Overview

The manuscript "Development of a variational flux inversion system (INVICAT v1.0) within the TOMCAT chemical transport model" describes the construction of a new 4D-Var data assimilation system for inverse modeling of different atmospheric tracers. Two adjoint tests were performed to validate the newly ATOMCAT adjoint model. Finally the variational system INVICAT accurately reproduced a set of surface fluxes in a twin experiment framework using SF6 pseudo-observations. Overall the manuscript is well written but several issues need to be addressed:

- The authors must replace the standard 3D-Var equations with the ones valid for 4D-Var approach. Consequently the observation term inside the cost function should include a summation reflecting observations at different times and this change should be reflected also in the cost function gradient formula. Thus a consistent part of Section 3 must be rewritten clarifying also some confusion aspects regarding TOMCAT model and the atmospheric transport component.
- A third adjoint test should be performed and the results should be plotted. This is called alpha test (see Navon et al. [6]) and gives the readers a chance to graphically notice the adjoint model accuracy. The authors must compute the expression $E(\alpha) = \frac{J(\mathbf{x}+\alpha\delta\mathbf{x})-J(x)}{\alpha\delta\mathbf{x}\nabla J(\mathbf{x})}$. For values of α that are small but not too close to the machine zero, one should expect to obtain a value of $E(\alpha)$ that is close to 1. A figure depicting the values of E with respect to α is desired.

I recommend the publication of the present manuscript with major revision.

Specific Comments

- The authors should revise the title of the manuscript since the data assimilation system INVICAT is making use of the forward and adjoint TOMCAT models and consequently INVICAT is not a part of TOMCAT. A more suitable title would be "Development of a variational flux inversion system (INVICAT v1.0) using the TOMCAT chemical transport model".
- 7118.25 The phrase starting with "In order to accurately model .." is too long and the authors should split it in two sentences.
- 7119.10 Remove "an" before "atmospheric species".
- 7120.15 The authors should mention what type of 4D-Var approach did use to build the INVICAT variational system. Is it the incremental method proposed by Courtier et al. [2] or the standard 4D-Var introduced in meteorology by Le-Dimet and Talagrand [5].
- 7120.20 Replace "similar that" with "similar to that".
- 7120.20 NWP data assimilation schemes employ variational approaches so the sentence makes no sense.
- 7120.20 In the abstract the authors describe the construction of a new adjoint model. Why here it is referred as something already existing?
- 7121.10 "Earlier references" should be removed since it creates confusion.
- 7121.15 The phrase starting with "The variational inverse modelling .." is too long and need to be revised. Perhaps splitting it is a good idea.
- 7122.10 As already mention above, the authors should cite the work of Courtier et al. [2] where low resolution linearized model replaces the high fidelity non-linear model during the optimization process. Also cite the work of Stefanescu et al. [10] and Amsallem et al. [1] where non-linear reduced models are used as constrains instead of the full non-linear models. Both approaches are capable to alleviate the computational burden of data assimilation experiments.
- 7123.5 Remove "an" before "atmospheric species".
- 7123.10 Split into two parts the phrase starting with "The state vector" one detailing on the control variables and the other one describing the temporal and spatial resolutions.
- 7123.15 If the authors refer to $\mathbf{M}$ as a matrix than the atmospheric transport model in the TOMCAT model is linear. Is it really the case?
- 7123.15 Since this is a 4D-Var data assimilation strategy the observation term should include a summation reflecting observations at different times.
- 7123.20 What is the nature of the observation operator $\mathbf{H}$? Is it linear or non-linear? If the observation operator role is only to interpolate from the modelled atmospheric concentrations to the observation space than the state variables include also the atmospheric concentrations. I think the authors should clarify this in the text and describe the state and control variables.

- 7124.5 The observational term in equation (4) should be rewrite according to 4D-Var. Otherwise the variational method considered here is 3D-Var
- 7124.15 The contribution of the observational term to the cost function gradient in equation (5) should be modified in order to reflect the 4D-Var algorithm.
- 7124.20 Here the authors mention that the adjoint version of the TOMCAT model is equivalent to the transpose of the matrix $\mathbf{M}$. If this is correct this means that the TOMCAT model is the same as the atmospheric transport model $\mathbf{M}$ which is linear. This is hard to believe since the chemical transport models deal with very high nonlinearities. The authors should clearly explain what is the difference between TOMCAT and atmospheric transport model $\mathbf{M}$. Enumerate also the models variables. I also think that it is more difficult to compute the entire adjoint M^* that only a part of it $\mathbf{M^T}$. Thus the entire 7124.15 and 7124.20 should be rewritten.
- 7125.25 The authors should clarify what are the models used for computing the gradient $\nabla_{\mathbf{x}} J(\mathbf{x})$. If the authors employ the TOMCAT and the new proposed adjoint model than the equations (1,4,5) should be written accordingly using M and M^* notations.
- 7125.5 The authors should cite the pioneer work on LBFGS optimization approach such as Perry [8], Shanno [9], Nocedal [7].
- 7125.20 The phrase starting with "Assuming that ..." should be rewritten since the tangent linear model is just a first order approximation of the first derivative of the operator M.
- 7126.20 Remove "which remains a possibility for the discrete adjoint".
- 7127.5 It would be nice if the authors can comment (outside the manuscript) on automatic differentiation software performances for generating tangent linear and adjoint models for non-linear Newton based discrete models.
- 7127.15 The authors should also refer to WRFDA variational data assimilation system. It makes use of an adjoint model and it is employed to generate atmospheric analysis (see Huang et al. [4]).
- 7128.10 Another way to incorporate model errors is to introduce an additional term in the cost function as in the weak constraint 4D-Var. While for the present manuscript the authors investigates the performance of the variational system using an idealized case for the future ones real SF_6 observations should be used. Consequently in addition to representativeness errors the authors should also consider measurements errors inside $\mathbf{R}$. Do you have any idea how to quantify these measurements errors for SF_6 observations?
- 7129.10 What do you mean by it "performed well"? This particular text should be removed since it is not a rigorous statement.
- 7129.15-20 It is not straightforward what is the space resolution of the TransCom initial conditions. Did you need interpolation from fine to coarse resolutions?
- 7129.25 The authors should define "ppt" first before using it.
- 7129.20-25 Rephrase the sentence starting with "The model replicates ...".
- 7132.20 Why do you need a new version of TOMCAT model? It is enough only to store the forward variables required by the adjoint model (this is called checkpointing - see Griewank and Walther [3]). If your model is linear you don't need checkpointing.
- 7133.5 Another way to test the accuracy of the adjoint model is to compute

$$E(\alpha) = \frac{J(\mathbf{x} + \alpha\delta\mathbf{x}) - J(x)}{\alpha\delta\mathbf{x}\nabla J(\mathbf{x})}.$$

  For values of α that are small but not too close to the machine zero, one should expect to obtain a value of $E(\alpha)$ that is close to 1. The adjoint model helps to compute the gradient justifying the usefulness of the test. This linearity test is derived from the alpha test described by Navon et al. [6]. It would be nice to introduce a figure describing the results of this test.
- 7134.5 I think that only one iteration is not enough to accurately test the adjoint subroutines and longer experiments are required.
- 7136.25 The phrase starting with "A quasi-Newtonian ..." is too long and needs to be splitted.
- 7137.5 (7136.25) Define MLISO.
- 7140.10 In 7124.25 the authors mentioned that the adjoint version of the entire TOMCAT model was written. However this contradicts the first sentence of Summary section which states that only adjoint version of the the transport section of the TOMCAT CTM was created.

References

[1] D. Amsallem, M. Zahr, Y. Choi, and C. Farhat. Design Optimization Using Hyper-Reduced-Order Models. Technical report, Stanford University, 2013.

[2] P. Courtier, J.-N. Thepaut, and A. Hollingsworth. A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120:1367–1387, 1994.

[3] Andreas Griewank and Andrea Walther. Algorithm 799: Revolve: An implementation of checkpointing for the reverse or adjoint mode of computational differentiation. *ACM Trans. Math. Softw.*, 26(1):19–45, March 2000. ISSN 0098-3500. doi: 10.1145/347837.347846. URL `http://doi.acm.org/10.1145/347837.347846`.

[4] Xiang-Yu Huang, Qingnong Xiao, Dale M. Barker, Xin Zhang, John Michalakes, Wei Huang, Tom Henderson, John Bray, Yongsheng Chen, Zaizhong Ma, Jimy Dudhia, Yongrun Guo, Xiaoyan Zhang, Duk-Jin Won, Hui-Chuan Lin, and Ying-Hwa Kuo. Four-dimensional variational data assimilation for wrf: formulation and preliminary results. *Mon. Wea. Rev.*, (137):299–314, 2009.

[5] F.X. Le-Dimet and O. Talagrand. Variational algorithms for analysis and assimilation of meteorological observations. *Tellus A*, 38:97–110, 1986.

[6] I. M. Navon, X. Zou, J. Derber, and J. Sela. Variational Data Assimilation with an Adiabatic Version of the NMC Spectral Model. *Monthly Weather Review*, 120(7):1433–1446, 1992.

[7] Jorge Nocedal. Updating Quasi-Newton Matrices with Limited Storage. *Mathematics of Computation*, 35(151):773–782, 1980. URL `http://www.jstor.org/stable/2006193`.

[8] Avinoam Perry. A Class of Conjugate Gradient Algorithms with a Two-Step Variable Metric Memory. Discussion Papers 269, Northwestern University, Center for Mathematical Studies in Economics and Management Science, January 1977. URL `http://ideas.repec.org/p/nwu/cmsems/269.html`.

[9] David F. Shanno. Conjugate gradient methods with inexact searches. *Mathematics of Operations Research*, 3(3):244–256, 1978. doi: 10.1287/moor.3.3.244. URL `http://pubsonline.informs.org/doi/abs/10.1287/moor.3.3.244`.

[10] R. Stefanescu, A. Sandu, and I.M. Navon. POD/DEIM Reduced-Order Strategies for Efficient Four Dimensional Variational Data Assimilation. Technical Report TR 3, Virginia Polytechnic Institute and State University, March 2014, also submitted to Journal of Computational Physics.