

Earth Syst. Sci. Data Discuss., author comment AC3  
<https://doi.org/10.5194/essd-2021-165-AC3>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Reply on RC2

Kytt MacManus et al.

---

Author comment on "Estimating population and urban areas at risk of coastal hazards, 1990–2015: how data choices matter" by Kytt MacManus et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-165-AC3>, 2021

---

We thank the referee for their helpful remarks, and have addressed them below inline. Our responses begin with "A:", we have marked each referee question with "Q:" in order to enhance readability.

The manuscript describes a thoroughly methodology to define LECZs at the global level, and assess the exposure of the rural and urban population living within them. Exposure is assessed by using and combining multiple different gridded population, elevation, and urban/rural datasets (differing in terms of their native spatial resolution, the underlying assumptions made, and the methods used to produce them). The analysis includes a rigorous sensitivity analysis and a key discussion about the uncertainty/variation associated to the exposure estimates. Results show the importance of accounting for such uncertainty/variation to better identify actual patterns and trends.

I am very supportive of the Author's effort and would like to highlight that more studies like this one, should be conducted to better inform and support policy and decision makers.

The manuscript is timely, appropriate for the journal, and potentially of interest for its readers. It is well written, articulated and presented, and offer an original contribution in the field of coastal hazard, as well as valuable insights and considerations into the advantages and challenges of using and combining multiple datasets to globally assess population and building exposure to sea level rise, and the associated hazards, in LECZs.

In my opinion, the manuscript should be published after minor revisions aimed at addressing the detailed comments provided below.

I have really enjoyed reading the manuscript and want to congratulate the Authors for their work.

Detailed comments:

Q: Lines 132-134: "To our knowledge, no targeted or multi-criteria evaluation of these global elevations data sets in the urban setting has been made (but see related analysis of the built environment in cities (Esch et al., 2020))." - What do the Authors mean by

saying "but see related analysis of the built environment in cities (Esch et al., 2020)"?

A: Added clarifying text that Esch et al. 2020 used TanDEM-X to characterize the built environment.

Q: "2.1.4 CoastalDEM90" - I would suggest to rename this subtitle as "CoastalDEM90 and ALOS World 3D"

A: We have adopted this suggestion, thank you!

Q: Lines 276-280: "Importantly, the covariate data used to delineate WorldPop estimates are static for the year that they were collected (even though some represent time-varying characteristics, like the night-time lights), and therefore the spatial distributions of population estimates are also static." - This statement is inaccurate and should be rephrased - most of the covariates used to delineate WorldPop estimates are actually either temporally-explicit (including the use of interpolated/extrapolated build-up areas based on GHSL and GUF data; please refer to Lloyd et al. 2019), representative of the whole modeling period (climate data), or assumed to be time-invariant, with respect to the modeling framework and objective (ie, elevation, slope, rivers, coastline, and landarea), with the only exception represented by OSM roads and derivatives.

Q: Lines 302-304: "Therefore, because GHS-POP is the only data representing a true time series in regards to the underlying spatial structure, and was acceptable in other regards as mentioned above, it was chosen as our core population data set." - While I agree with the Authors about the selection of GHS-POP as the core population dataset for this study (eg, its temporal coverage, the fact that it is produced without using elevation data as a modeling covariate, etc), similarly to my previous comment, stating that "GHS-POP is the only data representing a true time series" is highly inaccurate and should be rectified.

A: Thank you for your input here. We have addressed these two comments together, first by correcting the text to account for the fact that WorldPop does use time varying covariate data. As you mentioned, WorldPop (as well as LandScan) does use elevation covariate layers, and therefore may not be the best choice for estimating populations in LECZs, and the 1990 time point included in GHS-Pop was also a driving factor in our decision to make it a core data choice in this work.

Q: Line 324: "three large classes" - Which are these three classes? The ones listed in Table 4? If yes, I would suggest to refer to Table 4 at the end of this paragraph.

A: Yes these are in table 4...we have revised the text to this effect.

Q: Lines 397-398: "Of the four data sets included, only GHS-SMOD and GRUMP claim by design to represent urban extents." - Please add relevant references supporting such statement.

A: We have added references here.

Q: Line 603: "by a factor of 100" - Please explain.

A: We added the following parenthetical clarification here as follows: "... disaggregated by a factor of 100 (e.g., 1 pixel was divided into 100 pixels, given its 1km resolution)"

Q: Lines 605-606: "there was no need to upsample. These data were simply resampled at 9 arc second resolution" - Resampling from 30 to 9 arcsec is upsampling, right? Please elaborate further.

A: We have clarified the text to replace the term "upsample". What was meant was that the Mean Administrative Unit Area layer did not require further disaggregation during the resampling to a higher resolution pixel size, since it is a non-continuous surface.

Q: Line 614: The link is not working.

A: This works now.

Q: Line 629: "where possible." – Remove the full stop

A: Thank you, edited.

Q: Line 638 & 640 and in general through the whole manuscript: "proportion" – Maybe "percentage"

A: Full text reviewed and edited to make this replacement where appropriate. Thank you.

Q: Lines 655-657: "The relationship is clear to discern in the 5-10m LECZ, where GPW consistently estimates the lowest percentage, WorldPop the second lowest, LandScan the third lowest, and GHS-POP the highest percentage regardless of the elevation source used to define the LECZ." - In Figure 9, such relationship seems to be the other way around when considering <5m LECZs defined by MERIT, SRTM and TanDEM? Maybe a plotting error?

A: It is true that the relationship reverses in 0-5m based on MERIT, SRTM, and TanDEM-X. This can be explained by a number of factors: 1) CoastalDEM has a very large 0-5m zone, and 2) GPW allocates people indiscriminately based on land area, which might include wetlands, beaches, etc...Another point is that that the range of estimates is more condensed in the 0-5m zone.

Q: Figure 10: What does the pink area represent? Is the 0-5 blue area a non-contiguous 0-5 LECZ? I would suggest to use a different color, rather than gray, for the 10+ or not contiguous zones

A: Revised figure color and presentation. The area represented the LECZ where the blue shades represented raw elevation data.

Q: Line 699: "banks of rivers" - Does this mean that all these banks are higher than 5m?

A: No, we have revised the text to better clarify the point that CoastalDEM extends further inland than other DEMs when evaluating LECZs, due to the fact that other DEMs sometimes capture elevations of inland water >5m, whereas CoastalDEM sets those values to 0. This more inclusive criteria extends further inland and leads to capturing more river tributaries. Thus, we have generalized this point.

Q: Lines 719-720: "at least when comparing with urban and quasi-urban data not based on city lights." – Unclear, please explain/elaborate further

A: Edited to "at least this is observed when comparing..."

Q: Lines 752-753: "suggesting that some of the official definitions are drawn from areas that have a more quasi-urban character (such as towns, suburbs, etc)." - Unclear - please elaborate further

A: Added clarifying text.

Q: Caption of Fig 12B: "and Urban Proxy data sets" – I guess it should be "and GHS-SMOD"

A: We have made this correction.

Q: Line 846: "quasi-urban and rural" – Should be deleted?

A: Yes, thank you. We have deleted it.

Q: Lines 853-854: "as to be expected given its particularly low urban and quasi urban populations outside of the zone." - Unclear - please elaborate further.

A: Replaced with "However, GPW has the highest shares of the urban and quasi-urban populations in the  $\leq 10\text{m}$  LECZ, which is notable given its particularly low urban and quasi-urban populations outside of the zone, and provides additional evidence that the zone is disproportionately urban.

Q: Line 915: "53.86%" – Should be "46.16"?

A: Thank you, corrected.

Q: Lines 944-945: "there are no equivalent relationships with elevation levels and the LECZ zones." - Unclear, please explain/elaborate further

A: We have clarified the text.

Q: Lines 963-964: "All of the population densities in the 5-10m LECZ are higher than those outside of the LECZ regardless of which population source is used." - This does not seem the case when considering GHS-POP though, right?

A: Thank you, we have corrected and clarified.

Q: Line 1003: "we've" – "we have"

A: Corrected.

Q: Line 1032: ", and 2019)." – Should be deleted?

A: Corrected.

Q: Line 1074: "some of the assumptions made in the process are obscured to end users or non-experts." - Unclear - please elaborate further.

A: Thank you. We have clarified that although many of the data we evaluated were peer reviewed, the complexity of the models or given that peer-reviewed papers are often written for technical and discipline-specific audiences, details may be less accessible for interdisciplinary or non-experts users.

Q: Line 1256: "LECZ based on MERIT" – Should be the same for LECZ based on SRTM, right?

A: This is at line 1156. Yes you are correct. We have inserted the following: "(SRTM can also be distributed but is not here due to its known limitations.)"

Q: Line 1202: The link is not working.

A: The link seems to be working.

Q: Line 1205: "MERIS" – What about the same layers produced using SRTM?

A: SRTM can also be distributed but is not here due to its known limitations.

Q: Line 2012: The link is not working.

A: The link seems to be working. Note, line 1212.

Q: Line 1287: "'Globally" – should be "globally"

A: Corrected.

Q: Lines 1301-1302: "Given that both population and built-density differ by urban proxy data sets, even within the LECZ, we caution users to consider carefully what a given measure means to their analysis." - Unclear - please elaborate further.

A: We've deleted this sentence since it largely restates, with respect to built-up densities, what was articulated at the end of the prior paragraph. The rest of this paragraph remains as important value added, but no need for this last sentence.

Q: Lines 1362-1364: "One key explanation for the variation across population data sources is driven by the input resolution of the administrative units of the underlying census data that are made available (by national statistical offices)." - I do not fully agree with this statement, given that at least three population products (GPW, GHS-POP and WorldPop) are produced using the same input data - unless the Authors are referring to the fact that the disaggregation of larger, "less constrained", input admin units may produce very different results when disaggregated using different methods. Please explain/ elaborate further.

A: Thank you. We were not clear in our write-up. We were referring to the combination of input resolution and modelling and have revised the text as follows:

"This variation across population data sources is explained by a combination of the input resolution of the administrative units of the underlying census data that are made available (by national statistical offices) and modelling choices used to reallocate population within administrative units: Three of the four data sets share the same underlying population data units, but countries vary considerably in their administrative level of the data that is publicly available. When the data are available at a high resolution, variation in the modelling choices (both types and number of ancillary data) makes little difference. But where data are coarse or even of moderate resolution, the modelling choices matter."

Q: Line 1445: "Worldpop" – Should be "WorldPop"

A: Corrected.

Q: Line 1463-1464: "which at the subnational level may be less transparent inherently."

A: This sentence has been removed from the text.