

## ***Interactive comment on “A simple empirical model estimating atmospheric CO<sub>2</sub> background concentrations” by M. Reuter et al.***

**Anonymous Referee #1**

Received and published: 6 April 2012

The authors present a simple empirical model (called "SECM") to generate mixing ratio profiles for atmospheric CO<sub>2</sub>. The SECM could then provide a traceable a priori for remote sensing retrievals. The research presented is not particularly novel: for several years, the TCCON science team has been using a similar approach to determine a priori profiles for CO<sub>2</sub> based on GLOBAL-VIEW and balloon data. Given that the documentation for the TCCON approach for generating profiles is not widely available, it may be appropriate to document the approach used by Reuter et al. in AMT.

To merit publishing the authors need to give more consideration to actual data. Throughout the paper, the authors use CarbonTracker (CT2010) as "truth". I think it is important to remember that CarbonTracker is, in itself, an assimilation product. Just as the authors discuss the fact that remote sensing observations don't "see" all the at-

C563

mosphere and rely on a prior to fill in vertical information to which the observation has minimal sensitivity, CarbonTracker also relies on a prior set of fluxes that essentially smooth the information provided by the assimilated observations.

I have a couple comments on the way the authors have parameterized the profile shape in the SECM. On page 1297, they state that "The parameters  $c_0$  and  $c_1$  are determined similarly as  $a_{00}$ - $a_{14}$  by least squares fitting to CT2010 mixing ratio profiles." This means that the vertical mixing bias in TM5 is propagated into SECM. It appears then that the authors have not validated the shapes of their profiles against data. Such data is available (e.g., HIPPO, NOAA GMD aircraft program, ACE satellite). Instead, the authors both train the SECM with and primarily validate the SECM against CarbonTracker.

The shapes of the profiles could be improved if variations in tropopause height were taken into account. Although this adds complexity to the model, for remote sensing retrievals this is particularly important because the averaging kernel often shows a lot of curvature in the vicinity of the tropopause (see Fig. 5).

On page 1299, the authors state that the error covariance matrix can only be a reasonable approximation of the total error if the total error is dominated by the differences between SECM and CT2010. This assumption is testable using TCCON data. In fact, there are biases in the seasonal amplitude in Fig. 4, which suggests that the total error is not dominated by differences between SECM and CT2010 (assuming that the CT2010 timetraces overlaid on Figure 4 would look similar to the red SECM traces). It appears that the SECM overestimates summer CO<sub>2</sub> at the northern hemisphere sites of Bialystok, Bremen, and Park Falls. At these locations, the phasing of the seasonal oscillation also seems different, with the spring drawdown and the fall increase too "slow" compared to the data. At Darwin, the curvature in the SECM time traces does not agree with the data well. The authors should expand their discussion of these differences since this is the only observational data used to validate the SECM.

C564

The rest of the validation occurs against CarbonTracker profiles. Given that the TCCON comparison shows that there are seasonal biases in the comparisons, and that there is quite different behavior in the northern hemisphere, the tropics, and the southern hemisphere, it would be helpful if the authors broke down how well the SECM compares against CT2010 in different regions, rather than simply aggregating 10,000 randomly chosen global profiles to give the statistics at the end of section 2.

For the TCCON comparison the authors have trained the SECM through 2009, and allow the model to predict the values for 2010. Could the same SECM parameters also reasonably give priors for 2012, or will the model need to be re-trained with data up to 2011? If this is the case, the utility becomes less since frequent retraining may hinder the authors' stated goal of a "simple, traceable a priori" or conflict with the author's conclusion that "SECM is always available".

Minor points: Beginning in the abstract, but repeated in the Conclusions, the authors state that the SECM can be used as a quick check for obvious retrieval errors. This sort of thinking is dangerous. Retrieval errors should be identified based on the statistics of the fit and the ensemble of retrievals, not by discarding retrievals that deviate from the prior.

In the abstract: TCCON stands for Total Carbon Column Observing Network. The authors have reversed the two C's in the name and the first letters should be capitalized. (again on page 1295)

The authors conclude with a statement that "it is remarkable how well a simple empirical equation depending only on date and latitude can reproduce atmospheric CO<sub>2</sub> concentrations". I think the authors are somewhat overselling their model here, and need to keep in mind that CO<sub>2</sub> is quite well mixed and that it is variations of order 1-2% in atmospheric carbon dioxide from which seasonal and net carbon fluxes are inferred.

---

Interactive comment on Atmos. Meas. Tech. Discuss., 5, 1293, 2012.