Geosci. Model Dev. Discuss., 8, C820–C826, 2015
www.geosci-model-dev-discuss.net/8/C820/2015/

Geoscientific Model Development Discussions

Interactive comment on "Development and application of the WRFPLUS-Chem online chemistry adjoint and WRFDA-Chem assimilation system" *by* J. J. Guerrette and D. K. Henze

J. J. Guerrette and D. K. Henze

jonathan.guerrette@colorado.edu

Received and published: 19 May 2015

Thank you to the anonymous referee for their comments and questions. We edited the manuscript and have additional responses below.

1 Responses to specific comments

1. **Page 2316, line 28: WRF-4-DVar -> WRF-4D-Var**

   This error is corrected.

2. **Page 2318, line 8: WRF is spelled out here, but it appeared earlier in text. Some other abbreviations and symbols (e.g. FWM, Qv, ∼) are not spelled out or explained.**

   We rephrased the sentence starting on page 2315, line 15 to read: "Grell et al. (2004) used the Weather Research and Forecasting Model with chemistry (WRF-Chem) (Skamarock et al., 2008; Grell et al., 2005) to show that vertical mass transport..."

   Later citations and locations where "Weather Research and Forecasting Model" appears have been similarly adjusted.

   FWM: FWM first appeared on page 2316, line 29 of the original manuscript, as referencing "forward model".

   Q_v: The sentence starting on page 2326, line 4 is corrected to: "The CVs include initial conditions for BC_1, zonal wind (U), temperature (T), and water vapor mixing ratio (Q_v), and also BC emission scaling factors (α_{BC})."

   Tilde ($\sim$): If the referee is referring to the tilde used for $\tilde{r}_k$, this notation was described on page 2339 of the original manuscript. The tilde simply means that some form of residual error must be chosen by the user. For instance, it may be the mean or median of model ensembles for a particular observation location, if that information is available to the user.

3. **Page 2327, lines 1-5: Please specify how many 3-D state variables are there for the example that requires 1.46 GB per core on 64 cores.**

   The sentence starting on page 2327, line 3 is modified as follows: "For an illustrative domain, ..., 42 levels, a 5 cell boundary width, and twenty-eight 3-D state variables, the trajectory ..."

4. **Page 2331, line 29: Missing reference in"()".**

   We have fixed the reference in the manuscript.

5. **Page 2336: Equation (19) and $L_{max} = 9$ seems pretty arbitrary. The authors need to justify their choices here. In addition, this is not how representative errors are defined.**

   $L_{max} = 9$ is the result of comparing 10 s frequency observations to model variables available each 90 s time step. The DC-8 aircraft flight speed (150 to 200 $\mathrm{m\,s^{-1}}$) coincides with traversing an 18 km model column edge in 90 to 120 s. Thus we bin every 9 observations, centered around each model time step, and compare their average to the model concentration in the nearest model grid cell to the average location of those 9 observations. This method is described in the first paragraph on page 2332.

   The representative error is mislabeled in Eq. 20, which is actually a description of the average instrument error. The quantity in Eq. 19 is the representative error, because it captures the variability of measured concentrations within the model grid cell. For sure, the observations along an aircraft transect may not represent the full variability within the grid cell, but we do not know that information. We made an attempt to account for missing information by scaling the instrument error in Eq. 20. We have relabeled the variance definitions and adjusted the related text for Eqs. 18, 19, and 20.the following equations:

6. **Page 2343: Apparently, Eq.(30) does not hold for the weekday/weekend anthropogenic emissions, which d=1,...,7 does not apply.**

   Thank you for catching this mistake. The text and equation are modified to apply to a general number of days, n_d, so as to apply to weekday/weekend emission populations.

7. **Figure 1: A table would be more appropriate for this.**

   Refer to the new Table 1 in the resubmitted manuscript which replaces Fig. 1. Note that the remaining figures will have their numbers reduced by one.

8. **Figure 3: Please specify the meaning of "m" (slope) in caption.**

   We added the following sentence at the end of the Figure 3 caption: “The slope (m) and R^2 statistic for the linear fit are shown for each CV.”

9. **Figure 5: Can plots be arranged in a way that the same row/column represents the same J/x ? The case [J=BC1,x=U] looks really bad.**

   The arrangement of plots puts $J = \mathrm{BC}_1$ in the first row. Then derivatives of meteorological variables with respect to $x = Q_v$ and $x = U$ are in the second and third rows, respectively. Since only BC_1 has a nontrivial dependence on emissions (without radiative coupling), it prevents a different arrangement without showing another control variable. $J = [U, Q_v, T]$ were tested across $x = [Q_v, U]$ specifically because the cases [$J = \mathrm{BC}_1$,$x = Q_v$] and [$J = \mathrm{BC}_1$,$x = U$] looked bad. Please see our response to M. Krol, comment #8, for additional discussion of revisions to Fig. 5.

10. **Figure 5: Will smaller perturbations (δx <1%) generate better results? Are the authors confident that there are no mistakes made in this calculation? For instance, the adjoint boundary conditions could be wrong.**

    Smaller perturbations will not necessarily generate improved finite difference approximations, if the nonlinear models is discontinuous. We discussed this behavior in the reply to M. Krol, and repeat the explanation here:

    > The referee is right that continuous model equations should lead to finite difference approximations becoming more accurate as step size is decreased, which is a benefit to using them to approximate derivatives for nonlinear systems. In their Fig. 4, Henze et al. (2007) showed that this is not the case for discontinuous model equations, where larger perturbations may lead to smaller errors. This phenomena is described by Thuburn and Haine (2001), and likely arises in WRFPLUS-Chem

due to flux limiters in the 5th order, monotonic, horizontal tracer advection.

We are confident that the ADM and TLM are calculating exact derivatives. First, this is a comparison between tangent linear and nonlinear models, not the adjoint. We have ensured that state variables (e.g., U, V, T, Q_v, BC_1) in both models are identical in all time steps. This is not a big challenge in the TLM since it integrates forward in time. Boundary conditions, initial conditions, and emissions are not sources of error, because those were checked first, and are necessary to get the state variables matching.

On a related note, we found that there was one mistake in the caption and labels in Fig. 5. All of the derivatives are fully normalized, so that their magnitudes are comparable. Refer to the new Fig. 5 caption for the correction.

2 References

- D. K. Henze, A. Hakami, J. H. Seinfeld, and others, "Development of the adjoint of GEOS-Chem," Atmospheric Chemistry and Physics, vol. 7, no. 9, pp. 2413–2433, 2007.
- J. Thuburn and T. W. N. Haine, "Adjoints of Nonoscillatory Advection Schemes," Journal of Computational Physics, vol. 171, no. 2, pp. 616 – 631, 2001.

Interactive comment on Geosci. Model Dev. Discuss., 8, 2313, 2015.

f05.pdf

Fig. 1. Fully normalized time variant sensitivities calculated with the TLM with second order checkpointing and with multiple finite difference perturbation sizes. Each plot is for a single pair of source and receptor locations, q and p.

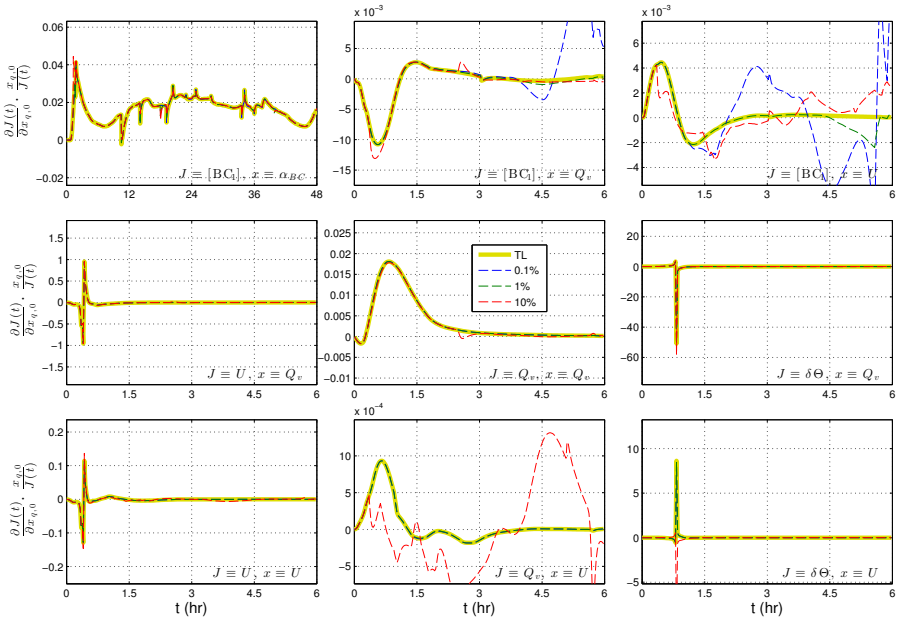

Fig. 2. Fully normalized time variant sensitivities calculated with the TLM with second order checkpointing and with multiple finite difference perturbation sizes. Each plot is for a single pair of source an