June 5, 2020

Responses to Reviewer #1

Thank you for your review of our paper: ESSD-2019-252. Below we paraphrase your comments in bold and provide our responses in regular text. We also recognize the challenging context of the COVID-19 pandemic as well, and are most appreciative of your review.

Needs substantial revision before acceptable to ESSD.
Data access comments, related to difficulty downloading the data from the Dryad repository, including suggestions to provide a "teaser data product" and a more prompt delivery of data.
We became aware of these issues when a few scientists contacted us about the data repository. We found that Dryad does in fact respond promptly, but because of the large file size sends an email with a URL for users to download the data. We have discovered that this email, unfortunately, is frequently filtered into the Spam folder. We added a note in the data description portion of the Dryad repository material to alert users to this situation. We note that the Dryad repository meets the requirements specified by ESSD. Part of the challenge is simply due to the fine-resolution global datasets and floating-point values that do not compress well. Also, while we recognize it isn't a permanent solution, we provided a URL:
(https://davidtheobald8.users.earthengine.app/view/global-human-modification-change) for a dynamic mapping website that allows rapid visualization of our data, and comparison to a few other commonly referenced datasets.

I am worried about source data availability, particularly in relation to the use of Google Earth Engine.
You are correct that we did implement our analysis in Google Earth Engine (GEE) and did upload source data into the GEE platform to conduct our analysis. But, all source data used are open source and are accessible externally via the permanent DOIs that we provided. Furthermore, the formulas are carefully and clearly described, following the guidelines of ESSD in providing DOI permanent links to all source data and "recipe" used in the analysis to create the data product presented in the paper.
[Note to Editor: please provide guidance on this issue, as well as advising on what arrangement within the paper is preferred -- during the creation of the Discussion Paper we originally had a list of all datasets in the Data Availability section, but moved the citations with DOI into the References section in response to a request to streamline and provide acronyms in Table 1. We are open to either structure in the paper.]

Manuscript fails to present comprehensive estimates of uncertainty.
Thank you for your detailed comments on this important issue. In our response we address four aspects from the issue you raised regarding "uncertainty": (1) our ability to capture dynamic events, such as wildfire or climate change; (2) understanding the uncertainty of our results related to

measurement error; (3) the precision with which we report results; and (4) including uncertainty in our validation analysis. We address each of these in order:

1. Uncertainties associated with dynamic events, land uses, and activities that we did not attempt to capture. We address this briefly in the caveat section and by citing our previous work where we discuss these challenges, particularly around wildlife and climate change. We revised our text in the caveat section to read, on lines 560-570:

   > *"As with any model, we recognize there are limitations of our work. We did not include data for all human stressors, largely because of incomplete global coverage or coarse mapping units (Klein Goldewijk et al. 2007; Geldmann et al., 2014), an inability to discern human-induced versus natural disturbances, or uncertainty in the location and directionality of its impact (e.g.; climate change on terrestrial systems; Geldmann et al., 2014). In particular and discussed in Kennedy et al. (2019a, 2019b), changes to land cover due to ecological disturbance events, such as wildfires or flooding, are not included in our analysis because of the difficulty in separating natural from human-caused disturbances -- yet, we recognize that because of the broad extent of wildfire, in particular, would have strong implications. We did not include climate data as a stressor in this product to keep our analysis manageable and tractable. For more integrated analyses, our data product should be used in combination with datasets of impacts due to climate change (e.g., Parks et al. 2020)."*

2. We revised our manuscript to improve how we address how uncertainty affects our results for 2017 by conducting an additional analysis of the per-pixel variability (standard deviation) and adding summary results of it along with a map (in Figure 4), providing values across the randomized iterations of the mode.

   Additionally, we realize that a few of the uncertainty measures we incorporated were dispersed in the methodology section when describing the modeling approach. We therefore added text in the revised Uncertainty and validation analysis section (quoted below in italics) to reiterate key aspects of the methodology that directly include uncertainty in the formulas used to calculate the human modification for each stressor. In particular we: (a) used the results from the accuracy assessment of the land cover dataset, to adjust weights associated with land cover types when estimating the degree of human modification (H); (b) similarly, we weighted our estimates for the urban/built-up stressor when calculating H, as a function of the degree of confidence of the modeled estimates provided by the GHSL dataset, on a per-pixel basis; and (c) addressed the spatial uncertainty associated with stressors represented as points (e.g., mine locations, gas flares) and lines (i.e. roads) when calculating H.

   The above mentioned revisions are included in Figure 4 and the revised text on lines 502-512:

   > *"To address uncertainty in our results, we specifically included an estimate of uncertainty associated with each stressor in the calculations of human modification for 2017 conditions (Equations 4-26). For example, we adjusted $p(C_{cp})$ by directly incorporating measured confusion among land cover types using the results from the accuracy assessment of the land cover dataset (from Eq. 4).*

We summarized the uncertainty maps by calculating the global mean for each of the 50 randomizations, and found the mean of the 50 global mean values was 0.1434 (SD= ±0.0076) and ranged from 0.1243 to 0.1612, thus, in line with the global mean of 0.1461 obtained using our "best-estimate" intensity values. We also mapped the per-pixel variance (standard deviation) to examine the spatial pattern of uncertainty (Figure 4). The locations of the highest levels of uncertainty tend to be in more highly developed landscapes."

3. We responded to your comment about our reporting of results without providing variance measures and overly-high precision by modifying our text to include a measure of human modification (which ranges from 0.0 to 1.0) using 4 orders of precision (i.e., +/- 0.0001 rather than +/- 0.00001) to be consistent with reporting percentages. As suggested, we also added a +/- when reporting our estimates of human modification in terms of area (i.e. square kilometers). This includes removing our statement that you found troubling regarding the 100 km^2 unit of analysis area. Changes occur on lines 28, 422, 435, 450-462, 508-9, and Tables 4-6.

4. We addressed your comment about validation: "Neither do the authors assign any uncertainty to so-called validation products HF or THPI", by clarifying the purpose of our validation analysis and how we accomplished it. We revised our manuscript text to clarify our steps, and included further citations that provide additional details. To be clear, we did not validate our results against the modeled outputs from the human footprint (HF) or human pressure index (THPI), rather, we simply *compared* our data to them because they are typically perceived as being similar, are readily-available and frequently used datasets, and we anticipated such a comparison will be a common and reasonable question of readers and data users. In fact, a central reason we have produced the work in this manuscript is to build on and provide a more refined and improved way to spatially represent and measure the degree of human modification on landscapes. That said, while we compared our data to other available products, we note that we did indeed validate our data by calculating and reporting the coefficient of determination (i.e. r^2) against "ground truth" data described in Kennedy et al. (2019a, 2019b) i.e., "our validation data" mentioned in lines 514-520:

   "We found strong agreement between H for ~2017 and our validation data (r=0.783), with an average root-mean-square-error of 0.22 and a mean-absolute-error of 0.04, for the 926 ~1 km^2 plots (9,260 sub-plots). There were 726 plots within ±20% agreement, while for 161 plots H was estimated higher than our visual estimate from the validation data (and 39 plots lower). Estimates of H were biased high, likely because the stressors for the "human intrusion" and electrical infrastructure (based on nighttime lights) are not readily observable from the aerial imagery used to generate the validation data. Our results here are consistent with our earlier findings (Kennedy et al. 2019a, 2019b, 2019c)."

Specific comments and suggestions below:

Page 2, lines 23-24, question about the duration of a breath:
Thank you -- following your questions and Reviewer #2's suggestion, we removed this "real-world" comparison from the manuscript.

Page 3, Line 49: Have [we] dismissed too many prior studies or contemporary work on human impact issues?
Thank you -- at your suggestion, we cited a few additional important works in the field, in particular the work on HYDE 3.2 product (https://doi.org/10.5194/essd-9-927-2017), Ellis' work on mapping the Anthromes, and recent work (Riggio et al. 2020) that compares human modification (Kennedy et al. 2019a), Human Footprint, Anthromes, as well as Jacobsen's (Jacobsen et al. 2019) and Riggio's work (Riggio et al. 2020). It is always a challenge to balance providing enough context for when developing new science. We chose to provide a more focused, technical description as the purpose of our paper is to develop a specific data resource that examines recent change (1990-2015) and relatively high-resolution for global work (0.3 km) -- rather than a broader review of similar previous efforts.

Page 3, line 60: Clarify wording: "...obstructions by vegetation canopy (e.g., some roads, trails)".
Good suggestion, we modified lines 59-62 to read:

> *"This is because remotely sensed imagery has limitations for this application -- especially prior to ~2010 -- because it can require human-interpretation to classify adequately and can miss development features that are obstructed by vegetation canopy or are small or narrow features (e.g., towers, wind turbines, powerlines)."*

Page 3, Line 69: clarify explanation of "...additive but monotonic relationships…"
Thanks, we simplified this sentence to clarify it on line 70, to read: "...measure that assumes additive relationships among stressors..."

Page 8, line 264: be consistent when listing metals with common names.
Thanks, we modified our text on line 266 to state "uranium oxide".

Page 10, line 337: Clarify why wildfires, if excluded in the analysis, do not show up as an uncertainty.
Good point, we added text to describe how wildfire (and other dynamic ecological processes) are considered within our work on lines 339-342:

> *"(Note that we excluded wildfire as a stressor because of the challenges of attributing wildfires to human causation-- especially over global extent, and urbanization because it is measured directly by the built-up stressor)."*

and lines 564-570:

> *"In particular and discussed in Kennedy et al. (2019a, 2019b), changes to land cover due to ecological disturbance events, such as wildfires or flooding, are not included in our analysis because of the difficulty in separating natural from human-caused disturbances -- yet, we recognize that the broad extent of wildfire in particular, could have strong implications. We did not include climate data as a stressor in this product to keep our analysis manageable and*

tractable. For more integrated analyses, our data product should be used in combination with datasets of impacts due to climate change (e.g., Parks et al. 2020).”

Page 12, line 418: clarify and be consistent with area estimate, and remove reference to football pitches.
Thanks, done.

Page 12, lines 423-4. Clarify why available climate change datasets are not used.
Thanks, this is an important point. We agree that climate change effects are happening, and there are numerous climate data products and a burgeoning field of science. To address this point, we clarified our decision not to include it in our analysis on lines 116-123:

> *“We note that we did not map stressors for invasive species or pathogens and genes, geologic events, or climate change. This was because suitable temporal global data were not available to capture stressors due to invasive species or pathogens and genes; the majority of geological events are not directly caused by humans; and climate change is better modeled as separate process distinct from the effects of direct human activities and has a plethora of research on this topic (Geldmann et al. 2014; Titeux et al. 2016).”*

and on lines 568-570:

> *“We did not include climate data as a stressor in this product to keep our analysis manageable and tractable. For more integrated analyses, our data product should be used in combination with datasets of impacts due to climate change (e.g., Parks et al. 2020).”*

Sincerely,

David M. Theobald, Ph.D., on behalf of co-authors