Dear referee#3, thank you again for your thoughts.
(comment and answers of the first response are cursive, new answers are marked in red)

I will follow up here with only those issues I feel are unresolved from the authors first response.

Main Issues 3. Reading of the paper suffers from the lack of an at least, semiquantitative definition of the term 'long-term consistency'. It is easy enough for the reader to have a view of what this is, but since it is the primary point of the paper it should be defined probably in the introduction.

We want to check our datasets for shifts or biases caused by e.g. instrumental failure or change of instruments for the whole dataset of all used sites. The duration of these datasets varies from several to up to 20 years. We also want to compare datasets recorded continuously and datasets that have been only recorded during campaigns

Right, will this statement be made more complete and added into the manuscript? If so then this issue will be considered resolved.

> In our opinion, a good description of 'long-term consistency' in our manuscript would be "multiyear stability of instrument precision and accuracy"
> We'll add this to complete the statement above and add it into the manuscript.

4. The most important issue: The 'NDACC-IR retrieval' used here employs profile scaling even when the data have high spectral resolution in the MIR. It is shown that the scaled single a-priori (WACCM) inflicted a bias. Why was the NDACC-IR standard procedure of profiling with appropriate constraints not used? This seems a glaring misuse of the standard NDACC-IR techniques when its primary attribute could be used to obvious advantage. Section 4.2 discusses the difficulty of gaining the seasonal cycle
amplitudes by not employing the MIR data to its full extent. Here a correction is found to mitigate it but could it not have been avoided?

We agree that a profiling retrieval would likely improve the sensitivity of the NDACC XCO2 product. Such further development of our here presented approach could be presented in a future study. For the current study we decided to present retrieval as simple as possible. This simplicity is an important aspect and assures that the method can be correctly and consistently applied for different sites and time periods:

Given that the NDACC-IRWG raw data is very consistent and consistent retrievals are also implemented for other species it should be the case that a consistent profile retrieval for CO2 is readily achievable. But that this is considered beyond the scope of the paper is reasonable but again it should be stated in the manuscript. This statement is very important given the problems faced by the scaling technique.

> We'll add a statement in the manuscript.

The retrieval strategy is kept rather simple, which is an advantage:
- perform a simple scaling retrieval (TCCON also only performs a scaling retrieval).

This retrieval means a clear and simple constraint and can be very easily used at any measurement site. A profile means a more sophisticated constraint, which has to be clearly specified in order to assure consistency between the sites.

The fact that the TCCON procedure uses scaling hides the fact that the scaled profile varies daily and widely through the year.

We think the fact that TCCON uses a daily and widely varying a priori profile is well described in the manuscript. To examine the effect of the varying a priori, we also show NDACC retrievals with the daily varying TCCON a priori in the manuscript, but our aim is to have a simple and easily traceable retrieval setup.

- we use a fixed a priori. We want to use the retrieval result as stability criterion for the NDACC/FTIR measurement and by a fixed and well traceable a priori (WACCM) we can attribute the variations in the retrieved XCO2 to the NDACC/FTIR measurement.

This is precisely why a constrained profile would be a better choice even if it is more complex to implement.

Again we would like to point out that a simple and easily traceable retrieval setup is very important for our purpose. In the manuscript we demonstrate that our simple retrieval is sufficient for achieving our objectives. For this reason we don't see an argument for using a more complex retrieval strategy (like for instance a profile retrieval), especially since more complexity might also mean a higher risk of inconsistency. A profile retrieval is outside the scope of this manuscript and a profile retrieval could easily be instigated and investigated in the future.

Minor Technical Issues
2. L192: NCEP data for pressure and temperature for NDACC are supplied at the NDACC DHF at www.ndacc.org. p-T data for TCCON are supplied elsewhere. The t-P data you use is one or the other or a third?
This must be stated clearly.

The NCEP p-T data used for the NDACC retrieval are available e.g. via the Goddard automailer (Lait, 2005). Lait, L.R.: Using the Goddard Automailer, available at: http://code916.gsfc.nasa.gov/Dataservices; 2014:
This information will be added.

Thank you this should be added. But have you verified that the automailer NCEP data is identical to the p-T data supplied at the NDACC DHF? It may not necessarily be so.

The source of the p-T data supplied at the NDACC DHF is also NCEP. We only found these data for stations formerly called "primary site", i.e. only for a small subset of the NDACC-IR stations, we therefore would recommend using the above mentioned source (e.g. Goddard automailer).