Atmos. Meas. Tech. Discuss., 6, C2207–C2210, 2013
www.atmos-meas-tech-discuss.net/6/C2207/2013/

Atmospheric Measurement Techniques Discussions

Interactive comment on "Semi-autonomous sounding selection for OCO-2" *by* L. Mandrake et al.

Anonymous Referee #2

Received and published: 23 August 2013

In this manuscript, the authors present an algorithm to semi-autonomously create a set of filtering rules for selecting valid soundings from a large data set of satellite-based total column tracer measurements. This was developed keeping in mind the large data volume of the future OCO-2 mission, and tested on XCO2 soundings from the currently operational satellite GOSAT. The paper is very well written and tackles a problem relevant to future satellite-based Earth monitoring missions. I would therefore recommend publication. There are no technical changes or corrections needed as such (except one, mentioned below), but I do have a few questions that I'd like answered.

1 Questions

1. Reduction of the mean monthly scatter (MMS) as defined in equation (1) is close to – if not identical to – the assumption made by *Wunch et al.* (2011) for deriving a global bias correction, and is a physically valid assumption. However, unlike *Wunch et al.* (2011), the authors here do not demonstrate that this assumption "works" globally, i.e., that it results in reasonable selection of soundings elsewhere. Since the authors demonstrate their GA technique with ACOS XCO2 retrievals, it should be easy for them to show that their choice of filtering scheme yields soundings near TCCON stations that match TCCON XCO2 measurements (modulo the ACOS bias correction, perhaps). In other words, I would like to see a demonstration that as they lower the transparency of their filter set, or increase the complexity (i.e., go towards a "better" or "more stringent" filter set), the TCCON-coincident soundings that pass match the TCCON soundings better.

2. In § 4.1 the authors mention that they limit the complexity to two. Is it the case that the two dominant filters in that case are always dPc and CO2 ratio across the entire population of dominant solutions, or could it vary from solution to solution, and it's just that dPc and CO2 ratio happen to be the two most commonly important filters? Also, does a complexity of N always mean N physically meaningful fixed filters (such as SNR, dPc, etc.), or could one of those "degrees" of complexity be a combination of multiple interpretable fixed filters in some vector space?

3. In § 4.2.1, the authors describe the two features that show up as highly significant in Tables 2 and 3. These are features from the IMAP-DOAS preprocessor. From a full physics retrieval, would the authors expect expect a different set of features to be more significant, such as the total aerosol loading or the surface albedo?

4. In the methods so far employed by different groups to filter GOSAT XCO2 (e.g., *Guerlet et al.* (2013) or *Wunch et al.* (2011)), it's very often the case that some of the filters end up co-varying with XCO2, and as a result filtering leads to elimination of certain variations in XCO2 that could be due to surface fluxes. This is usually countered by bias correction schemes to get rid of the covariations. Is there a mechanism in place within the genetic algorithm described in this manuscript to filter by parameters that do not covary with XCO2? If not, how would the authors get rid of those covariances, given that within a population of "most fit" filters, different fixed filters may be the most dominant ones, i.e., the parameters against which one devises a bias correction may vary within the population?

5. In § 4.3, during the discussion of warn levels, are the other filter thresholds kept fixed, i.e., is the complexity always two?

6. In Figure 16, I would expect the fit to full data to increase monotonically with the temporal length of the training data set. Why then are the 10% and 90% transparency curves non-monotonic?

2 Technical corrections

1. Page 5894, line 7: "Intel Xeon" instead of "Intel Xenon"

References

Guerlet, S., et al. (2013), Impact of aerosol and thin cirrus on retrieving and validating XCO2 from GOSAT shortwave infrared measurements, *Journal of Geophysical Research: Atmospheres*, doi:10.1002/jgrd.50332.

Wunch, D., et al. (2011), A method for evaluating bias in global measurements of CO2 total columns from space, *Atmospheric Chemistry and Physics*, *11*(23), doi:10.5194/acp-11-12317-2011.

Interactive comment on Atmos. Meas. Tech. Discuss., 6, 5881, 2013.