General comments on corrections done

We would like to thank all five reviewers for their valuable input. A major revision of the structure of the document has been done as suggested by most. The technical details related to the code were moved to the Appendix.
Also, their remarks lead to the rephrasing of three paragraphs and the discussion part in the conclusions was extended. More details have been provided on the setup of the experiments with different modeling of B. Some plots and captions have been corrected and completed. The first part of this document gives an overview of the modifications done, followed by the answers to the comments of each reviewer.

(1) Modifications of the structure of the document.

Reviewers asked modifications of the structure of the document:

In Sect. 2.0, Sect. 2.2.3 doesn't exist anymore.

As asked by different reviewers, the technical details of Sect 3. has been moved in Appendices. Thus, Sect 3. is renamed "Five stages to generate the background error covariance statistics (GEN_BE code version 2.0). and subsections from 3.1.1 to 3.1.4 renumbered from 3.1 to 3.4.
The previous Sect. 3.2 does not exist anymore:

- Sect. 3.2.1 have been included in the Appendix A (FORTRAN code and input/output description)
- The first part of the Sect. 3.2.2 has been merged to the new Sect 3.2.
- Section 3.2.3 becomes Appendix C (Installation, compilation, set up and visualization).
- The description of the namelist options goes in Appendix B (Description of the namelist options)

Section 5.0 includes now the results related to chemistry data assimilation previously shown in Appendix A. Sect. 5.0 is renamed "Cloud and chemistry variational data assimilation"

- Sect. 5.1 is named "Generation of a multivariate background error covariance for hydrometeors.
- Sect. 5.1.1 is added and is composed by the part related to the balance operator previously presented in Sect. 3.2.2. Section 5.1.1 is named "Generation of a multivariate background error covariance for hydrometeors.
- Previous Sect 5.1, and 5.2 becomes 5.1.2 and 5.1.3
- Previous Appendix B becomes Sect. 5.2 and is named "Background Error for Chemical Species"

(2) Modification Equations

Some Equations has been corrected, added and renumbered. We give an update below of the different modifications done.

- Eq. (1) J_b and J_o terms are added
- Eq. (2) new equation added to present a general definition of B
- Eq. (3) the definition of $\delta x=(x_b-x)$ added and renumbered
- Eq. (4) renumbered
- Eq. (5) $B^{1/2}$ is presented instead of δx
- Eq. (6) new equation to present the calculation of the regression coefficient.
- Eq. (7) presents of the calculation of the unbalanced part of the perturbations δt_u
- Eq. (8a) presentation of the Daley's formula that define the vertical length scale for one dimension along the vertical (z).
- Eq. (8b) presentation of an approximation of the formula of Daley along the vertical
- Eq. (9a) presentation of the Gaussian formula that define the vertical length scale for one dimension along the vertical (z).
- Eq. (9b) inverted expression of 9(a)
- Eq. (10a) corrected and renumbered
- Eq. (10b) corrected and renumbered
- Eq. (11) corrected and renumbered
- Eq. (12a-c) Identical

(3) Modification Figures

Previous Fig. 14, that shows the distribution of the vertical model level in function of pressure level, is presented earlier in the document (in the first paragraph of section 3.0 and becomes Fig. 3).
It allows visualizing the density of the vertical model in function of pressure and switch from vertical model level to pressure accurately when results are presented in sect 3.0, 4.0 and 5.1.

Fig. 9, 10,11,12,13,15,16,18a added right vertical axis in hPa pressure levels.

(4) Modification Tables

Table are renumbered:
Table 4 becomes Table 1
Table 2 is created to gather the setup information about the different modeling of B.
The other Tables are moved into the appendix:
- Previous Tables B1, B2, and B3 become Tables A1, A2 and A3.

- Previous Tables 1, 2, 3, 6, 7 and 5 become respectively Tables B1, B2, B3, B4, B5, and B6.

(5) Major revision in the text

Description of the experiments:

(a) The description of the D-ensemble dataset (50 members over the CONUS domain) coming from DART is done in the second paragraph of Sect. 3. : "Figures shown in … Romine et al. (2014) to generate the ensemble and … Table contains detailed information setup of the data assimilation experiment."

Reference about DC3 experiment of Romine et al. 2012 is replaced by: Romine G., S., Schwartz C., S., Berner J., Fossell, R., K., Snyder C., Anderson J. and Weisman M., L.: Representing forecast error in a convection-permitting ensemble system, Mon. Weather Rev., doi: http://dx.doi.org/10.1175/MWR-D-14-00100.1, 2014.

(b) A new table 2 is presented Section 4.0, to give details about the benchmark performed.

Table 2: Description of the setup of the background error matrix modeling diagnosed over the CONUS Domain. $\mathbf{B}_{eof}$ and $\mathbf{B}_{rcf}$ are diagnosed using GEN_BE code version 2.0 and the D-Ensemble method while $\mathbf{B}_{nam}$ is performed by NCEP using the NMC method.

Paragraphs rephrased:

(a) In the introduction, the first paragraph has been corrected, the second and the third rephrased following the remarks of the different reviewers.
(b) Section 2.2.2, the order of the description of the different transform match the Eq. 5:
- The $\mathbf{U}_p$ matrix, called physical transform or balance operator, …
- The $\mathbf{S}$ matrix is …
- The $\mathbf{U}_v$ matrix, called vertical transform, …
- The $\mathbf{U}_h$ matrix, called horizontal transform, …

(c) First paragraph of Section 3.0 has been rephrased.
(d) Section 3.2 has been rephrased (merge of previous sections).
(e) First paragraph of Sect 4.0 is rephrased and additional information is given to the general setup of the different modeling of B ($\mathbf{B}_{eof}$, $\mathbf{B}_{rcf}$ and $\mathbf{B}_{nam}$). References have been added: Romine et al. 2014, Rogers et al. 2009 and Wu 2005.
(f) Section 4.2 has been rephrased
(g) Section 5.1.1 coming from the previous Sect. 3.2.2 is partially rephrased to become independent.

(h) The discussion has been extended in Section 6, which is partially rephrased.

(6) Direct answers are given on the different referee below, in the following document.

Corrections Referee 5

Comments on “Generalized Background Error covariance matrix model (GEN_BE v2.0)” by G. Descombes, T. Auligné, F. Vandenberghe, and D. M. Barker
The paper ‘Generalized Background Error covariance model (GEN_BE v2.0)’ presents a tool for the diagnosis of the background error covariance matrix for meteorological and atmospheric chemistry data assimilation applications. The code is based on existing techniques and does not present novel algorithms. However, GEN_BE v2.0 is of potential interest for many researchers in the field of geophysical data assimilation and the presentation is supported by several examples of scientific interest.
The paper lacks of scientific rigour in some sections, the structure is not optimal and it contains multiple language mistakes or approximations. Therefore, I recommend a major revision prior to publication in GMD. The main comments are detailed below.

General comments:
Introduction
no particular emphasis on the scientific aspects that are examined later in the paper (e.g. the analysis of meteorological and chemistry error covariances). These applications are listed in the content of sections, with lack of important details, like the ensemble specifications, or too much detail, like the specification of the CV5 set of variables or the CONUS domain. I suggest to the authors to better introduce the scientific framework of the examined cases (e.g. multivariate meteorological analyses), with corresponding references, then introduce the numerical experiments. The reader should understand why those experiments are done at the introduction level. Details about the single experiences (e.g. the geographical domain, the ensemble...) could be given later in the corresponding sections.

Section 3
Section 3 describes the details of the employed algorithms, the code utilization and presents some results from the numerical experiments (mostly error correlation plots). This makes a very long section, difficult to be read. I suggest the authors to remove all the technical details like names of FORTRAN variables and routines from the text. Some sub-sections could also be removed (e.g. 3.2.1 and 3.2.3). The code instructions should be moved in an appendix and reference the main text when needed. Second, I suggest to move the discussion of the correlation plots (Fig 3,4,5) to section 4, adding a detailed description of the model configuration used to calculate the ensemble statistics, which was missing

in Section 3. In this way the reader can find the complete discussion of the numerical experiments in the same section. Moreover, the analysis of error correlations will be directly followed by the length scale/EOF analysis.Finally, section 3.2.2 could be merged with section 3.1.2, since they are strictly related.

Section 4
Please avoid switching frequently from grid point to km when discussing the length scales (e.g. page 4309, lines 20-22, page 4313, lines 7-9). Physical units like km for horizontal distances or hPa for vertical distances are preferable. Otherwise put always grid points and corresponding physical values in brackets. All plots should provide axes in physical units as well (Figure 4-5-6-9-11-12-13-15-16-18). Figure 14 would not be necessary anymore.

Answers to the general comments

We want to thank Referee 5 for the numerous remarks that lead to major revisions in the structure and presentation of the document. Additional information has been given to improve the presentation and discussion of the section about chemistry data assimilation.

- ***Introduction***
  Several part of the introduction has been rephrased to introduce why we do a focus on cloud and chemistry data assimilation.

- ***Section 3.0***
  We follow most of the recommendation to change the structure of the document: technical information are moved in three appendices and some paragraph have been merged. (see the structure presented at the beginning of the document)

- ***Section 4.0***
  Discussion on length scale are done preferably done in their physical units. Additional information to switch easily from grid units to physical units for the Figure 8-9-11-12-13-15-16-18.

Detailed comments
1) Page 4292, lines 5-10: This sentence is too long and does not clarify what the GEN_BE does. From the title and the previous lines (3-6) the reader expects a generic or generalized code conceived to model background error covariances

for data assimilation applications. Here GEN is used for GENerate, which is indeed the purpose of the presented code e.g. generate B parameterizations for further use in data assimilation systems (like WRFDA and GSI). The abstract should clearly state this and the authors should decide between 'generate' and 'generalized'.

The title "Generalized Background Error Covariance Matrix Model (GEN_BE v2.0)" refers the ability of the new code version 2.0. The first version was designed mainly to handle variables and linear regression coefficient hard coded in the code. As explained in the abstract and the introduction, the new framework allows to handle different control variables defines and cross-correlated errors defined as an input. Also, implementation of new models should be straightforward. Finally, the code version 2.0 gathers different transform such Uv defined by EOF or with recursive filter, variance 3D (S).

Also, the sentence have been spit: " …Forecasting (WRF) community model. GEN_BE allows for a simpler, flexible, robust, and community-oriented framework that gathers methods used by some meteorological …"

2) Page 4292, lines13:'...performing benchmarks...', please precise what kind of benchmark you considered in the study (e.g. multivariate meteorological analyses) before introducing the hydrometeors and atmospheric chemistry applications.

The benchmark involves different modeling of **B**.*The sentence has been completed: " … by performing benchmarks of different modeling of **B** and showing …"*

Additional modifications are done:
L17: "a tool flexible enough to involve" replaced by "a tool flexible enough to implement"

3) Page 4292, line20:'...chosen as a testbed for diagnostic and new modelling of B' Do you mean that GEN_BE can be used to verify the results of similar codes? Or that new variables and error covariances can be implemented and tested easily? Please clarify or remove.

The sentence has been replaced by: "L20:" replaced by " … (GEN_BE v2.0) can be easily applied to other domains of science and be chosen to diagnose and to model B."
.

4) Page 4292, lines 25-26 and page 4293 lines 1-5: I find this affirmation too strong, the performances of data assimilation can be improved also by considering more advanced assimilation algorithms or by improving observation error estimations. Moreover, I can't see the logical link with the end of the sentence '...assuming that the underlying probability errors are normally distributed'.

The first sentence has been modified and split: "Since the best estimate of the background error covariances matrix (B) is a key component for data assimilation improvements, various … within a variational framework"
The second sentence is: "The probability errors are supposed to be normally distributed and B is determined for a limited set of variables, called control variables"

5) Page 4293, lines 5-7: '...are usually...' Please either add a reference or explain the reason of choosing variables with uncorrelated errors.

The reason is that we want B block diagonal after Up to be able to model a full B matrix. When the control variables are uncorrelated, there is no need to model their cross-correlated errors. Otherwise, they will be model by linear regressions in our case. This part has been moved into section 2.2.2, in paragraph of Up, to explain more in details.

6) Page 4293, line17-18-19:'MM5, NCAR, WRF', Please add the full name of every model or institute the first time they appear in the text, and possibly a reference in case of a model (e.g. for WRFDA).

We added the description and the references.

7) Page 4293, line26: '...unite them'. Clarify what should be unified.

It has been rephrased:
"This new flexibility associated with the possibility to define a set of control variables and their covariance errors as an input should reduce future developments of the code considerably and should benefit to a larger community in geophysical science."

8) Page 4294, line5: '...using different transforms...'.
The concept of transform was not introduced before, which makes the sentence obscure for the reader.

It has been defined now in the sentence. "Section 2.0 presents the role of the background error covariance and how a series of different operators (i.e. balance, vertical and horizontal transforms) can model B."

9) Page 4295, line7: the errors are supposed uncorrelated, not the observations themselves.

We replaced "uncorrelated observations" by "uncorrelated observation errors".

10) Page 4295, line 17: please specify that x=($\mathbf{x}_b$-$\mathbf{x}$)

We replaced $\delta x = B^{1/2}u$ *by* $\delta x = (x_b - \mathrm{x}) = \mathbf{B}^{1/2}u$.

11) Page 4295, line 24: you could probably mention that the rewritten cost function in Eq. 3 is quadratic, which allows a global minimization

We replaced "H is the tangent linear operator" by "H is the linearized observation operator which makes the cost function quadratic and easier to minimize."

12) Page 4296, line 10-14: please define what does it mean balanced and unbalanced before, or add a reference.

The paragraph describing the Up matrix is rephrased and reference to section 3.2 is added.

13) Page 4296, line 17: please clarify how horizontal diffusion is used in the framework of B modelling or remove it. The reader is anyhow addressed to other studies on the subject of covariance modelling few lines later.

We removed "which are affordable approximations of horizontal diffusion."

14) Page 4297: Section 2.2.3 seems more as part of the introduction or should be reduced and merged with 2.2.2.

The section 2.2.3 doesn't exist anymore.

15) Page 4297, lines 23-27: This was already said at line 2-3 and in the introduction. Please consider removing it.

The section 2.2.3 doesn't exist anymore.

16) Page 4298, line 4: define 'raw model perturbations of the analysis variables'. Do 'analysis variables' correspond to the 'control variables'?

We replaced "raw model perturbations of the analysis variables" by "model perturbations of the control variables."

17) Page 4299, lines 8-16: The explication of the reasons to perform spatial averaging, or 'binning', are not clear. I don't see how spatial averaging can 'increase the number of samples' or 'reduce the dimensional of statistical output parameter' or 'add heterogeneity and anisotropy in B'. I suppose that the authors want to say that, since the number of samples of the ensemble is limited, a strategy to filter the sampling noise is needed. The paragraph should be rephrased with the aid of some of the numerous references that exists in term of ensemble filtering.

Answer: In variational methods, B needs to be estimated for the entire domain. Since it is not possible to compute a full rank B matrix, different hypothesis are taken to filter the sampling noise coming from a limited number of perturbations and to reduce its dimensions in order to model a static error covariances. The focus of this paragraph is to have a background error statistics (coming from an ensemble or a NMC method) model for the entire domain: Binning is also a way to model a B matrix for specifics needs, to filter the statistics, and reduce its dimensions (Regression coefficients computed by grid point Reg_coeff(i,j,k1,k2) becomes Reg_coeff(b,k1,k2) where b in the bin class of a grid point at the location (i,j))

Correction: the paragraph is rephrased
"Since the number … characterize convection events"

18) Page 4299, line 20. Please add a reference about the resulting skewness of hydrometeors statistics.

We added the reference "can be skewed (Michel et al. 2011)".

19) Page 4300, line 8. What do you mean by 'estimation error'?

The corrected sentence is "Analysis increment for one variable may impact an another if they have correlated errors."

20) Page 4300, lines 11-15. Either give a reference to the NCEP method or write more clearly the steps that lead to the calculation of the regression coefficients.
In section 3.2, Eq. (5) is added and the appendix B gives some details about the calculation (no references found).

Similarly for lines 16-19. Are linear regressions calculated on perturbations or variables themselves?

Linear regressions are applied to the perturbations to compute the statistics on

uncorrelated control variables. They are also applied later in the data assimilation process on the variables themselves.

Is Up block diagonal or Uh and Uv? Please clarify.
Up is block diagonal.

21) Page 4300, line 20. Stage 2 has changed with respect to GEN_BE v1.0? Is it necessary to be written?
It has been removed.

22) Page 4301, line 20. L should be squared, x should be $\delta\mathbf{x}$ and the equation seems not numbered.
Corrected to δz and numbered.

23) Page 4302, line 2. The correct equation seems 5 or the one which is not numbered.
The equation has been numbered since.

24) Page 4302, line 8-9. What does it means 'by bin'? Do you mean, without spatial averaging? And why it is not useful for data assimilation? Please clarify.

We removed this sentence and rephrase this paragraph:
Pannekoucke et al. (2008) studied … the horizontal length scale for stage 4.

25) Page 4302, line 9-11. Which regression coefficient? Does it mean that the binning can be decided independently at each stage? Please clarify

For example, the regression coefficients can be binned and the vertical length scale computed uniform by vertical level. This is the case for the Bnam defined for regional applications in GSI provided by NCEP.

26) Page 4302, lines 19-24. Quantify larger, smaller and local in term of kilometres.

This has been included in the text:
"The stream function (3a) and velocity potential control variables have larger and more isotropic spatial correlations while the temperature (3b) and the humidity (3c) control variables show smaller and anisotropic correlations at different locations. The radius of the area where the correlation overpasses 0.9 is within a range of 100 km to 400 km for stream function while this radius reaches its maximum around 100 km for temperature and humidity. Hydrometeors mixing ratio show even more local structures due to their sparse location on the

horizontal and the vertical (3d)."

27) Page 4303, line 5. I could not find the explanation in Sect. 3.1.2
Wrong reference, it should be Sect. 3.3 where we introduced the decomposition by EOF modes.
We replaced "by EOF mode or by vertical level as explained in Sect. 3.1.2 by "by vertical level or by EOF mode as explained in Sect 3.3"

28) Page 4303, lines 6-7: Is the solution calculated considering the nearest grid points?

A radius r0 can be defined to consider only the points which are distant of a distance r inferior to r0. Moreover as it is mentioned in the same section, the use of the second formula (ls_method=2,Wu et al. (2002)) is advised.

29) Page 4303, line 11: what is it meant by 'pseudo correlation'?
We removed pseudo.

30) Page 4303, line 20-21. What does it mean 'at best it can be statistically binned'? Moreover, horizontal length scales for a given vertical level are 'usually' not uniform, as also shown in the example in Figure 3. Please clarify.

We agree, length scale can be computed uniform or binned (which include diagnosed by grid point). Moreover, in practice, operational centers such as NCEP, used statistics averaged by vertical level and binned for some of them. There are potentially some issues to handle heterogeneous length scales with recursive filters as mentioned in this rephrased paragraph: "The horizontal length scale can be … be required because of recursive normalization issues (Michel and Auligne 2010). "

31) Page 4304, line 4. Please add a reference about the poor results of recursive filters.
See point (30)

32) Page 4305, line 9. What does 'Generalized' stands for in the section title? As suggested in the general comment I would merge this section with the 3.1.2.

This paragraph has been merged to the section mentioned and rephrased. Also, the part related to data assimilation of the multivariate hydrometeors experiment is presented now in the new section 5.1. This merged section has been rephrased.

33) Page 4305, lines 19-21. The sentence is not clear, what kind of benchmark is done? Which are the other series of operators?

The sentence has been replaced by "Benchmark results of pseudo temperature test involving different modeling of B and the same Up transform (CV5) are shown Sect 4."

34) Page 4305, line 22. 'Recent studies' should be referenced.
As mentioned point (32), the paragraph has been rephrased.

Now the studies dedicated to better estimate the background error covariances matrix in cloudy areas are first presented and then discussed.
They are introduced by the sentence "Thus, various studies have been dedicated to better estimate the background error of humidity in cloudy areas (Carron and Fillon 2010, Montmerle et Berre 2010, Ménétrier and Montmerle 2011)."

35) Page 4305, lines 26-28. The statement is not really supported by Figure 4 because, as far as I understood, the statistics are shown for the entire CONUS domain (dry and wet areas). Or does the statement refer only to the cited study?

The statistics are shown Fig. 4 for the entire CONUS domain. The statement refers only to the cited studies. The new paragraph presents first these figures and then discusses about the application of binning.

36) Page 4306, line 1. Please avoid using probably, if the results are suggesting the conclusion that condensation and precipitation process determine the observed statistics clarify it, add a reference otherwise.

The sentence is now:
"At saturation, these statistics likely rely on processes of condensation and precipitation when the released latent heat flux warms the atmosphere (Holm 2002)."

37) Page 4306, line 5. 'They explain that imbalance in precipitating areas'. Please clarify the imbalance between which variables.

This sentence has been completed: "For a winter test-case where stratiform-type precipitation is predominant, they explain that geostrophic imbalance in precipitation areas, can be characterized by the linear balance operator between the stream function and the mass fields (t and ps)"

38) Page 4306, lines 17-18. ‘As the dynamic control variable...do not explain statically the presence of fog’ The authors probably want to say that dynamical variables such as vorticity and divergence do not drive fog formation processes.

The sentence is rephrased: ” Dynamical variables such as vorticity and divergence are not included in the balance humidity operator since they do not drive fog formation processes.”

39) Page 4306, line 22. ‘...dry and humid atmosphere’ . I imagine the authors mean for both a dry and a humid atmosphere. Again, is this statement supported by the Figure 4, and if yes please clarify. Otherwise add a reference.
The rephrased paragraph should clarify it.

The paragraph has been rephrased and the sentence is:
“For example, Fig. 4 shows the cross-correlation between humidity and temperature for all atmosphere conditions (mixing dry and wet conditions).”

40) Page 4306, lines 24-25. Which is the transform used in real time at NCEP? For real time do the authors mean operational analyses?
The same Up transform is used. Kleist et al. (2009) described this transform used in GFS-GDAS system. This sentence has been removed.

41) Page 4308, lines 10-18. As far as I understood a non-cloudy/cloudy mask is used to restrict the statistical sample of perturbations. Which values of cloudiness or other relevant variables are considered to perform this filtering? ‘Such filter may overestimate the vertical correlation around a given vertical level’. Please clarify the reason and which levels are affected by this issue.

As suggested in previous comment, this part has been moved to a new section 5.1.1. The sentence has been replaced by: ”However, we may want to localized this balance around a given vertical model level.”

42) Page 4308, lines 21-26 and page 4309 lines 1-4. In the general comments I suggested to move here the description of the numerical experiences setting. Some additional details should however be given or appropriate references should be provided for a better interpretation of the results. Which is the NCEP real time configuration (e.g. assimilated datasets)? What are the main features of the WRF ensemble (type and magnitude of perturbations, initialization...)? What kind of horizontal and vertical grid do GSI and WRFDA use (degrees, hybrid sigma-pressure levels, resolution)? What does NAM stands for? Is the NCEP real time configuration differing also on the vertical grid? What kind of data is assimilated in the NCEP operational system?

Table 2 has been added section 4.0 to explain in details the different modeling of B. The Acronym NAM has been described.

43) Page 4309, line 22. '...decreases more monotonically' is not a clear statement, unless a degree of 'monotonicity' is defined. Please rephrase.
The new sentence is "Relative humidity length scale remains small, decreasing from approximately 30 km to 15 km as a function of the EOF mode."

44) Page 4310, line 12. '... representing more synoptic events at high altitude' is not scientifically sound. What it is meant by 'more synoptic' and 'high altitude'? Please rephrase.
The end of the sentence has been replaced by "from the bottom to the top of the model as they represent larger scale events."

45) Page 4310, lines 22-24. First define the experiment setting (innovation and observation error values, location of the observation) then describe briefly what do the plots represent (horizontal and vertical slices of the resulting increment).
The full section 4.2 has been rephrased.

46) Page 4311, lines 4-7. The sentence is too long and not very clear. What is the link with the fact that the domain is of limited area? Please rephrase.
The full section 4.2 has been rephrased.

47) Page 4311, line 9. '...show close results'. It is difficult to verify this statement on the plots. Values of contour lines in Fig. 11-12-13 are in very small letters and it seems that the contour ranges are different among the different experiences. The plot range should be uniformed, the physical units for the contour lines added and I might suggest adding a color scale to ease the evaluation of the maximum and minimum values of the increment.
The full section 4.2 has been rephrased.

48) Page 4311, line 12. Can you provide some insights about the observed differences in the horizontal length scale between the EOF and the level by level estimation?
The full section 4.2 has been rephrased.

49) Page 4311, line 15-21. 'More climatological' is not scientifically sound, please rephrase. I also think that a deeper discussion of the differences between the NMC derived B and the ensemble derived B would greatly improve the paper. But this should be probably done when horizontal and vertical length scales are

discussed (currently Sec. 3.1.4 and 4.1.2 currently).
The full section 4.2 has been rephrased.

50) Page 4311, line 24-25. 'The XZ plan follows the isocontour of 0 m s-1 for U' means that the U increment is negligible? Are the 'complex structures' observed for V realistic in term of the modelled balance?
The full section 4.2 has been rephrased.

51) Page 4311, line 28. As noted in points 48-49-50, these differences should be better presented and discussed before affirming that they are well explained.
The full section 4.2 has been rephrased.

52) Page 4312, line 16. Please clarify why recursive filters make the analysis of length scale 'easier'.

The sentence is rephrased: "The vertical and horizontal transforms retained are the recursive filters making the interpretation of the length scale parameter easier as they are directly associated to a vertical model level."

In addition, in section 4.1.1, a sentence has been added first at the end of the first paragraph: "Also, the EOF decomposition allows optionally some filtering as the largest variances (i.e. eigen values) are associated with the first EOFs, the latest EOFs may be not taken into account if they mostly represent vertical noise in the system"
Then a second sentence at the end of the second paragraph:
"As the horizontal length scale is associated to EOF mode and not directly related to a vertical model level, futher discussions on the association of length scale with physiscal event may be difficult."

53) Page 4313, lines 3-14 Figure 16 seems to be identical to figure 15. Please check.
Figures should be different, mistake corrected.

54) Page 4313, line 24. Change Fig 18a with 18b (and b with a at page 4314 line 2). Is the variance profile in Fig. 18 coming from the ensemble?
We changed Fig 18a with 18b. Both figures come from the same D-ensemble.

55) Page 4314, line 3-5. 'The increment is most likely important' is not correct. Please put larger, smaller or significant and quantification in physical units. Are the observed increments over the dry area not

realistic?

The sentence: "The increment is most likely important where the variability of cloud presence exists …" is replaced by 'The increment is most likely greater than 10^{-3} g/kg where the variability of cloud presence exists."

Answer to the question: There is no increment if the background error standard deviation (diag(B)) are equal to zero. This is the reason of the sentence that follows "A minimum value would likely certainly need to be set to retain the possibility of increments in the dry area". If a minimum value is set up, increment will be possible.

56) Page 4314, line 12-13. Is this result specific to the examined case or is it expected in general?
The sentence is removed.

57) Page 4315, line 16. 'similar results with comprehensive differences' is not correct. Please rephrase considering the new elements arising from the discussion in Sec. 4.2
It has been rephrased in the text: "second, .. using Bnam."

58) Page 4315, lines 26-27. This statement is too generic. Non-linearity exists in meteorology as well and it does not hamper data assimilation.

We removed this statement from the text.

59) Page 4316, lines 19-21. The reference to Barré et al. 2013 is not very pertinent to the discussion. Either add a comprehensive list of studies that performed chemical data assimilation or cite only the studies that focused on the modelling of the B matrix (e.g. Massart et. al 2012, Jaumouillé et al 2013, Gaubert et al. 2014). Since this is not a review article the second option should be considered.

We add a short review of the B matrix characterization for atmospheric chemical data assimilation. We mention studies that use different orders of detail for the B matrix modelization: from using static estimated B matrix to hourly length scales variations. We should mention that Gaubert et al. 2014 uses a ensemble Kalman filter technique that implicitly characterize the B matrix, as opposed to variational technique where B has to be specified.

60) Page 4316, lines 23-25. Taking a realistic background error into account

does not depend on the complexity and the accuracy of the chemical models. Consider removing this sentence.

The statement has been removed from the text.

61) Page 4317, line 1 and previous line. Either detail how the aerosol optical depth is used or do not mention it.

The statement has been removed from the text.

62) Page 4317, lines 9-14. Please detail what kind of chemical scheme is used and/or add a reference for WRF-CHEM, MOZART and MEGAN. Provide also some information or reference about the ensemble perturbations (variance, spatial/temporal correlation etc.)

The text has been detailed.

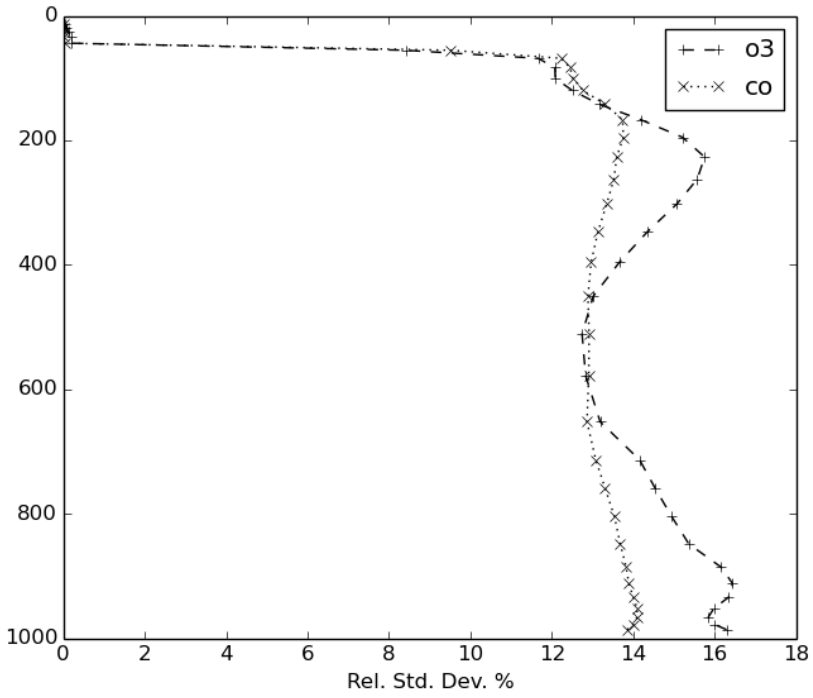
*In the last following comments the reviewer arises important questions on chemical **B** matrix characterization. Since this paper is a general presentation on the GENBV2 system and the chemistry part a proof of concept that the code can be employed on specific cases as chemistry, we do not decide to dig into details this part. We agree with the reviewer that is an interesting topic and needs further diagnosing in a possible following paper. This has been clarified at the end of the section.*

63) Page 4317, lines 14-16. The relative variability should also be displayed in Figure A1, at least for ozone. It would allow to better detect the boundary layer variability of ozone due to the perturbed emissions.

In this section, we chose to put the standard deviations on their original units to show if the calculated standard deviation looked physical and (also because a data assimilation system will deal with those absolute values and not the relative ones).

Enhanced boundary layer values due to perturbed emissions appear close to the source regions (mostly anthropogenic sources over urbanized regions). Where emissions are strong enough, the emission perturbation will produce a standard deviation that is stronger than the standard deviation produced by the lateral boundary conditions. This is most likely the case for relatively long-lived species as ozone and carbon monoxide (couple weeks of life-time). The spread values relative to the ensemble mean averaged along the domain will not necessarily reveal enhanced values on the boundary layer more than absolute spread

values. To convince the reviewer we provide below a plot showing relative standard deviations profiles for ozone and carbon monoxide. Enhanced boundary layer relative variability is not obviously observed for carbon monoxide. Only slight increase for ozone relative variability is observed toward the surface. Because of reason stated above ozone standard deviations is not showing a clear enhancement due to the averaging of different regions (e.g. sea versus land, high emission regions versus remote region, high PBL height versus low PBL height). This point needs further regional detailed investigations that are out of the scope of the paper.



64) Page 4317, lines 22-26. Vertical mixing in the planetary boundary layer is supposed to introduce a vertical error correlation, not to decrease it. Since the vertical mixing decreases above the boundary layer, this is probably the reason of the decrease of the vertical length scale above 850 hPa. On the other hand, surface emissions are generally injected over the first levels of chemical transport models, which might increase the error correlation close to the surface. The authors should verify the way emissions are treated in WRF-CHEM.

We agree with the reviewer with the first part of this comment and we have clarified the text accordingly. However we do not fully explain the strong decrease of vertical correlations close to the surface, since emissions mostly impact the model lowest level (closest to the surface) in WRF-Chem. The text has been clarified accordingly.

65) Page 4318, line 10. Since one of the main content of the paper is the balance between control variables it would have been very interesting to check whether the linear regression approach provides meaningful results applied to interacting chemical species like NOx, CO and O3. Can the authors comment on this?

We agree on the reviewer that the chemical balance between variables is a very important and interesting topic for chemical data assimilation purposes. However as stated above, this section is a proof of concept that the GEN_BE v2.0 code can be directly adapted to chemical variables. Diagnosing the chemical balance would require an extensive study on the B matrix for tropospheric chemistry, which is not the scope of this paper. Chemical balance (on various atmospheric chemical models at different scales) then could be diagnosed by using GEN_BE V2.0 in following studies.

Minor corrections:

1) Page 4293, line 9: change 'dataset observations' to 'observational datasets'
changed

2) Page 4293, lines11-12: do you mean that the availability of more observations involve the control of new model variables? Please rephrase.
Cloud and Chemistry data assimilation may involve new variables such as hydrometeors and chemical species. It has been added the word may " … large set of sensors that may involve more variables, which are .."

3) Page 4293, line 27. Change 'the two first sections' with 'Section 2.1 and 2.2'
It has been rephrased.

4) Page 4295, line 10: 'comprised of' should be 'being comprised of'
done

5) Page 4296, line 3: Change 'decomposed to' in 'decomposed into'
changed "decomposed into"

6) Page 4296, line9: Please add 'foreach grid point'.
It has been added (grid point space)

7) Page 4298, line 15: change '24 minus...' with '24 h minus...'
changed to "e.g. 24 hour minus 12 hour forecasts"

8) Page 4300, line 24: change ‘do not depend of the control variables’ to ‘do not *depend on the particular choice of the control variables’*

9) Page4301,line 5: specify that the length scale is horizontal

10) Page 4303, line 5: change ‘by EOF mode’ with ‘for each EOF mode’

11) Page 4305, line 25: change ‘correlated errors between...’ with ‘error correlation between...’

12) Page 4310, line 3. Change ‘applied by vertical level...’ to ‘applied for each vertical level...’
changed by “applied at every vertical model level for each variables”

Bibliography

Gaubert, B., Coman, a., Foret, G., Meleux, F., Ung, a., Rouil, L., ... Beekmann, M. (2014). Regional scale ozone data assimilation using an ensemble Kalman filter and the CHIMERE chemical transport model. Geoscientific Model Development, 7(1), 283–302. doi:10.5194/gmd-7-283-2014

Jaumouillé, E., Massart, S., Piacentini, A., Cariolle, D., & Peuch, V.-H. (2012). Impact of a time-dependent background error covariance matrix on air quality analysis. Geoscientific Model Development, 5(5), 1075–1090. doi:10.5194/gmd-5-1075-2012

Massart, S., Piacentini, A., & Pannekoucke, O. (2012). Importance of using ensemble estimated background error covariances for the quality of atmospheric ozone analyses. Quarterly Journal of the Royal Meteorological Society, 138(665), 889–905. doi:10.1002/qj.971