

Interactive comment on “Generalized Background Error covariance matrix model (GEN_BE v2.0)” by G. Descombes et al.

Anonymous Referee #3

Received and published: 15 August 2014

This paper presents a code called GENERate the Background Errors version 2.0 (GEN_BE v2.0), which goes together with data assimilation tools using the Weather Research and Forecasting (WRF) model. It allows different modeling of the background error covariance matrix B. Those elements can be very useful and this should be published in GMD. However, this paper lacks of scientific (or technical and research reports) references for both the description of the system and results interpretation. I think this is an important issue that needs to be addressed, some elements can be found in both minor and major comments.

Major comments:

Firstly, the main objectives and methods need to be clarifying in the introduction section. The authors present a “D-Ensemble” in the introduction. It seems that it is used

C1389

as a benchmark to be compare to. It is important to present how this setup is different to other simulations conducted using the variational approach and in different way of modeling B. Please provide some insight on first of the data assimilation setup, which and how many observations are effectively assimilated (even if it is only one), the analysis period and what are the differences in model simulations (grid spacing, boundary conditions, physics). This can be done using a table that summarizes the different experiments and set-up. Figures about horizontal and vertical domain dimensions can be presented at the same time. Also, the authors might consider some minor reorganization of the paper as indicated below:

Present a brief review of the 5 stages (Sect. 3) and the code structure in the introduction. I suggest to move all the code/software aspects in an appendix, especially “Sect. 3.2.1 FORTRAN code and input/output” and “3.2.3 Installation, compilation, set up and visualization”. This appendix must expose the general structure of the code (options, names of files...). It must be an intermediate between results and methods exposed in the main text (add references to the appendix if necessary) and the code in the supplementary material.

I think the appendix on chemistry can be really interesting for that community, and those results should be presented in the main text if it goes with appropriate references (in addition to the one already presented in the paper). First, reviews of chemical data assimilation applied to air quality modeling can be found in Carmichael et al. (2008), Sandu and Chai (2011), and Lahoz et al. (2007) for stratospheric application. Then, you should refer to publications where assimilation has been done using WRF/CHEM and GSI or a 3D-Var (e.g. Pagowski et al. 2010, Schwartz et al. 2012, Li et al. 2013, Pagowski et al. 2014). Finally, some papers gives some estimation of similar quantities presented in figure A1, A2, and A3, such as errors length scale and variances (e.g. Constantinescu et al. 2007, Schwinger and Elbern 2010, Jaumouillé et al. 2012, Gaubert et al. 2014, Robichaud and Ménart 2014).

Minor comments:

C1390

P4293 L3-4: This sentence is not clear to me: “assuming that the underlying probability errors are normally distributed”, I would suggest “assuming that errors are normally distributed”.

P4293 L5-7: “that minimize the error covariance between variables”, which variables is it, the control variable? The verb “minimize” is confusing since this “determination” is done a priori. P4293 L19: Please provide a reference for WRF, UM and WRFDA as it is done for GSI.

P4294 L13 Can you add an appropriate reference for DART, e.g. Anderson et al. (2009).

P4295 L2: “to the irregularly distributed observation locations”, you can remove “irregularly distributed”.

P4295 L3: Note that the exact knowledge of R and B would theoretically require the knowledge of the true state of the atmosphere...”. I would say “By definition, exact values of B and R would requires the knowledge...”.

P4295 L7: “i.e. uncorrelated observations, ...”, you can say “i.e. uncorrelated observation errors” or “i.e. observations are assumed to be independent, ...”.

P4295 L9 to 15: This paragraph should be more detailed and presented before the description of the different section. “All the results presented in the different sections were obtain from a numerical experiment with the WRF model”, this statement does not seems to be true, see for example in Sect. 5.2 (P4313, L17): “...we conducted a series of tests in which pseudo-observations of hydrometeors were assimilated into WRF-DA...”.

P4295 L23: “using a non linear observation operator. H is the tangent linear operator”. Please clarify, H is the observation operator and can be linear.

P4297 L8-9: “The new version of the code allows modeling a real time configuration of B like NCEP does using five control variables”. Can you clarify what do you mean by

C1391

“real time configuration”? Please remove “like NCEP does” or give a reference.

P4297 L14: “statistics of chemistry species to model B”, it is “error statistics of chemical species needed to model B”. “The community system Data Assimilation Research Test (DART)”, it is “Data Assimilation Research Testbed”

P4297 L24: “The version 2.0 of the code includes more physics options and flexibility has been added making all the algorithm in the different stages independent of the choice of control variable and model input”. This sentence is not clear and should be split like: “The version 2.0 of the code includes more physics options. In addition, the use of different stages, independent of control variable and model input allows more flexibility.

P4298 L5: please correct “proxi”.

P4298 L20: Add some references (e.g. fisher 2003, Pereira et al. 2006).

P4299 L28 to P4300 L5: “Stage 1 creates the NetCDF file bin.nc ... module io_input.f90” Please move theses sentences in an appendix dedicated to the code description.

P4300 L12 “The NCEP method”, please provide a reference.

P4302 L10-L15 Please rephrase and move the algorithm description in an appendix dedicated to the code description.

P4302 L18: Can you provide values in their unit and the level in parentheses?

P4303 L6: Please rephrase “horizfunc = gauss”, put this options in parentheses.

P4303 L10: Idem.

P4304 L1 to L7: Please avoid the use of codes variables. You need to clarify the different available options.

P4304-4305 Sect. 3.2.1: please move this paragraph dedicated to the code descrip-

C1392

tion.

P4305 L3: Change "variational a data assimilation" to "a variational data assimilation.

P4304-4305: please move this paragraph dedicated to the code description.

P4305 L12: "For example, NCEP operates", please give a reference.

P4305 L15-19: Please clarify and describe code options in parentheses.

P4306 L22: "A univariate version . . .", please provide a reference for this statement.

P4308 Sect. 3.2.3: Please move that section in an appendix dedicated to the code description.

P4309 L3: can you describe the NAM acronym.

P4309 L12: "The first five eigenvectors are shown Fig. 6", "are shown in Fig. 6".

P4310 L9: Can you indicate distances first and the grid point in parentheses?

P4310 L15: "parabolic approximation Eq. (6)." Please rephrase.

P4310 L24: "and the observation error of 1K." Is it not "an observation error of 1K"?

P4311 L9: "Bnam matrix coming from NAM". Can you specify, like "constructed from NAM forecast error statistics".

P4312 L12: The Fig. 14 should have been presented before.

P4312 L20: Please indicate the distance in km first.

P4314 L15: Define the acronym NWP.

P4314 L17-18: Give references about Meteo-France and the Met-Office system.

P4315 L28: "even if data assimilation of chemical species and aerosols remains difficult due to strong non-linearities". This statement is imprecise and needs referencing. Table B2: Last row, last column, 'readable', you mean readable.

C1393

Figure 1: the last word is statistics.

Figure 3: 'Sparecly'. The sentence is not clear.

Figure 6-8-9-10-16-A2-A3: Can you indicate (remind), at least an approximate value, how much distance is representing by a grid point.

Figure A1, can you redo this figure using ppbv instead of ppmv, especially for the NOx

References:

Anderson, Jeffrey, Tim Hoar, Kevin Raeder, Hui Liu, Nancy Collins, Ryan Torn, and Avelino Avellano. "The data assimilation research testbed: A community facility." *Bulletin of the American Meteorological Society* 90, no. 9 (2009): 1283-1296.

Carmichael, Gregory R., Adrian Sandu, Tianfeng Chai, Dacian N. Daescu, Emil M. Constantinescu, and Youhua Tang. "Predicting air quality: Improvements through advanced methods to integrate models and measurements." *Journal of Computational Physics* 227, no. 7 (2008): 3540-3571.

Constantinescu, Emil M., Adrian Sandu, Tianfeng Chai, and Gregory R. Carmichael. "Ensemble-based chemical data assimilation. II: Covariance localization." *Quarterly Journal of the Royal Meteorological Society* 133, no. 626 (2007): 1245-1256.

Fisher, M.: Background error covariance modeling, Proc. Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean, Reading, UK, ECMWF, available at: http://old.ecmwf.int/newsevents/meetings/annual_seminar/seminar2003_presentations/Fisher.pdf (last access: 8 August 2014), 45–63, 2003.

Gaubert, B., A. Coman, G. Foret, F. Meleux, A. Ung, L. Rouil, A. Ionescu, Y. Candau, and M. Beekmann. "Regional scale ozone data assimilation using an ensemble Kalman filter and the CHIMERE chemical transport model." *Geoscientific Model Development* 7, no. 1 (2014): 283-302.

Jaumouill e, E., S. Massart, A. Piacentini, D. Cariolle, and V-H. Peuch. "Impact of a

C1394

time-dependent background error covariance matrix on air quality analysis." *Geoscientific Model Development* 5, no. 5 (2012): 1075-1090.

Lahoz, W. A., Q. Errera, R. Swinbank, and D. Fonteyn. "Data assimilation of stratospheric constituents: a review." *Atmospheric Chemistry and Physics* 7, no. 22 (2007): 5745-5773.

Li, Z., Z. Zang, Q. B. Li, Y. Chao, D. Chen, Z. Ye, Y. Liu, and K. N. Liou. "A three-dimensional variational data assimilation system for multiple aerosol species with WRF/Chem and an application to PM 2.5 prediction." *Atmospheric Chemistry and Physics* 13, no. 8 (2013): 4265-4278.

Pagowski, Mariusz, G. A. Grell, S. A. McKeen, S. E. Peckham, and D. Devenyi. "Three-dimensional variational data assimilation of ozone and fine particulate matter observations: some results using the Weather Research and Forecasting Chemistry model and Gridpoint Statistical Interpolation." *Quarterly Journal of the Royal Meteorological Society* 136, no. 653 (2010): 2013-2024.

Pagowski, M., Liu, Z., Grell, G. A., Hu, M., Lin, H.-C., and Schwartz, C. S.: Implementation of aerosol assimilation in Gridpoint Statistical Interpolation (v. 3.2) and WRF-Chem (v. 3.4.1), *Geosci. Model Dev.*, 7, 1621-1627, doi:10.5194/gmd-7-1621-2014, 2014.

Pereira, Margarida Belo, and Loik Berre. "The use of an ensemble approach to study the background error covariances in a global NWP model." *Monthly weather review* 134, no. 9 (2006): 2466-2489.

Robichaud, A., and R. Ménard. "Multi-year objective analyses of warm season ground-level ozone and PM 2.5 over North America using real-time observations and Canadian operational air quality models." *Atmospheric Chemistry and Physics* 14, no. 4 (2014): 1769-1800.

Sandu, Adrian, and Tianfeng Chai. "Chemical data assimilation: an overview." *Atmosphere* 2, no. 3 (2011): 426-463.

C1395

Schwartz, Craig S., Zhiquan Liu, Hui-Chuan Lin, and Stuart A. McKeen. "Simultaneous three-dimensional variational assimilation of surface fine particulate matter and MODIS aerosol optical depth." *Journal of Geophysical Research: Atmospheres* (1984–2012) 117, no. D13 (2012).

Schwinger, J., and H. Elbern. "Chemical state estimation for the middle atmosphere by four-dimensional variational data assimilation: A posteriori validation of error statistics in observation space." *Journal of Geophysical Research: Atmospheres* (1984–2012) 115, no. D18 (2010).

Interactive comment on *Geosci. Model Dev. Discuss.*, 7, 4291, 2014.

C1396