Authors' Response to Reviews of

Evaluation of tropospheric water vapour and temperature profiles retrieved from Metop-A by the Infrared and Microwave Sounding scheme

Tim Trent, Richard Siddens, Brian Kerridge, Marc Schrörder, Noëlle A. Scott, and John Remedios
EGUsphere, `https://doi.org/10.5194/egusphere-2022-757, 2022`

RC: *Reviewers' Comment*,   AR: Authors' Response,   ☐ Manuscript Text

1. Reviewer #2

RC: *This paper describes the validation of the RAL IMS water vapor and temperature profiles retrieved from infrared hyperspectral and microwave sounders. This is useful if this dataset is to be used as a CDR for climate applications. The paper is well written and well structured. I therefore deserves publishing.*

 AR: We would like to thank the referee for taking the time to review our manuscript. Below we reply to the issues raised by the referee. The original reviewer comments (RC) are given in bold italics, with the author's responses (AR) in plain text. Where we have updated the manuscript, the extract is included in a quote box with the original removed text in red and struck out. New text appears in blue and is underlined.

1.1. Page 3 line 63. Section 2.Data

RC: *The retrievals are not only IASI based, so you might want to add something like "and microwave sounder data".*

 AR: We thank the reviewer for highlighting this point. Indeed, the IMS scheme uses infrared and microwave sounder data to retrieve water vapour and temperature profiles. Therefore, we have updated the text to make sure this point is clear:

> This section describes the algorithm used to retrieve water vapour and temperature profiles from IASI ~~, along with~~ with companion microwave sounder data and details of the radiosonde data sets used for their assessment.

1.2. Page 4. line 99.

RC: *You might want to add something like "This is justified by the finding that WV inhomgeneities within the FOV do cause a significant modification in the results of the radiative transfer modeling (`https://amt.copernicus.org/articles/11/6409/2018/`,`https://cimss.ssec.wisc.edu/itwg/itsc/itsc23/presentations/oral.2.01.calbet.pdf`).*

 AR: We have added the reference and a sentence to reflect the links between the two bodies of work. Inhomogeneities and their impacts will have an increasing significance as a source of uncertainty for profile climate data records. Therefore it is correct to acknowledge the previous study here.

> The bias correction is needed to account for systematic differences between RTTOV and the IASI observations, including errors in RTTOV. Allowing the retrieval to fit scale factors x_{b0} and x_{b1}, instead of assuming a fixed scan-angle dependence improves the fit (gives lower cost) over a wide range of observing conditions. The recent study by Calbet et al. (2018) supports this approach as they demonstrated that the inhomogeneities in water vapour within a satellite Instantaneous Field of View (IFOV) cause a significant modification in the results from radiative transfer modelling.

1.3.

RC: *I would be interested in seeing the results of the bias corrections. If the inhomogeneity correction hypothesis is true, the biggest contribution to this bias will come from inhomegeneities in WV within the FOV. So this could be a way to measure them. They potentially can be correlated with turbulence.*

AR: We thank the reviewer for raising this interesting point. Firstly, scene inhomogeneity across IASI IFOVs is accounted for by EUMETSAT in generating L1 data through the use of co-located AVHRR images. This inhomogeneity is predominantly due to the presence of clouds. For the retrieval, water vapour spectral features present in the systematic bias spectra are most likely due to inconsistencies in H2O spectroscopic data between different wavelength intervals used by IMS. Though not impossible that scene inhomogeneity might contribute in certain cases, this would require a detailed study which would not be merited at this stage in our view. To reflect the points we have added additional text to the manuscript and direct the reader to examples of the bias-corrected spectra from RTTOV that can be found in the EUMETSAT report, Siddans & Gerber, 2015.

> The bias correction is needed to account for systematic differences between RTTOV and the IASI observations, including errors in RTTOV. Allowing the retrieval to fit scale factors x_{b0} and x_{b1}, instead of assuming a fixed scan-angle dependence improves the fit (gives lower cost) over a wide range of observing conditions. Examples of these corrections to systematically biased spectra are given in Siddans & Gerber (2015). The recent study by Calbet et al. (2018) supports this approach as they demonstrated that the inhomogeneities in water vapour within a satellite Instantaneous Field of View (IFOV) cause a significant modification in the results from radiative transfer modelling. Observational inhomogeneities across the IASI IFOV are predominantly due to clouds within the scene. These effects are accounted for at the L1 data stage by EUMETSAT through the collocation of AVHRR images within the IFOV (EUMETSAT, 2019)

1.4.

RC: *I do not completely understand what DOFS is. And along with it Fig. 3. Also in Fig. 3 you say you plot DOFS but in the scale you show "% of profiles per bin". Please explain this better for people not familiar with DOFS. One or two sentences should be enough.*

AR: Firstly, we appreciate that the concepts around averaging kernels and degrees of freedom for signal (DOFS) may not be familiar to many people as they are not commonly used by the tropospheric water vapour community. Secondly, we thank the reviewer for highlighting possible confusion around Figure 3. Therefore, we propose to update the text (line 165 onwards) on averaging kernels and DOFS to make this point clearer:

> The practical uses of the square averaging kernel matrix (**A**) ~~is of relatively little practical use, and averaging kernels in this form are only used in IMS~~ are to i) smooth atmospheric profiles from models,

reanalysis and *in situ* measurements (discussed further in Section 3.2), and ii) to obtain the degrees of freedom for signal (DOFS) for specific retrieval products. ~~This is~~ The DOFS are given by the trace (sum of the diagonal elements) of the sub-matrix of $\mathbf{A}$ corresponding to a specific product and represent the total number of independent pieces of information in the profile. The averaging kernel of retrieved water vapour profiles (defined on the RTTOV levels) with respect to perturbations on the fine atmospheric grid is given by:

$$\mathbf{A}_{qf} = \mathbf{M}_{qf}\mathbf{A}_{f:q}, \tag{10}$$

where $\mathbf{A}_{f:q}$ is the averaging kernel for the water vapour state vector elements with respect to perturbations in ln(ppmv) on the fine atmospheric levels. The averaging kernel for temperature is derived similarly using the corresponding matrices. An understanding of ~~the~~ a profile vertical resolution ~~is~~ can be inferred from the DOFS value, as it describes the number of independent pieces of information resolved (Rodger, 2000). Figure 3 shows the range of DOFS for IMS temperature and water vapour profiles as a function of latitude, with the tropopause height (TPH) overlaid. ~~Nadir~~ The 2D histograms show the distribution of profile DOFS, from which we can see that in the tropics, most profiles sit between 6-7 and 11-12 DOFS for water vapour and temperature, respectively. Moving outwards through the mid-tropics to the high latitudes, the DOFS values reduce, with the distribution becoming more variable. Comparing the water vapour distribution to the cold point tropopause height (black dashed line), we can observe that they hold similar shapes, while for temperature, this is less so. This result is expected as nadir IR+MW sounders are predominately sensitive to the emissions from the troposphere. ~~Examples of~~, especially for water vapour.

The next conceptual step is how the DOFS relate to the vertical resolution of IMS profiles. Examples of averaging kernels for water vapour and temperature ~~averaging kernels~~ profiles from the IMS L2 product are given in Figure 4. ~~Examination of the cumulative degrees of freedom~~, where we see that most of the information for water vapour is situated in the lowest 10 km of the atmosphere while for temperature is more continuous. Therefore, an examination of the cumulative degrees of freedom for signal (CDOFS) from these averaging kernels can be used to describe the vertical resolution of the retrieved profiles. The gradients of the CDOFS as a function of altitude can then be interpreted as the profile resolution at given heights. The desired performance for vertical resolutions from IASI is 1 & 2 km for temperature and water vapour profiles, respectively (Hilton et al., 2012). What can be seen from Figure 4 is that vertical resolution is not necessarily constant throughout the troposphere. Indeed, examining a different sounding over the same radiosonde site would show subtle differences. A key observation here is that the information from the water vapour profile terminates (vertical gradient) at the tropopause. Therefore, any use of the IMS water vapour profile above this height is meaningless.

To complement this new text, we also modify the caption of Figure 3 to orientate the reader better to the information given.
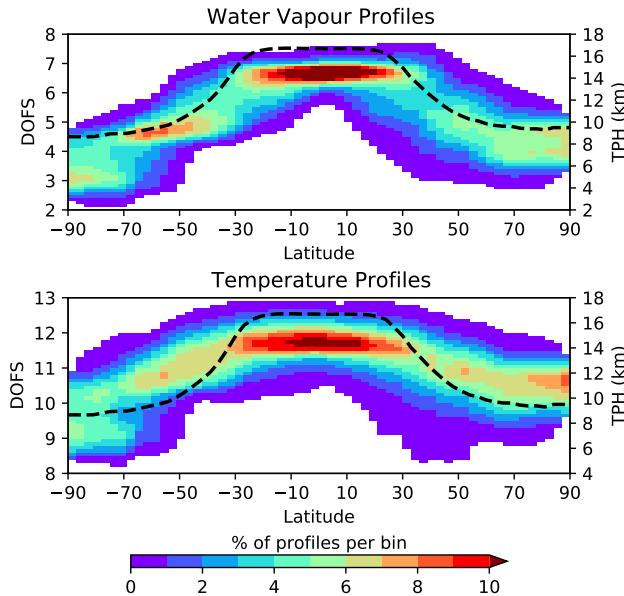
3

Figure 3: Visualisation of IMS water vapour and temperature profile degrees-of-freedom for signal (DOFS). This figure illustrates the latitudinal distribution of DOFS variability for both IMS water vapour and temperature profiles. DOFs were collected from the IMS L2 files between 2007-2016 and binned as a function of latitude. Values were then normalised using the total number of profiles in their respective latitude bin. ~~The dashed line represents the~~ DOFs vary approximately between 2-7 for water vapour and 8-13 for temperature with strong peaks in the tropics. The spread in the data resembles the cold point tropopause height (TPH), especially for water vapour. The dashed black line represents the cold point TPH calculated from ERA5 temperature profiles (Hersbach et al., 2020).

1.5.

RC: *Please explain which radiosondes are used in GRUAN. There is a relatively big difference in WV between RS92 and RS41, being the latter much more precise.*

AR: In this study we only use the RS92 soundings. Therefore, we add a statement to the end of Line 189 to clarify this:

> An advantage of the higher resolution of GRUAN measurements is that it captures changes in humidity gradients and temperature inversions which can be missed or underrepresented by standard and significant pressure levels. It should be noted that the soundings from GRUAN feature only the Vaisala RS92 radiosondes measurements and not the more recent RS41.

1.6. Page 10. line 133.

RC: *I believe for $x_{(z)}$ you mean "layer mean profile" and not "weighted layers". This is confusing since in the next couple of sentences the term "layer mean profile" is used. Please correct or explain better.*

AR: The term $x_{(z)}$ in the text refers to the mean layer within a profile over which we are averaging. We thank the

4

reviewer for pointing out the confusion and acknowledge that the term needs to be updated to be consistent with others later in the text. Therefore, we propose to update the following text to clarify things for the reader:

Next, ~~profile~~ values are calculated for weighted layers ($\bar{\mathbf{x}}_{(z)}$) within each new layer mean profile ($\bar{\mathbf{x}}$), where the layer boundaries are defined by standard pressure levels defined at 1000, 925, 850, 700, 500, 400 and 300 hPa:

$$\bar{\mathbf{x}}_{(z)} = \frac{\sum_{l=1}^{n} \mathbf{x}_{(l)} \mathbf{P}_{(l)}}{\sum_{l=1}^{n} \mathbf{P}_{(l)}}. \tag{13}$$

Where $\mathbf{x}_{(l)}$ is the convolved radiosonde or IMS profile value at level l, $\mathbf{p}_{(l)}$ is the pressure profile value at level l, and n is the numbers of levels in the layer. ~~Layer~~ Weighted layer mean profiles are not calculated for altitudes higher than 300 hPa because ARSA profile values are taken from ERA-Interim in the ~~upper-troposphere~~ upper troposphere/stratosphere. All statistics used by this study are calculated from the layer mean profiles.

1.7.

RC: *Please note that RS92 sondes do not measure very well with low WV. This usually happens above tropopause. This is probably why the biases are much bigger at higher altitudes. You might want to repeat the statistics using only levels below tropopause. Or, equivalently, with small GRUAN uncertainties. This will most likely reduce the biases at high altitudes. Something to consider.*

AR: We thank the reviewer for this suggestion; this issue was something we were aware of when designing the analysis. The issue we faced is that the ARSA radiosondes do not have any uncertainty information associated with the profiles. Therefore, we could not apply the same criteria to these measurements. Experience from Trent et al. 2019 has shown us that for water vapour averaging kernels with sensitivity to the upper troposphere/lower stratosphere (UT/LS) actually have a bigger impact than the radiosonde biases due to the lack of sensitivity of IR sounder to moisture in this region. What we observe from Figure 2 is that above the tropopause the water vapour averaging kernels are very close to zero, and when applied to the difference between the radiosonde profile and a priori (equation 12) this term approaches also approaches zero. Therefore, above the tropopause the convolved profile essentially matches the a priori. This discussion is included with the updated text for averaging kernels (RC No. 1.4). Finally, the cut off of the profiles at 300 hPa also helps remove layers with UT/LS sensitivity or include profile levels above the tropopause.

1.8. Page 10. line 206.

RC: *Change "i) the dataset is being validated" with "i) the dataset being validated".*

AR: Done.

The match-up database (MUDB) is generated by supplying the MUP a driver file containing information on i) the dataset ~~is~~ being validated,...

1.9.

RC: *Please explain in the text what MAD is when it first appears.*

AR: We have updated the text to clarify this point

> Therefore, we can think of the variability of the median as an estimate of the precision of the bias. To quantify the spread about the median, we calculate the median absolute deviation ($\sigma_{(z)}$), a robust measure of the data variability:
>
> $$\sigma_{(z)} = median|(\mathbf{x}_{(z)} - \mathbf{x}_{est(z)}) - b_{(z)}|. \tag{14}$$
>
> As we use robust statistics, the ~~MAD~~ median absolute deviation (MAD) values cannot be treated in the same way as standard deviation and used to calculate the standard error by dividing through by $\sqrt{N}$.

1.10. Fig 5. and 6.

RC: *If you use a smaller collocation window than 3 hrs and 100 km, the MAD will certainly disminish. Something to consider as en exercise.*

AR: Due to the static launch times of radiosondes when we reduce the number of sites with which we make collocations. This is especially true for the IMS product because we do not get averaging kernels for every IFOV. The scope of this study is to present the IMS product performance at global and sub-global scales, which requires collocation criteria that give enough scope to achieve this. Therefore, reducing these windows can introduce a significant sampling bias, with results becoming highly localised. Our strategy was to gather as many collocated pairs as possible and use robust statistics to minimise the impact of individual comparison performance arising from the broad collocation criteria. This approach works well in general, although suffers in areas of low coverage, as we would expect. We demonstrated this is a previous study and have added the reference to the text where we explain our strategy (lines 202-220):

> (line 216)...However, this is minimised by calculating global or per latitude band statistics which uses large numbers of matched pairs, unlike for site comparisons with a low number of matched cases. When using broad collocation criteria, any mismatch introduced during the match-up will affect the performance of individual comparison performance (Sun et al. (2010) & Sun et al.(2017)). Therefore, a robust statistics approach was adopted to minimise this effect as demonstrated in Trent et al. 2019.

1.11. Fig. 8.

RC: *Why is there less cases with higher cloud fractions? Can you give an explanation or hypothesis?*

AR: We thank the reviewer for raising this point as it was not clear in the manuscript. The combination of infrared and microwave instruments allows for the retrieval of profiles in the presence of cloud which would not be possible if the infrared alone were used. That said, cloud can still impact the retrieval and as such the product contains a flag that allows for cloudy scenes which the retrieval struggles to work for. Therefore, we see a disproportionate effect on scenes containing cloud especially this with higher cloud fractions. This results in the distribution we see in Figure 8. We propose adding this explanation to the text:

> Therefore, understanding the impact of cloud fraction within the IASI ~~Instantaneous Field of View (IFOV)~~ IFOV on IMS profile biases is also of interest to this study. IMS profile biases were binned according to cloud fraction at intervals of 0.1 for all sites for all matches (day & night cases) and for the separate day and night cases. Water vapour and temperature bias results as a function of cloud fraction are presented in Figure 8 along with the difference between day and night cases to the 'all cases' result. It should be noted that a brightness temperature difference (BTD) flag is used to remove cloudy

scenes that significantly impact the retrieval. While IMS can produce profiles for cloudy IFOVs, the BTD flag will remove some of the profiles, disproportionately across increasing cloud fractions. This explains the distribution we observe in Figure 8.

1.12.

RC: *GRUAN WV measurements are bias corrected with, mainly, an estimation of the incident radiation from the Sun. This is critical for RS92 sondes. Not so much for RS41. Are ARSA sondes also bias corrected in WV?*

AR: ARSA is based on raw radiosondes observations extracted from the ECMWF archive. The required minimal information is to have radiosonde measured values from the surface up to 30 hPa for temperature profiles and from surface to 300 hPa for water vapor profiles. These radiosonde observations are extended above their highest measured point with time and space ERA-Interim co-located data up to 0.1 hPa(temperature, water vapour, and ozone up to 0.1 hPa) and then with SciSat ACE FTS level2 data (from 0.1hPa to the top the atmosphere: 0.0026 hPa). The validation of every surface to top ARSA profile relies upon the study of statistics (bias, standard deviation) between simulated and observed satellite radiances: TOVS and ATOVS, as well as, in more recent years, the MetOp A&B IASI, HIRS4, and MHS observations. The simulated data are generated by the 4A/OP radiative transfer model (Scott and Chédin, 1981, Chéruy et al, 1995) fed with the closest time and space ARSA colocated profile. Due to the excellent stability of the IASI radiances, the quality of the ARSA profiles may then be assessed. It is during such validation procedures that the water vapour profiles above 300hPa (i.e. in the area concerned by the extension of radiosondes by ERA-Interim co-located time-space profiles) turned out to be modified: the procedure is detailed in QUASAR (QUality Assessment of SAtellite and Radiosonde data) report (Scott, 2015). No correction is made to the water vapour due to an estimation of the incident radiation from the Sun".

1.13.

RC: *Are ARSA sondes RS92? Are they corrected from WV biases are they are in GRUAN (taking into account radiation)? Please explain in the paper. This would explain the day night bias difference in ARSA.*

AR: Because of its global coverage (see Figure 4) from nearly 1500 stations, ARSA includes RS92 but not only. No correction is made to the water vapour due to an estimation of the incident radiation from the Sun. The quality of ARSA profiles is assessed using match-ups with satellite radiance measurements with the development, validation, and characteristics of the ARSA database are fully described in the QUASAR (QUality Assessment of SAtellite and Radiosonde data) report (Scott, 2015). It is possible that this could explain the day night differences we observe between the two archives, though we do get different global sampling which will also impact the result. To clarify the differences between ARSA and GRUAN, we propose modifying the text on ARSA (line 190-199) to :

> The second source of radiosonde data is taken from the Analyzed RadioSoundings Archive (ARSA). Produced at the Laboratoire de Météorologie Dynamique (LMD) since the late ~~90's~~90s, ARSA is designed for the processing and validation of level 1 (L1) and level 2 (L2) satellite data and applications. This includes forward and inverse radiative transfer simulations and inter-comparison of retrieved satellite geophysical parameters. The ARSA database is a global archive with observations from approximately 1450 stations. In the first instance, raw radiosondes observations with measurements between the surface and 300 hPa for water vapour and 30 hPa for temperature profiles are extracted from the ECMWF archive. These radiosonde observations are then extended above their highest

measured point to 0.1 hPa with collocated data from ERA-Interim. Finally, level profile data from the SciSat Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS) is used to complete the profile between 0.1 and 0.0026 hPa. The vertical resolution of ARSA varies within the profile, where the lowest part of the troposphere ranging from the surface to 800 hPa, has a resolution of 0.5 km. Between 800 and 200 hPa, the resolution is 0.8 km, increasing to 1.5 km from 200 hPa to 100 hPa. Above 100 hPa to the top-of-atmosphere (TOA), the resolution further reduces to 2.5 km. Unlike GRUAN, which applies a number of corrections to the raw measurement, e.g. correction to water vapour due to incident solar radiation on the radiosonde casing, rather the validation of every ARSA profile relies upon analysing the bias and standard deviation between observed satellite and simulated radiances (Scott, 2015). The ARSA measurement record ~~starts~~ started in January 1979 ~~,~~ and is regularly updated on a monthly basis. Locations of 587 sites present in the archive during the study period can be seen in Figure 4b. For this study, we use the current version 2.7 archive, which has been in use since 2005.

We also add a short discussion on the limitations of both data sets to the end of section 2.2:

Finally, it is worth noting that while radiosondes provide a source of reference data for profile validation, they are not without their own limitations and caveats of use:

- Model type: Corrections made to radiosondes are highly dependent on the make and model type, especially with older radiosondes (e.g. Miloshevich et al. (2001, 2006)). Both archives used in this study have different approaches to correct radiosondes, with GRUAN applying empirical corrections (Dirksen et al., 2014) and ARSA using a radiative transfer modelling to test for consistency between stable satellite radiances (Scott, 2015; Calbet et al., 2017).

- Time series consistency: Radiosonde archives are subject to semi-regular observation system changes, some of which are recorded by the WMO. For GRUAN, their certified sites undergo periodic auditing of their measurement programs in addition to annual reviews to ensure all sites continue to meet practice standards. It is not clear how well this approach scales would scale from ≈30 sites to 500+ found in a global network. ARSA uses the long-term statistics from the radiance intercomparisons to ensure quality consistency across the archive. This approach allows for a common method to be applied to a global network of up to 1450 sites, however, this relies on the radiometric stability of the reference satellite instrument.

- Sources of uncertainty: Radiosondes are subject to a number of sources of uncertainty which can be difficult to characterise fully. The GRUAN provides a comprehensive error budget for their products as their correction process allows estimates for each step. However, ARSA, like other global datasets, does not give an uncertainty on the profiles it provides do the complexity of such an exercise. In Trent et al. (2019), it was demonstrated that the uncertainty on operational records reduces to a few % ppmv with large collocation numbers.

- Distribution of sites: One of the strengths of operational radiosonde records is a large number of global sites available for match-ups. While ARSA does quality filter these, it still has over 500 sites within the study period. For GRUAN, there are only a small number of sites, though they try to sample major climate regimes to provide some global representation. A key weakness for any radiosonde archive is the lack of sites in the southern hemisphere, especially for GRUAN (Figure 4a).

1.14.

RC: *I would separate discussions and conclusions sections. The section is too long with too many comments for a conclusion.*

AR: We accept that the combined discussion and conclusion section benefits from being separated and have done so. To close off the new discussion section, we add the following text:

> . . . With our study, we have adopted a methodology to that outlined in Trent et al., (2019) so that the results are comparable within a common framework. This is especially important for future validation exercises when considering combining different platforms to create a CDR. Capability from other international collaborative efforts has been incorporated to extend the scope of this validation framework. Tools developed within G-VAP Schröder et al. (2019) have been employed to investigate the stability of observed biases, which is vital for the long-term characterisation of CDR performance.

However, we disagree that there are too many concluding points as this study performs an extensive analysis of temperature and water vapour profiles. Where previous studies will focus on one reference dataset, with a diurnal split and maybe some latitudinal analysis, we do this for two reference data records and investigate cloud effects and bias stability. Therefore, we feel the number of main conclusions (6) is appropriate. The final statement of this section was originally included as part of the main conclusions (bullet point); this has now been added to the main flow of the text, and the section flows better as a result.

1.15.

RC: *I would recommend to draw conclusions from night time sondes only. Since we know day time sondes have biases. Especially if they are RS92 sondes and even more if they are not bias corrected in ARSA. Your conclusions on bias trends might vary.*

AR: While we recognise that the nigh-time match-ups will give us the most robust results, we also feel that the 'all' and 'daytime' results also have value for potential users of the IMS data. With the exception of the the bias stability all other results are split diurnally. Therefore, we have updated the these result to also include the day night split:

Table 3. Stability of IMS water vapour profile biases at ARSA sites given as the trend in % ppmv/decade. Trends are reported for all global matches, split between all matches, daytime and night-time only IMS soundings, and a breakdown for each of the five broad latitude bands. It should be noted that the separation between day and night cases is more representative of seasonal differences (e.g. polar winter/summer) than diurnal ones for high latitudes. Values denoted with * are outside the 95% confidence interval.

Site-Pressure layer	Global	90°S-60°S	60°S-30°S	30°S-30°N	30°N-60°N	60°N-90°N
All soundings						
400-300 hPa	0.79±0.83	-10.71±4.45*	-0.29±0.73	0.57±0.66	2.04±0.99	-0.29±1.10
500-400 hPa	-1.76±1.19	-10.18±5.36*	-2.65±1.64	-2.52±1.23	-0.38±1.32	0.56±1.41
700-500 hPa	-1.34±1.34	0.11±4.91*	-3.32±1.55	-2.23±1.03	-0.74±1.38	1.20±1.81
850-700 hPa	0.93±1.01	-5.94±4.63*	0.74±1.40	-1.32±0.89	2.16±1.62	3.38±1.84
925-850 hPa	0.21±0.69	0.97±3.34*	1.42±1.09	-1.22±0.78	0.86±0.97	1.41±1.22
1000-925 hPa	-0.25±0.91	2.61±3.53*	1.31±0.93	-2.31±1.05	0.61±0.84	0.96±1.26
Daytime soundings						
400 to 300 hPa	1.16±0.95	-7.38±6.69*	-0.38±0.88	1.27±0.63	1.84±0.85	2.00±2.02
500 to 400 hPa	-2.46±1.35	-4.76±8.29*	-2.49±1.64	-0.66±0.54	-2.68±1.55	0.95±2.40
700 to 500 hPa	-2.45±1.66	-1.72±9.17*	-3.88±1.67	-1.47±0.74	-2.67±1.40	2.04±2.45
850 to 700 hPa	1.30±1.38	-11.37±7.24*	0.45±1.47	0.06±0.46	1.73±1.71	3.11±1.80
925 to 850 hPa	1.09±1.12	2.14±4.27*	1.49±0.95	0.80±1.07	1.13±1.09	1.35±1.78
1000 to 925 hPa	0.81±0.94	3.01±4.48*	1.89±0.95	-0.41±0.73	1.50±1.15	-0.75±2.04
Night-time soundings						
400 to 300 hPa	0.38±1.17	-7.54±6.89*	0.51±1.15	-0.74±1.10*	2.36±1.25	-0.35±1.14
500 to 400 hPa	-0.55±1.14	-17.96±8.26*	-1.74±1.83	-5.81±2.30	2.94±1.18	1.85±1.56
700 to 500 hPa	0.42±0.83	-3.17±6.85*	-2.29±2.01	-3.85±1.85	2.25±1.66	2.25±1.84
850 to 700 hPa	0.45±0.59	-6.24±7.27*	2.71±1.64	-3.45±1.65	2.76±1.31	3.10±1.65
925 to 850 hPa	-0.83±0.85	11.22±12.29*	1.44±1.28	-4.89±2.14	0.73±1.31	2.06±1.33
1000 to 925 hPa	-1.57±1.10	8.11±9.02*	0.21±0.90	-5.98±3.08	-0.34±1.11	0.76±1.28

Table 4. As Table 3 but for temperature profile biases. Trends are given in K/decade.

Site-Pressure layer	Global	90°S-60°S	60°S-30°S	30°S-30°N	30°N-60°N	60°N-90°N
All soundings						
400-300 hPa	-0.12±0.12	-0.18±0.15	-0.18±0.12	-0.26±0.16	-0.11±0.09	0.31±0.24
500-400 hPa	-0.23±0.21	0.00±0.38	-0.28±0.11	-0.33±0.31	-0.17±0.25	0.28±0.79
700-500 hPa	-0.32±0.18	0.03±0.22	-0.34±0.15	-0.41±0.36	-0.28±0.13	0.19±0.87
850-700 hPa	-0.14±0.21	0.22±0.39	-0.13±0.16	-0.21±0.14	-0.12±0.27	0.27±0.97
925-850 hPa	0.08±0.15	0.26±0.49	0.13±0.18	0.11±0.24	0.07±0.17	0.30±0.59
1000-925 hPa	0.10±0.27	0.15±0.41	0.28±0.18	0.12±0.10	0.13±0.24	0.31±0.28
Daytime soundings						
400 to 300 hPa	-0.14±0.13	-0.11±0.23	-0.23±0.14	-0.24±0.14	-0.14±0.11	0.24±0.18
500 to 400 hPa	-0.24±0.13	0.06±0.39	-0.28±0.10	-0.28±0.23	-0.20±0.21	0.26±0.53
700 to 500 hPa	-0.36±0.18	0.01±0.33	-0.38±0.16	-0.38±0.29	-0.37±0.13	0.22±0.44
850 to 700 hPa	-0.17±0.17	-0.01±0.53	-0.18±0.22	-0.22±0.14	-0.20±0.28	0.37±0.53
925 to 850 hPa	0.13±0.22	-0.01±0.60	0.11±0.27	0.18±0.31	0.08±0.32	0.43±0.38
1000 to 925 hPa	0.17±0.28	-0.11±0.48	0.22±0.23	0.25±0.09	0.16±0.25	0.33±0.21
Night-time soundings						
400 to 300 hPa	-0.09±0.08	0.46±0.35	-0.09±0.07	-0.30±0.13	-0.06±0.06	0.36±0.21
500 to 400 hPa	-0.20±0.29	0.44±0.38	-0.26±0.10	-0.42±0.31	-0.12±0.27	0.35±0.79
700 to 500 hPa	-0.24±0.29	0.44±0.30	-0.26±0.08	-0.46±0.32	-0.17±0.24	0.27±0.96
850 to 700 hPa	-0.07±0.41	0.62±0.47	-0.14±0.09	-0.17±0.11	-0.03±0.43	0.30±0.98
925 to 850 hPa	0.03±0.13	0.59±0.73	0.02±0.08	-0.03±0.24	0.06±0.14	0.31±0.54
1000 to 925 hPa	0.00±0.23	0.46±0.70	0.21±0.28	-0.15±0.68	0.06±0.40	0.29±0.27

With these results we still observe some differences and similarities between the day and night results. However, 50% of the layers analysed are still within the trend uncertainty mainly due to the length of the time series. Therefore, we feel our final conclusion, that this is (to our knowledge) the first published looking at bias stability of IASI profiles and that we require longer time series to really investigate this properly. This is especially true when other IR sounders are processed with the IMS algorithm. Additional platforms with different sounding characteristics and overpass times will likely introduce other features that add to the noise, which in turn require longer time series for signal detection (Weatherhead et al. 1998). The new diurnal results point to a broader study where the effects which impact radiosonde biases also need to be accounted for in context satellite profile bias stability.