Not suited for ESSD.

Well written. Good data access at Zenodo. Authors apply multiple ML techniques. But effort as described fails to meet internal metrics and journal expectations. Recommend rejection.

Authors propose to share “the first remote sensing-based global high-precision long-term XCO2 dataset”. Not true by a long shot. A large contingent of researchers, inside and out of NASA, seek to explore and confirm OCO-2 measurements of XCO2. This manuscript unfortunately reports very little of that other work (only Conner et al. 2016?) evaluating, among other complexities, aerosols, clouds, sensor sensitivity and degradation, orbit degradation, etc. A full estimation of accuracy (low on some scales) and precision (highly variable) for XCO2 from OCO-2 exists, not cited here. Another manuscript in ESSD (2023-449) attempts a similar evaluation; also not cited here. Latest version of TCCON also in ESSD, likewise not cited here.

Authors promise new distinction, terrestrial vs ocean. They relate downloading MODIS EVI, CHL-a, and reflectance data (lines 128 to 131). But reader finds only weak reports around lines 275 to 285, with nothing firm nor quantitative. Certainly no uncertainties, signal to noise, etc. Authors fail to meet their own expectations?

Validation seems uncertain and, unfortunately, inconclusive. Their product as good as, or closely correlated to, CarbonTracker (CT, as but one example)? But a user community knows and trusts e.g. CT (in part because of extensive reports on CT uncertainties - see below), so with extra effort (and computing) this product adds what to CT? Nothing, evidently. Authors here correlate (validate?) with both CT and other model syntheses, and with TCCON, but CT and those other simulations also validate against TCCON. So how can these authors validate against a product that also serves as reference to other products they also validate against? NOAA does not, so far as this reader knows, report any XCO2, only surface and ‘above marine boundary layer’ measurements. With extreme attention to precision and representativeness, both completely missing here.

What, if anything, did we gain here? Authors have not made a case for real improvements. They like to claim “spatial resolution of 0.05 degrees and a temporal resolution of 8 days, from 2000 to 2020”? Beg pardon but don’t we already know that? 411 ppm for 2020? This reader can learn that easily from daily co2. Trend of 2 to 2.2 ppm per year over those decades? Again, this reader can learn that from NOAA’s flask network. Plus we know that trend of atmos CO2 has changed (increased) over recent decades. But not here? If authors have made real advancement, they need to prove it. Too often they err by citing ‘surface’ CO2 data when in fact they report XCO2 data.

Biggest failure: no uncertainties. Authors claim “high-precision” but in fact ignore precision/uncertainty entirely. ‘Uncertainty’ as a term never appears in their text. Never an error bar. Reader encounters RMSE but those assume linear correlations and apply, on their own scales, only within specific figures. OCO-2 data products come with well-described much-discussed uncertainty matrix (unfortunately ignored here). On top of that, with abundant un-cited evidence, authors insert additional uncertainty with every step and modification: cross-fold validation (introduces uncertainties); spatial expandability (introduces substantial additional uncertainties); temporal extensions (further substantial uncertainties). Who might know these uncertainties better than these authors? Reader finds and learns nothing. Personally, this reader doubts, after cumulating collective uncertainties across multiple steps, that they can claim better than $\pm$ 50% in annual XCO2 increments or $\pm$100% in 2000 to 2020 trends. Have they done better than my estimates? RMSE’s do not answer these questions; author provide no evidence one way or another.

Not clear from evidence presented here that this product adds any value.