Anonymous Referee #1

Update of IASI channel selection with correlated observation-errors for NWP

By Coopmann et al.

Suggestions for revision or reasons for rejection (will be published if the paper is accepted for final publication)

This paper is substantially better than it was before, and has a sensible purpose and the more limited scope is much better presented. I also really appreciate that this paper is now a lot less verbose.

I would like to thank the referee for their valuable comments and suggestions that have significantly improved the scientific scope of this paper.

The review is as follows:

The original text from the referee is in black, the answers in blue and the modifications incorporated into the paper in red.

However, there are two things that I would like to see expanded before publication

1) The method is still not clear enough: the bit about the link between ARPEGE and MOCAGE is now much better explained, but what's lacking is enough detail for someone else to be able to replicate your method. What I understand from what's written is that you select 400 channels for each of 60 profiles. What's not clear is how you pick the final 400 channels.

You are right, it is important that the method can be replicate. I'm sorry if the explanation isn't clear yet. For each of the 60 profiles, we selected 400 channels. Then, we computed statistics on the channels selected. Among the 5499 available channels, 1779 have been selected for at least one profile. Their selection frequency has been calculated and they were ranked according to this frequency.

I have added these additional descriptions in the paper for more clarity on the selection method:

new P11 L30: “ As described in Section 2.2, for each of the 60 profiles, we looked for the channel with the highest DFS value, then the channel pair with the highest DFS value, and so on. “

new P12 L58: " From these results we can sort the channels selected at least once (1779) according to their selection frequency. Thus the n most frequently selected channels will form a new selection of n channels. "

2) You talk about the fact that your Jacobians have quite mixed sensitivities as a positive - but do not address the fact that that was exactly what Collard was trying to avoid - and a major reason why PC Scores have not become operationally assimilated. Also, the avoidance of large parts of the window and water vapour channel in his selection was due to intentional removal of channels that have sensitivity to other species, which you don't address in your selection. It would be useful to understand the sensitivities to these other species in the channels that you have selected. Do you have channels sensitive to Methane, CO, HCN, NOx, etc, that might need to be screened out in the case of a volcanic eruption, biomass burning or other event?

Indeed, Collard tried to avoid selecting channels with mixed sensitivities. This is complicated since there are very few channels sensitive to a single variable in the IASI spectrum. Thus, the question is to know to which species IASI channels are sensitive and whether the variability of these species has an effect on the brightness temperature. To these questions, Collard answers that only those channels that are sensitive to the following 3 species (CH_4, CO and N_2O), have a significant enough impact on the brightness temperature (> 1 K) to be blacklisted.

In our study, the use of an R matrix accounting for inter-channel error correlations allows us to exploit the potential of mixed sensitivities of IASI channels mainly in terms of temperature and humidity information.

As we use spectral bands 1 and 2 (645 - 2019.75 cm^{-1}) in this study, we did not select CO-sensitive channels (2100 - 2150 cm^{-1}). However, it is possible that some selected channels in the water vapour sensitivity band between 1210 and 1650 cm^{-1} are also sensitive to CH_4 and N_2O.

Although it is important to know whether the selected channels are sensitive to other species, particularly in the event of a volcanic eruption, biomass burning or other event, most NWP centres have effective quality control of the observations and are able to reject potentially problematic channels.

The following texts were added to the paper. We also propose to add a short appendix that describes the sensitivity of the spectrum to various species and the location of the channels for the CS275, NS275

and NS400, if you and the Associate Editor find it has some added value for the paper.

new P3 L56: " Indeed, unlike the selection method chosen by Collard, the use of an R matrix accounting for inter-channel error correlations, allows us to consider all the channels sensitive to several variables (temperature from the CO_2 band, water vapour, ozone, skin temperature in the atmospheric window). Note that the IASI spectrum also is sensitive to main absorbing gases (CH_4, CO and N_2O) and weaker absorbers (CCl_4, CFC-11, CFC-12, CFC-14, HNO_3, NO_2, OCS, NO and SO_2). "

new P16 L74: " It should be noted that some channels selected in this study may be sensitive to minor gases and others selected between 1210 and 1650 cm^{-1} may be sensitive to CH_4 and N_2O. However, none are sensitive to CO (2100 - 2150 cm^{-1}) since the selection was limited to channels up to 2019.75 cm^{-1} (more details in Appendix A). "

Referee Report: amt-2019-242-referee-report.pdf

P1 L61: The Collard paper was based on Rodgers, not Rabier. There is an important difference, as mentioned in my previous review

Corrected

P2 L7: instrument health monitoring

Corrected

P2 L14: short-range

Corrected

P3 L69: it?

Clarified

P4 L17: but in this case, as in operational NWP,

Corrected

P4 L23: change in a

Corrected

P6 L17: and background error!

Added

P6 L60: SD also because of representation error?
These different values of standard deviations derive mainly from the accumulated errors in the approximations made on the humidity conversions and also from the difference in the representation of humidity between the troposphere, where humidity is well represented in the models, and the stratosphere, where humidity is difficult to model. You are right that the representation error may also vary depending on where each channel is peaking (text added).

P6 L66: Not an appropriate reference for the NWPSAF 1D-Var
Corrected

P7 L15: I would say, began to be used for operational NWP: 1986 is a long time ago!
Added

P7 L24: Again "have explored" sounds a bit passive. They're already using them
Clarified

P7 L55: Maybe mentioned later, but which Desrozier assumptions are broken by doing this in 1D-Var? You haven't mentioned the background error matrix yet. I think it might be better to reorder the sections as the B-matrix is an input into the calculation
At the beginning of the 1D-Var process, the R and B matrices are quite independent. Indeed, it makes more sense to put the section on background-errors before observation-errors. Section 3.3 and section 3.4 have been swapped.

P8 L25: Do you have something to cite here?
Added (Bormann el at 2016, Migliorini 2015, Stewart et al 2014 and Weston et al 2014).

P9 L2: Presumably this is the same forecast is in the second row of your flow chart? I wonder if it's possible to make that clearer in the diagram?
You are right: the first 24 hours are the same, but for further forecast range, the meteorological forcing is not the same in the two simulations. It's complicated to show it on the diagram.

P9 L14: It's not completely clear to me - MOCAGE +ARPEGE has 60 levels; RTTOV has 54, of which you ignore 6 because too high or too low, so you use 60 levels to interpolate onto 48 levels, which sounds reasonable in theory, but are there parts of the profile where RTTOV is more dense than MOCAGE? Thus leaving your matrix under-determined? Have you had to regularise the matrix at all? And if so, how?
Concretely, I interpolated the fields from 60 to 54 levels and I calculated the matrix B on these 54 vertical levels. Contrary to the matrix R, the large statistical number of cases makes it possible to directly obtain a definite positive matrix. Then from this matrix B over 54 levels, I constructed a new matrix B with 48 levels for temperature and ozone and 27 levels for humidity to be in agreement with the levels I wish to minimize during the 1D-Var process. That is true that the 60 MOCAGE levels are not spaced the same way as the 54 RTTOV

levels. Nevertheless, they both have higher density close to the surface and around the tropopause.

P9 L40: And for the top few levels?!

Added:

new P9 L95: " Correlations are weaker in the stratosphere and increase in the upper stratosphere probably due to interpolation as mentioned above. "

P10 L10: You don't say anything about your Tskin background error?

Added:

new P9 L11: " The background-errors of skin temperature, surface temperature and surface humidity used in this study are derived from the values available in the reference B-matrix of the 1D-Var. The background error standard deviation value for skin temperature is 2.0 K. "

P11 L6: What if you want to retrieve surface emissivity?

Indeed, we do not address surface emissivity retrieval in this paper, as we are using emissivity atlases at Meteo-France, as in some other NWP centres. For those who retrieve surface emissivity in their assimilation process, some channels would need to be added to the current selection.

P11 L29: So it's not clear to me how you actually did the channel selection. You do the iterative calculation for each profile, and then what? Pick the channels from the right hand side of the bar chart above, working towards the left, until you have 400 channels?

As explained in the main question.

P11 L33: I don't think more accurate is really the correct term. "allows more effective identification of the most informative channels"?

Fully in agreement with this suggestion. Corrected

P12 L15: This is partly because of blacklisting to avoid channels sensitive to trace gases that aren't modelled

It is right that some of the ozone channels also are sensitive to CO_2 and H_2O.

P13 Fig 11: I note your water vapour channels are more sensitive to the lowest layers of the temperature profile - will that be significant?

Indeed, we can notice that further improvement can be obtained with NS275 in the lowest layers in figure 12.d.

P13 L3: This shouldn't be a surprise - the Collard selection chose pure window channels and weak water vapour lines manually. You've used a temperature retrieval that removes the bottom two levels of the profile.

We did reject the lowest to levels as they are most of the time below the surface (pressures are 1033 and 1050 hPa).

P13 L8: You could argue that with a univariate B-matrix, it would be better to choose channels which have less sensitivity to species other than their primary sensitivity - it is not demonstrated anywhere that these features you point out are advantageous.
That is correct.

P13 L20: sensitive
Corrected

P13 L23: Or, it could mean that you find it harder to retrieve surface emissivity down the line because your window channels are sensitive to water vapour...
Again, we do not address surface emissivity retrieval here.

P14 L21: 900? 800? Hard to tell for 1000!
Corrected

P14 L26: I don't think you mean that exactly - it sounds like you are saying 400 channels is less good than 275. I think you mean that 400 offers a small improvement over 275 in some parts of the atmosphere.
Modified

P15 L19: we provide a more detailed description
Corrected

P15 L34: I would have been interested to see whether the results were different if you used the B-matrix that Collard used. You have shown convincingly that your selection is better for the problem you have defined, but not necessarily that it is better for the problem that he defined.
If I can access the B-matrix data he used at the time, indeed, it would be interesting to carry out new experiments to compare the results. The further evaluation we plan in the ARPEGE 4DVAR will probably help to have more evidence (or not).

P15 L74: I'm not sure what the relevance of this is? The average Jacobians aren't used anywhere (and really shouldn't be either).
Average Jacobians have been calculated to illustrate the different channel sensitivities in Figure 3, but are not used for channel selection. We modified the text "we calculated the Jacobians" instead of "we calculated the means of the Jacobians"

P16 L41: You shouldn't really throw this in here at this stage of the paper, especially without a reference... And presumably you are referring to the work done for the previous version of the paper, which you yourselves have chosen not to rework into this version of the paper, suggesting that you agree it's a bit iffy...
We were referring to a previous paper [Coopmann et al., 2018] and not to the previous version of the present paper. We decided to drop this sentence.

Anonymous Referee #2

Update of IASI channel selection with correlated observation-errors for NWP

By Coopmann et al.

Suggestions for revision or reasons for rejection (will be published if the paper is accepted for final publication)

Summary

The paper describes a new method for IASI channel selection which is based on a statistically derived observation error covariance matrix rather than the simple diagonal models assumed in the past. This representation is intended to allow for spectral correlations introduced by the forward model errors as well as instrument noise.

The determination of this OE covariance follows the method suggested by Desroziers which is essentially to evaluate the NWP model background covariance, the IASI-background covariance and attribute the difference to observation error. However, evaluating an appropriate background covariance it itself a difficult task.

It is noted that OE covariances are now regularly evaluated and used in NWP assimilation schemes, but these are all based on predefined channel selections. So the main purpose of this work is to try to incorporate the OE covariance into the channel selection itself.

The authors demonstrate that their new selection leads to improved performance in a data assimilation context compared to the 'standard' channel selection of Collard.

I would like to thank the referee for their valuable comments and suggestions that have significantly improved the scientific scope of this paper.

The review is as follows:

The original text from the referee is in black, the answers in blue and the modifications incorporated into the paper in red.

Main Comments

(for these, I do not expect the authors to revise their experiment, but I would like to see either some explicit answer or at least acknowledgement of the questions in their revised paper).

1) The Collard selection had a number of restrictions, particularly

a) exclusion of spectrally adjacent channels

b) selection of CO_2-only, then allowing channels with H_2O and O_3 sensitivity Both these restrictions are relaxed in this case so it is not clear how much improvement is simply due to this relaxation and how much due to the more complex representation of the OE covariance. In particular I would expect (a) to limit the ability to sound at higher altitudes since most of the high-altitude CO_2 channels are concentrated around the Q-branch at 667cm-1, while (b) would limit the channel selection to the edge of the main CO_2 band rather than extending across the window region where water vapour starts to contribute. There should at least be some explicit acknowledgement that removing these restrictions would, by themselves, be expected to lead to some improvement.

Indeed, it is true that having access to more channels provides additional information. We do acknowledge this in some sections of the paper, such as when describing the Jacobians of the channels. But it is the use of an R matrix that takes into account the inter-channel error correlations that makes it possible to efficiently exploit all of these channels.

2) Local Jacobian matrices (H) were computed for each of the 60 profiles, but (it seems) the same global R and B matrices were used. Ideally, local R and B matrices would also be used unless there is good reason to assume that these are constant (which seems unlikely). Was this considered?

Constant global R and B matrices have indeed been used in this study. In NWP models, constant R matrices are used also. On the contrary, most (not to say all) NWP centres used at least varying sigma_b and, in most of cases, a localized B of the day. The latter is obtained from Ensemble Data Assimilation systems or Ensemble Var systems. In our study, the B matrix is derived from a NMC method, as there is no coupled data assimilation between NWP and chemistry at Météo-France yet. We fully agree that errors in the prior state vary depending on location (and time of the year / day). Even if our sample for the NMC method is large (1 year), it was not large enough to obtain robust statics for a given area and a given month. Which is why we decided to use a global B matrix as a reasonable approximation.

3) It appears that 60 separate channel selections were derived (Fig 9) but it was unclear how these were combined into a single selection. It could be that the selection across all 60 profiles was performed simultaneously using an aggregated or averaged DFS but this isn't stated.

I'm sorry if the explanation isn't clear yet. For each of the 60 profiles, we selected 400 channels. Then, we computed statistics on the channels selected. Among the 5499 available channels, 1779 have been selected for at least one profile. Their selection frequency has been calculated and they were ranked according to this frequency.
I have added these additional descriptions in the paper for more clarity on the selection method:

new P11 L30: " As described in Section 2.2, for each of the 60 profiles, we looked for the channel with the highest DFS value, then the channel pair with the highest DFS value, and so on. "

new P12 L58: " From these results we can sort the channels selected at least once (1779) according to their selection frequency. Thus the n most frequently selected channels will form a new selection of n channels. "

4) The assimilation formalism (eg Eq 6) assumes that the background and observation error covariances are completely uncorrelated with each other. That might be true until the first IASI measurements are assimilated but thereafter the background covariance updated for the next assimilation step will have some component from the OE error so no longer strictly independent. Where a large amount of information for the NWP is coming from other sources this may be negligible, but it should be stated somewhere that this has been assumed (although the authors also note P2 L20) that 60% of all observations assimilated in ARPEGE are from IASI.

You're right, background errors and observations errors may not be really independent. As IASI represents a large part of the assimilated observations in ARPEGE, the quality of the forecast is linked to the assimilation of the these observation, as can be claimed by the FSOI metrics (Forecast sensitivity to observations, which is widely used in NWP centres). Nevertheless, the link between forecast and observation error is never evaluated and most NWP centres still neglect these correlations. We can also add that, even if IASI represents 60% of the assimilated observations, the DFS brought by IASI in ARPEGE only represents 20-25%.

5) The authors show the main features of their derived OE covariance (Fig 5) but there is very little discussion of whether they believe this is truly representative or how much of the magnitude may be ascribed to uncertainties in the evaluation of the background covariance. Why, for example, would the observation error be larger in the ozone band than in the H2O bands? On the other hand it seems quite plausible that the ozone contribution to the background error covariance has not been properly characterised. If the background and observation errors are indeed uncorrelated, is it just coincidence that where the observation error is large the background error is also large?

We agree with you. You can notice that the diagnosed observation error (blue curve) in the ozone band have values similar to some channels in the water vapour band. On the contrary, the standard deviations of Obs – Guess are much higher in the ozone band, as this quantity aggregates both observation error and background error contributions. I added these explanations:

P11 L33: " The higher values observed in the ozone and water vapour band for observation-error standard deviations are probably due to errors in the radiative transfer modelling because of larger biases for these variables. Indeed, the ozone profiles from MOCAGE used as an input variable to RTTOV are more realistic than the single profile but they have biases that can affect the quality of the simulations. Similarly, the humidity profiles from ARPEGE are more realistic in the troposphere than in the stratosphere, which can lead to poor simulations of sensitive water vapour channels in the stratosphere. Hence these high and low standard deviations in the water vapour band. "

Minor/Typographical Comments

Abstract

P1 L2: IASI = Infrared Atmospheric Sounding Interferometer

Corrected

P1 L24: I suggest temperature error more conventionally expressed in K rather than % (also elsewhere)

The impact on atmospheric levels is different. Moreover, we do not evaluate the retrievals against verification data. We find more efficient to have a global view of the improvement on the analysis errors in percentage.

P1 L40: 'The IASI spectrum ...'

Corrected

P2 L17: I think "This 'analysis' state is thus ..." would read better.

Corrected

P2 L25: What does uncorrelated 'vertically' mean with regard to IASI? There is no vertical coordinate in the measurement space.

Corrected: spectrally

P2 L41: For the opening sentence of this paragraph it is not clear whether the authors are referring to previous work or what they will be presenting in this paper.

Indeed, the sentence was unclear and is replaced by:

P2 L41: " Currently, cross-channel observation error correlations are estimated for infrared sounders whose channel selections have already been made. "

There should be some clarification here: if the only 'observation error' considered is the instrument noise, then it is quite valid to consider only the diagonal elements of the covariance matrix, with the additional precaution of discarding adjacent channels (a side-effect of the Gaussian apodisation being to introduce correlations in the noise between adjacent spectral points).

P2 L64: Jacobian (with capital J), and numerous other instances.

Corrected

P2 L88: For completeness, which IASI instrument? And were the profiles restricted to near-nadir views or did you sample the full across-track swath? Did you use different FOV elements from the set of four?

Added:

P2 L88: " Then, from this setup, we selected 6123 IASI pixels at near-nadir views (Metop-A and B) ... "

P2 L98: At nighttime is the AVHRR cloud flag reliable? How would it distinguish between clear surface and a stratiform cloud top? Usually there is an additional test based on a comparison of the retrieved skin temperature compared with the model forecast.

The AVHRR cloud information that is embedded in IASI L1C data does have defects. At Météo-France, in the global and regional assimilation systems, the only test which is used is the McNally and Watts algorithm, that flags each channels as clear or cloudy in each pixel. On the other hand, when retrieving the land surface temperature from IASI, the AVHRR cloud flag is used, both during day and night.

P3 L33: 'the the'. Also 'Degrees of Freedom' (plural).

Corrected

P3 L33: Strictly speaking DFS is just one of a variety of different criteria that could be used and, since it ignores off-diagonal information, isn't actually the one that provides the 'largest information content'. It is, however, the one that is conventionally used for channel selection, so I have no argument with the choice of DFS here as well.

P3 L49: 'represents the Jacobian matrix ...'

Corrected

P3 L45: Eq (2) should be 'I + ' ....

It is indeed a "+" in this equation.

P3 L61: If you're using 645-2000 that should be 5421 channels, or 5420 without channel 1194. I assume 645-2019.75cm-1 is intended here. also P15 L40

I agree with you. Corrected

P3 Table 1: For completeness give ranges of band 1 and band 2 separately in the table caption

Added

P4 Fig 2 caption: 'subset of the 60 atmospheric profile database'.

Corrected

P4 L1: I expect that once m gets beyond a fews 10s this becomes computationally expensive. Can you give some indication of the time required?

Indeed, the selection required a significant IT cost. A selection of 400 channels for 60 profiles required several days of calculation.

P4 L8: Why is this section titled 'Preliminary' work? That suggests further work will be presented.

Replaced by "Preparatory work"

P4 L23: Use consistent font (ie math font) for BT and X in the text and in Eq(4)

Corrected

P4 L42: 'also to water vapour' - this, and indeed the whole of panel b, are inconsistent with Table 1 which suggests H2O is only retrieved from band 2.

Indeed it is not clear. In table 1 I wanted to indicate the main absorption bands and their main applications. I added "main absorption band".

P4 L51: Temperature sensitivity is, of course, a necessary accompaniment to sensitivity to *any* absorber, so every part of the spectrum will be sensitive either to surface or to atmospheric temperature.

Agreed.

P5 Fig 3: caption missing explanation of panel (d). 'Jacobians' should have capital J. The pressure axis on panel (d) is inappropriate and would be better converted to show the size of the Jacobian.

Corrected

I have chosen to represent the sensitivity of skin temperature Jacobians in order to be consistent with the graphs a, b and c and not to confuse the reader.

P6 L13: The usefulness of a 'poly-sensitive' channel depends mainly on the assimilation scheme modelling all the relevant parameters. This would be true even for a single channel, so inter-channel correlations are no more (or less) important for 'poly-sensitive' channels than for single parameter channels.

We agree that the background error description and the assimilation scheme are important too.

P6 L35: Apart from the large departure of 3K which I guess is at the centre of the CO_2 Q-branch at 667cm-1

That's correct, it is a sensitive channel in the very upper stratosphere. This value is probably due to a misrepresentation of the temperature at this altitude in the model.

P6 L80: 'B and H are ...'

Corrected

P7 Fig 5: It would be better if Fig(a) were flipped so IASI channel number was on the x-axis

The direction of the axes is deliberate in order to be able to compare the left and right figure with the same y-axis.

P7 L44: It is not clear what use is made of the diagonal R matrix. I assume it forms part of the pseudo-inversion procedure for the full R?

The diagnosed R-matrix is calculated in part from the analyses from the 1D-Var data assimilations. The assimilation process needs a starting point which in our case is our diagonal R matrix.

P10 Fig 8: These would be better with logarithmic x-axes.

I want to show a linear evolution of the selection.

P11 Fig 9: y-axis title 'Number of channels' slightly clipped in my PDF file, although that may just be a local issue.

It was clipped too tight. Corrected

P11 L8: Since DFS has a physical interpretation as 'number of independent pieces of information that can be retrieved' (ie effectively independent profile levels) it makes more sense to present these results as actual DFS values rather than percentage of the total. For example I would expect skin temperature to be very close to 1, but '6.5% of total' means nothing to me.

DFS values have added.

P11 L12: I am unclear at this point as to whether you now have 60 separate channel selections, one based on each profile, or just a single selection based on solving for an aggregated/averaged DFS over 60 profiles

Explained earlier.

P11 L30: I believe the Collard selection considered only a single atmospheric profile (?). Are you using a selection based on the same single profile here?

Collard used six AFGL standard atmospheres. I do not use these profiles and I have selected a new database of 60 typical and extreme atmospheric profiles taking into account several atmospheric scenarios.

P11 L44: It does not surprise me that there is not much overlap with the Collard channels. With such high redundancy there are usually a large number of channels available at each step in the iteration, each differing only very slightly in information content. Of course Collard also excludes adjacent chanels to those already selected. But does Collard use the same set of profile levels, which would also modify a selection based on DFS?

Collard made his selection in 2007 using the RT-IASI radiative transfer model which had a different level number of RTTOV. For his selection he considered 82 levels (ignoring the top 8).

P12 L21: 'Jacobians' (capital J).

P13 Fig 11: 'Jacobian' (capital J).

Corrected

P13 L13: In terms of coverage represented purely by the Jacobians I see no reason why inter-channel errors should encourage a more homogeneous coverage. Could this be related to density of profile levels in different regions of the atmosphere? Or ability to select adjacent channels in the new algorithm (I imagine this would be a particular limitation for the Collard algorithm since high altitude CO_2 channels are limited to the 667cm-1 region). Also the Collard selection starts with a limitation to CO_2-only channels so would exclude channels which could be used jointly.

Letting the selection of adjacent channel possible (which is possible because we can take into account interchannel OE correlations) probably helps having a better coverage.

P13 L19/20: '... channels in the first ...'

Corrected

P14 Fig 12: I assume the profile of DFS values are evaluated from (1 - (sigma^a)^2/(sigma_b)^2) for the different selections? So, closely related to the (sigma^a-sigma^b) plots. It would be helpful to make this clear.

DFS profiles are taken from the diagonal of $I\text{-}AB^{-1}$ which is different from (1 - (sigma^a)^2/(sigma_b)^2).