

## ***Interactive comment on “Ozone-sensitive channel selection over IASI full spectrum with correlated observation errors for NWP” by Olivier Coopmann et al.***

### **Anonymous Referee #1**

Received and published: 27 August 2019

Review of: Ozone-sensitive channel selection over IASI full spectrum with correlated observation errors for NWP

by Coopman et al. July 2019

#### Overview

The paper proposes a new method of IASI channel selection based on a full representation of the observational error covariance matrix rather than the standard diagonal (ie uncorrelated error) matrix assumption.

This observational error matrix is derived from the difference between a total covari-

C1

ance matrix based on the residuals from (IASI - model simulations), and a separate estimate of the model-only (or background) error covariance matrix.

The results are analysed with respect to the reduction in standard deviation of the residuals after (effectively) a retrieval has been performed using these radiances. The performance of the new channel selection is also compared with an old channel selection based on the uncorrelated error assumption.

#### General Comments

The essence of the method presented is to derive an observation error covariance  $S_o$  (or  $R$ ) from  $S_t = S_o + S_b$  where  $S_t$  is the total error covariance based on the difference between IASI observations and a RTM calculation based on MOCAGE model output, and  $S_b$  (or  $B$ ) is the background error covariance associated with MOCAGE. The tricky part is to estimate  $S_b$  with sufficient accuracy that  $S_o$  can be derived. As the authors mention, this is routinely performed in data assimilation. In that framework usually I would expect  $S_b \ll S_o$  so that  $S_o \sim S_t$  (ie differences are dominated by the observational error) and inaccuracies in  $S_b$  are less critical.

In this case it also seems to be concluded that  $S_b \ll S_o$  (judging by the small differences between the red and blue curves in Fig 9a), meaning that most of the total covariance is ascribed to  $S_o$ . And these errors - 1 or 2 K - really are large. Do the authors believe these come from the IASI instrument or the RTTOV model? And surely if they are this large someone else would have noticed by now? Is there any other evidence to support this magnitude of error?

However whether Fig 9(b) really represents  $S_o (=B)$  depends crucially on whether  $S_b$  has been correctly evaluated, and I suspect it has been underestimated. It is not clear from the paper whether the IASI ozone channel observations have been assimilated into MOCAGE in the first place. If not that might explain why MOCAGE produces an artificially uniform ozone field which would be misinterpreted (in the NMC method) as a low background error covariance. Even if ozone channels are assimilated, the  $B$  matrix

C2

seems to be estimated as a global mean (again, not clear from the paper) rather than specifically for each site or for the mean of the set of sites.

The complexities of establishing the observational error covariance matrix aside, it is not demonstrated that the channels (or results) obtained with such as matrix represent any improvement on the channels obtained with a diagonal covariance assumption. This, and the irregular behaviour of the RED as a function of the number of channels used (Fig 11) both seem symptomatic of a misrepresentation of the R matrix, presumably originating from the estimate of the B matrix. Although the authors argue that their selection is also designed to improve temperature and humidity, I am unconvinced that it represents any improvement in channel selection. The improvements in T,q seem small and may even be achievable with a random selection of additional channels in this spectral region.

The authors confine themselves to considering only additional O3 channels. A more robust test of their method would be a complete channel reselection and a demonstration that this does indeed lead to improved results for T,q as well as ozone.

So, as it stands, while the proposed method is plausible, I am not convinced that it has been correctly applied (or indeed if it can be) and neither am I convinced that the additional complexity of this method results in any objective benefit.

#### Minor Comments

I found it difficult to keep track of all the different sets of channels referred to in various parts of the paper. Perhaps a summary table would help?

I find the text reads better if references which are to be read as part of the text are presented, for example, as "... was performed by Collard (2007) ..." rather than put the complete citation in brackets "... was performed by (Collard, 2007)". But that's just my personal preference.

P1 L16: quoting these percentages without explanation of context is a bit misleading.

C3

For example it might be understood that the humidity error has been reduced from, say, 50% uncertainty (or whatever it was) down to 30.1% uncertainty - which is certainly not the case.

P2 L6: 'uses 75% of observations from infrared sounders'. Does this mean, of all the observations made by infrared sounders, it uses 75% of these. Or does it mean that 75% of the observations used come from infrared sounders? Also, does an observation count as a single IASI pixel or a single IASI channel?

P2 L10: surplus comma after 'last one'.

P2 L11-12: I suggest rearranging to make it clear that the 0.25cm<sup>-1</sup> sampling is what leads to 8461 measurements rather than the 0.5cm<sup>-1</sup> resolution. It's also probably worth mentioning at this point that the reduction from 0.25cm<sup>-1</sup> to 0.5cm<sup>-1</sup> is largely due to the Gaussian apodisation routinely applied to the spectra as part of the processing rather than an inherent property of the interferometer itself.

P2 L18: If you take 300+14 channels from 8461, I make that 3.7%. Perhaps you are actually referring to the 123 channels mention in the Fig 2 caption, but that is not clear at this point.

P2 L22: observationS

P2 L25: I suggest '... because ozone-sensitive channels ...'

P2 L27: '... a realistic ozone information ...' ? Do you mean a more realistic ozone representation in the assimilation model?

P3 L4: One point which should perhaps be mentioned is that although Collard selected channels assuming a diagonal observation-error covariance, he also imposed a requirement that adjacent channels are not selected. This was specifically to avoid the noise correlation between adjacent channels that is introduced by the apodisation.

P3 L10: ' ... which assimilateS ...'. And presumably IASI L1C data rather than L1B?

C4

P3 L10: '... carry out AN ozone-sensitive ...'

P3 L19: RTTOV is a notorious example of a 3rd-order acronym, probably best left unexpanded, especially since the TOVS part is now largely historical (and TIROS even more so). However, you should provide a reference at this point.

P3 L24: later you say you discard the inter-variable correlations in favour of a univariate B matrix (P13 L1)?

P4 L2: I suggest 'radiosonde launch sites' However, are these all actually ozonesondes rather than radiosondes (or both?). Otherwise how else do you get your ozone profiles for later?

P4 Fig1: The caption is confusing. The map \*only\* shows the radiosonde launch sites from the WOUDC network. The mention of 345 profiles selected from these sites is better left in the main text.

P4 L15: I assume you are referring to the cloud fraction reported in the IASI L1C spectra (rather than, say the Eumetsat L2 product). But how do you know that the matched radiosonde measurements are not taken in cloud?

P6 L21: '... multiplied by 10% ... etc' - I don't understand what this means or why you have done it.

P6 L20: It is not clear what units are used for humidity and ozone. This will affect the definition (and shape) of the Jacobian.

P7 Fig4: I'm not convinced that this figure is helpful since the magnitude of the Jacobian also depends on the thickness of the model layers and the use of channel index as the x-axis is confusing. What do the two vertical lines indicate?

I suggest it would be more informative to have a plot similar to Fig 2 but showing the pressure at which transmittance to the top of the atmosphere reaches 1/e since the accompanying text largely discusses the altitude from which the information comes (this

C5

would be the same for temperature or composition). For the window channels, where the total transmittance is always greater than 1/e, the appropriate pressure would be a weighted average of the atmospheric and surface contributions.

P8 L13: '... high on mid tropospheric ...' doesn't make sense.

P8 L14: 'where there is on average the most humidity' ? Is this relative humidity or H<sub>2</sub>O mixing ratio. In any case I would expect Jacobians for any species (not just H<sub>2</sub>O) to be most sensitive in the mid-troposphere simply because this is where the combination of the product of the temperature contrast (against the earth surface background) and number of molecules of absorber reaches a maximum. Lower down the number of molecules is larger but the temperature contrast vanishes, so you see nothing. And higher up the temperature contrast is larger but the number of molecules becomes vanishingly small, so you see nothing.

P8 Fig 4: Since this is the first reference to IASI bands, the caption should at least say where Band 1 and Band 2 lie on this plot.

P8 L16: There should be some mention of which molecules have been included in the RTTOV calculation. I believe that RTTOV lumps a number of these together as well-mixed gases, so the concentrations are presumably constrained to some fixed value (appropriate for a particular year, if it includes CO<sub>2</sub> and CFCs?).

P9 L15: Without knowing the details of the selection of the lines from 645-770 used for temperature content, I would assume that these would have been screened to \*exclude\* any ozone-sensitive lines and just restricted to CO<sub>2</sub> lines. In that case I would expect the ozone-sensitive lines in this spectral range to be distinct from the temperature-sensitive lines.

P9 L6: From a brief glance, the method of Gambacorta and Barnett seems to be a very basic threshold test of channels which are most sensitive to ozone perturbations, without any consideration of interdependence or redundancy. Is that what you've done

C6

here? If so, why such a crude method compared with the other possibilities such as the Rodgers method?

P9 L22: '... due to the high concentration of tropospheric ozone' ? Firstly, why would tropospheric ozone be high - normally it is low. But how do you know these temperature Jacobians have anything to do with tropospheric ozone at all? It seems more likely that they just represent sensitivity to temperature perturbations in tropospheric CO<sub>2</sub> or H<sub>2</sub>O visible through optically thin stratospheric ozone. More likely H<sub>2</sub>O given the similar structure to the H<sub>2</sub>O Jacobians.

P10 L2: National Meteorological Center (singular).

P10 L3: rangeS

P13 Fig 8: Given the inversion of the y-axis to have level 0 at the top, it seems more natural to similarly invert the x-axis so the main diagonal extends from bottom left to top right.

P13 L1: I note the restriction to univariate B matrices but, even so, it would have been interesting to see the full correlation matrix for T, q and O<sub>3</sub>.

P13 L8: Bormann et al discuss only microwave instruments, which are very different to IASI. They make no comment on the applicability to hyperspectral infrared sounders.

P13 L10: Ventress and Dudhia constructed their R matrix using a 'bottom up' approach of estimating separate sources of forward model uncertainty, as opposed to the 'top down' approach used here.

P13 L12: Perhaps I have misunderstood, but the SD represented in Fig 4 is surely a combination of observation error (by which I mean instrument noise and RTTOV modelling error) \*and\* background error represented by the failure of MOCAGE to represent the real atmosphere?

P14 Fig 9: It would be helpful to have a reference in the figure caption on how to relate

C7

the 428 channels to the actual IASI spectrum.

P15 L1: I can't reconcile this form of R with the version quoted in Bormann et al (which makes more sense to me).

P17 L2: so on (rather than 'soon').

P17 L30: The RED considers only the SD. However wouldn't you also expect the bias to be reduced by adding more ozone channels?

P18 Fig 11 caption: I suggest 'red bar' etc rather than 'red line' when referring to the histogram.

P19 L11: The non-monotonic evolution could also be a symptom of the observation error covariance being misrepresented.

P21 Fig13: Presumably in Panel (c) the black and red lines are almost superimposed?

P21 L3: ratio (rather than ration).

---

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-242, 2019.

C8