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

2 Dear Referee #2,

3 We thank you for your review of our manuscript and your detailed

4 remarks. We would like to improve the article here in an immediate reply.

5 Please, find our answers/comments on your notes below:

6 (1) p. 5, l. 108: ... from the thousands of channels' observations... please

7 rephrase, but I have no suggestions

8 Yes, you are right.

9 "It is important to properly ... from the thousands of channels'
10 observations ...." has been modified to "In order to improve the
11 calculation efficiency and retrieval quality, it is very important to properly
12 select a set of channels that can provide as much information as possible."
13 (L 104-L 107)

14 Thanks.

15 (2) p. 8, l. 155: Channel selection mostly uses the information content

and delievers the largest amount of information for the selected channel

17 combination during the retreival.

18 Yes, you are right.

19 "Currently, information content is often employed in channel selection.

20 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

- 23 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
- 26 al., 2018)." (L 153-L 156)

27 Thanks.

- 28 (3) p. 9, l. 194: after the retrieval
- 29 Yes, you are right.
- 30 "...after retrieval" has been modified to "...after the retrieval" (L 193)
- 31 Thanks.
- 32 (4) p. 11, l. 236: (2) The number looks like an equation number, so it is a
- 33 bit confusing.
- 34 Sorry for the confusion.
- The step number has been modified to (I), (II), (III) and (IV). (L 222; L
- 36 234; L 255; L 270; L 272)
- 37 Thanks.
- 38 (5) p. 18, l. 389: The footprint size is 13.5 km
- 39 Yes, you are right.
- 40 This has been modified. (L 288)
- 41 Thanks.
- 42 (6) p. 19, 1. 401/402: in Fig. 1. The measurement error is not below 0.2K
- 43 for all the instrument channels. There are...
- 44 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)
- 51 Thanks.
- 52 (7) p. 21, l. 434: ranking -> in the range
- 53 Yes, you are right.
- 54 This has been modified. (L 432)
- 55 Thanks.
- 56 (8) p. 27, l. 532: temperature, a better observation can be obtained for
- 57 higher temperatures
- 58 Yes, you are right.
- <sup>59</sup> "...temperature, the higher temperature is, the better observation can be
- obtained;" has been modified to "...temperature, a better observation can
- 61 be obtained for higher temperatures" (L 524-L 525)
- 62 Thanks.
- 63 (9) p. 28, l. 547: to the -> at
- 64 Yes, you are right.
- 65 This has been modified. (L 538)
- 66 Thanks.

- 67 (10) p. 30, l. 567: reach to -> reaches
- 68 Yes, you are right.
- 69 This has been modified. (L 560)
- 70 Thanks.
- (11) p. 36, l. 650: humidity and 10m wind speed (remove dot after speed)
- 72 Sorry for my carelessness.
- 73 This has been modified. (L 643)
- 74 Thanks.
- 75 (12) p. 43, l. 778: to the -> at
- 76 Yes, you are right.
- This has been modified. (L 771; L 828)
- 78 Thanks again for your careful review.
- 79

80 Dear Referee #4,

- 81 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.
- 83 Please, find our answers/comments on your notes below:
- 84 (1) p1, 116-20: Suggest to replace the first two sentences of the abstract
- 85 by: "This study introduces an effective channel selection method for
- <sup>86</sup> hyperspectral infrared sounders. The method is illustrated for the
- 87 Atmospheric InfraRed Sounder (AIRS) instrument."
- 88 Yes, we agree with you.

- 89 This has been modified. (L 16-L 18)
- 90 Thanks.
- 91 (2) p1, l21: "improved method" -> "improved channel selection (ICS)
- 92 method"
- 93 Yes, you are right.
- 94 This has been modified. (L 19-L 21)
- 95 Thanks.
- 96 (3) p2, l24: suggest to delete "and closer to the actual atmosphere"
- 97 Yes, you are right.
- 98 This has been deleted. (L 22)
- 99 Thanks.
- 100 (4) p2, 130: suggest to replace "In general, ..." by "Also at lower
- 101 heights, ..."
- 102 Yes, you are right.
- 103 This has been modified. (L 28)
- 104 Thanks.
- 105 (5) p2, l31-32: delete "(Improved Channel Selection)"
- 106 Yes, you are right.
- 107 This has been deleted. (L 29)
- 108 Thanks.
- (6) p2, l32-36: suggest to rephrase this as "Statistical inversion
- 110 comparison experiments for four different regions illustrate latitudinal

- and seasonal variations and better performance of ICS compared to the
- 112 NWP Channel Selection (NCS) and Primary Channel Selection (PCS)

methods. The ICS method shows potential for future applications."

114 Yes, we agree with you.

"Statistical inversion comparison experiments in four typical regions 115 indicate that ICS in this paper is significantly better than NCS (NWP 116 Channel Selection) and PCS (Primary Channel Selection) in different 117 regions and shows latitudinal variations, which shows potential for future 118 applications." has been modified to "Statistical inversion comparison 119 experiments for four different regions illustrate latitudinal and seasonal 120 variations and better performance of ICS compared to the NWP Channel 121 Selection (NCS) and Primary Channel Selection (PCS) methods. The ICS 122 method shows potential for future applications." (L 30-L 34) 123

- 124 Thanks.
- 125 (7) p4, 170: please fix: "Infrared Atmospheric Sounding Interferometer"
- 126 Yes, you are right.
- 127 This has been modified. (L 68-L 69)
- 128 Thanks.
- (8) p5, 1110: replace "the channel selection algorithm" by "channel
- 130 selection algorithms"
- 131 Yes, you are right.
- 132 This has been modified. (L 108)

- 133 Thanks.
- (9) p11, l236: replace "p matrix" by "for calculating the p value"?
- 135 Yes, you are right.
- 136 This has been modified. (L 234)
- 137 Thanks.
- 138 (10) p11, 1237: replace "matrixes, it" by "matrices, they"
- 139 Yes, you are right.
- 140 This has been modified. (L 235)
- 141 Thanks.
- 142 (11) p12, l257: replace "According to" by "Using"
- 143 Yes, you are right.
- 144 This has been modified. (L 257)
- 145 Thanks.
- (12) p18, 1383: The reference to Hoffmann and Alexander (200) can be
- deleted here, I think. Maybe move reference to Susskind et al. (2003)
- 148 from line 389 to this place.
- 149 Yes, we agree with you.
- 150 "...(Aumann et al., 2003; Hoffmann and Alexander, 2009)." has been
- 151 modified to "...(Aumann et al., 2003; Susskind et al., 2003).". (L
- 152 382-L383)
- "Susskind et al. (2003)" has been deleted. (L 388)
- 154 Thanks.

(13) p27, 1515-534: 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

158 the 10 um band..."

159 Yes, you are right.

"(1) When the wavenumber approaches 1000, the wavelength is 10  $\mu$ m 160 (1/1000 cm-1). Near this band, fewer channels are selected by PCS 161 because the retrieval of ground temperature is considered by NCS; (2) 162 When the wavenumber is near 1200, the wavelength is 9  $\mu$ m (1/1200 163 cm-1). Near this band, no channels are selected by PCS because the 164 retrieval of O3 is not considered in this paper; (3) When the wavenumber 165 approaches 1500, the wavelength is 6.7  $\mu$ m (1/1500 cm-1). As is known, 166 the spectral range from 6 µm to 7 µm corresponds to water vapor 167 absorption bands, but fewer channels are selected by NCS; (4) When the 168 wavenumber is close to 2000, it derives a wavelength of 5  $\mu$ m (1/2000) 169 cm-1), which includes 4.2 µm for N2O and 4.3 µm for CO2 absorption 170 bands." has been modified to "(1) Near 10 µm band, fewer channels are 171 selected by PCS because the retrieval of ground temperature is considered 172 by NCS; (2) Near 9 µm band, no channels are selected by PCS because 173 the retrieval of O3 is not considered in this paper; (3) As is known, the 174 spectral range from 6 µm to 7 µm corresponds to water vapor absorption 175 bands, but fewer channels are selected by NCS; (4) Near 5 µm band, it 176

- 177 includes 4.2  $\mu$ m for N2O and 4.3  $\mu$ m for CO2 absorption bands." (L 178 514-L 521)
- 179 "(5) In the near infrared area, the wavenumber exceeds 2200, deriving a
- wavelength of less than 4  $\mu$ m (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  $\mu$ m band, a small number of channels is selected

by NCS, but no channels are selected by PCS." (L 526-L 527)

184 Thanks.

- (14) p27, 1529-533: It is still not clear what is meant by "high temperature
- zone", I think. Does this refer to temperatures at high altitudes?
- 187 Sorry for the confusion.
- 188 "high temperature zone" is not clear for readers which we want to say
- 189 "under the higher temperature conditions" But it can be deleted in this
- sentence, because we have explained it in the following sentences.
- <sup>191</sup> "PCS is favorable for atmospheric temperature observation in the high
- temperature zone. Because 4.2  $\mu$ m and 4.3  $\mu$ m 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  $\mu$ m and 4.3  $\mu$ m bands are sensitive to high
- temperature, a better observation can be obtained for higher temperatures;"
- 197 (L 522-L 525)
- <sup>198</sup> "Due to the method selected in this paper, there are more channels at 4.2

- $\mu$ m for N2O and 4.3  $\mu$ m for CO2 absorption bands, and the channel
- 200 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
- $4.2 \ \mu m$  for N2O and  $4.3 \ \mu m$  for CO2 absorption bands, and the channel
- combination of PCS is superior to that of NCS for atmospheric
- temperature observation." (L 689-L 693)
- 206 Thanks.
- 207 (15) p28, 1537: replaced "used in this paper" by "considered in this study"
- 208 Yes, you are right.
- 209 This has been modified. (L 528)
- 210 Thanks.
- (16) p29, Fig. 5: I would like to suggest to combine the curves of plot a)
- and b) into a single plot.
- 213 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)
- 217 Thanks again for your careful review.
- 218

220	(b) The list of all relevant changes made in the manuscript
221	A channel selection method for hyperspectral
222	atmospheric infrared sounders based on
223	layering
224	Shujie Chang <sup>1, 2,3</sup> , Zheng Sheng <sup>1,2</sup> , Huadong Du <sup>1,2</sup> , Wei Ge <sup>1,2</sup> and
225	Wei Zhang <sup>1,2</sup>
226	<sup>1</sup> College of Meteorology and Oceanography, National University of
227	Defense Technology, Nanjing, China
228	<sup>2</sup> Collaborative Innovation Center on Forecast and Evaluation of
229	Meteorological Disasters, Nanjing University of Information
230	Science and Technology, Nanjing, China
231	<sup>3</sup> South China Sea Institute for Marine Meteorology, Guangdong
232	Ocean University, Zhanjiang, China
233	
234	Correspondence: Zheng Sheng (19994035@sina.com)
235	
236	Abstract. Because a satellite channel's ability to resolve
237	hyperspectral data varies with height, This study introduces an
238	improved channel selection method is proposed based on
239	information content. An effective channel selection scheme-method
240	for a hyperspectral atmospheric infrared sounder using sounders.
241	The method is illustrated for the Atmospheric InfraRed Sounder

(AIRS data based on layering is proposed.) instrument. The results 242 are as follows: (1) Using the improved method, channel selection 243 (ICS), the atmospheric retrievable index is more stable, the value 244 reaching 0.54. The coverage of the weighting functions is more 245 evenly distributed over height with this method and closer to the 246 actual atmosphere; (2) Statistical inversion comparison experiments 247 show that the accuracy of the retrieval temperature, using the 248 improved channel selection method in this paper, is consistent with 249 that of 1Dvar channel selection. In the stratosphere and mesosphere 250 especially, from 10 hPa to 0.02 hPa, the accuracy of the retrieval 251 temperature of our improved channel selection method is improved 252 by about 1 K. In general Also at lower heights, the accuracy of the 253 retrieval temperature of ICS (Improved Channel Selection) is 254 improved; (3) Statistical inversion comparison experiments infor 255 four typical regions indicate that ICS in this paper is significantly 256 different regions illustrate latitudinal and seasonal variations and 257 better than NCS (performance of ICS compared to the NWP Channel 258 Selection (NCS) and PCS (Primary Channel Selection) in different 259 regions and shows latitudinal variations, which (PCS) methods. The 260 **ICS** method shows potential for future applications. 261

262

## 263 **1 Introduction**

Since the successful launch of the first meteorological satellite, 264 TIROS in the 1960s, satellite observation technology has developed 265 rapidly. Meteorological satellites observe the Earth's atmosphere 266 from space and are able to record data from regions which are 267 otherwise difficult to observe. Satellite data greatly enrich the 268 content and range of meteorological observations, and consequently, 269 atmospheric exploration technology and meteorological observations 270 have taken us to a new stage in our understanding of weather 271 systems and related phenomena (Fang, 2014). From the perspective 272 of vertical atmospheric observation, satellite instruments are 273 developing rapidly. In their infancy, the traditional infrared 274 measurement instruments for detecting atmospheric temperature and 275 moisture profiles, such as TOVS (Smith et al., 1991) or HIRS in 276 ATOVS (Chahine, 1972; Li et al., 2000; Liu, 2007), usually 277 employed filter spectrometry. Even though such instruments have 278 played an important role in improving weather prediction, it is 279 difficult to continue to build upon improvements in terms of 280 observation accuracy and vertical resolution due to the limitation of 281 low spectral resolution. By using this kind of filter-based 282 spectroscopic measurement instrument, therefore, it is difficult to 283 meet today's needs in numerical weather prediction (Eyre et al., 284 1993; Prunet et al., 2010; Menzel et al., 2018). To meet this 285

286	challenge, a series of plans for the creation of high-spectral
287	resolution atmospheric measurement instruments has been executed
288	in the United States and in Europe in recent years: One example is
289	the AIRS (Atmospheric InfraRed Sounder) on the Earth Observation
290	System, "Aqua", launched on May 4, 2002 from the United States.
291	AIRS has 2378 spectral channels providing sensitivity from the
292	ground to up to about 65 km of altitude (Aumann et al., 2003;
293	Hoffmann and Alexander, 2009; Gong et al., 2011). The United
294	States and Europe, in 2010 and in 2007, also installed the CRIS
295	(Cross-track Infrared Sounder) and the IASI
296	(Inter-AttractiveInfrared Atmospheric Sounding Interferometer) on
297	polar-orbiting satellites.
298	China also devotes great importance to the development of such

advanced sounding technologies. In the early 1990s, the National 299 Satellite Meteorological Center began to investigate the principles 300 and techniques of hyperspectral resolution atmospheric observations. 301 China's development of interferometric atmospheric vertical 302 detectors eventually led to the launch of Fengyun No. 3, on May 27, 303 2008, and Fengyun No. 4 on December 11, 2016, both of which 304 were equipped with infrared atmospheric instruments. How best to 305 use the hyperspectral resolution observation data obtained from 306 these instruments, to obtain reliable atmospheric temperature and 307

humidity profiles, is an active area of study in atmospheric inversiontheory.

Due to technical limitations, only a limited number of channels 310 could at first be built into the typical satellite instruments. In this 311 case, channel selection generally involved controlling the channel 312 weighting function by utilizing the spectral response characteristics 313 of the channel (such as center frequency and bandwidth). With the 314 development of measurement technology, increasing numbers of 315 hyperspectral detectors were carried on meteorological satellites. 316 Due to the large number of channels and data supported by such 317 instruments today (such as AIRS with 2378 channels and IASI with 318 8461 channels), it has proven extremely cumbersome to store, 319 transmit, and process such data. Moreover, there is often a close 320 correlation between the channel, causing an ill-posedness of the 321 inversion, potentially compromising accuracy of the retrieval 322 product based on hyperspectral resolution data. 323

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 330 large enough to meet the demands of simulating all the channels 331 simultaneously within the forecast time. It order to improve the 332 calculation efficiency and retrieval quality, it is very important to 333 properly select a groupset of channels that can provide as much 334 information as possible from the thousands of channels' observations 335 to improve the calculation efficiency and retrieval quality. 336 Many researchers have studied the channel selection 337 algorithm.algorithms. Menke (1984) first chose channels using a data 338 precision matrix method. Aires et al. (1999) made the selection using 339 the Jacobian matrix, which has been widely used since then (Aires et 340 al., 2002; Rabier et al., 2010). Rodgers (2000) indicated that there 341 are two useful quantities in measuring the information provided by 342 the observation data: Shannon information content and degrees of 343 freedom. The concept of information capacity then became widely 344 used in satellite channel selection. In 2007, Xu (2007) compared the 345 Shannon information content with the relative entropy, analyzing the 346 information loss and information redundancy. In 2008, Du et al. 347 (2008) introduced the concept of the atmospheric retrievable index 348 (ARI) as a criterion for channel selection, and in 2010, Wakita et al. 349 (2010) produced a scheme for calculating the information content of 350 the various atmospheric parameters in remote sensing using 351

Bayesian estimation theory. Kuai et al. (2010) analyzed both the 352 Shannon information content and degrees of freedom in channel 353 selection when retrieving CO<sub>2</sub> concentrations using thermal infrared 354 remote sensing and indicated that 40 channels could contain 75% of 355 the information from the total channels. Cyril et al. (2003) proposed 356 the optimal sensitivity profile method based on the sensitivity of 357 different atmospheric components. Lupu et al. (2012) used degrees 358 of freedom for signals (DFS) to estimate the amount of information 359 contained in observations in the context of observing system 360 experiments. In addition, the singular value decomposition method 361 has also been widely used for channel selection (Prunet et al., 2010; 362 Zhang et al., 2011; Wang et al., 2014). In 2017, Chang et al. (2017) 363 selected a new set of Infrared Atmospheric Sounding Interferometer 364 (IASI) channels using the channel score index (CSI). Richardson et 365 al. (2018) selected 75 from 853 channels based on the high 366 spectral-resolution oxygen A-band instrument on NASA's Orbiting 367 Carbon Observatory-2 (OCO-2), using information content analysis 368 to retrieve the cloud optical depth, cloud properties, and position. 369 Today's main methods for channel selection use only the 370 weighting function to study appropriate numerical methods, such as 371 the data precision matrix method (Menke, 1984), singular value 372 decomposition method (Prunet et al., 2010; Zhang et al., 2011; Wang 373

374	et al., 2014), and the Jacobi method (Aires et al., 1999; Rabier et al.,
375	2010). The use of the methods allows sensitive channels to be
376	selected. The above-mentioned studies also take into account the
377	sensitivity of each channel to atmospheric parameters during channel
378	selection, while ignoring some factors that impact retrieval results.
379	The accuracy of retrieval results depends not only on the channel
380	weighting function but also on the channel noise, background field,
381	and the retrieval algorithm.

Currently, Channel selection mostly uses the information content isoften employed in channel selection. During retrieval, this
methodand delivers the largest amount of information for the selected
channel combination <u>during the retrieval</u> (Rodgers, 1996; Du et al.,
2008; He et al., 2012; Richardson et al., 2018).

This method has made great breakthroughs in both theory and 387 practice, and the concept of information content itself does consider 388 all the height dependencies of the kernel matrix K (Rodgers, 2000). 389 However, earlier works have neglected the height dependencies of K 390 for simplicity. This paper uses the atmospheric retrievable index 391 (ARI) as the index, which is based on information content (Du et al., 392 2008; Richardson et al. 2018). Channel selection is made at different 393 heights, and an effective channel selection scheme is proposed 394 which fully considers various factors, including the influence of 395

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.

# 401 2 Channel selection indicator, scheme and method

#### 402 **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:

407

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

408

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

410

where  $S_a$  is the error covariance matrix of the background or the estimated value of atmospheric profile,  $S_{\varepsilon}$  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:

420 
$$p = 1 - \exp(\frac{1}{2n} ln |\hat{S}S_a^{-1}|),$$
 (2)

421

419

Assuming that before and after the retrieval, the ratio of the root 422 mean square error of each element in the atmospheric state vector is 423 1-p, then  $|\hat{S}S_a^{-1}| = (1-p)^{2n}$  is derived. By inverting the equation, 424 the ARI that is p can be obtained in Eq. (2), which indicates the 425 relative portion of the error that is eliminated by retrieval. In fact, 426 before and after retrieval, the ratio of the root mean square error of 427 each element cannot be 1-p. Therefore, p defined by Eq. (1) is 428 actually an overall evaluation of the retrieval result. 429

430

#### 431 **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 atmosphereand 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  $C_N^n$ 439 kinds of p results. Furthermore, there are M layers in the vertical 440 direction of the atmosphere. Therefore, the entire atmosphere must 441 be calculated  $\mathbf{M} \cdot \mathbf{C}_{\mathbf{N}}^{\mathbf{n}}$  times. However, the calculation  $\mathbf{M} \cdot \mathbf{C}_{\mathbf{N}}^{\mathbf{n}}$  times 442 will be particularly large, which makes this approach impractical in 443 calculating p for all possible combinations. Therefore, it is necessary 444 to design an effective calculation scheme, and such a scheme, i.e., a 445 channel selection method, using iteration is proposed, called the 446 "sequential absorption method" (Dudhia et al., 2002; Du et al., 2008). 447 The method's main function is to select ("absorb") channels one by 448 one, taking the channel with the maximum value of p. Through n 449 iterations, n channels can be selected as the final channel 450 combination. The steps are as follows: 451

(41) 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:

457 
$$\hat{S} = S_a - S_a k^T (s_{\varepsilon} + k S_a k^T)^{-1} k S_a.$$
 (3)

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

460 
$$\hat{S} = \left(I - \frac{S_a k^T k}{\left(s_{\varepsilon} + k S_a k^T\right)}\right) S_a = \left(I - \frac{\left(k S_a\right)^T k}{\left(s_{\varepsilon} + k \left(k S_a\right)^T\right)}\right) S_a.$$
(4)

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

462 
$$p = 1 - \exp(\frac{1}{2n}ln(\left|I - \frac{(kS_a)^T k}{(s_{\varepsilon} + k(kS_a)^T)}\right|)).$$
 (5)

463

464 (2II) Simplification of Eq. (5) for calculating the p matrix value: 465 Since  $S_a$  and  $S_{\varepsilon}$  are positive definite symmetric matrixes, 466 itmatrices, they can be decomposed into  $S_a = (S_a^{1/2})^T (S_a^{1/2})$  and 467  $S_{\varepsilon} = (S_{\varepsilon}^{1/2})^T (S_{\varepsilon}^{1/2})$ .

468

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

470

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,  $r = s_{\varepsilon}^{-1/2} k S_a^{1/2}$ , represents the normalized weighting function matrix of a single channel. Substituting r into Eq. (5) gives:

476 
$$p = 1 - \exp(\frac{1}{2n} ln\left(\left|I - \frac{rr^T}{1 + r^T r}\right|\right)).$$
 (7)

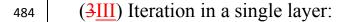
477

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

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

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

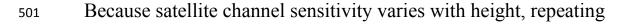
482 = 
$$1 - \exp\left(-\frac{1}{2n}\ln(1+r^T r)\right).$$
 (8)



First, the iteration in a single layer requires the calculation of R. 485 According to Using  $S_a$ ,  $S_\epsilon$ , K and Eq. (6), R can be calculated. 486 Second, using Eq. (8), p of each candidate channel can be calculated. 487 Moreover, the channel corresponding to maximum p is the selected 488 channel for this iteration. After a channel has been selected, 489 according to Eq. (3) we can use  $\hat{S}$  to get  $S_a$  for the next iteration. 490 Finally, channels which are not selected during this iteration are used 491 as the candidate channels for the next iteration. 492 When selecting n from N channels, it is necessary to calculate 493  $(N-n/2)n \approx Nn p$  values, which is much smaller than  $C_N^n$ . In addition 494 to high computational efficiency by using this method, another 495 advantage is that all channels can be recorded in the order in which 496 they are selected. In the actual application, if n'channels are 497 needed, and n' < n, we will not need to select the channel again, 498 but record the selected channel only. 499

#### 500

(4<u>IV</u>) Iteration for different altitudes:



the iterative process of step (3111), 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· (N – n/2)n  $\approx$  M·Nn p values, a much smaller number than M·C<sup>n</sup><sub>N</sub>. In this way, different channel sets can be used to evaluate corresponding height in the retrieved profiles.

508

**2.3 Statistical inversion method** 

The inversion methods for the atmospheric temperature profiles can 510 be summarized in two categories: statistical inversion and physical 511 inversion. Statistical inversion is essentially a linear regression 512 model which uses a large number of satellite measurements and 513 atmospheric parameters to match samples and calculate their 514 correlation coefficient. Then, based on the correlation coefficient, the 515 required parameters of the independent measurements obtained by 516 the satellite are retrieved. Because the method does not directly solve 517 the radiation transfer equation, it has the advantages of fast 518 calculation speed. In addition, the solution is numerically stable, 519 which makes it one of the highest precision methods (Chedin et al., 520 1985). Therefore, the statistical inversion method will be used for 521 our channel selection experiment and a regression equation will be 522 established. 523

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

528 
$$\mathbf{T} = T^* \cdot A,\tag{9}$$

529

$$530 \quad T_b = T_b^* \cdot A, \tag{10}$$

531

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:

538

 $S39 \quad \mathbf{A} = \mathbf{V} \cdot \mathbf{B},\tag{11}$ 

540

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

542

545 
$$\mathbf{B} = (T_b^*)^T T_b,$$
 (13)

547 
$$A = (T^*)^T T.$$
 (14)

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

552

553 
$$\widehat{T} = \overline{T} + \widehat{T}' = \overline{T} + GT_{b}' = \overline{T} + G(T_{b} - \overline{T_{b}}),$$
 (15)

where  $\widehat{T}$  stands for the retrieval atmospheric temperature.  $\overline{T}$  and  $\overline{T_b}$  are the corresponding average values of the elements,

respectively.  $\widehat{T}'$  and  $\overline{T}_{b}'$  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:

562 
$$T' = (t'_1, t'_2, \dots, t'_k),$$
 (16)

563

564 
$$T_{b}^{'} = (t_{b1}^{'}, t_{b2}^{'}, \cdots, t_{bk}^{'}).$$
 (17)

565

566 Define the inversion error matrix as:

568 
$$\delta = \overline{T} - \widehat{T} = \widehat{T}' - T' .$$
 (18)

570 The retrieval error covariance matrix is:

571

$$S_{\delta} = \frac{1}{k - n - 1} \delta \delta^{T}$$

$$= \frac{1}{k - n - 1} (T' - GT_{b}') (T' - GT_{b}')^{T}$$

$$= \frac{k - 1}{k - n - 1} (S_{e} - G^{T}S_{xy} - S_{xy}G^{T} + GS_{y}G^{T}), \qquad (19)$$

574

575 where

576

577 
$$S_e = \frac{1}{k-1} T' T'^{T}$$
,  
578  $S_y = \frac{1}{k-1} T_b' T_b^{T}^{T}$ ,  
579  $S_{xy} = \frac{1}{k-1} T' T_b^{T}^{T}$ . (20)

580

 $S_e$  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,  $S_\delta$ , 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:

$$\delta^{2} = tr(S_{\delta}) = min.$$
(21)
Taking a derivative of Eq. (21) with respect to G,  $\frac{\partial}{\partial G} tr(S_{\delta}) = 0 =$ 
(-2S<sub>xy</sub> + 2GS<sub>y</sub>), which means that:
$$G = S_{xy}S_{y}^{-1}.$$
(22)
Substituting Eq. (22) into Eq. (15) finally gives the least squares
solution as:
$$\hat{T} = \bar{T} + S_{xy}S_{y}^{-1}(T_{b} - \overline{T_{b}}).$$
(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<sub>xy</sub> and S<sub>y</sub> can be determined.

**3. Channel selection experiment** 

609 3.1 Data and model

588

610	The Atmospheric Infrared Sounder (AIRS) is primarily designed to
611	measure the Earth's atmospheric water vapor and temperature
612	profiles on a global scale (Aumann et al., 2003; Hoffmann and
613	Alexander, 2009). Susskind et al., 2003). AIRS is a continuously
614	operating cross-track scanning sounder, consisting of a telescope that
615	feeds an echelle spectrometer. The AIRS infrared spectrometer
616	acquires 2378 spectral samples at a resolution $\lambda/\Delta\lambda$ , ranging from
617	1086 to 1570, in three bands: 3.74 $\mu m$ to 4.61 $\mu m,$ 6.20 $\mu m$ to 8.22
618	$\mu$ m, and 8.8 $\mu$ m to 15.4 $\mu$ m. The footprint size <u>is</u> 13.5 km <del> at nadir</del>
619	(Susskind et al., 2003). The spectral range includes 4.3 µm and 15.5
620	$\mu m$ for important temperature observation and CO <sub>2</sub> , 6.3 $\mu m$ for
621	water vapor, and 9.6 $\mu$ m for ozone absorption bands (Menzel et al.,
622	2018). The root mean square error (RMSE) of the measured
623	radiation is better than 0.2 K (Susskind et al., 2003). Moreover,
624	global atmospheric profiles can be detected every day. Due to
625	radiometer noise and faults, there are currently only 2047 effective
626	channels. However, compared with previous infrared detectors,
627	AIRS boasts a significant improvement in both the number of
628	channels and spectral resolution (Aumann, 1994; Huang et al., 2005;
629	Li et al., 2005).
630	The root mean square error of an AIRS infrared channel is shown
631	in Fig. 1, with black spots, indicating that. The measurement error is

not below 0.2K for all the instrument channels possess a 632 measurement error of less than 0.2 K. There are a few channels with 633 extremely large measurement errors, which reduce the accuracy of 634 prediction to some extent. Among them, some extremely large 635 measurement errors reduce the accuracy of prediction to some extent 636 (Susskind et al., 2003). At present, more than 300 channels have not 637 been used because their errors exceed 1 K. If data from these 638 channels were to be used for retrieval, the accuracy of the retrieval 639 could be reduced. Therefore, it is necessary to select a group of 640 channels to improve the calculation efficiency and retrieval quality. 641 In this paper we study channel selection for temperature profile 642 retrieval by AIRS. 643

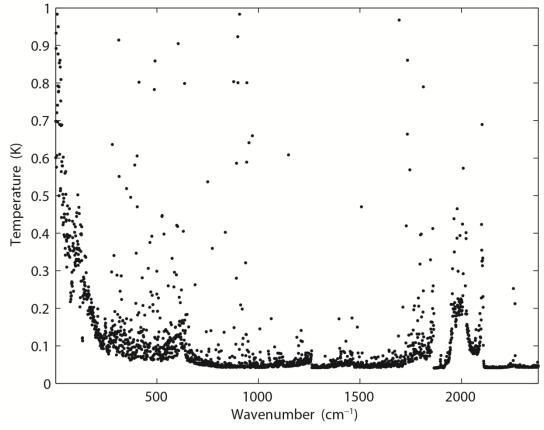


Figure 1. Root mean square error of AIRS infrared channel (blackspots).

647

For the calculation of radiative transfer and the weighting function 648 matrix, K, the RTTOV (Radiative Transfer for TOVS) v12 fast 649 radiative transfer model is used. Although initially developed for the 650 TOVS (TIROS Operational Vertical Sounder) radiometers, RTTOV 651 can now simulate around 90 different satellite sensors measuring in 652 the MW (microwave), IR (infrared) and VIS (visible) regions of the 653 spectrum (Saunders et al., 2018). The model allows rapid 654 simulations (1 ms for 40 channel ATOVS (Advanced TOVS) on a 655 desktop PC) of radiances for satellite visible, infrared, or microwave 656 nadir scanning radiometers given atmospheric profiles of 657 temperature and trace gas concentrations, and cloud and surface 658 properties. The only mandatory gas included as a variable for 659 RTTOV v12 is water vapor. Optionally, ozone, carbon dioxide, 660 nitrous oxide, methane, carbon monoxide, and sulfur dioxide can be 661 included, with all other constituents assumed to be constant. RTTOV 662 can accept input profiles on any defined set of pressure levels. The 663 majority of RTTOV coefficient files are based on the 54 levels (see 664 Table A1 in Appendix A), rankingin the range from 1050 hPa to 0.01 665 hPa, though coefficients for some hyperspectral sounders are also 666

available on 101 levels.

In order to correspond to the selected profiles, the atmosphere is 668 divided into 137 layers, each of which contains corresponding 669 atmospheric characteristics, such as temperature, pressure, and the 670 humidity distribution. Each element in the weighting function matrix 671 can be written as  $\partial yi/\partial xj$ . The subscript i is used to identify the 672 satellite channel, and the subscript j is used to identify the 673 atmospheric variable. Therefore,  $\partial yi/\partial xj$  indicates the variation in 674 brightness temperature in a given satellite channel, when a given 675 atmospheric variable in a given layer changes. We are thus able to 676 establish which layer of the satellite channel is particularly sensitive 677 to which atmospheric characteristic (temperature, various gas 678 contents) in the vertical atmosphere. The RTTOV K (the K mode), 679 is used to calculate the matrix H(X0) (Eq. (1)) for a given 680 atmospheric profile characteristic. 681

682

## **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

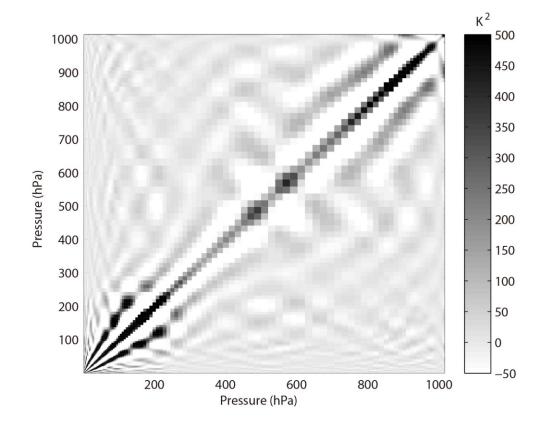
NWPSAF 1DVar (one-dimensional variational analysis) scheme, in
accordance with the requirements of the NWPSAF (Saunders et al.,
2018). Second, 324 channels were selected using the information
capacity method. This method was adopted by Du et al. (2008)
without the consideration of layering. PCS is short 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,  $S_{\varepsilon}$ , in the experiment is 702 provided by NWP SAF 1Dvar. In general, it can be converted to a 703 diagonal matrix, the elements of which are the observation error 704 standard deviation of each hyperspectral detector channel, which is 705 the square of the root mean square error for each channel. The root 706 mean square error of the AIRS channels is shown in Fig. 1. The error 707 covariance matrix of the background,  $S_a$ , is calculated using 5000 708 samples of the IFS-137 data provided by the ECMWF dataset (The 709 detailed information will be introduced in Sect. 4). The last access 710

<sup>711</sup> date is April 26th, 2019 (download address:

- 712 <u>https://www.nwpsaf.eu/site/update-137-level-nwp-profile-dataset/</u>,
- 2019). The covariance matrix of temperature is shown in Fig. 2. The
- results are consistent with the previous study by Du et al. (2008).
- 715

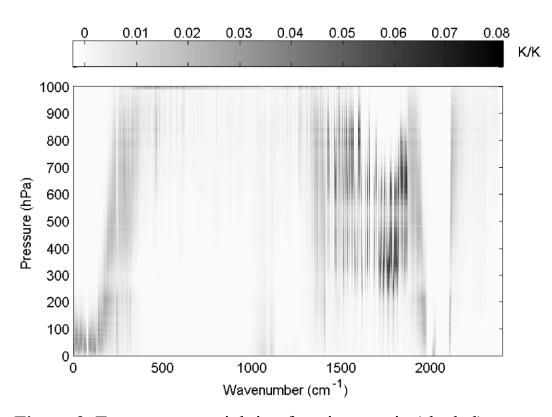


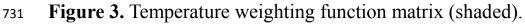
716

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

718

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.





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

- <sup>736</sup> 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.

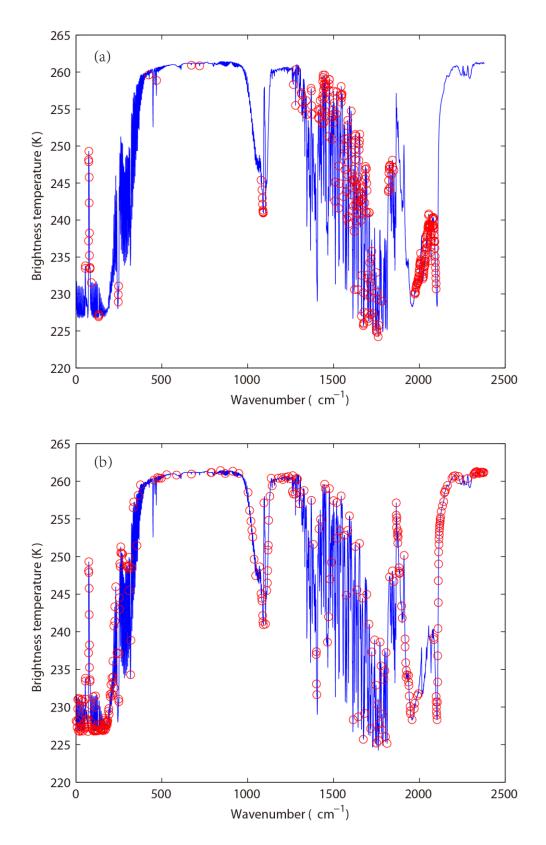


Figure 4. The distribution of different channel selection methods without considering layering in the AIRS brightness temperature

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

746	Without considering layering, the main differences between the
747	324 channels selected by PCS and NCS are as follows: (1) When the
748	wavenumber approaches 1000, the wavelength is 10 $\mu$ m (1/1000-
749	$em^{-1}$ ). Near this <u>Near 10 µm</u> band, fewer channels are selected by
750	PCS because the retrieval of ground temperature is considered by
751	NCS; (2) When the wavenumber is near 1200, the wavelength is 9
752	$\mu m (1/1200 \text{ cm}^{-1})$ . Near this Near 9 $\mu m$ band, no channels are
753	selected by PCS because the retrieval of O3 is not considered in this
754	paper; (3) When the wavenumber approaches 1500, the wavelength-
755	is 6.7 μm (1/1500 cm <sup>-1</sup> ). As is known, the spectral range from 6 μm
756	to 7 $\mu$ m corresponds to water vapor absorption bands, but fewer
757	channels are selected by NCS; (4) When the wavenumber is close to-
758	2000, it derives a wavelength of 5 μm (1/2000 cm <sup>-1</sup> ), which Near 5
759	$\mu$ m band, it includes 4.2 $\mu$ m for N <sub>2</sub> O and 4.3 $\mu$ m for CO <sub>2</sub> absorption
760	bands. As is shown in Fig. 4, fewer channels are selected by PCS in
761	those bands. PCS is favorable for atmospheric temperature
762	observation in the high temperature zone. Because 4.2 $\mu$ m and 4.3
763	$\mu$ m bands are sensitive to high temperature, the higher temperature is,
764	thea better observation can be obtained for higher temperatures; (5) In-
765	the near infrared area, the wavenumber exceeds 2200, deriving a

wavelength of less thanNear 4 µm (1/2000 cm<sup>-1</sup>). Aband, 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 769 paperstudy only takes the temperature profile retrieval into 770 consideration, so the channel combination of PCS is inferior to that 771 of NCS for the retrieval of surface temperature and the O<sub>3</sub> profile. 772 The advantages of the channel selection method based on 773 information content in this paper are mainly reflected in: (1) 774 Stratosphere and mesosphere is less affected by the ground surface, 775 so the retrieval result of PCS is better than that of NCS. (2) Due to 776 the method selected in this paper there are more channels at  $4.2 \,\mu m$ 777 for  $N_2O$  and 4.3 µm for  $CO_2$  absorption bands; the channel 778 combination of PCS is better than that of NCS for atmospheric 779 temperature observation to theat higher temperature. 780 By comparing channel selection without considering layering, 781 we note the general advantages and disadvantages of PCS and NCS 782 for the retrieval of temperature and can improve the channel 783 selection scheme. First, the retrieval of the temperature profile for 784 324 channels selected by PCS is obtained. The relationship between 785

the number of iterations and the ARI is shown in Fig. 5.

787

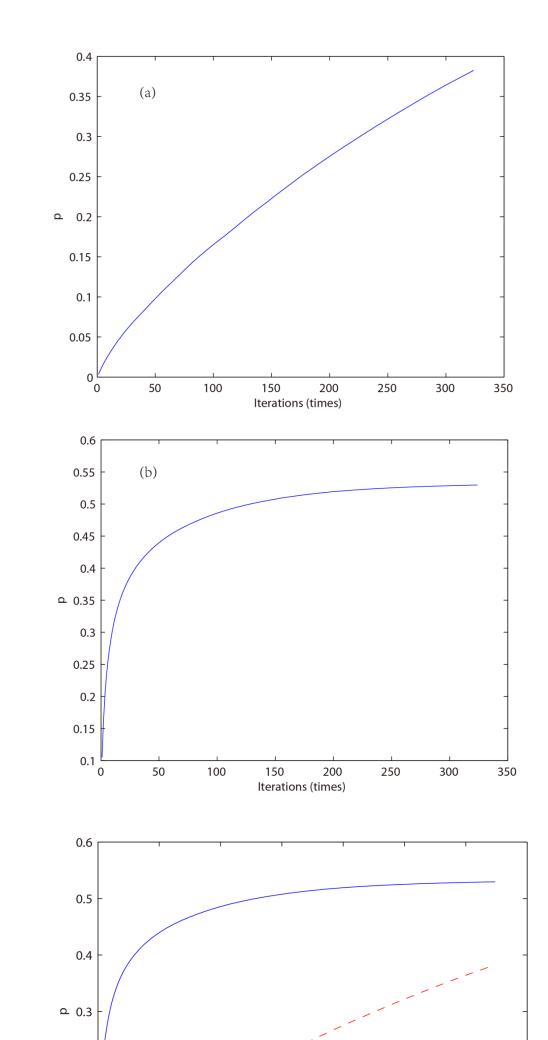


Figure 5. The relationship between the number of iterations and ARI.
(a)Blue line represents the result of ICS. Red dotted line stands for
the result of PCS. (b) ICS.

793

The ARI for PCS tends to be 0.38 and is not convergent, so the 794 PCS method needs to be improved. In this paper, the atmosphere is 795 divided into 137 layers, and based on the information content and 796 iteration, 324 channels are selected for each layer. Then, the 797 temperature profile of each layer can be retrieved based on statistical 798 inversion (see at Sect. 4). The relationship between the number of 799 iterations and the ARI for ICS is shown in Fig. 5b. When the number 800 of iterations approaches 100, the ARI of ICS tends to be stable, and 801 reach toreaches 0.54. Thus, in terms of the ARI and convergence, the 802 ICS method is better than that of PCS. 803

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.

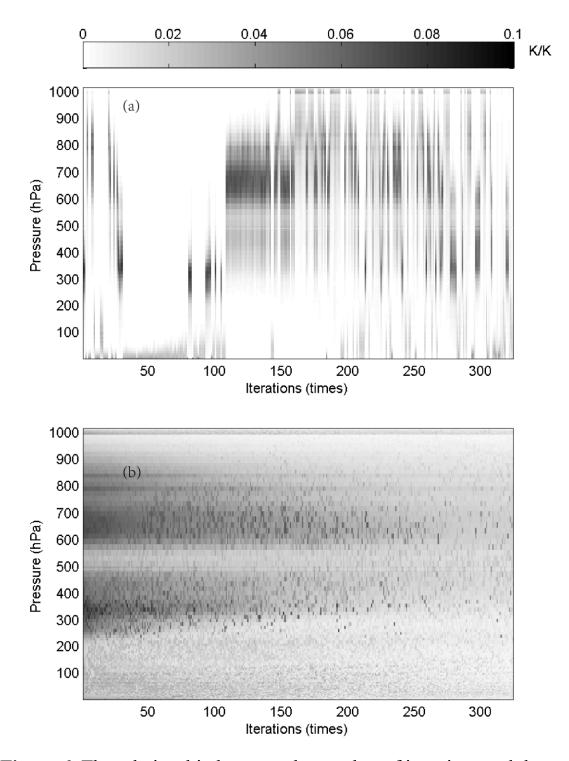


Figure 6. The relationship between the number of iterations and the
weighting function of the top 324 selected channels (shaded). (a)
ICS. (b) PCS.

As illustrated in Fig. 6, in the first 100 iterations, the distribution 814 of the temperature weighting function for PCS is relatively scattered; 815 it does not reflect continuity between the adjacent layers of the 816 atmosphere. Besides, the ICS result is better than that of PCS, 817 showing that: (1) the distribution of the temperature weighting 818 function is more continuous and reflects the continuity between 819 adjacent layers of the atmosphere; (2) regardless of the number of 820 iterations, the maximum value of the weighting function is stable 821 near 300–400 hPa and 600–700 hPa, without scattering, which is 822 closer to the situation in real atmosphere. 823

824

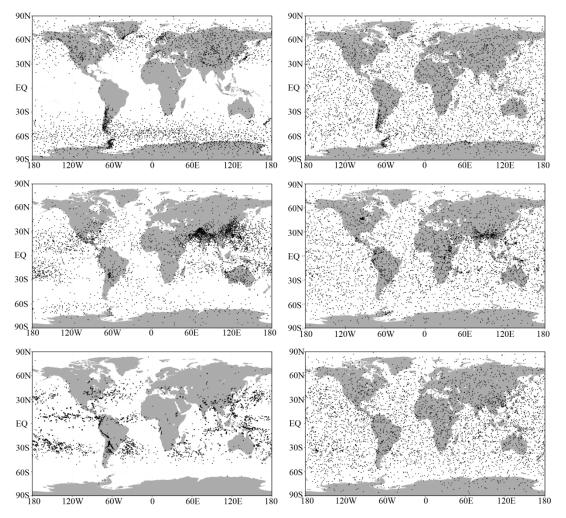
# **4. Statistical multiple regression experiment**

#### **4.1 Temperature profile database**

A new database including a representative collection of 25,000 827 atmospheric profiles from the European Centre for Medium-range 828 Weather Forecasts (ECMWF) was used for the statistical inversion 829 experiments. The profiles were given in a 137-level vertical grid 830 extending from the surface up to 0.01 hPa. The database was divided 831 into five subsets focusing on diverse sampling characteristics such as 832 temperature, specific humidity, ozone mixing ratio, cloud 833 condensates, and precipitation. In contrast with earlier releases of the 834

ECMWF diverse profile database, the 137-level database places 835 greater emphasis on preserving the statistical properties of sampled 836 distributions produced by the Integrated Forecasting System (IFS) 837 (Eresmaa and McNally, 2014; Brath et al., 2018). IFS-137 spans the 838 period from September 1, 2013 to August 31, 2014. There are two 839 operational analyses each day (at 00z and 12z), and approximately 840 13 000 atmospheric profiles over the ocean. The pressure levels 841 adopted for IFS-137 are shown in Table A2 (see Table A2 in 842 Appendix A). 843

The locations of selected profiles of temperature, specific 844 humidity, and cloud condensate subsets of the IFS-91 and IFS-137 845 databases are plotted on the map in Fig. 7. In the IFS-91 database, 846 the sampling is fully determined by the selection algorithm, which 847 makes the geographical distributions very inhomogeneous. Selected 848 profiles represent those regions where gradients of the sampled 849 variable are the strongest: in the case of temperature, mid- and 850 high-latitudes dominate, while humidity and cloud condensate 851 subsets concentrate at low latitudes. However, the IFS-137 database 852 shows a much more homogeneous spatial distribution in all the 853 sampling subsets, which is a consequence of the randomized 854 selection. 855



**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
- 860 <u>https://www.nwpsaf.eu/site/update-137-level-nwp-profile-dataset/</u>,
- 861 2019).

- <sup>863</sup> 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.

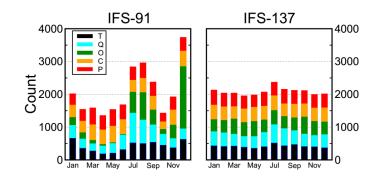


Figure 8. Distribution of profiles within the calendar months in 869 IFS-91 (left) and IFS-137 (right) databases. Different subsets are 870 shown in different colors. Black parts stand for temperature. Blue 871 parts represent specific humidity. Green parts indicate ozone subset. 872 Orange parts stand for cloud condensate. Red parts represent 873 precipitation. The last access date is April 26th, 2019. (from 874 https://www.nwpsaf.eu/site/update-137-level-nwp-profile-dataset/, 875 2019). 876

877

# 878 **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

882 follows:

(1) 5000 profiles and their corresponding surface factors, 883 including surface air pressure, surface temperature, 2 m temperature, 884 2 m specific humidity, 10 m wind speed- are put into the RTTOV 885 mode. Then, the simulated AIRS spectra are obtained. 886

(2) The retrieval of temperature is carried out in accordance with 887 Eq. (23). The 5000 profiles are divided into two groups. The first 888 group of 2500 profiles is used to obtain the regression coefficient, 889 and the second group of 2500 is used to test the result. 890

(3) Verification of the results. The test is carried out based on the 891 standard deviation between the retrieval value and the true value. 892

893

894

# 4.3 Results and Discussion

For the statistical inversion comparison experiments, the standard 895 deviation of temperature retrieval is shown in Fig. 9. First, because 896 PCS does not take channel sensitivity as a function of height into 897 consideration, the retrieval result of PCS is inferior to that of ICS. 898 Second, by comparing the results of ICS and NCS we found that 899 below 100 hPa, since the method used in this paper considers near 900 ground to be less of an influencing factor, the channel combination 901 of ICS is slightly inferior to that of NCS, but the difference is small. 902

From 100 hPa to10 hPa, the retrieval temperature of ICS in this 903 paper is consistent with that of NCS, slightly better than the channel 904

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 912 improvement value of the ICS and its percentages compared with the 913 PCS and NCS at different heights are shown in Table 1. Because 914 PCS does not take channel sensitivity as a function of height into 915 consideration, the retrieval result of PCS is inferior to that of ICS. In 916 general, the accuracy of the retrieval temperature of ICS is improved. 917 Especially, from 100 hPa to 0.01 hPa, the mean value of ICS is 918 evidently improved by more than 0.5 K which means the accuracy 919 can be improved by more than 11%. By comparing the results of ICS 920 and NCS we found that below 100 hPa, since the method used in this 921 paper considers near ground to be less of an influencing factor, the 922 channel combination of ICS is slightly inferior to that of NCS, but 923 the difference is small. From 100 hPa to 0.01 hPa, the mean value of 924 ICS is improved by more than 0.36 K which means the accuracy can 925 be improved by more than 9.6%. 926

# **Table 1.** The mean improvement value of the ICS and its

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%	

## percentages compared with the PCS and NCS at different heights.

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  $\mu$ m for N<sub>2</sub>O and

 $4.3 \ \mu m$  for CO<sub>2</sub> 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

939 retrieval result is improved.

940

927

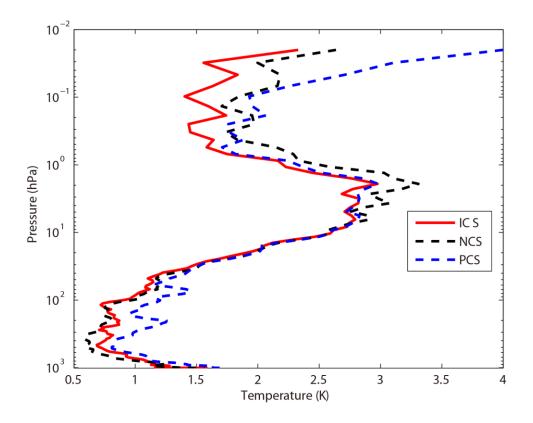




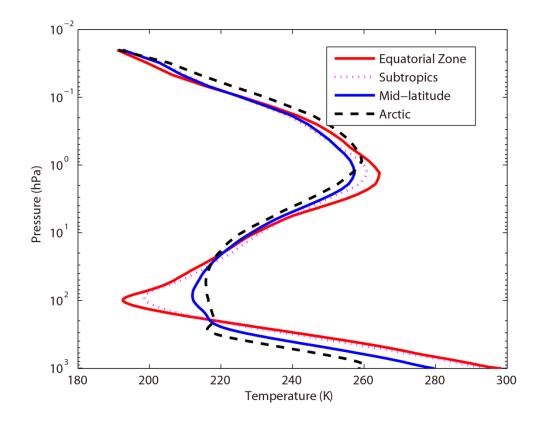
Figure 9. The temperature profile standard deviation of statistical
inversion comparison experiments. Red line indicates the result of
ICS. Black dotted line stands for the result of NCS. Blue dotted line
represents the result of PCS.

# 947 5 Statistical inversion comparison experiments in four typical

```
948 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.



961

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.

#### 968 **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 regressioncoefficient.

(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.

981

# 982 5.2 Results and Discussion

<sup>983</sup> Using statistical inversion comparison experiments in four typical <sup>984</sup> regions, the standard deviation of temperature retrieval is shown in <sup>985</sup> Fig. 11. Generally, the retrieval temperature by ICS is better than <sup>986</sup> that of NCS and PCS. In particular, above 1 hPa (the stratosphere <sup>987</sup> and mesosphere), the standard deviation of atmospheric temperature <sup>988</sup> can be improved by 1 K with PCS and NCS. Thus, ICS shows a

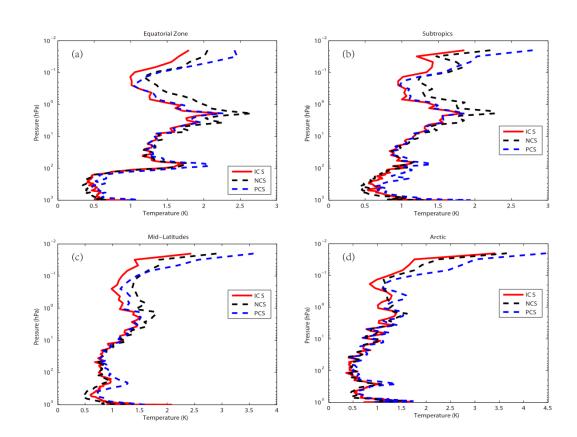


Figure 11. The temperature profile standard deviation of statistical
inversion comparison experiments in four typical regions. Red line
indicates the result of ICS. Black dotted line stands for the result of
NCS. Blue dotted line represents the result of PCS. (a) Equatorial
zone. (b) Subtropics. (c)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

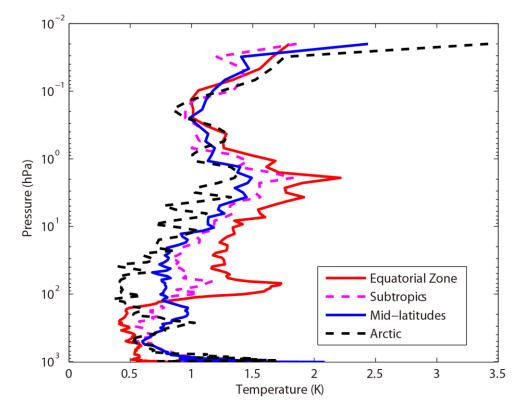


Figure 12. The temperature standard deviation of ICS in four typical
regions. Red line indicates the result of equatorial zone. Pink dotted
line represents the result of Subtropics. Blue line represents the
result of Mid-latitudes. Black dotted line stands for the result of
Arctic.

1002

1009The temperature standard deviations of the ICS in the four typical1010regions are large (Fig. 12). Below100 hPa, due to the high1011temperature in the equatorial zone, the channel combination of ICS1012is better than that of PCS and NCS for atmospheric temperature1013observation to theat higher temperature. The standard deviation is10140.5K. Due to the method selected in this paper there are more

channels at 4.2  $\mu$ m for N<sub>2</sub>O and 4.3  $\mu$ m for CO<sub>2</sub> absorption bands 1015 which has been previously described in Sect. 3. Near the tropopause, 1016 the standard deviation of the equatorial zone increases sharply. It is 1017 also due to the sharp drops in temperature. However, the standard 1018 deviation of the Arctic is still around 0.5K. From 100hPa to 1hPa, 1019 the standard deviation of ICS is 0.5 K to 2K. With the increase of 1020 latitude, the effectiveness considerably increases. According to Fig. 1021 11, ICS takes channel sensitivity as a function of height into 1022 consideration, so its retrieval result is better. 1023

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.

1030

## **1031** 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 1037 exploration, therefore, on and of the near space layer are of great 1038 importance. A new channel selection scheme and method for 1039 hyperspectral atmospheric infrared sounder AIRS data based on 1040 layering is proposed. The retrieval results of ICS concerning the near 1041 space atmosphere are particularly good. Thus, ICS aims to provide a 1042 new and an effective channel selection method for the study of the 1043 near space atmosphere using the hyperspectral atmospheric infrared 1044 sounder. 1045

An improved channel selection method is proposed, based on 1046 information content in this paper. A robust channel selection scheme 1047 and method are proposed, and a series of channel selection 1048 comparison experiments are conducted. The results are as follows: 1049 (1) Since ICS takes channel sensitivity as a function of height into 1050 consideration, the ARI of PCS only tends to be 0.38 and is not 1051 convergent. However, as the 100<sup>th</sup> iteration is approached, the ARI of 1052 ICS tends to be stable, reaching 0.54, while the distribution of the 1053 temperature weighting function is more continuous and closer to that 1054 of the actual atmosphere. Thus, in terms of the ARI, convergence, 1055 and the distribution of the temperature weighting function, ICS is 1056 better than PCS. 1057

1058 (2) Statistical inversion comparison experiments show that the

retrieval temperature of ICS in this paper is consistent with that of 1059 NCS. In particular, from 10 hPa to 0.02 hPa (the stratosphere and 1060 mesosphere), the retrieval temperature of ICS is obviously better 1061 than that of NCS at about 1 K. In general, the accuracy of the 1062 retrieval temperature of ICS is improved. Especially, from 100 hPa 1063 to 0.01 hPa, the accuracy of ICS can be improved by more than 11%. 1064 The reason is that stratosphere and mesosphere are less affected by 1065 the ground surface, so the retrieval result of ICS is better than that of 1066 NCS. Additionally, due to the method selected in this paper there are 1067 more channels at 4.2  $\mu$ m for the N<sub>2</sub>O and at 4.3  $\mu$ m for the CO<sub>2</sub> 1068 absorption bands; the channel combination of ICS is better than that 1069 of NCS for atmospheric temperature observation to theat higher 1070 temperature. 1071

(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.

1076

1077 *Data availability*. The data used in this paper are available from the 1078 corresponding author upon request.

1079

1080 *Appendices* 

1081 Appendix A

**Table A1.** Pressure levels adopted for RTTOV v12 54 pressure level

- 1083 coefficients and profile limits within which the transmittance
- 1084 calculations are valid. Note that the gas units here are ppmv.
- 1085 (From <u>https://www.nwpsaf.eu/site/software/rttov/</u>, RTTOV Users
- 1086 guide, 2019).

Level	Pressure	Tmax	Tmin	Qmax	Qmin	Q2max	Q <sub>2</sub> min	Q <sub>2</sub> Ref
Number	hPa	К	к	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

1089 (in hPa).

Level	pressure	Level	pressure	Level	pressure	Level	pressure	Level	pressure
number	hPa	number	hPa	number	hPa	number	hPa	number	hPa
1	0.02	31	12.8561	61	106.4153	91	424.019	121	934.766
2	0.031	32	14.2377	62	112.0681	92	441.5395	122	943.139
3	0.0467	33	15.7162	63	117.9714	93	459.6321	123	950.908
4	0.0683	34	17.2945	64	124.1337	94	478.3096	124	958.103
5	0.0975	35	18.9752	65	130.5637	95	497.5845	125	964.758
6	0.1361	36	20.761	66	137.2703	96	517.4198	126	970.904
7	0.1861	37	22.6543	67	144.2624	97	537.7195	127	976.573
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	986.603
10	0.4288	40	29.0039	70	167.045	5 100	600.1668	130	991.02
11	0.5496	41	31.3512	71	175.2731	101	621.1624	131	995.082
12	0.6952	42	33.8174	72	183.8344	102	642.0764	132	998.808
13	0.869	43	36.4047	73	192.7389	103	662.8084	133	1002.22
14	1.0742	44	39.1149	74	201.9969	) 104	683.262	134	1005.35
15	1.3143	45	41.9493	75	211.6186	5 105	703.3467	135	1008.22
16	1.5928	46	44.9082	76	221.6146	5 106	722.9795	136	1010.84
17	1.9134	47	47.9915	77	231.9954	107	742.0855	137	1013.2
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	5 110	795.6396		
21	3.6898	51	61.5607	81	277.5852	2 111	812.0847		
22	4.2759	52	65.2695	82	290.0548	8 112	827.7756		
23	4.9262	53	69.1187	83	302.9762	2 113	842.6959		
24	5.6441	54	73.1187	84	316.3607	<b>'</b> 114	856.8376		
25	6.4334	55	77.281	85	330.2202	2 115	870.2004		
26	7.2974	56	81.6182	86	344.5663	3 116	882.791		
27	8.2397	57	86.145	87	359.4111	117	894.6222		
28	9.2634	58	90.8774	88	374.7666	5 118	905.7116		
29	10.372	59	95.828	89	390.645	5 119	916.0815		

*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.

1095

1096 *Competing interests*. The authors declare that they have no conflict 1097 of interest.

1098

Acknowledgements. The study was supported by the National Key 1099 Research Program of China: Development of high-resolution data 1100 assimilation technology and atmospheric reanalysis data set in East 1101 Asia (Research on remote sensing telemetry data assimilation 1102 technology, Grant no. 2017YFC1501802). The study was also 1103 supported by the National Natural Science Foundation of China 1104 (Grant no. 41875045) and Hunan Provincial Innovation Foundation 1105 for Postgraduate (Grant no. CX2018B033 and no. CX2018B034). 1106

1107

#### 1108 **References**

Aires, F., Schmitt, M., Chedin, A., and Scott, N.: The "weighting

smoothing" regularization of MLP for Jacobian stabilization,

- 1111 IEEE. T. Neural. Networks., 10, 1502-1510,
  1112 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.,
- 1125 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.
- 1131 Chahine, M. I.: A general relaxation method for inverse solution of 1132 the full radiative transfer equation, J. Atmos. Sci., 29, 741-747,
  - 63

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 1139 improved initialization inversion method: a high resolution 1140 physical method for temperature retrievals from satellites of the 1141 tiros-n series. J. Appl. Meteor. 24. 128-143. 1142 https://doi.org/10.1175/1520-0450(1985)024<0128:TIIIMA>2.0.C 1143 O;2, 1985. 1144
- 1145 Cyril, C., Alain, C., and Scott, N. A.: Airs channel selection for CO<sub>2</sub>
- and other trace-gas retrievals, Q. J. Roy. Meteor. Soc., 129,

1147 2719-2740, https://doi.org/10.1256/qj.02.180, 2003.

- 1148 Du, H. D., Huang, S. X., and Shi, H. Q.: Method and experiment of
- channel selection for high spectral resolution data, Acta. Physica.
- 1150 Sinica., 57, 7685-7692, 2008.
- 1151 Dudhia, A., Jay, V. L., and Rodgers, C. D.: Microwindow selection
- 1152 for high-spectral-resolution sounders, Appl. Opt. 41, 3665-3673,
- 1153 https://doi.org/10.1364/AO.41.003665, 2002.
- 1154 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 1157 satellite sounding radiances in numerical weather prediction. High 1158 Spectral Resolution Infrared Remote Sensing for Earth's Weather 1159 Springer, Studies, and Climate Berlin, Heidelberg. 1160 https://doi.org/10.1007/978-3-642-84599-4 25, 1993. 1161
- 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,

1168 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.: 1177 Evaluation of cloud-cleared radiances for numerical weather 1178 cloud-contaminated sounding prediction and applications. 1179 Atmospheric and Environmental Remote Sensing Data Processing 1180 Atmospheric Numerical Utilization: Prediction and and 1181 Monitoring, I. S. Photonics., Environmental О. 1182 https://doi.org/10.1117/12.613027, 2005. 1183

Kuai, L., Natraj, V., Shia, R. L., Miller, C., and Yung, Y. L.: Channel
selection using information content analysis: a case study of CO<sub>2</sub>
retrieval from near infrared measurements. J. Q. S. Radiative.
Transfer., 111, 1296-1304,

1188 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., 1189 Huang, H. L., and Achtor, T. H.: Global soundings of the 1190 atmosphere from ATOVS measurements: the algorithm and 1191 validation, J. Appl. Meteor., 39, 1248-1268, 1192 https://doi.org/10.1175/1520-0450(2000)039<1248:GSOTAF>2.0. 1193 CO:2, 2000. 1194

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
- 1197 MODIS, IEEE. Trans. GRS. , 43, 1266-1278, http://dx.doi.org/
- 1198 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.
- 1202 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.
- 1206 Menke, W.: Geophysical Data Analysis: Discrete Inverse Theory,
- 1207 Acad. Press., Columbia University, New York,
- 1208 https://doi.org/10.1016/B978-0-12-397160-9.00019-9, 1984.
- 1209 Menzel, W. P., Schmit, T. J., Zhang, P. and Li, J.: Satellite-based
- atmospheric infrared sounder development and applications, Bull.
- 1211 Amer. Meteor. Soc., 99, 583–603,
- 1212 https://doi.org/10.1175/BAMS-D-16-0293.1, 2018.
- 1213 Prunet, P., Thépaut J. N., and Cass, V.: The information content of
- 1214 clear sky IASI radiances and their potential for numerical weather
- 1215 prediction, Q. J. Roy. Meteor. Soc., 124, 211-241,
- 1216 https://doi.org/10.1002/qj.49712454510, 2010.
- 1217 Xu, Q.: Measuring information content from observations for data
- assimilation: relative entropy versus shannon entropy difference,
- 1219 Tellus. A., 59, 198-209,
- 1220 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,
- 1231 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.
- 1235 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,
- <sup>1239</sup> 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,

- 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,
- 1259 1, 36-42, http://dx.doi.org/10.3969/j.issn.1674-7097.2011.01.005,
- 1260 2011.
- 1261
- 1262
- 1263
- 1264

<sup>41, 390-409,</sup> https://doi.org/10.1109/TGRS.2002.808236, 2003.

1265	(c) The marked-up manuscript version
1266	A channel selection method for hyperspectral
1267	atmospheric infrared sounders based on
1268	layering
1269	Shujie Chang <sup>1, 2,3</sup> , Zheng Sheng <sup>1,2</sup> , Huadong Du <sup>1,2</sup> , Wei Ge <sup>1,2</sup> and
1270	Wei Zhang <sup>1,2</sup>
1271	<sup>1</sup> College of Meteorology and Oceanography, National University of
1272	Defense Technology, Nanjing, China
1273	<sup>2</sup> Collaborative Innovation Center on Forecast and Evaluation of
1274	Meteorological Disasters, Nanjing University of Information
1275	Science and Technology, Nanjing, China
1276	<sup>3</sup> South China Sea Institute for Marine Meteorology, Guangdong
1277	Ocean University, Zhanjiang, China
1278	
1279	Correspondence: Zheng Sheng (19994035@sina.com)
1280	
1281	Abstract. This study introduces an effective channel selection
1282	method for hyperspectral infrared sounders. The method is
1283	illustrated for the Atmospheric InfraRed Sounder (AIRS) instrument.
1284	The results are as follows: (1) Using the improved channel selection
1285	(ICS), the atmospheric retrievable index is more stable, the value
1286	reaching 0.54. The coverage of the weighting functions is more

evenly distributed over height with this method; (2) Statistical 1287 inversion comparison experiments show that the accuracy of the 1288 retrieval temperature, using the improved channel selection method 1289 in this paper, is consistent with that of 1Dvar channel selection. In 1290 the stratosphere and mesosphere especially, from 10 hPa to 0.02 hPa, 1291 the accuracy of the retrieval temperature of our improved channel 1292 selection method is improved by about 1 K. Also at lower heights, 1293 the accuracy of the retrieval temperature of ICS is improved; (3) 1294 Statistical inversion comparison experiments for four different 1295 regions illustrate latitudinal and seasonal variations and better 1296 performance of ICS compared to the NWP Channel Selection (NCS) 1297 and Primary Channel Selection (PCS) methods. The ICS method 1298 shows potential for future applications. 1299

1300

### 1301 **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

1309	have taken us to a new stage in our understanding of weather
1310	systems and related phenomena (Fang, 2014). From the perspective
1311	of vertical atmospheric observation, satellite instruments are
1312	developing rapidly. In their infancy, the traditional infrared
1313	measurement instruments for detecting atmospheric temperature and
1314	moisture profiles, such as TOVS (Smith et al., 1991) or HIRS in
1315	ATOVS (Chahine, 1972; Li et al., 2000; Liu, 2007), usually
1316	employed filter spectrometry. Even though such instruments have
1317	played an important role in improving weather prediction, it is
1318	difficult to continue to build upon improvements in terms of
1319	observation accuracy and vertical resolution due to the limitation of
1320	low spectral resolution. By using this kind of filter-based
1321	spectroscopic measurement instrument, therefore, it is difficult to
1322	meet today's needs in numerical weather prediction (Eyre et al.,
1323	1993; Prunet et al., 2010; Menzel et al., 2018). To meet this
1324	challenge, a series of plans for the creation of high-spectral
1325	resolution atmospheric measurement instruments has been executed
1326	in the United States and in Europe in recent years: One example is
1327	the AIRS (Atmospheric InfraRed Sounder) on the Earth Observation
1328	System, "Aqua", launched on May 4, 2002 from the United States.
1329	AIRS has 2378 spectral channels providing sensitivity from the
1330	ground to up to about 65 km of altitude (Aumann et al., 2003;

Hoffmann and Alexander, 2009; Gong et al., 2011). The United 1331 States and Europe, in 2010 and in 2007, also installed the CRIS 1332 (Cross-track Infrared Sounder) and the IASI (Infrared Atmospheric 1333 Sounding Interferometer) on polar-orbiting satellites. 1334 China also devotes great importance to the development of such 1335 advanced sounding technologies. In the early 1990s, the National 1336 Satellite Meteorological Center began to investigate the principles 1337 and techniques of hyperspectral resolution atmospheric observations. 1338 China's development of interferometric atmospheric vertical 1339 detectors eventually led to the launch of Fengyun No. 3, on May 27, 1340 2008, and Fengyun No. 4 on December 11, 2016, both of which 1341 were equipped with infrared atmospheric instruments. How best to 1342 use the hyperspectral resolution observation data obtained from 1343 these instruments, to obtain reliable atmospheric temperature and 1344 humidity profiles, is an active area of study in atmospheric inversion 1345 theory. 1346

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. 1353 Due to the large number of channels and data supported by such 1354 instruments today (such as AIRS with 2378 channels and IASI with 1355 8461 channels), it has proven extremely cumbersome to store, 1356 transmit, and process such data. Moreover, there is often a close 1357 correlation between the channel, causing an ill-posedness of the 1358 inversion, potentially compromising accuracy of the retrieval 1359 product based on hyperspectral resolution data. 1360 However, hyperspectral detectors have many channels and 1361 provide real-time mode prediction systems with vast quantities of 1362 data, which can significantly improve prediction accuracy. But, if all 1363 the channels are used to retrieve data, the retrieval time considerably 1364 increases. Even more problematic are the glut of information 1365 produced, and the unsuitability of the calculations for real-time 1366 forecasting. Concurrently, the computer processing power must be 1367 large enough to meet the demands of simulating all the channels 1368 simultaneously within the forecast time. In order to improve the 1369 calculation efficiency and retrieval quality, it is very important to 1370 properly select a set of channels that can provide as much 1371 information as possible. 1372 Many researchers have studied channel selection algorithms. Menke

1373 Many researchers have studied channel selection algorithms. Menke1374 (1984) first chose channels using a data precision matrix method.

Aires et al. (1999) made the selection using the Jacobian matrix, 1375 which has been widely used since then (Aires et al., 2002; Rabier et 1376 al., 2010). Rodgers (2000) indicated that there are two useful 1377 quantities in measuring the information provided by the observation 1378 data: Shannon information content and degrees of freedom. The 1379 concept of information capacity then became widely used in satellite 1380 channel selection. In 2007, Xu (2007) compared the Shannon 1381 information content with the relative entropy, analyzing the 1382 information loss and information redundancy. In 2008, Du et al. 1383 (2008) introduced the concept of the atmospheric retrievable index 1384 (ARI) as a criterion for channel selection, and in 2010, Wakita et al. 1385 (2010) produced a scheme for calculating the information content of 1386 the various atmospheric parameters in remote sensing using 1387 Bayesian estimation theory. Kuai et al. (2010) analyzed both the 1388 Shannon information content and degrees of freedom in channel 1389 selection when retrieving CO<sub>2</sub> concentrations using thermal infrared 1390 remote sensing and indicated that 40 channels could contain 75% of 1391 the information from the total channels. Cyril et al. (2003) proposed 1392 the optimal sensitivity profile method based on the sensitivity of 1393 different atmospheric components. Lupu et al. (2012) used degrees 1394 of freedom for signals (DFS) to estimate the amount of information 1395 contained in observations in the context of observing system 1396

experiments. In addition, the singular value decomposition method 1397 has also been widely used for channel selection (Prunet et al., 2010; 1398 Zhang et al., 2011; Wang et al., 2014). In 2017, Chang et al. (2017) 1399 selected a new set of Infrared Atmospheric Sounding Interferometer 1400 (IASI) channels using the channel score index (CSI). Richardson et 1401 al. (2018) selected 75 from 853 channels based on the high 1402 spectral-resolution oxygen A-band instrument on NASA's Orbiting 1403 Carbon Observatory-2 (OCO-2), using information content analysis 1404 to retrieve the cloud optical depth, cloud properties, and position. 1405 Today's main methods for channel selection use only the 1406 weighting function to study appropriate numerical methods, such as 1407 the data precision matrix method (Menke, 1984), singular value 1408 decomposition method (Prunet et al., 2010; Zhang et al., 2011; Wang 1409 et al., 2014), and the Jacobi method (Aires et al., 1999; Rabier et al., 1410 2010). The use of the methods allows sensitive channels to be 1411 selected. The above-mentioned studies also take into account the 1412 sensitivity of each channel to atmospheric parameters during channel 1413 selection, while ignoring some factors that impact retrieval results. 1414 The accuracy of retrieval results depends not only on the channel 1415 weighting function but also on the channel noise, background field, 1416 and the retrieval algorithm. 1417

1418 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 1422 practice, and the concept of information content itself does consider 1423 all the height dependencies of the kernel matrix K (Rodgers, 2000). 1424 However, earlier works have neglected the height dependencies of K 1425 for simplicity. This paper uses the atmospheric retrievable index 1426 (ARI) as the index, which is based on information content (Du et al., 1427 2008; Richardson et al. 2018). Channel selection is made at different 1428 heights, and an effective channel selection scheme is proposed 1429 which fully considers various factors, including the influence of 1430 different channels on the retrieval results at different heights. This 1431 ensures the best accuracy of the retrieval product when using the 1432 selected channel. In addition, statistical inversion comparison 1433 experiments are used to verify the effectiveness of the method. 1434 1435

### 1436 2 Channel selection indicator, scheme and method

1437 **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

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

1443

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

1445

where  $S_a$  is the error covariance matrix of the background or the 1446 estimated value of atmospheric profile,  $S_{\varepsilon}$  represents the 1447 observation error covariance matrix of each hyperspectral detector 1448 channel,  $\hat{S} = (S_a - S_a K^T (K S_a K^T + S_{\varepsilon})^{-1} K S_a)$  denotes the 1449 covariance matrix after retrieval, K is the weighting function matrix. 1450 In order to describe the accuracy of the retrieval results visually 1451 and quantitatively, the atmospheric retrievable index (ARI), p, (Du et 1452 al., 2008) is defined as follows: 1453

1454

1455 
$$p = 1 - \exp(\frac{1}{2n} ln |\hat{S}S_a^{-1}|),$$
 (2)

1456

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  $|\hat{S}S_a^{-1}| = (1-p)^{2n}$  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 ofeach element cannot be 1-p. Therefore, p defined by Eq. (1) is

actually an overall evaluation of the retrieval result.

1465

#### 1466 **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 1472 and N satellite channels. Selecting n from N channels, there will be 1473  $C_N^n$  combinations in each layer, leading  $C_N^n$  calculations to get  $C_N^n$ 1474 kinds of p results. Furthermore, there are M layers in the vertical 1475 direction of the atmosphere. Therefore, the entire atmosphere must 1476 be calculated  $\mathbf{M} \cdot \mathbf{C}_{\mathbf{N}}^{n}$  times. However, the calculation  $\mathbf{M} \cdot \mathbf{C}_{\mathbf{N}}^{n}$  times 1477 will be particularly large, which makes this approach impractical in 1478 calculating p for all possible combinations. Therefore, it is necessary 1479 to design an effective calculation scheme, and such a scheme, i.e., a 1480 channel selection method, using iteration is proposed, called the 1481 "sequential absorption method" (Dudhia et al., 2002; Du et al., 2008). 1482 The method's main function is to select ("absorb") channels one by 1483

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:

1487 (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:

1492 
$$\hat{S} = S_a - S_a k^T (s_{\varepsilon} + k S_a k^T)^{-1} k S_a.$$
 (3)

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

1495 
$$\hat{S} = \left(I - \frac{S_a k^T k}{\left(s_{\varepsilon} + k S_a k^T\right)}\right) S_a = \left(I - \frac{\left(k S_a\right)^T k}{\left(s_{\varepsilon} + k\left(k S_a\right)^T\right)}\right) S_a.$$
(4)

1496 Substituting Eq. (4) into Eq. (2) gives:

1497 
$$\mathbf{p} = 1 - \exp(\frac{1}{2n} ln(\left|I - \frac{(kS_a)^T k}{(s_{\varepsilon} + k(kS_a)^T)}\right|)).$$
 (5)

1498

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

1504 Define 
$$R = S_{\varepsilon}^{1/2} K S_{a}^{1/2}$$
. (6)

The matrix R can then be regarded as a weighting function matrix, 1506 normalized by the observed error and a priori uncertainty. A row 1507 vector of R,  $r = s_{\varepsilon}^{-1/2} k S_a^{1/2}$ , represents the normalized weighting 1508 function matrix of a single channel. Substituting r into Eq. (5) gives: 1509 1510

1511 
$$p = 1 - \exp(\frac{1}{2n}ln\left(\left|I - \frac{rr^{T}}{1 + r^{T}r}\right|\right)).$$
 (7)

1512

For arbitrary row vectors, a and b, using the matrix property 1513  $det(I + a^{T}b) = 1 + ba^{T}$ , the new expression for p is: 1514

1515

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

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

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

(8)

1517

(III) Iteration in a single layer: 1519

First, the iteration in a single layer requires the calculation of R. 1520 Using  $S_a$ ,  $S_\epsilon$ , K and Eq. (6), R can be calculated. Second, using Eq. 1521 (8), p of each candidate channel can be calculated. Moreover, the 1522 channel corresponding to maximum p is the selected channel for this 1523 iteration. After a channel has been selected, according to Eq. (3) we 1524

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.

<sup>1528</sup> When selecting n from N channels, it is necessary to calculate <sup>1529</sup>  $(N-n/2)n \approx Nn p$  values, which is much smaller than  $C_N^n$ . In addition <sup>1530</sup> to high computational efficiency by using this method, another <sup>1531</sup> advantage is that all channels can be recorded in the order in which <sup>1532</sup> they are selected. In the actual application, if n' channels are <sup>1533</sup> needed, and n' < n, we will not need to select the channel again, <sup>1534</sup> but record the selected channel only.

#### 1535 (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

1542 corresponding height in the retrieved profiles.

1543

## 1544 **2.3 Statistical inversion method**

1545 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 1547 model which uses a large number of satellite measurements and 1548 atmospheric parameters to match samples and calculate their 1549 correlation coefficient. Then, based on the correlation coefficient, the 1550 required parameters of the independent measurements obtained by 1551 the satellite are retrieved. Because the method does not directly solve 1552 the radiation transfer equation, it has the advantages of fast 1553 calculation speed. In addition, the solution is numerically stable, 1554 which makes it one of the highest precision methods (Chedin et al., 1555 1985). Therefore, the statistical inversion method will be used for 1556 our channel selection experiment and a regression equation will be 1557 established. 1558

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

1562

1563 
$$\mathbf{T} = T^* \cdot A,\tag{9}$$

1564

$$1565 T_b = T_b^* \cdot A, (10)$$

1566

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

1569	A and B stand for the corresponding expansion coefficient vectors of		
1570	temperature (humidity) and brightness temperature.		
1571	Using the least squares method and the orthogonal property, the		
1572	coefficient conversion matrix, V, is introduced:		
1573			
1574	$\mathbf{A} = \mathbf{V} \cdot \mathbf{B},\tag{11}$		
1575			
1576	where $V = AB^T (BB^T)^{-1}$ . (12)		
1577			
1578	Using the orthogonality, we get:		
1579			
1580	$\mathbf{B} = (T_b^*)^T T_b, \tag{13}$		
1581			
1582	$\mathbf{A} = (T^*)^T T. \tag{14}$		
1583			
1584	For convenience, the anomalies of the state vector (atmospheric		
1585	temperature), T, and the observation vector (brightness temperature),		
1586	$T_b$ , are taken:		
1587			
1588	$\widehat{\mathbf{T}} = \overline{\mathbf{T}} + \widehat{\mathbf{T}}' = \overline{\mathbf{T}} + \mathbf{G}\mathbf{T}_{\mathbf{b}}' = \overline{\mathbf{T}} + \mathbf{G}(\mathbf{T}_{\mathbf{b}} - \overline{\mathbf{T}_{\mathbf{b}}}), \qquad (15)$		
1589			
1590	where $\widehat{T}$ stands for the retrieval atmospheric temperature. $\overline{T}$ and		

 $\overline{T_b}$  are the corresponding average values of the elements,

respectively.  $\widehat{T}'$  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:

1597 
$$T' = (t'_1, t'_2, \dots, t'_k),$$
 (16)

1599 
$$T_{b}' = (t_{b1}', t_{b2}', \cdots, t_{bk}').$$
 (17)

1601 Define the inversion error matrix as:

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

1605 The retrieval error covariance matrix is:

$$S_{\delta} = \frac{1}{k - n - 1} \delta \delta^{T}$$

$$= \frac{1}{k - n - 1} (T' - GT_{b}') (T' - GT_{b}')^{T}$$

$$= \frac{k - 1}{k - n - 1} (S_{e} - G^{T}S_{xy} - S_{xy}G^{T} + GS_{y}G^{T}), \qquad (19)$$

1610 where

1612 
$$S_e = \frac{1}{k-1} T' T'^{T}$$
,  
1613  $S_y = \frac{1}{k-1} T_b' T_b^{T}^{T}$ ,  
1614  $S_{xy} = \frac{1}{k-1} T' T_b^{T}^{T}$ . (20)

S<sub>e</sub> stands for the sample covariance matrix of T, S<sub>y</sub> 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<sub> $\delta$ </sub>, 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:

1623

1624 
$$\delta^2 = \operatorname{tr}(S_{\delta}) = \min.$$
 (21)

1625

Taking a derivative of Eq. (21) with respect to G,  $\frac{\partial}{\partial G} \operatorname{tr}(S_{\delta}) = 0 =$ 1627 (-2S<sub>xy</sub> + 2GS<sub>y</sub>), which means that:

1628

1629 
$$G = S_{xy}S_y^{-1}$$
. (22)

1630

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

1634 
$$\widehat{T} = \overline{T} + S_{xy}S_y^{-1}(T_b - \overline{T_b}).$$
(23)

1635

1636	It should be noted that the least squares solution obtained here
1637	aims to minimize the sum of the error variance for each element in
1638	the atmospheric state vector after retrieval for several times. At
1639	present, statistical multiple regression is widely used in the retrieval
1640	of atmospheric profiles based on atmospheric remote sensing data.
1641	As long as there are enough data, $S_{xy}$ and $S_y$ can be determined.
1642	

**3. Channel selection experiment** 

#### 1644 **3.1 Data and model**

1645 The Atmospheric Infrared Sounder (AIRS) is primarily designed to

1646 measure the Earth's atmospheric water vapor and temperature

1647 profiles on a global scale (Aumann et al., 2003; Susskind et al.,

1648 2003). AIRS is a continuously operating cross-track scanning

sounder, consisting of a telescope that feeds an echelle spectrometer.

1650 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

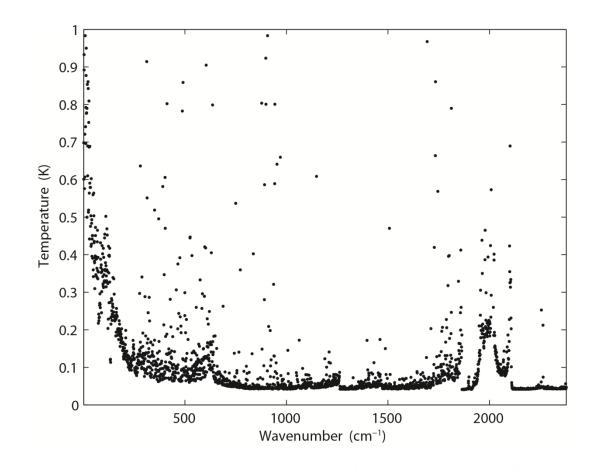
to 4.61  $\mu$ m, 6.20  $\mu$ m to 8.22  $\mu$ m, and 8.8  $\mu$ m to 15.4  $\mu$ m. The

footprint size is 13.5 km. The spectral range includes 4.3  $\mu$ m and

1654 15.5  $\mu$ m for important temperature observation and CO<sub>2</sub>, 6.3  $\mu$ m for

water vapor, and 9.6 µm for ozone absorption bands (Menzel et al., 1655 2018). The root mean square error (RMSE) of the measured 1656 radiation is better than 0.2 K (Susskind et al., 2003). Moreover, 1657 global atmospheric profiles can be detected every day. Due to 1658 radiometer noise and faults, there are currently only 2047 effective 1659 channels. However, compared with previous infrared detectors, 1660 AIRS boasts a significant improvement in both the number of 1661 channels and spectral resolution (Aumann, 1994; Huang et al., 2005; 1662 Li et al., 2005). 1663

The root mean square error of an AIRS infrared channel is shown 1664 in Fig. 1. The measurement error is not below 0.2K for all the 1665 instrument channels. There are a few channels with extremely large 1666 measurement errors, which reduce the accuracy of prediction to 1667 some extent. Among them, some extremely large measurement 1668 errors reduce the accuracy of prediction to some extent (Susskind et 1669 al., 2003). At present, more than 300 channels have not been used 1670 because their errors exceed 1 K. If data from these channels were to 1671 be used for retrieval, the accuracy of the retrieval could be reduced. 1672 Therefore, it is necessary to select a group of channels to improve 1673 the calculation efficiency and retrieval quality. In this paper we study 1674 channel selection for temperature profile retrieval by AIRS. 1675



1676

1677 Figure 1. Root mean square error of AIRS infrared channel (black1678 spots).

For the calculation of radiative transfer and the weighting function 1680 matrix, K, the RTTOV (Radiative Transfer for TOVS) v12 fast 1681 radiative transfer model is used. Although initially developed for the 1682 TOVS (TIROS Operational Vertical Sounder) radiometers, RTTOV 1683 can now simulate around 90 different satellite sensors measuring in 1684 the MW (microwave), IR (infrared) and VIS (visible) regions of the 1685 spectrum (Saunders et al., 2018). The model allows rapid 1686 simulations (1 ms for 40 channel ATOVS (Advanced TOVS) on a 1687 desktop PC) of radiances for satellite visible, infrared, or microwave 1688

nadir scanning radiometers given atmospheric profiles of 1689 temperature and trace gas concentrations, and cloud and surface 1690 properties. The only mandatory gas included as a variable for 1691 RTTOV v12 is water vapor. Optionally, ozone, carbon dioxide, 1692 nitrous oxide, methane, carbon monoxide, and sulfur dioxide can be 1693 included, with all other constituents assumed to be constant. RTTOV 1694 can accept input profiles on any defined set of pressure levels. The 1695 majority of RTTOV coefficient files are based on the 54 levels (see 1696 Table A1 in Appendix A), in the range from 1050 hPa to 0.01 hPa, 1697 though coefficients for some hyperspectral sounders are also 1698 available on 101 levels. 1699

In order to correspond to the selected profiles, the atmosphere is 1700 divided into 137 layers, each of which contains corresponding 1701 atmospheric characteristics, such as temperature, pressure, and the 1702 humidity distribution. Each element in the weighting function matrix 1703 can be written as  $\partial yi/\partial xj$ . The subscript i is used to identify the 1704 satellite channel, and the subscript j is used to identify the 1705 atmospheric variable. Therefore,  $\partial yi/\partial xj$  indicates the variation in 1706 brightness temperature in a given satellite channel, when a given 1707 atmospheric variable in a given layer changes. We are thus able to 1708 establish which layer of the satellite channel is particularly sensitive 1709 to which atmospheric characteristic (temperature, various gas 1710

- 1711 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.

1714

#### 1715 **3.2 Channel selection comparison experiment and results**

1716 In order to verify the effectiveness of the method, three sets of

1717 comparison experiments were conducted. First, 324 channels used

<sup>1718</sup> by the EUMETSAT Satellite Application Facility on Numerical

1719 Weather Prediction (NWP SAF) were selected. NCS is short for

1720 NWP channel selection in this paper. NCS were released by the

1721 NWPSAF 1DVar (one-dimensional variational analysis) scheme, in

accordance with the requirements of the NWPSAF (Saunders et al.,

1723 2018). Second, 324 channels were selected using the information

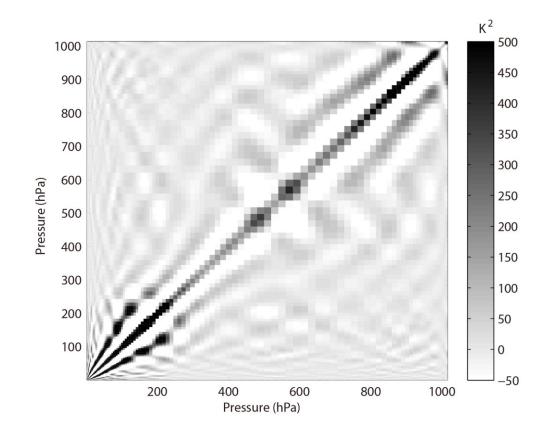
capacity method. This method was adopted by Du et al. (2008)

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.

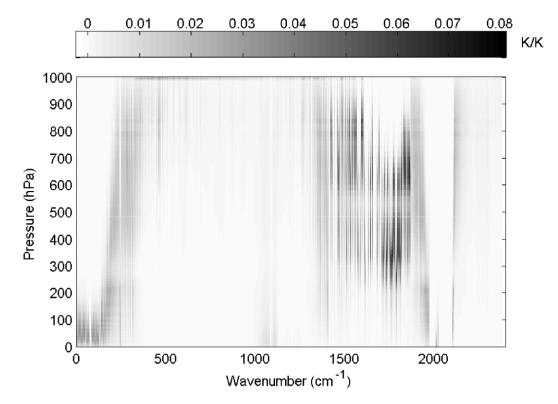
The observation error covariance matrix,  $S_{\varepsilon}$ , in the experiment is 1734 provided by NWP SAF 1Dvar. In general, it can be converted to a 1735 diagonal matrix, the elements of which are the observation error 1736 standard deviation of each hyperspectral detector channel, which is 1737 the square of the root mean square error for each channel. The root 1738 mean square error of the AIRS channels is shown in Fig. 1. The error 1739 covariance matrix of the background,  $S_a$ , is calculated using 5000 1740 samples of the IFS-137 data provided by the ECMWF dataset (The 1741 detailed information will be introduced in Sect. 4). The last access 1742 date is April 26th, 2019 (download address: 1743 https://www.nwpsaf.eu/site/update-137-level-nwp-profile-dataset/, 1744 2019). The covariance matrix of temperature is shown in Fig. 2. The 1745 results are consistent with the previous study by Du et al. (2008). 1746



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

1748

The reference atmospheric profiles are from the IFS-137 database, 1751 and the temperature weighting function matrix is calculated using 1752 the RTTOV K mode, as shown in Fig. 3; the results are consistent 1753 with those of the previous study by Du et al. (2008). For the 1754 air-based passive atmospheric remote sensing studied in this paper, 1755 when the same channel detects the atmosphere from different 1756 observation angles, the value of the weighting function matrix K 1757 changes due to the limb effect. The goal of this section is focusing 1758 on the selection methods of selecting channels; therefore the biases 1759 produced from different observation angles can be ignored. 1760



**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.

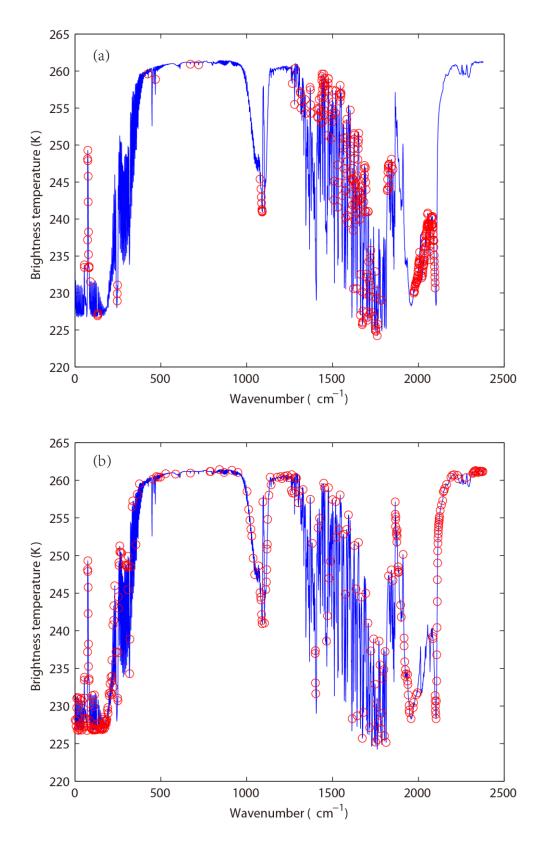


Figure 4. The distribution of different channel selection methods
without considering layering in the AIRS brightness temperature

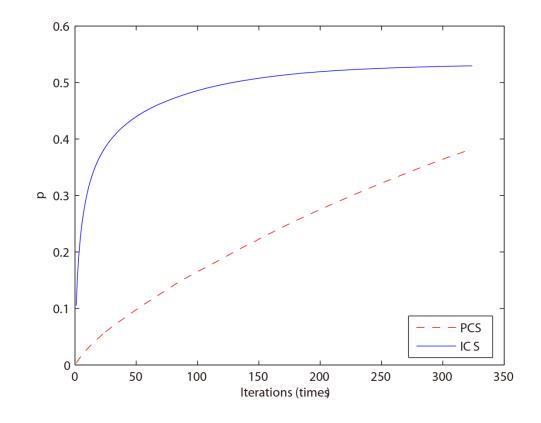
spectrum (blue line). (a) 324 channels selected by PCS (red circles).(b) 324 channels selected by NCS (red circles).

Without considering layering, the main differences between the 1778 324 channels selected by PCS and NCS are as follows: (1) Near 10 1779 µm band, fewer channels are selected by PCS because the retrieval 1780 of ground temperature is considered by NCS; (2) Near 9 µm band, 1781 no channels are selected by PCS because the retrieval of O3 is not 1782 considered in this paper; (3) As is known, the spectral range from 6 1783 μm to 7 μm corresponds to water vapor absorption bands, but fewer 1784 channels are selected by NCS; (4) Near 5 µm band, it includes 4.2 1785  $\mu$ m for N<sub>2</sub>O and 4.3  $\mu$ m for CO<sub>2</sub> absorption bands. As is shown in 1786 Fig. 4, fewer channels are selected by PCS in those bands. PCS is 1787 favorable for atmospheric temperature observation. Because 4.2 µm 1788 and 4.3 µm bands are sensitive to high temperature, a better 1789 observation can be obtained for higher temperatures; (5) Near 4  $\mu$ m 1790 band, a small number of channels is selected by NCS, but no 1791 channels are selected by PCS. 1792

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<sub>3</sub> profile. The advantages of the channel selection method based on information content in this

1798paper are mainly reflected in: (1) Stratosphere and mesosphere is1799less affected by the ground surface, so the retrieval result of PCS is1800better than that of NCS. (2) Due to the method selected in this paper1801there are more channels at 4.2  $\mu$ m for N<sub>2</sub>O and 4.3  $\mu$ m for CO<sub>2</sub>1802absorption bands; the channel combination of PCS is better than that1803of NCS for atmospheric temperature observation at higher1804temperature.

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.



1812

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 1817 PCS method needs to be improved. In this paper, the atmosphere is 1818 divided into 137 layers, and based on the information content and 1819 iteration, 324 channels are selected for each layer. Then, the 1820 temperature profile of each layer can be retrieved based on statistical 1821 inversion (see at Sect. 4). The relationship between the number of 1822 iterations and the ARI for ICS is shown in Fig. 5b. When the number 1823 of iterations approaches 100, the ARI of ICS tends to be stable, and 1824 98

- reaches 0.54. Thus, in terms of the ARI and convergence, the ICSmethod is better than that of PCS.
- 1827 Furthermore, because an iterative method is used to select
- channels, the order of each selected channel is determined by the
- 1829 contribution from the ARI. The weighting function matrix of the top
- 1830 324 selected channels, according to channel order, is shown in Fig.
- 1831 6.

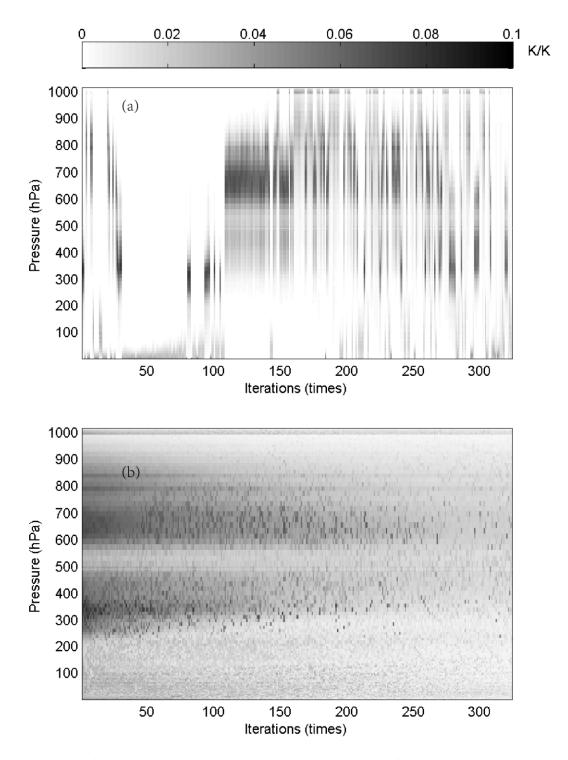


Figure 6. The relationship between the number of iterations and the
weighting function of the top 324 selected channels (shaded). (a)
ICS. (b) PCS.

1837	As illustrated in Fig. 6, in the first 100 iterations, the distribution		
1838	of the temperature weighting function for PCS is relatively scattered;		
1839	it does not reflect continuity between the adjacent layers of the		
1840	atmosphere. Besides, the ICS result is better than that of PCS,		
1841	showing that: (1) the distribution of the temperature weighting		
1842	function is more continuous and reflects the continuity between		
1843	adjacent layers of the atmosphere; (2) regardless of the number of		
1844	iterations, the maximum value of the weighting function is stable		
1845	near 300–400 hPa and 600–700 hPa, without scattering, which is		
1846	closer to the situation in real atmosphere.		
1843 1844 1845	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		

1847

# 1848 **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

1852 Weather Forecasts (ECMWF) was used for the statistical inversion

experiments. The profiles were given in a 137-level vertical grid

1854 extending from the surface up to 0.01 hPa. The database was divided

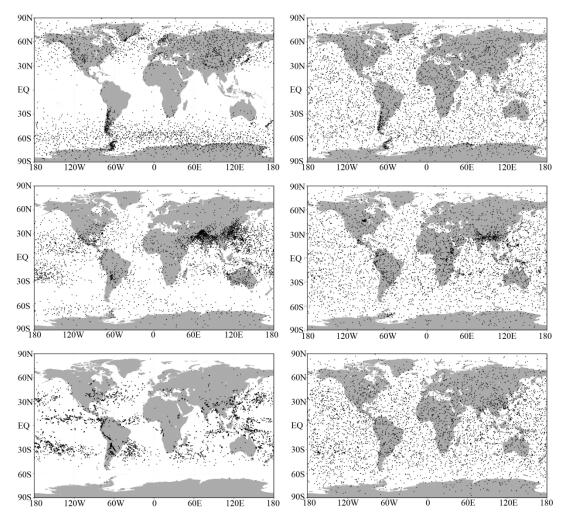
into five subsets focusing on diverse sampling characteristics such as

temperature, specific humidity, ozone mixing ratio, cloud

1857 condensates, and precipitation. In contrast with earlier releases of the

ECMWF diverse profile database, the 137-level database places 1858 greater emphasis on preserving the statistical properties of sampled 1859 distributions produced by the Integrated Forecasting System (IFS) 1860 (Eresmaa and McNally, 2014; Brath et al., 2018). IFS-137 spans the 1861 period from September 1, 2013 to August 31, 2014. There are two 1862 operational analyses each day (at 00z and 12z), and approximately 1863 13 000 atmospheric profiles over the ocean. The pressure levels 1864 adopted for IFS-137 are shown in Table A2 (see Table A2 in 1865 Appendix A). 1866

The locations of selected profiles of temperature, specific 1867 humidity, and cloud condensate subsets of the IFS-91 and IFS-137 1868 databases are plotted on the map in Fig. 7. In the IFS-91 database, 1869 the sampling is fully determined by the selection algorithm, which 1870 makes the geographical distributions very inhomogeneous. Selected 1871 profiles represent those regions where gradients of the sampled 1872 variable are the strongest: in the case of temperature, mid- and 1873 high-latitudes dominate, while humidity and cloud condensate 1874 subsets concentrate at low latitudes. However, the IFS-137 database 1875 shows a much more homogeneous spatial distribution in all the 1876 sampling subsets, which is a consequence of the randomized 1877 selection. 1878



**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
- 1883 <u>https://www.nwpsaf.eu/site/update-137-level-nwp-profile-dataset/</u>,
- 1884 2019).

- 1886 The temporal distribution of the selected profiles is illustrated in Fig.
- 1887 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.

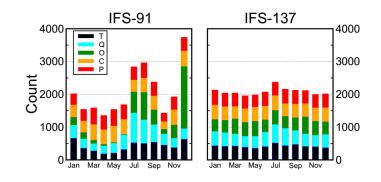


Figure 8. Distribution of profiles within the calendar months in 1892 IFS-91 (left) and IFS-137 (right) databases. Different subsets are 1893 shown in different colors. Black parts stand for temperature. Blue 1894 parts represent specific humidity. Green parts indicate ozone subset. 1895 Orange parts stand for cloud condensate. Red parts represent 1896 precipitation. The last access date is April 26th, 2019. (from 1897 https://www.nwpsaf.eu/site/update-137-level-nwp-profile-dataset/, 1898 2019). 1899

1900

### 1901 **4.2 Experimental scheme**

1902 In order to verify the retrieval effectiveness of ICS, 5000

temperature profiles provided by the IFS-137 were used for

1904 statistical inversion comparison experiments. The steps are as

1905 follows:

(1) 5000 profiles and their corresponding surface factors, 1906

including surface air pressure, surface temperature, 2 m temperature, 1907

2 m specific humidity, 10 m wind speed are put into the RTTOV 1908

mode. Then, the simulated AIRS spectra are obtained. 1909

(2) The retrieval of temperature is carried out in accordance with 1910

Eq. (23). The 5000 profiles are divided into two groups. The first 1911

group of 2500 profiles is used to obtain the regression coefficient, 1912

and the second group of 2500 is used to test the result. 1913

(3) Verification of the results. The test is carried out based on the 1914 standard deviation between the retrieval value and the true value. 1915

1916

1917

## **4.3 Results and Discussion**

For the statistical inversion comparison experiments, the standard 1918 deviation of temperature retrieval is shown in Fig. 9. First, because 1919 PCS does not take channel sensitivity as a function of height into 1920 consideration, the retrieval result of PCS is inferior to that of ICS. 1921 Second, by comparing the results of ICS and NCS we found that 1922 below 100 hPa, since the method used in this paper considers near 1923 ground to be less of an influencing factor, the channel combination 1924 of ICS is slightly inferior to that of NCS, but the difference is small. 1925

From 100 hPa to10 hPa, the retrieval temperature of ICS in this 1926 paper is consistent with that of NCS, slightly better than the channel 1927

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 1935 improvement value of the ICS and its percentages compared with the 1936 PCS and NCS at different heights are shown in Table 1. Because 1937 PCS does not take channel sensitivity as a function of height into 1938 consideration, the retrieval result of PCS is inferior to that of ICS. In 1939 general, the accuracy of the retrieval temperature of ICS is improved. 1940 Especially, from 100 hPa to 0.01 hPa, the mean value of ICS is 1941 evidently improved by more than 0.5 K which means the accuracy 1942 can be improved by more than 11%. By comparing the results of ICS 1943 and NCS we found that below 100 hPa, since the method used in this 1944 paper considers near ground to be less of an influencing factor, the 1945 channel combination of ICS is slightly inferior to that of NCS, but 1946 the difference is small. From 100 hPa to 0.01 hPa, the mean value of 1947 ICS is improved by more than 0.36 K which means the accuracy can 1948 be improved by more than 9.6%. 1949

# **Table 1.** The mean improvement value of the ICS and its

	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%
1953			
1954	This is becau	Stratosphere and	
	1 • 1	CC ( 1.1 (1 ) 1	

### <sup>1952</sup> percentages compared with the PCS and NCS at different heights.

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  $\mu$ m for N<sub>2</sub>O and

 $4.3 \mu m$  for CO<sub>2</sub> absorption bands, and the channel combination of

1959 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

1962 retrieval result is improved.

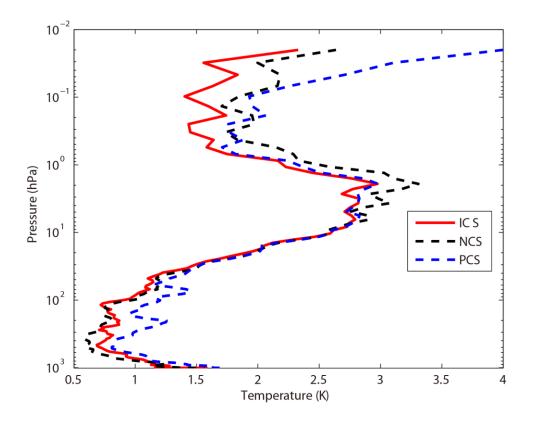




Figure 9. The temperature profile standard deviation of statistical
inversion comparison experiments. Red line indicates the result of
ICS. Black dotted line stands for the result of NCS. Blue dotted line
represents the result of PCS.

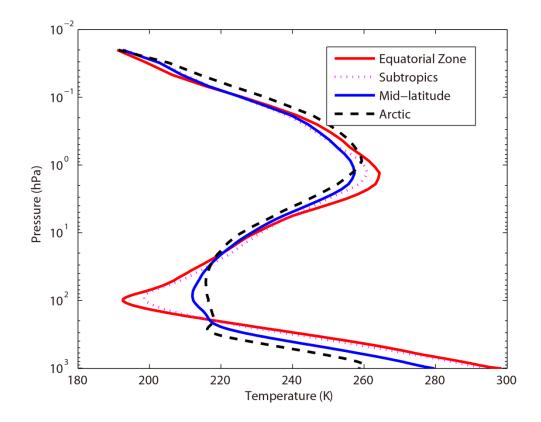
## 1970 **5** Statistical inversion comparison experiments in four typical

1971 regions

<sup>1972</sup> The accuracy of the retrieval temperature varies from place to place

- and changes with atmospheric conditions. Therefore, in order to
- 1974 further compare the inversion accuracy under different atmospheric
- 1975 conditions, this paper has divided the atmospheric profile from the
- 1976 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.



1984

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.

#### 1991 **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 regressioncoefficient.

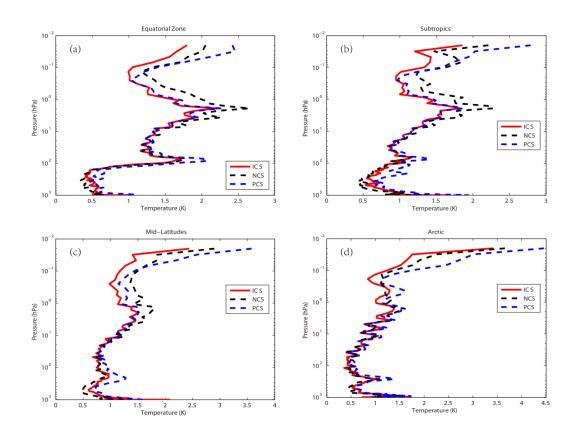
(2) The atmospheric profiles of the four typical regions: equatorial
zone, subtropical region, mid-latitude region and Arctic are used for
statistical inversion comparison experiments and test the result.(3)
Verification of the results. The test is carried out based on the
standard deviation between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

2004

## 2005 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





2014

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.

2024

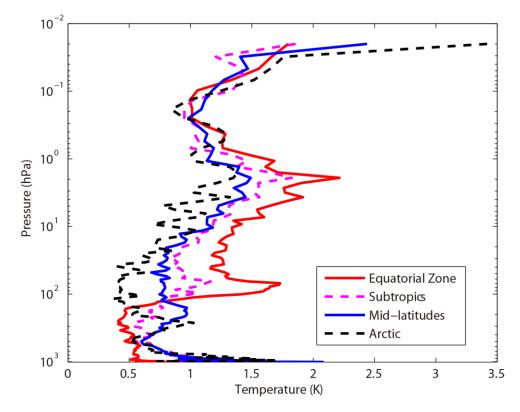


Figure 12. The temperature standard deviation of ICS in four typical regions. Red line indicates the result of equatorial zone. Pink dotted line represents the result of Subtropics. Blue line represents the result of Mid-latitudes. Black dotted line stands for the result of Arctic.

2025

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  $\mu$ m for N<sub>2</sub>O and 4.3  $\mu$ m for CO<sub>2</sub> absorption bands which has 2038 been previously described in Sect. 3. Near the tropopause, the 2039 standard deviation of the equatorial zone increases sharply. It is also 2040 due to the sharp drops in temperature. However, the standard 2041 deviation of the Arctic is still around 0.5K. From 100hPa to 1hPa, 2042 the standard deviation of ICS is 0.5 K to 2K. With the increase of 2043 latitude, the effectiveness considerably increases. According to Fig. 2044 11, ICS takes channel sensitivity as a function of height into 2045 consideration, so its retrieval result is better. 2046

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.

2053

## 2054 7 Conclusions

In recent years, the atmospheric layer in the altitude range of about 2056 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 2060 exploration, therefore, on and of the near space layer are of great 2061 importance. A new channel selection scheme and method for 2062 hyperspectral atmospheric infrared sounder AIRS data based on 2063 layering is proposed. The retrieval results of ICS concerning the near 2064 space atmosphere are particularly good. Thus, ICS aims to provide a 2065 new and an effective channel selection method for the study of the 2066 near space atmosphere using the hyperspectral atmospheric infrared 2067 sounder. 2068

An improved channel selection method is proposed, based on 2069 information content in this paper. A robust channel selection scheme 2070 and method are proposed, and a series of channel selection 2071 comparison experiments are conducted. The results are as follows: 2072 (1) Since ICS takes channel sensitivity as a function of height into 2073 consideration, the ARI of PCS only tends to be 0.38 and is not 2074 convergent. However, as the 100<sup>th</sup> iteration is approached, the ARI of 2075 ICS tends to be stable, reaching 0.54, while the distribution of the 2076 temperature weighting function is more continuous and closer to that 2077 of the actual atmosphere. Thus, in terms of the ARI, convergence, 2078 and the distribution of the temperature weighting function, ICS is 2079 better than PCS. 2080

2081 (2) Statistical inversion comparison experiments show that the

retrieval temperature of ICS in this paper is consistent with that of 2082 NCS. In particular, from 10 hPa to 0.02 hPa (the stratosphere and 2083 mesosphere), the retrieval temperature of ICS is obviously better 2084 than that of NCS at about 1 K. In general, the accuracy of the 2085 retrieval temperature of ICS is improved. Especially, from 100 hPa 2086 to 0.01 hPa, the accuracy of ICS can be improved by more than 11%. 2087 The reason is that stratosphere and mesosphere are less affected by 2088 the ground surface, so the retrieval result of ICS is better than that of 2089 NCS. Additionally, due to the method selected in this paper there are 2090 more channels at 4.2  $\mu$ m for the N<sub>2</sub>O and at 4.3  $\mu$ m for the CO<sub>2</sub> 2091 absorption bands; the channel combination of ICS is better than that 2092 of NCS for atmospheric temperature observation at higher 2093 temperature. 2094

(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.

2099

2100 *Data availability*. The data used in this paper are available from the 2101 corresponding author upon request.

2102

2103 Appendices

2104 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.
- 2108 (From <u>https://www.nwpsaf.eu/site/software/rttov/</u>, RTTOV Users
- 2109 guide, 2019).

Level	Pressure	Tmax	Tmin	Qmax	Qmin	Q2max	Q <sub>2</sub> min	Q <sub>2</sub> Ref
Number	hPa	к	к	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

2112 (in hPa).

	proceiling		prosourc	Loval	proceiling	Lovel	proceuro	Lovel	prosouro
Level number	pressure hPa	Level number	pressure hPa		•		pressure hPa	Level number	pressure hPa
1	0.02	31	12.8561		106.4153		424.019		
2	0.031	32	14.2377		112.0681		441.5395		943.1399
3	0.0467	33	15.7162		117.9714		459.6321		950.9082
4	0.0683	34	17.2945		124.1337		478.3096		958.1037
5	0.0975	35	18.9752		130.5637		497.5845		964.7584
6	0.1361	36	20.761		137.2703		517.4198		970.9046
7	0.1861	37	22.6543		144.2624		537.7195		976.5737
8	0.2499	38	24.6577		151.5493		558.343		981.7968
9	0.3299	39	26.7735		159.1403		579.1926		986.6036
10	0.4288	40	29.0039		167.045		600.1668		991.023
11	0.5496	41	31.3512		175.2731		621.1624		995.0824
12	0.6952	42	33.8174		183.8344		642.0764		998.8081
13	0.869	43	36.4047		192.7389		662.8084		1002.225
14	1.0742	44	39.1149		201.9969		683.262		1005.356
15	1.3143	45	41.9493		211.6186		703.3467		1008.224
16	1.5928	46	44.9082		221.6146		722.9795		1010.849
17	1.9134	47	47.9915		231.9954		742.0855		1013.25
18	2.2797	48	51.199		242.7719		760.5996		
19	2.6954	49	54.5299		253.9549		778.4661		
20	3.1642	50	57.9834		265.5556		795.6396		
21	3.6898	51	61.5607		277.5852		812.0847		
22	4.2759	52	65.2695	82	290.0548	3 112	827.7756		
23	4.9262	53	69.1187	83	302.9762	2 113	842.6959		
24	5.6441	54	73.1187	84	316.3607	7 114	856.8376		
25	6.4334	55	77.281	85	330.2202	2 115	870.2004		
26	7.2974	56	81.6182	86	344.5663	3 116	882.791		
27	8.2397	57	86.145	87	359.4111	I 117	894.6222		
28	9.2634	58	90.8774	88	374.7666	6 118	905.7116		
29	10.372	59	95.828	89	390.645	5 119	916.0815		

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.

2118

2119 *Competing interests*. The authors declare that they have no conflict2120 of interest.

2121

Acknowledgements. The study was supported by the National Key 2122 Research Program of China: Development of high-resolution data 2123 assimilation technology and atmospheric reanalysis data set in East 2124 Asia (Research on remote sensing telemetry data assimilation 2125 technology, Grant no. 2017YFC1501802). The study was also 2126 supported by the National Natural Science Foundation of China 2127 (Grant no. 41875045) and Hunan Provincial Innovation Foundation 2128 for Postgraduate (Grant no. CX2018B033 and no. CX2018B034). 2129

2130

#### 2131 **References**

Aires, F., Schmitt, M., Chedin, A., and Scott, N.: The "weighting

smoothing" regularization of MLP for Jacobian stabilization,

- 2134 IEEE. T. Neural. Networks., 10, 1502-1510,
  2135 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 2162 improved initialization inversion method: a high resolution 2163 physical method for temperature retrievals from satellites of the 2164 tiros-n series. J. Appl. Meteor. 24. 128-143. 2165 https://doi.org/10.1175/1520-0450(1985)024<0128:TIIIMA>2.0.C 2166 O;2, 1985. 2167
- <sup>2168</sup> Cyril, C., Alain, C., and Scott, N. A.: Airs channel selection for CO<sub>2</sub>
- and other trace-gas retrievals, Q. J. Roy. Meteor. Soc., 129,

2170 2719-2740, https://doi.org/10.1256/qj.02.180, 2003.

- 2171 Du, H. D., Huang, S. X., and Shi, H. Q.: Method and experiment of
- channel selection for high spectral resolution data, Acta. Physica.
- 2173 Sinica., 57, 7685-7692, 2008.
- 2174 Dudhia, A., Jay, V. L., and Rodgers, C. D.: Microwindow selection
- for high-spectral-resolution sounders, Appl. Opt. 41, 3665-3673,
- 2176 https://doi.org/10.1364/AO.41.003665, 2002.
- 2177 Eresmaa, R. and McNally, A. P.: Diverse profile datasets from the

ECMWF 137-level short-range forecasts, Tech. rep., ECMWF, 2179 2014.

- Eyre, J. R., Andersson E., and McNally, A. P.: Direct use of 2180 satellite sounding radiances in numerical weather prediction. High 2181 Spectral Resolution Infrared Remote Sensing for Earth's Weather 2182 Springer, Studies, and Climate Berlin, Heidelberg. 2183 https://doi.org/10.1007/978-3-642-84599-4 25, 1993. 2184
- 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.: 2200 Evaluation of cloud-cleared radiances for numerical weather 2201 cloud-contaminated sounding prediction and applications. 2202 Atmospheric and Environmental Remote Sensing Data Processing 2203 Atmospheric Numerical Utilization: Prediction and and 2204 Monitoring, I. S. Photonics., Environmental О. 2205 https://doi.org/10.1117/12.613027, 2005. 2206

Kuai, L., Natraj, V., Shia, R. L., Miller, C., and Yung, Y. L.: Channel
selection using information content analysis: a case study of CO<sub>2</sub>
retrieval from near infrared measurements. J. Q. S. Radiative.
Transfer., 111, 1296-1304,

2211 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., 2212 Huang, H. L., and Achtor, T. H.: Global soundings of the 2213 atmosphere from ATOVS measurements: the algorithm and 2214 validation, J. Appl. Meteor., 39, 1248-1268, 2215 https://doi.org/10.1175/1520-0450(2000)039<1248:GSOTAF>2.0. 2216 CO:2, 2000. 2217
- 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
- 2220 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.
- 2225 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.
- 2232 Menzel, W. P., Schmit, T. J., Zhang, P. and Li, J.: Satellite-based
- atmospheric infrared sounder development and applications, Bull.
- 2234 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
- 2238 prediction, Q. J. Roy. Meteor. Soc., 124, 211-241,
- https://doi.org/10.1002/qj.49712454510, 2010.
- 2240 Xu, Q.: Measuring information content from observations for data
- assimilation: relative entropy versus shannon entropy difference,
- 2242
   Tellus.
   A.,
   59,
   198-209,

   2243
   https://doi.org/10.1111/j.1600-0870.2006.00222.x, 2007.
  - 124

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.
- 2252 Rodgers, C. D.: Information content and optimisation of high 2253 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.
- 2258 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,
- 2262 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,
- 2283 2011.
- 2284