Author's Response

J. J. Guerrette[1] and D. K. Henze[1]

[1]Department of Mechanical Engineering, University of Colorado, Boulder, CO, 80309, USA

General comments:

(1) After replacing Figure 1 with Table 1, all of the figure numbers are reduced in the new manuscript. The old figure numbers are used in the discussion.
(2) During the discussion period, the authors identified an error in the aircraft observation operator described in the manuscript Section 5.1.2. The SP2 observations are reported at 1013 hPa, 273K, but the original observation operator did not account for the temperature and pressure conversion. This error has since been corrected in WRFPLUS-Chem. Figures 7-12 (old manuscript figure numbers) have been updated accordingly. The discussion for Figs. 11 and 12 is also adjusted, including text in the abstract and conclusions.

The remainder of the document is organized as follows:

(1) Reply to M. Krol
(2) Reply to Anonymous Referee #2
(3) latexdiff output showing manuscript changes

Geosci. Model Dev. Discuss., 8, C801–C819, 2015
www.geosci-model-dev-discuss.net/8/C801/2015/

Geoscientific Model Development Discussions
Open Access

GMDD
8, C801–C819, 2015

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

We are thankful for all of the constructive comments and questions by M. Krol. The corresponding responses and manuscript changes are given below.

1 Responses to specific comments

1. **In the introduction, no reference is made to the pioneering work of Elbern et al. with the EURAD model, who worked on 4D-VAR chemical data assim-**

ilation for more than two decades.

The introduction to this paper focused on literature describing data assimilation methods that either utilized an online model or were for aerosols. Bocquet at al. (2014) review existing online and offline chemical data assimilation capabilities in more detail. As such, we have modified page 2315, line 25-26 as follows:

> To address this, chemical data assimilation can be used to improve short-term forecasts. Bocquet at al. (2014) review existing methods and previous applications of chemical data assimilation in CTMs and NWP-chemistry models.

2. **In general, it would be interesting to compare the approach described here to other approaches. For instance, some 4DVAR approaches (e.g. Bergamaschi, P., Frankenberg, C., Meirink, J. F., Krol, M., Villani, M. G., Houweling, S., et al. (2009). Inverse modeling of global and regional CH 4emissions using SCIAMACHY satellite retrievals. Journal of Geophysical Research, 114(D22), D22301. doi:10.1029/2009JD012287) use a two-step inversion. Observations that are not fitted within 3σ after the first optimization are left out with an argument that the model is not able to reproduce these observations. In the current study, some of the high aircraft observations may be due to specific layered outflow from a specific convection event, which is not (and might never be) adequately resolved by the model. Nevertheless, the advanced estimation of model error with the different settings in WRF is impressive. Without a true inversion, however, it is not possible to assess how well the observations finally will be matched. My main point here is that a discussion of this work in the context of existing techniques would be of added value.**

We agree that the online data assimilation method will need to be compared against existing CTM 4D-Var systems, especially to evaluate the benefits of in-

GMDD
8, C801–C819, 2015

cluding online physics. Still, the purpose of this paper is to describe development of a chemical adjoint with online meteorology, which is the first step toward enabling the online 4D-Var system in WRFDA-Chem. The WRFDA-Chem platform is not at a state where a comparison might be made between online and offline approaches.

Additionally, the weighting scheme presented here cannot be compared to some other observation filtering method outside the context of a 4D-Var inversion. As stated on pp. 2339-2340, we introduce the weighting scheme, but do not exhaustively test it. You are correct that future 4D-Var studies will require distinction between residual error due to emission inventory and physical parameterization errors. As you mention, Bergamaschi et al. (2009) remove observations outside 3σ after an initial 4D-Var optimization that includes four outer loop iterations. Analyzing residual errors after an inversion is very useful to determine where model descriptions are weak. As such, we have amended the paragraph starting on page 2337, line 13 as follows:

> Observations with significant model bias would require the largest perturbation in control variables to alleviate, and would seem to inform the inversion process the greatest. However, they must also have low total variance to contribute to an inversion. Figure 9 shows the surface and aircraft SD plotted versus residual error. Also plotted in that figure are one and two SD zones, as well as lines of constant $\lambda^*_{k,o}$ for all $w_k = 1$. Any residual falling outside the 2σ zone has a combined model and observation SD that is small enough to determine with 95 % confidence ($p < 0.05$) that the residual error deviates from zero (i.e., the model and observation disagree). These statistically significant model errors indicate that some kind of inversion is worthwhile. In their multicycle 4D-Var approach, Bergamaschi et al. (2009) eliminate observations outside three SD's after an initial 4D-Var cycle, with the thought

GMDD
8, C801–C819, 2015

that incorrect model physics prevent those residual errors from being fixed with 4D-Var. Thus, while statistically significant residuals are important to driving a 4D-Var inversion, that they remain afterward is a strong indication of errors in the model description that cannot be fixed through adjustments to emissions. Figure 9 shows that the relative contributions of observation and model variances is in general proportional to the relative magnitudes of observed and modeled concentration. Specifically, model (observation) variation contributes to a large fraction of uncertainty in positive (negative) residuals.

3. **I do not see why the summation is split in eq. 2a and does not simply run to n. Please explain.**

   The summation is split to distinguish the inner loop control variable increment, $\vec{\delta x}$. We have changed Eqs. (2a) and (2b) to the following:

$$\mathsf{J_b} = \frac{1}{2}\left[\vec{\delta x} + \sum_{i=1}^{n-1}\left(\vec{x}^{i} - \vec{x}^{i-1}\right)\right]^{\top} \mathbf{B}^{-1} \left[\vec{\delta x} + \sum_{i=1}^{n-1}\left(\vec{x}^{i} - \vec{x}^{i-1}\right)\right] \text{ and } \mathsf{J_o} = \frac{1}{2}\sum_{k=1}^{K}\left\{H_k\left[M_k(\vec{x}^{n})\right] - \vec{y}_k\right\}^{\top} \mathbf{R}_k^{-1} \left\{H_k\left[M_k(\vec{x}^{n})\right] - \vec{y}_k\right\}$$
$$\approx \frac{1}{2}\sum_{k=1}^{K}\left[\mathbf{H}_k\mathbf{M}_k\vec{\delta x} - \vec{d}_k\right]^{\top} \mathbf{R}_k^{-1} \left[\mathbf{H}_k\mathbf{M}_k\vec{\delta x} - \vec{d}_k\right].$$

4. **Page 2321, line 2: "earliest emission time": up to now "x" was a general variable, that is now linked suddenly to emissions. Please explain this better.**

   Indeed, this sentence should be kept more general. We will replace "the earliest emission time" with "earlier simulation times".

5. **Page 2321, line 17: "nonlinear to a quadratic form". I see what you mean,**

GMDD
8, C801–C819, 2015

but strictly speaking quadratic is also nonlinear.

This sentence was incorrectly paraphrased. Page 2321, lines 16-17 will be corrected to: "The purpose of the two-level optimization is that approximating M with $\mathbf{M}$ simplifies the full problem to a quadratic problem, and guarantees a unique solution to the minimization..."

6. **Page 2325: "some cost function at location p and time step f with". Up to now, the cost function was introduced as a global variable (eq. 1). Defining it here as a space and time dependent variable is confusing.**

   In order to make the new definition less confusing, we have explicitly stated the new definition instead of only implying it. We have changed the text on page 2325 from lines 4 to 13 as follows:

   > ... We define a new cost function equal to a single predicted state variable, locally defined in grid cell p and at the end of time step f, $J = SV_{p,f}$. We use the TLM, ADM, and a centered finite difference approximation from the FWM to evaluate derivatives $\chi_{p,q} = \frac{\partial J}{\partial x_{q,0}}$, with respect to some CV at location q and the initial time, 0. The finite difference derivatives are calculated from $\chi^{NL}_{p,q} \approx \frac{J(x_{q,0}+\delta x)-J(x_{q,0}-\delta x)}{2\delta x}$, where each evaluation of J results from a FWM simulation with some perturbed value of $x_{q,0}$. δx varies between 0.1 and 10% of the value of $x_{q,0}$. The adjoint and tangent linear derivatives are found by forcing the model gradient fields, λ^* and λ, at J and x_q, respectively. ...

7. **Page 2325: Comparing the results of eqs. 5 and 6 is known as the gradient test (see e.g. ECMWF documentation). Normally, you take dx that approaches zero and the finite difference gradient will approach the true gradient until numerical rounding errors become important. To my experience, for double precision calculations, derivatives can be approximated**

GMDD
8, C801–C819, 2015

within 10^{-7} before rounding errors kick in. In my applications there is convergence until dx is about 10^{-6}. I do not see why one has to fiddle around with different values of dx (0.1, 1, 10%, i.e. relatively large values) to see what value performs best.

We appreciate the referee pointing out that the gradient test is not specific to any modeling system, and have corrected the sentence beginning page 2325, line 3 to be: "Here we use an alternative verification approach based on Taylor series derivative approximations, and similar to that used by, e.g., Henze et al. (2007), to verify WRFPLUS-Chem."

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 numerical algorithms, where larger perturbations may lead to smaller errors. This phenomena is described by Thuburn and Haine (2001), and it likely arises in WRFPLUS-Chem due to flux limiters in the 5th order, monotonic, horizontal tracer advection.

8. **In the paper (page 2326) it is written: "A range of finite difference perturbations dx is used for U, T , and Q_v control variables in order to find a value of χ_{NL} with the best compromise between truncation and roundoff error." Another problem might be that perturbations to U in the forward model perturb the physics (atmospheric flow is normally defined in vorticity and divergence), and that this violates some mass-conservation constraints. This might be the reason for the strange behavior presented later in figure 5, which look rather suspect in my opinion. The sensitivities for something linear as emissions (figure 5, first panel) look perfectly fine and what would be expected.**

In response to your questioning perturbations in U, the standard WRFPLUS ver-

GMDD
8, C801–C819, 2015

ification test described by Zhang et al. (2013) perturbs U, T, and Q_v, but does so across the entire domain simultaneously. Their approach results in a decrease of error as the domain-wide perturbation is reduced, however it likely averages out any discontinuities introduced at specific perturbation locations. To be clear, the T used in the model is actually the perturbation potential temperature. Thus we have changed T to $\delta\Theta$ in the paper.

We agree that the finite difference results in Fig. 5 reflect substantial nonlinear responses. As such, we repeated the test with new source code and post processing. Unfortunately, we were unable to determine the initial locations, q, that we used previously due to a data loss. However, we repeated these tests in locations likely to be equivalent; a revised version of Fig. 5 is shown below. We think that the resulting derivatives are within the tangent linear regime expected in an online model. Also note that the derivatives with respect to U and Q_v are evaluated very near the coastline, which could be a source of additional advective discontinuities (as discussed in response 7 above).

We have also modified the text starting on page 2329, line 3 until the end of that paragraph as follows:

> However, the TLM has inflection points at the same times as the finite difference approximations, including during fast transient periods, such as for $\frac{\partial U}{\partial U}$ and $\frac{\partial U}{\partial Q_v}$. The duration over which the tangent linear assumption is valid for chemical responses to U and Q_v depends on the size of the perturbation and on the local regime of meteorology. For instance, the test location shown is very near the California coast, but better agreement was found for an inland response location, p. Further testing of the coupled derivatives will be necessary to determine over what time period they are suitable for inverse modeling, and under what conditions the model nonlinearities cease to be a limiting factor. Future inversions with coupled physics and chemistry will need to verify that

GMDD
8, C801–C819, 2015

$\frac{\partial J}{\partial \alpha}$ has a near linear response over the time frame considered. The behaviors noted here are similar or improved across the other thirteen pairs of q and p.

9. **Page 2326: The adjoint test presented is figure 2 was compromised by some errors, as mentioned later in the paper (page 2328). So it seems logical to replace figure 2 by a corrected one.**

   The improved agreement between the ADM and TLM described on page 2328 resulted from changing single precision variable definitions to double precision in the TL version of the dry deposition velocity calculation. This changed the TLM derivatives in the 8th digit and beyond. There would be no perceptible change to Fig. 2, other than the Max Rel. Err. result. Because running these verification tests would take several weeks to repeat, and the ADM/TLM agreement is already within the bounds of a single precision calculation in Fig. 2, we believe it is unnecessary to repeat the verification.

10. **Page 2326, line 5: Q_v has not been introduced in the paper.**

    The sentence starting on page 2326, line 4 has been 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})."

11. **Page 2328, line 26: "BC concentrations respond linearly to a 1% perturbation of emissions for at least 48 h". Is there any reason that a non-linear response can be expected when coupling with BC and radiation is turned off?**

    Since the emission and transport processes are linear, we did not expect nonlinearities. However, it was not known whether discontinuities in the advection would introduce non-physical nonlinearities. We believe it is important to state this was not the case, at least for these tests. We modified the sentence beginning on

GMDD
8, C801–C819, 2015

page 2328, line 25 as follows: “Most importantly for multi-day 4D-Var emissions inversions, and as would be expected, BC concentrations respond linearly to ...”

12. **Page 2331, line 28: reference missing**

    We have fixed the reference in the manuscript.

13. **Page 2332: I am a bit worried that you use two different measurement techniques for BC. BC is particularly tricky to measure and LAC and TOR might have different biases. For sure, BC and EC cannot be compared directly, because they are defined differently. Using ARCTAS and IMPROVE data in the same inversion might have to deal with a bias of one method to the other. Maybe it is important to highlight how comparable the data are. There is a wealth of literature available on bias correction of particular data streams (e.g. satellite data).**

    This comment is important, because it calls upon the relatively few comparisons that have been made between the many BC/EC measurement techniques. This is an important point to consider for any model-observation comparison, especially when utilizing multiple in-situ measurement approaches. Yelverton et al. (2014) compared SP2 BC and IMPROVE EC from TOR, in addition to eight other BC or EC measurement protocols, for a single concentration of carbonaceous particles. Those authors found that time-averaged mass concentrations of SP2 BC and IMPROVE EC agree within 7%, with the EC values being larger, and the two averages being within 2σ. Although we have not applied any correction to either observation set, the experiment performed here is meant as a demonstration. A 7% change in concentrations from either device would not change our qualitative conclusions. For future work, we will be sure to apply the correction. Additionally, refractory coatings in aged aerosol or near biomass burning sources may introduce additional measurement bias. In light of this discussion, we added the following text to the paragraph ending on page 2331:

GMDD

8, C801–C819, 2015

While we use IMPROVE elemental carbon (EC) and SP2 absorbing carbon as equivalents herein, Yelverton et al. (2014) found that the former is approximately 7% higher than the latter, but that their error bars overlap. For the qualitative analysis performed in this demonstration, bias correction would not change any of the final conclusions.

14. **Page 2335, eq. 16: Sure you divide by L^2?. Anyhow, it would be better to have the non-summed part before the summation sign for clarity. Also for eqs. 17, 20, 24.**

The variance of the mean observation is $\mathrm{Var}(\bar{y}_k) = Var\left(\frac{1}{L_k}\sum_{l=1}^{L_k} y_k\right)$ $= \frac{1}{L_k^2} Var\left(\sum_{l=1}^{L_k} y_{l_k}\right)$. The variance of the sum of any two variables, X and Y is (e.g., http://en.wikipedia.org/wiki/Variance#Basic_properties) $\mathrm{Var}(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$. Since in our case, the observations are independent, the variance of the mean is then the sum of the variances, divided by the number of observations squared: $\mathrm{Var}(\bar{y}_k) = \frac{1}{L_k^2}\sum_{l=1}^{L_k} Var(y_{l_k})$ $= \frac{1}{L_k^2}\sum_{l=1}^{L_k} \sigma_{l_k}^2$.

In order to improve clarity, we moved the non-summed denominators before the summation symbol for Eqs. 14, 16, 17, 19, 20, and 24.

15. **Page 2338. In the discussion of the model-data mismatches and associated adjoint forcings, I miss a discussion of the role of the adjoint model. The H^T operator projects the mismatches to $\frac{dJ}{dc_k}$ and indicates how sensitive a particular observation is for a particular emission change. By only discussion the adjoint forcings and their magnitudes, I cannot see how you can defend the need for a w_k scaling in the covariance matrix. In my opinion, the need of this scaling only appears after a full inversion, and I would like to hear the authors' opinion about this.**

GMDD
8, C801–C819, 2015

We agree that the paragraph on page 2338 is unclear as to the justification for applying w_k. We devised the weighting scheme based on past experience with 4D-Var, and from considering how the $\mathbf{H}^\top$ operator informs the optimization. When the adjoint forcing, $\lambda^*_{k,o}$, increases in magnitude, there is greater potential for the adjoint operator, $\mathbf{H}^\top$, to generate large sensitivities (positive or negative). Consider Fig. 9, where we show that for positive residuals, the model uncertainty dominates, while for negative residuals, the observation uncertainty dominates. This is a result of each uncertainty being proportional to the respective concentration.

The rest of the justification is given below, and replaces the paragraph on page 2338 starting with "There are several outlier negative...":

> When both the observed and modeled concentrations are small, the total variance decreases to the minimum possible value, governed by the MML and MDL. This generally happens in remote regions, where small concentrations result from some combination of small nearby sources and transport from many distant sources. If the total variance is small enough relative to the residual error, $\lambda^*_{k,o}$ will be very large, often larger than in cases with larger residual errors (see Fig. 10a). The adjoint model propagates a relatively large forcing from a small residual backward, resulting in large sensitivities to emission scaling factors. These sensitivities then translate to large emission perturbations in the optimization process.
>
> The residual errors in remote locations are likely within combined model and observation uncertainty, but the model variance at these locations is unrealistically small. The ensemble will underestimate variance at observations near low-biased prior sources due to the absence of tracer mass. The opposite may be true for a high-biased prior. The challenge then is to define the concentration uncertainty introduced by

GMDD
8, C801–C819, 2015

the model physics, independent of the magnitude of emissions, which we attempt to do with a weighting scheme. The weights are used only to inflate variance, which when very low is thought to misinform the adjoint about concentration errors. Variance reduction may be necessary for observations near high-biased sources. Also, while we apply the weights to the total variance, they could be applied to only the model portion. Here we are developing a philosophy for scaling the variances, of which the following description is but one example.

To clarify the impact of adding the weights, we also will add the following sentences after line 13 on page 2340: "This is a considerable change from the unity weights where $|\lambda^*_{k,o}|$ was as large as $200\,\mu g^{-1}m^3$ in the region between the 1σ and 2σ zones."

16. **Page 2341, discussion figure 11. Figure is unclear to me. Results are shown for anthropogenic (red?) and biomass burning (blue?) BC emissions. I understand that the black markers highlight boxes with biomass burning, but how do I see where the anthropogenic emissions are and how this relates to the blue and red colors?**

We see where there might be some confusion. We have changed the legend to make it the figure more clear (see below). Hopefully the changes are sufficient, and you are able to understand the following. The color bar on the right indicates the magnitude of the sensitivities, whereas the shapes indicate the type of emissions. The squares apply to anthropogenic emissions, and the circles are for biomass burning sources. Black of either shape indicates a sensitivity near zero, as shown in the color bar. Because the grid cells with the largest anthropogenic and fire sources are not collocated, showing the largest magnitude normalized sensitivity in each surface grid cell does not obscure any critical information. Please indicate if any of these points are unclear.

GMDD
8, C801–C819, 2015

17. **Page 2343, line 21. The authors write: "The increased burning sensitivity magnitude indicate the weighting scheme is successful at generating a cost function that is more robustly sensitive to emission perturbations." I think this is not a valid reasoning. It is not surprising that the inversion is somewhat sensitive to the settings of the physical model parameters, simply because the boundary layer scheme determines how emissions are transported in the atmosphere and how the simulated observations look like. By using a different weighting these sensitivities become more alike, but this is not a proof that the new weighting scheme is better or worse: it simply gives different results because outliers receive more weight compared to better simulated observations. Like stated before: it is unclear why the authors felt the need to deviate from Bayesian statistics. Only after a true inversion and calculation of the associated statistics (e.g. χ^2 values) one might conclude that the weighting scheme gives more favorable results.**

The inclusion of a weighting term does not deviate from Bayesian statistics, but instead potentially introduces a non-Gaussian distribution of observation error. Alternatively, and as described in the response to comment 13 above, the statistics may still be Gaussian if we are inflating variance where it is not fully captured by the model ensemble. Still, we have reworded Page 2339, beginning on line 17, as follows:

> However, care must be taken when selecting γ, β, and $\tilde{r}_k$ to ensure convergence in 4D-Var. Use of these weights may imply that residual errors do not fit a Gaussian distribution.

Page 2343, line 21 does not state the results are more favorable and we agree that such an assessment cannot be made until after a true inversion. As robust may be a strong word choice, we have made the following changes:

GMDD
8, C801–C819, 2015

Abstract, page 2314, starting on line 11: "A cost function weighting scheme was devised to increase adjoint sensitivity consistency in future inverse modeling studies."

Page 2343, starting on line 21: "The increased burning sensitivity magnitudes indicate the weighting scheme is successful at generating a cost function that is more consistently sensitive to emission perturbations across multiple model configurations."

Page 2345, starting on line 10: "Results indicate that the weighting scheme is effective at generating consistent sensitivities of the cost function to emissions across multiple model configurations."

We believe adjoint sensitivity consistency is demonstrated. In the forward model simulation, both the SLAB and PX LSM schemes result in a net negative model bias. Because the anthropogenic sensitivities are net positive in both cases, and we assume the emissions are the only source of bias in the demonstration, then the burning emissions must be under-predicted. Without the weighting scheme, the two configurations disagree about the sign of burning emission perturbations that would reduce model bias. With the weights, the sensitivities between SLAB and PX LSM schemes agree that burning emissions should be increased. An inversion using the weighting scheme would not necessarily result in more correct emissions, but such an assessment could be made after respective 4D-Var inversions.

2 References

- P. Bergamaschi, C. Frankenberg, J. F. Meirink, M. Krol, M. G. Villani, S. Houweling, F. Dentener, E. J. Dlugokencky, J. B. Miller, L. V. Gatti, A. Engel, and I. Levin, "Inverse modeling of global and regional CH4 emissions using SCIAMACHY

GMDD

8, C801–C819, 2015

satellite retrievals," Journal of Geophysical Research: Atmospheres, vol. 114, no. D22, p. n/a–n/a, 2009.

- M. Bocquet, H. Elbern, H. Eskes, M. Hirtl, R. ?abkar, G. R. Carmichael, J. Flemming, A. Inness, M. Pagowski, J. L. Pérez Camaño, P. E. Saide, R. San Jose, M. Sofiev, J. Vira, A. Baklanov, C. Carnevale, G. Grell, and C. Seigneur, "Data assimilation in atmospheric chemistry models: current status and future prospects for coupled chemistry meteorology models," Atmospheric Chemistry and Physics Discussions, vol. 14, no. 23, pp. 32233–32323, Dec. 2014.
- 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.
- T. L. B. Yelverton, M. D. Hays, B. K. Gullett, and W. P. Linak, "Black Carbon Measurements of Flame-Generated Soot as Determined by Optical, Thermal-Optical, Direct Absorption, and Laser Incandescence Methods," Environmental Engineering Science, vol. 31, no. 4, pp. 209–215, Apr. 2014.
- X. Zhang, X.-Y. Huang, and N. Pan, "Development of the Upgraded Tangent Linear and Adjoint of the Weather Research and Forecasting (WRF) Model," Journal of Atmospheric and Oceanic Technology, vol. 30, no. 6, pp. 1180–1188, Feb. 2013.

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

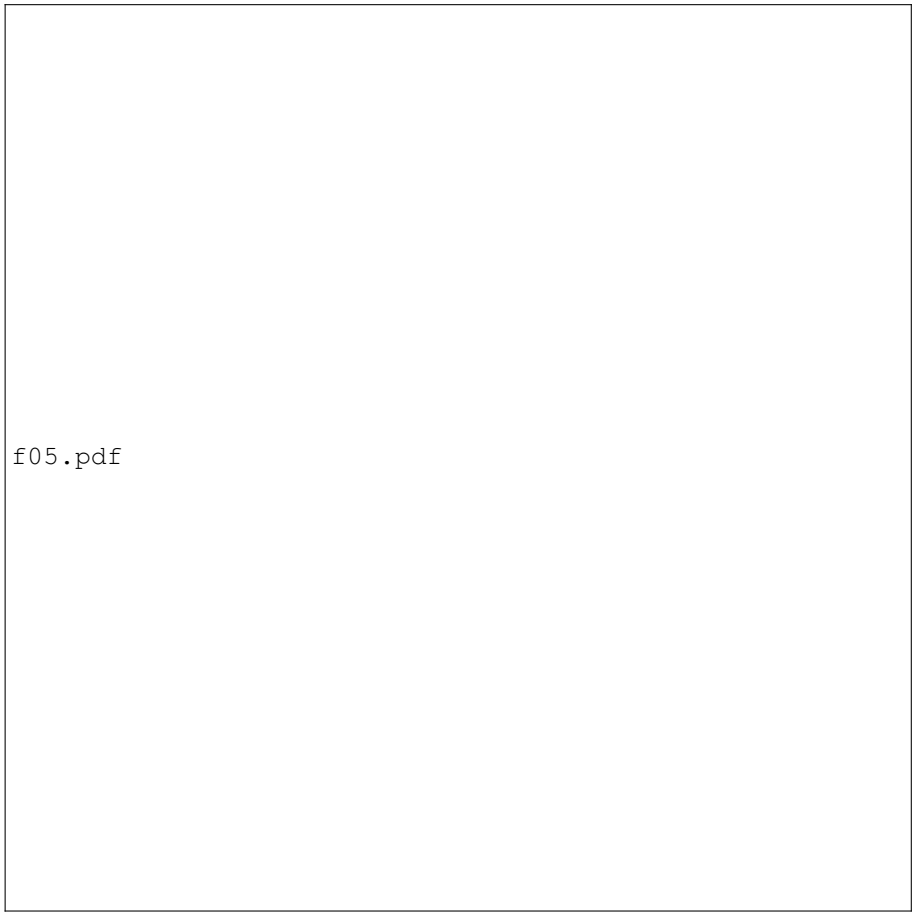

Fig. 1. Fully normalized time variant sensitivities calculated with the TLM with second order checkpointing and with multiple finite difference perturbation sizes. Each plot is for a single pair of source and receptor locations, q and p.

GMDD
8, C801–C819, 2015

f11.pdf

Fig. 2. Normalized sensitivities ($\frac{\partial ln J}{\partial ln E_{i,j,d}}$) of the 4D-Var cost function (for surface and aircraft observations) with respect to anthropogenic and burning emission scaling factors overlaid on MODIS Aqua true color images for six days during the simulation. Anthropogenic sensitivities with magnitudes less than 1 % of the maximum anthropogenic sensitivity magnitude are removed. There is a marker for all grid cells with non-zero burning emissions.

GMDD
8, C801–C819, 2015

GMDD
8, C801–C819, 2015

Interactive Comment

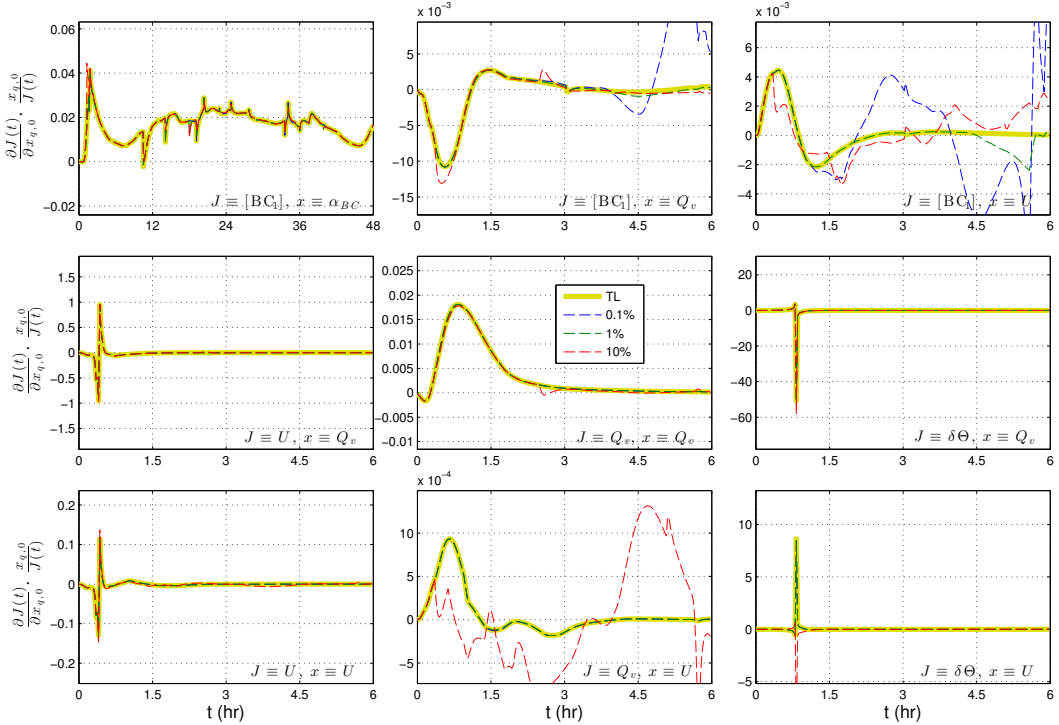

Fig. 3. Fully normalized time variant sensitivities calculated with the TLM with second order checkpointing and with multiple finite difference perturbation sizes. Each plot is for a single pair of source an

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

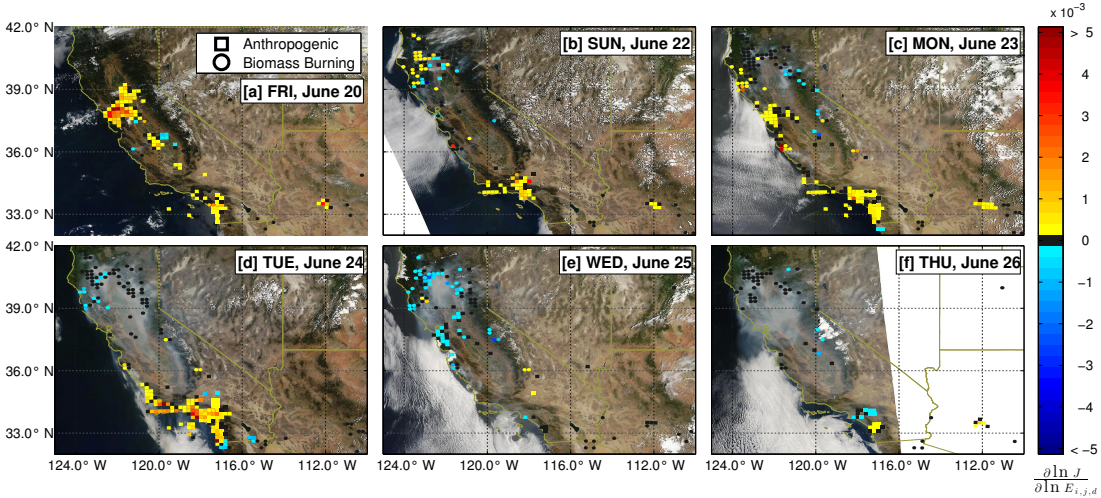

Fig. 4. Normalized sensitivities ($\frac{\partial \text{ln }J}{\partial \text{ln }E_{i,j,d}}$) of the 4D-Var cost function (for surface and aircraft observations) with respect to anthropogenic and burning emi

GMDD
8, C801–C819, 2015

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

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

Geoscientific Model Development Discussions
Open Access

GMDD
8, C820–C826, 2015

Interactive Comment

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, $\sim$) 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.

GMDD

8, C820–C826, 2015

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 $m\,s^{-1}$) 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.

GMDD

8, C820–C826, 2015

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

GMDD
8, C820–C826, 2015

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.

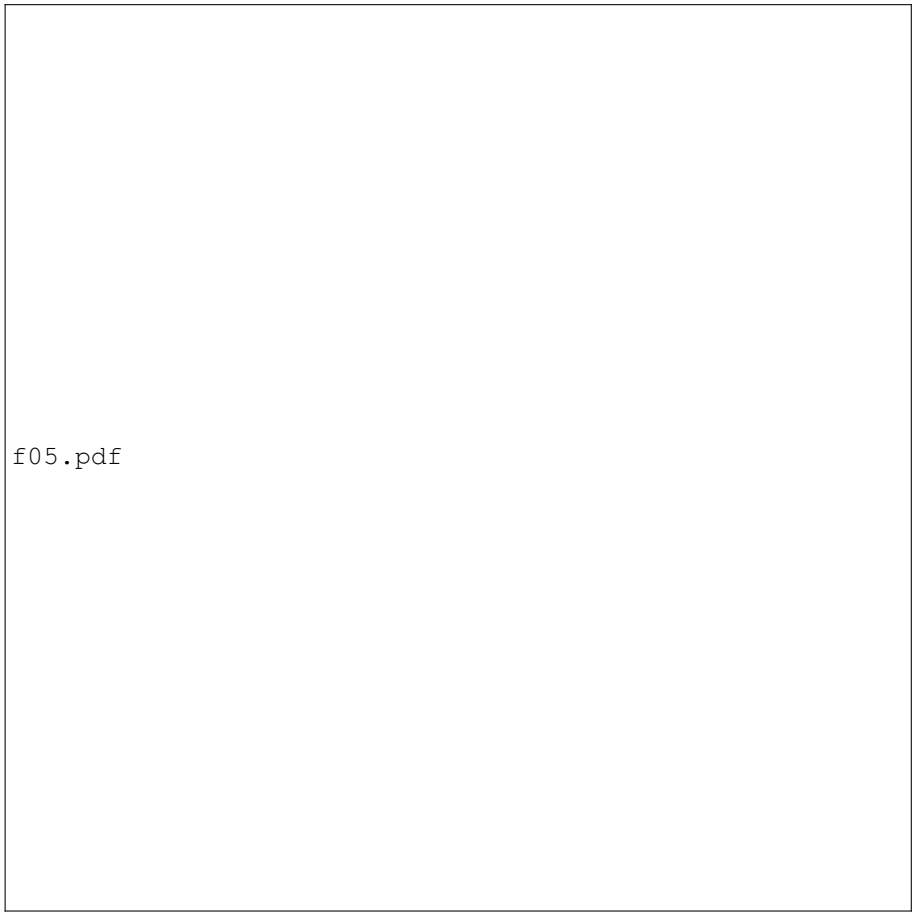

Fig. 1. Fully normalized time variant sensitivities calculated with the TLM with second order checkpointing and with multiple finite difference perturbation sizes. Each plot is for a single pair of source and receptor locations, q and p.

GMDD
8, C820–C826, 2015

GMDD
8, C820–C826, 2015

Interactive Comment

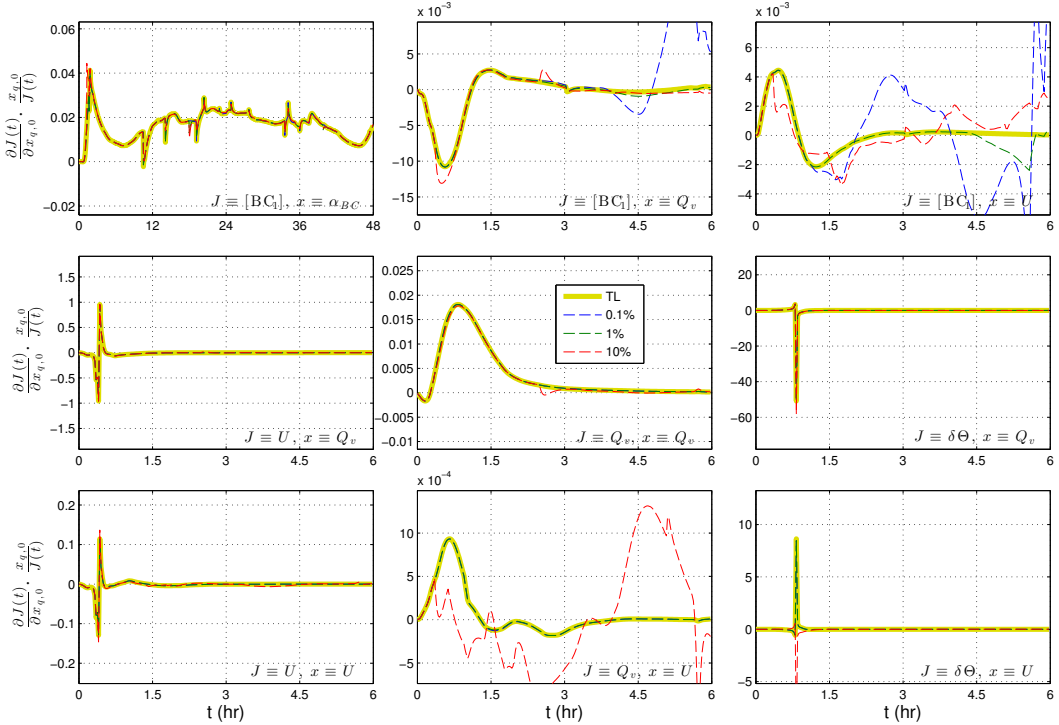

Fig. 2. Fully normalized time variant sensitivities calculated with the TLM with second order checkpointing and with multiple finite difference perturbation sizes. Each plot is for a single pair of source an

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

Manuscript prepared for Geosci. Model Dev. Discuss.
with version 2014/07/09 7.01 Copernicus papers of the LaTeX class copernicus.cls.
Date: 19 May 2015

Development and application of the WRFPLUS-Chem online chemistry adjoint and WRFDA-Chem assimilation system

J. J. Guerrette and D. K. Henze

Department of Mechanical Engineering, University of Colorado, Boulder, CO, 80309, USA

Correspondence to: J. J. Guerrette (jonathan.guerrette@colorado.edu)

Abstract

Here we present the online meteorology and chemistry adjoint and tangent linear model, WRFPLUS-Chem, which incorporates modules to treat boundary layer mixing, emission, aging, dry deposition, and advection of black carbon aerosol. We also develop land surface and surface layer adjoints to account for coupling between radiation and vertical mixing. Model performance is verified against finite difference derivative approximations. A second order checkpointing scheme is created to reduce computational costs and enable simulations longer than six hours. The adjoint is coupled to WRFDA-Chem, in order to conduct a sensitivity study of anthropogenic and biomass burning sources throughout California during the 2008 Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS) field campaign. A cost function weighting scheme was devised to ~~increase adjoint sensitivity robustness~~ reduce the impact of statistically insignificant residual errors in future inverse modeling studies. Results of the sensitivity study show that, for this domain and time period, anthropogenic emissions are over predicted, while wildfire ~~emissions are under predicted~~emission error signs vary spatially. We consider the diurnal variation in emission sensitivities to determine at what time sources should be scaled up or down. Also, adjoint sensitivities for two choices of land surface model indicate that emission inversion results would be sensitive to forward model configuration. The tools described here are the first step in conducting four-dimensional variational data assimilation in a coupled meteorology-chemistry model, which will potentially provide new constraints on aerosol precursor emissions and their distributions. Such analyses will be invaluable to assessments of particulate matter health and climate impacts.

1 Introduction

Fine particulate matter impacts human health (Schwartz et al., 2007; Krewski et al., 2009) and climate (Myhre et al., 2013). Atmospheric climate forcing from aerosols is potentially large, but also highly uncertain owing to a complex spatial-temporal distribution of con-

centration, mixing state, and particle size for multiple species, each emitted from varying precursor sources, both anthropogenic and natural (Textor et al., 2006; Schulz et al., 2006). Depending on the species and quality of records, a nation's annual aerosol precursor and primary emissions have uncertainties anywhere between 7 % and a factor of four, with larger variation on seasonal to diurnal scales for particular sectors (Streets et al., 2003; Suutari et al., 2001). Over these shorter time scales, aerosols impact meteorology through the semi-direct (Hansen et al., 1997; Koch and Del Genio, 2010) and indirect (Twomey, 1977; Lohmann and Feichter, 2005) cloud effects, which are both dependent on aerosol vertical profiles (e.g., Samset et al., 2013) governed by mixing.

Atmospheric models are used to improve our understanding of aerosol sources, distributions, and processes. Online numerical weather prediction and chemistry (NWP-chemistry) models integrate dynamic and chemical equations simultaneously, whereas offline chemical transport models (CTMs) interpolate meteorological fields from 3 to 6 h reanalyses. Grell et al. (2004) ~~showed~~ 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 of chemical tracers is highly sensitive to the choice of online ~~vs.~~ versus offline modeling methodologies due to variations in boundary layer mixing strength. Additionally, NWP-chemistry models account for moisture and temperature perturbations to dynamics due to aerosol microphysics and radiative forcing, while CTMs can not account for these feedbacks.

There are numerous online models with aerosol-meteorology feedbacks (e.g., WRF-Chem, COSMO-ART (Vogel et al., 2009), GEM-AQ (Kaminski et al., 2008), and IFS-MOZART (Kinnison et al., 2007; Flemming et al., 2009; Morcrette et al., 2009)). Better descriptions of sources, loss mechanisms, and vertical transport in coupled models are needed to increase accuracies in short-term climate modeling (Baklanov et al., 2014). To address this, chemical data assimilation can be used to improve short-term forecasts. Bocquet et al. (2014) review methods and applications of chemical data assimilation in CTMs and NWP-chemistry models. In WRF, three-dimensional variational data assimilation (3D-Var) (Pagowski et al., 2010; Liu et al., 2011; Schwartz et al., 2012; Saide et al., 2012,

2013), ensemble Kalman filter (EnKF) (Pagowski and Grell, 2012), and hybrid approaches (Schwartz et al., 2014) have all been used to improve chemical initial conditions. The limitation of these studies, using sequential methods, has been the decay of chemical concentrations back to the emissions-driven values following the characteristic loss rate of each species, necessitating periodic reinitialization with new observations. Using data assimilation solely to perturb initial conditions leaves behind underlying deficiencies in model description, emissions, or other input parameters.

In contrast to 3-D approaches, 4-D data assimilation attempts to minimize the discrepancy between model predicted values and observations at the same time observations are acquired. Variational 4-D data assimilation (4D-Var) requires an adjoint, which calculates the sensitivity of a model metric to all input parameters, such as resolved aerosol precursor emissions. Several offline CTMs already have adjoints for constraining aerosol and aerosol precursor emissions, including GEOS-Chem (Henze et al., 2007), STEM (Sandu et al., 2005; Hakami et al., 2005), CMAQ (Turner et al., 2015), GOCART (Dubovik et al., 2008), and LMDz (Huneeus et al., 2009). Inverse modeling has been used to constrain aerosol emissions with 4D-Var, but only in offline models (e.g., Hakami et al., 2005; Dubovik et al., 2008; Henze et al., 2009; Wang et al., 2012). In addition to inverse modeling, derivatives calculated from CTM adjoints have been used to analyze sensitivities of model estimates to emissions (e.g., Turner et al., 2012). Online chemical 4-D variational data assimilation (4D-Var) has been performed with the global IFS-MOZART model, although without two-way coupling, to improve aerosol (Benedetti et al., 2009) and gas-phase (Inness et al., 2013) initial conditions. To our knowledge, 4D-Var still has not been used in a regional NWP-chemistry model with online coupling to constrain aerosol precursor emissions or other important model parameters, such as vertical mixing coefficients.

Here we present the first such system, building on existing capabilities of the WRF data assimilation (WRFDA) framework. WRFDA includes both 3D-Var (Barker et al., 2004) and incremental 4D-Var (Barker et al., 2005; Huang et al., 2009) algorithms, which are designed for constraining meteorological initial conditions (e.g., wind fields, temperature, moisture). For WRFDA v3.2 and later, ~~WRF-4-DVar~~ WRF-4DVar requires calling the WRFPLUS for-

ward (FWM), tangent linear (TLM) and adjoint (ADM) models. These models include adiabatic WRF dynamics, along with simplified surface friction (i.e., boundary layer), cumulus, and microphysics packages (Zhang et al., 2013). Here we integrate aerosol chemistry and vertical mixing from WRF-Chem into WRFPLUS, including complementary TLM and ADM components. While existing CTMs are capable of aerosol emission inversions, this development promises to introduce new insights into meteorology-chemistry couplings. We apply this system to black carbon (BC) aerosol, because of its important implications for climate (Bond et al., 2013) and health (Grahame et al., 2014). Additionally, the widespread use and development of WRF furthers the potential for continued model improvement and a community of future users.

BC is emitted from incomplete combustion of fuels. Major anthropogenic sources include residential cookstoves in developing countries, open crop burning, diesel transportation, and coal power plants with poor emission controls. Wildfires, or biomass burning, are the largest natural source. The major limitations to devising accurate bottom-up emissions inventories are poor activity data in developing countries and difficulty parameterizing complex biomass burning sources. Even in developed countries, changing economic landscapes affect real year-to-year emissions. ~~Black carbon (BC)~~ BC is unique among atmospheric aerosols as being radiatively absorptive, relatively inert, primary emitted, and having potentially complex cloud interactions. BC is possibly the second most important human emitted pollutant in terms of climate forcing in the present-day atmosphere, with a net forcing of $+1.1\,\mathrm{W\,m^{-2}}$, but with 90 % uncertainty ($+0.17$ to $+2.1\,\mathrm{W\,m^{-2}}$) (Bond et al., 2013). Also, reductions in BC emissions have been shown to reduce fine particulate health impacts (e.g., Anenberg et al., 2011).

The new TLM and ADM – referred to collectively herein as "AD/TL models" – aerosol treatments lay the groundwork for constraining aerosol precursor emissions using 4D-Var in a NWP-chemistry model. In Sect. 2, we describe the WRFPLUS-Chem and WRFDA-Chem model architectures. In Sect. 3, we describe the construction and verification of the AD/TL models of specific WRF-Chem forward model components. In Sect. 4, we describe a special checkpointing scheme that enables adjoint and tangent linear simulations longer

than 6 h which are required for accumulating sensitivities of sparse chemical observations with respect to emissions. In Sect. 5, we demonstrate the capability of the adjoint model to calculate sensitivities of BC observation errors in WRFDA-Chem. Finally, we discuss future developments for WRFPLUS-Chem and WRFDA-Chem.

2 Methods

Creating the foundation for WRFDA-Chem required managing relationships between five related, but separate models. These include ~~the~~ (1) ~~Weather Research and Forecast Model (WRF)~~WRF, (2) its "-Chem" variant, and the (3, 4) WRFPLUS AD/TL models. Finally, (5) WRFDA 4D-Var requires communication of critical namelist and state variables to the FWM, TLM, and ADM. ~~Figure ?? shows the relationships between these different models, including all~~ Table 1 lists the WRF-Chem components that previously had AD/TL ~~code that was previously developed , and code that we have added, modified, or plan to add~~descriptions in WRFPLUS, those with new code developed for WRFPLUS-Chem, and those that need future development to enable fully coupled chemical 4D-Var.

2.1 Forward model

For this work, we use WRF version 3.6. The WRFPLUS-Chem code repository (https://svn-wrf-model.cgd.ucar.edu/branches/WRFPLUSV3-Chem) contains the most current version. Interested users can contact NCAR ~~to request user~~ or the authors for access to the code. WRF contains multiple non-hydrostatic dynamic cores and parameterization options for modeling unresolved physical processes. The FWM is identical in WRF and WRFPLUS~~, though typically only very simple~~ . The simplified treatments for unresolved physics are ~~applied in WRFPLUS~~typically only used in the AD/TL models. In addition, WRF-Chem simulates the emission, deposition, transport, turbulent and cumulus mixing, wet scavenging, cloud interactions, and chemical transformation of trace gasses and aerosols. All of these processes are modeled at the same spatial and temporal resolution, which enables coupling WRF radiation and microphysics calculations directly with chemical processes.

The forward model configuration for which we have developed the corresponding TLM and ADM will be referred to as the "adjoint model configuration," because we use the same settings when running the adjoint. We use GOCART aerosols (chem_opt = 300), wherein the chem array has 19 aerosol (e.g., SO_2, sulfate, black carbon, dust, sea salt) and zero gas-phase members. This option includes bulk mass sulfate chemistry and black carbon oxidative aging. We employ combined local and non-local ACM2 PBL mixing (Pleim, 2007b, a), with surface interactions handled by the Pleim-Xiu (PX) LSM (Xiu and Pleim, 2001; Pleim and Xiu, 2003; Pleim and Gilliam, 2009) and surface layer (Pleim, 2006) mechanisms (all options seven). Soil moisture and temperature nudging are not used within the PXLSM. Prior to version 3.6, the WRF-Chem vertical mixing scheme solely carried out PBL mixing and dry deposition for chemical species. That vertical mixing depended on a (local) turbulent eddy mixing coefficient from a user-selected PBL scheme and a dry deposition velocity. There is new capability to calculate tracer turbulent mixing and dry deposition within the ACM2 subroutine itself, enabling non-local mixing. Trace gas and particle deposition velocities are calculated using characteristic resistances found using methods from Wesely (1989). Microphysics and radiation AD/TL models with aerosol feedbacks have not been incorporated into WRFPLUS-Chem yet. These crucial components will be partially adapted from previous work (e.g. Saide et al., 2012, 2013), while others still need to be developed. Both microphysics and radiation are turned off for Sect. 3.3 verification simulations. In order to ensure appropriate radiative fluxes at the land-air boundary, the GSFCSW and Goddard LW radiation compute ground-incident radiation for the Sect. 5 adjoint sensitivity demonstration. However, online coupling between radiation and chemical species is deactivated.

2.2 Incremental 4D-Var

WRFDA uses an incremental 4D-Var method (Courtier et al., 1994) for finding the minimum of the cost function, J, by adjusting control variables (CV), $\boldsymbol{x}$. As described by Huang et al. (2009), the WRFDA cost function has three terms

$$J = J_{\mathrm{b}} + J_{\mathrm{o}} + J_{\mathrm{c}}, \tag{1}$$

where J_b, J_o, and J_c are the background, observation, and balancing cost functions, respectively. J_c is not relevant to the current work. The background and observation cost functions are

$$J_\mathrm{b} = \frac{1}{2}\left[(\boldsymbol{x}^n - \boldsymbol{x}^{n-1})\boldsymbol{\delta x} + \sum_{i=1}^{n-1}\left(\boldsymbol{x}^i - \boldsymbol{x}^{i-1}\right)\right]^\top \mathbf{B}^{-1} \left[(\boldsymbol{x}^n - \boldsymbol{x}^{n-1})\boldsymbol{\delta x} + \sum_{i=1}^{n-1}\left(\boldsymbol{x}^i - \boldsymbol{x}^{i-1}\right)\right] \tag{2a}$$

and

$$\begin{aligned} J_\mathrm{o} &= \frac{1}{2}\sum_{k=1}^{K}\left\{H_k\left[M_k\left(\boldsymbol{x}^n\right)\right] - \boldsymbol{y}_k\right\}^\top \mathbf{R}_k^{-1} \left\{H_k\left[M_k\left(\boldsymbol{x}^n\right)\right] - \boldsymbol{y}_k\right\} \\ &\approx \frac{1}{2}\sum_{k=1}^{K}\left[\mathbf{H}_k\mathbf{M}_k(\boldsymbol{x}^n\boldsymbol{\delta x} - \boldsymbol{x}^{n-1}) - \boldsymbol{d}_k\right]^\top \mathbf{R}_k^{-1} \left[\mathbf{H}_k\mathbf{M}_k(\boldsymbol{x}^n\boldsymbol{\delta x} - \boldsymbol{x}^{n-1}) - \boldsymbol{d}_k\right]. \end{aligned} \tag{2b}$$

The background cost function is a penalty term, which ensures the departure of the posterior, $\boldsymbol{x}^n$, from the prior, $\boldsymbol{x}^0 = \boldsymbol{x}^\mathrm{b}$, remains within the bounds justified by the background error covariance, $\mathbf{B}$. The observation cost function measures the distance between the 4D-Var model solution, $\boldsymbol{x}^n$, and the observations, $\boldsymbol{y}$. M and H are the nonlinear model and observation operators, while $\mathbf{M}$ and $\mathbf{H}$ are their linearized forms, or tangent linear operators, used to propagate analysis increments $\boldsymbol{\delta x} = \boldsymbol{x}^n - \boldsymbol{x}^{n-1}$ from ~~the earliest emission time~~ earlier simulation times to the kth observation. $\mathbf{R}$ is the observation error covariance matrix. The innovation,

$$\boldsymbol{d}_k = \boldsymbol{y}_k - H_k\left[M_k\left(\boldsymbol{x}^{n-1}\right)\right], \tag{3}$$

is the residual error between the real and modeled observations k at the end of 4D-Var iteration $n-1$. ~~This notation slightly differs from Huang et al. (2009), who employed K observation windows, each containing multiple observations.~~

For each iteration of incremental 4D-Var, the model is linearized about a trajectory, which is a collection of stored values of all model state variables at all time steps within the assimilation window. This trajectory enables propagation of sensitivities forward and backward in time within the TLM and ADM. Each of these models are called in an inner loop to calculate the gradient of the observation cost function, $\nabla_{\boldsymbol{x}} J_{\mathrm{o}}$. An optimization algorithm uses the gradients to calculate optimal analysis increments to the CVs, which minimize the observation cost function. If the CVs, $\boldsymbol{x}^n$, depart too much from the initial guess for the current outer loop iteration, $\boldsymbol{x}^{n-1}$, the model must be relinearized about the new state, $\boldsymbol{x}^n$, using M. The purpose of the two-level optimization is that approximating M with $\mathbf{M}$ ~~transforms the cost function from a nonlinear~~ simplifies the full problem to a quadratic ~~form~~problem, and guarantees a unique solution ~~$\boldsymbol{x}^*$~~ to the minimization (Courtier et al., 1994). Refer to Huang et al. (2009) for more details on the WRFDA incremental method, including a full expression for $\nabla_{\boldsymbol{x}} J$ given by Eq. (7) of that article. The main purpose of this work is to introduce the AD/TL model components of WRFPLUS-Chem.

3 Tangent linear and adjoint model construction and verification

We have developed and tested adjoint and tangent linear code to represent aerosol-relevant processes in WRFPLUS-Chem. This development required a four step process:

1. Automatically differentiate specific WRF-Chem modules using TAPENADE (Hascoët and Pascual, 2013) version 3.6.

2. Verify standalone TLM and ADM derivatives against finite difference approximations; debug as necessary.

3. Incorporate code manually into WRFPLUS.

4. Repeat step 2 for fully integrated WRFPLUS-Chem model.

TAPENADE takes discrete Fortran or C source code as input, then generates either TLM or ADM code using a user-generated list of independent and dependent parameters. In addition to creating the differential code, TAPENADE reduces adjoint computational cost by eliminating unnecessary lines of code. Similar to Xiao et al. (2008) and Zhang et al. (2013), integrating the automatically differentiated adjoint code into WRFPLUS required significant manual intervention and debugging. Methods for constructing discrete adjoints are well-documented (Giering and Kaminski, 1998; Hascoët and Pascual, 2013). For the remainder of this section, we discuss the particular mechanisms for which we have created AD/TL models, and then we provide verification results for WRFPLUS-Chem.

3.1 Transport mechanisms

PBL physics and dry deposition in a column are handled by ACM2. The simple surface friction previously developed for WRFPLUS does not perform vertical mixing of tracers, which is a minimum requirement of any PBL scheme used in WRFPLUS-Chem. The ACM2 PBL depends on ground-atmosphere interactions that necessitate additional surface layer and land surface model (LSM) AD/TL code. For example, the ACM2 PBL scheme depends on the friction velocity, U^*, calculated in a surface layer scheme, which itself depends on wind speed, and the state variables u and v. ACM2 also depends on surface heat (HFX) and moisture (QFX) fluxes, which can be calculated within the surface layer or LSM code, but also depend on U^*. The dependence of HFX and QFX on ground-incident shortwave radiation (GSW) is calculated in the LSM. GSW is calculated in the radiation scheme, and depends on the aerosol composition and atmospheric moisture phase and distribution. Because we have not developed radiation AD/TL code, this coupling is not represented in WRFPLUS-Chem yet. The dependencies themselves are illustrative of how ACM2, and indeed most any other PBL scheme available in WRF, is appropriate for representing chemistry-meteorology interactions critical to understanding short-term climate impacts from aerosols. ACM2 is compatible with the Monin-Obukhov and PX (options 91

and 7) surface layer options, as well as the SLAB and PX (options 1 and 7) LSM options. TLM and ADM code is developed for all of these choices, and have been tested in standalone verification tests. In the interest of brevity, complete model verification in Sect. 3.3 has been limited to the two PX options.

Advection of inert tracers was added to WRFPLUS by X. Zhang (2012, personal communication). The same treatment has been applied to the "chem" array, with additional checkpointing and parallel communications. We generated standalone TLM and ADM code for deep cumulus convection as handled by the Grell–Freitas cumulus scheme (Grell and Freitas, 2014). One of the major benefits of this cumulus scheme is the ability to use online calculated cloud condensation nuclei (CCN) to account for the effect of aerosols on liquid and vapor water mass fractions. These parameters directly impact convection, including tracer transport. The ability of the standalone AD/TL codes to produce the relevant members of the Jacobian has been verified for a single set of column conditions using similar methods as described in Sect. 3.3. However, the FWM, TLM, and ADM do not yet account for vertical transport of chemical tracers, and thus have not been integrated into WRFPLUS-Chem.

3.2 Aerosol-specific components

GOCART is a bulk aerosol scheme that treats reactive species (BC, OC, sulphate) using a total mass approach and divides non-reactive species (dust, sea salt) into multiple size bins (Chin et al., 2000). Oxidative aging for both BC and OC is handled by a first-order decay from hydrophobic to hydrophilic forms using a time constant of 2.5 days. Sulphate (SO_4^{2-}) is produced from SO_2 and dimethyl sulfide precursor gases in GOCART. Sulphate chemistry also requires offline-calculated values for nitrate and OH radical, which are taken from climatologies available from the PREP-CHEM-SRC preprocessor (Freitas et al., 2011). WRFPLUS-Chem includes both the carbon and sulfate chemistry AD/TL codes, but only the BC component is tested and applied here.

Emissions of aerosol precursors in WRF-Chem is a linear process corresponding to specific chemistry and emission inventory options. Emission magnitudes are calculated, then

distributed spatially and temporally, in offline preprocessors. Typically, emissions are read in hourly following some diurnal pattern. In order to make the emissions code easily differentiable, scaling factors are added to the emissions such that

$$\mathbf{E}_{c,i_{\text{sc}}} = \boldsymbol{\alpha}_{c,i_{\text{sc}}} \tilde{\mathbf{E}}_c. \tag{4}$$

At any simulation time, $\tilde{\mathbf{E}}_c$ are the emissions most recently read in from file for chemical species c. $\boldsymbol{\alpha}_{c,i_{\text{sc}}}$ and $\mathbf{E}_{c,i_{\text{sc}}}$ are the emission scaling factors and effective emissions, respectively, during scaling period i_{sc}. For emission inversions, the CVs, $\boldsymbol{x}$, are spatial-temporal resolved emission scaling factors. At the beginning of 4D-Var or during an adjoint sensitivity study, the scaling factors are set to unity. The scaling factors are applied in the FWM if environment option WRFPLUS $==1$ is set during compilation.

Dry deposition velocities are calculated in WRF-Chem within the dry deposition driver. In order to ease adjoint code construction and reduce checkpointing requirements, the dry deposition velocity calculation is moved to immediately precede the PBL driver as depicted in Fig. **??**. The new source code is similar to the dry deposition driver, except that only code corresponding to the GOCART aerosol option remains. The dry deposition AD/TL code accounts for dependencies of the dry deposition velocity on physical parameters (e.g., temperature, water vapor, U^*). As mentioned previously, the chemical concentrations are sensitive to dry deposition velocity within the PBL scheme.

3.3 Verification and linearity test

WRFPLUS FWM, TLM, and ADM performance were previously verified by Zhang et al. (2013). Here we use an alternative ~~verification approach~~ approach based on Taylor series derivative approximations, and similar to that used by, e.g., Henze et al. (2007)~~. We~~ , to verify WRFPLUS-Chem. We define a new cost function equal to a single predicted state variable, locally defined in grid cell p and at the end of time step f, $J = SV_{p,f}$. We use the TLM, ADM, and a centered finite difference approximation from the FWM to evaluate

derivatives

$$\chi_{p,q} = \frac{\partial J_{p,f}}{\partial x_{q,0}} \frac{\partial J}{\partial x_{q,0}}, \tag{5}$$

~~of some cost function at location p and time step f~~ with respect to some CV at location q and the initial time, 0. The finite difference derivatives are calculated from

$$\chi^{\mathsf{NL}}_{p,q} \approx \frac{J_{p,f}\left(x_{q,0}+\delta x\right) - J_{p},f\left(x_{q,0}-\delta x\right)}{2\delta x} \frac{J\left(x_{q,0}+\delta x\right) - J\left(x_{q,0}-\delta x\right)}{2\delta x}, \tag{6}$$

where each evaluation of J results from a FWM ~~evaluation~~ simulation with some perturbed value of $x_{q,0}$. δx varies between 0.1 and 10 % of the value of $x_{q,0}$. The adjoint and tangent linear derivatives are found by forcing the model gradient fields, λ^* and λ, at ~~J_p~~ J and x_q, respectively. The tangent linear gradient and adjoint gradient variables are analogous to state variables in the FWM. We force gradients of 1.0, indicating a 100 % perturbation of the variable, and the resulting derivatives are retrieved from the model output gradient fields, such that

$$\chi^{\mathsf{TL}}_{p,q} = \lambda_{p,f} = \mathbf{M}\left(\lambda_{q,0}\right) \tag{7}$$

and

$$\chi^{\mathsf{AD}}_{p,q} = \lambda^*_{q,0} = \mathbf{M}^{\top}\left(\lambda^*_{p,f}\right), \tag{8}$$

where $\mathbf{M}^{\top}$ is the adjoint operator.

In order to evaluate our additions to WRFPLUS-Chem, we test cost functions equal to hydrophobic (BC_1) and hydrophilic (BC_2) black carbon concentrations in 100 different grid cells. We evaluate derivatives with respect to five ~~state variables~~ CVs at three initial locations for each of those 200 cost functions. The CVs include BC emission scaling factors (α_{BC}) and initial conditions for BC_1, zonal wind (U~~, T, and~~), perturbation potential

temperature ($\delta\Theta$), and water vapor mixing ratio (Q_v, ~~and also BC emission scaling factors, α_{BC}~~). All sensitivities apply over a 3 h duration for a domain covering the southwest United States. For a full domain and model setup description refer to Sect. 5.1.1. Figure 1 shows that the maximum relative error between the TLM and ADM is in the 8th significant digit. Thus we only need to compare the nonlinear model to the TLM to verify both the TLM and ADM. Those results are given in Fig. 2. The slope and R^2 statistic for a linear fit of those comparisons are very nearly unity for all CVs tested. Each of the plots in Figs. 1 and 2 depicts 600 derivative evaluations. A range of finite difference perturbations δx is used for U, ~~T~~$\delta\Theta$, and Q_v control variables in order to find a value of χ^{NL} with the best compromise between truncation and roundoff error. We test derivatives with respect to meteorological variables in order to show the AD/TL models will be functional in a setting with coupled chemistry and physics. In such a system, the emissions will impact meteorology, which in turn impacts concentrations. These results illustrate the capability of the AD/TL models to represent the latter part of that relationship. All of the verification results apply to a three hour simulation period, but longer simulations are needed to calculate the average influence of emissions on the modeled state-space.

4 Second order checkpointing

As discussed in Sect. 2.2, the nonlinear model trajectory is an integral component for propagating gradients in the AD/TL models. As one might imagine, the trajectory contains a large amount of information. WRFPLUS stores the entire double precision trajectory in memory in order to eliminate expensive I/O time. This is very helpful with regards to storage, but presents a challenge in terms of memory. The system is designed for 6 h operational assimilation windows. In a typical WRFPLUS-Chem simulation there are at least twenty-eight 3-dimensional state variables (8 physical, 1 to 3 moisture, and 19 GOCART species), and numerous other 2- and 1-dimensional state variables that must be included in the trajectory. For an illustrative domain, simulating 3 h with a 90 s time step (18 km resolution), 79×79 columns, 42 levels, ~~and~~ a 5 cell boundary width, and 28 3-D state variables, the trajec-

tory would require 1.46 GB per core on 64 cores. This final cost per core includes a 50 % storage growth per doubling of the number of cores. Because the trajectory is stored for all time steps, required memory scales linearly with simulation duration and the number of simulated chemical species. For multi-day and multi-week inversions, as is typical in non-operational chemical data assimilation, the memory requirements become impractical for most cluster computing systems.

To solve this problem we implement a second order checkpointing scheme that shares the storage burden between the hard disks and memory. In a standard WRFPLUS adjoint simulation, the FWM is called first in order to calculate the trajectory. The FWM integrates the nonlinear equations from the initial to the final time, and stores the model trajectory at each time step. The ADM integrates the transpose of the linearized model equations backward in time, and at each time step reads the trajectory previously stored by the FWM. This process is depicted as "1st-order checkpoint" in Fig. 3. Since the storage limitation is driven by the duration of a simulation, we break the simulation into smaller segments, while maintaining continuity in the adjoint derivatives. The checkpointed adjoint simulation begins with a full FWM simulation beginning at the initial time, t_0, and ending at the final time, t_f. WRF restart files are written at time intervals equal to the checkpoint interval, Δt_c. Once the simulation is completed, the FWM is restarted at initial time equal to $t_\mathrm{f} - \Delta t_\mathrm{c}$. During that simulation, the trajectory is stored in memory. The trajectory is then recalled in an adjoint simulation that proceeds backward toward the current initial time. The checkpoint system alternately calls the FWM and ADM until returning to t_0. The major hurdle to integrating this second order checkpointing system into WRF-4-DVar is that the trajectory is no longer readily available to WRFDA for calculating modeled observations, $H_k\left[M_k\left(\boldsymbol{x}^n\right)\right]$, between the calls to the forward and adjoint models. Instead, these values must be calculated during either the full FWM (Step 1) or checkpoint FWM (Steps 2, 4, 6, etc.) simulations. We take the former approach. A similar checkpointing system is also implemented for the TLM in order to enable long duration incremental 4D-Var.

In order to ensure the checkpointing method delivers consistent derivatives to the non-checkpointed version, we again compare AD/TL derivatives to finite difference approxima-

tions. Because of the wall time required to calculate derivatives across extended time periods, we limit our tests to fourteen pairs of initial and final locations, q and p. For all of the J and x pairs tested in Sect. 3.3, the ADM and TLM agree to 13 or more digits over a 9 h test. The improved performance relative to the previous 3 h test came about after ~~a few minor bug fixes~~increasing the precision of several variables in the TLM dry deposition subroutine. Because of this machine precision AD/TL agreement, we only compare the finite difference approximations to the TLM. For these checkpointed simulations, we analyze the derivative of a time variant cost function with respect to multiple control variables

$$\chi_{p,q}(t) = \frac{\partial J_p(t)}{\partial x_{q,0}}. \tag{9}$$

Doing so ensures that the derivatives are continuous across multiple checkpoint intervals and we are able to see the transient behavior of multiple finite difference perturbation sizes at times when there are large discrepancies with the TLM. The finite difference approximations of derivatives of BC with respect to the physical variables grow more unstable with time. Thus, we calculate those derivatives only for a 6 h period, while we test derivatives with respect to emissions for 48 h. Here we also include derivatives of U, ~~T~~ $\delta\Theta$ and Q_v with respect to U and Q_v to ensure that those relationships are represented properly in the surface layer, LSM, and PBL AD/TL schemes, so that they may be used in a meteorological 4D-Var setting.

Figure 4 shows the resulting derivatives for nine different pairs of J and x for a single pair of q and p. Most importantly for multi-day 4D-Var emissions inversions, and as would be expected, BC concentrations respond linearly to a 1 % perturbation of emissions for at least 48 h. Next, it becomes apparent why derivatives with respect to U and Q_v require multiple finite difference perturbation sizes to ensure one of them matches the TLM at a particular cost function evaluation time. There are times when either the smallest, largest, or no value for δx agrees with the TLM. However, the TLM has inflection points at the same times as the finite difference approximations, including during ~~periods of intense oscillation~~fast transient periods, such as for $\frac{\partial U}{\partial U}$ and $\frac{\partial U}{\partial Q_v}$. The ~~chemical concentrations respond nonlinearly to all~~

duration over which the tangent linear assumption is valid for chemical responses to U and Q_v ~~perturbation sizes for periods longer than 1in the plots shown, and longer than 3for all test scenarios considered~~depends on the size of the perturbation and on the local meteorological regime. For instance, the test results shown here are for a response location very near the California coast, but better agreement was found farther inland. Further testing of ~~these~~ the coupled derivatives will be necessary to determine over what time period they are suitable for inverse modeling, and under what conditions the model nonlinearities cease to be a limiting factor. Future ~~emission inversion work~~ inversions with coupled physics and chemistry will need to verify that $\frac{\partial J}{\partial \alpha}$ has a near linear response over the time frame considered. The behaviors noted here are ~~consistent~~ similar or improved across the other thirteen pairs of q and p.

5 Sensitivities to BC emissions in California

Here we demonstrate the new WRFPLUS-Chem capabilities in an adjoint sensitivity study. For the present example, the 4D-Var cost function is the model response metric and the biomass burning, and weekday and weekend anthropogenic emissions are the model parameters of interest. This framework is used to analyze where and when these parameters most impact the model performance and are thus in need of improvement.

5.1 Approach

For this demonstration, we calculate the sensitivity of the 4D-Var cost function in the first iteration. The background term is zero and there has been no prior CV increment (i.e., $\boldsymbol{\delta x} = 0$). Therefore, the cost function, Eq. (1), simplifies to

$$J = \frac{1}{2}\sum_{k=1}^{K}\{H_k\left[M_k\left(\boldsymbol{x}_b\right)\right] - \boldsymbol{y}_k\}^{\top}\mathbf{R}_k^{-1} \{H_k\left[M_k\left(\boldsymbol{x}_b\right)\right] - \boldsymbol{y}_k\}. \quad (10)$$

All off-diagonal covariances in **R** are assumed to be zero in order to enable timely matrix inversion.

5.1.1 Model configuration

The model domain encompasses California and other southwest US states from 20 June 2008, 00:00:00 UTC to 27 June 2008, 09:00:00 UTC. We generated chemical initial conditions by running WRF-Chem for five days prior to the adjoint time period. We used the default WRF-Chem boundary condition for BC concentration of $0.02\,\mu g\,kg^{-1}$. This is consistent with a single upwind Pacific ocean transect taken during the 22 June flight. Meteorological initial and boundary conditions are interpolated from 3 h, 32 km North American Regional Reanalysis (NARR) fields. The horizontal resolution is 18 km throughout, and there are 42 vertical levels between the surface and model top at 100 hPa. The eta levels are 1.000, 0.997, 0.993, 0.987, 0.977, 0.967, 0.957, 0.946, 0.934, 0.921, 0.908, 0.894, 0.880, 0.860, 0.840, 0.820, 0.800, 0.780, 0.750, 0.720, 0.690, 0.660, 0.620, 0.570, 0.520, 0.470, 0.430, 0.390, 0.350, 0.310, 0.270, 0.230, 0.190, 0.150, 0.115, 0.090 , 0.07 , 0.052, 0.035, 0.020, 0.010, and 0.000. For a column where the ground is at sea level, there are 13 levels below 1 km and an additional 5 levels below 2 km. The subgrid physics options used are described in Sect. 2.1.

Anthropogenic emissions are taken from the US EPA's 2005 National Emissions Inventory (NEI2005). Fire emissions are provided by the Fire INventory from NCAR (FINN Version 1) (Wiedinmyer et al., 2011, 2006). FINN uses Moderate Resolution Imaging Spectroradiometer (MODIS) active fire locations and radiative power from NASA Terra and Aqua satellites, as well as speciated emission factors for four vegetation types, to calculate daily total 1 km resolution emissions. Burned areas are scaled to the combined fractional coverage of each 1 km^2 fire pixel by tree and herbaceous vegetation types assigned by the MODIS Vegetation Continuous Fields product (Hansen et al., 2003). Repeated fire detections in a single fire pixel are removed according to Al-Saadi et al. (2008). Plume rise injection heights are calculated in WRF-Chem by an embedded one-dimensional cloud-resolving model (Freitas et al., 2007, 2010; Grell et al., 2011).

5.1.2 Model-observation comparison

We compare the model to observations in individual time steps, which differs from previous data assimilation approaches with WRF. In the standard WRFDA 4D-Var architecture, observations are binned over intervals, or windows, typically of one hour or longer duration. Whereas WRFDA typically has k observation windows, here WRFDA-Chem and WRFPLUS-Chem handle k ~~observations possibly each at a different time~~ observation time steps, each of which might have multiple measurements available. In order to reduce memory requirements, the adjoint forcing is stored in a column array, instead of the 2-D and 3-D arrays that were required for each state variable for each window, k in WRFDA. Also, while WRFDA includes meteorological observation operators to be called offline, ~~a~~ the new fine temporal resolution observation operator ~~must be~~ is called directly within WRFPLUS. The traditional approach made communication between WRFDA and WRFPLUS less cumbersome, but also limited the ability to use dynamic observations recorded across broad temporal scales in an inversion.

In-situ observations were collected throughout California during the June 2008 portion of the Arctic Research of the Composition of the Troposphere from Aircraft and Satellites field campaign in collaboration with the California Air Resources Board (ARCTAS-CARB) (Jacob et al., 2009). Instruments aboard the DC-8 aircraft measured trace gas and aerosol concentrations over four days, including ~~elemental carbon (EC)~~ absorbing carbonaceous aerosol from the single particle soot photometer (SP2) at 10 s intervals (Sahu et al., 2012). Additionally, 41 Interagency Monitoring of Protected Visual Environments (IMPROVE) sites measured daily average surface light absorbing carbon (LAC) on 20, 23, and 26 June by thermal/optical reflectance (TOR) analysis of quartz filters (Malm et al., 1994). Surface and aircraft observation locations during the campaign are indicated in Figs. 5 and 6. The aircraft trajectories are overlaid on MODIS Aqua true color images (Gumley, 2008), and locations of MODIS active fires (NASA , 2014). While we use IMPROVE elemental carbon (EC) and SP2 absorbing carbon as equivalents herein, Yelverton et al. (2014) found that the former is approximately 7% higher than the latter, but that their error bars overlap. For the qualitative

analysis performed in this demonstration, bias correction would not change any of the final conclusions.

The observation operators for aircraft and surface observations require temporal averaging. The 10 s resolution ARCTAS observations of BC concentration, pressure, latitude, and longitude are averaged to the 90 s model time step, which is approximately the time the DC-8 would take to traverse a single 18 km × 18 km column. However, the 10 s resolution ARCTAS BC concentrations are revision 2 (R2), while a later revision 3 (R3) product was released at 60 s resolution only. The later revision includes additional mass in the 50–900 nm size range as a result of applying a lognormal fit. In order to utilize this improved product, as well as leverage the finer resolution observations, the 10 s BC mass is scaled by the mass ratio between the 60 s R3 and the 60 s average R2 datasets. The scaled 90 s average observations are compared directly with the nearest model grid cell so that the model values are not interpolated. The pressure measurements are compared to online model pressures to determine the model level of each observation. For 24 h average surface measurements from IMPROVE, the observation operator averages the nearest model surface grid cell concentration over all time steps within the observation period. For the few surface sites that have two air samplers simultaneously measuring, they are averaged together to prevent nonzero correlation in the cost function (i.e., off-diagonal terms in **R**). After all averaging, there are 995 aircraft observations and 107 surface observations.

As depicted in Fig. 7, the WRF-Chem simulation is, on average, biased low for both the surface and aircraft observations. The lowest biased aircraft observations tend to be at higher altitudes, although this is not true in all cases. There are many high biased observations, and they tend to be at lower altitudes and to occur earlier in the simulation period when anthropogenic emissions dominate. Both surface and aircraft model predictions exhibit a wide spread of positive and negative errors. In order to determine potential causes for bias in specific locations, we consider the model residual errors, or simply "residuals,"

$$r_k = H_k \left[M_k \left(\boldsymbol{x}_b \right) \right] - \boldsymbol{y}_k, \tag{11}$$

for each aircraft observation k. Figure 6 shows the statistically significant ($p < 0.32$) residuals for observations above and below the top of the model PBL. Section 5.1.3 describes relevant measures of observation variance and statistical significance.

Negative residuals, and hence low model bias, are most prevalent in northern California on 22 and 26 June, most likely due to under prediction of biomass burning sources. There is also low bias above the PBL in the southern San Joaquin Valley on 20 June and below the PBL inland from San Diego on 24 June. Although neither case has visual smoke in the MODIS images, there were fires detected within 300 km. The largest positive residual occurs in Palmdale, CA close to landing on 24 June. It could be indicative of either an emission error or the coarse horizontal resolution that collocates the airport with other significant nearby sources. Other notable high model biases aloft occur near cities during the flights on 20, 22, and 24 June. Similarly, surface site biases are higher near cities, and along the coast. As might be expected, proximity to sources is a strong indicator of error magnitude, as that is where the highest concentrations occur. The error sign ~~appears to be consistent above and below the PBL where such observations are collocated. Still,~~ and magnitude on 24 June differs in the PBL and free troposphere. That and the spatial error pattern could reflect some combination of meteorology and emissions deficiencies. For the positive residuals off the coast of Los Angeles on 22 and 24 June, there could be errors in predicting vertical mixing associated with the land-sea circulation or predominant near-surface wind direction. Discerning errors caused by emissions from those caused by meteorological mechanisms would require a separate in-depth study.

5.1.3 Variance and residual error significance

When **R** is assumed to be diagonal, each residual in the 4D-Var cost function is weighted inversely proportional to the observation error variance. The form of the cost function is based in Bayesian statistics, with an aim of converging on posterior control variables in a maximum-likelihood sense. However, using the variance alone to weight the residuals may result in very large cost function terms for relatively small residual errors. As our interest in this study is to determine how errors in emission estimates may be leading to model bias,

we wish to ensure the largest residuals have the greatest weight, while also accounting for differences in statistical significance of particular errors. Thus we define the diagonal terms of **R** as

$$R_{k,k} = \frac{w_k}{\sigma^2_{k,k}}, \tag{12}$$

where w_k is an additional weighting term and $\sigma^2_{k,k}$ is the variance.

The variance is comprised of components due to both observation and model uncertainty as

$$\sigma^2_{k,k} = \sigma^2_k = \sigma^2_{k,m} + \sigma^2_{k,o}. \tag{13}$$

The model variance at each observation location is found from an ensemble of $N_c = 156$ WRF-Chem configurations during the modeling period. Each ensemble member, c, uses a different combination of PBL, surface layer, LSM, and longwave and shortwave radiation options. Also, there are configurations both with and without microphysics and cumulus convection. From the ensemble, we use the population of residuals at each observation, k, to calculate the model variance

$$\sigma^2_{k,m} = \mathrm{MAX}\left(\frac{1}{N_c - 1} \sum_{c=1}^{N_c} \frac{(r_{k,c})^2}{N_c - 1} r^2_{k,c}, \mathrm{MML}^2 \right), \tag{14}$$

where MML is the minimum model limit. The minimum possible modeled BC concentration is limited by the boundary condition, which fills the entire model domain during the five day warm-up simulation. The MML is simply taken as the minimum model concentration for all observation locations and all model configurations, and is found to be 0.01 and 0.02 μg m^{-3} for aircraft and surface measurements, respectively, after rounding to the observation precision.

The IMPROVE instrument variance combines both relative and absolute uncertainties, the latter of which arises due to the minimum detection limit (MDL) (UC-Davis, 2002). For

a single filter analysis, the variance (in $\mu g^2\,m^{-6}$) is

$$\sigma^2_{l_k,\text{inst.}} = \left[\frac{\sqrt{34^2 + [(1000)\,(0.07)\,y_{l_k}]^2}}{1000}\right]^2 . \tag{15}$$

The sub-observation index l_k is useful at sites with more than one air sampler. When a site has data from multiple instruments in a single day, we take their average and combine their instrument variances as

$$\sigma^2_{k,o} = \frac{1}{L_k^2}\sum_{l_k}^{L_k} \frac{\sigma^2_{l_k,\text{inst.}}}{L_k^2}\sigma^2_{l_k,\text{inst.}}, \tag{16}$$

where L_k is the observation count. We assume ~~the IMPROVE measurements~~ each IMPROVE measurement fully represent the encompassing grid cell, since all sites are in remote locations and the samples are averaged over a 24 h period.

In contrast, the aircraft variance must capture the representativeness uncertainty associated with comparing the average of an entire model grid cell with an average of multiple short duration segments of a sparse aircraft transect. According to commercial literature for the SP2 device, it has an MDL of 0.01 $\mu g\,m^{-3}$, which we assume applies over the 10 s observation interval used during the ARCTAS campaign. The observations available through the NASA ARCTAS data archive have a BC mass concentration uncertainty of $\pm 30\,\%$. Although Sahu et al. (2012) report $\pm 10\,\%$ BC mass uncertainty, that range is given by Kondo et al. (2011), who state their results are applicable in regions not impacted by refractory organic compounds, such as from biomass burning sources. Because there are significant burning sources in this domain, we adopt the more conservative range. We utilize the instrument uncertainties in a definition for total observation variance with components due to both averaging and representativeness, such that for each average aircraft measurement,

$$\bar{y}_k = \frac{1}{L_k}\sum_{l_k=1}^{L_k} \frac{y_{l_k}}{L_k} y_{l_k}, \tag{17}$$

the total variance is

$$\sigma_{k,o}^2 = \mathrm{MDL}^2 + \sigma_{k,\mathrm{avg.}}^2 + \sigma_{k,\mathrm{rep.}}^2. \tag{18}$$

Adding the minimum variance associated with the MDL prevents the total variance from trending toward zero for any particular observation. This is important when using the variance in the cost function to ensure that near zero observations – which have low variances – with small residuals do not dominate the inversion. The ~~averaging~~ representative variance is the variance of the y_{l_k}'s that makeup $\bar{y}_k$, which is an attempt to capture the spread of true concentrations in a model grid cell. In the case that there is only a single observation, the ~~averaging~~ representative uncertainty is taken as double the instrument uncertainty. Thus,

$$\sigma_{k,\mathrm{avg.rep.}}^2 = \begin{cases} \frac{1}{L_k - 1} \sum_{l_k=1}^{L_k} \frac{\left(y_{l_k} - \bar{y}_k\right)^2}{L_k - 1} \left(y_{l_k} - \bar{y}_k\right)^2 & \text{if } L_k > 1; \\ \left(2\sigma_{k,\mathrm{inst.}}\right)^2 & \text{if } L_k = 1 \end{cases} \tag{19}$$

For any time step where $L_k < L_{\max} = 9$, there is an additional variance penalty proportional to the sum of the individual instrument variances,

$$\sigma_{k,\mathrm{rep.inst.}}^2 = \sqrt{\frac{L_{\max} - L_k}{L_{\max}}} \frac{1}{L_k^2} \sum_{l_k=1}^{L_k} \frac{\sigma_{l_k,\mathrm{inst.}}^2}{L_k^2} \sigma_{l_k,\mathrm{inst.}}^2, \tag{20}$$

where

$$\sigma_{l_k,\mathrm{inst.}} = \mathrm{MAX}\left(\mathrm{MDL}, 0.3 \cdot y_{l_k}\right). \tag{21}$$

The square root term in Eq. (20) inflates the instrument error in cases when there are fewer than $L_{\max}$ samples in the mean.

In order to motivate the weight, w_k, applied to each residual model error, let us consider the primary inputs to the adjoint simulation, which are the adjoint forcings

$$\lambda_{k,m}^* = \frac{\partial J}{\partial c_k}$$

$$= \mathbf{H}_k^\top \sigma_k^{-2} \{H_k [M_k (\boldsymbol{x}_b)] - \boldsymbol{y}_k\}$$
$$= \mathbf{H}_k^\top \lambda_{k,o}^*. \tag{22}$$

c_k is any state variable on which H_k depends and which $M_k(\boldsymbol{x}_b)$ predicts. For our purposes, the state variables are modeled BC concentrations. The adjoint of the observation operator, $\mathbf{H}_k^\top$ transforms the forcing from observation space ($\lambda_{k,o}^*$) back to model space ($\lambda_{k,m}^*$). Thus, the forcing in observation space is

$$\lambda_{k,o}^* = \frac{r_k}{\sigma_k^2}. \tag{23}$$

Observations with significant model bias would require the largest perturbation in control variables to alleviate, and would seem to inform the inversion process the greatest. However, they must also have low total variance to contribute to an inversion. Figure 8 shows the surface and aircraft SD plotted ~~vs.~~ versus residual error. Also plotted in that figure are one and two SD zones, as well as lines of constant $\lambda_{k,o}^*$ for all $w_k = 1$. Any residual falling outside the 2σ zone ~~is one where the~~ has a combined model and observation SD that is small enough to ~~say~~ determine with 95 % confidence ($p < 0.05$) that the residual error deviates from zero (i.e., the model and observation disagree). These statistically significant model errors indicate that some kind of inversion is worthwhile. ~~The~~ In their multi-cycle 4D-Var approach, Bergamaschi et al. (2009) eliminate observations outside three SD's after an initial 4D-Var cycle, with the thought that incorrect model physics prevents those residual errors from being fixed with 4D-Var. Thus, while statistically significant residuals are important to driving a 4D-Var inversion, that they remain afterward is a strong indication of errors in the model description that cannot be fixed through adjustments to emissions. Figure 8 shows that the relative contributions of observation and model variances is in general proportional to the relative magnitudes of observed and modeled concentration. ~~Thus~~Specifically, model (observation) variation contributes to a large fraction of uncertainty in positive (negative) residuals.

~~There are several outlier negative residuals with magnitudes much larger than the remainder of the population. A large portion of these have large enough uncertainty that~~

their adjoint forcing is much less than that of other lower magnitude residuals.Consider the region where $|r_k| < 0.5$and $\sigma_k < 0.3$When both the observed and modeled concentrations are small, the total variance decreases to the minimum possible value, governed by the MML and MDL. This generally happens in remote regions, where small concentrations result from some combination of small nearby sources and transport from many distant sources. If the total variance is small enough relative to the residual error, $\lambda^*_{k,o}$ will be very large, often larger than in cases with larger residual errors (see Fig. 10a). The adjoint forcing magnitude is between 10 and 200, varying the mean forcing magnitude to the maximum for any observation in the whole population. model propagates a relatively large forcing from a small residual backward, resulting in large sensitivities to emission scaling factors. These sensitivities then translate to large emission perturbations in the optimization process.

The residual errors within the 1σ and 2σ zones are not statistically significant, yet they might have larger adjoint forcing than observations with larger residual error at higher significance levels. Applying these adjoint forcings as-is could drive the inversion to fitting data points with smallabsolute residual error. This adjoint forcing imbalance between high and low significance observations can be alleviated by a counteracting in remote locations are likely within combined model and observation uncertainty, but the model variance at these locations is unrealistically small. The ensemble will underestimate variance at observations near low-biased prior sources due to the absence of tracer mass. The opposite may be true for a high-biased prior. The challenge then is to define the concentration uncertainty introduced by the model physics, independent of the magnitude of emissions, which we attempt to do with a weighting scheme. In order to devise such a scheme, we consider which forms of statistical significance are important to this inverse problem. The weights are used only to inflate variance, which when very low is thought to misinform the adjoint about concentration errors. Variance reduction may be necessary for observations near high-biased sources. Also, while we apply the weights to the total variance, they could be applied to only the model portion. Here we are developing a philosophy for scaling the variances, of which the following description is but one example.

Because our goal in an emission inversion is to reduce model bias by perturbing emissions, model bias is itself an important characteristic. We use the ensemble of model configurations to calculate the variance in all residual errors, that is

$$\sigma_{\mathsf{r}}^2 = \frac{1}{N_c K - 1} \sum_{c=1}^{N_c} \sum_{k=1}^{K} \frac{r_{k,c}^2}{N_c K - 1} r_{k,c}^2. \tag{24}$$

The residual SD, σ_{r}, are 0.69 and 0.29 $\mu g\, m^{-3}$ for surface and aircraft observation populations, respectively. After confirming that the residual errors are approximately normally distributed, the significance of the bias of a single observation relative to the entire population is

$$f_{\mathsf{POP},k} = \mathrm{erf}\left(\frac{|\tilde{r}_k|}{\sqrt{2\sigma_{\mathsf{r}}^2}}\right). \tag{25}$$

In statistics, the ratio of $\frac{|\tilde{r}_k|}{\sigma_{\mathsf{r}}}$ is called the z value, and denotes the number of SD between $\tilde{r}_k$ and the expected value of zero. The variable $\tilde{r}_k$ indicates the user must select a specific form of residual error. Two examples are the mean or median of r_k. A third approach, and the one taken here is to use the residual found in the first 4D-Var iteration, $r_{k,n=0}$. $f_{\mathsf{POP},k}$ is a continuously variable p value, or the percentage of the population of all $r_{k,c}$ that is less significant than $\tilde{r}_k$. Another measure of significance is visualized in the σ zones of Fig. 8, and was discussed previously. That is, for an individual residual error and variance, what is the probability that there will always be mismatch between the model and observation? The individual error significance is

$$f_{\mathsf{IND},k} = \mathrm{erf}\left(\frac{|\tilde{r}_k|}{\sqrt{2\sigma_k^2}}\right). \tag{26}$$

The population and individual error significances are combined to derive the adjoint forcing weight,

$$w_k = \left[\left(f_{\mathrm{POP},k} \right)^{\gamma} \left(f_{\mathrm{IND},k} \right)^{1-\gamma} \right]^{\beta}. \tag{27}$$

The weighting scheme can be tuned for a specific application using the γ and β parameters to reshape the adjoint forcing contours. However, care must be taken when selecting γ, β, and $\tilde{r}_k$ to ensure convergence in 4D-Var~~to the mean of the Gaussian distribution of residual errors~~. Use of these weights may imply that residual errors do not fit a Gaussian distribution. Here we only introduce the weighting scheme and use it in a demonstration, but do not verify its validity. We use $\gamma = 0.5$ to provide some balance between the two measures of significance and $\beta = 2$ to ensure the weighting has a large impact. After calculating the w_k's according to Eq. (27), the new effective adjoint forcings are compared to the original values in Fig. 9. The weighting scheme is successful at reducing the impact of observation errors with low significance on the cost function.

After applying the new weighting scheme, the $\lambda^*_{k,o}$ contours no longer converge on the y axis as depicted in Fig. 8. Instead, they exit radially from the origin in all directions. As both the population and individual z values approach zero, the adjoint forcing converges toward

$$\lambda^*_{k,o} \approx \frac{r_k}{\sigma_k^2} \left(0.8 \frac{|\tilde{r}_k|}{\sigma_{\mathrm{r}}^{\gamma} \sigma_k^{1-\gamma}} \right)^{\beta} = 0.64 \frac{r_k \tilde{r}_k^2}{\sigma_{\mathrm{r}} \sigma_k^3}. \tag{28}$$

For our specific values of σ_{r}, all residual errors within the 2σ zone satisfy $|\lambda^*_{k,o}| \lesssim 5\,\mu\mathrm{g}^{-1}\mathrm{m}^3$ for surface, and $|\lambda^*_{k,o}| \lesssim 10\,\mu\mathrm{g}^{-1}\mathrm{m}^3$ for aircraft observations. This is a considerable change from the unity weights where $|\lambda^*_{k,o}|$ was as large as 200 $\mu\mathrm{g}^{-1}\mathrm{m}^3$ in the region between the 1σ and 2σ zones.

5.2 Results and discussion

With the weighting function applied, we calculate sensitivities of the 4D-Var cost function with respect to emissions for determining potential sources of model bias. The weights

reduce the cost function from 5374 to 3784, which increases the normalized cost function sensitivity to emission perturbations. Figure 10 shows fully normalized sensitivities,

$$\frac{\partial \ln J}{\partial \ln E_{i,j,d}} = \sum_{n=1}^{24} \frac{\partial \ln J}{\partial \ln E_{i,j,d,n}}, \tag{29}$$

for six days of the simulation. The sensitivity in a particular grid cell is summed over the local diurnal cycle for hours $n = [1, \ldots, 24]$ on day d. For anthropogenic emissions, the local time is calculated for discrete 15° time zones, whereas for biomass burning emissions, local time corresponds to the continuous sun cycle. Undoubtedly, there are locations with positive and negative sensitivities at different times of day that will cancel, but this temporally aggregated sensitivity is an attempt to obtain average daily relationships across the domain. Although the color bar has been saturated at $\pm 5 \times 10^{-3}$, the full range of sensitivities are from ~~-2.7×10^{-3} to $+5.6 \times 10^{-3}$ and -5.4×10^{-3} to $+6.3 \times 10^{-3}$~~ -2.3×10^{-3} to $+7.1 \times 10^{-3}$ and -4.9×10^{-3} to $+11.3 \times 10^{-3}$ for anthropogenic and biomass burning emissions, respectively.

The magnitude of a normalized sensitivity corresponds to the fractional response in the cost function given a 100 % perturbation of emissions in a grid cell. If the model were perfect, the sensitivity magnitudes would be proportional to the difference between the background emission estimate and the true value. Thus a negative sensitivity indicates a location where estimated sources are too low, and vice versa. Because the sensitivities themselves depend on the emission magnitudes, they will change in each 4D-Var iteration, eventually converging on a minimum of the cost function where the sensitivities are zero. We use the sensitivities here as a qualitative indicator of emission errors, and not a quantitative conclusion as might be provided with a complete inversion.

The sensitivities exhibit a similar spatial-temporal pattern as the residual errors in Figs. 5 and 6, ~~in general indicating that~~ with the exception of sources far upwind of observations. Near observations, estimated anthropogenic emissions are too high, and ~~that~~ estimated fire emissions are too low. ~~A more complex depiction of all BC emission errors arises in the sensitivities than the residuals alone might reveal. While most~~ Indeed, most of the non-

negligible (not black) burning sensitivities are negative ~~, emissions from the Los Padres National Forest and northern redwoods on 23 Juneare potentially too high. The relative contributions of those fire estimates and simultaneous anthropogenic sources to positive coastal surface residuals near L.A. on~~ on 24 June and 25 June, likely due to the high-altitude negative residuals on 26 June. The positive coastal fire sensitivities on 22 June and 23 June are ~~difficult to disentangle. However, the 24 June positive residuals from ARCTAS are more likely due to the anthropogenic sources. That is because the model transports smoke into the flight path of the DC-8 south of San Pedro, where BC concentrations are under predicted. Still, some anthropogenic source regions are under predicted as well.~~attributable to positive forcing at the Point Reyes National Seashore IMPROVE site on 23 June and along the DC-8 flight track on 24 June. The influence of those fires on Los Angeles BC concentrations 24 hours or more after the emission was determined through a sensitivity test where a perturbed residual error and adjoint forcing on 24 June were propagated through the adjoint.

The spatial variations in sensitivities ~~are indicative of~~ reveal two phenomena. First, appreciable sensitivities will only arise in emissions that influence the particular observations available. Thus, full observation coverage is imperative to a successful inversion. Second, emission errors are heterogeneous in space and time. For ~~biomass burning sources~~the FINN inventory, heterogeneity arises due to missed detections in the MODIS active fire product, as well as potential errors in vegetation classification or attribution of a particular vegetation class to one of four land cover types~~used in FINN~~. Anthropogenic source error heterogeneity could be due to a static inventory from 2005 being used to describe emissions in 2008, or to spatial variations in BC emission factors for a particular source sector.

~~Comparative adjoint sensitivities are calculated using the SLAB~~ All of the conclusions that might be drawn from the sensitivity maps about emission errors are subject to the assumption that the transport is correct in this model configuration. The SLAM LSM scheme (option 1) is used in place of the PX option ~~. In these results , the same positive coastal sensitivities are even more pronounced and widespread on~~ to calculate comparative adjoint sensitivities. Relative to the PX option, these results exhibit non-negligible negative

sensitivities to fires in the Shasta-Trinity National Forest on 23~~and 24~~ June, but much larger positive sensitivities to those same fires on 22 June. ~~Negative sensitivities to~~ Sensitivities with respect to coastal fires in the ~~Sequoia and Inyo National Forests are larger in magnitude than those in the Sierras on 23 June, but the~~ Los Padres National Forest also increase on 22 June. The spatial sensitivity patterns between SLAB and PX options are consistent on 25 June. The differences are presumably due to changes in the residual error between the two configurations~~, since the weights and variances used are identical. $\tilde{r}_k$ was not recalculated for the SLAB case.~~ . The differing spatial sensitivity patterns indicate that the surface heat and moisture fluxes calculated by each LSM scheme contributes non-negligibly to the vertical mixing of BC to aircraft measurement altitudes.

We also consider temporal sensitivity patterns to compare the two LSM schemes. Figure 11 shows the diurnal distribution of biomass burning, and weekday and weekend anthropogenic BC emission sensitivities for both of the LSM configurations, and for unity weights, $w_k = 1$ and w_k from Eq. (27). Each bar in that plot represents a summation of sensitivities across the whole domain from 20 June, 00:00:00 UTC to 26 June, 23:00:00 UTC (~~$d = [1, \ldots, 7]$~~$d = [1, \ldots, n_d]$) within a particular local hour, n, such that

$$\frac{\partial \ln J}{\partial \ln E_n} = \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \sum_{d=1}^{7n_d} \frac{\partial \ln J}{\partial \ln E_{i,j,d,n}}. \tag{30}$$

The signs and magnitudes of sensitivities fit the previous description for the spatially distributed temporal aggregation. The time period of emissions to which an observation is most sensitive depends on the altitude of that observation and the flow mechanisms that transport emitted aerosol mass to that observation. Thus, any conclusions drawn could be biased if observations do not have full temporal coverage, especially near sources. Since normalized sensitivities are proportional to emissions, it is to be expected that sensitivities at peak emission hours are magnified. Also, each hour of sensitivity is a sum of many diverse source locations. So while the net sensitivity in a single hour may be positive, the

spatial distribution of sensitivities is much more varied, as was ~~previously discussed.~~shown in Fig. 10.

The FINN biomass burning inventory applies an identical diurnal emission apportionment for all fires, regardless of vegetation, shading due to slopes, wind speed, or relative humidity. This scaling is applied in preprocessing. Both the PX and SLAB LSM setups seem to agree that ~~the timing of the FINN burning emissions peak is correct within ±1and that the peak should be sharper~~fire emissions between 10:00 and 18:00 local time are over predicted. Without the weighting scheme, the PX configuration indicates that ~~burning emissions are too low in peak hours~~the peak should be smoothed out, while the SLAB configuration concludes that ~~burning emissions are too high in~~ the peak should be made sharper by reducing off-peak ~~hours~~emissions. With the weighting scheme applied, both configurations agree that ~~fire emissions need to be increased to reduce the cost function. The increased burning sensitivity magnitudes indicate the weighting scheme is successful at generating a cost function that is more robustly sensitive to emission perturbations~~the peak is timed correctly. While the two setups disagree about the magnitude of fire emission correction required, their differences are small in comparison to the implied anthropogenic correction. The relative disagreement in burning sensitivity magnitude between the two LSM configurations is attributable to differences in residual errors, r_k, and the resulting adjoint forcings, $\lambda^*_{k,o}$. ~~Both configurations seem to agree that the timing of emissions is correct, and in fact the midday peak should be sharpened. However, the normalized sensitivities are proportional to emissions, meaning an emission peak should correspond to a sensitivity peak. Some work still needs to be done to interpret diurnal sensitivity patterns for use in a full inversion.~~

In contrast to FINN, NEI applies a variety of diurnal patterns to point, area, and traffic sources. The weekend and weekday emission profiles shown in Fig. 11 are the emission weighted averages for the entire domain. Individual sources may have a profile closer to flat, or alternatively zero overnight, and flat during daylight hours. The weighted average profile shown is close to the one used for commercial diesel traffic, since that is the largest BC source within the domain. Attributing sensitivities, or errors, to specific sectors is not straight-forward and doing so may require a smaller horizontal grid spacing to reduce the

number of sectors per grid cell. Results for the weighted and unweighted cost functions are very similar. In general, anthropogenic emissions are too high throughout all times of the day on both weekdays and weekends. Both LSM configurations indicate the weekday profile peak should be sharper near 14:00 LT, and not at 16:00 LT, but also that emissions from 06:00 to 16:00 LT should be closer to the late evening and early morning magnitudes. The weekend sensitivities indicate the evening and morning emissions are too high, and that the daytime peak is timed about right, with the exception of the 18:00 LT spike. However, the relatively small magnitude of weekend sensitivities could also indicate there were not enough observations of anthropogenic sources on 21 June (SAT) and 22 (SUN) to draw definitive conclusions about emission timing.

Results for the two LSM options reveal the potential for model configuration to introduce bias in a 4D-Var inversion. For these particular observations, the posterior emissions from the PX option would likely be higher than those from the SLAB option, because of their relative sensitivity values. Model variability must be taken into consideration in 4D-Var sensitivity studies of high resolution emissions, because model variation represents a large fractional contribution to observation error variance for positive residuals, as shown in Fig. 8.

6 Conclusions

We have implemented, verified, and demonstrated the WRFPLUS-Chem coupled meteorology and chemical adjoint and tangent linear models for PBL mixing, emission, aging, dry deposition, and advection of BC aerosol. A second order checkpointing scheme enables tangent linear and adjoint model runs longer than six hours. The adjoint was used in the first iteration of a 4D-Var inversion within WRFDA-Chem, where model-observation residual errors are compared for low- and high-temporal resolution IMPROVE surface and ARCTAS-CARB aircraft observations during one week of June 2008. A novel cost function weighting scheme was devised to ~~increase~~ reduce the impact of ~~high significance~~ low-significance observations in future 4D-Var inversions. ~~Results indicate that the weighting scheme is effective at generating robust sensitivities of the cost function to emissions.~~ The adjoint

sensitivities also indicate that anthropogenic emissions are over predicted and biases in burning emissions are ~~under predicted for the domain and time period considered~~spatially and temporally heterogenous. The diurnal sensitivities would seem to indicate that burning emission profiles should be steeper midday, while anthropogenic emission profiles should be flattened on weekdays and sharpened on weekends. A full inversion is necessary to quantify the magnitude of the errors in the emissions. Additionally, adjoint sensitivities found using two different LSM options indicate that the results of such inversions will be sensitive to the choice of model configuration.

The next steps are as follows. We intend to incorporate tangent linear and adjoint observation operators for useful remote sensing products (e.g., aerosol optical depth (Saide et al., 2013) and absorbing aerosol optical depth). This addition will enable WRFDA-Chem to be applied to a wider range of domains and time periods and operationally. The WRFDA-Chem optimization algorithm still needs to be applied to control variables for chemical species initial conditions and emission scaling factors. Future development and incorporation of radiation and microphysics adjoints (e.g. Saide et al., 2012) will provide coupling between aerosols and meteorology, and provide new insights into sensitivities of direct, indirect, and semi-direct radiative forcing to emission sectors and locations. In addition to the aerosol applications discussed, WRFDA-Chem 4D-Var will also be suited to emission inversions for green house gases and other chemical tracers.

7 Code availability

Although an annual code release may be available in the future, WRFPLUS-Chem and WRFDA-Chem are continually being developed. A static version would not include the most recent bug fixes. Interested users can obtain the code as-is by contacting the authors: J. J. Guerrette (jonathan.guerrette@colorado.edu) and D. K. Henze (daven.henze@colorado.edu). Potential developers may also contact National Center for Atmospheric Research (NCAR) scientist, H. C. Lin (hclin@ucar.edu) for access to the WRFPLUS-Chem repository. Any questions should be directed to the authors.

The Supplement related to this article is available online at doi:10.5194/gmdd-0-1-2015-supplement.

Acknowledgements. This research has been supported by a grant from the U.S. Environmental Protection Agency's Science to Achieve Results (STAR) program. Although the research described in the article has been funded wholly or in part by the U.S. Environmental Protection Agency's STAR program through grant R835037, it has not been subjected to any EPA review and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred. We are thankful for the ARCTAS mission, which was supported by NASA. We thank Y. Kondo for making the SP2 observations available through the NASA LaRC Airborne Science Data for Atmospheric Composition database. We acknowledge the use of FIRMS data from the Land Atmosphere Near-real time Capability for EOS (LANCE) system operated by the NASA/GSFC/Earth Science Data and Information System (ESDIS) with funding provided by NASA/HQ.

References

Al-Saadi, J., Soja, A. J., Pierce, R. B., Szykman, J., Wiedinmyer, C., Emmons, L., Kondragunta, S., Zhang, X., Kittaka, C., Schaack, T., and Bowman, K.: Intercomparison of near-real-time biomass burning emissions estimates constrained by satellite fire data, J. Appl. Remote Sens., 2, 021504, 1–24, doi:10.1117/1.2948785, 2008.

Anenberg, S. C., Talgo, K., Arunachalam, S., Dolwick, P., Jang, C., and West, J. J.: Impacts of global, regional, and sectoral black carbon emission reductions on surface air quality and human mortality, Atmos. Chem. Phys., 11, 7253–7267, doi:10.5194/acp-11-7253-2011, 2011.

Baklanov, A., Schlünzen, K., Suppan, P., Baldasano, J., Brunner, D., Aksoyoglu, S., Carmichael, G., Douros, J., Flemming, J., Forkel, R., Galmarini, S., Gauss, M., Grell, G., Hirtl, M., Joffre, S., Jorba, O., Kaas, E., Kaasik, M., Kallos, G., Kong, X., Korsholm, U., Kurganskiy, A., Kushta, J., Lohmann, U., Mahura, A., Manders-Groot, A., Maurizi, A., Moussiopoulos, N., Rao, S. T., Savage, N., Seigneur, C., Sokhi, R. S., Solazzo, E., Solomos, S., Sørensen, B., Tsegas, G., Vignati, E., Vogel, B., and Zhang, Y.: Online coupled regional meteorology chemistry models in Europe: current status and prospects, Atmos. Chem. Phys., 14, 317–398, doi:10.5194/acp-14-317-2014, 2014.

Barker, D., Lee, M.-S., Guo, Y.-R., Huang, W., Huang, H., and Rizvi, Q.: WRF-Var – a unified 3/4D-Var variational data assimilation system for WRF, in: Sixth WRF/15th MM5 Users' Workshop, Boulder, CO, NCAR, 17 pp., available at: http://www2.mmm.ucar.edu/wrf/users/workshops/WS2005/presentations/session10/1-Barker.pdf (last access: 20 February 2015), 2005.

Barker, D. M., Huang, W., Guo, Y.-R., Bourgeois, A. J., and Xiao, Q. N.: A three-dimensional variational data assimilation system for MM5: implementation and initial results, Mon. Weather Rev., 132, 897–914, 2004.

Benedetti, A., Morcrette, J.-J., Boucher, O., Dethof, A., Engelen, R. J., Fisher, M., Flentje, H., Huneeus, N., Jones, L., Kaiser, J. W., Kinne, S., Mangold, A., Razinger, M., Simmons, A. J., and Suttie, M.: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: 2. Data assimilation, J. Geophys. Res., 114, D13205, doi:10.1029/2008JD011115, 2009.

Bergamaschi, P., Frankenberg, C., Meirink, J. F., Krol, M., Villani, M. G., Houweling, S., Dentener, F., Dlugokencky, E. J., Miller, J. B., Gatti, L. V., Engel, A., and Levin, I.: Inverse modeling of global and regional CH4 emissions using SCIAMACHY satellite retrievals, J. Geophys. Res.-Atmos, 114, D22301, doi:10.1029/2009JD012287, 2009.

Bocquet, M., Elbern, H., Eskes, H., Hirtl, M., Žabkar, R., Carmichael, G. R., Flemming, J., Inness, A., Pagowski, M., Pérez Camaño, J. L., Saide, P. E., San Jose, R., Sofiev, M., Vira, J., Baklanov, A., Carnevale, C., Grell, G. and Seigneur, C.: Data assimilation in atmospheric chemistry models: current status and future prospects for coupled chemistry meteorology models, Atmos. Chem. Phys. Discuss., 14, 32233–32323, doi:10.5194/acpd-14-32233-2014, 2014.

Bond, T. C., Doherty, S. J., Fahey, D. W., Forster, P. M., Berntsen, T., DeAngelo, B. J., Flanner, M. G., Ghan, S., Kärcher, B., Koch, D., Kinne, S., Kondo, Y., Quinn, P. K., Sarofim, M. C., Schultz, M. G., Schulz, M., Venkataraman, C., Zhang, H., Zhang, S., Bellouin, N., Guttikunda, S. K., Hopke, P. K., Jacobson, M. Z., Kaiser, J. W., Klimont, Z., Lohmann, U., Schwarz, J. P., Shindell, D., Storelvmo, T., Warren, S. G., and Zender, C. S.: Bounding the role of black carbon in the climate system: a scientific assessment: black carbon in the climate system, J. Geophys. Res.-Atmos., 118, 5380–5552, doi:10.1002/jgrd.50171, 2013.

Chin, M., Savoie, D. L., Huebert, B. J., Bandy, A. R., Thornton, D. C., Bates, T. S., Quinn, P. K., Saltzman, E. S., and De Bruyn, W. J.: Atmospheric sulfur cycle simulated in the global model GOCART: comparison with field observations and regional budgets, J. Geophys. Res.-Atmos., 105, 24689–24712, doi:10.1029/2000JD900385, 2000.

Courtier, P., Thépaut, J.-N., and Hollingsworth, A.: A strategy for operational implementation of 4D-Var, using an incremental approach, Q. J. Roy. Meteor. Soc., 120, 1367–1387, doi:10.1002/qj.49712051912, 1994.

Dubovik, O., Lapyonok, T., Kaufman, Y. J., Chin, M., Ginoux, P., Kahn, R. A., and Sinyuk, A.: Retrieving global aerosol sources from satellites using inverse modeling, Atmos. Chem. Phys., 8, 209–250, doi:10.5194/acp-8-209-2008, 2008.

Flemming, J., Inness, A., Flentje, H., Huijnen, V., Moinat, P., Schultz, M. G., and Stein, O.: Coupling global chemistry transport models to ECMWF's integrated forecast system, Geosci. Model Dev., 2, 253–265, doi:10.5194/gmd-2-253-2009, 2009.

Freitas, S. R., Longo, K. M., Chatfield, R., Latham, D., Silva Dias, M. A. F., Andreae, M. O., Prins, E., Santos, J. C., Gielow, R., and Carvalho Jr., J. A.: Including the sub-grid scale plume rise of vegetation fires in low resolution atmospheric transport models, Atmos. Chem. Phys., 7, 3385–3398, doi:10.5194/acp-7-3385-2007, 2007.

Freitas, S. R., Longo, K. M., Trentmann, J., and Latham, D.: Technical Note: Sensitivity of 1-D smoke plume rise models to the inclusion of environmental wind drag, Atmos. Chem. Phys., 10, 585–594, doi:10.5194/acp-10-585-2010, 2010.

Freitas, S. R., Longo, K. M., Alonso, M. F., Pirre, M., Marecal, V., Grell, G., Stockler, R., Mello, R. F., and Sánchez Gácita, M.: PREP-CHEM-SRC – 1.0: a preprocessor of trace gas and aerosol emission fields for regional and global atmospheric chemistry models, Geosci. Model Dev., 4, 419–433, doi:10.5194/gmd-4-419-2011, 2011.

Giering, R. and Kaminski, T.: Recipes for adjoint code construction, ACM T. Math. Software, 24, 437–474, doi:10.1145/293686.293695, 1998.

Grahame, T. J., Klemm, R., and Schlesinger, R. B.: Public health and components of particulate matter: the changing assessment of black carbon, J. Air Waste Manage., 64, 620–660, doi:10.1080/10962247.2014.912692, 2014.

Grell, G., Freitas, S. R., Stuefer, M., and Fast, J.: Inclusion of biomass burning in WRF-Chem: impact of wildfires on weather forecasts, Atmos. Chem. Phys., 11, 5289–5303, doi:10.5194/acp-11-5289-2011, 2011.

Grell, G. A. and Freitas, S. R.: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling, Atmos. Chem. Phys., 14, 5233–5250, doi:10.5194/acp-14-5233-2014, 2014.

Grell, G. A., Knoche, R., Peckham, S. E., and McKeen, S. A.: Online versus offline air quality modeling on cloud-resolving scales, Geophys. Res. Lett., 31, L16117, doi:10.1029/2004GL020175, 2004.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled "online" chemistry within the WRF model, Atmos. Environ., 39, 6957–6975, 2005.

Gumley, L.: MODIS Today, Space Science and Engineering Center, University of Wisconsin-Madison, Madison, WI, USA, available at: http://ge.ssec.wisc.edu/modis-today (last access: 20 January 2015), 2008.

Hakami, A., Henze, D. K., Seinfeld, J. H., Chai, T., Tang, Y., Carmichael, G. R., and Sandu, A.: Adjoint inverse modeling of black carbon during the Asian Pacific Regional Aerosol Characterization Experiment, J. Geophys. Res.-Atmos., 110, D14301, doi:10.1029/2004JD005671, 2005.

Hansen, J., Sato, M., and Ruedy, R.: Radiative forcing and climate response, J. Geophys. Res.-Atmos., 102, 6831–6864, doi:10.1029/96JD03436, 1997.

Hansen, M., DeFries, R., Townshend, J., Carroll, M., Dimiceli, C., and Sohlberg, R.: 500 m MODIS Vegetation Continuous Fields, The Global Land Cover Facility, College Park, Maryland, 2003.

Hascoët, L. and Pascual, V.: The Tapenade Automatic Differentiation tool: principles, model, and specification, ACM T. Math. Software, 39, 20, 1–43, doi:10.1145/2450153.2450158, 2013.

Henze, D. K., Hakami, A., and Seinfeld, J. H.: Development of the adjoint of GEOS-Chem, Atmos. Chem. Phys., 7, 2413–2433, doi:10.5194/acp-7-2413-2007, 2007.

Henze, D. K., Seinfeld, J. H., and Shindell, D. T.: Inverse modeling and mapping US air quality influences of inorganic $PM_{2.5}$ precursor emissions using the adjoint of GEOS-Chem, Atmos. Chem. Phys., 9, 5877–5903, doi:10.5194/acp-9-5877-2009, 2009.

Huang, X.-Y., Xiao, Q., Barker, D. M., Zhang, X., Michalakes, J., Huang, W., Henderson, T., Bray, J., Chen, Y., Ma, Z., Dudhia, J., Guo, Y., Zhang, X., Won, D.-J., Lin, H.-C., and Kuo, Y.-H.: Four-dimensional variational data assimilation for WRF: formulation and preliminary results, Mon. Weather Rev., 137, 299–314, doi:10.1175/2008MWR2577.1, 2009.

Huneeus, N., Boucher, O., and Chevallier, F.: Simplified aerosol modeling for variational data assimilation, Geosci. Model Dev., 2, 213–229, doi:10.5194/gmd-2-213-2009, 2009.

Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C., Coheur, P., Engelen, R. J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, V., Hurtmans, D., Jones, L., Kaiser, J. W., Kapsomenakis, J., Lefever, K., Leitão, J., Razinger, M., Richter, A., Schultz, M. G., Simmons, A. J., Suttie, M., Stein, O., Thépaut, J.-N., Thouret, V., Vrekoussis, M., Zerefos, C., and the MACC team: The MACC reanalysis: an 8 yr data set of

atmospheric composition, Atmos. Chem. Phys., 13, 4073–4109, doi:10.5194/acp-13-4073-2013, 2013.

Jacob, D. J., Crawford, J. H., Maring, H., Clarke, A. D., Dibb, J. E., Emmons, L. K., Ferrare, R. A., Hostetler, C. A., Russell, P. B., Singh, H. B., Thompson, A. M., Shaw, G. E., McCauley, E., Pederson, J. R., and Fisher, J. A.: The Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS) mission: design, execution, and first results, Atmos. Chem. Phys., 10, 5191–5212, doi:10.5194/acp-10-5191-2010, 2010.

Kaminski, J. W., Neary, L., Struzewska, J., McConnell, J. C., Lupu, A., Jarosz, J., Toyota, K., Gong, S. L., Côté, J., Liu, X., Chance, K., and Richter, A.: GEM-AQ, an on-line global multiscale chemical weather modelling system: model description and evaluation of gas phase chemistry processes, Atmos. Chem. Phys., 8, 3255–3281, doi:10.5194/acp-8-3255-2008, 2008.

Kinnison, D. E., Brasseur, G. P., Walters, S., Garcia, R. R., Marsh, D. R., Sassi, F., Harvey, V. L., Randall, C. E., Emmons, L., Lamarque, J. F., Hess, P., Orlando, J. J., Tie, X. X., Randel, W., Pan, L. L., Gettelman, A., Granier, C., Diehl, T., Niemeier, U., and Simmons, A. J.: Sensitivity of chemical tracers to meteorological parameters in the MOZART-3 chemical transport model, J. Geophys. Res., 112, D20302, doi:10.1029/2006JD007879, 2007.

Koch, D. and Del Genio, A. D.: Black carbon semi-direct effects on cloud cover: review and synthesis, Atmos. Chem. Phys., 10, 7685–7696, doi:10.5194/acp-10-7685-2010, 2010.

Kondo, Y., Sahu, L., Moteki, N., Khan, F., Takegawa, N., Liu, X., Koike, M., and Miyakawa, T.: Consistency and traceability of black carbon measurements made by laser-induced incandescence, thermal-optical transmittance, and filter-based photo-absorption techniques, Aerosol Sci. Tech., 45, 295–312, doi:10.1080/02786826.2010.533215, 2011.

Krewski, D., Jerrett, M., Burnett, R. T., Ma, R., Hughes, E., Shi, Y., Turner, M. C., Pope III, C. A., Thurston, G., Calle, E. E., and Eugenia, E.: Extended Follow-Up and Spatial Analysis of the American Cancer Society Study Linking Particulate Air Pollution and Mortality, Tech. Rep. 140, Health Effects Institute, available at: http://www.healtheffects.org/Pubs/RR140-Krewski.pdf (last access: 15 August 2014), 2009.

Liu, Z., Liu, Q., Lin, H.-C., Schwartz, C. S., Lee, Y.-H., and Wang, T.: Three-dimensional variational assimilation of MODIS aerosol optical depth: implementation and application to a dust storm over East Asia: AOD DATA ASSIMILATION, J. Geophys. Res.-Atmos., 116, D23206, doi:10.1029/2011JD016159, 2011.

Lohmann, U. and Feichter, J.: Global indirect aerosol effects: a review, Atmos. Chem. Phys., 5, 715–737, doi:10.5194/acp-5-715-2005, 2005.

Malm, W. C., Sisler, J. F., Huffman, D., Eldred, R. A., and Cahill, T. A.: Spatial and seasonal trends in particle concentration and optical extinction in the United States, J. Geophys. Res.-Atmos., 99, 1347–1370, doi:10.1029/93JD02916, 1994.

Morcrette, J.-J., Boucher, O., Jones, L., Salmond, D., Bechtold, P., Beljaars, A., Benedetti, A., Bonet, A., Kaiser, J. W., Razinger, M., Schulz, M., Serrar, S., Simmons, A. J., Sofiev, M., Suttie, M., Tompkins, A. M., and Untch, A.: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: forward modeling, J. Geophys. Res., 114, D06206, doi:10.1029/2008JD011235, 2009.

Myhre, G., Shindell, D., Breon, F., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H.: Anthropogenic and natural radiative forcing, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, UK and New York, NY, USA, 129–234, 2013.

NASA: MCD14ML MODIS Active Fire Detections, available at: https://earthdata.nasa.gov/active-fire-data#tab-content-6, last access: 18 December 2014.

Pagowski, M. and Grell, G. A.: Experiments with the assimilation of fine aerosols using an ensemble Kalman filter, J. Geophys. Res.-Atmos., 117, D21302, doi:10.1029/2012JD018333, 2012.

Pagowski, M., Grell, G. A., McKeen, S. A., Peckham, S. E., and Devenyi, D.: Three-dimensional variational data assimilation of ozone and fine particulate matter observations: some results using the Weather Research and forecasting-chemistry model and grid-point statistical interpolation, Q. J. Roy. Meteor. Soc., 136, 2013–2024, doi:10.1002/qj.700, 2010.

Pleim, J. E.: A simple, efficient solution of flux–profile relationships in the atmospheric surface layer, J. Appl. Meteorol. Clim., 45, 341–347, doi:10.1175/JAM2339.1, 2006.

Pleim, J. E.: A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: Model description and testing, J. Appl. Meteorol. Clim., 46, 1383–1395, doi:10.1175/JAM2539.1, 2007a.

Pleim, J. E.: A combined local and nonlocal closure model for the atmospheric boundary layer. Part II: Application and evaluation in a mesoscale meteorological model, J. Appl. Meteorol. Clim., 46, 1396–1409, doi:10.1175/JAM2534.1, 2007b.

Pleim, J. E. and Gilliam, R.: An indirect data assimilation scheme for deep soil temperature in the Pleim–Xiu Land Surface Model, J. Appl. Meteorol. Clim., 48, 1362–1376, doi:10.1175/2009JAMC2053.1, 2009.

Pleim, J. E. and Xiu, A.: Development of a land surface model. Part II: Data assimilation, J. Appl. Meteorol., 42, 1811–1822, 2003.

Sahu, L. K., Kondo, Y., Moteki, N., Takegawa, N., Zhao, Y., Cubison, M. J., Jimenez, J. L., Vay, S., Diskin, G. S., Wisthaler, A., Mikoviny, T., Huey, L. G., Weinheimer, A. J., and Knapp, D. J.: Emission characteristics of black carbon in anthropogenic and biomass burning plumes over California during ARCTAS-CARB 2008, J. Geophys. Res., 117, D16302, doi:10.1029/2011JD017401, 2012.

Saide, P. E., Carmichael, G. R., Spak, S. N., Minnis, P., and Ayers, J. K.: Improving aerosol distributions below clouds by assimilating satellite-retrieved cloud droplet number, P. Natl. Acad. Sci. USA, 109, 11939–11943, available at: http://www.pnas.org/content/109/30/11939.short (last access: 16 August 2014), 2012.

Saide, P. E., Carmichael, G. R., Liu, Z., Schwartz, C. S., Lin, H. C., da Silva, A. M., and Hyer, E.: Aerosol optical depth assimilation for a size-resolved sectional model: impacts of observationally constrained, multi-wavelength and fine mode retrievals on regional scale analyses and forecasts, Atmos. Chem. Phys., 13, 10425–10444, doi:10.5194/acp-13-10425-2013, 2013.

Samset, B. H., Myhre, G., Schulz, M., Balkanski, Y., Bauer, S., Berntsen, T. K., Bian, H., Bellouin, N., Diehl, T., Easter, R. C., Ghan, S. J., Iversen, T., Kinne, S., Kirkevåg, A., Lamarque, J.-F., Lin, G., Liu, X., Penner, J. E., Seland, Ø., Skeie, R. B., Stier, P., Takemura, T., Tsigaridis, K., and Zhang, K.: Black carbon vertical profiles strongly affect its radiative forcing uncertainty, Atmos. Chem. Phys., 13, 2423–2434, doi:10.5194/acp-13-2423-2013, 2013.

Sandu, A., Daescu, D. N., Carmichael, G. R., and Chai, T.: Adjoint sensitivity analysis of regional air quality models, J. Comput. Phys., 204, 222–252, doi:10.1016/j.jcp.2004.10.011, 2005.

Schulz, M., Textor, C., Kinne, S., Balkanski, Y., Bauer, S., Berntsen, T., Berglen, T., Boucher, O., Dentener, F., Guibert, S., Isaksen, I. S. A., Iversen, T., Koch, D., Kirkevåg, A., Liu, X., Montanaro, V., Myhre, G., Penner, J. E., Pitari, G., Reddy, S., Seland, Ø., Stier, P., and Takemura, T.: Radiative forcing by aerosols as derived from the AeroCom present-day and pre-industrial simulations, Atmos. Chem. Phys., 6, 5225–5246, doi:10.5194/acp-6-5225-2006, 2006.

Schwartz, C. S., Liu, Z., Lin, H.-C., and McKeen, S. A.: Simultaneous three-dimensional variational assimilation of surface fine particulate matter and MODIS aerosol optical depth, J. Geophys. Res., 117, D13202, doi:10.1029/2011JD017383, 2012.

Schwartz, C. S., Liu, Z., Lin, H.-C., and Cetola, J. D.: Assimilating aerosol observations with a "hybrid" variational-ensemble data assimilation system, J. Geophys. Res.-Atmos., 119, 4043–4069, doi:10.1002/2013JD020937, 2014.

Schwartz, J., Coull, B., Laden, F., and Ryan, L.: The effect of dose and timing of dose on the association between airborne particles and survival, Environ. Health Persp., 116, 64–69, doi:10.1289/ehp.9955, 2007.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., and Powers, J. G.: A description of the advanced research WRF version 3, Tech. rep., DTIC Document, available at: http://nldr.library.ucar.edu/repository/collections/TECH-NOTE-000-000-000-855 (last access: 28 January 2014), 2008.

Streets, D. G., Bond, T. C., Carmichael, G. R., Fernandes, S. D., Fu, Q., He, D., Klimont, Z., Nelson, S. M., Tsai, N. Y., Wang, M. Q., Woo, J.-H., and Yarber, K. F.: An inventory of gaseous and primary aerosol emissions in Asia in the year 2000, J. Geophys. Res.-Atmos., 108, D20, GTE30/1–GTE30/23, doi:10.1029/2002JD003093, 2003.

Suutari, R., Amann, M., Cofala, J., Klimont, Z., Posch, M., and Schöpp, W.: From economic activities to ecosystem protection in Europe – an uncertainty analysis of two scenarios of the RAINS integrated assessment model, Tech. Rep. CIAM/CCE Rep. 1/2001, Int. Inst. for Appl. Syst. Anal., Laxenburg, Austria, 2001.

Textor, C., Schulz, M., Guibert, S., Kinne, S., Balkanski, Y., Bauer, S., Berntsen, T., Berglen, T., Boucher, O., Chin, M., Dentener, F., Diehl, T., Easter, R., Feichter, H., Fillmore, D., Ghan, S., Ginoux, P., Gong, S., Grini, A., Hendricks, J., Horowitz, L., Huang, P., Isaksen, I., Iversen, I., Kloster, S., Koch, D., Kirkevåg, A., Kristjansson, J. E., Krol, M., Lauer, A., Lamarque, J. F., Liu, X., Montanaro, V., Myhre, G., Penner, J., Pitari, G., Reddy, S., Seland, Ø., Stier, P., Takemura, T., and Tie, X.: Analysis and quantification of the diversities of aerosol life cycles within AeroCom, Atmos. Chem. Phys., 6, 1777–1813, doi:10.5194/acp-6-1777-2006, 2006.

Turner, A. J., Henze, D. K., Martin, R. V., and Hakami, A.: The spatial extent of source influences on modeled column concentrations of short-lived species: modeled source influences, Geophys. Res. Lett., 39, L12806, doi:10.1029/2012GL051832, 2012.

Turner, M., Henze, D., Hakami, A., Zhao, S., Resler, J., Carmichael, G., Stanier, C., Baek, J., Sandu, A., Russell, A., Jeong, G., Capps, S., Percell, P., Pinder, R., Napelenok, S., Bash, J., and Chai, T.: Differences between magnitudes and health impacts of BC emissions across the US using 12 km scale seasonal source apportionment, submitted, 2015.

Twomey, S.: The influence of pollution on the shortwave Albedo of clouds, J. Atmos. Sci., 34, 1149–1152, doi:10.1175/1520-0469(1977)034<1149:TIOPOT>2.0.CO;2, 1977.

UC-Davis: Interagency Monitoring of Protected Visual Environments Quality Assurance Project Plan, Tech. rep., available at: http://vista.cira.colostate.edu/improve/Publications/QA_QC/IMPROVE_QAPP_R0.pdf (last access: 16 November 2012), 2002.

Vogel, B., Vogel, H., Bäumer, D., Bangert, M., Lundgren, K., Rinke, R., and Stanelle, T.: The comprehensive model system COSMO-ART – Radiative impact of aerosol on the state of the atmosphere on the regional scale, Atmos. Chem. Phys., 9, 8661–8680, doi:10.5194/acp-9-8661-2009, 2009.

Wang, J., Xu, X., Henze, D. K., Zeng, J., Ji, Q., Tsay, S.-C., and Huang, J.: Top-down estimate of dust emissions through integration of MODIS and MISR aerosol retrievals with the GEOS-Chem adjoint model, Geophys. Res. Lett., 39, L08802, doi:10.1029/2012GL051136, 2012.

Wesely, M. L.: Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models, Atmos. Environ., 23, 1293–1304, doi:10.1016/0004-6981(89)90153-4, 1989.

Wiedinmyer, C., Quayle, B., Geron, C., Belote, A., McKenzie, D., Zhang, X., O'Neill, S., and Wynne, K. K.: Estimating emissions from fires in North America for air quality modeling, Atmos. Environ., 40, 3419–3432, doi:10.1016/j.atmosenv.2006.02.010, 2006.

Wiedinmyer, C., Akagi, S. K., Yokelson, R. J., Emmons, L. K., Al-Saadi, J. A., Orlando, J. J., and Soja, A. J.: The Fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning, Geosci. Model Dev., 4, 625–641, doi:10.5194/gmd-4-625-2011, 2011.

Xiao, Q., Kuo, Y.-H., Ma, Z., Huang, W., Huang, X.-Y., Zhang, X., Barker, D. M., Michalakes, J., and Dudhia, J.: Application of an Adiabatic WRF Adjoint to the investigation of the May 2004 McMurdo, Antarctica, severe wind event, Mon. Weather Rev., 136, 3696–3713, doi:10.1175/2008MWR2235.1, 2008.

Xiu, A. and Pleim, J. E.: Development of a land surface model. Part I: Application in a mesoscale meteorological model, J. Appl. Meteorol., 40, 192–209, 2001.

Yelverton, T. L. B., Hays, M. D., Gullett, B. K., and Linak, W. P.: Black carbon measurements of flame-generated soot as determined by optical, thermal-optical, direct absorption, and laser icandescence methods, Environ. Eng. Sci., 31, 209-215, doi:10.1089/ees.2014.0038, 2014.

Zhang, X., Huang, X.-Y., and Pan, N.: Development of the upgraded tangent linear and adjoint of the weather research and forecasting (WRF) model, J. Atmos. Ocean. Tech., 30, 1180–1188, doi:10.1175/JTECH-D-12-00213.1, 2013.

			WRFPLUS	WRFPLUS-Chem	Future Work
solve_em	Subgrid Processes	Radiation			X
		SFCLAY (options 91 & 7)		X	
		LSM (options 1 & 7)		X	
		Dry dep. velocity only		X	
		PBL (NEW: option 7 w/ dry dep)	X	X	
		Cumulus convection	X		X
	Advection		X		
	Microphysics (need aerosols)		X		X
chem_driver	Emissions			X	
	Aerosol optical obs. operators				X
	Dry dep. velocity and vertical mixing				
	Cumulus convection of chem. species				X
	Gas chem. (sulfate precursors)			X	
	Aerosol chem. (GOCART)			X	
	Sum PM			X	

Table 1. Status of AD/TL development for WRF-Chem processes.

~~Dependencies between WRF, WRF-Chem, WRFPLUS AD/TL, and WRFDA. AD/TL development status is also noted.~~

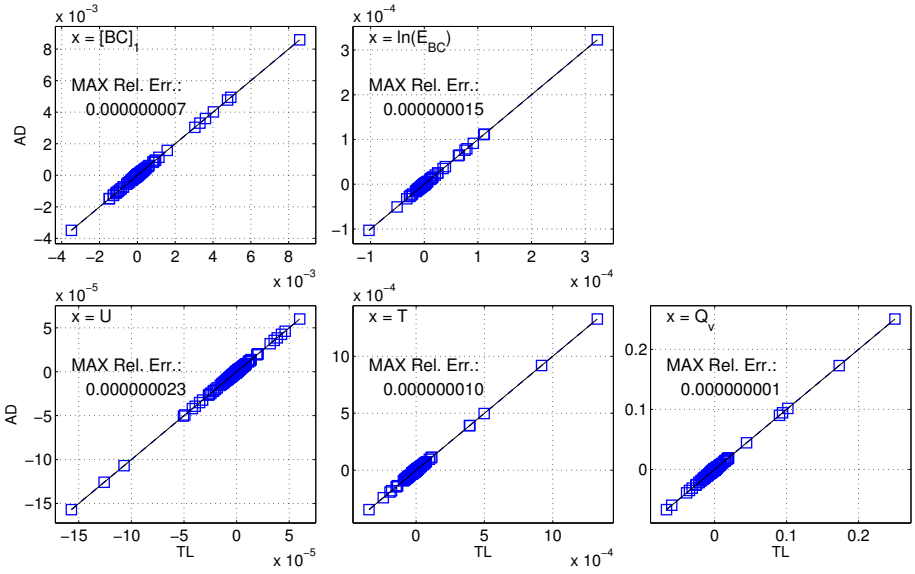

Figure 1. Comparison of ADM to TLM evaluations of $\frac{\partial[\mathrm{BC}_1]}{\partial \boldsymbol{x}}$ and $\frac{\partial[\mathrm{BC}_2]}{\partial \boldsymbol{x}}$ for 300 derivatives for each denominator variable.

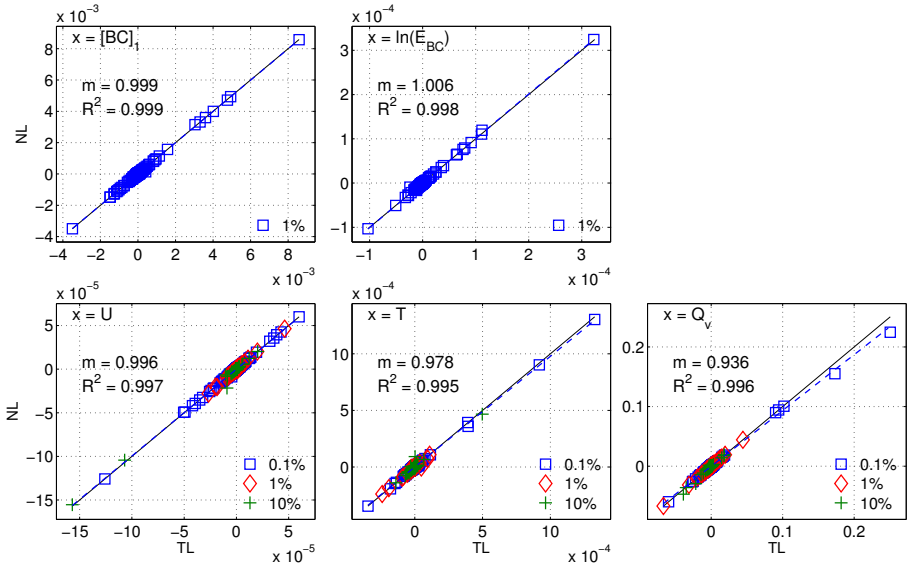

Figure 2. Comparison of nonlinear finite difference approximations to TLM evaluations of $\frac{\partial[\mathrm{BC}_1]}{\partial \boldsymbol{x}}$ and $\frac{\partial[\mathrm{BC}_2]}{\partial \boldsymbol{x}}$ for 300 derivatives for each denominator variable. The different markers for $\boldsymbol{x} = [U, T, Q_v]$ indicate the δx percentage that yielded a finite difference derivative closest to the tangent linear value. The slope (m) and R^2 statistic for the linear fit are shown for each CV.

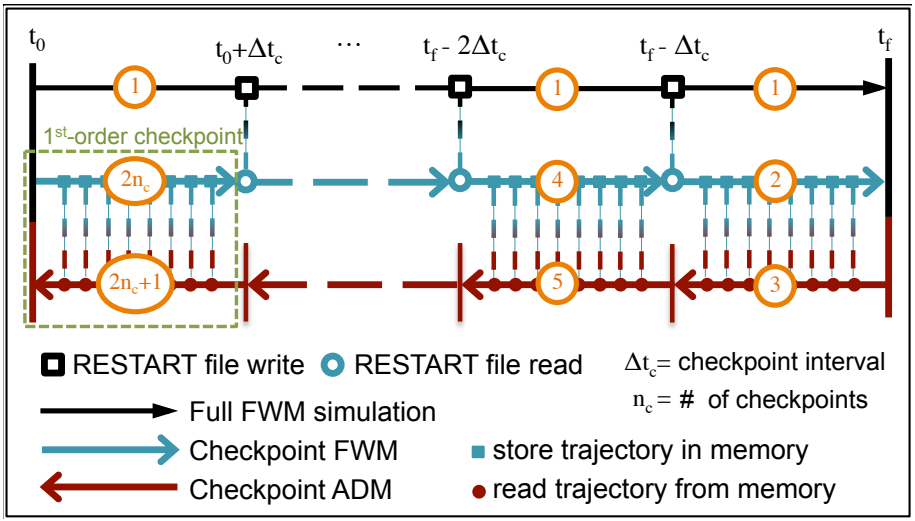

Figure 3. Second order checkpointing scheme implemented in WRFPLUS-Chem.

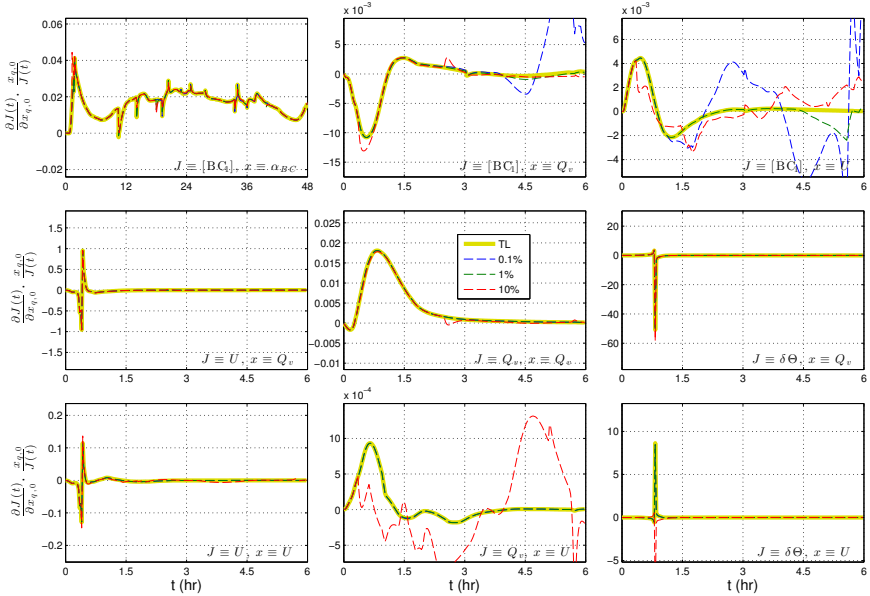

Figure 4. ~~Time~~ Fully normalized time variant sensitivities ~~of cost function J~~ calculated with ~~respect to control variable x for multiple perturbations and~~ the TLM with second order checkpointing and with multiple finite difference perturbation sizes. Each plot is for a single pair of source and receptor locations, q and p.

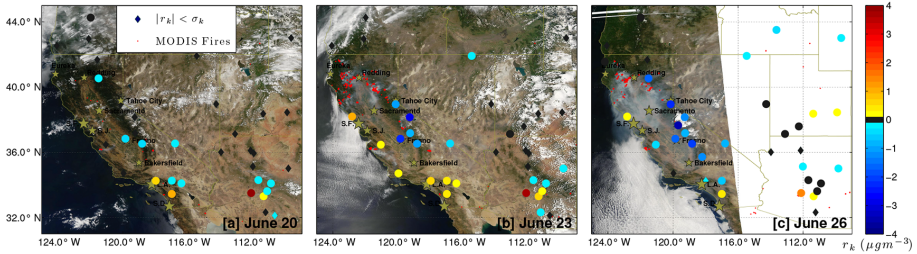

Figure 5. Surface site residual model error, r_k, overlaid on MODIS Aqua true color images and active fire retrievals. Observations with a bias less than one SD are also indicated.

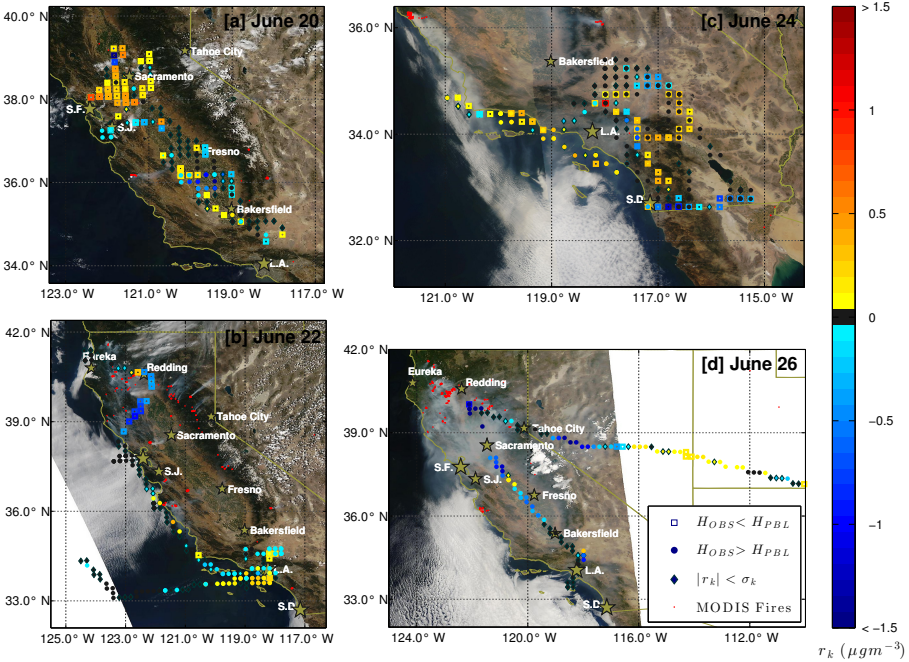

Figure 6. Aircraft residual model error, r_k, with indication for the observation height relative to the model PBL height overlaid on MODIS Aqua true color images and active fire retrievals. Observations with a bias less than one SD are also indicated.

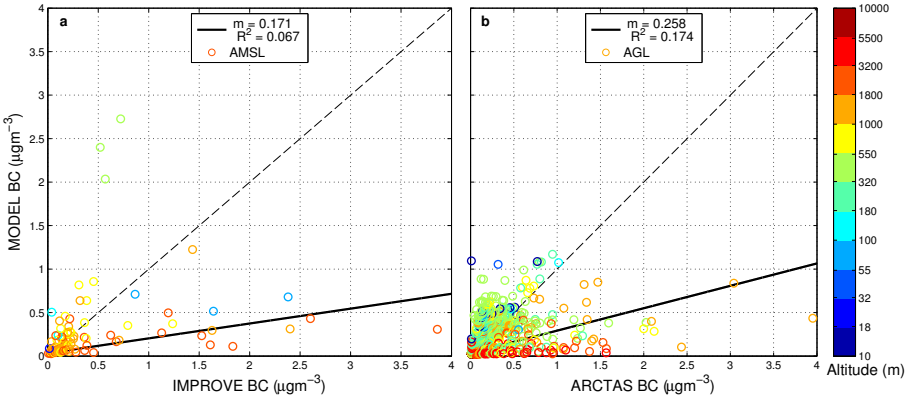

Figure 7. Linear fits between model BC concentrations with slope (m) and coefficient of determination R^2 for **(a)** IMPROVE surface and **(b)** ARCTAS-CARB aircraft observations colored by model height above mean sea level (a.m.s.l.) and above ground level (a.g.l.).

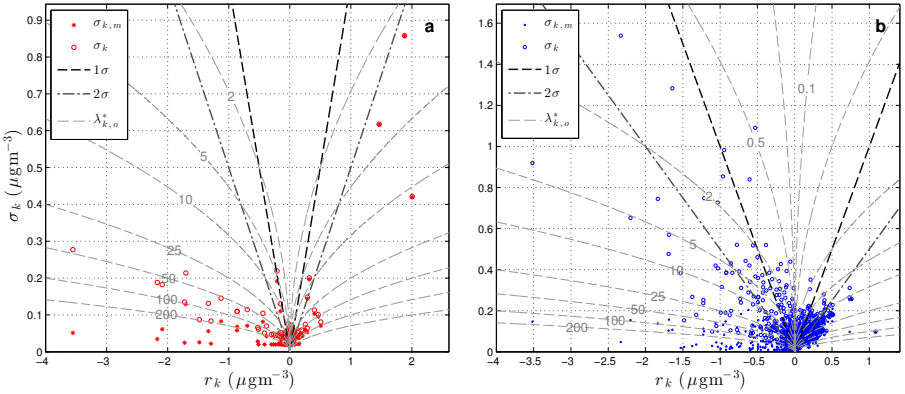

Figure 8. Model and total observation error SD ($\sigma_{k,m}, \sigma_k$) ~~vs.~~ versus model residual error (r_k) with adjoint forcing ($\lambda^*_{k,o}$) contours corresponding to $w_k = 1$ for **(a)** surface and **(b)** aircraft observations. 1σ and 2σ zones reflect regions of increasing statistical significance.

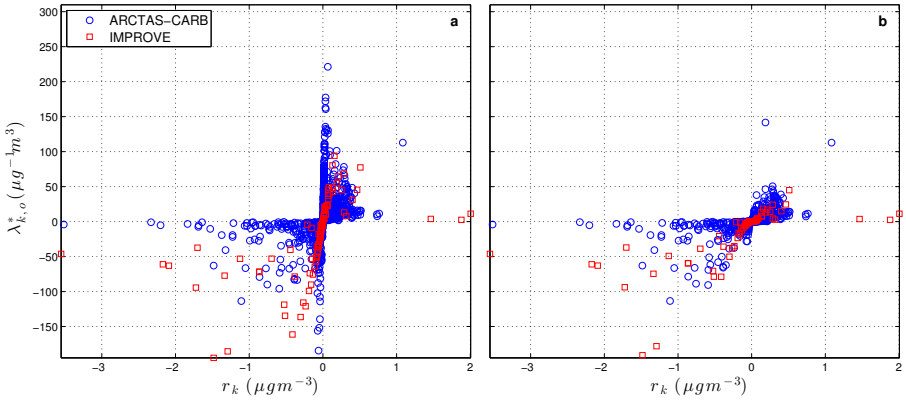

Figure 9. Adjoint forcing ($\lambda^*_{k,o}$) ~~vs.~~ versus residual error (r_k) for ARCTAS and IMPROVE observations using weights of **(a)** $w_k = 1$ and **(b)** w_k from Eq. (27).

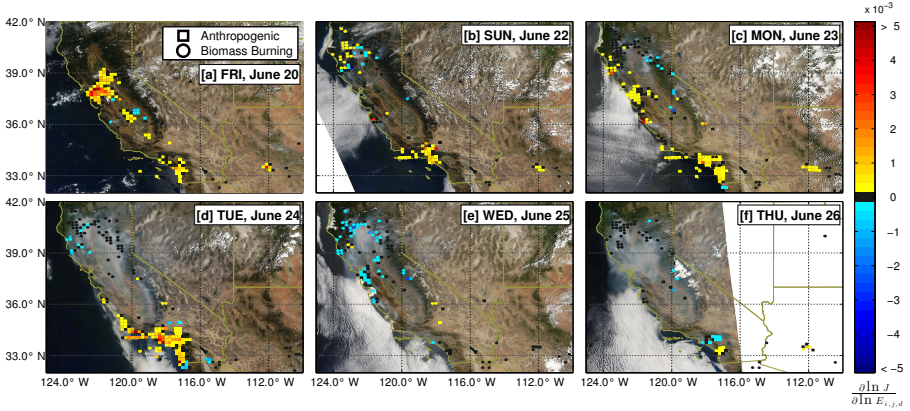

Figure 10. Normalized sensitivities ($\frac{\partial \ln J}{\partial \ln E_{i,j,d}}$) of the 4D-Var cost function (for surface and aircraft observations) with respect to anthropogenic and burning emission scaling factors overlaid on MODIS Aqua true color images for six days during the simulation. Anthropogenic sensitivities with magnitudes less than 1 % of the maximum anthropogenic sensitivity magnitude are removed. There is a marker for all grid cells with non-zero burning emissions.

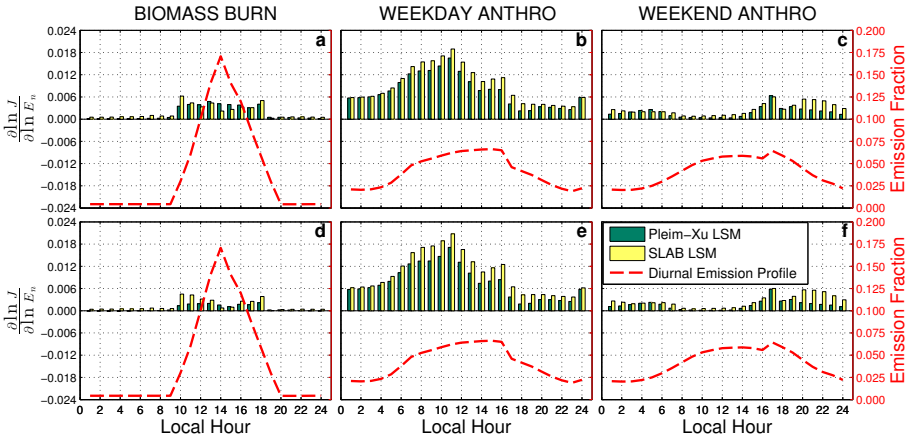

Figure 11. Diurnal normalized sensitivities ($\frac{\partial \ln J}{\partial \ln E_n}$) of the 4D-Var cost function with respect to emissions scaling factors for (**a**, **b**, and **c**) $w_k = 1$ and (**d**, **e**, and **f**) w_k from Eq. (27). Also plotted are diurnal emission fractions. Sensitivities were calculated for two different WRF LSM options and are shown separately for biomass burning, and weekend and weekday anthropogenic emissions.