scheme for calculating the information content of the various atmospheric parameters in remote sensing using Bayesian estimation theory. Kuai et al. (2010) analyzed both the Shannon information content and degrees of freedom in channel selection when retrieving CO_2 concentrations using thermal infrared remote sensing and indicated that 40 channels could contain 75% of the information from the total channels. Cyril et al. (2003) proposed the optimal sensitivity profile method based on the sensitivity of different atmospheric components. Lupu et al. (2012) used degrees of freedom for signals (DFS) to estimate the amount of information contained in observations in the context of observing system experiments. In addition, the singular value decomposition method has also been widely used for channel selection (Prunet et al., 2010; Zhang et al., 2011; Wang et al., 2014). In 2017, Chang et al. (2017) selected a new set of Infrared Atmospheric Sounding Interferometer (IASI) channels using the channel score index (CSI). Richardson et al. (2018) selected 75 from 853 channels based on the high spectral-resolution oxygen A-band instrument on NASA's Orbiting Carbon Observatory-2 (OCO-2), using information content analysis to retrieve the cloud optical depth, cloud properties, and position.

Today's main methods for channel selection use only the weighting function to study appropriate numerical methods, such as the data precision matrix method (Menke, 1984), singular value decomposition method (Prunet et al., 2010; Zhang et al., 2011; Wang et al., 2014), and the Jacobi method (Aires et al., 1999; Rabier et al., 2010). The use of the methods allows sensitive channels to be selected. The above-mentioned studies also take into account the sensitivity of each channel to atmospheric parameters during channel selection, while ignoring some factors that impact retrieval results. The accuracy of retrieval results depends not only on the channel weighting function but also on the channel noise, background field, and the retrieval algorithm.

Channel selection mostly uses the information content and delivers the largest amount of information for the selected channel combination during the retrieval (Rodgers, 1996; Du et al., 2008; He et al., 2012; Richardson et al., 2018).

This method has made great breakthroughs in both theory and practice, and the concept of information content itself does consider all the height dependencies of the kernel matrix K (Rodgers, 2000). However, earlier works have neglected the height dependencies of K for simplicity. This paper uses the atmospheric retrievable index (ARI) as the index, which is based on information content (Du et al., 2008; Richardson et al. 2018). Channel selection is made at different heights, and an effective channel selection scheme is proposed which fully considers various factors, including the influence of different channels on the retrieval results at different heights. This ensures the best accuracy of the retrieval product when using the selected channel. In addition, statistical inversion comparison experiments are used to verify the effectiveness of the method.

2 Channel selection indicator, scheme and method

2.1 Channel selection indicator

According to the concept of information content, the information content contained in a selected channel of a hyperspectral instrument can be described as H (Rodgers, 1996; Rabier et al., 2010). The final expression of H is:

$$\mathrm{H} = -\frac{1}{2} ln\left|\hat{S}S_a^{-1}\right|$$

$$= -\frac{1}{2} ln|(S_a - S_aK^T(KS_aK^T + S_\varepsilon)^{-1}KS_a)S_a^{-1}|, \quad (1)$$

where S_a is the error covariance matrix of the background or the estimated value of atmospheric profile, S_ε represents the observation error covariance matrix of each hyperspectral detector channel, $\hat{S} = (S_a - S_aK^T(KS_aK^T + S_\varepsilon)^{-1}KS_a)$ denotes the covariance matrix after retrieval, K is the weighting function matrix.

In order to describe the accuracy of the retrieval results visually and quantitatively, the atmospheric retrievable index (ARI), p, (Du et al., 2008) is defined as follows:

$$\mathrm{p} = 1 - \exp(\frac{1}{2n} ln\left|\hat{S}S_a^{-1}\right|), \quad (2)$$

Assuming that before and after the retrieval, the ratio of the root mean square error of each element in the atmospheric state vector is 1-p, then $\left|\hat{S}S_a^{-1}\right| = (1 - \mathrm{p})^{2n}$ is derived. By inverting the equation, the ARI that is p can be obtained in Eq. (2), which indicates the relative portion of the error that is eliminated by retrieval. In fact, before and after retrieval, the ratio of the root mean square error of each element cannot be 1-p. Therefore, p defined by Eq. (1) is actually an overall evaluation of the retrieval result.

2.2 Channel selection scheme

The principle of channel selection is to find the optimum channel combination after numbering the channels. This combination makes the information content, H, or the ARI defined in this paper as large as possible, in order to maintain the highest possible accuracy in the retrieval results.

Let there be M layers in the vertical direction of the atmosphere and N satellite channels. Selecting n from N channels, there will be C_N^n combinations in each layer, leading C_N^n calculations to get C_N^n kinds of p results. Furthermore, there are M layers in the vertical direction of the atmosphere. Therefore, the entire atmosphere must be calculated $\mathrm{M} \cdot \mathrm{C}_\mathrm{N}^\mathrm{n}$ times. However, the calculation $\mathrm{M} \cdot \mathrm{C}_N^n$ times will be particularly large, which makes this approach impractical in calculating p for all possible combinations. Therefore, it is necessary to design an effective calculation scheme, and such a scheme, i.e., a channel selection method, using iteration is proposed, called the "sequential absorption method" (Dudhia et al., 2002; Du et al., 2008). The method's main function is to select ("absorb") channels one by one, taking the channel with the maximum value of p. Through n iterations, n channels can be selected as the final channel combination. The steps are as follows:

(I) The expression of information content in a single channel:

First, we use only one channel for retrieval. A row vector, k, in the weighting function matrix, K, is a weighting function corresponding to the channel. After observation in this channel, the error covariance matrix is:

$$\hat{S} = S_a - S_a k^T (s_\varepsilon + k S_a k^T)^{-1} k S_a. \tag{3}$$

It should be noted that $(s_\varepsilon + k S_a k^T)$ is a scalar value in Eq. (3), so Eq. (3) can be converted to:

$$\hat{S} = \left(I - \frac{S_a k^T k}{(s_\varepsilon + k S_a k^T)}\right) S_a = \left(I - \frac{(k S_a)^T k}{(s_\varepsilon + k (k S_a)^T)}\right) S_a. \tag{4}$$

Substituting Eq. (4) into Eq. (2) gives:

$$\mathrm{p} = 1 - \exp\left(\frac{1}{2n} ln\left(\left|I - \frac{(k S_a)^T k}{(s_\varepsilon + k (k S_a)^T)}\right|\right)\right). \tag{5}$$

(II) Simplification of Eq. (5) for calculating the p value:

Since S_a and S_ε are positive definite symmetric matrices, they can be decomposed into $S_a = (S_a^{1/2})^T (S_a^{1/2})$ and $S_\varepsilon = (S_\varepsilon^{1/2})^T (S_\varepsilon^{1/2})$.

Define $\mathrm{R} = S_\varepsilon^{1/2} K S_a^{1/2}$. (6)

The matrix R can then be regarded as a weighting function matrix, normalized by the observed error and a priori uncertainty. A row vector of R, $\mathrm{r} = s_{\varepsilon}^{-1/2} k S_a^{1/2}$, represents the normalized weighting function matrix of a single channel. Substituting r into Eq. (5) gives:

$$\mathrm{p} = 1 - \exp\left(\frac{1}{2n} ln\left(\left|I - \frac{rr^T}{1+r^T r}\right|\right)\right). \tag{7}$$

For arbitrary row vectors, a and b, using the matrix property $\det(\mathrm{I} + a^T \mathrm{b}) = 1 + \mathrm{b}a^T$, the new expression for p is:

$$\begin{aligned}\mathrm{p} &= 1 - \exp\left(\frac{1}{2n} ln\left(1 - \frac{r^T r}{1 + r^T r}\right)\right)\\ &= 1 - \exp\left(\frac{1}{2n} ln\left(\frac{1}{1+r^T r}\right)\right)\\ &= 1 - \exp\left(-\frac{1}{2n} ln(1 + r^T r)\right).\end{aligned} \tag{8}$$

(III) Iteration in a single layer:

First, the iteration in a single layer requires the calculation of R. Using $\mathrm{S_a}$, $\mathrm{S_\varepsilon}$, K and Eq. (6), R can be calculated. Second, using Eq. (8), p of each candidate channel can be calculated. Moreover, the channel corresponding to maximum p is the selected channel for this iteration. After a channel has been selected, according to Eq. (3) we can use $\hat{S}$ to get S_a for the next iteration. Finally, channels which are not selected during this iteration are used as the candidate channels for the next iteration.

When selecting n from N channels, it is necessary to calculate (N-n/2)n≈Nn p values, which is much smaller than C_N^n. In addition to high computational efficiency by using this method, another advantage is that all channels can be recorded in the order in which they are selected. In the actual application, if n' channels are needed, and $n' < n$, we will not need to select the channel again, but record the selected channel only.

(IV) Iteration for different altitudes:

Because satellite channel sensitivity varies with height, repeating the iterative process of step (III), selects the optimum channels at different heights. Assuming there are M layers in the atmosphere and selecting n from N channels, it is necessary to calculate $M \cdot (N - n/2)n \approx M \cdot Nn$ p values, a much smaller number than $M \cdot C_N^n$. In this way, different channel sets can be used to evaluate corresponding height in the retrieved profiles.

2.3 Statistical inversion method

The inversion methods for the atmospheric temperature profiles can be summarized in two categories: statistical inversion and physical inversion. Statistical inversion is essentially a linear regression model which uses a large number of satellite measurements and atmospheric parameters to match samples and calculate their correlation coefficient. Then, based on the correlation coefficient, the required parameters of the independent measurements obtained by the satellite are retrieved. Because the method does not directly solve the radiation transfer equation, it has the advantages of fast calculation speed. In addition, the solution is numerically stable, which makes it one of the highest precision methods (Chedin et al., 1985). Therefore, the statistical inversion method will be used for our channel selection experiment and a regression equation will be established.

According to an empirical orthogonal function, the atmospheric temperature (or humidity), T, and the brightness temperature, T_b, are expanded as:

$$\mathrm{T} = T^* \cdot A, \tag{9}$$

$$T_b = T_b^* \cdot A, \tag{10}$$

where T^* and T_b^* are the eigenvectors of the covariance matrix of temperature (or humidity) and brightness temperature, respectively.

A and B stand for the corresponding expansion coefficient vectors of temperature (humidity) and brightness temperature.

Using the least squares method and the orthogonal property, the coefficient conversion matrix, V, is introduced:

$$\mathrm{A} = \mathrm{V} \cdot B, \tag{11}$$

where $V = AB^T(BB^T)^{-1}$. (12)

Using the orthogonality, we get:

$$\mathrm{B} = (T_b^*)^T T_b, \tag{13}$$

$$\mathrm{A} = (T^*)^T T. \tag{14}$$

For convenience, the anomalies of the state vector (atmospheric temperature), T, and the observation vector (brightness temperature), T_b, are taken:

$$\widehat{\mathrm{T}} = \overline{\mathrm{T}} + \widehat{\mathrm{T}}' = \overline{\mathrm{T}} + \mathrm{G}\mathrm{T}_{\mathrm{b}}' = \overline{\mathrm{T}} + \mathrm{G}(\mathrm{T}_{\mathrm{b}} - \overline{\mathrm{T}_{\mathrm{b}}}), \tag{15}$$

where $\widehat{\mathrm{T}}$ stands for the retrieval atmospheric temperature. $\overline{\mathrm{T}}$ and

$\overline{T_b}$ are the corresponding average values of the elements, respectively. $\hat{T}'$ and T_b' represent the corresponding anomalies of the elements, respectively.

Assuming there are k sets of observations, a sample anomaly matrix with k vectors can be constructed:

$$T' = (t_1', \ t_2', \ \cdots, \ t_k'), \tag{16}$$

$$T_b' = (t_{b1}', \ t_{b2}', \ \cdots, \ t_{bk}'). \tag{17}$$

Define the inversion error matrix as：

$$\delta = \overline{T} - \hat{T} = \hat{T}' - T'. \tag{18}$$

The retrieval error covariance matrix is:

$$\begin{aligned} S_\delta &= \frac{1}{k-n-1}\delta\delta^T \\ &= \frac{1}{k-n-1}(T' - GT_b')(T' - GT_b')^T \\ &= \frac{k-1}{k-n-1}(S_e - G^T S_{xy} - S_{xy}G^T + GS_yG^T), \end{aligned} \tag{19}$$

where

$$S_e = \frac{1}{k-1} \mathrm{T}' \, \mathrm{T}'^{\mathrm{T}} ,$$

$$S_y = \frac{1}{k-1} \mathrm{T}_b' \, \mathrm{T}_b'^{\mathrm{T}} ,$$

$$S_{xy} = \frac{1}{k-1} \mathrm{T}' \, \mathrm{T}_b'^{\mathrm{T}} . \qquad (20)$$

S_e stands for the sample covariance matrix of T, S_y denotes the sample covariance matrix of T_b, and S_{xy} represents the covariance matrix of T and T_b. The elements on the diagonal of the error covariance matrix, S_δ, represent the retrieval error variance of T. The matrix G that minimizes the overall error variance is the least squares coefficient matrix of the regression equation (15), which meets the criteria:

$$\delta^2 = \mathrm{tr}(S_\delta) = \min. \qquad (21)$$

Taking a derivative of Eq. (21) with respect to G, $\frac{\partial}{\partial G}\mathrm{tr}(S_\delta) = 0 = (-2S_{xy} + 2GS_y)$, which means that:

$$G = S_{xy} S_y^{-1}. \qquad (22)$$

Substituting Eq. (22) into Eq. (15) finally gives the least squares solution as:

$$\widehat{\mathrm{T}} = \overline{\mathrm{T}} + \mathrm{S_{xy}} \mathrm{S_y^{-1}} (\mathrm{T_b} - \overline{\mathrm{T_b}}). \tag{23}$$

It should be noted that the least squares solution obtained here aims to minimize the sum of the error variance for each element in the atmospheric state vector after retrieval for several times. At present, statistical multiple regression is widely used in the retrieval of atmospheric profiles based on atmospheric remote sensing data. As long as there are enough data, $\mathrm{S_{xy}}$ and $\mathrm{S_y}$ can be determined.

3. Channel selection experiment

3.1 Data and model

The Atmospheric Infrared Sounder (AIRS) is primarily designed to measure the Earth's atmospheric water vapor and temperature profiles on a global scale (Aumann et al., 2003; Susskind et al., 2003). AIRS is a continuously operating cross-track scanning sounder, consisting of a telescope that feeds an echelle spectrometer. The AIRS infrared spectrometer acquires 2378 spectral samples at a resolution λ/Δλ, ranging from 1086 to 1570, in three bands: 3.74 µm to 4.61 µm, 6.20 µm to 8.22 µm, and 8.8 µm to 15.4 µm. The footprint size is 13.5 km. The spectral range includes 4.3 µm and 15.5 µm for important temperature observation and CO_2, 6.3 µm for water vapor, and 9.6 µm for ozone absorption bands (Menzel et al., 2018). The root mean square error (RMSE) of the measured radiation is better than 0.2 K (Susskind et al., 2003). Moreover, global atmospheric profiles can be detected every day. Due to radiometer noise and faults, there are currently only 2047 effective channels. However, compared with previous infrared detectors, AIRS boasts a significant improvement in both the number of channels and spectral resolution (Aumann, 1994; Huang et al., 2005; Li et al., 2005).

The root mean square error of an AIRS infrared channel is shown in Fig. 1. The measurement error is not below 0.2K for all the instrument channels. There are a few channels with extremely large measurement errors, which reduce the accuracy of prediction to some extent. Among them, some extremely large measurement errors reduce the accuracy of prediction to some extent (Susskind et al., 2003). At present, more than 300 channels have not been used because their errors exceed 1 K. If data from these channels were to be used for retrieval, the accuracy of the retrieval could be reduced. Therefore, it is necessary to select a group of channels to improve the calculation efficiency and retrieval quality. In this paper we study channel selection for temperature profile retrieval by AIRS.

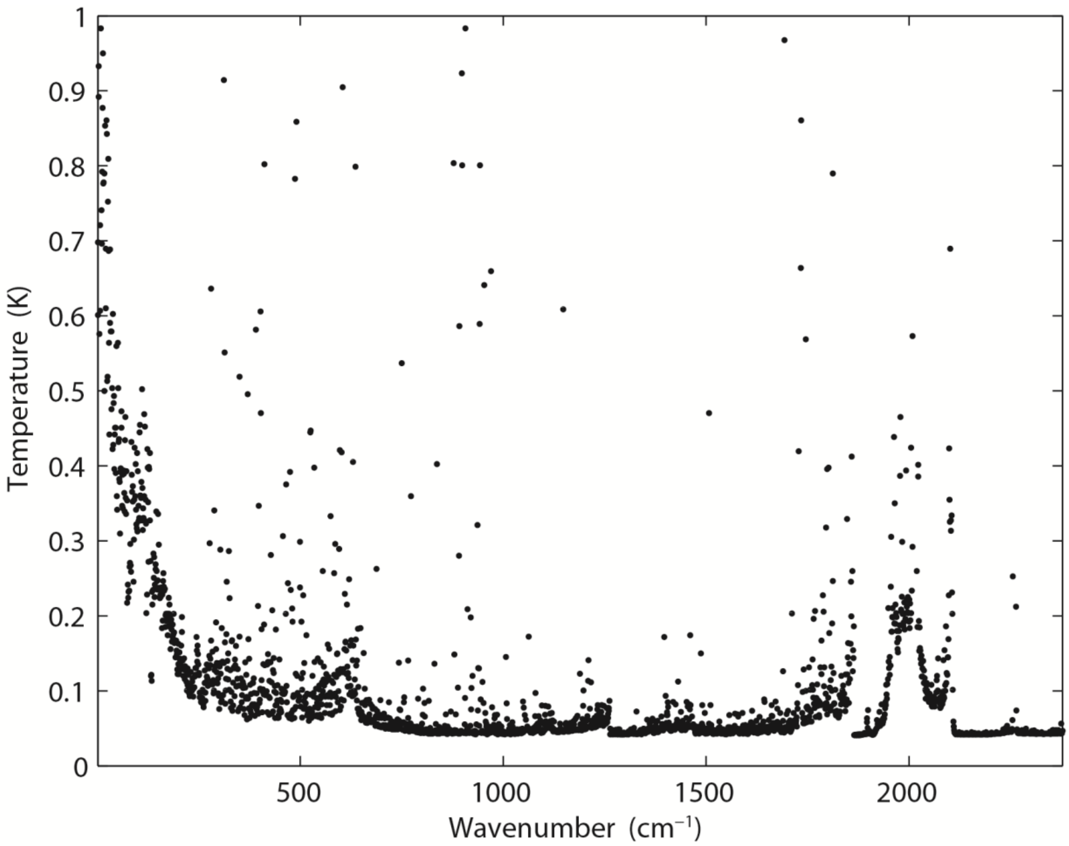


Figure 1. Root mean square error of AIRS infrared channel (black spots).

For the calculation of radiative transfer and the weighting function matrix, K, the RTTOV (Radiative Transfer for TOVS) v12 fast radiative transfer model is used. Although initially developed for the TOVS (TIROS Operational Vertical Sounder) radiometers, RTTOV can now simulate around 90 different satellite sensors measuring in the MW (microwave), IR (infrared) and VIS (visible) regions of the spectrum (Saunders et al., 2018). The model allows rapid simulations (1 ms for 40 channel ATOVS (Advanced TOVS) on a desktop PC) of radiances for satellite visible, infrared, or microwave nadir scanning radiometers given atmospheric profiles of temperature and trace gas concentrations, and cloud and surface properties. The only mandatory gas included as a variable for RTTOV v12 is water vapor. Optionally, ozone, carbon dioxide, nitrous oxide, methane, carbon monoxide, and sulfur dioxide can be included, with all other constituents assumed to be constant. RTTOV can accept input profiles on any defined set of pressure levels. The majority of RTTOV coefficient files are based on the 54 levels (see Table A1 in Appendix A), in the range from 1050 hPa to 0.01 hPa, though coefficients for some hyperspectral sounders are also available on 101 levels.

In order to correspond to the selected profiles, the atmosphere is divided into 137 layers, each of which contains corresponding atmospheric characteristics, such as temperature, pressure, and the humidity distribution. Each element in the weighting function matrix can be written as $\partial yi/\partial xj$. The subscript i is used to identify the satellite channel, and the subscript j is used to identify the atmospheric variable. Therefore, $\partial yi/\partial xj$ indicates the variation in brightness temperature in a given satellite channel, when a given atmospheric variable in a given layer changes. We are thus able to establish which layer of the satellite channel is particularly sensitive to which atmospheric characteristic (temperature, various gas contents) in the vertical atmosphere. The RTTOV_K (the K mode), is used to calculate the matrix H(X0) (Eq. (1)) for a given atmospheric profile characteristic.

3.2 Channel selection comparison experiment and results

In order to verify the effectiveness of the method, three sets of comparison experiments were conducted. First, 324 channels used by the EUMETSAT Satellite Application Facility on Numerical Weather Prediction (NWP SAF) were selected. NCS is short for NWP channel selection in this paper. NCS were released by the NWPSAF 1DVar (one-dimensional variational analysis) scheme, in accordance with the requirements of the NWPSAF (Saunders et al., 2018). Second, 324 channels were selected using the information capacity method. This method was adopted by Du et al. (2008) without the consideration of layering. PCS is short for primary channel selection in this paper.

Third, 324×M channels were selected using the information capacity method for the M layer atmosphere. ICS is short for improved channel selection in this paper. In order to verify the retrieval effectiveness after channel selection, statistical inversion comparison experiments were performed using 5000 temperature profiles provided by the ECMWF dataset, which will be introduced in Sect. 4.

The observation error covariance matrix, S_ε, in the experiment is provided by NWP SAF 1Dvar. In general, it can be converted to a diagonal matrix, the elements of which are the observation error standard deviation of each hyperspectral detector channel, which is the square of the root mean square error for each channel. The root mean square error of the AIRS channels is shown in Fig. 1. The error covariance matrix of the background, S_a, is calculated using 5000 samples of the IFS-137 data provided by the ECMWF dataset (The detailed information will be introduced in Sect. 4). The last access date is April 26th, 2019 (download address: https://www.nwpsaf.eu/site/update-137-level-nwp-profile-dataset/, 2019). The covariance matrix of temperature is shown in Fig. 2. The results are consistent with the previous study by Du et al. (2008).

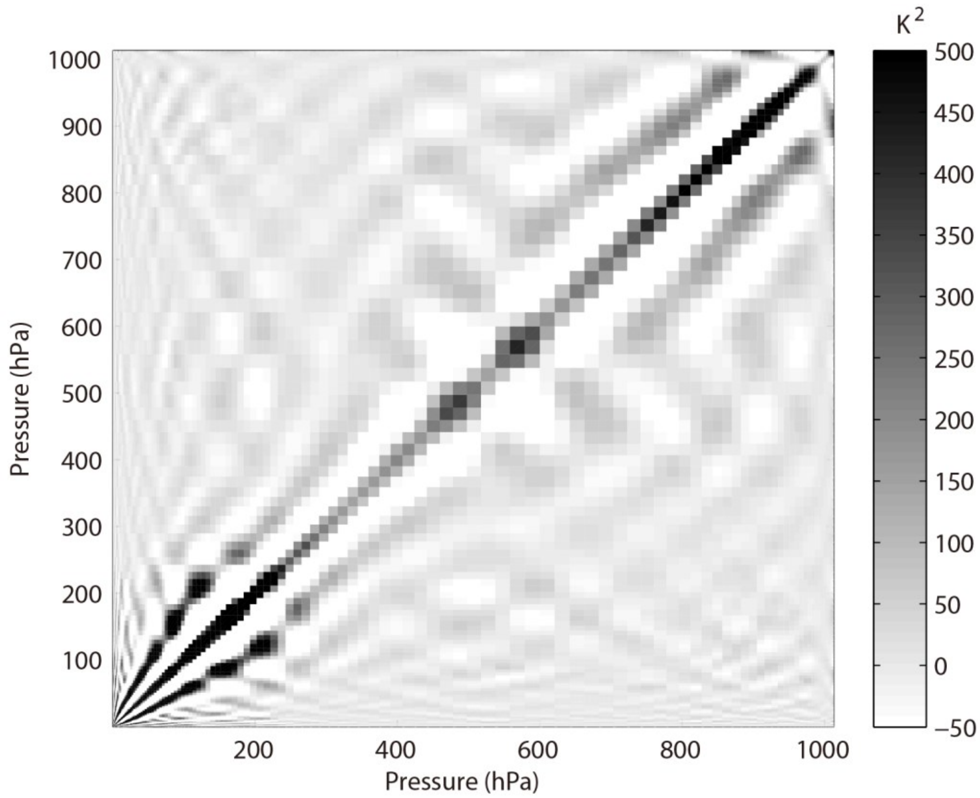


Figure 2. Error covariance matrix of temperature (shaded).

The reference atmospheric profiles are from the IFS-137 database, and the temperature weighting function matrix is calculated using the RTTOV_K mode, as shown in Fig. 3; the results are consistent with those of the previous study by Du et al. (2008). For the air-based passive atmospheric remote sensing studied in this paper, when the same channel detects the atmosphere from different observation angles, the value of the weighting function matrix K changes due to the limb effect. The goal of this section is focusing on the selection methods of selecting channels; therefore the biases produced from different observation angles can be ignored.

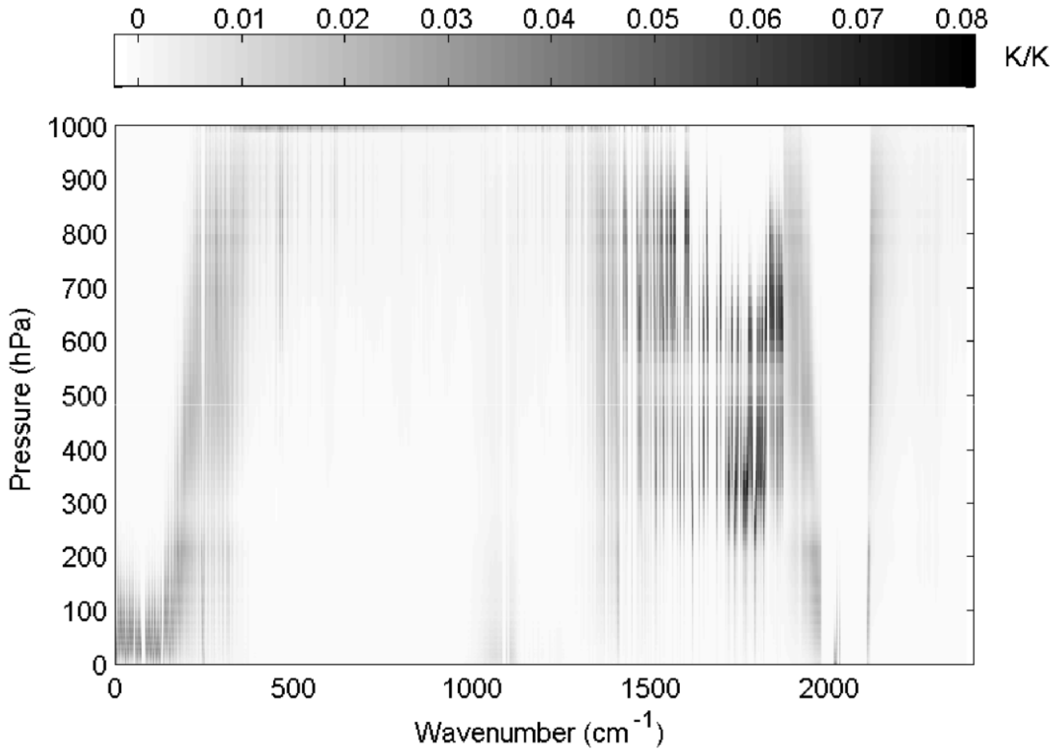


Figure 3. Temperature weighting function matrix (shaded).

In order to verify the effectiveness, the distribution of 324 channels, without considering layering, in the AIRS brightness temperature spectrum is indicated in Fig. 4. The background brightness temperature is the simulated AIRS observation brightness temperature, which is from the atmospheric profile in RTTOV put into the model. Figure 4(a) shows the 324 channels selected by PCS, while Fig. 4(b) shows the 324 channels selected by NCS.

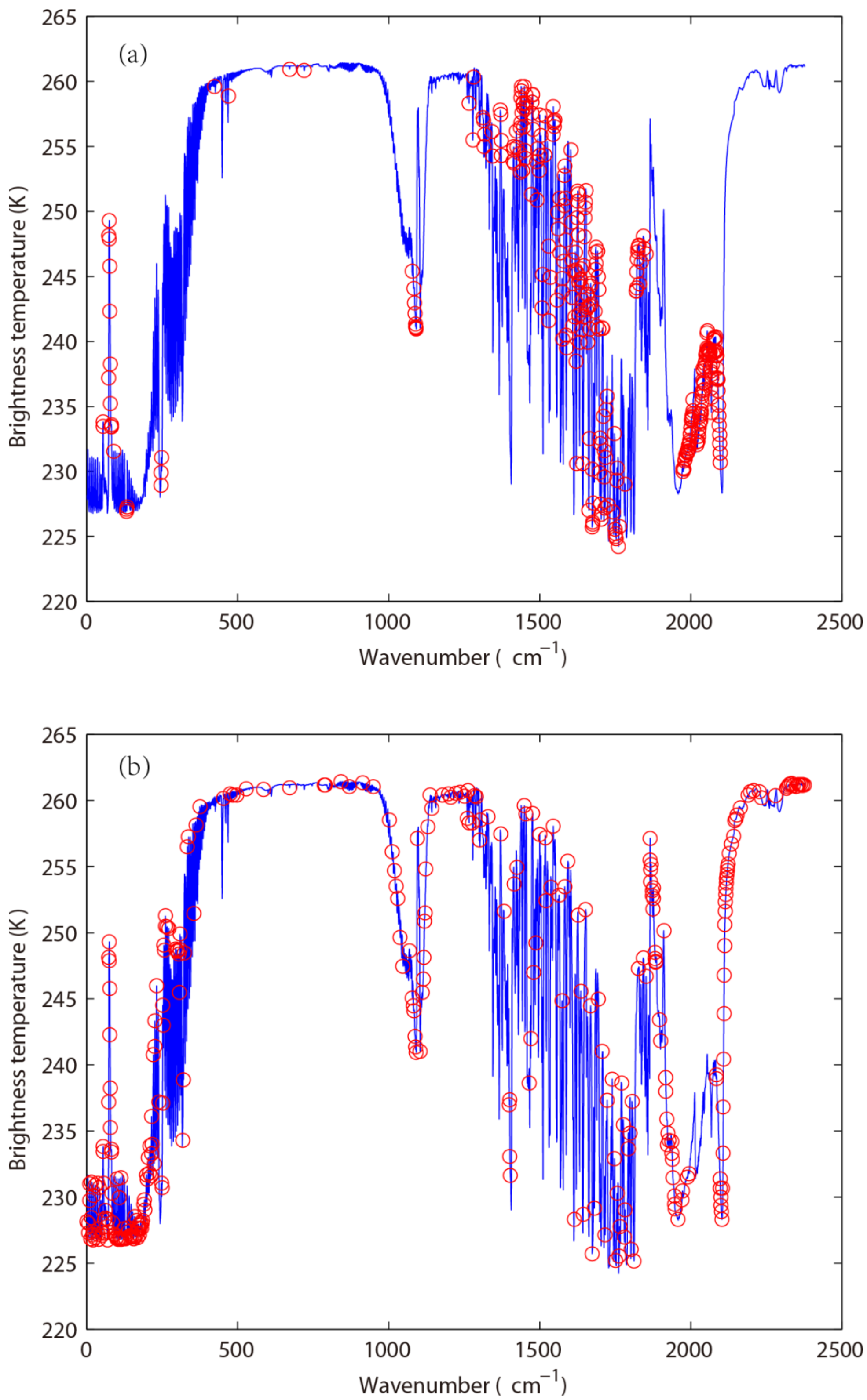


Figure 4. The distribution of different channel selection methods without considering layering in the AIRS brightness temperature

spectrum (blue line). (a) 324 channels selected by PCS (red circles). (b) 324 channels selected by NCS (red circles).

Without considering layering, the main differences between the 324 channels selected by PCS and NCS are as follows: (1) Near 10 μm band, fewer channels are selected by PCS because the retrieval of ground temperature is considered by NCS; (2) Near 9 μm band, no channels are selected by PCS because the retrieval of O3 is not considered in this paper; (3) As is known, the spectral range from 6 μm to 7 μm corresponds to water vapor absorption bands, but fewer channels are selected by NCS; (4) Near 5 μm band, it includes 4.2 μm for N_2O and 4.3 μm for CO_2 absorption bands. As is shown in Fig. 4, fewer channels are selected by PCS in those bands. PCS is favorable for atmospheric temperature observation. Because 4.2 μm and 4.3 μm bands are sensitive to high temperature, a better observation can be obtained for higher temperatures; (5) Near 4 μm band, a small number of channels is selected by NCS, but no channels are selected by PCS.

Above all, the information content considered in this study only takes the temperature profile retrieval into consideration, so the channel combination of PCS is inferior to that of NCS for the retrieval of surface temperature and the O_3 profile. The advantages of the channel selection method based on information content in this paper are mainly reflected in: (1) Stratosphere and mesosphere is less affected by the ground surface, so the retrieval result of PCS is better than that of NCS. (2) Due to the method selected in this paper there are more channels at 4.2 μm for N_2O and 4.3 μm for CO_2 absorption bands; the channel combination of PCS is better than that of NCS for atmospheric temperature observation at higher temperature.

By comparing channel selection without considering layering, we note the general advantages and disadvantages of PCS and NCS for the retrieval of temperature and can improve the channel selection scheme. First, the retrieval of the temperature profile for 324 channels selected by PCS is obtained. The relationship between the number of iterations and the ARI is shown in Fig. 5.

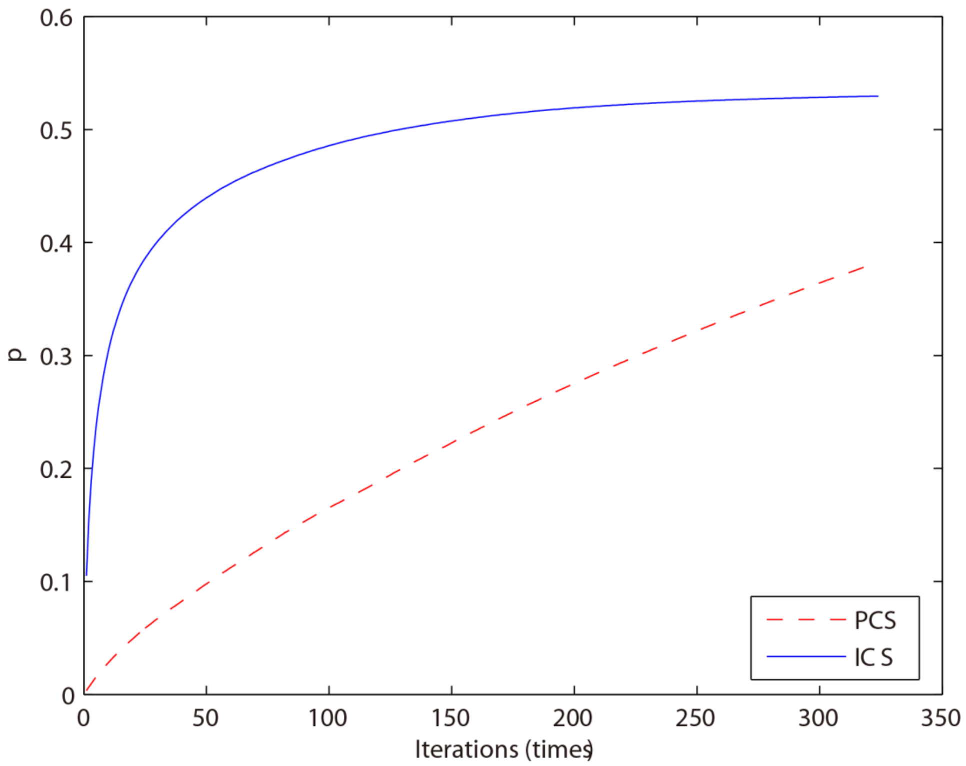


Figure 5. The relationship between the number of iterations and ARI. Blue line represents the result of ICS. Red dotted line stands for the result of PCS.

The ARI for PCS tends to be 0.38 and is not convergent, so the PCS method needs to be improved. In this paper, the atmosphere is divided into 137 layers, and based on the information content and iteration, 324 channels are selected for each layer. Then, the temperature profile of each layer can be retrieved based on statistical inversion (see at Sect. 4). The relationship between the number of iterations and the ARI for ICS is shown in Fig. 5b. When the number of iterations approaches 100, the ARI of ICS tends to be stable, and reaches 0.54. Thus, in terms of the ARI and convergence, the ICS method is better than that of PCS.

Furthermore, because an iterative method is used to select channels, the order of each selected channel is determined by the contribution from the ARI. The weighting function matrix of the top 324 selected channels, according to channel order, is shown in Fig. 6.

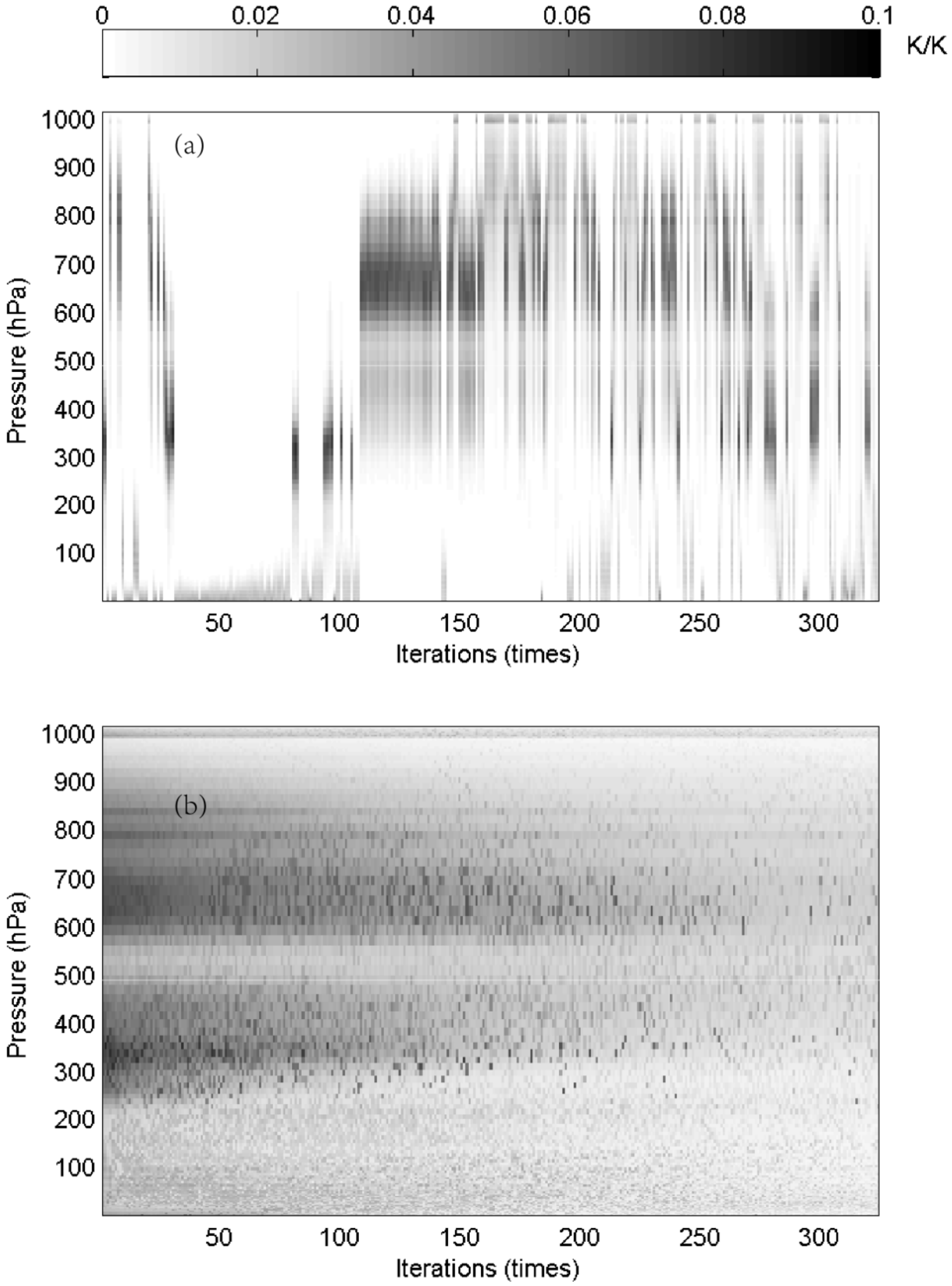


Figure 6. The relationship between the number of iterations and the weighting function of the top 324 selected channels (shaded). (a) ICS. (b) PCS.

As illustrated in Fig. 6, in the first 100 iterations, the distribution of the temperature weighting function for PCS is relatively scattered; it does not reflect continuity between the adjacent layers of the atmosphere. Besides, the ICS result is better than that of PCS, showing that: (1) the distribution of the temperature weighting function is more continuous and reflects the continuity between adjacent layers of the atmosphere; (2) regardless of the number of iterations, the maximum value of the weighting function is stable near 300–400 hPa and 600–700 hPa, without scattering, which is closer to the situation in real atmosphere.

4. Statistical multiple regression experiment

4.1 Temperature profile database

A new database including a representative collection of 25,000 atmospheric profiles from the European Centre for Medium-range Weather Forecasts (ECMWF) was used for the statistical inversion experiments. The profiles were given in a 137-level vertical grid extending from the surface up to 0.01 hPa. The database was divided into five subsets focusing on diverse sampling characteristics such as temperature, specific humidity, ozone mixing ratio, cloud condensates, and precipitation. In contrast with earlier releases of the

ECMWF diverse profile database, the 137-level database places greater emphasis on preserving the statistical properties of sampled distributions produced by the Integrated Forecasting System (IFS) (Eresmaa and McNally, 2014; Brath et al., 2018). IFS-137 spans the period from September 1, 2013 to August 31, 2014. There are two operational analyses each day (at 00z and 12z), and approximately 13 000 atmospheric profiles over the ocean. The pressure levels adopted for IFS-137 are shown in Table A2 (see Table A2 in Appendix A).

The locations of selected profiles of temperature, specific humidity, and cloud condensate subsets of the IFS-91 and IFS-137 databases are plotted on the map in Fig. 7. In the IFS-91 database, the sampling is fully determined by the selection algorithm, which makes the geographical distributions very inhomogeneous. Selected profiles represent those regions where gradients of the sampled variable are the strongest: in the case of temperature, mid- and high-latitudes dominate, while humidity and cloud condensate subsets concentrate at low latitudes. However, the IFS-137 database shows a much more homogeneous spatial distribution in all the sampling subsets, which is a consequence of the randomized selection.

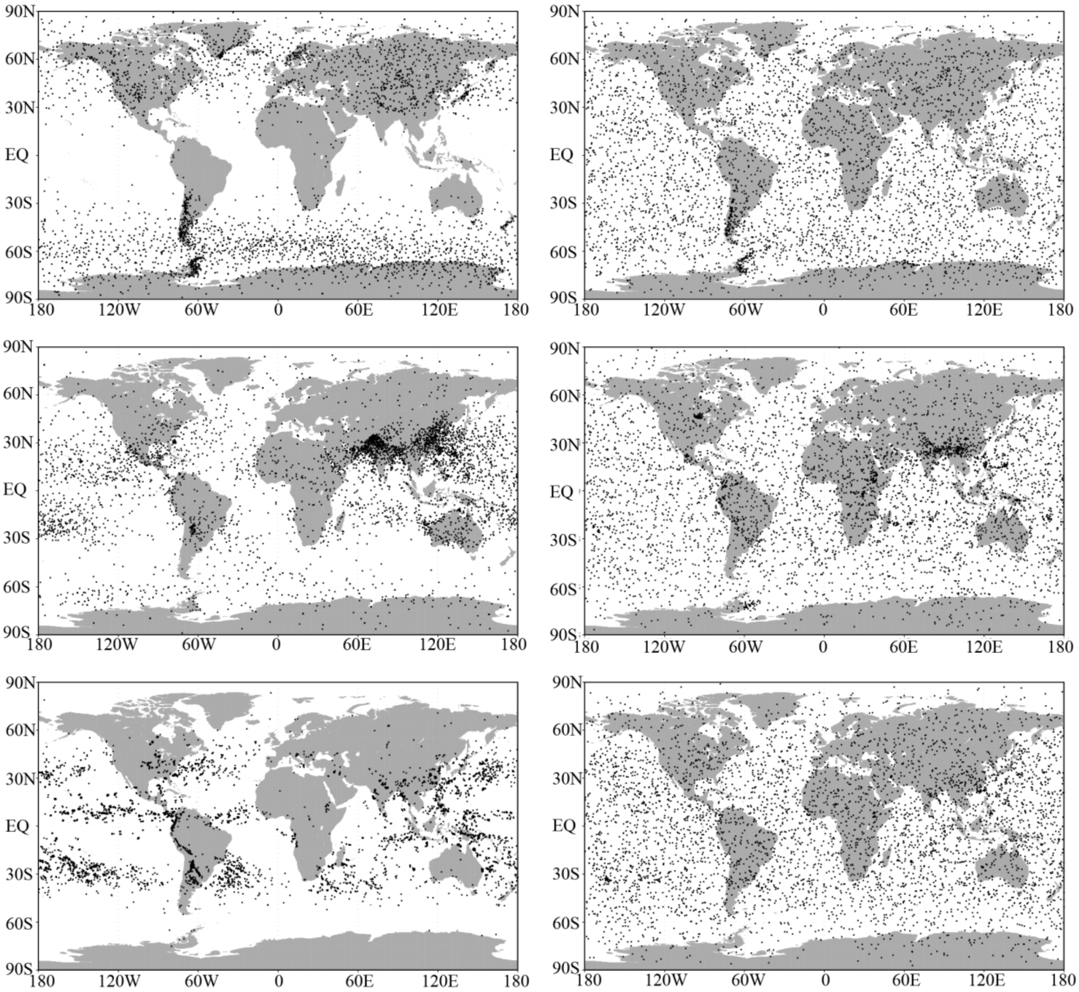


Figure 7. Locations of selected profiles in the temperature (top), specific humidity (middle), and cloud condensate (bottom), sampled subsets of the IFS-91 (left) and IFS-137 (right) databases (from https://www.nwpsaf.eu/site/update-137-level-nwp-profile-dataset/ , 2019).

The temporal distribution of the selected profiles is illustrated in Fig. 8. The coverage of the IFS-137 data set is more homogeneous than the IFS-91 data set. Moreover, the IFS-137 database supports the mode with input parameters, such as detection angle, 2 m temperature, and cloud information. Therefore, it is feasible to use the selected samples in a statistical multiple regression experiment.

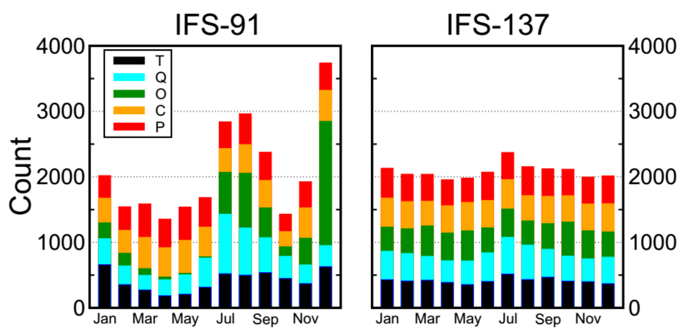


Figure 8. Distribution of profiles within the calendar months in IFS-91 (left) and IFS-137 (right) databases. Different subsets are shown in different colors.Black parts stand for temperature. Blue parts represent specific humidity. Green parts indicate ozone subset. Orange parts stand for cloud condensate. Red parts represent precipitation. The last access date is April 26th, 2019. (from https://www.nwpsaf.eu/site/update-137-level-nwp-profile-dataset/ , 2019).

4.2 Experimental scheme

In order to verify the retrieval effectiveness of ICS, 5000 temperature profiles provided by the IFS-137 were used for statistical inversion comparison experiments. The steps are as follows:

(1) 5000 profiles and their corresponding surface factors, including surface air pressure, surface temperature, 2 m temperature, 2 m specific humidity, 10 m wind speed are put into the RTTOV mode. Then, the simulated AIRS spectra are obtained.

(2) The retrieval of temperature is carried out in accordance with Eq. (23). The 5000 profiles are divided into two groups. The first group of 2500 profiles is used to obtain the regression coefficient, and the second group of 2500 is used to test the result.

(3) Verification of the results. The test is carried out based on the standard deviation between the retrieval value and the true value.

4.3 Results and Discussion

For the statistical inversion comparison experiments, the standard deviation of temperature retrieval is shown in Fig. 9. First, because PCS does not take channel sensitivity as a function of height into consideration, the retrieval result of PCS is inferior to that of ICS. Second, by comparing the results of ICS and NCS we found that below 100 hPa, since the method used in this paper considers near ground to be less of an influencing factor, the channel combination of ICS is slightly inferior to that of NCS, but the difference is small.

From 100 hPa to10 hPa, the retrieval temperature of ICS in this paper is consistent with that of NCS, slightly better than the channel selected for NCS. From 10 hPa to 0.02 hPa, near the space layer, the retrieval temperature of ICS is better than that of NCS. In terms of the standard deviation, the channel combination of ICS is slightly better than that of PCS from 100 hPa to 10 hPa. From 10 hPa to 0.02 hPa, the standard deviation of ICS is lower than that of NCS at about 1 K, meaning that the retrieval result of ICS is better than that of NCS.

In order to further illustrate the effectiveness of ICS, the mean improvement value of the ICS and its percentages compared with the PCS and NCS at different heights are shown in Table 1. Because PCS does not take channel sensitivity as a function of height into consideration, the retrieval result of PCS is inferior to that of ICS. In general, the accuracy of the retrieval temperature of ICS is improved. Especially, from 100 hPa to 0.01 hPa, the mean value of ICS is evidently improved by more than 0.5 K which means the accuracy can be improved by more than 11%. By comparing the results of ICS and NCS we found that below 100 hPa, since the method used in this paper considers near ground to be less of an influencing factor, the channel combination of ICS is slightly inferior to that of NCS, but the difference is small. From 100 hPa to 0.01 hPa, the mean value of ICS is improved by more than 0.36 K which means the accuracy can be improved by more than 9.6%.

Table 1. The mean improvement value of the ICS and its percentages compared with the PCS and NCS at different heights.

Pressure	Improved mean value /Percentage compared with PCS	Improved value /Percentage compared with NCS
hPa	K/%	K/%
surface-100hPa	0.24/10.77%	-0.04/-3.27%
100hPa-10hPa	0.15/5.08%	0.06/2.4%
10hPa-1hPa	0.04/0.64%	0.17/2.99%
1hPa-0.01hPa	0.52/11.92%	0.36/9.57%

This is because, as shown in Fig. 4: (1) Stratosphere and mesosphere is less affected by the ground surface, so the retrieval result of PCS is better than that of NCS. (2) Due to the method selected in this paper, there are more channels at 4.2 µm for N_2O and 4.3 µm for CO_2 absorption bands, and the channel combination of PCS is superior to that of NCS for atmospheric temperature observation in the high temperature zone. Moreover, ICS takes channel sensitivity as a function of height into consideration, so its retrieval result is improved.

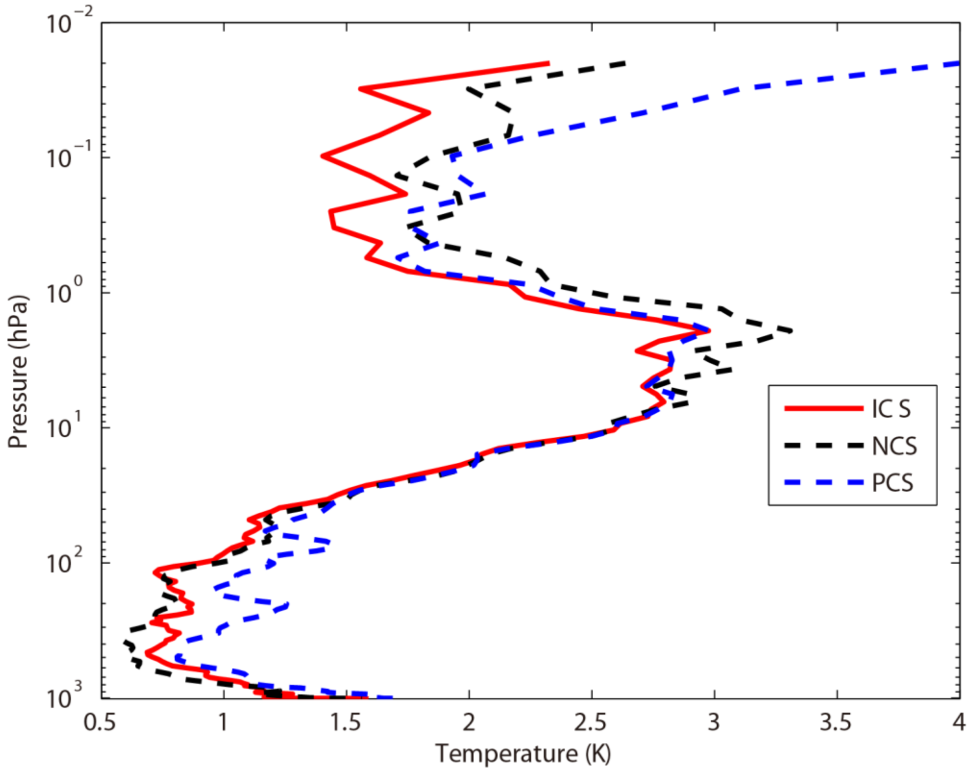


Figure 9. The temperature profile standard deviation of statistical inversion comparison experiments. Red line indicates the result of ICS. Black dotted line stands for the result of NCS. Blue dotted line represents the result of PCS.

5 Statistical inversion comparison experiments in four typical regions

The accuracy of the retrieval temperature varies from place to place and changes with atmospheric conditions. Therefore, in order to further compare the inversion accuracy under different atmospheric conditions, this paper has divided the atmospheric profile from the IFS-137 database introduced in Sect. 4 into four regions: equatorial zone, subtropical region, mid-latitude region and Arctic. The average temperature profiles in these four regions are shown in Fig. 10. The retrieval temperature varies from place to place and changes with atmospheric conditions. In order to further compare the regional differences of inversion accuracy, the temperature standard deviations of ICS in four typical regions are compared in Sect. 5.2.

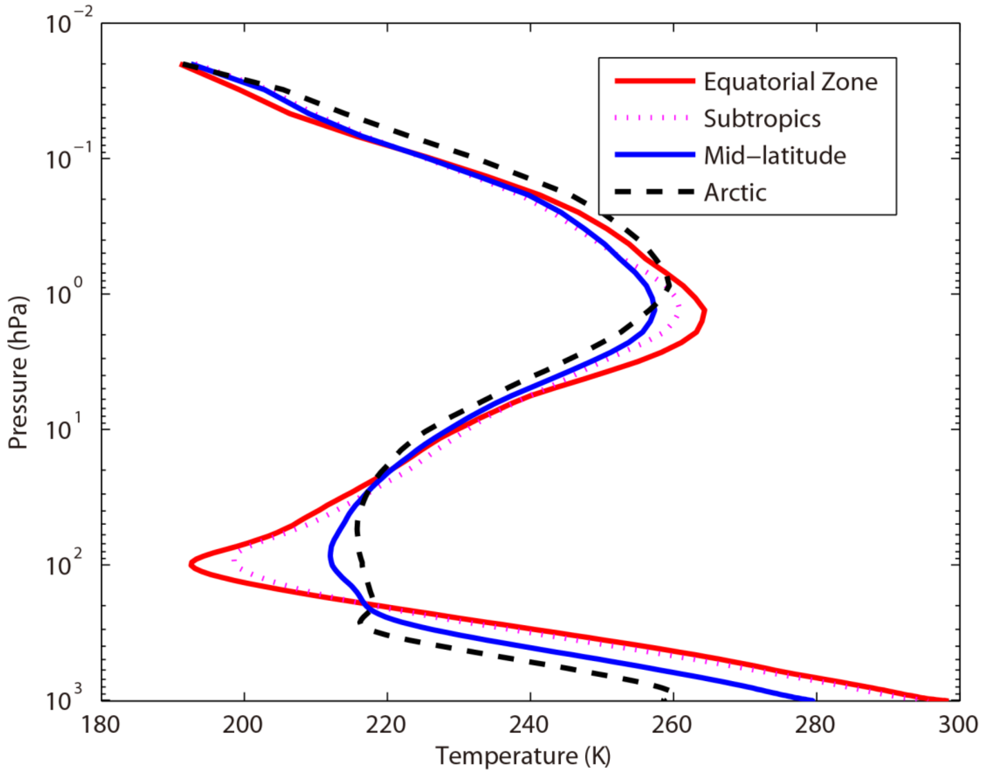


Figure 10. The average temperature profiles in four typical regions. Red line indicates the equatorial zone. Pink dotted line stands for the subtropics. Blue dotted line represents the mid-latitude region. Black dotted line stands for the Arctic.

5.1 Experimental scheme

In order to further illustrate the different accuracy of the retrieval temperature using our improved channel selection method under different atmospheric conditions, the profiles in four typical regions were used for statistical inversion comparison experiments. The experimental steps are as follows:

(1) 2500 profiles in Sect. 4 are used to work out the regression coefficient.

(2) The atmospheric profiles of the four typical regions: equatorial zone, subtropical region, mid-latitude region and Arctic are used for statistical inversion comparison experiments and test the result.(3) Verification of the results. The test is carried out based on the standard deviation between the retrieval value and the true value.

5.2 Results and Discussion

Using statistical inversion comparison experiments in four typical regions, the standard deviation of temperature retrieval is shown in Fig. 11. Generally, the retrieval temperature by ICS is better than that of NCS and PCS. In particular, above 1 hPa (the stratosphere and mesosphere), the standard deviation of atmospheric temperature can be improved by 1 K with PCS and NCS. Thus, ICS shows a great improvement. The results were consistent with Sect. 4.

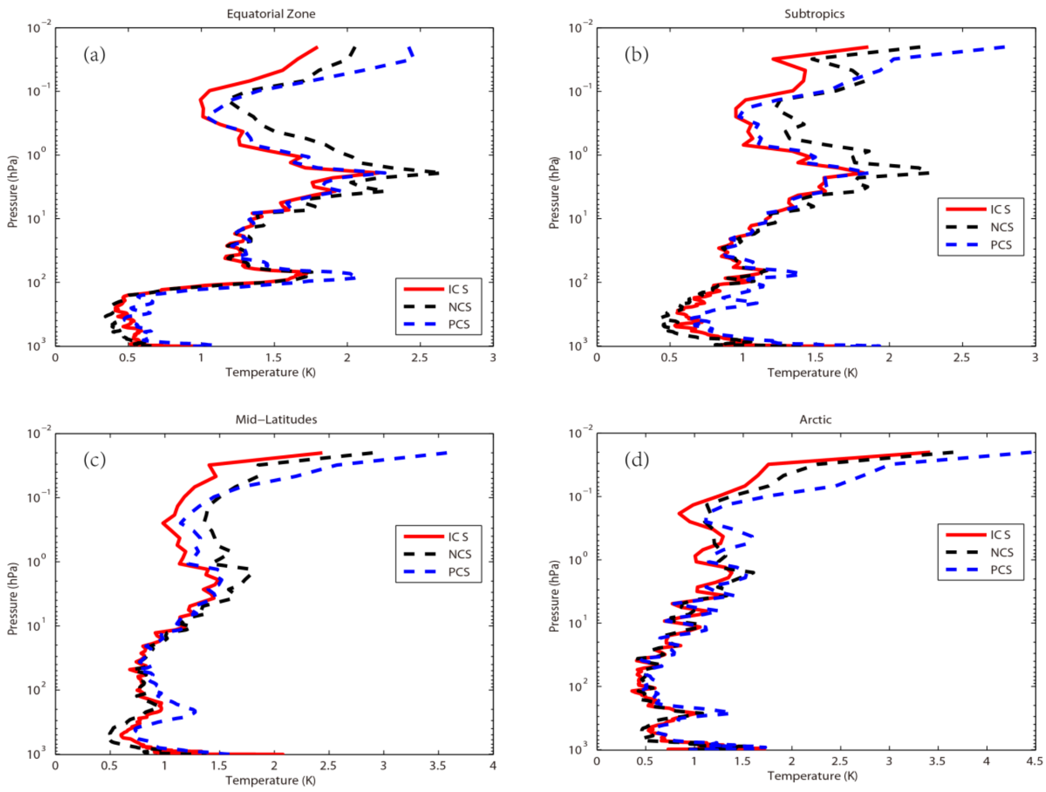


Figure 11. The temperature profile standard deviation of statistical inversion comparison experiments in four typical regions. Red line indicates the result of ICS. Black dotted line stands for the result of NCS. Blue dotted line represents the result of PCS. (a) Equatorial zone. (b) Subtropics. (c) Mid-latitudes. (d) Arctic.

In order to further compare the regional differences of inversion accuracy, the temperature standard deviation of ICS in four typical regions are compared in Fig. 12.

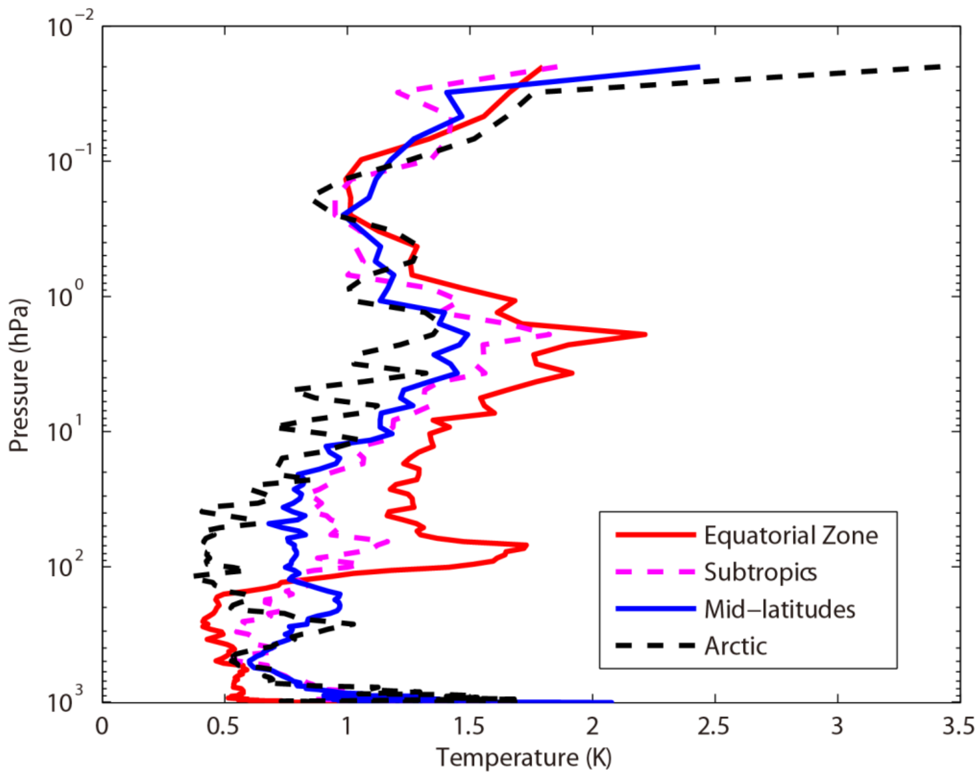


Figure 12. The temperature standard deviation of ICS in four typical regions. Red line indicates the result of equatorial zone. Pink dotted line represents the result of Subtropics. Blue line represents the result of Mid-latitudes. Black dotted line stands for the result of Arctic.

The temperature standard deviations of the ICS in the four typical regions are large (Fig. 12). Below100 hPa, due to the high temperature in the equatorial zone, the channel combination of ICS is better than that of PCS and NCS for atmospheric temperature observation at higher temperature. The standard deviation is 0.5K. Due to the method selected in this paper there are more channels at

4.2 μm for N_2O and 4.3 μm for CO_2 absorption bands which has been previously described in Sect. 3. Near the tropopause, the standard deviation of the equatorial zone increases sharply. It is also due to the sharp drops in temperature. However, the standard deviation of the Arctic is still around 0.5K. From 100hPa to 1hPa, the standard deviation of ICS is 0.5 K to 2K. With the increase of latitude, the effectiveness considerably increases. According to Fig. 11, ICS takes channel sensitivity as a function of height into consideration, so its retrieval result is better.

Although the improvements of ICS in the four typical regions are different, in general, the accuracy of the retrieval temperature of ICS is improved. Because PCS does not take channel sensitivity as a function of height into consideration, the retrieval result of PCS is inferior to that of ICS. In general, the accuracy of the retrieval temperature of ICS is improved.

7 Conclusions

In recent years, the atmospheric layer in the altitude range of about 20–100 km has been named "the near space layer" by aeronautical and astronautical communities. It is between the space-based satellite platform and the aerospace vehicle platform, which is the transition zone between aviation and aerospace. Its unique resource has attracted a lot of attention from many countries. Research and exploration, therefore, on and of the near space layer are of great importance. A new channel selection scheme and method for hyperspectral atmospheric infrared sounder AIRS data based on layering is proposed. The retrieval results of ICS concerning the near space atmosphere are particularly good. Thus, ICS aims to provide a new and an effective channel selection method for the study of the near space atmosphere using the hyperspectral atmospheric infrared sounder.

An improved channel selection method is proposed, based on information content in this paper. A robust channel selection scheme and method are proposed, and a series of channel selection comparison experiments are conducted. The results are as follows:

(1) Since ICS takes channel sensitivity as a function of height into consideration, the ARI of PCS only tends to be 0.38 and is not convergent. However, as the 100^{th} iteration is approached, the ARI of ICS tends to be stable, reaching 0.54, while the distribution of the temperature weighting function is more continuous and closer to that of the actual atmosphere. Thus, in terms of the ARI, convergence, and the distribution of the temperature weighting function, ICS is better than PCS.

(2) Statistical inversion comparison experiments show that the

retrieval temperature of ICS in this paper is consistent with that of NCS. In particular, from 10 hPa to 0.02 hPa (the stratosphere and mesosphere), the retrieval temperature of ICS is obviously better than that of NCS at about 1 K. In general, the accuracy of the retrieval temperature of ICS is improved. Especially, from 100 hPa to 0.01 hPa, the accuracy of ICS can be improved by more than 11%. The reason is that stratosphere and mesosphere are less affected by the ground surface, so the retrieval result of ICS is better than that of NCS. Additionally, due to the method selected in this paper there are more channels at 4.2 μm for the N_2O and at 4.3 μm for the CO_2 absorption bands; the channel combination of ICS is better than that of NCS for atmospheric temperature observation at higher temperature.

(3) Statistical inversion comparison experiments in four typical regions indicate that ICS in this paper is significantly better than NCS and PCS in different regions and shows latitudinal variations, which shows potential for future applications.

Data availability. The data used in this paper are available from the corresponding author upon request.

Appendices

Appendix A

Table A1. Pressure levels adopted for RTTOV v12 54 pressure level coefficients and profile limits within which the transmittance calculations are valid. Note that the gas units here are ppmv. (From https://www.nwpsaf.eu/site/software/rttov/, RTTOV Users guide, 2019).

Level	Pressure	Tmax	Tmin	Qmax	Qmin	Q_2max	Q_2min	Q_2Ref
Number	hPa	K	K	ppmv*	ppmv*	ppmv*	ppmv*	ppmv*
1	0.01	245.95	143.66	5.24	0.91	1.404	0.014	0.296
2	0.01	252.13	154.19	6.03	1.08	1.410	0.069	0.321
3	0.03	263.71	168.42	7.42	1.35	1.496	0.108	0.361
4	0.03	280.12	180.18	8.10	1.58	1.670	0.171	0.527
5	0.13	299.05	194.48	8.44	1.80	2.064	0.228	0.769
6	0.23	318.64	206.21	8.59	1.99	2.365	0.355	1.074
7	0.41	336.24	205.66	8.58	2.49	2.718	0.553	1.471
8	0.67	342.08	197.17	8.34	3.01	3.565	0.731	1.991
9	1.08	340.84	189.50	8.07	3.30	5.333	0.716	2.787
10	1.67	334.68	179.27	7.89	3.20	7.314	0.643	3.756
11	2.50	322.5	17627	7.75	2.92	9.191	0.504	4.864
12	3.65	312.51	175.04	7.69	2.83	10.447	0.745	5.953
13	5.19	303.89	173.07	7.58	2.70	12.336	1.586	6.763
14	7.22	295.48	168.38	7.53	2.54	12.936	1.879	7.109
15	9.84	293.33	166.30	7.36	2.46	12.744	1.322	7.060
16	13.17	287.05	16347	7.20	2.42	11.960	0.719	6.574
17	17.33	283.36	161.49	6.96	2.20	11.105	0.428	5.687
18	22.46	280.93	161.47	6.75	1.71	9.796	0.278	4.705
19	28.69	282.67	162.09	6.46	1.52	8.736	0.164	3.870
20	36.17	27993	162.49	6.14	1.31	7.374	0.107	3.111

21	45.04	27315	164.66	5.90	1.36	6.799	0.055	2.478
22	55.44	265.93	166.19	6.21	1.30	5.710	0.048	1.907
23	67.51	264.7	167.42	9.17	1.16	4.786	0.043	1.440
24	81.37	261.95	159.98	17.89	0.36	4.390	0.038	1.020
25	97.15	262.43	163.95	20.30	0.01	3.619	0.016	0.733
26	114.94	259.57	168.59	33.56	0.01	2.977	0.016	0.604
27	134.83	259.26	169.71	102.24	0.01	2.665	0.016	0.489
28	156.88	260.13	169.42	285.00	0.01	2.351	0.013	0.388
29	181.14	262.27	17063	714.60	0.01	1.973	0.010	0.284
30	207.61	264.45	174.11	1464.00	0.01	1.481	0.013	0.196
31	236.28	270.09	177.12	2475.60	0.01	1.075	0.016	0.145
32	267.10	277.93	181.98	4381.20	0.01	0.774	0.015	0.110
33	300.00	285.18	184.76	6631.20	0.01	0.628	0.015	0.086
34	334.86	293.68	187.69	9450.00	1.29	0.550	0.016	0.073
35	371.55	300.12	190.34	12432.00	1.52	0.447	0.015	0.063
36	409.89	302.63	194.40	15468.00	2.12	0.361	0.015	0.057
37	449.67	304.43	198.46	18564.00	2.36	0.284	0.015	0.054
38	490.&5	307.2	201.53	21684.00	2.91	0.247	0.015	0.052
39	532.56	31217	202.74	24696.00	3.67	0.199	0.015	0.050
40	572.15	31556	201.61	27480.00	3.81	0.191	0.012	0.050
41	618.07	318.26	189.95	30288.00	6.82	0.171	0.010	0.049
42	661.00	321.71	189.95	32796.00	6.07	0.128	0.009	0.048
43	703.59	327.95	189.95	55328.00	6.73	0.124	0.009	0.047
44	745.48	333.77	189.95	37692.00	8.71	0.117	0.009	0.046
45	786.33	336.46	189.95	39984.00	8.26	0.115	0.008	0.045
46	825.75	338.54	189.95	42192.00	7.87	0.113	0.008	0.043
47	863.40	342.55	189.95	44220.00	7.53	0.111	0.007	0.041
48	898.93	346.23	189.95	46272.00	7.23	0.108	0.006	0.040
49	931.99	34924	189.95	47736.00	6.97	0.102	0.006	0.038
50	962.26	349.92	189.95	51264.00	6.75	0.099	0.006	0.034

51	989.45	350.09	189.95	49716.00	6.57	0.099	0.006	0.030
52	1013.29	360.09	189.95	47208.00	6.41	0.094	0.006	0.028
53	1033.54	350.09	189.95	47806.00	6.29	0.094	0.006	0.027
54	1050.00	350.09	189.95	47640.00	6.19	0.094	0.006	0.027

Table A2. Pressure levels adopted for IFS-137 137 pressure levels (in hPa).

Level number	pressure hPa	Level number	pressure hPa	Level number	pressure hPa	Level number	pressure hPa	Level number	pressure hPa
1	0.02	31	12.8561	61	106.4153	91	424.019	121	934.7666
2	0.031	32	14.2377	62	112.0681	92	441.5395	122	943.1399
3	0.0467	33	15.7162	63	117.9714	93	459.6321	123	950.9082
4	0.0683	34	17.2945	64	124.1337	94	478.3096	124	958.1037
5	0.0975	35	18.9752	65	130.5637	95	497.5845	125	964.7584
6	0.1361	36	20.761	66	137.2703	96	517.4198	126	970.9046
7	0.1861	37	22.6543	67	144.2624	97	537.7195	127	976.5737
8	0.2499	38	24.6577	68	151.5493	98	558.343	128	981.7968
9	0.3299	39	26.7735	69	159.1403	99	579.1926	129	986.6036
10	0.4288	40	29.0039	70	167.045	100	600.1668	130	991.023
11	0.5496	41	31.3512	71	175.2731	101	621.1624	131	995.0824
12	0.6952	42	33.8174	72	183.8344	102	642.0764	132	998.8081
13	0.869	43	36.4047	73	192.7389	103	662.8084	133	1002.225
14	1.0742	44	39.1149	74	201.9969	104	683.262	134	1005.356
15	1.3143	45	41.9493	75	211.6186	105	703.3467	135	1008.224
16	1.5928	46	44.9082	76	221.6146	106	722.9795	136	1010.849
17	1.9134	47	47.9915	77	231.9954	107	742.0855	137	1013.25
18	2.2797	48	51.199	78	242.7719	108	760.5996		
19	2.6954	49	54.5299	79	253.9549	109	778.4661		
20	3.1642	50	57.9834	80	265.5556	110	795.6396		
21	3.6898	51	61.5607	81	277.5852	111	812.0847		
22	4.2759	52	65.2695	82	290.0548	112	827.7756		
23	4.9262	53	69.1187	83	302.9762	113	842.6959		
24	5.6441	54	73.1187	84	316.3607	114	856.8376		
25	6.4334	55	77.281	85	330.2202	115	870.2004		
26	7.2974	56	81.6182	86	344.5663	116	882.791		
27	8.2397	57	86.145	87	359.4111	117	894.6222		
28	9.2634	58	90.8774	88	374.7666	118	905.7116		
29	10.372	59	95.828	89	390.645	119	916.0815		

30	11.5685	60	101.0047	90	407.0583	120	925.7571

Author contributions. ZS contributed the central idea. SC, ZS and HD conceived the method, developed the retrieval algorithm and discussed the results. SC analyzed the data, prepared the figures and wrote the paper. WG contributed to refining the ideas, carrying out additional analyses. All co-authors reviewed the paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The study was supported by the National Key Research Program of China: Development of high-resolution data assimilation technology and atmospheric reanalysis data set in East Asia (Research on remote sensing telemetry data assimilation technology, Grant no. 2017YFC1501802). The study was also supported by the National Natural Science Foundation of China (Grant no. 41875045) and Hunan Provincial Innovation Foundation for Postgraduate (Grant no. CX2018B033 and no. CX2018B034).

References

Aires, F., Schmitt, M., Chedin, A., and Scott, N.: The "weighting smoothing" regularization of MLP for Jacobian stabilization, IEEE. T. Neural. Networks., 10, 1502-1510, https://doi.org/10.1109/72.809096, 1999.

Aires, F., Chédin, Alain., Scott, N. A., and Rossow, W. B.: A regularized neural net approach for retrieval of atmospheric and surface temperatures with the IASI instrument, J. Appl. Meteorol., 41,144-159, https://doi.org/10.1175/1520-0450(2002)041<0144:ARNNAF>2.0.CO;2, 2002.

Aumann, H. H.: Atmospheric infrared sounder on the earth observing system, Optl. Engr., 33, 776-784, https://doi.org/10.1117/12.159325, 1994.

Aumann, H. H., Chahine, M. T., Gautier, C., and Goldberg, M.: AIRS/AMSU/HSB on the Aqua mission: design, science objective, data products, and processing systems, IEEE. Trans. GRS., 41,253-264, http://dx.doi.org/10.1109/TGRS.2002.808356, 2003.

Brath, M., Fox, S., Eriksson, P., Harlow, R. C., Burgdorf, M., and Buehler, S. A.: Retrieval of an ice water path over the ocean from ISMAR and MARSS millimeter and submillimeter brightness temperatures, Atmos. Meas. Tech., 11, 611–632, https://doi.org/10.5194/amt-11-611-2018, 2018.

Chahine, M. I.: A general relaxation method for inverse solution of the full radiative transfer equation, J. Atmos. Sci., 29, 741-747, https://doi.org/10.1175/1520-0469(1972)029<0741:AGRMFI>2.0.CO;2, 1972.

Chang, K. W, L'Ecuyer, T. S., Kahn, B. H., and Natraj, V.: Information content of visible and midinfrared radiances for retrieving tropical ice cloud properties, J. Geophys. Res., 122, https://doi.org/10.1002/2016JD026357, 2017.

Chedin, A., Scott, N. A., Wahiche, C., and Moulinier, P.: The improved initialization inversion method: a high resolution physical method for temperature retrievals from satellites of the tiros-n series, J. Appl. Meteor., 24, 128-143, https://doi.org/10.1175/1520-0450(1985)024<0128:TIIIMA>2.0.CO;2, 1985.

Cyril, C., Alain, C., and Scott, N. A.: Airs channel selection for CO_2 and other trace-gas retrievals, Q. J. Roy. Meteor. Soc., 129, 2719-2740, https://doi.org/10.1256/qj.02.180, 2003.

Du, H. D., Huang, S. X., and Shi, H. Q.: Method and experiment of channel selection for high spectral resolution data, Acta. Physica. Sinica., 57, 7685-7692, 2008 .

Dudhia, A., Jay, V. L., and Rodgers, C. D.: Microwindow selection for high-spectral-resolution sounders, Appl. Opt. 41, 3665-3673, https://doi.org/10.1364/AO.41.003665, 2002.

Eresmaa, R. and McNally, A. P.: Diverse profile datasets from the

ECMWF 137-level short-range forecasts, Tech. rep., ECMWF, 2014.

Eyre, J. R., Andersson E., and McNally, A. P.: Direct use of satellite sounding radiances in numerical weather prediction, High Spectral Resolution Infrared Remote Sensing for Earth's Weather and Climate Studies, Springer, Berlin, Heidelberg, https://doi.org/10.1007/978-3-642-84599-4_25, 1993.

Fang, Z. Y.: The evolution of meteorological satellites and the insight from it, Adv. Meteorol. Sci. Technol., 4, 27-34, https://doi.org/ 10.3969/j.issn.2095-1973.2014.06.003, 2014.

Gong, J., Wu, D. L., and Eckermann, S. D.: Gravity wave variances and propagation derived from AIRS radiances, Atmos. Chem. Phys., 11, 11691-11738, https://doi.org/10.5194/acp-12-1701-2012, 2011.

He, M. Y., Du, H. D., Long, Z. Y., and Huang, S. X.: Selection of regularization parameters using an atmospheric retrievable index in a retrieval of atmospheric profile, Acta. Physica Sinica., 61, 024205-160, 2012.

Hoffmann, L. and Alexander, M. J.: Retrieval of stratospheric temperatures from atmospheric infrared sounder radiance measurements for gravity wave studies, J. Geophys. Res. Atm., 114, https://doi.org/10.1029/2008JD011241, 2009.

Huang, H. L., Li, J., Baggett, K., Smith, W. L., and Guan, L.: Evaluation of cloud-cleared radiances for numerical weather prediction and cloud-contaminated sounding applications, Atmospheric and Environmental Remote Sensing Data Processing and Utilization: Numerical Atmospheric Prediction and Environmental Monitoring, I. S. O. Photonics., https://doi.org/10.1117/12.613027, 2005.

Kuai, L., Natraj, V., Shia, R. L., Miller, C., and Yung, Y. L.: Channel selection using information content analysis: a case study of CO_2 retrieval from near infrared measurements. J. Q. S. Radiative. Transfer., 111, 1296-1304, https://doi.org/10.1016/j.jqsrt.2010.02.011, 2010.

Li, J., Wolf, W. W., Menzel, W. P., Paul, Menzel. W., Zhang, W. J., Huang, H. L., and Achtor, T. H.: Global soundings of the atmosphere from ATOVS measurements: the algorithm and validation, J. Appl. Meteor., 39, 1248-1268, https://doi.org/10.1175/1520-0450(2000)039<1248:GSOTAF>2.0.CO;2, 2000.

Li, J., Liu, C. Y., Huang, H. L., Schmit, T. J., Wu, X., Menzel, W. P., and Gurka, J. J.: Optimal cloud-clearing for AIRS radiances using MODIS, IEEE. Trans. GRS. , 43, 1266-1278, http://dx.doi.org/10.1109/tgrs.2005.847795, 2005.

Liu, Z. Q.: A regional ATOVS radiance-bias correction scheme for rediance assimilation, Acta. Meteorologica. Sinica., 65, 113-123, 2007.

Lupu, C., Gauthier, P., and Laroche, Stéphane.: Assessment of the impact of observations on analyses derived from observing system experiments, Mon. Weather. Rev., 140, 245-257, https://doi.org/10.1175/MWR-D-10-05010.1, 2012.

Menke, W.: Geophysical Data Analysis: Discrete Inverse Theory, Acad. Press., Columbia University, New York, https://doi.org/10.1016/B978-0-12-397160-9.00019-9, 1984.

Menzel, W. P., Schmit, T. J., Zhang, P. and Li, J.: Satellite-based atmospheric infrared sounder development and applications, Bull. Amer. Meteor. Soc., 99, 583–603, https://doi.org/10.1175/BAMS-D-16-0293.1, 2018.

Prunet, P., Thépaut J. N., and Cass, V.: The information content of clear sky IASI radiances and their potential for numerical weather prediction, Q. J. Roy. Meteor. Soc., 124, 211-241, https://doi.org/10.1002/qj.49712454510, 2010.

Xu, Q.: Measuring information content from observations for data assimilation: relative entropy versus shannon entropy difference, Tellus. A., 59, 198-209, https://doi.org/10.1111/j.1600-0870.2006.00222.x, 2007.

Rabier, F., Fourrié, N., and Chafäi, D.: Channel selection methods for infrared atmospheric sounding interferometer radiances, Q. J. Roy. Meteor. Soc., 128, 1011-1027, https://doi.org/10.1256/003590002321042036, 2010.

Richardson, M. and Stephens, G. L.: Information content of oco-2 oxygen a-band channels for retrieving marine liquid cloud properties, Atmospheric Measurement Techniques, 11, 1-19, https://doi.org/10.5194/amt-11-1515-2018, 2018.

Rodgers, C. D.: Information content and optimisation of high spectral resolution remote measurements, Adv. Spa. Research, 21, 136-147, https://doi.org/10.1016/S0273-1177(97)00915-0, 1996.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding, Inverse methods for atmospheric sounding, World Scientific, https://doi.org/10.1142/3171, 2000.

Saunders, R., Hocking, J., Turner, E., Rayer, P., Rundle, D., Brunel, P., Vidot, J., Roquet, P., Matricardi, M., Geer, A., Bormann, N., and Lupu, C.: An update on the RTTOV fast radiative transfer model (currently at version 12), Geosci. Model Dev., 11, 2717-2737, https://doi.org/10.5194/gmd-11-2717-2018, 2018.

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

Smith, W. L., Woolf, H. M., and Revercomb, H. E.: Linear simultaneous solution for temperature and absorbing constituent profiles from radiance spectra, Appl. Optics., 30, 1117, https://doi.org/10.1364/AO.30.001117, 1991.

Wakita, H., Tokura, Y., Furukawa, F., and Takigawa, M.: Study of the information content contained in remote sensing data of atmosphere, Acta. Physica. Sinica., 59, 683-691, 2010.

Wang, G., Lu, Q. F., Zhang, J. W., and Wang, H. Y.,.: Study on method and experiment of hyper-spectral atmospheric infrared sounder channel selection, Remote Sensing Technology and Application., 29, 795-802 , 2014.

Zhang, J. W., Wang, G., Zhang, H., Huang J., Chen J., and Wu, L. L.: Experiment on hyper-spectral atmospheric infrared sounder channel selection based on the cumulative effect coefficient of principal component, Journal of Nanjing Institute of meteorology, 1, 36-42, http://dx.doi.org/10.3969/j.issn.1674-7097.2011.01.005, 2011.