Quantifying the UK's Carbon Dioxide Flux: An atmospheric inverse modelling approach using a regional measurement network (White et al., 2018)

General comments

The authors present a comprehensive quantitative study of the UK's CO_2 flux in 2013 and 2014, using an inverse approach that separately scales GPP and respiration. Their data assimilation system makes use of regional ground-based observations, two different models for the biospheric prior, and a LPDM for the atmospheric transport (which they modified to include the option of using time-disaggregated footprints). In addition to the net fluxes, the impact of the biospheric prior on the obtained posterior fluxes is studied in detail.

Overall the work is well written and of high quality. The manuscript provides sufficient details about the different components of the assimilation framework and contains an extensive discussion of the prior and posterior fluxes. To strengthen the presented conclusions, however, it would be good to pay some more attention to the technical details of the inversion results.

Specific comments

- On p.6, the authors assess the impact of the timespan for which footprints are disaggregated, and they conclude that the effect of going from 24h to 72h of time-disaggregation is negligible. However, I do not find the provided evidence for this statement very convincing. Indeed, the mole fractions calculated at Ridge Hill coincide nicely for 24h, 48h and 72h, but is this result representative for the entire UK? And how do the results compare to a simulation without time-disaggregation? The expected range of net annual biospheric fluxes has shifted significantly between the simulations with 24h of aggregation and with 72h of disaggregation, compared to the prior. It would be helpful if the results for no and for 48h of time-disaggregated footprints are included in table S1 as well. In addition, the uncertainty of the obtained net flux seems to be unaffected by the timespan of the disaggregated footprint. Can you comment on this?
- On p.12 the basis functions for the inversion are shortly discussed. This part could be made more clear by including a figure showing the clustering of the scaling factors.
- The inversion framework allows to scale the prior estimates for respiration and GPP separately. In a synthetic test it is shown that this approach indeed allows to compensate for biases in either respiration or GPP, which is obviously not possible with the NEE inversion. To strengthen the point of the superior behavior of the TER+GPP inversion over the NEE inversion, it would be nice to instead include a synthetic test with a truth that less obviously favors one approach over the other, e.g. by using a combination of the JULES and CARDAMOM fluxes as truth.
- A different anthropogenic flux map (EDGAR) is used outside the UK. Are these fluxes in agreement with the NAEI fluxes for the UK?
- The data statistics in table S2 cover the year 2014 only. Are the statistics for 2013 similar, and why are they not included? It would be nice to see these statistics for the entire range of presented inversion results, i.e. both 2013 and 2014. Moreover, for the inversion with JULES prior the bias increases for 2 out of 6 measurement

sites. Can you comment on this?
In addition, I find it peculiar that the prior RMSE and prior bias differ between the GPP-TER and the NEE inversion. How did you evaluate these statistics?
I would also suggest to include the statistics for the NEE-inversion in table S2 (maybe put them in a smaller font and/or grey) and to move the table to the main paper.

- The optimized state vector and hyper-parameters are not discussed, while it would be interesting to have at least a look at the number of time periods resolved in the inversion, how scaling factors for GPP and TER vary, and to what extent the boundary conditions required scaling during the inversion. Please include such a discussion.
- Due to the use of satellite remote-sensing data in CARDAMOM, one would expect a benefit of using this prior over the JULES model, at least with respect to spatial distribution of the fluxes. This point does not seem to be addressed.
- Figure 6 can be improved by adding two additional subfigures that show the difference between CARDAMOM and JULES posterior fluxes. Also the addition of maps showing the related uncertainty in posterior NEE would be helpful.

Technical corrections

- P.2, the abbreviation GHG is used before it is introduced.
- P.5, please include an indication of the uncertainty related to the CO_2 measurements
- P.10, please mention the spatiotemporal resolution of the prior ocean flux estimates.
- P.11, to reduce initial confusion, it might be useful to move the paragraph starting at line 350 (about the variable time instead of space dimension) to before equation 4.
- Please be consistent in the colors you use: e.g. the colors for CARDAMOM and JULES are reversed between figures 2 and 5.
- The section numbering in chapter 4 should be updated, now 4.2 comes after 4.3