

Dear Editor,

Pleased find below how we have addressed the comments of both referees. I would like to mention that we would like to change the title of the paper « Homogeneity criteria from AVHRR information within IASI pixels in a Numerical Weather Prediction context ». We have added a study on a homogeneity evaluation with independent SEVIRI data, changed most of the figures.

Best regards,

Nadia Fourrié

Reviewer 1 :

We thank Reviewer 1 for his/her comments which helped to improved, we hope, the quality of the manuscript. Reviewer 1's comments are in bold font, our answers are written with normal font.

This manuscript looks at an interesting area and some new results are presented but without getting into much depth of analysis. There are gaps to be filled to support the conclusions. In particular it is hard to make an informed choice between the different inhomogeneity screening options based on this work, as only one of the viable options was tested in an NWP system (the other tested option already showed clear defects even before NWP system testing). There is also no independent validation of whether the screening achieves its goal, which is homogeneous scenes.

An independent validation was added. The cloud homogeneity was compared with an homogeneity criterion based on cloud types retrieved from SEVIRI observations. We found similar results as those obtained with the AVHRR cloud cover. In addition we have removed one of the data assimilation experiment in order to keep only the experiment with the COMPR criteria.

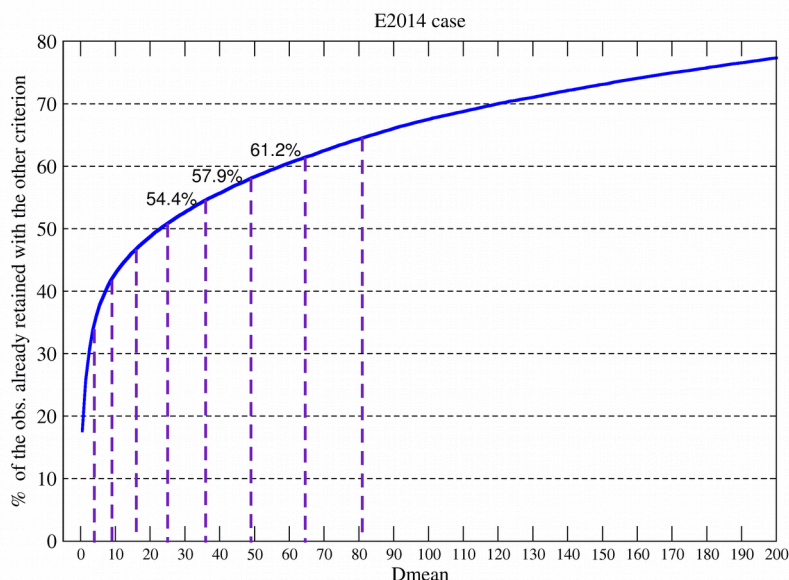
Major points

1) P6 L9: "we plan to assimilate clear or cloudy observations that are completely covered in the IASI FOV discarding fractional cloud observations". This still allows the possibility of fully clear obs being assimilated in a fully cloudy model (or vice-versa). Is that the intention? How do these screening methods treat cloudy modeled scenes if at all? The text should explain. In the context of the preparation of all-sky assimilation, we plan to assimilate indiscriminately clear or cloudy observations that are completely covered in a homogeneous way, discarding the cases of fractional cloud observations : the clear cases would be assimilated as in the current operational version and the cloudy ones in manner to be determined. The comparison to modeled scenes is ensured with the background departure check. In this case it will not be possible to keep a clear observation with a fully cloudy model.

2) The choice of $49K^2$ departure threshold in 3.2.2 is unsupported and uninvestigated in the text. Important questions are what this threshold means in terms of retained cloudy scenes, and how do its effects differ from those of the 7K AVHRR departure check in the Martinet et al. (2013) approach? Ultimately it should be investigated why the adapted (it is not the original) Eresmaa (2014) technique provides poor cloud screening here. Maybe it is the adaptation of this departure check? Finally, it is probably a case of poor wording rather than science, but it seems incorrect to claim the departure threshold as a check on homogeneity rather than just on cloud (P10 L19).

The choice of $49K^2$ departure was studied with the graph of leaving observations as a function of the departure. As shown below, this threshold allows to keep more than 50 % of the observations. That is why this threshold was kept. In addition it fits the 7K AVHRR departure check of M2013.

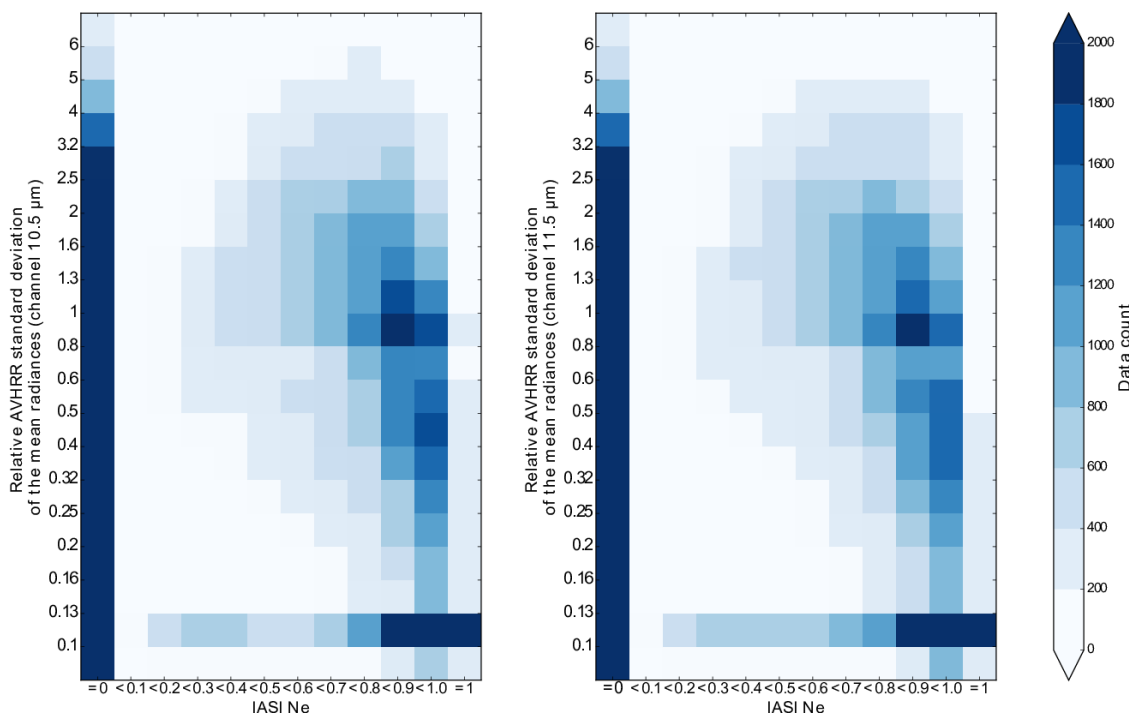
It is proposed in the text the following sentence « is less than $49K^2$. This particular value of threshold allows to keep more than 50 % of the observations compared to the initial threshold of $1K^2$ by Eresmaa (2014) which retains only 10 % of the observations. In addition this threshold compares well with the one applied by M2013, but it is applied over the 2 IR AVHRR channels. » The compromise selection method is an adaptation of the E2014 and is strongly based on it as we used the 2 IR AVHRR channels and the distance D_{mean} proposed by E2014 is used for the background check.



We agree that the D_{mean} based check cannot be considered as a homogeneity check and the sentence (initially P10 l19) has changed : « In this test, we used the D_{mean} proposed by Eresmaa (2014) to perform a kind of cloudiness consistency check between the observation and the model simulation »

3) The choice of the 0.8% threshold on p11 is barely supported in the text or by Figure 2. It may be the colour scale but there seem to be no highly inhomogeneous scenes according to the IASI Ne (e.g. between 0.3 and 0.7 on Fig. 2) and there certainly seems no correlation between the relative cluster standard deviation and the Ne.

We recognise that the original Figure 2 was not clear with the interpolation. Figure 2 has been plotted with another color scale, a logarithm scale on y axis and as a function of data count.



As can be seen in the Figure, the threshold of 0.8% allows to remove 39,8% and 42,0% of observations for each AVHRR channel.

4) P11 L17 "Similarly in model space the D_mean..." Since D_mean is based on observation minus background it is not in model space and neither is it a direct indication of the model cloudiness.

We agree that D_mean is computed in the observation space. We propose the following rewording : « Similarly, we used the background departure check in the observation space D_mean (presented in section 3.2.2).

5) "The percentage of cloudy AVHRR pixels" - if this is a good enough indicator of fractional cloud and/or inhomogeneity to use it to validate the screening criteria, why is it not used as part of the screening criteria? We agree that this percentage of cloudy AVHRR pixels is not so good indicator of the presence of homogeneous clouds to assess the screening criteria. As said above, we propose here an independent evaluation against SEVIRI cloud type. The results obtained with these data are similar to those obtained with the AVHRR cloud cover.

6) Section 5, the intercomparison of selection criteria, does not fully make the case for the proposed selection method. M2013 keeps 29% of data in table 2 compared to 21% in the compromise method, with only slightly higher standard deviations. That could be a good choice, but it has been rejected at this stage. The balance between a slight increase in std. dev. and gaining extra data has hence not been properly explored. It seems odd to instead test the "Obs_HOM" approach in data assimilation as already from the intercomparison it is clear it does not work well. Further, without an exploration of its sensitivity to threshold choices, the E2014 test does not have much of a chance in this intercomparison. Finally, each of the previous techniques M2013 and E2014, as well as the newly proposed compromise technique are composed of two tests, and it would be good to know how many rejections each is responsible for and how much overlap there is between the two tests.

We agree with the reviewer that the issue here is to find a compromise between good statistics and a sufficient number of observations for the assimilation. Here the Obs_HOM approach clearly keeps too much observations for the assimilation. In section 5, this experiment with the Obs_Hom Criteria was removed and we keep only the COMPR criteria for the data assimilation experiment. The COMPR method results from the E2014 as the D_mean criterion was used.

Please find below a table summarizing the impact of each criteria for each method.

	% of remaining observations	When 2 criteria applied	When all criteria applied
M2013			
Sigma inter 1	89,8%	84,8%	53,9%
Sigma intra 1	87,8%		
Background check 1	59,4%		
E2014			
Crit 1 can 1	39,8%	39,4%	22,8%
Crit1 can 2	42,9%		
Dmean<49K	42,7%		
COMPR			
obs_hom 1	68,2%	67,3%	36%
Obs hom 2	69,6%		

Dmean<49K ²	42,7%		
OBS_Hom 1+Dmean 1	43,3%	36%	
obs_Hom 2+Dmean 2	44,7%		

As can be seen in the previous table, if tests based on a single channel allow to keep more or less the same percentage of observations, the combination of both channels lead to a more restrictive choice.

7) Much of the conclusions will need to be updated to reflect a more thorough comparison of the different methods but particularly problematic is the statement P19 L4-6, saying that the M2013 and E2014 techniques were unsatisfactory due to a large loss of observations. Since M2013 provides more observations than the proposed method according to table 2, this statement is wrong. Also E2014 was not well explored in terms of thresholds and other possible adaptations to make it work in the current framework, using all-sky forward radiative transfer. Any rejection of this technique needs to be carefully qualified.

We have removed the sentence. The conclusions of this study were updated.

A new method using of collocated AVHRR cluster information to improve the selection of homogeneous IASI observation scenes within the numerical weather prediction ARPEGE model has been developed at Météo-France for data assimilation purposes and has been presented in this study.

The first step consisted in adapting the IASI observation operator based on the RTTOV radiative transfer model by using the RTTOV-CLD module with cloudy microphysical parameters (liquid water content (ql), ice content (qi) and cloud fraction) for the simulation of cloudy radiances. A qualitative evaluation of such module showed realistic simulated cloud structures at various locations around the globe with a quite good agreement against IASI observations.

The second and main step of this work was to assess the impact of several methods used to select homogeneous IASI observations using AVHRR clusters. Two selection methods (derived from the literature : Martinet et al. (2013) and Eresmaa (2014))) were preliminarily evaluated. Despite a good improvement in terms of biases and standard deviations of the FG departures, **it was found that these two methods were not satisfactory in an operational context (in assimilation) to a large IASI observation reduction. the criteria from the Martinet et al. (2013)'s method favours the homogeneous cloudy observations and retains more than a half of the observations while the Eresmaa (2014)'s method gives priority to clear observations and keeps only 22% of the observations. Then, two new sets of criteria were defined from these two previous methods in order to have a more balanced choice of clear and cloudy observations and good statistics in terms of background departures and implemented within the ARPEGE model :**

- The first criterion **derived from Martinet et al. (2013) method** looks for the consistency between different clusters occupying the same IASI FOV by examining this homogeneity relative to the weighted average brightness temperature of the AVHRR clusters; it is only based on observations **and computed for both infrared AVHRR channels as in Eresmaa (2014). This criterion allows to retain 67% of observations.**

- In addition, the second criterion **is derived from Eresmaa (2014)'s test and** assesses the coherence of each cluster compared to the background **brightness temperature simulation** ; it is in fact a good compromise between **the previous criterion and** the two "historical" ones with accurate statistics and a sufficient number of observations **36% that passed the check. It also allows to retain the same proportion of homogeneous clear and cloudy observations contrary to the derived Martinet et al. (2013) and Eresmaa (2014) methods.**

Therefore, assimilation experiments were conducted to assess the impact of these new selecting homogeneous IASI observation features in the current clear sky assimilation. This revised check was added to the McNally and Watts (2003) cloud detection. The results obtained in this case show that the scenes categorization has been facilitated and cloudy observations can be better filtered out compared to what is done in the operational ARPEGE version. 3% of all observations are rejected with the compromise method **for the assimilation**. The impacts on the first guess and analysis departures (showing more Gaussian shape) are generally low but with a beneficial reduction on the standard deviation of first guess departures mainly on the IASI and AMSU-A observations. Regarding the forecasts scores, neutral impact is reported when these selection criteria are taken into account on top of

the McNally and Watts (2003) algorithm.

Minor points

1) Introduction: an up-to-date reference on homogeneity criteria for all-sky is the following. It should be discussed: Okamoto, K. (2017), Evaluation of IR radiance simulation for all-sky assimilation of Himawari-8/AHI in a mesoscale NWP system. Q.J.R. Meteorol. Soc., 143: 1517-1527. doi:10.1002/qj.3022 A reference to the work by Okamoto (2017) was added in the text, after the section on the selection of cloudy scenes based on cloud homogeneity. « Okamoto (2017) studied the impact of the super-observation homogeneity quality control on the Advanced Himawari Imager brightness temperature simulation. He concluded that for larger size of super-observations, the standard deviation threshold should be relaxed in order to keep sufficiently low brightness temperatures associated with high-level cloud because of the presence of more cloud heterogeneity in large size observations»

2) P3 L31 "the first level at 10m" sounds odd; surely the lowest 10m of the atmosphere is also included in the model? The first level of the ARPEGE is well at 10m. Fields below this level are interpolated from the model and the surface.

3) P5 L28 "Stratus Continental and Stratus Maritime" are cloud microphysical options in RTTOV; this should be stated; also it should be made clear how the choice is made, even if it is as obvious as using the land-sea mask. The cloud type chosen depends on the land-sea mask of the model. Over land, the cloud type stratus continental is chosen, the stratus maritime is used over sea. The sentence was reworded as follows : « To simulate the radiances observed in cloudy conditions using RTTOV-CLD, we use two main cloud types : firstly liquid water cloud which corresponds to two RTTOV-CLD cloud microphysical options depending on the land sea mask of the model (Stratus continental over land and Stratus Maritime over the sea), secondly the ice water cloud of the Cirrus type, using Baran parameterisation (Baran et al 2014 and Vidot et al 2015) to define the optical properties

4) P10 L1 "Background brightness temperature for AVHRR" - for clarity explain if this is clear-sky or all-sky. Here this background brightness temperature is simulated with clear sky as in Eresmaa (2014). This was clarified in the text « where R_i^{BG} is the clear-sky background brightness temperature for AVHRR channel i ».

5) P10 L7 If f_j is different from C_j please explain how; otherwise use consistent notation. f_j and C_j are the same quantity, the fractional coverage of the cluster. f_j was changed into C_j .

6) P11 L10 Instead of "relationship" is "ratio" intended? At present the text is imprecise. Relationship was changed into ratio.

7) P12 L15 "Mean standard deviation" is always a confusing phrase and needs explanation of what samples got meant or standard deviated and in what order. The paragraph was rewritten and there is no more mean standard deviations.

8) Table 2 should additionally include statistics for the observations that are fractionally cloudy according to AVHRR. This column was added in Table 2.

9) P12 L15 to P14 L10 is hard to read as it is overloaded with numerical results and bereft of much interpretation. Many of the numerical results are already listed in table 2 and do not need exhaustive restating in the text, which needs to be rewritten with a higher level of

analysis for the reader. Where numerical results given in this text are not already in tables, they should be.

This part of the paper was rewritten and numbers scattered along the text were removed. "The percentage of cloudy AVHRR pixels in the IASI field was also used to assess the choice of homogeneity criteria (Table 3).

Our global data set is made of 50% of the observations entirely covered by clouds and 12% of clear observations according to the AVHRR cloud cover. These results obtained over the globe set agree well with the ones obtained with SEVIRI data over the Atlantic Ocean. Except for M2013 method, the percentage of selected observations for each selection method is larger (+2/4 %) with the SEVIRI data evaluation than with the AVHRR cloud cover. The bias and standard deviation of observations minus simulations (O-G), are shown in Figure 3.(a) for the 314 IASI channels. As expected, the best statistics are obtained for channels less affected by clouds (e. g. CO₂ and water vapour high peaking channels).

For the whole dataset, window channels present a bias of around -0.6K. The standard deviations are larger (around 12 K) for window channels sensitive to the surface. With the M2013 selection method (figure 3. (b)), the standard deviation of window channels is reduced to around 4 K as well as the bias close to zero. The standard deviation of the other channels (680-780cm⁻¹ is also well decreased. The E2014 selection method (figure 3. (c)) improves the bias and the standard deviation (2.0 K for window channels) for all the channels. As expected, the impact is larger for surface sensitive (and thus cloud sensitive) channels than for the tropospheric channels (680-780cm⁻¹). On the contrary, with the Obs_HOM method (figure 3. (d)), small statistics improvement is obtained for the standard deviation and the bias. The statistics obtained with the COMPR method (Fig. 3.(e)) are reduced compared to the whole data set and slightly less good than with the initial E2014 method (for window channels the standard deviation is around 2.2 K instead of 2 K for E2014)."

10) P14 L22 "Gaussian" is overly strong here and is not backed up by any statistical tests of Gaussianity. We propose the following modification « the O-G distributions become symmetrical and get closer to the gaussian distribution »

11) P14 L30 "we keep more observations" - which method is referred to by "we"? This corresponds to E2014 method and it was changed in the text.

Response to Reviewer 2.

We thank Reviewer 2 for his/her comments which helped to improved, we hope, the quality of the manuscript. Reviewer 2's comments are in bold font, our answers are written with normal font.

General comments:

The English is generally very good, the paper is well structured and nicely concise. I am proposing major revisions because some of the plots should be remade, and the scope of the paper does not seem to fit the ambitions of the title. I found the choice of experiments a little strange, and was rather confused on the details. However, whilst I am suggesting some extensive changes, I think that these matters could be fixed quite easily by the authors.

To agree with the content of the paper, we suggest to change the title of the paper with « Homogeneity criteria from AVHRR information with IASI pixels in a Numerical Weather Prediction context ». In addition we have removed the data assimilation experiment with the only Obs-Hom criteria and only kept the experiment with the compromise criteria. Some parts of section 4 were rewritten as well as the conclusions.

I realised when I reached the final sections of this paper that I had totally misunderstood its intention. I had expected that the homogeneity criteria would be used to select additional observations, that are homogeneously cloudy, to assimilate in addition to the clear-sky

channels accepted by the McNally and Watts check - i.e. to do something similar to McNally (2009). It took me almost the whole paper to realise that what is being proposed is additional quality control on radiances already accepted by McNally and Watts. This does not seem like “preparation for all-sky assimilation” and I think that the scope of the paper should be revised.

We recognised that the title of the paper may be confusing that is why we propose to change the title of the paper with « Homogeneity criteria from AVHRR information within IASI pixels in a Numerical Weather Prediction context » The objective of the paper in the last paragraph of the introduction was also modified into : « Our objective is to determine homogeneity criteria valid for both clear and cloudy conditions, suitable to an NWP context using collocated AVHRR and IASI information. »

At least, I assume that it is the case that this is just extra QC on clear sky calculations I confess that I found the paper surprisingly confusing! There is no mention that RTTOV-CLD is being used in the assimilation experiments, therefore I assume the calculations are clear sky, and there is no mention of the use even of a single grey-cloud layer scheme in use as in McNally (2009).

We propose an ensemble of homogeneity criteria. In case of clear sky, the McNally and Watts is applied as an additional QC test afterwards. In case of cloudy scenes, these criteria could pave the way to an all-sky assimilation (methodology not decided yet). You are right in the assimilation experiments, only RTTOV was used (clear sky assimilation). RTTOVCLD was only used to compute the homogeneity criteria based on cloudy AVHRR simulations. In the operational version of ARPEGE , a single layer grey cloud scheme is used (Guidard et al 2011) but in the experiments carried out in this paper, this possibility has been switched off to focus on clear sky assimilation.

This has been specified in the text : « In these experiments, no cloudy observations detected with the CO2-slicing method and used with a single grey-cloud layer scheme was assimilated unlike in Guidard et al. (2011). RTTOVCLD was only used to compute the homogeneity criteria based on cloudy AVHRR simulations and RTTOV was used for the clear sky assimilation. »

Some curious choices are made throughout the paper: I believe that the E2014 method was adapted because Eresmaa's intention was to keep only clear scenes, whereas you wish to allow through homogeneously cloudy scenes also, but in fact because you are rejecting observations that had already been allowed through by McNally and Watts, I don't see why you don't just apply Eresmaa's method without modification. Why do you not include the scheme of McNally (2009)?

The objective of Eresmaa (2014) method is to keep clear pixels. Our objective is to keep homogeneous scenes both in clear and cloudy conditions. The Eresmaa method is not suitable in our case and this why we propose the modification based on this method. The single layer grey cloud scheme could have been used but the selection of cases to be assimilated this way would need a dedicated study.

And finally, perhaps most surprisingly, you use the AVHRR clear/cloudy pixel fraction as a measure of whether the homogeneity criteria have “correctly” picked out homogeneous scenes, and on p15 you state that you are happy to accept a reasonable proportion of observations with >90% cloud cover. Why not just test the use of the AVHRR clear pixel fraction? And yet, you performed assimilation experiments with a scheme that you had seemed to reject based on the O-B statistics presented in Figure 4.

As suggested by reviewer 1 we added an evaluation with independent data of cloud type from SEVIRI on board MSG satellite in section 4 because the cloud cover from the AVHRR has some defects. We agree that observations with >90 % cloud cover are not necessarily homogeneous.

We have now removed the data assimilation experiments with obs_HOM criteria and we kept only the compromise method.

It is not really surprising that there is little impact, as very little seems to have changed in the experiments relative to the control. The work therefore seems rather immature for a publication. You still apply the CO2-slicing method (p16 line 17) - presumably this is designed to reject cloudy scenes? What effect does this have on the homogeneous scenes?

The sentence p16 l 17 was confusing and thus modified as said above : « In these experiments, no cloudy observations detected with the CO2-slicing method and used with a single grey-cloud

layer scheme was assimilated unlike Guidard et al. (2011). RTTOVCLD was only used to compute the homogeneity criteria based on cloudy AHVRR simulations and RTTOV was used for the clear sky assimilation. »

3 % of observations are rejected between Exp and Ref, so you are correct that it is not very surprising that there is very little impact. Nevertheless, this impact is slightly positive. Thus it seems that this method is reliable enough to be used in a NWP assimilation. These criteria will reveal their full potential in an all-sky assimilation. The design of a IR all-sky assimilation has still to be done.

Specifics:

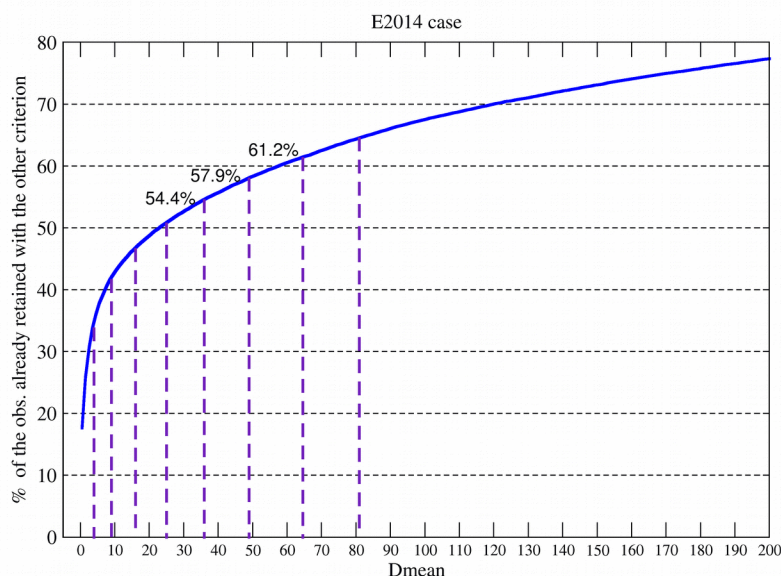
P9 line 7: it is not clear whether this 7K check is an additional criterion over the 8% criterion in the previous paragraph. Why 7K? I also don't understand "interpolated using 12 points" - is that 12 points in 3D? This 7K check is an additional check verifying that both the observation and the model observe the same cloudy scene. In the original Martinet et al (2013) it was applied to the difference between the mean AVHRR brightness temperatures from the observed and simulated clusters. In the case of ARPEGE the horizontal mesh is coarser and we replaced this fine scale check with the difference between the brightness temperature from the guess profile and the observation. The guess profile results from a horizontal interpolation of 12 profiles surrounding the observation position coming from a 6-hour forecast.

We have added this point in the text : « In the original Martinet et al (2013) study, a third check verifying that both the observation and the model observe the same cloudy scene was done with the difference between the mean AVHRR brightness temperatures from the observed and simulated clusters less than 7 K. Here, the ARPEGE model has a coarser resolution and it is not possible to simulate the AVHRR clusters. This check was adapted with the difference between the AVHRR observation and the AVHRR simulation from the guess, which come from a horizontal interpolation of the 12 profiles surrounding the observation position coming from a 6-hour forecast. »

P10 line 14-17: I don't understand how this first bullet point is different from the original method. The difference lies in the fact that the simulation computation were done with RTTOV-CLD. This is mentioned now in the text. « All AVHRR simulations from background are made with RTTOV-CLD and the threshold of the background departure check was modified » The description of the threshold for inter-cluster homogeneity was thus removed.

P10 line 18-19: Why 49K²? Other than that it fits the 7K applied to M2013, it seems quite high relative to 1K.

Many trials were done for the value of D_{mean} . Below you will find the function of leaving observations as a function of D_{mean} . It appears that this value allows to keep more than 50 % of the observations. In addition it also fits the 7K threshold of M2013 but over the 2 IR AVHRR channels.



We used the D mean proposed by Eresmaa (2014) to perform here a kind of cloudiness consistency check between the observation and the model simulation if D mean is less than 49 K^2 . This particular value of threshold allows to keep more than 50% of the observations compared to the initial threshold of 1 K^2 by Eresmaa (2014) which retains only 18% of the observations. In addition, this threshold compares well with the one applied by M2013, but it is applied over the 2 IR AVHRR channels. The text was modified accordingly.

P11 line 5: This is the same as the first test of Martinet but with 2 channels. It would be clearer if this was stated. Why change the L to R in the equation? What does the addition of the second channel bring in practical terms?

It was done : « It is the same test as in M2013 but in the brightness temperature space ». The change of L into R is practical as it is easier to work with brightness temperatures. We chose to have all value in brightness temperatures. If we only consider channel 1, we keep 68,2 % of the observations, with channel 2 69,6 % are remaining but with both 67,3 % of the observations are selected. This was mentioned in the text : « If this test is only applied over channel (10:5 m, 68,2% of the observations are selected, if applied over channel 11:5 m), 69,6% of the data are kept and if it is applied over both channels, 67,3% pass the test. »

P12: What is this dataset of 59 million observations? Is it just 24 hours' worth of observations? You state that 50% of the observations are 100% cloudy - that sounds potentially high for a normal dataset? There was a confusion between the number of observations and the number of channels. In fact the statistics were computed over 188090 IASI observations for 30 January 2017 and the number was corrected in the text and in the table. Half of the observations having a cloud cover of 100 % may seem a very high percentage, but with the independent validation we proposed with SEVIRI data, we find the same result.

P12/Figure 3: Are the numbers in the text for bias and SD an average over a number of channels, or the maximum value from the windows? It should probably be the latter. I cannot match the figures in the text with the plots - the numbers do not seem to match (e.g. 11.7K bias -> the bias looks over 12 K in the figure). It would also be better to just plot Band 1 so we can see the effect on the temperature channels. I think the numbers scattered over several paragraphs and two pages would be better in a table.

Figure 3 was changed and represents now only the spectral range of band 1. Section 4 was partly rewritten. This section was re-arranged and there are now less numbers in the text. Tables were also modified in order to show statistics on window channels and CO2 channels.

P14 line 15-20: I would disagree that the distribution asymmetry is small. I also disagree that the Obs_HOM approach reduces the range of the tropospheric water vapour channel distribution. Both sentences were removed and changed with. « The distribution asymmetry is reduced for mid and low tropospheric water vapour channels with M2013 and E2014 selection. » for the first one.

P15: The discussion focuses on letting through the most data - this isn't necessarily the best criterion, as you may be letting through inhomogeneous scenes. There is trade off between more data and better data. M2013 and Obs_HOM let through a lot of partially cloudy scenes (and even 100% cloudy scenes may have different cloud types in one pixel). We agree that 100% cloudy scenes may have different cloud types in one pixel. Now the discussion is based on SEVIRI cloud which is an independent validation datum. We also agree that there is a trade-off between more data provided by M2013 and Obs_HOM and better data. That is why we propose a fourth method based on the Eresmaa (2014) test.

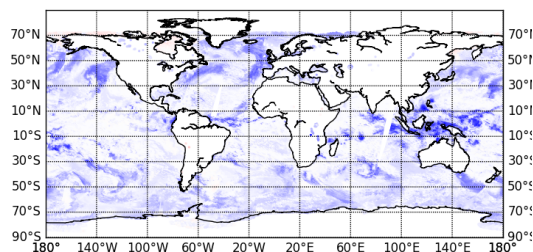
P16: It is not clear what the set-up for the assimilation experiment is - you do not mention RTTOVCLD - presumably this is still clear sky.

Now it is clearly stated : « RTTOVCLD was only used to compute the homogeneity criteria based on cloudy AHVRR simulations and RTTOV was used for the clear sky assimilation. »

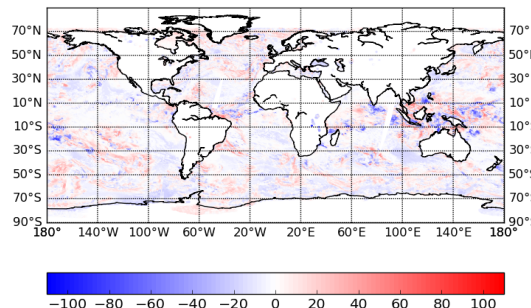
Figures:

- **Figure 1: would be better as two bigger O-B plots.** We have replaced the simulation plots with the suggested O-B plots. The comments on Figure 1 have been changed accordingly :

b) Observations minus RTTOV (clear-sky) simulation



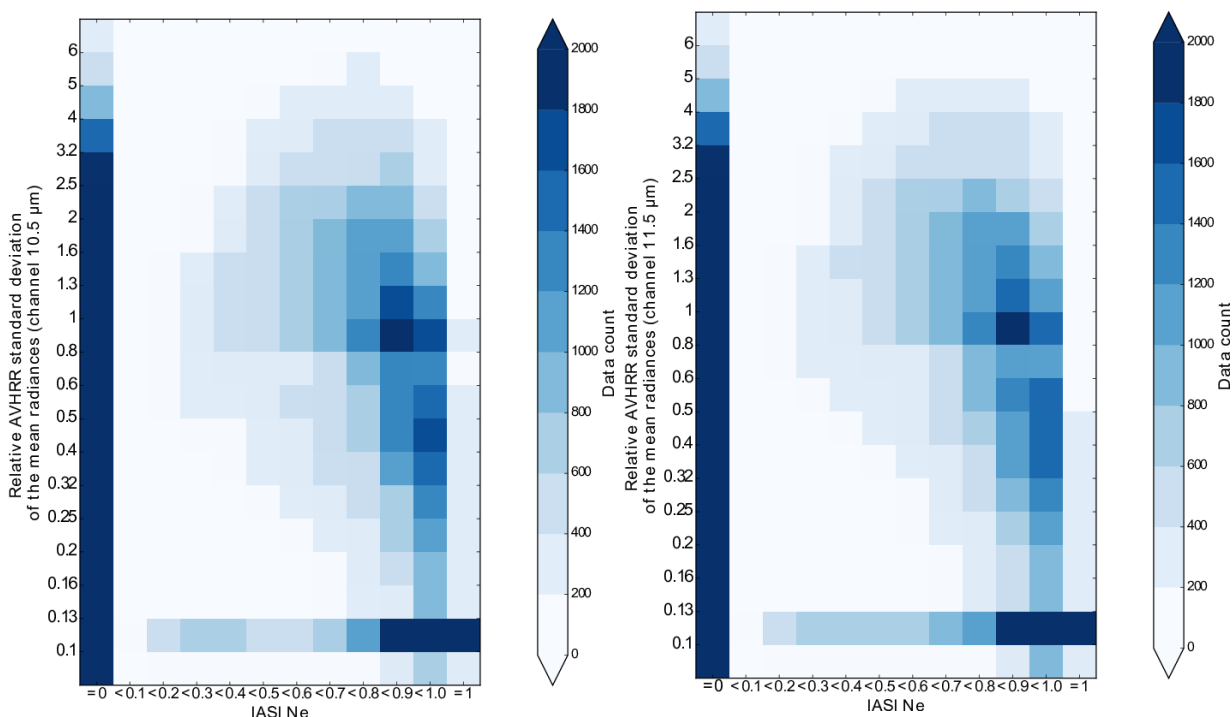
c) Observations minus RTTOV-CLD simulations



« To illustrate the benefit brought by RTTOV-CLD, Figure (1) shows IASI brightness temperature observations of a cloud-sensitive surface channel (1271, 962.5 cm⁻¹) and differences between the observations and the simulations computed with RTTOV considering clear-sky and with RTTOV-CLD. Brightness temperatures less than 250 K are usually associated with higher elevation cloud structures. By using RTTOV in clear sky (figure 1.b) to simulate IASI observations, the observation departures are mainly below zero and may reach up to -60 K. This can be

explained by the fact that the main cloud structures associated with low values of brightness temperature for the surface channel are missing in the simulation. On the contrary, differences obtained with the RTTOV-CLD simulations are in overall in better agreement with lower positive and negative values (figure 1.c). No major differences are found for example for the cloud structures located over the North Atlantic (30N-70N, 40W-0W) and above (30S-70S, 60W-0W) the Southern Atlantic Ocean. Large difference values are mainly obtained in the Tropics region. This may be explained by the fact that clouds are better simulated in the ARPEGE model for mid-latitudes than in the Tropics. »

- **Figure 2: very strange Y-axis. You can't see much on these plots. Is the Y-axis expressed as % as in the criterion on p11 line 11?** Yes the Y-axis is expressed as % as in the criterion line 10 page 11. The figure changed, with a log scale on the y axis, another color scale and as a function of data count.



- **Figure 4: Why not plot a temperature sounding channel? The x-axis has strange divisions. It would be better symmetrical.** Panels of Figure 4 were drawn again in order to have symmetrical x axis as shown below. We do not consider temperature sounding channel as we are more interested in humidity and clouds.

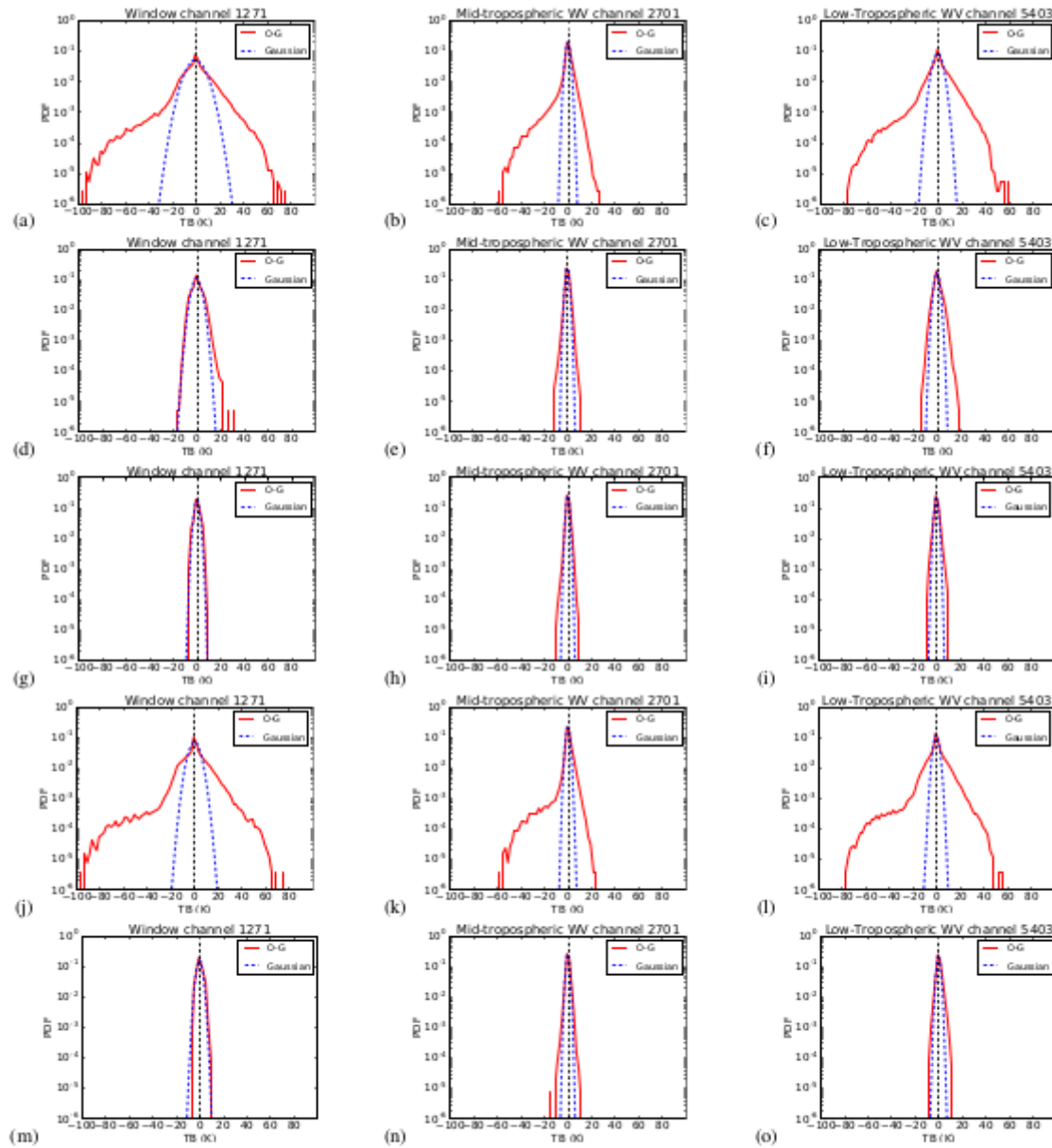


Figure 4. Frequency distribution of brightness temperature difference between observation and background (O-G) for all observations (a, b, c), after applying the homogeneity criteria derived from Martinet et al 2013 (d, e, f), the homogeneity criteria derived from Eresmaa 2014 (g, h, i), the third method based on observation space method (j, k, l) and the compromised approach (m, n, o). The PDF are presented for three channels: window channel 1271, low-tropospheric water vapour channel 5403, and mid-tropospheric water vapour channel 2701). The Gaussian distributions with the same error characteristics (mean and standard deviation) are also shown in blue dashed lines.

- **Figure 6: I honestly cannot see any difference between these three plots. You need to revise the colour scale to highlight the differences.**
As the figure is a bit redundant with figure 7 we propose to remove it.
- **Figure 7: I cannot see the REF line: is it under the green line or the red line? This is an important figure as it is the first time I realised this paper was about improved QC (more obs are assimilated with Experiment B than A).**

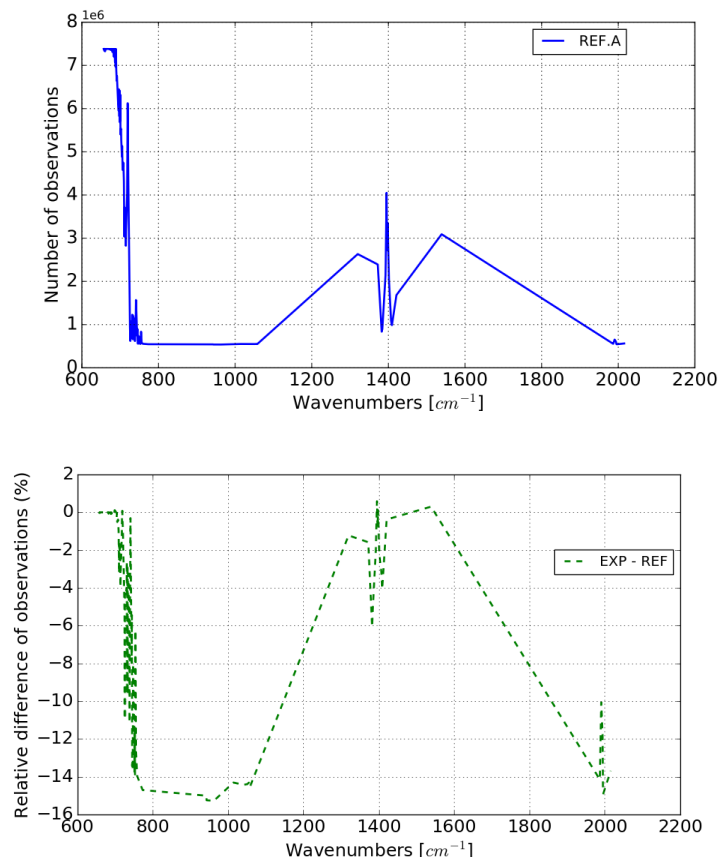


Figure 7 (a) has been changed with this upper-panel representing the number of assimilated data in REF and the second panel represent the the relative difference of number of observations for EXP (EXP.B) compared with REF (EXP.A being removed from this section). The caption was changed accordingly.

Tables:

- **Table 2: Should include the % partially cloudy**
The column of partially cloudy observations was added in Table 2.
- **Tables A1 and A2 are unnecessary - this information is presumably included elsewhere. If not, a simple list of channel numbers would suffice.**
Tables were removed.

Minor points:

- **P2 line 9: seems to be the first use of IR without the abbreviation being expanded.**
Done
- **Section 2: this section is a light-touch description of the model and IASI, as it should be, but it is important to get the details correct in that case and make sure the writing is clear: P4 has a few poorly worded sentences, or poorly explained concepts.**
 - **P4 Line 7 - the background error statistics are not “derived from a climatological matrix”- it isn’t actually a matrix, and you do not explain how the ensemble information is incorporated.** We agree that this sentence was not well written and we replaced it with « The background errors are computed at each analysis time based on the 25-member assimilation ensemble (see Berre et al 2015 for further details). »
 - **P4 Line 22 - this area needs rewriting - Presumably you mean that the accuracy of the forward model calculation is limited by the accuracy of the NWP model, and that for some variables this is not sufficient to correctly model the observations?** This paragraph was rewritten : The observation operator allows to simulate observations from the model variables for comparison with the actual measurements. For satellites radiances, it includes a radiative transfer model.

The accuracy of the forward model calculation could be limited by the accuracy of the NWP model, for some variables this is not sufficient to correctly model the observations and these observations have to be discarded.

- **“Modelisation” -> “Modelling” in English!** Corrected.
- **P4 Line 29 - The McNally & Watts scheme is not clearly described. This scheme is now better described in the text :** The McNally and Watts (2003) scheme intends to detect clear channels and to assimilate channels unaffected by clouds even in a cloud-affected pixel. The channels are first re-ordered according to a ranking with respect to the altitude that reflects their relative sensitivity to the presence of cloud. After having applied a low-pass filter a search for the channel at which a monotonically growing departure can first be identified. Having found this channel all channels ranked more sensitive are flagged as cloudy and those ranked less sensitive are flagged clear.
- **P4 Line 31: In this section, there are numerous references to CTOP and Ne, but suddenly you switch to PTOP - maybe Pangaud (2009) used PTOP instead of CTOP but this switch is not necessary.** PTOP was modified into CTOP.
- **P5 line 1: “IASI is a key element of the payload of the Metop series of European: :” P5 para 1: you may as well update this with Metop-C launch date.** The Metop-C launch data has been specified in then text: November 2018. « The third instrument was mounted on the Metop-C satellite, which was launched in November 2018. »
- **P5 line 22: It’s a bit far to say that failing to assimilate cloudy IR observations is a source of error.** This sentence was modified : “Assimilation of cloudy radiances is a crucial challenge for NWP centres as the cloudy observations discard represent an under-exploitation of hyperspectral sounders especially in sensitive meteorological areas (McNally, 2002; Fourrié and Rabier, 2004).”
- **P5 line 25: “allows to better describe” - not good english “allows a better description of: :”** Change made.
- **P5 line 29: You should reference the Baran parameterisation if you are going to mention it.** The reference to Baran, A. J., Cotton, R., Furtado, K., Havemann, S. , Labonnote, L.-C., Marengo, F., Smith, A. and Thelen, J.-C. 2014: A self-consistent scattering model for cirrus. II: The high and low frequencies. Q.J.Roy. Meteorol. Soc., 140: 1039–1057. doi:10.1002/qj.2193
- **P6 line 8: “an innovative challenge”? Remove the word innovative.** Done
- **P6 line 8: The sentence “In the context of...” doesn’t make sense.** The sentence was correted : « In the context of the preparation of all-sky assimilation, we plan to assimilate clear and cloudy observations that are completely covered in a homogeneous way, discarding the cases of fractional cloud observations. »
- **P6 line 21: They are not IASI L1c products - they are components of the L1c product.** This was corrected.
- **P6 line 24: this sentence is not clear either.** It has been reworded: For each class and each AVHRR channel, the cluster product provides the coverage of the class within the IASI pixel, the mean and the standard deviation of AVHRR brightness temperatures within the class.
- **P6: line 28: this sentence is not clear. I think it is a stretch to say something with one class can be less homogeneous than something with multiple classes - this is a bit subjective.** Our thought was that an important parameter is the standard deviation inside each class. The sentence was rewritten : A IASI FOV with several

classes, each one having a small standard deviation and a mean radiance close to the ones of the other classes, can thus be more homogeneous than a FOV with a single class but with very large value of standard deviations.

- **P9 line 11: “aimed to propose” - that is a bit of a negative slant on this reference! “Proposed” would be better!** The change was made.
- P17 line 15: No need to reference Table 3 here - it is a very basic table and you describe it all in the text, The Table and the reference were removed.

Homogeneity criteria from AVHRR information within IASI pixels in a Numerical Weather Prediction context

Imane Farouk¹, Nadia Fourri ¹, and Vincent Guidard¹

¹CNRM UMR 3589, Universit  de Toulouse, M t o-France, CNRS, Toulouse, France

Correspondence: Nadia Fourri  (nadia.fourrie@meteo.fr)

IASI, homogeneous scenes, clouds, ARPEGE model

Abstract. This article focuses on a selection of satellite infra-red IASI observations and their simulation in the global Numerical Weather Prediction (NWP) system ARPEGE (Action de Recherche Petite Echelle Grande Echelle), using the sophisticated radiative transfer model RTTOV-CLD which takes into account the multi-layer clouds and the cloud scattering from atmospheric profiles and cloudy microphysical parameters (liquid water content, ice content and cloud fraction). The aim of this work is to select homogeneous scenes by using information of the collocated Advanced Very High Resolution Radiometer (AVHRR) pixels inside each IASI field of view and to retain the most favourable cases for the assimilation of IASI infrared radiances. Two methods to select homogeneous scenes using homogeneity criteria already proposed in the literature were employed: criteria derived from Martinet et al. (2013) for cloudy sky selection in the French mesoscale model AROME (Applications of Research to Operations at MEscale), and the criteria from Eresmaa (2014) for clear sky selection in the global model IFS (Integrated Forecasting System). A comparison between these methods reveals considerable differences, both in the method to compute the criteria and in the statistical results. From this comparison a revised method is proposed that is a compromise between the different tested methods, using the two infrared AVHRR channels to define the homogeneity criteria in the brightness temperature space. This revised method has a positive impact on the observation minus the simulation statistics, while retaining 36% observations for the assimilation. It was then tested in the NWP system ARPEGE and tested for the clear-sky assimilation. These criteria were added to the current data selection based on the McNally and Watts (2003) cloud detection. It appears that the impact on analyses and forecast is rather neutral.

1 Introduction

Satellite observations are currently the dominant source of information for Numerical Weather Prediction (NWP) systems. Their assimilation together with in-situ observations give the atmosphere analysis, which is a necessary step in the definition of the initial conditions of the forecast. This analysis consists in finding a state of the atmosphere that is compatible with the different sources of observations, the dynamics of the atmosphere and a previous state of the model. In the M t o-France global model ARPEGE (Action de Recherche Petite Echelle Grande Echelle, Courtier et al. (1991)), 70% of used observations come from infrared hyperspectral sounders, of which IASI (Infrared Atmospheric Sounding Interferometer, Cayla (2001)) fills a large part. This sounder provides information about the atmospheric temperature and humidity, and through its window

channels information about the land surface parameters in clear sky and cloudy parameters can be obtained. However, the wealth of information provided by this type of sensor with its large number of channels or radiances (8461 per pixel in the case of IASI) and its overall coverage with a horizontal resolution of 12 km at nadir, is far from being fully exploited. Indeed, the presence of clouds in the instrument field of view, which affects the majority of observations, prevents from an accurate simulation of the radiances. In fact, NWP centres use only a small amount of observations from these sounders mostly in clear sky above clouds. Previous studies have shown that sensitive areas are often covered by clouds (McNally (2002), Fourrié and Rabier (2004)) and different techniques have been developed in the frame of global models to use infrared radiances in these regions.

In the past, different approaches have been proposed for cloud detection. A method to detect clear channels from high-resolution **InfraRed (IR)** spectral instruments was proposed by McNally and Watts (2003) to assimilate channels unaffected by clouds.

At the Met Office, Pavelin et al. (2008) showed that it was possible to assimilate cloud-affected infrared radiances when retrieved cloud parameters are used as set constraints. The cloud-top pressure (CTOP) and the effective cloud fraction (Ne) are firstly retrieved by a one-dimensional variational data assimilation system (1D-Var) and then provided to four-dimensional variational data assimilation (4D-Var) for the assimilation of cloud-affected infrared radiances. The analysis is significantly improved over the first guess by this method and it is used operationally to assimilate AIRS and IASI cloud-affected radiances. At ECMWF (European Centre for Medium-Range Weather Forecasts), McNally (2009) proposed a method based on two cloud parameters (CTOP and Ne) to assimilate cloud-affected IR radiances directly. In that case, the cloud parameters are determined with two channels and they are then introduced into the analysis control vector of the 4D-Var system of the global NWP model to constrain the minimization.

At Météo-France, the cloud parameters (CTOP and Ne) are retrieved for AIRS and IASI cloud-affected radiances with the CO_2 -slicing method Menzel et al. (1983). Channels affected by clouds, the cloud top pressure (CTOP) of ranges from 650 to 900 hPa with an effective cloud fraction (Ne) of 1, are assimilated in addition to clear ones in the ARPEGE 4D-Var and the AROME (Applications of Research to Operations at MEscale) 3D-Var Pangaud et al. (2009) and Guidard et al. (2011).

As pointed out by Errico et al. (2007), studies on the assimilation of clouds and precipitation from satellite sensors have started in the 80s and despite the encountered difficulties to implement them, operational weather centres are now assimilating them with a clear benefit for the forecast quality. Efforts started with microwave radiances and direct all-sky microwave radiance assimilation is effective at ECMWF (Bauer et al., 2010) since 2009 and at NOAA (National Oceanic and Atmospheric Administration) NCEP (National Centers for Environmental Prediction) since 2016. Even though ECMWF focussed on the assimilation of microwave imaging and humidity-sounding channels and NOAA NCEP, on the contrary, of temperature channels from Advanced Microwave Sounding Unit-A (AMSU-A), both centres noticed benefits of such an all-sky assimilation on the forecast quality (Geer et al., 2017; Zhu et al., 2016).

Concerning infrared radiance all-sky assimilation, no operational centre is yet to assimilate infra-red observations but research has still started in this area. Many aspects have already been studied as the information on cloud microphysics brought by the adjoint sensitivity in the assimilation (Greenwald et al., 2002) or by the retrieval of cloudy infrared radiances (Martinet

et al., 2013). In addition the sensitivity, the reproducibility and the nonlinearity of the simulation of IR radiances in the presence of multi-layer clouds was studied using diagnosed cloud schemes (Chevallier et al., 2004) and (Stengel et al., 2010). These studies also showed beneficial results.

A step further was achieved with the study by Okamoto et al. (2014). They studied the assimilation of multi-layer cloud-affected infrared radiances using the all-sky assimilation approach already implemented for microwave imager at ECMWF. They particularly investigated the cloud effects on the differences between observations and simulations and thus proposed an appropriate quality check and dedicated observation errors.

In this study we are interested in IASI observations, where the radiances are considered with colocated clusters statistical properties of the Advanced Very High Resolution Radiometer (AVHRR) also on board the Metop platform. AVHRR has a horizontal resolution of 1km at nadir (Cayla, 2001). Intuitively, colocated AVHRR data provide information on surface properties and the presence of clouds in the IASI Field Of View (FOV). They can therefore be used for cloud detection. The AVHRR cluster information associated with IASI has already proven to be useful for selection purposes in the context of cloud data assimilation, with an explicit treatment of microphysical variables in the AROME model by Martinet et al. (2013). Eresmaa (2014) at ECMWF also used AVHRR cluster information for cloud detection and observation selection in clear sky.

Martinet et al. (2013) selected cloudy scenes based on cloud homogeneity. This study was done in a 1D-Var framework using an advanced radiative transfer model (RTTOV-CLD) including profiles for liquid water content, ice water content and cloud fraction to simulate cloud-affected radiances as background equivalents to AROME fields. The persistence of cloud information brought by the analysis of cloud variables during a 3h forecast has then been evaluated successfully with an one-dimensional model AROME version (Martinet et al., 2014). Okamoto (2017) studied the impact of the super-observation homogeneity quality control on the Advanced Himawari Imager brightness temperature simulation. He concluded that for larger size of super-observations, the standard deviation threshold should be relaxed in order to keep sufficiently low brightness temperatures associated with high-level cloud because of the presence of more cloud heterogeneity in large size observations.

In this article, our objective is to determine homogeneity criteria valid for both clear and cloudy conditions, suitable to a NWP context using colocated AVHRR and IASI information. we try to determine whether or not colocated AVHRR and IASI information would facilitate the selection of homogeneous scenes which could be potentially used in an all-sky assimilation approach. Section 2 describes the ARPEGE NWP system, the IASI instrument and the radiative transfer model RTTOV-CLD in cloudy sky conditions. In section 3, information about the AVHRR clusters is detailed, the strengths and weaknesses of the different methods to select homogeneous observations are discussed, the chosen method is presented together with a description of the selected observations. Section 4 depicts the impacts on analyses and forecasts of selected clear and cloudy IASI observations. Conclusion and perspectives are given in section 5.

2 Experimental framework

2.1 The ARPEGE model and its 4D-Var system

The ARPEGE model is the global NWP model at Météo-France, used operationally since the early 1990s (Courtier et al., 1991). This system is fully integrated within the ARPEGE-IFS software that was conceived, developed and maintained in collaboration with ECMWF.

This model is a spectral global model with a stretched grid having a horizontal resolution around 7.5 km over France and 37 km over the antipodes. It has 105 vertical levels according a following-terrain pressure hybrid coordinate, with the first level located at 10 m above the surface and an upper level at around 70 km. Clouds and precipitation are described by using three different scheme in the ARPEGE model. The stratiform clouds in terms of cloud profile and precipitation are explicitly modeled from the microphysical condensation scheme by Lopez (2002). The large-scale effects of deep convection are parametrized from a mass-flux scheme derived from Bougeault (1985) and the shallow convection ones with the Bechtold et al. (2001) one. In these last two cases, the cloud fraction and the liquid water, ice and precipitation profiles are diagnosed.

ARPEGE has four analyses per day at 00, 06, 12 and 18 UTC. Since June 20th, 2000 the operational data assimilation system of the ARPEGE model is a 4D-Var. This implementation, as detailed in Janiskova et al. (1999) and Rabier et al. (2000), is used to provide an analysis which corresponds to the best atmospheric state knowing observations, an a-priori state, dynamical and physical constraints.

The background errors are computed at each analysis time based on the 25-member assimilation ensemble (see Berre et al. (2015) for further details). The control variables considered are temperature, specific humidity, vorticity, divergence and the logarithm of the surface pressure.

At each analysis around 7 million observations are assimilated. They include conventional observations (from radiosounding, aircraft, ground stations, ships, buoys, etc.) and satellite data. These latter include radiances in the infrared and microwave spectra such as AIRS (Atmospheric InfraRed Sounder), IASI, CrIS (Cross-track Infrared Sounder), SEVIRI (Spinning Enhanced Visible and InfraRed Imager), AMSU-A (Advanced Microwave Sounding Unit-A), MHS (Microwave Humidity Sounder), ATMS (Advanced Technology Microwave Sounder) and atmospheric motion vectors. Scatterometers provide information on ocean surface wind. Zenithal total delay signals and from radio-occultation measurements from the Global Navigation Satellite System (GNSS) are also assimilated.

With the advent of hyperspectral sounders such as AIRS and IASI, a variational bias correction (VarBC) method (Auligné et al., 2007) has been operationally implemented at Météo France and notably in the ARPEGE model. The VarBC scheme aims to minimize systematic innovations in radiances while preserving the differences between the background and other observations in the analysis system.

The observation operator allows to simulate observations from the model variables for comparison with the actual measurements. For satellites radiances, it includes a radiative transfer model. The accuracy of the forward model calculation could be limited by the accuracy of the NWP model, for some variables this is not sufficient to correctly model the observations and these observations have to be discarded.

The assimilation of clear radiances at Météo France is based on the McNally and Watts cloud detection scheme (McNally and Watts (2003)) The McNally and Watts (2003) scheme intends to detect clear channels and to assimilate channels unaffected by clouds even in a cloud-affected pixel. The channels are first re-ordered according to a ranking with respect to the altitude that reflects their relative sensitivity to the presence of cloud. After having applied a low-pass filter a search for the channel at which a monotonically growing departure can first be identified. Having found this channel all channels ranked more sensitive are flagged as cloudy and those ranked less sensitive are flagged clear.

In addition, a cloud characterization is made using cloud parameters (a cloud top pressure (CTOP) and an effective cloud fraction (Ne)) deduced from a CO_2 -slicing algorithm (Pangaud et al. (2009)). These two parameters are used to model the radiative impact of cloud as a single layer cloud, with an emissivity set to 1 using a clear sky radiative transfer model.

2.1.1 Main features of the IASI instrument

IASI is a key element of the Metop series payload of European polar orbiting meteorological satellites (Cayla, 2001). It was designed by CNES (Centre National d'Etudes Spatiales) in cooperation with EUMETSAT. The first flight model was launched in 2006 on board the first European meteorological satellite Metop-A in polar orbit. The second instrument, mounted on the Metop-B satellite, was launched in September 2012. The third instrument was mounted on the Metop-C satellite, which was launched in November 2018. The horizontal resolution of the instrument is 12 km at the nadir. IASI is dedicated to operational meteorological soundings with a high level of accuracy (specifications on temperature accuracy: 1 K for 1 km and 10% for humidity (Chalon et al., 2001)). Its measurements are also useful for atmospheric chemistry to estimate and monitor different trace gases such as ozone, methane or carbon monoxide on a global scale (Hilton et al., 2012).

IASI is a passive IR remote-sensing instrument using an accurately calibrated Fourier transform spectrometer to cover the spectral range from $3.62 \mu\text{m}$ (2760 cm^{-1}) up to $15.5 \mu\text{m}$ (645 cm^{-1}) with 8461 channels. Its spectral resolution is 0.5 cm^{-1} with a spectral sampling of 0.25 cm^{-1} . The IASI spectrum can be divided into three major bands:

- from $645\text{--}1210 \text{ cm}^{-1}$: CO_2 , window and ozone channels mainly sensitive to temperature, called long-wave (LW) channels;
- from $1210\text{--}2040 \text{ cm}^{-1}$: channels mainly sensitive to humidity, called water-vapour (WV) channels;
- from $2040\text{--}2700 \text{ cm}^{-1}$: named short-wave (SW) channels.

Only a subset of 314 channels (300 channels selected by Collard (2007) and 14 additional channels for monitoring purposes) used in operations at Météo-France, is considered in this study

2.1.2 Towards the assimilation of cloudy infrared IASI radiances

Assimilation of cloudy radiances is a crucial challenge for NWP centres as the cloudy observations discard represent an under-exploitation of hyperspectral sounders especially in sensitive meteorological areas (McNally, 2002; Fourrié and Rabier, 2004).

As mentioned in the introduction, studies about all-sky infrared assimilation have started. The radiative transfer model RTTOV-CLD for cloudy sky, included in RTTOV version 11 (Saunders et al., 2013), offers a realistic modeling of the cloud scattering. This model also allows a better description of the cloud emissivity as well as cloud scattering, using the microphysical cloud profiles (water content, cloud ice content and cloud cover).

To simulate the radiances observed in cloudy conditions using RTTOV-CLD, we use two main types of clouds: firstly liquid water cloud which corresponds to two RTTOV-CLD cloud microphysical options depending on the land sea mask of the model (Stratus continental over land and Stratus Maritime over the sea) secondly the ice water cloud of the Cirrus type, using Baran parameterisation (Baran et al., 2014; Vidot et al., 2015) to define the optical properties.

To illustrate the benefit brought by RTTOV-CLD, Figure (1) shows IASI brightness temperature observations of a cloud-sensitive surface channel ($1271, 962.5 \text{ cm}^{-1}$) and differences between the observations and the simulations computed with RTTOV considering clear-sky and with RTTOV-CLD. Brightness temperatures less than 250 K are usually associated with higher elevation cloud structures. By using RTTOV in clear sky (figure 1.b) to simulate IASI observations, the observation departures are mainly below zero and may reach up to -60 K. This can be explained by the fact that the main cloud structures associated with low values of brightness temperature for the surface channel are missing in the simulation. On the contrary, differences obtained with the RTTOV-CLD simulations are in overall in better agreement with lower positive and negative values (figure 1.c). No major differences are found for example for the cloud structures located over the North Atlantic (30N-70N, 40W-0W) and above (30S-70S, 60W-0W) the Southern Atlantic Ocean. Large difference values are mainly obtained in the Tropics region. This may be explained by the fact that clouds are better simulated in the ARPEGE model for mid-latitudes than in the Tropics.

3 Selection method of homogeneous observations

The assimilation of cloudy radiances in NWP models remains a innovative challenge. In the context of the preparation of all-sky assimilation, we plan to assimilate clear and cloudy observations that are completely covered in a homogeneous way, discarding the cases of fractional cloud observations.

These scenes are supposed to be better characterized and simulated than fractional cloudy scenes in NWP models. Indeed, by selecting homogeneous cloudy scenes in both model and observation spaces, we improve the agreement between observations and background simulations. This selection of cases seen as homogeneous by both IASI and the model avoids misplacement errors. In this section, limited to cases over sea to avoid problems related to the land surface properties, we describe several methods for analysing the homogeneity of the scene in the observation and model space. However these methods were applied over all surface in the assimilation experiments of section 4.

3.1 AVHRR clusters

In order to select homogeneous pixels, the AVHRR imager information collocated within IASI on the MetOp platform is used. The spatial resolution of AVHRR observations is around 1 km at nadir and measures the radiation emitted in six broad-band

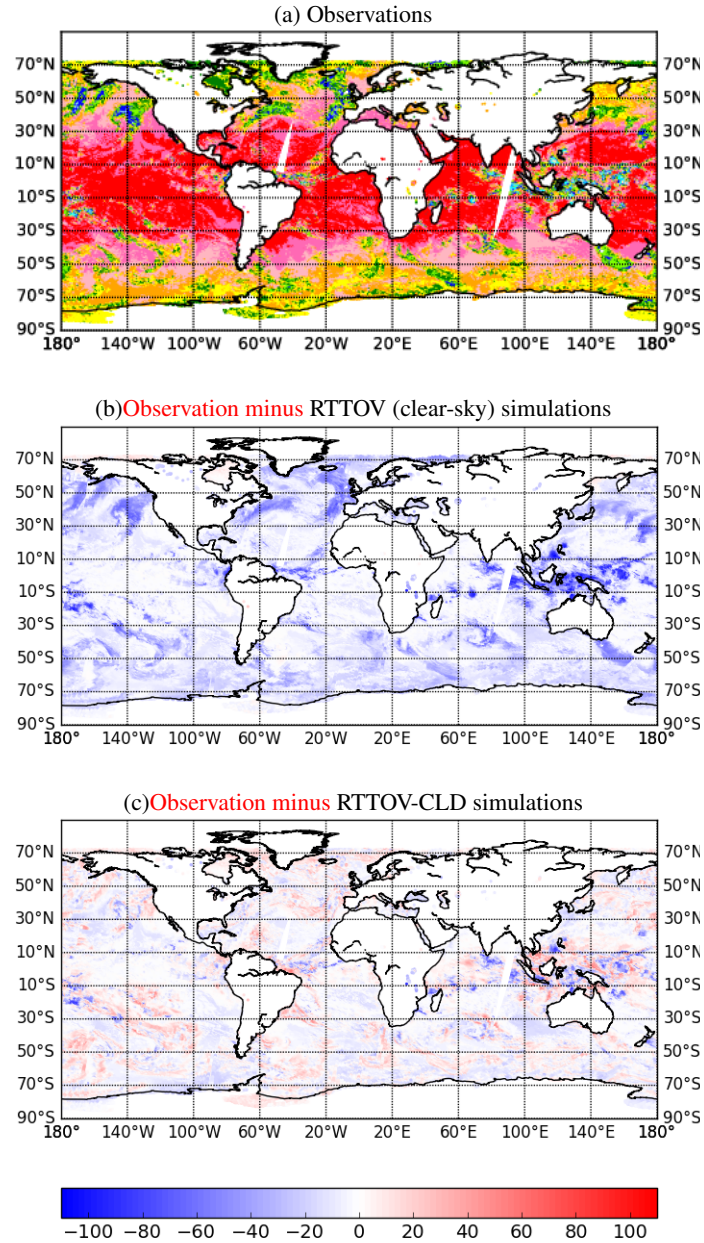


Figure 1. IASI brightness temperature (K) observations (a) from Metop A and B satellites and differences between observations and simulations using RTTOV (b) and RTTOV-CLD (c) for surface channel ($1271, 962.5 \text{ cm}^{-1}$) for 30 January 2017 daytime over sea from ARPEGE 6-hour forecast fields.

channels: one visible channel, two near-infrared channels, a shortwave infrared channel and two long-wave infrared channels (10.5 μm and 11.5 μm). Two **components of the IASI Level 1c products** provided by EUMETSAT were used: the AVHRR clusters (Cayla, 2001) and the percentage of cloudy AVHRR pixels in the IASI FOV (product GEUMAvhrr1BCldFrac:Pequignot and Lonjou (2009)). The AVHRR pixels are clustered into homogeneous classes in the radiance space, (visible and infrared channels) using the K-mean classification algorithm. **For each class and each AVHRR channel, the cluster product provides the coverage of the class within the IASI pixel, the mean and the standard deviation of AVHRR brightness temperatures within the class.**

3.2 Selection criteria for homogeneous observations

This study intends to focus on those IASI pixels that contain only one cluster, which corresponds to a homogeneous scene. However only 2% of daytime IASI observations over sea contain only one class. The aggregation is built with all available AVHRR channels (visible, NIR, IR), several classes can be produced with the K-mean classification even with relatively small standard deviations for the IR channels. A IASI FOV with several classes, each one having a small standard deviation and a mean radiance close to the ones of the other classes, can thus be more homogeneous than a FOV with a single class **but with very large value of standard deviations.**

For this reason, the number of AVHRR clusters within each IASI pixel has not been used as a homogeneity criterion, but these characteristics have been used to calculate the overall AVHRR cluster statistics, aggregating the information provided by all clusters in the IASI FOV.

We tested four methods for selecting homogeneous scenes by calculating homogeneity criteria in the observation space as well as in the model space, using the AVHRR channels. The first two ones are described in the literature and we propose two other ones which are detailed below.

3.2.1 Homogeneity criteria derived from Martinet and al., (2013)

These homogeneity criteria are based on a single AVHRR infrared channel 11.5 μm , which is used to compute three homogeneity tests, the first two tests are calculated in the observation space and the third one in the model space:

Intercluster homogeneity

The intercluster homogeneity is based on σ_{inter} defined as:

$$\sigma_{inter} = \sqrt{\frac{1}{\sum C_j} \sum_{j=1}^N C_j (L_j - L_{mean})^2} \quad (1)$$

Where L_j is the mean radiances of cluster j at channel 11.5 μm , L_{mean} represents the radiance weighted average. The weighting is determined by C_j is the cluster fraction of each class inside the IASI pixel. N is the number of classes in the IASI pixel.

A small calculated standard deviation σ_{inter} means that all classes observe a similar cloudy scene in the infrared channel. If this standard deviation is too high, each class observes a different scene (clear or cloudy) and the IASI pixel is very heterogeneous.

Intracluster homogeneity

In order to finalize the homogeneity criterion in the observation space, it is also necessary to check if each class itself is sufficiently homogeneous, using the following formula:

$$\sigma_{intra} = \sqrt{\frac{1}{\sum C_j} \sum_{j=1}^N C_j \sigma_j^2} \quad (2)$$

Where the σ_j are the standard deviations of each cluster j calculated for the infrared channel $11.5 \mu\text{m}$. The IASI observation is considered homogeneous if it verifies the following criteria:

- Ratio between intracluster homogeneity σ_{intra} and mean radiance $L_{mean} < 4\%$.
- Ratio between intercluster homogeneity σ_{inter} and mean radiance $L_{mean} < 8\%$.

Background departure check

Finally, in order to obtain a similar criterion in the model space, each AROME grid point within IASI FOV was used to simulate the equivalent AVHRR channel $11.5 \mu\text{m}$ with RTTOV-CLD. Homogeneous IASI observations are preserved if the ratio of the standard deviation of the AVHRR simulations and the simulated mean radiance of the AVHRR is less than 8%.

Adaptation of the method

In the original Martinet et al. (2013) study, the third check verifying that both the observation and the model observe the same cloudy scene was done with the difference between the mean AVHRR brightness temperatures from the observed and simulated clusters less than 7 K. Here, the ARPEGE model has a coarser resolution and it is not possible to simulate the AVHRR clusters. This check was adapted with the difference between the AVHRR observation and the AVHRR simulation from the guess, which come from an horizontal interpolation of the 12 profiles surrounding the observation position coming from a 6-hour forecast.

~~In our study, which focuses on the ARPEGE global model, we chose to use the simulated brightness temperature from the guess profiles coming from a 6-hour forecast and interpolated using 12 points surrounding the observation position.~~ The homogeneous cases are retained as long as the difference between AVHRR observations and simulations is less than 7 K. This method will be noted M2013.

3.2.2 Homogeneity criteria derived from Eresmaa (2014)

The study of Eresmaa (2014) proposed an imager assisted cloud detection for the global ECMWF NWP system and was based on the hypothesis that each AVHRR cluster are made of fully clear or fully cloudy pixels.

Therefore, his selection criteria is only intended to diagnose and retain observations when they were completely clear, using the last two infrared channels of AVHRR (10.5 μm and 11.5 μm). This detection is based on three checks called the homogeneity check, the intercluster consistency check and the background departure check. If a IASI pixel do not satisfy one of these checks, it is not free of cloud and is rejected.

- 5 The standard deviation of the brightness temperatures of the two infrared channels from all pixels present in the FOV is used for the first check. If both standard deviations are over the pre-determined threshold values (0.75 and 0.80 K, respectively), it means that a cloud is potentially observed and the IASI observation is rejected. The intercluster consistency check relies on the comparison between the properties of the different clusters within the IASI FOV. The distance of each cluster to the background in both infrared AVHRR channels as well as the distance between each pair of clusters. A cloud is detected if there
10 is a pair of clusters covering more than 3% of the IASI FOV and for which the intercluster distance exceeds the minimum value of the distances between these clusters and the background.

The distance between 2 clusters j and k is computed as the squared-summed intercluster departure:

$$D^{jk} = \sum_{i=4}^5 (R_i^j - R_i^k)^2 \quad (3)$$

- where R_i^j is the mean brightness temperature of cluster j for channel i . In addition the distance of the cluster j to the
15 background is computed with:

$$D^j = \sum_{i=4}^5 (R_i^j - R_i^{BG})^2 \quad (4)$$

Where R_i^{BG} is the **clear-sky** background brightness temperature for AVHRR channel i . The observation is rejected due to the diagnostics of the presence of a cloud if the following inequality is true and the coverage of clusters j and k is over 3%:

$$D^{jk} > \min(D^j, D^k) \quad (5)$$

- 20 The last check on the background departure is computed as a fractional-weighted mean of the squared-summed background departures:

$$D_{mean} = \sum_{j=1}^N D^j C_j \quad (6)$$

where N is the number of clusters in the IASI FOV and C_j is the fractional coverage of cluster j . The presence of cloud is diagnosed if D_{mean} exceeds the threshold value of 1K².

25 **Adaptation of the method**

Since this method assumes that each cluster is made of pixels that are either all clear or cloudy, its homogeneity tests have been adapted to the selection of clear and cloudy pixels, with criteria that would fit our purpose, with the first test in the observation

space and the second one **background departure check**. All AVHRR simulations from background are made with RTTOV-CLD and the threshold of the background departure check was modified.

-Inter-cluster homogeneity.

This test uses the standard deviation of the infrared brightness temperature, calculated on all clusters occupying the IASI field of view. The standard deviation is calculated in the same way as Eresmaa (2014) and the IASI pixel is considered homogeneous if the two standard deviations (one for each channel) are below their predetermined threshold values of 0.75 K and 0.8 K respectively.

-Background departure check. In this test, We used the D_{mean} proposed by Eresmaa (2014) **but the IASI pixel is considered as homogeneous to perform here a kind of cloudiness consistency check between the observation and the model simulation if D_{mean} is less than 49 K². This particular value of threshold allows to keep more than 50% of the observations compared to the initial threshold of 1K² by Eresmaa (2014) which retains only 18% of the observations. In addition this threshold compares well with the one applied by M2013, but it is applied over the 2 IR AVHRR channels.** This method is referenced as E2014 in the following.

The threshold values of the homogeneity criteria derived from Martinet et al. (2013) and Eresmaa (2014), are based on the analysis of statistics, applied to all IASI FOVs of the different situations (day/night at sea). Threshold values are specified in such a way that the standard deviation between the observations and simulations is not too large while keeping a fair amount of the observations.

3.2.3 Selecting homogeneous scenes in observation space

This third method, (called Obs_HOM thereafter) proposes a homogeneity check in the brightness temperature space calculated only on the observation space, using both infrared AVHRR channels (10.5 μm and 11.5 μm). **It is the same test as in M2013 but in the brightness temperature space.** This inter-cluster homogeneity criterion is based on the relative standard deviation of AVHRR clusters inside the IASI pixel. This test is satisfied when all classes observe a very similar scene in the AVHRR infrared channels. To evaluate the interclass homogeneity, the standard deviation of the mean brightness temperature of clusters which occupy the IASI FOV has been calculated using the following formula:

$$\sigma_{inter} = \sqrt{\frac{1}{\sum C_j} \sum_{j=1}^N C_j (R_{i,j} - R_{i,mean})^2} \quad (7)$$

Where: $R_{i,j}$ is the mean brightness temperature of cluster j on channel i , $R_{i,mean}$ represents the weighted average on channel i , N is the number of classes in the IASI pixel, and C_j is the **cluster** fraction.

Figure 2 provides a calibration to determine the thresholds to be used to define homogeneous scenes. These thresholds should lead to a sufficient size of the selected dataset and avoid selecting the fractional cloud as much as possible. Therefore we decided to select an observation if the **ratio** between intercluster homogeneity and mean radiance for both AVHRR IR channels (10.5 μm and 11.5 μm) are less than 0.8%. **If this test is only applied over channel (10.5 μm), 68,2% of the observations are**

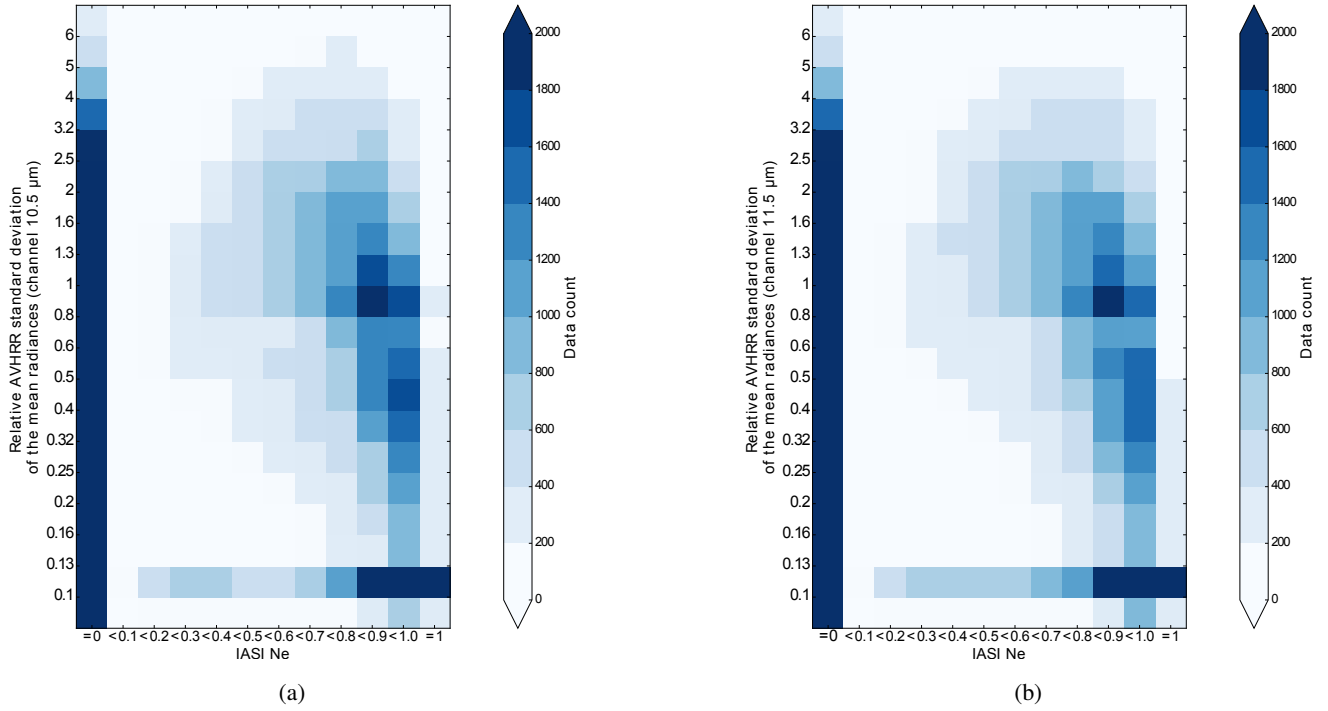


Figure 2. Density plot (number of observations) of the values of effective cloud fraction retrieved from IASI by a CO_2 -slicing algorithm (on the abscissa) with respect to the relative cluster standard deviation of the mean radiances (% , on the y-axis) for intercluster homogeneity for (a) the AVHRR IR channel ($10.5 \mu m$) and (b) the AVHRR IR channel ($11.5 \mu m$)

selected, if applied over channel ($11.5 \mu m$), 69,6% of the data are kept and if it is applied over both channels, 67,3% pass the test.

3.2.4 Compromise for the homogeneous scene selection

Based on the previous methods, we propose a fourth one which represents a compromise between them. Two AVHRR infrared channels ($10.5 \mu m$ and $11.5 \mu m$) are used, and we define two homogeneity criteria in the observed and simulated brightness temperature spaces.

The first criterion for homogeneity is the interclass homogeneity check which was used in the third method, calculated in the observation space (presented in section 3.2.3). Similarly, we used the background departure check in the observation space D_{mean} (presented in section 3.2.2).

10 Only observations that fulfilled the two following criteria were selected:

- Ratio between intercluster homogeneity and mean brightness temperature for two AVHRR IR channels ($10.5 \mu m$ and $11.5 \mu m$) $< 0.8\%$.
- Sum of the average distances between each cluster and the background $< 49 K^2$.

Methods	Literature	AVHRR channels used	Homogeneity criteria in observation space	Test on background simulation
M2013	Martinet et al. (2013)	11.5 μ m	intra and intercluster	distance with observation
E2014	Eresmaa (2014)	10.8 and 11.5 μ m	intercluster	average distance with each cluster
Obs_HOM		10.8 and 11.5 μ m	intercluster	No
COMPR		10.8 and 11.5 μ m	intercluster	average distance with each cluster

Table 1. Summary of the criteria for homogeneous IASI observation selection used in this study.

This method is named COMPR in the following. All the four methods are summarized in Table 1.

4 Inter-comparison of selection criteria

We applied our selection criteria on January 30, 2017 and result from an observation sample composed of 188090 IASI FOV during the daytime over the sea are presented. Same conclusions were found for the other cases (night-time and/or over land).

	Number of Observations	Homogeneous Clear Observations	Homogeneous Cloudy Observations	Heterogeneous Observations
All observations	67599	12%	51%	37%
M2013	57%	11%	27%	20%
E2014	26%	10%	9%	7%
Obs_HOM	69%	12%	35%	22%
COMPR	40%	10%	17%	13%

Table 2. Evaluation of heterogeneity inside the IASI pixel with respect to the SEVIRI cloud type.

- 5 An evaluation of the homogeneity inside the IASI pixels according the various selection criteria was performed with independent SEVIRI data over the sea to check that the retained observations are selected for clear or overcast scenes. Here the considered data were the cloud type from the geostationary SEVIRI sounder which is a product from the Satellite Application Facility on support to NoWCasting and very-short range forecasting of EUMETSAT. For each IASI pixel, the cloud types of the 4 SEVIRI pixels closest to the IASI centre were compared. If the 4 cloud types were sea, the IASI pixel is considered as clear,
- 10 if the same cloud type is found for SEVIRI pixels, the IASI observation is set to homogeneous cloudy observation, otherwise it is considered as a heterogeneous observation. Table 2 presents the results of this comparison. 67599 IASI observations were colocated with SEVIRI data during the day of 30 January 2017. In the global dataset, we found that 12% are clear homogeneous observations, 51% are cloudy homogeneous observations and 37% of IASI observations are made of different SEVIRI cloud type. This corresponds well to the results obtained with the results found with the percentage of cloudy AVHRR pixels
- 15 in the IASI field of view (Table 3). When applying the M2013 criteria, 11% of clear observations are retained. this number

	Number of observations	Cloudy observations CldCover=100	Clear observations CldCover=0	Partially cloudy observations 0<CldCover<100
All observations	188090	50%	12%	38%
M2013	54%	19%	10%	25%
E2014	22%	6%	10%	6%
Obs_HOM	67%	32%	13%	22%
COMPR	36%	11%	10%	15%

Table 3. Overview table of statistics obtained with the different homogeneity criteria : number of observations retained, percentage of cloudy observations (cloudcover of 100), percentage of fully clear observations (cloudcover of 0). Percentage are given with respect to the whole number of observations.

does not vary a lot when applying the other criteria. The number of homogeneous cloudy observation is more sensitive to the criteria because it varies between 9% for E2014 and 35% for the Obs_HOM one. The COMPR criteria provide a compromise between M2013 and E2014 in terms of selected observations.

The percentage of cloudy AVHRR pixels in the IASI field was also used to assess the choice of homogeneity criteria (Table 3).

Our global dataset is made of 50% of the observations entirely covered by clouds and 12% of clear observations according to the AVHRR cloud cover. These results obtained over the globe set agree well with the ones obtained with SEVIRI data over the Atlantic Ocean. The percentage of selected observations for each selection method is larger (+2/4 %) with the SEVIRI data evaluation than with the AVHRR cloud cover.

The bias and standard deviation of observations minus simulations (O-G), are shown in Figure 3.(a) for the 314 IASI channels. As expected, the best statistics are obtained for channels less affected by clouds (e. g. CO₂ and water vapour high peaking channels).

For the whole dataset, window channels present a bias of around -0.6K. The standard deviations are larger (around 12 K) for window channels sensitive to the surface. With the M2013 selection method (figure 3. (b)), the standard deviation of window channels is reduced to around 4 K as well as the bias close to zero. The standard deviation of the other channels (680-780cm⁻¹) is also well decreased. The E2014 selection method (figure 3. (c)) improves the bias and the standard deviation (2.0 K for window channels) for all the channels. As expected, the impact is larger for surface sensitive (and thus cloud sensitive) channels than for the tropospheric channels (680-780cm⁻¹). On the contrary, with the Obs_HOM method (figure 3. (d)), small statistics improvement is obtained for the standard deviation and the bias. The statistics obtained with the COMPR method (Fig. 3.(e)) are reduced compared to the whole dataset and slightly less good than with the initial E2014 method (for window channels the standard deviation is around 2.2 K instead of 2 K for E2014).

To complete the comparison, the Probability Density Function (PDF) of the O-G differences was studied (Fig.4). Three channels were assessed: the window channel 1271 (962.5 cm⁻¹, whose weighting function peaks at around 1000 hPa), the

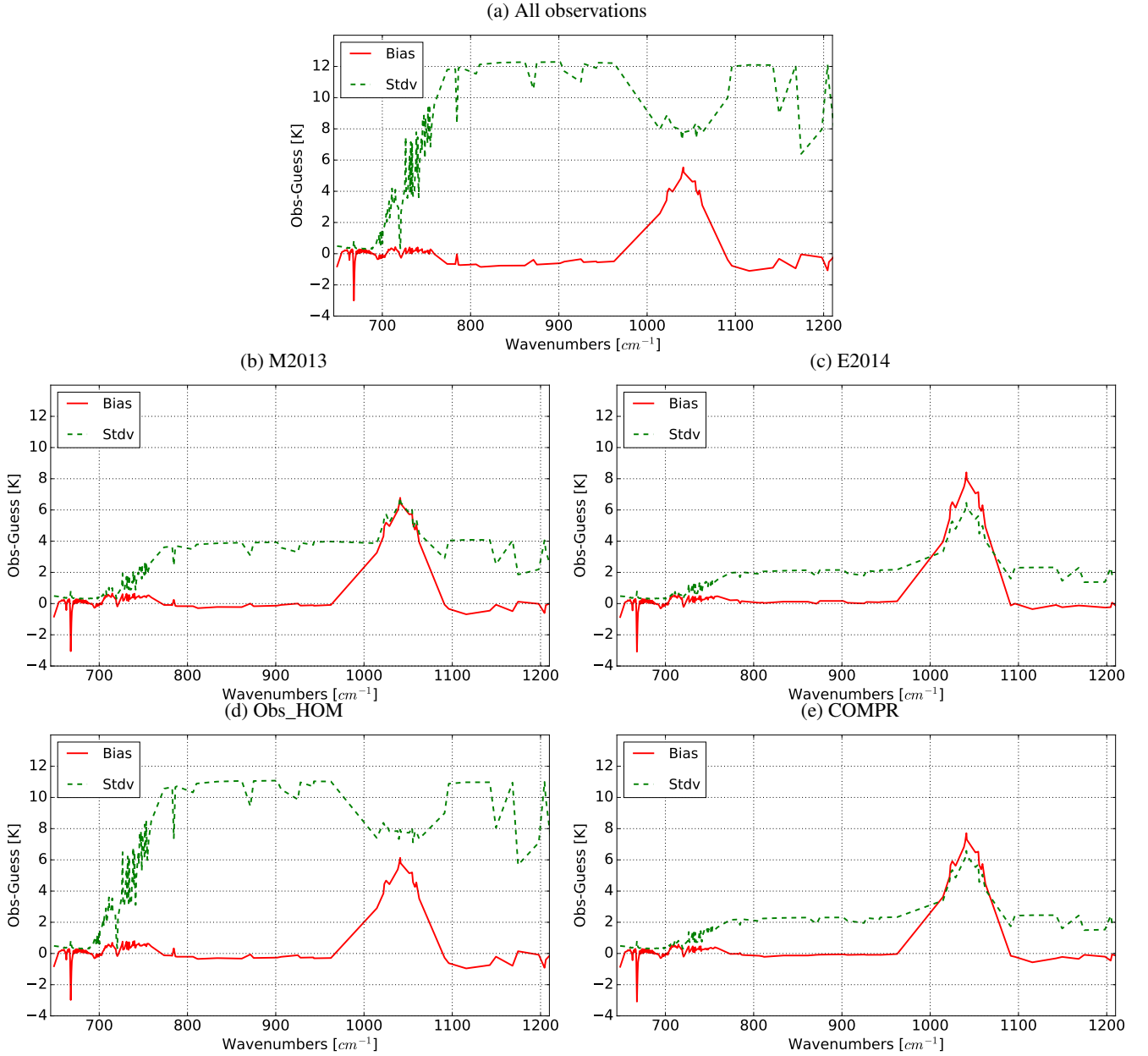


Figure 3. Bias (red solid line) and standard deviation (green dashed line) in Kelvin (K) of the differences between IASI observations and background simulations using RTTOV-CLD and a 6-hour forecast: (a) for the whole dataset, (b) after applying the homogeneity criteria derived from (Martinet et al., 2013), (c) after applying the homogeneity criteria derived from (Eresmaa, 2014), (d) after applying the selecting homogeneous scenes based on observation space, (e) after applying the compromise to select the homogeneous scenes. Observations are for January 30, 2017.

mid-tropospheric water vapour channel 2701 (1320 cm^{-1} , weighting function maximum at around 400 hPa) and the low-tropospheric water vapour channel 5403 (1955 cm^{-1} , weighting function peaking at around 900 hPa),

The distribution asymmetry is relatively small for mid and low tropospheric water vapour channels. The distribution asymmetry is reduced for mid and low tropospheric water vapour channels with M2013 and E2014 selection. The impact of clouds is evident on the window channel, with differences ranging from -90 to 64 K. After the homogeneity criteria is been applied, narrower Gaussian distributions are observed for all channels with a significant improvement for the window channel. Using the M2013 criteria, differences in O-G for the window channel are significantly reduced, from -18 K to 20 K, and from -7 to 9 K using the E2014 criteria (Figure 4.g, h, i).

With Obs_HOM criteria (Figure 4.(j, k, l)), the O-G distribution is not much improved for all channels ~~except for the low-tropospheric water vapour channel where the range is reduced from 60 K to 40 K.~~ When the homogeneity criterion in the model space is added using the COMPR selection, the O-G distributions become symmetrical, get closer to the gaussian distribution (and Gaussian) and centered around zero for the three previously selected channels(Figure 4.(m, n, o)), which indicates the data are correctly diagnosed as homogeneous.

	Number of observations	Bias temperature channels ($650\text{-}770\text{ cm}^{-1}$)	Stdev temperature channels ($650\text{-}770\text{ cm}^{-1}$)	Bias (Window channels) ($770\text{-}980\text{ cm}^{-1}$)	Stdev (Window channels) ($770\text{-}980\text{ cm}^{-1}$)
All observations	188090	0.06 K	2.53 K	-0.60 K	11.7 K
M2013	54%	0.14 K	0.82 K	-0.16 K	3.7 K
E2014	22%	0.13 K	0.59 K	0.11 K	2.0 K
Obs_HOM	67%	0.19 K	2.22 K	-0.20 K	10.5 K
COMPR	36%	0.12 K	0.64 K	-0.09 K	2.1 K

Table 4. Overview table of statistics obtained with the different homogeneity criteria : number of observations retained, bias and standard deviation computed for the channels included in the range between $650\text{-}770\text{ cm}^{-1}$ and $770\text{-}980\text{ cm}^{-1}$.

Table 4 summarizes statistics in terms of bias and standard deviation of background departure for the different datasets. The bias and standard deviation obtained by the M2013 method have some reasonable statistics before the assimilation (-0.6 K for the bias and 3.7 K for the standard deviation, for the window channels). The E2014 selection method seems relevant for selecting homogeneous scenes in terms of bias and standard deviation (0.11 K and 2.0 K respectively, for the window channels). However, the number of selected observations presents a disadvantage for this selection method, since E2014 method keeps only 22% of the observations of which 10% are totally clear, 6% are totally covered by clouds and 6% are heterogeneous. These observations are distributed throughout the globe, but we keep more observations on high latitudes.

The Obs_HOM method allows to keep 67% of observations, which 12% are totally clear and 32% are totally covered by clouds, but this method does not give acceptable statistics (bias of -0.2 K and standard deviation of 10.5K). When the test on observations minus simulations of the infrared channels AVHRR are added by the COMPR method, results are improved. For window channels the bias is reduced to -0.09 K and the standard deviation to 2.1 K compared to -0.6 K and 11.7 K for all obser-

vations, which presents a good score compared to the M2013 and Obs_HOM methods. In addition 36% of the observations are retained, compared to the whole dataset, with 10% of clear observations and 11% of cloudy observations of the total amount, which represents twice the amount of cloudy observations selected by E2014, which removes many more observations, and shows that the proposed methodology is effective.

- 5 The cloud cover distribution corresponding to the amount of observations that is kept (36%) is made of 28% of clear observations and 29% of the observations totally covered by clouds. In addition, 14% of the observations have a cloud cover of less than 10% and 4% of the observations have a cloud cover exceeding 90%. The observations kept are distributed in different parts of the globe (Figure 5.a) although we have been able to retain different cloud types, including high clouds even in the tropics for few cases only (Figure 5.b). This may be explained by the weakness of the model clouds in these areas.
- 10 The main objective of the study is to select homogeneous IASI observations in clear and cloudy sky which are well simulated with RTTOV-CLD and could be used in data assimilation. Comparison of different methods of selecting homogeneous scenes showed that the M2013 method improves the first guess departure statistics (bias of -0.16 K and standard deviation of 3.17 K) but it keeps more heterogeneous observations (25%) according AVHRR cloud cover than the E2014 method, which significantly improves the statistics (bias of 0.11 K and standard deviation of 2 K) and favours more clear observations but
- 15 keeps only 22% of the observations. The Obs_HOM method, which focuses only on homogeneity in the observation space, does not strongly improve the statistics but it filters 33% of heterogeneous observations. However the addition of the criterion on the simulated observations in the COMPR method improves the scores on IASI simulations (bias of 0.09 K and standard deviation of 2 K), while retaining 36% of the observations and among them a similar part of homogeneous clear and homogeneous cloudy observations. This data selection representing a compromise between M2013 and E2014 is chosen for a data
- 20 assimilation experiment.

5 Impact on NWP analyses and forecasts

After the selection criteria were implemented in the assimilation system of Météo France, their impact was tested through a 4D-Var assimilation experiments in the ARPEGE global model. The impact of the homogeneity criteria for data selection on all observation simulations, on analyses and forecasts is evaluated.

25 5.1 Experimental design

To evaluate the impact of our homogeneity criteria on the assimilation process over sea and land, during daytime and night, four experiments were performed over one month from 06/12/2017 to 17/01/2018. 314 IASI channels were used in the simulation, and 129 channels (Tables ?? and ?? in Appendix) were used for assimilation as operationally.

- 30 The first experiment is the reference (REF), where IASI observations are assimilated with all other observation type as in the operational system at Météo-France.

~~The Obs_HOM criteria for selecting homogeneous IASI observations (presented in the 3.2.3 section) are applied in the second experiment called (EXP.A) on top of the Mc Nally and Watts cloud detection. Finally in the third In a second experiment~~

called (EXP), we applied our COMPR approach (presented in the 3.2.4) on top of the Mc Nally and Watts cloud detection. These ~~sets of~~ experiments aim to evaluate the impact of the COMPR method of selecting homogeneous IASI observations on simulation and assimilation processes. ~~The following Table ?? details the main features of each experiment.~~

In these experiments, no cloudy observations detected with the CO2-slicing method ~~and used with a single grey-cloud layer~~ scheme was assimilated ~~unlike~~ Guidard et al. (2011). We focus on clear sky assimilation. RTTOV-CLD was only used to compute the homogeneity criteria based on cloudy AHVRR simulations and RTTOV was used for the clear sky assimilation.

5.2 Impact on observation

~~The number of assimilated channels for each observation is different depending the areas, e.g. in the tropics, there are less clear channels (between 30 and 40 channels) in the REF (Figure ??a), EXP.A (Figure ??b) and EXP.B (Figure ??c), which is explained by the presence of high clouds in this region. In EXP.A (Figure ??b) with the Obs_HOM criteria, some observations were filtered in the tropical area, and even more in EXP.B (Figure ??c) where the criterion used is even more stringent, more observations are filtered in areas corresponding to high clouds.~~

Figure 6.a gives the number of assimilated observations into the ARPEGE model as function of the IASI wavenumber for REF ~~, EXP.A and EXP.B~~ experiment. The order of magnitude is below 8.10^6 whichever the wavenumber considered and for ~~both~~ experiments showing that among the 129 IASI channels selected for the assimilation, the occurrence proportion are well balanced. However, the assimilated number changes between experiments depending upon the spectral band considered. Four spectral band can be mentioned:

- $[657, 687.25] \text{ cm}^{-1}$ wavenumbers range corresponding to the stratospheric temperature channels keeps the same number of observations in ~~both~~ experiments because these channels are not affected by the presence of clouds.
- $[726.5, 1421] \text{ cm}^{-1}$ wavenumber range, corresponding to tropospheric and surface temperature channels and also to mid and high tropospheric water vapour channels, the number of observations is decreased by 15% for EXP (Figure 6.b).
- $[1800, 2015.5] \text{ cm}^{-1}$, finally the number of low tropospheric water vapour sensitive channels is slightly decreased ~~for EXP.A by 0.5% and~~ between 8% and 14% for EXP (figure 6.b).

5.3 Impact on background and analyses

The analysis departure data discussed below are made by comparing the analysis between the REF and the EXP experiments ~~(EXP.A and EXP.B)~~ to evaluate the impact of the criteria for selecting homogeneous IASI observations ~~(refer to Table ??)~~.

Figures 7.a and 7.b present the impact of COMPR criteria in the temperature and humidity analyses ~~of the first assimilation cycle~~. This implementation removes some IASI observations from the assimilation and this reduction has an impact on the analysis. In Figure 7.a, a negative temperature difference is located in the Atlantic Ocean near to the South-West African coast. ~~A similar behaviour on temperature is noticed from the impact of COMPR (EXP.B) as shown in Figure 7.e..~~ Weaker and patchy impact is reported on specific humidity that is mainly located in the tropics. EXP seems to remove some temperature and humidity analysis increments from the REF experiment just at some isolated locations.

In order to assess the impact of the new selection of IASI observations on the analyses and forecast, first guess departures (FG departures) corresponding to the difference between the observations and the simulation from the 6-h forecast and the analysis departures (AN departures) are computed. As biases and standard deviations of FG and AN departures were very weak for IASI, CRIS and AMSU-A instruments, and humidity measurements performed by radiosondes, relative differences have been performed between experiments to highlight detailed comparisons (Figure 8). **Negative differences are related to a reduction of the standard deviation with respect to REF one and thus to an improvement.**

For IASI (Figure 8.a), regarding exclusively the significant differences with a 95% value, FG departure standard deviation was mainly reduced or unchanged in EXP compared to REF depending upon the wavenumber; increases are observed around 2000 cm^{-1} and 1400 cm^{-1} while significant reduction of 0.5-1 % is seen between 1000 and 1320 cm^{-1} . The standard deviation reduction on AN departures can be noted at around 2000 cm^{-1} and between 950 and 1320 cm^{-1} .

Concerning the CrIS observations (Fig. 8.b), the differences results are mainly not significant excepted at around the 850 water vapour channels where the standard deviation increases for both FG and AN departures, and a significant improvement for the AN departures at the channels 160.

Results obtained for AMSU-A (Fig. 8.c) are mainly satisfactory with FG departure standard deviation differences reduced by around 0.05 % for channels 6, 9, 10, 11 and 12 ~~with the EXP.A and EXP.B~~. However, a significant degradation of 0.2 % is observed for channel 8. The AN departure results follow more or less the same behaviour. Finally, no significant standard deviation difference is observed concerning the TEMP-q observations (Figure 8.d).

Results shown in this part report **small but** non-negligible impact of the homogeneous criteria implemented into EXP on the analyses and the short range forecasts. Indeed, as seen in the previous section (section 5.2), selected IASI observations are removed over the more cloudy locations and then impact the humidity and temperature analyses (as seen in Figure 7). Statistical results in Figure 8 report a non-negligible decrease of the dispersion within the FG departure and AN departure for IASI observation and AMSU-A for several channels but some negative impacts have to be noted for other wavenumbers. More attenuated and mainly non-significant impacts can be recorded for CrIS and TEMP-q observations. Thus, the analyses and short range forecasts have been slightly changed compared to REF.

5.4 Impact on forecast scores

The forecasts from EXP at 00UTC for the period 7 December 2017 to 17 January 2018 were compared to REF ones and evaluated against radiosondes and operational analyses from ECMWF. Rootmean square forecast errors at the 12-h forecast ranges with respect to the ECMWF analyses were computed for temperature, relative humidity and wind. Similar computations were made against radiosondes. No major difference can be found between **both** experiments. Very small improvements of the 12-hour forecast with respect to the ECMWF analyses were found in the Southern hemisphere for temperature and wind at around 700 hPa (Figs. 9). This reduction of 2% for temperature and 0.5% for the wind is significant according a Bootstrap test with a 99.5% confidence level. Other improvements are found at 200 hPa for temperature (1.5%) and at 500 hPa for wind (0.5%). Regarding the evaluation against radiosondes, very small, but not significant, improvements for the wind were found in the troposphere in the Southern hemisphere and in the Tropics.

6 Conclusion and perspectives

A new method using of collocated AVHRR cluster information to improve the selection of homogeneous IASI observation scenes within the numerical weather prediction ARPEGE model has been developed at Météo-France for data assimilation purposes and has been presented in this study.

- 5 The first step consisted in adapting the IASI observation operator based on the RTTOV radiative transfer model by using the RTTOV-CLD module with cloudy microphysical parameters (liquid water content (ql), ice content (qi) and cloud fraction) for the simulation of cloudy radiances. A qualitative evaluation of such module showed realistic simulated cloud structures at various locations around the globe with a quite good agreement against IASI observations.

The second and main step of this work was to assess the impact of several methods used to select homogeneous IASI observations using AVHRR clusters. Two selection methods (derived from the literature : Martinet et al. (2013) and Eresmaa (2014))) were preliminarily evaluated. Despite a good improvement in terms of biases and standard deviations of the FG departures, ~~it was found that these two methods were not satisfactory in an operational context (in assimilation) to a large IASI observation reduction.~~ the criteria from the Martinet et al. (2013)'s method favours the homogeneous cloudy observations and retains more than a half of the observations while the Eresmaa (2014)'s method gives priority to clear observations and keeps only 22% of the observations. Then, two ~~new-sets-of~~ criteria were defined from these two previous methods in order to have a more balanced choice of clear and cloudy observations and good statistics in terms of background departures and implemented within the ARPEGE model :

- The first criterion ~~derived from Martinet et al. (2013) method~~ looks for the consistency between different clusters occupying the same IASI FOV by examining this homogeneity relative to the weighted average brightness temperature of the AVHRR clusters; it is only based on observations ~~and computed for both infrared AVHRR channels as in Eresmaa (2014).~~ This criterion allows to retain 67% of observations.
- In addition, the second criterion ~~is derived from Eresmaa (2014)'s test and~~ assesses the coherence of each cluster compared to the background ~~brightness temperature simulation~~; it is in fact a good compromise between ~~the previous criterion and~~ the two "historical" ones with accurate statistics and a sufficient number of observations 36% that passed the check. It also allows to retain the same proportion of homogeneous clear and cloudy observations contrary to the derived Martinet et al. (2013) and Eresmaa (2014) methods.

Therefore, assimilation experiments were conducted to assess the impact of these new selecting homogeneous IASI observation features in the current clear sky assimilation. This revised check was added to the McNally and Watts (2003) cloud detection. The results obtained in this case show that the scenes categorization has been facilitated and cloudy observations can be better filtered out compared to what is done in the operational ARPEGE version. 3% of all observations are rejected with the compromise method ~~and only 1% for the method based only on homogeneity in observation space which is more convenient for the assimilation.~~ The impacts on the first guess and analysis departures (showing more Gaussian shape) are generally low but with a beneficial reduction on the standard deviation of first guess departures mainly on the IASI and AMSU-A observa-

tions. Regarding the forecasts scores, neutral impact is reported when these selection criteria are taken into account on top of the McNally and Watts (2003) algorithm.

However, this step has been necessary to prepare the future which will consist of the assimilation all sky within the ARPEGE model. **This method** of observation selection allows to separate the clear-sky and cloudy scenes and manage each route in an independent way. Then, it could be available to directly assimilate the cloudy radiances into the 4D-Var ARPEGE by adapting the observation errors for more cloudy situations. However, hydrometeors used in the RTTOV-CLD are not available into the background error covariance matrix and then cloudy and convective situations are badly represented and will penalise the cloudy direct assimilation. In order to bypass this problem a the second solution under study is to retrieve information within cloudy observation by a Bayesian inversion method, in a first step, and assimilate these retrieved products in terms of temperature and/or humidity profiles into the 4D-Var in a second step. This method called 1D-Bayesian + 4D-Var was already studied for microwaves radiances (Guerbette et al., 2016; Duruisseau et al., 2018) and is successfully used since 2010 for radar reflectivities (Wattrelot et al., 2014) assimilation within the AROME convective scale model.

Competing interests. No competing interests are present here.

Acknowledgements. CNES and Météo-France funded this work through a PhD grant for Imane Farouk. Cloud classification data from MSG are provided by Météo-FranceCMS. We thank the ICARE Data and Services Center for providing access to the data used in this study. Jean-Francois Mahfouf and Jean Maziejewski are warmly thanked for their help in revising a previous version of the manuscript.

References

- Auligné, T., McNally, A., and Dee, D.: Adaptive bias correction for satellite data in a numerical weather prediction system, *Quarterly Journal of the Royal Meteorological Society*, 133, 631–642, 2007.
- Baran, A. J., Cotton, R., Furtado, K., Havemann, S., Labonnote, L.-C., Marengo, F., Smith, A., and Thelen, J.-C.: A self-consistent scattering model for cirrus. II: The high and low frequencies, *Quarterly Journal of the Royal Meteorological Society*, 140, 1039–1057, <https://doi.org/10.1002/qj.2193>, <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2193>, 2014.
- 5 Bauer, P., Geer, A. J., Lopez, P., and Salmond, D.: Direct 4D-Var assimilation of all-sky radiances. Part I: Implementation, *Quarterly Journal of the Royal Meteorological Society*, 136, 1868–1885, 2010.
- Bechtold, P., Bazile, E., Gichard, F., Mascart, P., and Richard, E.: A mass-flux convection scheme for regional and global models, *Q. J. R. Met. Soc.*, 127, 869–886, 2001.
- 10 Berre, L., Varella, H., and Desroziers, G.: Modelling of flow-dependent ensemble-based background-error correlations using a wavelet formulation in 4D-Var at Météo-France, *Quarterly Journal of the Royal Meteorological Society*, 141, 2803–2812, <https://doi.org/10.1002/qj.2565>, 2015.
- Bougeault, P.: A simple parameterization of the large-scale effects of cumulus convection, *Mon. Wea. Rev.*, 113, 2108–2121, 1985.
- 15 Cayla, F.: L’interféromètre IASI; un nouveau sondeur satellitaire à haute résolution., *La météorologie 8ème série*, 32, 23–39, 2001.
- Chalon, G., Cayla, F., and Diebel, D.: IASI- An advanced sounder for operational meteorology, in: *IAF, International Astronautical Congress*, 52 nd, Toulouse, France, 2001.
- Chevallier, F., Lopez, P., Tompkins, A. M., Janiskova, M., and Moreau, E.: The capability of 4D-Var systems to assimilate cloud-affected satellite infrared radiances, *Q. J. R. Met. Soc.*, 130, 917–932, 2004.
- 20 Collard, A.: Selection of IASI channels for use in numerical weather prediction, *Quarterly Journal of the Royal Meteorological Society*, 133, 1977–1991, 2007.
- Courtier, P., Freydl, C., Rabier, F., and Rochas, M.: The ARPEGE Project at Météo-France, *ECMWF Seminar Proceedings*, 7, 193–231, 1991.
- Duruiseau, F., Chambon, P., Wattrelot, E., Barreyat, M., and Mahfouf, J.-F.: Assimilating cloudy and rainy microwave observations from SAPHIR on-board Megha-Tropiques within the ARPEGE global model, *Q. J. R. Met. Soc.*, submitted, 2018.
- 25 Eresmaa, R.: Imager-assisted cloud detection for assimilation of Infrared Atmospheric Sounding Interferometer radiances, *Quarterly Journal of the Royal Meteorological Society*, 140, 2342–2352, 2014.
- Errico, R. M., Bauer, P., and Mahfouf, J.-F.: Issues regarding the assimilation of cloud and precipitation data, *Journal of the Atmospheric Sciences*, 64, 3785–3798, 2007.
- 30 Fourrié, N. and Rabier, F.: Cloud characteristics and channel selection for IASI radiances in meteorologically sensitive areas, *Q. J. R. Met. Soc.*, 130, 1839–1856, 2004.
- Geer, A., Baordo, F., Bormann, N., Chambon, P., English, S., Kazumori, M., Lawrence, H., Lean, P., Lonitz, K., and Lupu, C.: The growing impact of satellite observations sensitive to humidity, cloud and precipitation, *Quarterly Journal of the Royal Meteorological Society*, 143, 3189–3206, 2017.
- 35 Greenwald, T. J., Hertenstein, R., and Vukićević, T.: An all-weather observational operator for radiance data assimilation with mesoscale forecast models, *Monthly weather review*, 130, 1882–1897, 2002.

- Guerbette, J., Mahfouf, J.-F., and Plu, M.: Towards the assimilation of all-sky microwave radiances from the SAPHIR humidity sounder in a limited area NWP model over tropical regions, *Tellus A: Dynamic Meteorology and Oceanography*, 68, 28 620, 2016.
- Guidard, V., Fourrié, N., Brousseau, P., and Rabier, F.: Impact of IASI assimilation at global and convective scales and challenges for the assimilation of cloudy scenes, *Quarterly Journal of the Royal Meteorological Society*, 137, 1975–1987, 2011.
- 5 Hilton, F., Armante, R., August, T., Barnet, C., Bouchard, A., Camy-Peyret, C., Capelle, V., Clarisse, L., Clerbaux, C., Coheur, P.-F., Collard, A., Crevoisier, C., Dufour, G., Edwards, D., Faijan, F., Fourrié, N., Gambacorta, A., Goldberg, M., Guidard, V., Hurtmans, D., Illingworth, S., Jacquinet-Husson, N., Kerzenmacher, T., Klaes, D., Lavanant, L., Masiello, G., Matricardi, M., McNally, A., Newman, S., Pavelin, E., Payan, S., Péquignot, E., Peyridieu, S., Phulpin, T., Remedios, J., Schlüssel, P., Serio, C., Strow, L., Stubenrauch, C., Taylor, J., Tobin, D., Wolf, W., and Zhou, D.: Hyperspectral Earth observation from IASI: Five years of accomplishments, *bulletin of the american meteorological Society*, 93, 347–370, <https://doi.org/10.1175/BAMS-D-11-00027.1>, 2012.
- 10 Janiskova, M., Thépaut, J.-N., and Geleyn, J. F.: Simplified and regular physical parametrizations for incremental 4Dvar assimilation, *Mon. Wea. Rev.*, 127, 26–45, 1999.
- Lopez, P.: Implementation and validation of a new prognostic large-scale cloud and precipitation scheme for climate and data-assimilation purposes, *Quarterly Journal of the Royal Meteorological Society*, 128, 229–257, <https://doi.org/10.1256/00359000260498879>, 2002.
- 15 Martinet, P., Fourrié, N., Guidard, V., Rabier, F., Montmerle, T., and Brunel, P.: Towards the use of microphysical variables for the assimilation of cloud-affected infrared radiances, *Quarterly Journal of the Royal Meteorological Society*, 139, 1402–1416, 2013.
- Martinet, P., Fourrié, N., Bouteloup, Y., Bazile, E., and Rabier, F.: Toward the improvement of short-range forecasts by the analysis of cloud variables from IASI radiances, *Atmospheric Science Letters*, 15, 342–347, 2014.
- McNally, A.: A note on the occurrence of cloud in meteorologically sensitive areas and the implications for advanced infrared sounders, *Q. J. R. Met. Soc.*, 128, 2551–2556, 2002.
- 20 McNally, A. and Watts, P.: A cloud detection algorithm for high spectral resolution infrared sounders, *Q. J. R. Met. Soc.*, 129, 3411–3423, 2003.
- McNally, A. P.: The direct assimilation of cloud-affected infrared satellites radiances for numerical weather prediction, *Q. J. R. Met. Soc.*, 135, 1214–1229, 2009.
- 25 Menzel, W., Smith, T., and Stewart, T.: Improved cloud motion wind vector and altitude assignment using VAS, *J. appl. Meteorol*, 22, 377–384, 1983.
- Okamoto, K.: Evaluation of IR radiance simulation for all-sky assimilation of Himawari-8/AHI in a mesoscale NWP system, *Quarterly Journal of the Royal Meteorological Society*, 143, 1517–1527, <https://doi.org/10.1002/qj.3022>, 2017.
- Okamoto, K., McNally, A., and Bell, W.: Progress towards the assimilation of all-sky infrared radiances: an evaluation of cloud effects, *Quarterly Journal of the Royal Meteorological Society*, 140, 1603–1614, 2014.
- 30 Pangaud, T., Fourrié, N., Guidard, V., Dahoui, M., and Rabier, F.: Assimilation of AIRS radiances affected by mid to low level clouds, *Mon. Wea. Rev.*, pp. 4276–4292 (17)DOI: 10.1175/2009MWR3020.1., 2009.
- Pavelin, E., English, S., and Eyre, J.: The assimilation of cloud-affected infrared satellite radiances for numerical weather prediction, *Q. J. R. Met. Soc.*, 134, 737–749, 2008.
- 35 Pequignot, P. and Lonjou, V.: Dossier de définition des algorithmes IASI, Internal report ia.df.0000.2006.cne., CNES, 2009.
- Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J.-F., and Simmons, A.: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. , *Q. J. R. Met. Soc.*, 126, 1143–1170, 2000.

- Saunders, R., Hocking, J., Rundle, D., Rayer, P., Matricardi, M., Geer, A., Lupu, C., Brunel, P., and Vidot, J.: RTTOV-11 Science and Validation Report, Tech. rep., Available at https://nwpsaf.eu/deliverables/rtm/docs_rttov11/rttov11_svr.pdf, 2013.
- Stengel, M., Lindskog, M., Undén, P., Gustafsson, N., and Bennartz, R.: An extended observation operator in HIRLAM 4D-VAR for the assimilation of cloud-affected satellite radiances, *Quarterly Journal of the Royal Meteorological Society*, 136, 1064–1074, 2010.
- 5 Vidot, J., Baran, A. J., and Brunel, P.: A new ice cloud parameterization for infrared radiative transfer simulation of cloudy radiances: Evaluation and optimization with IIR observations and ice cloud profile retrieval products, *Journal of Geophysical Research: Atmospheres*, 120, 6937–6951, 2015.
- Wattrelot, E., Caumont, O., and Mahfouf, J.-F.: Operational implementation of the 1D+ 3D-Var assimilation method of radar reflectivity data in the AROME model, *Monthly Weather Review*, 142, 1852–1873, 2014.
- 10 Zhu, Y., Liu, E., Mahajan, R., Thomas, C., Groff, D., Van Delst, P., Collard, A., Kleist, D., Treadon, R., and Derber, J. C.: All-Sky Microwave Radiance Assimilation in NCEP’s GSI Analysis System, *Monthly Weather Review*, 144, 4709–4735, 2016.

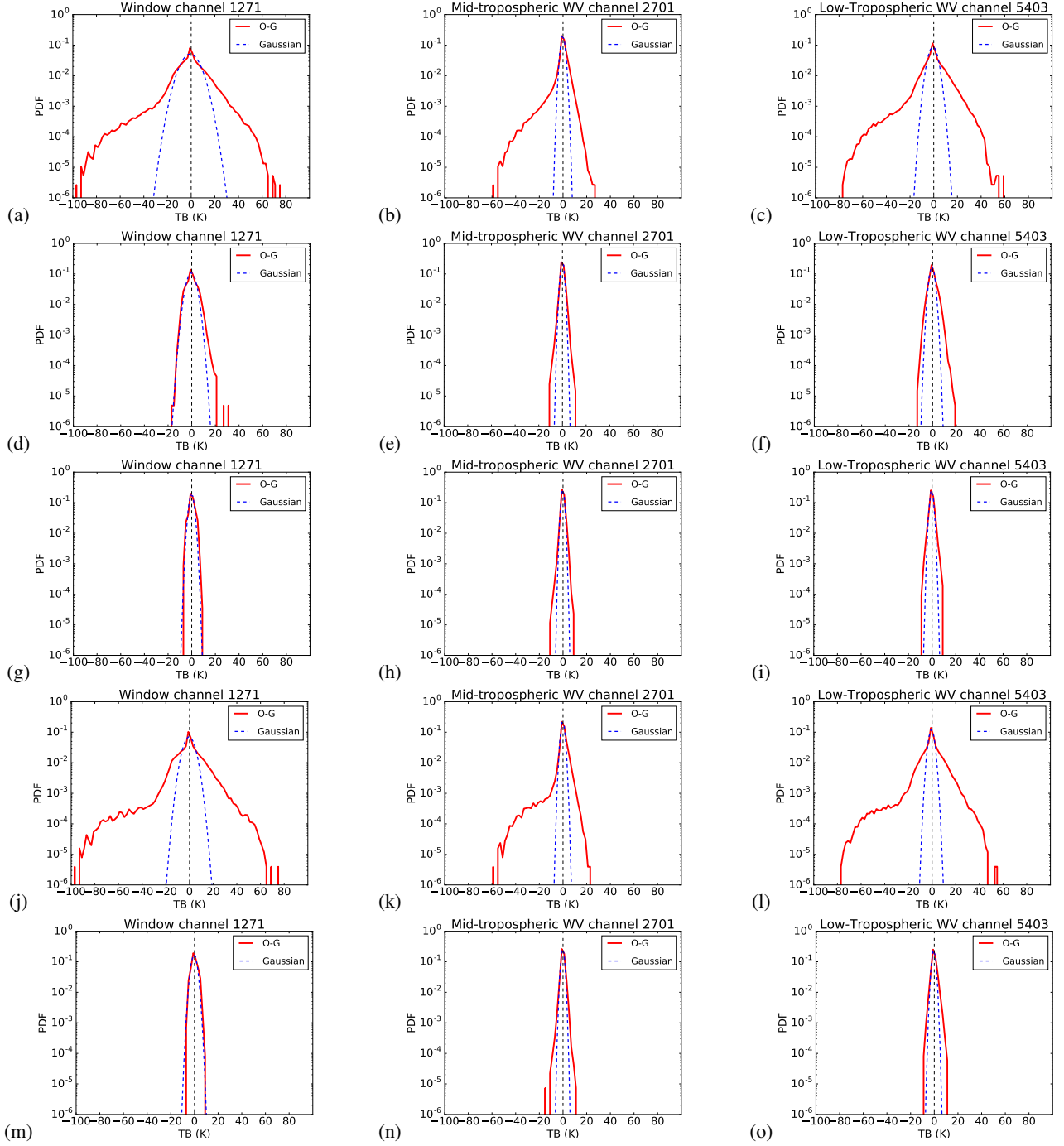


Figure 4. Frequency distribution of brightness temperature difference between observation and background (O-G) for all observations (a, b, c), after applying the homogeneity criteria derived from Martinet et al 2013 (d, e, f), the homogeneity criteria derived from Eresmaa 2014 (g, h, i), the third method based on observation space method (j, k, l) and the compromised approach (m, n, o). The PDF are presented for three channels: window channel 1271, low-tropospheric water vapour channel 5403, and mid-tropospheric water vapour channel 2701). The Gaussian distributions with the same error characteristics (mean and standard deviation) are also shown in blue dashed lines.

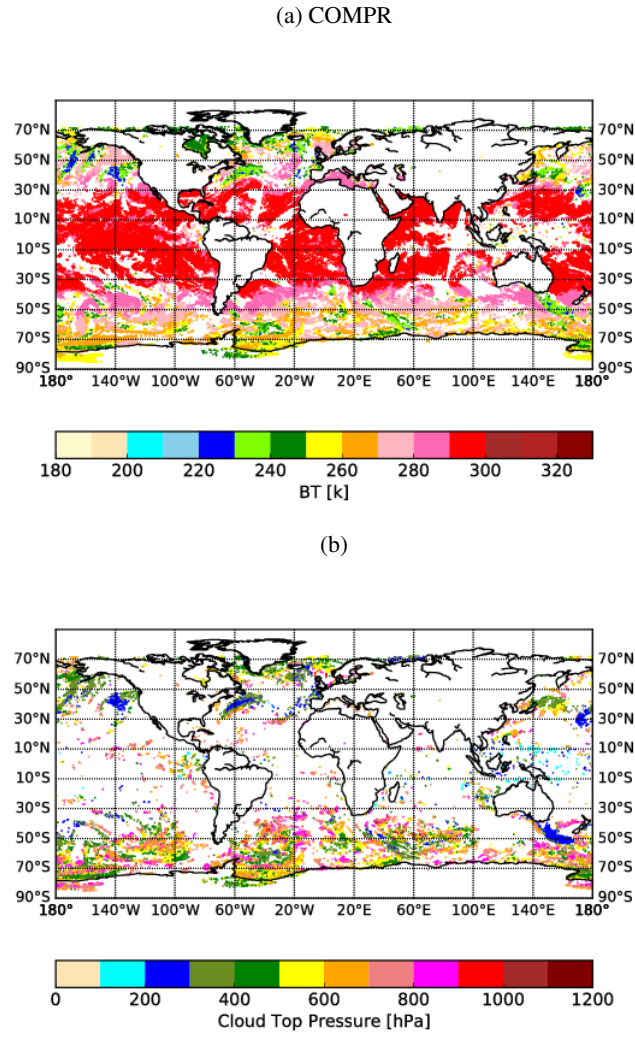


Figure 5. Map of IASI observations of brightness temperature (K) for surface channel ($1271, 962.5 \text{ cm}^{-1}$), after applying the COMPR method (a), cloud-top pressure (hPa) observations retrieved from a CO_2 -slicing algorithm applied on IASI data (b), for 30 January 2017 daytime over sea.

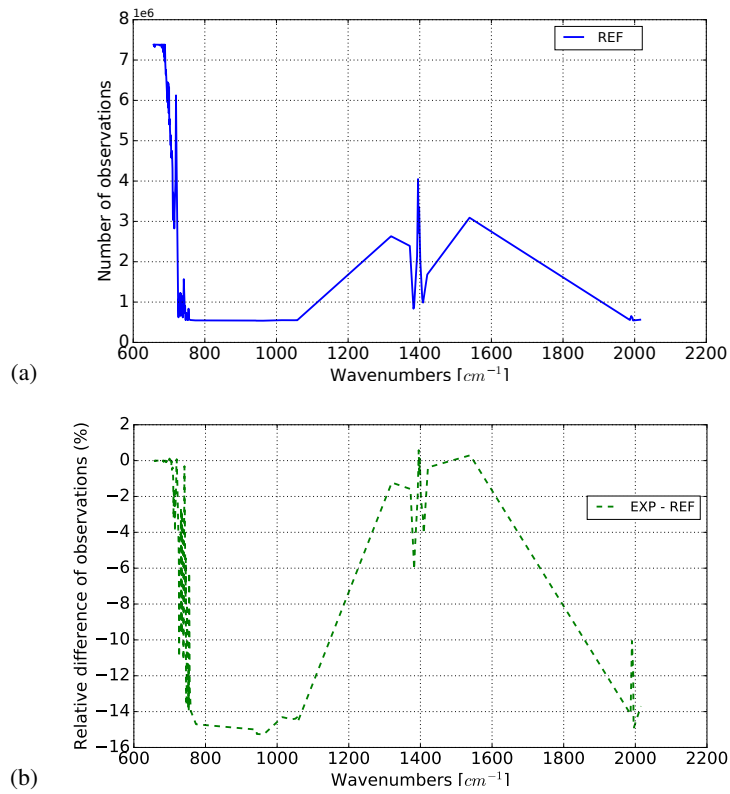


Figure 6. Number of assimilated IASI data over the whole experimental period (41 days) as a function of wavenumber of IASI for the REF experiment (a) and relative difference of number of assimilated observations (b) for the EXP compared with REF (dashed green line).

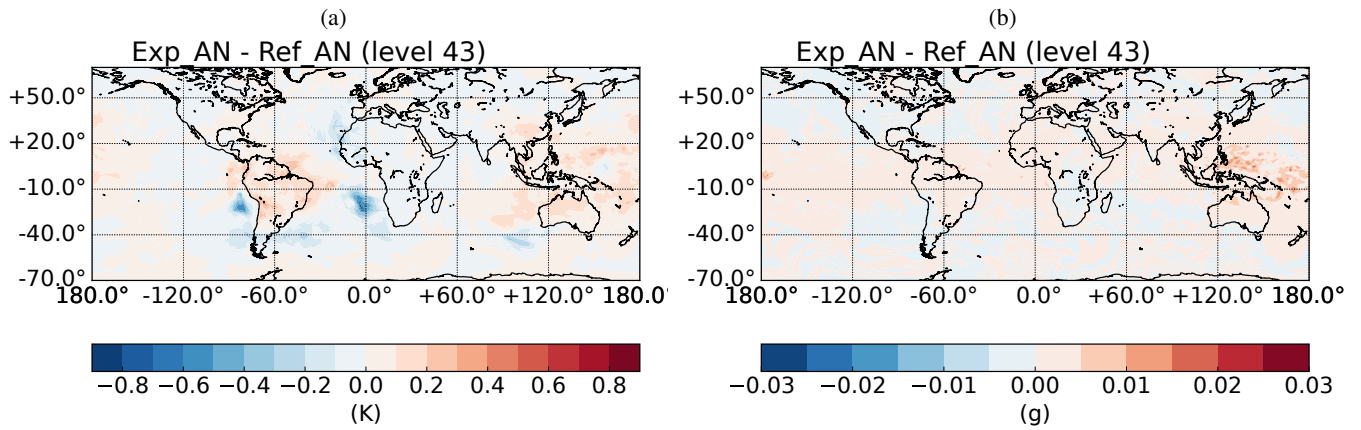


Figure 7. Temperature (a) and humidity (b) analyses difference between REF and EXP for the first assimilation cycle on 07 December 2017 at 00 UTC at ARPEGE model level 43 which corresponds to 200 hPa.

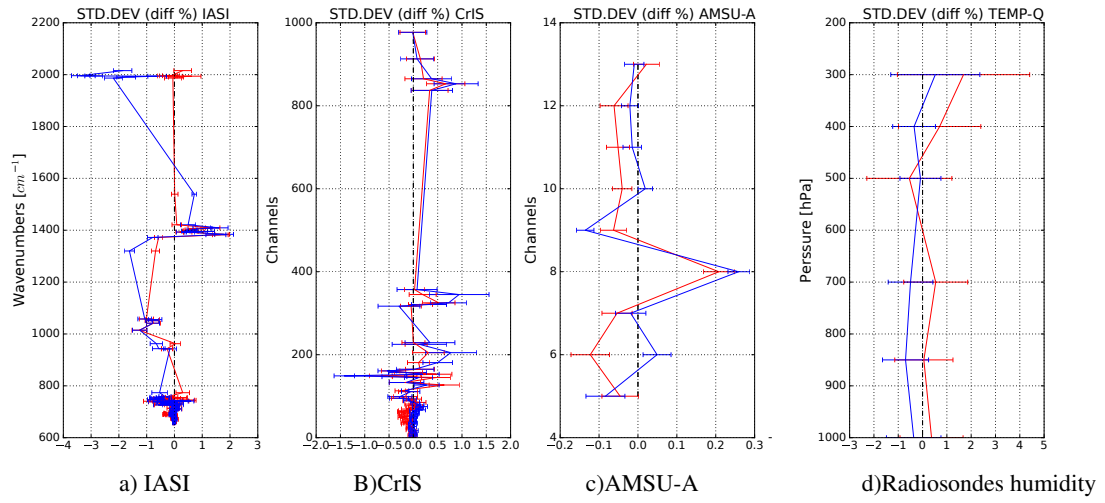


Figure 8. Relative differences of FG departure (red curve) and AN departure (blue curve) standard deviation between EXP and REF for IASI (a), CrIS (b), AMSU-A (c) and TEMP-q (d). The horizontal error-bars represent the 95% significance value for each difference.

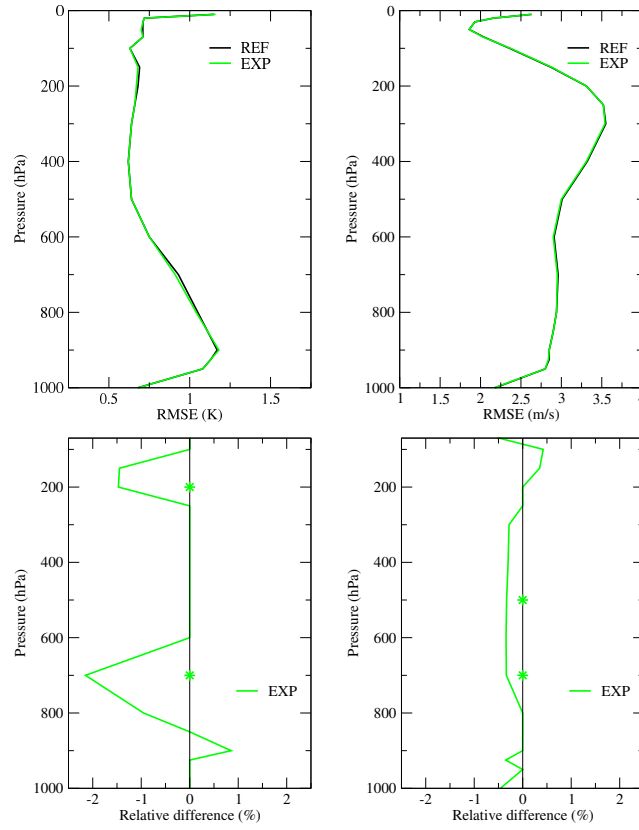


Figure 9. Root mean square error of the 12-hour forecast error for the Southern Hemisphere computed with respect to ECMWF analysis over the period from 7 December 2017 to 17 January 2018 for REF (in black) and EXP (in green) for temperature (left-hand side) and wind (right-hand side). The second line represents the relative difference with respect to the reference. Green stars indicate that the differences of EXP with REF are statistically significant according a Bootstrap test with a 99.5% confidence level.