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

The introduction clearly states the practical reasons for the development of GEN_BE, but there is 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.

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'.
2) Page 4292, lines 13: '…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.
3) Page 4292, line 20: '…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.
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'. Please rephrase.
5) Page 4293, lines 5-7: '…are usually…' Please either add a reference or explain the reason of choosing variables with uncorrelated errors.
6) Page 4293, line 17-18-19: '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).
7) Page 4293, line 26: '…unite them'. Clarify what should be unified.
8) Page 4294, line 5: '…using different transforms…'. The concept of transform was not introduced before, which makes the sentence obscure for the reader.

9) Page 4295, line 7: the errors are supposed uncorrelated, not the observations themselves.
10) Page 4295, line 17: please specify that $\delta x=(\mathbf{x}_b-\mathbf{x})$
11) Page 4295, line 24: you could probably mention that the rewritten cost function in Eq. 3 is quadratic, which allows a global minimization.
12) Page 4296, line 10-14: please define what does it mean balanced and unbalanced before, or add a reference.
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.
14) Page 4297: Section 2.2.3 seems more as part of the introduction or should be reduced and merged with 2.2.2.
15) Page 4297, lines 23-27: This was already said at line 2-3 and in the introduction. Please consider removing it.
16) Page 4298, line 4: define 'raw model perturbations of the analysis variables'. Do 'analysis variables' correspond to 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.
18) Page 4299, line 20. Please add a reference about the resulting skewness of hydrometeors statistics.
19) Page 4300, line 8. What do you mean by 'estimation error'?
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. Similarly for lines 16-19. Are linear regressions calculated on perturbations or variables themselves? Is $\mathbf{U_p}$ block diagonal or $\mathbf{U_h}$ and $\mathbf{U_v}$? Please clarify.
21) Page 4300, line 20. Stage 2 has changed with respect to GEN_BE v1.0? Is it necessary to be written?
22) Page 4301, line 20. L should be squared, x should be $\delta\mathbf{x}$ and the equation seems not numbered.
23) Page 4302, line 2. The correct equation seems 5 or the one which is not numbered.
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.
25) Page 4302, line 9-11. Which regression coefficient? Does it mean that the binning can be decided independently at each stage? Please clarify
26) Page 4302, lines 19-24. Quantify larger, smaller and local in term of kilometres.
27) Page 4303, line 5. I could not find the explanation in Sect. 3.1.2

28) Page 4303, lines 6-7: Is the solution calculated considering the nearest grid points?
29) Page 4303, line 11: what is it meant by 'pseudo correlation'?
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.
31) Page 4304, line 4. Please add a reference about the poor results of recursive filters.
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.
33) Page 4305, lines 19-21. The sentence is not clear, what kind of benchmark is done? Which are the other series of operators?
34) Page 4305, line 22. 'Recent studies' should be referenced.
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?
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.
37) Page 4306, line 5. 'They explain that imbalance in precipitating areas'. Please clarify the imbalance between which variables.
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. Please rephrase.
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.
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?
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.
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?

43) Page 4309, line 22. '…decreases more monotonically' is not a clear statement, unless a degree of 'monotonicity' is defined. Please rephrase.
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.
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).
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.
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.
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?
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).
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?
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.
52) Page 4312, line 16. Please clarify why recursive filters make the analysis of length scale 'easier'.
53) Page 4313, lines 3-14 Figure 16 seems to be identical to figure 15. Please check.
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?
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?

56) Page 4314, line 12-13. Is this result specific to the examined case or is it expected in general?
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
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. Please precise more or consider removing this sentence.
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.
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.
61) Page 4317, line 1 and previous line. Either detail how the aerosol optical depth is used or do not mention it.
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.)
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.
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.
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?

Minor corrections:

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

2) Page 4293, lines 11-12: do you mean that the availability of more observations involve the control of new model variables? Please rephrase.
3) Page 4293, line 27. Change ‘the two first sections’ with ‘Section 2.1 and 2.2’
4) Page 4295, line 10: ‘comprised of’ should be ‘being comprised of’
5) Page 4296, line 3: Change ‘decomposed to’ in ‘decomposed into’
6) Page 4296, line 9: Please add ‘for each grid point’.
7) Page 4298, line 15: change ’24 minus…’ with ’24 h minus…’
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) Page 4301, line 5: specify that the length scale is horizontal
10) Page 4303, line 5: change ‘by EOF mode’ with ‘for each EOF mode’ and for vertical level as well
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…’

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