Atmos. Chem. Phys. Discuss., 13, C4771–C4782, 2013
www.atmos-chem-phys-discuss.net/13/C4771/2013/

Atmospheric Chemistry and Physics
Discussions

Interactive comment on "Multi-year objective analyses of warm season ground-level ozone and $PM_{2.5}$ over North America using real-time observations and Canadian operational air quality models." *by* A. Robichaud and R. Ménard

Anonymous Referee #1

Received and published: 12 July 2013

Robichaud and Menard create a set of objective analysis data which combine modeled and measured North American ozone and PM2.5 over the period of a decade. They then use this data to describe and explore air pollution climatology, multi-year trends, and correlations to various meteorological and economic factors. The methods for creating their objective analysis appear to be technically sound and their work provides a valuable and extensive dataset which will be of interest to ACP readers. However, the text of the manuscript needs editing and I have some concerns about the metrics that they use for analyzing climatology and trends. This paper would also benefit from

including discussions of and comparisons to other relevant/similar work which is not currently cited (specific references given below). In addition, I have numerous specific comments about the authors' presentation of and interpretation of their results. After revisions, I think that this manuscript has the potential to substantially add to the scientific literature. Please see general and specific comments below.

General Comments:

The manuscript is in need of copy editing. There were numerous grammatical mistakes and awkward sentences

Sections 2.1-2.3 should be re-written so that they are understandable to the general ACP audience, many of whom are not data assimilation specialists. The authors should include a concise, high-level explanation of the method. Clearer links need to be made between the mathematical language/notation and the actual atmospheric data being analyzed. Section 2.1 seeks to give an abbreviated introduction to optimal interpolation including the notation and equations used for this data assimilation technique. Unfortunately in an attempt to summarize more thorough descriptions of these methods such as what is found in the Kalnay text, this section ends up being unnecessarily confusing for readers who are not as familiar with data assimilation techniques. Before jumping right into equation (1) the authors should begin by giving the observation equation and the state equation so that it is clear throughout this section what is meant by the observation error and the "background" error. To clarify the presentation it would also help to provide the dimensions of the vectors and matrices in equation 1. First of all, x_aˆn is a vector, correct? Describe for the reader the dimension of this vector in terms of number of grid cells and number of variables in the model output so that it is clear how this equation applies to your hourly CHRONOS and GEM-MACH output (you do not have to use specific numbers, it is just to clarify the matrix algebra in equation (1)). The state vector in OI can include multiple chemical species or meteorological variables across space, i.e. x_fˆn could be the vectorized version of gridded model output for 10s or 100s of variables. In this case the background error correlation is not just a function of distance in space but also defines how these variables are correlated with one another. From your description and the format of equations (3) and (4) it sounds like x_fˆn in your case is for a single pollutant. Again providing the dimension of x_fˆn would clarify this early on.

Throughout the paper, the authors present ozone results which include all hours in the summer season. I do not think this is the proper metric for ozone which is known to have a strong diurnal cycle and is regulated based on daily maximum values (generally max of 1-h or 8-h average for each day). I strongly suggest that the authors switch their analysis to look at climatology, correlations, and trends for the daily maximum 8-h average rather than the all hours average that they have been using. There are several reasons for this suggestion:

**Both Canadian and US regulations are based on the maximum daily 8-h average and their emissions reductions are targeted at this metric not at the 24-h average that the authors analyze.

**Most epidemiology studies use the 8-h daily max rather than 24-h average in correlating health outcomes to ozone.

**Daytime ozone values are likely to decrease in response to emissions control strategies while nighttime values are likely to increase in response to these same emissions reductions (due to ozone disbenefits from NOx which are common at night as well as during wintertime months). Consequently, any emissions-related trends which include both day and nighttime values are likely to be dampened relative to trends in daytime ozone. This behavior may also explain some of the reduction in standard deviation that the authors report, as increasing low nighttime values while decreasing high daytime values would cause such a result.

**Model biases are often opposite during the day and at night (for instance if the model has too little NOx, then daytime values will be underestimated while nighttime values will be overestimated). This makes interpretation of OA and model comparisons diffi-

cult.

**Due to the regulatory focus and daily maximum 8-h ozone values, most researchers also focus on this metric. Looking at climatology and trends in 24-h ozone will not be intuitive to most ACP readers or to the scientific community as a whole.

Section 5.1 deals with changes in average ozone. For ozone we are much more interested in the behavior of high values than in changes of the mean. Figure 11 would be more relevant if it showed changes to 98th percentile ozone or something similar.

Section 5.1: The choice of using two distinct years to map and evaluate air pollution trends leaves the analysis vulnerable to spurious meteorology-driven changes. Even though the years were chosen so that they "both show roughly similar weather regimes over many parts of North America", it is not possible to eliminate met-driven variability between the two years (i.e. no two years are going to have similar meteorology in all locations). It would more technically defensible to base this trends analysis on a multi-year trend to minimize the type of spurious differences that are likely to show up in comparing two distinct years. In figure 13, the authors show multi-year trend lines. It would be better if Figures 11 and 12 were based on location specific trends such as this (perhaps show gridded values as change in ppb/yr or ug/m3/ yr based on the same analysis used to create the trend lines in Fig 13). Additionally, the change in models between 2005 and 2012 adds another variable to the comparison of these two years. The year-to-year changes in Figure 13 make it look like perhaps the model switch (between 2009 and 2010) had noticeable impact on ozone but not on PM2.5. A comparison between the time series in Fig 13 and one based only on monitoring data might reveal whether the jump in ozone values between 2009 and 2010 is based on a real air quality change or is due to the switch in modeling systems.

page 13986, lines 16-19 and Sections 5.1 and 5.2: Although the analysis described here provides some advantages over past trends analyses, there are multiple reports and journal articles that are not mentioned in this article which evaluate trends in ozone and/or PM2.5 based on measurements made at monitors across North America. The US EPA regularly releases trends reports for ozone, PM2.5 and other pollutants (http://www.epa.gov/airtrends/reports.html) as does the corporate institute for research in the atmosphere (CIRA) at Colorado State for PM2.5 (http://vista.cira.colostate.edu/improve/Publications/improve_reports.htm). The article should acknowledge these other efforts and should compare results found here to other studies' findings:

**Our Nation's Air: Status and trends through 2010, US EPA, Office of Air Quality Planning and Standards, Research Triangle Park, NC, February 2012, EPA-454/R-12-001: http://www.epa.gov/airtrends/2011/report/fullreport.pdf

**Our Nation's Air: Status and trends through 2008, US EPA, Office of Air Quality Planning and Standards, Research Triangle Park, NC, February 2010, EPA-454/R-09-002: http://www.epa.gov/airtrends/2010/report/fullreport.pdf

**Interagency Monitoring of Protected Visual Environments, Spatial and seasonal patterns and temporal variability of haze and its constituents in the United States, Report V, June 2011, ISSN 0737-5352-87: http://vista.cira.colostate.edu/improve/Publications/Reports/2011/PDF/IMPROVE_V_FullRep

**Murphy, D.M., Chow, J.C., Leibensperger, E.M., Malm, W.C., Pitchford, M., Schichtel, B.A., Watson, J.G., White, W.H., Decrease in elemental carbon and fine particle mass in the United States, Atmospheric Chemistry and Physics, 11, 4679-4686, 2011.

**Hand, J.L., Schichtel, B.A., Malm, W.C., Pitchford, M.L., Particulate sulfate ion concentration and SO2 emission trends in the United States from the early 1990s through 2010, Atmospheric Chemistry and Physics, 12, 10353-10365, 2012.

Specific Comments:

Page 13970, line 10 and line 13: Technically the term "aerosol" refers to the mixture of gas and particles, not just the particles themselves. Also, particle comprising PM2.5

are often not just "solid" but include a liquid phase as well (both aqueous and organic phases may be liquid).

Page 13970, line 11-13: Secondary formation is also a major source. This sentence implies that all particles are primary in nature.

Page 13970, line 17-21: The authors leave off the most serious health effect that has been linked to PM2.5, death. As they state later in section 5.1.2, a 10 ug/m3 change in PM2.5 has been associated with a 6% change in death rate. It would be appropriate to mention that health outcome here.

Table 1 is unnecessary.

Page 13973, lines 2-13, page 13986, lines 16-19, and Page 14000, line 27-page 14001 line 2: The US EPA in collaboration with the US CDC have undertaken a similar project in which they have used a hierarchical bayesian model to combine air quality model results and measured concentrations of ozone and PM2.5 at 12km and 36km resolutions to look at multi-year trends (2001-2008 with more years in progress). Information about and data from this project is publically available (http://www.epa.gov/heasd/research/cdc.html). Several journal articles have been published on the hierarchical bayesian model developed for this purpose:

**Berrocal, V.J., Gelfand, A.E., Holland, D.M, A Spatio-temporal downscaler for output from numerical models, Journal of Agricultural, Biological, and Environmental Statistics, 15 (2), 167-197, 2010.

**McMillan, N.J., Holland, D.M., Morara, M., Feng, J. Combining numerical model output and particulate data using Bayesian space-time modeling, Environmetrics, 21, 48-65, 2010

**Berrocal, V.J., Gelfand, A.E., Holland, D.M. Space-Time data fusion under error in computer model output: an application to modeling air quality, Biometrics, 68, 837-848, 2012

Page 13976, line 23: References to Wikipedia are not acceptable in a peer-reviewed article. Please find a more reliable reference.

Page 13979, line 27 – Page 13980, line 4: What are the implications of assuming homogeneous background error given that ozone and PM2.5 have large spatial gradients in urban areas which cannot be fully captured by a 15km grid resolution? In addition to the rural versus urban difference it seems likely that across such a large domain the correlation structure of the errors probably changes in coastal areas vs interior locations, high elevation vs flat etc. A brief discussion of the implications of the assumption of homogenous background error would be appropriate.

Page 13984, line 18: Please add the following references:

**Herring, S. and Cass, G. The magnitude of bias in the measurement of PM2.5 arising from volatilization of particulate nitrate from Teflon filters, J. Air Waste, Manage., 49, 725-733, 1999

**Frank, N.H. Retained nitrate, hydrated sulfates, and carbonaceous mass in Federal Reference Method fine particulate matter for six eastern US cities, J. Air Waste Manage., 56, 500-511, 2006.

Page 13985, line 27: By using 18 UTC instead of the same local time everywhere, the authors have picked a time when ozone concentrations are typically higher on the East Coast (2pm LST) than on the West Coast (11am LST). This artificially inflates eastern ozone compared to western ozone in this figure and is visually misleading. It would be more appropriate to show this comparison either using local time or using daily maximum ozone values.

I suggest that you move Section 4 to come before Section 3. Validation of results should come before any interpretation of those results.

Page 13991, line 8: OA is not a monitoring system.

Page 13991, lines 10-12: Other likely explanations include:

**Lack of dense monitors in these areas making OA more uncertain

**Many studies have shown that coarse model resolution (15km and 21km) cannot capture meteorology in complex terrain. Models often need to be resolved at 4km, 1km or even finer resolutions to capture complex airflows and phenomena such as cold pools in mountainous regions.

**The satellite seems to be able to better capture certain air pollution features than OA, such as the high PM2.5 concentrations in Salt Lake City which are due to cold pool meteorology that regional air quality models have trouble simulating.

Page 13991, line 25 – Page 13992, line 1: I can understand why it is desirable to filter out outlier when they are due to erroneous measurements, but sometime the outliers may be real and important. The US ozone standard is based on the annual 4th highest ozone concentration ($\sim$98th percentile) and the US daily PM2.5 standard is based on the 98th percentile. So it does not seem desirable to filter out the very types of events that are the main focus of regulation and research.

Page 13992, lines 4-5: Again, locations that are close to primary PM sources (factories, agricultural burning etc) are the very locations where health effects from PM are of concern. Why is it desirable to have a technique that “disregard[s] data influenced too heavily by the proximity of local strong sources of PM2.5”? These are the very locations that we are most interested in for protecting public health.

Page 13993, lines 10-20:

**Figure 11 also shows increases in mean ozone in several urban areas (Detroit, Chicago and others). These ozone increases in urban areas may be due to reductions in NOx titration from mobile-source NOx decreases.

**Fire activity is highly variable from year-to-year, so by comparing two discrete years (2005 and 2012) you are more likely to see spurious effects from wildfires that may or may not represent a longer term trend in ozone from fires. Since models pick up strong localized fire effects and wildfires in 2005 and 2012 are not likely to have occurred in the exact same locations, it seems likely that the authors would see both increases and decreases in ozone due to differences in wildfires. If Figures 11 and 12 showed multi-year trends, as previously suggested, any conclusions about ozone due to increasing fire frequency would be more robust.

**The authors state that ozone decreased in the intermountain West, but from Figure 11 it appears that ozone increased in large areas of the intermountain West (CO, UT). The authors should be more specific and list the US states for which ozone decreased.

Page 13993, line 19: The authors give PM2.5 changes in ug/m3 per year based on two years of data. They should base these numbers on multiyear trends such as those shown in Fig 13.

Page 13994, lines 8-10: by "growing local socio-economic and industrial activities" do the authors mean increases in oil and gas drilling? If so, they should state this explicitly. If not, they should consider this as a potential driving force due to the rapid increase of emissions from these operations in the Western US.

Section 5.2: When creating inter-annual trends, it is important to use a consistent set of trends monitors with long-term observation records. Including monitors that operated for only part of the trends period may introduce bias. Have the authors filtered their monitoring data using this criterion before creating their OA surfaces?

Page 13994, lines 11-12: The authors use speculative arguments here about the expectation that wildfires will increase in the West with warming. Data is available on fire activity and burned area in 2005 and 2012. Instead of speculating, the authors should use this data to determine whether/where wildfires were more prevalent in 2012 compared to 2005. They could then make more definitive conclusions on whether wildfire activity actually contributed to the changes that they estimate.

Page 13993, lines 10-20 and Page 13994, lines 21-22: Why do the authors think that

the frequency or magnitude of stratospheric intrusions changed between 2005 and 2012? I am unaware of any studies which suggest that STE is increasing. If there is published literature on this phenomenon then the authors should cite it. If not, it seems very speculative to suggest STE is changing without any reason or evidence to suggest this might be the case. Since intrusions often happen at frontal boundaries, if the weather were quite different in these two years one might expect a difference in stratosphere-troposphere-exchange (STE), but the authors stated previously that these two years were specifically chosen for comparison due to similar weather patterns.

Page 13995, line 2-3: All the more reason to look at peak daily values instead of ozone from all hours, so that the authors can tease out different trends during the day (when ozone is of most concern from a health perspective) and at night.

Page 13995, line 20 - page 13998, line 2: I don't see the merit in trying to create a statistical model from datasets consisting of 9 and 11 obs. It might make sense if they were looking at a 30+ year trend, but the authors are really stretching their data too thin here. I would drop this entire section, and tables 7/8.

Page 13995, line 29: What is the source of the data on the wikimedia site? This site appears to be a data site in the vein of Wikipedia with no quality checks or peer-review. Can the authors find a more reputable sour for their data?

Page 13995, lin23 – Page 13996, line 8: Other studies have also been conducted that relate various meteorological parameters to air quality. How do your findings compare to theirs?

**Camalier, L., Cox, W., Dolwick, P., The effects of meteorology on ozone in urban areas and their use in assessing ozone trends, Atmospheric Environment, 41, 7127-7137, 2007.

**Zheng, J., Swall, J.L., Cox, W.M., Davis, J.M., Interannual variation in meteorologically adjusted ozone levels in the eastern United States: A comparison of two ap-

proaches, Atmospheric Environment, 41, 705-716, 2007.

**Davis, J., Cox, W., Reff, A., Dolwick, P., A comparison of CMAQ-based and observation-based statistical models relating ozone to meteorological parameters, Atmospheric Environment, 45, 3481-3487, 2011.

Page 13996 and 13997: All of the abbreviations make this section hard to read. I suggest eliminating the abbreviations in this section.

Tables 5a and 5b: Please describe the averaging time (all hours, daily max etc.) and units (ppb or %) in the table caption. Is this for summertime only? If so, state this as well.

Table 6: What statistical test was used to determine the p-value? Why is $P < 0.25$ considered significant? Standard practice in the research community is to consider $P < 0.05$ (or 0.01) significant. Please use the 0.05 cutoff or provide a compelling reason for this break from standard practice. It would also be useful if this table were broken out by region like Tables 5a and 5b.

Figures 4-14: Increase the font size. The scales on these tables are completely unreadable using the current font size. Please print out a copy of the article pdf before it is resubmitted to make sure that numbers/scales/axes on your figures are readable.

Figures 11/12: Make the areas labeled "Unreliable zone" white or gray. The red is distracting.

Figure 12: Are these annual or summertime values?

Figure 13: How were trend lines created? Please include a brief explanation in the text of what methods were used to create the decadal trend. Add y-axis labels to the right edges of these plots. This makes it easier to see the magnitude of the increases and decreases in the trend lines.

Figure 14: Consider using black and white (or gray) for the unmasked land/water to

eliminate colors which are similar to those included in the color bar.

Interactive comment on Atmos. Chem. Phys. Discuss., 13, 13967, 2013.