Response to Anonymous Referee #2 (RC2)

We would like to thank Referee #2 for his/her careful and balanced comments on our work and for their support for publication of this work. We reproduce quotes from Referee #2's report (in italics), and respond.

- Before the specific comments: *Whether it's actually practical is not clear, and the demo seems like an odd choice since it is so poorly observationally constrained.*

  - There are two ways that the practicality of a method can be defined. One is whether it is appropriate/applicable, and the other is whether it is efficient for use. We believe that the method set out in our paper is practical under both these definitions.
  On the first definition, the spectral method is associated with homogeneous correlations. While this property is not obeyed strictly in geophysical systems, it is a pragmatic assumption, which has traditionally been used successfully in many geophysical data assimilation problems. It is a first stab at the spectral method in a flux inversion problem. The paper does point out that the homogeneous property can be relaxed by using two spectral fields and masks, specifically to decouple land and sea points. Many other works in flux inversion assume either homogeneous and isotropic correlation functions (apart from avoiding land/sea coupling) for the flux, or assume that flux errors are uncorrelated, which are all approximations.
  On the second definition, the spectral method is extremely efficient (and would remain so using the two spectral fields to decouple land and sea). This is a very clear conclusion of the work.

  - The experiments shown, as the reviewer states, are just demos. The aim is to assimilate only in-situ data to avoid complications of satellite data. Although the available observations are relatively sparse, they are spread around the globe and we do believe are adequate to help validate the new method.

- Point 1: *Lines 38-40 ... It is not really fair to say that it is limited 'to a relatively small number of large-area surface regions' because the Jacobian matrix can be computed as an embarrassingly parallel problem ...*

  - Many thanks. We will include the reviewer's comments in a requested revision.

- Point 2: *Line 42 should also mention the LETKF approach used by Myazaki at JPL, ..., which has similarity to the transform done here.*

  - We will include this reference in a revision, but we don't believe that the LETKF data assimilation method uses the transform described in our paper.

- Point 3: *Lines 43-46 There should be some mention that the size of the state vector is limited not only by computational resources but also by information content. A problem with 4DVAR and EnKF methods is that information content is not directly characterized.*

  - Thank you for these interesting comments. We don't agree that the information content of the measurements should *necessarily* be a limiting factor in the size of the state vector, especially in the case of a non-diagonal B-matrix (see https://doi.org/10.5194/acp-16-14371-2016). Even if there are n elements to the state vector, and only $m \ll n$ observations (or even fewer degrees of freedom of the signal), a large state vector may be necessary, for instance, to provide enough resolution for the model to behave realistically. Information content may be characterised by the reduction in the error covariance from the prior to the posterior. This information is difficult to get from 4DVar (although not impossible), but easier to get from EnKF methods. The degrees of freedom of the signal is another measure of the information content. This is simply twice the value of the background term evaluated at the analysis, so is available immediately from 4DVar.

- Point 4: *Line 60 another problem in empirical construction of B is ensuring that it is positive definite – this typically requires massaging the matrix after empirical construction of the off-diagonals.*

– Indeed. A ridge regression (to ensure positive-definiteness) was necessary only for the error covariances of the prior methane initial conditions in the vertical direction. This is only a very minor part of this work anyway. The remaining covariance components are guaranteed to be positive-definite (see also point 6 below).

- Point 5: *Line 131 why does the change in grid-box size with latitude matter?*

  – For a diagonal B-matrix, the effective correlation length-scales are not zero, but of the order of the size of each grid box (and with structure functions that are steps). Hence as the grid box sizes of the TOMCAT model change with latitude, a diagonal B-matrix naturally imposes longer length-scales, which may not be realistic. This was only mentioned in passing in the manuscript, and can be explained in a revision.

- Point 6: *Line 152 will Bsp always be positive definite?*

  – Yes, we believe so, as long as the $\mathbf{\Lambda}$ and $\mathbf{\Sigma}$ (all diagonal) matrices in Eq. (7) are all positive, which they are. Example $\mathbf{\Lambda}_{\mathrm{h}c}$ and $\mathbf{\Lambda}_{\mathrm{h}\rho}$ spectra are shown in Fig. 2b.

- Point 7: *Lines 169-172 Spectral methods and spherical harmonics don't work great for atmospheric chemistry problems because the variability of chemical species does not follow wave structures . . .*

  – The spherical harmonics are just a linear basis to represent the fields, so can in principle represent any form that the grid can represent. It is not true therefore that wave structures are the only kinds of structures that can be represented. The concept that is exploited in the work is that the a-priori errors are uncorrelated between different wavenumbers. This naturally translates to homogeneous correlation functions in real space (and isotropic if the error variances are assumed a function of total wavenumber only as is done here).

- Point 8: *Line 218 'weak sink over Antarctica'. Where does this sink come from?*

  – This appears to be an artifact of the plotting code; in reality there are no a-priori methane fluxes over Antarctica. Apologies for this. We have found the problem and will produce new figures in a revision. The fix is unlikely to affect other plots, but we will check this.

- Point 9: *Line 226 errors in atmospheric chemistry problems generally do not show 'homogeneous and isotropic correlations'.*

  – This is the case in many geophysical systems, where data assimilation assumes homogeneous and isotropic correlation shapes (e.g. in many of the variational systems currently used for flux inversion, apart from at land-sea boundaries). It is an assumption that follows from the spectral method used in the paper, but as stated in first bullet point above, the land and sea points can be decoupled by using two spectral control variables and masks. It is also true to say, for instance, that atmospheric chemistry problems do not obey Gaussian statistics, but this does not make the Gaussian assumption inapplicable or useful (most atmospheric inversion systems assume Gaussian statistics). We believe the work in the paper is a reasonable first piece of work to test the spectral method for the problem of flux inversion.

- Point 10: *Fig. 2 The correlation length scale looks more like ten degrees, so ~1000 km. Seems long, particularly for methane fluxes which come from a diversity of uncorrelated source sectors. What is the rationale for a spatial error correlation in methane fluxes?*

  – A SOAR (Second Order Auto Regressive) function is used for the correlation functions in this work and are shown for length-scales of 400 km and 600 km in Fig. 2a. For a given length-scale parameter, ℓ, the SOAR may be compared to exponential and Gaussian correlations functions – SOAR: $(1 + |x/\ell|)\exp(-|x/\ell|)$, exponential: $\exp(-|x/\ell|)$, Gaussian: $\exp(-(x/\ell)^2/2)$. These are plotted in Fig. 1 below. For a given ℓ the SOAR 'looks' longer scale than the other two forms. This explains why the curves in Fig. 2a appear longer than the length-scales specified.

  – The reasons for adopting non-zero length-scales are many. (i) Even if emissions are physically uncorrelated, the *errors* between neighbouring locations can still be correlated (the correlation used in data assimilation rightly does not need to be associated with causation). (ii) A non-zero regularisation length-scale can be advantageous to produce a sufficiently smooth analysis. (iii) In the tests done in this paper, the results were consistently better when correlations were imposed.

- Point 11: *Figure 3 weird to have the vertical error correlation plot emphasize the stratosphere in a demo for methane fluxes.*
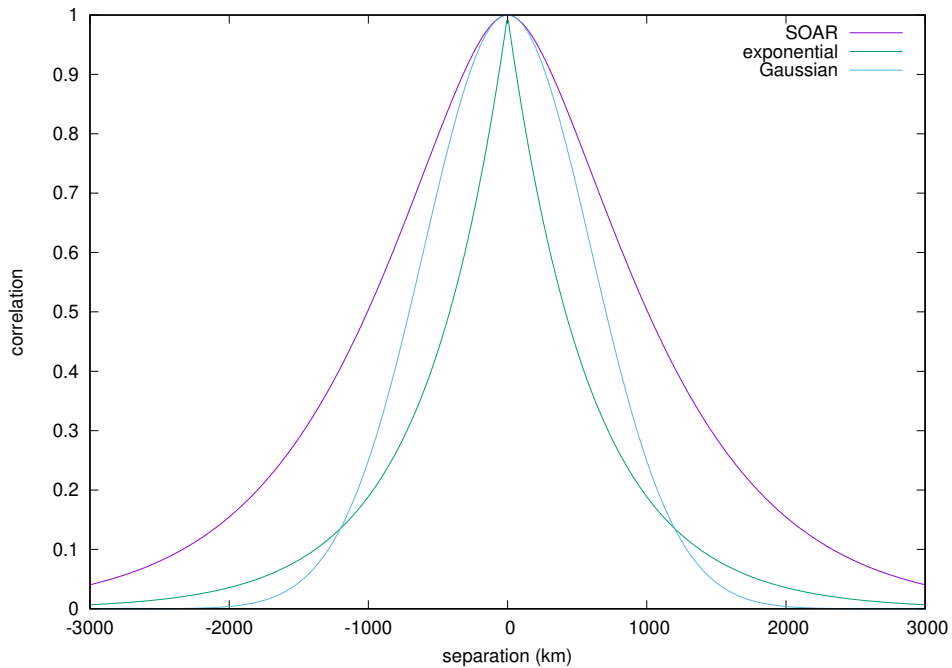
Fig. 1: Comparison of SOAR (Second Order Auto Regressive), exponential, and Gaussian correlation functions, all with $\ell = 600$ km.

– The vertical correlations in Fig. 3 are not for methane fluxes, but for the initial conditions of methane, which includes a representation in the stratosphere. Hence we think it is fair to show the stratospheric part.

• Point 12: *Line 284* (we assume the reviewer means Line 384) *The small information content to be obtained from the 60 NOAA stations is incommensurate to the size of the state vector. That would explain why the results are disappointing.*

– We don't regard the results as disappointing. As stated above, although there are fewer observations than used in a study using large amounts of satellite data, the data are still useful and effective for our purposes.