Atmos. Meas. Tech. Discuss., 6, C2565–C2567, 2013
www.atmos-meas-tech-discuss.net/6/C2565/2013/

Atmospheric Measurement Techniques Discussions

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

L. Mandrake et al.

lukas.mandrake@jpl.nasa.gov

Received and published: 13 September 2013

Question 1. We do indeed draw the tccon / southern hemisphere approximation equivalence from Debra's work as well as our own. It has been shown internally that the southern hemisphere approximation and the TCCON geographic match comparison yield nearly identical Warn Levels, and thus one will enforce the other. Adding TCCON comparison to this dataset requires more time than is currently available. I appeal to Dr. Wunch's prior work that the southern hemisphere results do indeed generalize over the planet, and have added some text to more strongly indicate this.

Question 2. A most astute question. We have added text to the end of 4.1 discussing the true statement that, initially a complexity two filter need not restrict itself to any

number of input features. Complexity two mearly indicates that only two at any transparency value are used simultaneously. Later, in section 4.2, we detail reducing the number of input features to match the desired filter complexity through the process of feature selection.

Question 3. This is outside the paper's sounding selection focus, but entirely germain to the concept. Indeed, we will be doing just what you suggest to determine sounding retrieval quality in the full product for OCO2. It turns out, through empirical trial, that co2_ratio is a dominant choice even when competing with the full set of the physics-based retrieval outputs. However, it is then more common to observe albedo_slope in a particular band as an excellent second pairing, although dP_cloud yet remains a strong contender.

Question 4. The question of covariance of XCO2 with the input features is most significant. We do indeed make no discussion of bias/covariance correction in this paper, as that is a larger topic deserving of many papers. Instead, during the filter creation phase between 4.1 and 4.2, there is a human interaction moment in which we attempt to estimate the Restrictiveness of the chosen features. One of the ways that Restriction would be present is if the xco2 and feature are highly correlated. We leave that to the human analyst to acknowledge which of the chosen features are most/least likely to be problemmatic in "all ways s/he knows" and choose the ones with the most desirable properties. This precise moment is why the method is only semi-autonomous and not a push-button operation. We have added more text to 4.2.1 making this addition to Restriction clearer.

Question 5. Yes, once fixed, the complexity of the filter is always two. However, we move from the space of 1000 individual thresholds as calculated by the genetic algorithm to a sub-sampling of only 19.

Question 6. What you suggest, monotinicity in warn level agreement with respect to increasing temporal scale, would be true if a slowly moving window of smaller size were

evaluated at every possible point sliding along the entire dataset. We instead chopped the dataset up into non-overlaping sections. Furthermore, due to irregularities in the data distribution temporaly, any subwindow that contained a large gap was thrown out so as not to over-bias those few points. This does then permit fluctuations in agreement given a small perentage of data being included/excluded as differing temporal scales are used. Figure 16 is only intended to give a general idea of the integrity loss due to data starvation. The details of the curve, while interesting, are not interpretted deeply here.

Many thanks for the technical corrections. They have all been accepted into the new version.

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