Comment on essd-2022-213
Anonymous Referee #1


General comments

Since this study is aimed at “informing countries’ carbon budgets”, as stated in the abstract, I am missing a bit a section on how the carbon stock changes defined in this study compare to what is required by the UNFCCC for national reporting. Moreover, this study calculates the Net Biosphere Exchange (NBE) from the inversions and then adjusts this using the terms for crop and wood trade as well as river export to derive changes in terrestrial carbon stock. The UNFCCC guidelines on reporting also correct carbon stocks for carbon losses due to harvest. For wood, the guidelines include consideration of the turn-over rate of wood products, since many wood products have lifetimes of decades to centuries. For crop, the turn-over time is generally much shorter, order of annual, thus I would think that this term should have little affect on carbon stock changes.

Specific comments

L135: Add a reference for the N-fertilization effect

L209: Concerning Eq. 2, why is the carbon in crop harvest considered in the change of carbon stocks, since crop harvest is only a relatively short-term stock of carbon (turn-over on annual time scales) since this will most be consumed or used for fuel?

Section 2: How does the carbon stock definition compare to what is required by UNFCCC (see also the general comment)? Also, are the turn-over times of wood products considered in calculating changes in carbon stock?
Eq. 4: May be there is some confusion on my side, but to correct the IQR to standard deviation (SD), I would think one would need to divide by \(~1.47\), since the IQR includes 50% of the values while one SD includes 34%. Or where does the value of 1.35 come from?

L358: What is the reasoning for assuming 30% uncertainty on the crop and wood exchange fluxes?

L415: Please change “data-model” to “observation-model” if that is what is meant since “data” can be either modelled or observed (also L419). Also, this sentence is a bit ambiguous, do the authors mean that the difference for the OG experiments is more negative in the evaluation data sets compared to experiments?