Editor Decision: Reconsider after minor revisions (Editor review) (17 Apr 2014) by William Lahoz

Comments to the Author:

The authors should address the comments from both referees, providing a detailed response. From an editorial point of view, the authors should do the following:

(i) Include at the end of the Introduction a brief paragraph outlining the subsequent discussion in the paper;

Paragraph added: "Section 2 outlines the OSSE framework including a description and comparison of the simulation models used, the present and future observing systems considered, the data assimilation system, and the quantification of the error correlation length scales. Section 3 describes the OSSE results showing improved monitoring of surface ozone across the Intermountain West from TEMPO observations and improved detection of high-ozone events in the Intermountain West by data assimilation. Section 4 presents a case study of a stratospheric intrusion demonstrating the detection of an exceptional ozone event by TEMPO its attribution to the North American background. Section 5 summarizes the results and discusses future research directions."

(ii) Make an effort to quantify statements, such as in L. 371.
Statement quantified: "We show that the TEMPO geostationary observations will provide a greatly improved observing system for monitoring such events, eliminating the *a priori* model bias, capturing 58% of surface MDA8 ozone variability, and capturing 82% of the distribution of high-ozone days."

(iii) Indicate in figure captions what the end range of the colour scale signify.
Added: "Orange and red values indicate ozone levels that would lead to exceedances of the current National Ambient Air Quality Standard (NAAQS) of 75 ppbv."

**Response to Reviewer #2:**

1-The description of the satellite instruments is almost nonexistent in the paper. This is crucial for the reader to know what these instruments are capable of measuring. I would suggest to add a paragraph or/and a table recapitulating the instrument characteristics. The figures with the averaging kernels are not sufficient to understand the impact of the data. For example, in Natraj et al, the averaging kernels are normalized averaging kernels and this is not specified in this paper. The instrument configuration used by Natraj et al. is probably different from the TEMPO one. How the averaging kernels presented in the paper are constructed? I haven't seen them in Natraj paper. I have a similar remark for the atmospheric (Temperature, species input, ..) and surface parameters (albedo, ..) used as input by Natraj et al. Are they relevant for the period and the surface of the OSSE? I would suggest the authors to comment this in the paper.

A comparison of TEMPO specifications has been added to the specifications from Natraj et al. : "The UV+Vis spectral ranges (290-340 nm, 560-620 nm) and spectral resolution (0.4 nm) assumed by Natraj et al. (2011) are comparable to the spectral ranges (290-490 nm, 540-740 nm) and spectral resolution (0.6 nm) planned for TEMPO."

*I may have misunderstood the reply from the authors, but I think the term "comparable" is not adequate in terms of comparison of these two instrument characteristics. The Natraj's instrument characteristics have a shorter band with a higher spectral resolution as mentioned by the authors. First, this already could give a different result in terms of ozone retrievals. Secondly, the visible channel (from Natraj et al) which is sensitive to the surface for ozone has a signal to noise ratio of 3 times the SNR of OMI (about 3000) while the TEMPO one is about 1000. The instrument of Natraj (probably difficult to build now) has a much higher performance than TEMPO. For these reasons the OSSE presented in the paper is definitely not using TEMPO instrument characteristics. So, I strongly suggest to give the instrument characteristics that correspond to the ones used in the OSSE; correct all the sentences mentioning that is TEMPO except if you could demonstrate that the Natraj's characteristics give similar averaging kernels than the TEMPO ones.*

Content added to clarify the relationship between the Natraj et al. simulations and projected TEMPO instrument: "The TEMPO instrument is still under development and thus does not have its characteristics fully finalized; Natraj et al. (2011) gives the published best estimate of TEMPO ozone sensitivities. We expect TEMPO ozone sensitivities to be similar to UV+Vis sensitivities from Natraj et al." We disagree with the statement 'The instrument of Natraj has a much higher performance than TEMPO' since the Natraj et al. simulations considered a very limited Vis spectral range.

The averaging kernel matrices are taken directly from the work by Natraj et al., as indicated in the paper.

*I think to be consistent with Natraj's paper the authors should mention which Natraj's atmospheric profile they used (see table 2 from Natraj's paper). They also should add the term "normalized" in the caption of Figure 2 to make a difference with the averaging kernels used in equation 1 (not normalized).*

Added reference to Natraj profile (see below response)

"Normalized" added to caption.

We have added to the conclusion on the effect of using fixed averaging kernel matrices on our OSSE results (see response to point 4 below).

*Finally, maybe I again misunderstood the reply of the authors but I was asking if the period, surface parameters and other atmospheric parameters used by Natraj et al. are comparable of that proposed in the OSSE. I strongly recommend to present a table with the different relevant parameters (surface albedo, atmosphere etc..) used in Natraj et al. (for example from table 2 in Natraj et al.) with these usually encountered over the US or/and from the AM-3 truth. That will put into the context the hypothesis made by the authors in the OSSE.*

Added discussion of realism of parameters: "The profile (index 5 from Natraj et al. 2011) used to generate these averaging kernels has moderate ozone (58 ppbv), moderate temperature contrast, and an intermediate viewing geometry, making it consistent with conditions in the Intermountain West. The assumed Vis surface albedo may be lower than the actual albedo which would result in an underestimation of TEMPO sensitivity to near-surface ozone." We have mentioned in the Summary that surface albedo is a subject needing additional research.

2-I found the use of the LEO data too quick to be convincing. I did not see if the authors used the nighttime data to conclude that LEO data do not add any significant contribution. The TIR should bring information during nighttime in the free troposphere and from long range transport. But the question is perhaps what is the information brought by the LEO satellite? For example what are the differences between the couple "ground based stations and TEMPO GEO" vs the couple "ground based stations and IASI-3 LEO"? and this for the two OSSEs proposed. I would suggest the authors to present the results of this OSSE to show the relevance of a GEO vs a LEO. We will see the real benefit of TEMPO vs the existing system.

We have attempted to make clear the use of nighttime LEO data twice in Section 2.2: "TIR has the advantage of providing observations at night that will be complementary to TEMPO."

and "We similarly generate synthetic LEO IASI-3 observations over the North American domain twice a day (local noon and midnight)."

We have modified the statement regarding the information provided by a LEO instrument in addition to having TEMPO observations: "The LEO instrument will thus be valuable for tracking transpacific transport of ozone plumes even when TEMPO is operational"

*This is fair enough concerning the addition of a LEO. But for the same reasons as mentioned above for TEMPO, this is not IASI-3 but, if I well understood, this is a typical TIR LEO with a special instrumental configuration taken by Natraj et al. The authors should also clarify this point and do not called it IASI throughout the paper.*

Sentence added to emphasize that the TIR LEO measurements are to be taken as representative of future LEO observations (IASI-3 being the only TIR LEO currently in planning): "These TIR measurements are intended as representative of ozone observations from LEO instruments operational during the TEMPO lifetime."

3-For the high-ozone events in the Intermountain West OSSE, I did not understand why there is no data that cover California. In the CASTNet surface network, there are stations located in California. Are they representative of the background? if not, this is a pity because one or two stations in this region or in the Las Vegas area would be sufficient to give better results with only surface data assimilated. In addition, I find the results of GEOS-Chem model too different from the CCM. Why GEOS-Chem model is so different? By using such simulations, the improvements by assimilating synthetic observations are highlighted too much. Please comment on this in the paper.

We have added a sentence explaining why California CASTNet observations were not used: "CASTNet stations outside of the Intermountain West are not used as they do not provide useful constraints for the region."

Added comment on how the differences between the models affect the OSSE results (see response to point 4 below).

*Maybe my first question was not clear enough but I think the answer is not sufficient. Why California CASTNet data are not useful constraints? This has to be shown. There are stations representative of the outflow of LA basin such as Joshua tree NP which could be interesting to use. It is interesting to know why the authors did not use California data or/and why did they only use the domain as shown in Figure 1 avoiding California? The authors should clarify in more details this point.*

Added further clarification: "CASTNet stations outside of the Intermountain West are not used; we assumed they do not provide useful constraints for the region but it is possible certain California sites might be exceptions."

4-Finally, I think the different assumptions taken by the author make the OSSE very likely overoptimistic. Above all the fix averaging kernel for the full period and the entire West of US area without taking into account the heterogeneity of the surface (surface albedo, surface temperature, etc ) for the GEO and the LEO is somehow questionable for the final results. Because if the OSSE is overoptimistic, how useful is the final result for concluding on a quantification of the benefit from GEO ozone measurements? I would suggest to comment on how overoptimistic (or pessimistic if it is the case) the OSSE could be.

We have added a paragraph to the conclusion on the effect of our assumptions on the OSSE results: “The use of invariant averaging kernel matrices is a limitation of this study. Preparation for TEMPO must include improved constraints on physical parameters, such as surface albedo, that can vary greatly over the North American domain and that affect the sensitivity of UV+Vis retrievals of near-surface ozone. Also, if the differences between the two models used in our OSSE are larger than future errors in modeled ozone, this study may overestimate the information TEMPO will provide.”

*I think the conclusion must be improved by using the comments above. Concerning the differences between the two models, there are probably validation studies between each model and real data. The statistics between real surface data and GEOS-Chem should be comparable to the statistics between AM-3 model and GEOS-Chem to make sure GEOS-Chem and AM-3 ozone outputs are not too far. Also, the authors should argue on how useful is an overoptimistic result from the addition of a GEO in a regional observing system over US?*

Validation studies of AM3 and GEOS-Chem in the Intermountain West are discussed in Section 2.1 (Simulation Models). Sentence added to clarify usefulness: “However, our OSSE demonstrates the large relative improvement of information provided by TEMPO over the current observing system.”