



Comment on **essd-2021-10**

Anonymous Referee #2

Referee comment on "Spatially explicit global gross cell product (GCP) data set consistent with the Shared Socioeconomic Pathways" by Tingting Wang and Fubao Sun, Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-10-RC2>, 2021

Review report for manuscript „Spatially explicit global gross domestic product (GDP) data set consistent with the Shared Socioeconomic Pathways“ by Wang and Sun

The present manuscript provides and describes a global dataset of national GDP data according to the SSPs downscaled to a resolution of 1km. The methodology to conduct this downscaling is very similar to already available approaches, without providing a comprehensive description and a rigorous assessment of issues, loopholes and uncertainties. I started out reading enthusiastically but more and more learned that this dataset and the very general description thereof does not provide a timely and novel contribution to the field and for ESSD. I do agree that provision of thoroughly downscaled “lit-pop” SSP GDP data can be relevant for people working in the field but I do not think that the present dataset provides more answers than questions. For example, the authors question the approach by Murakami and Yamagata (2019), which very thoroughly describe their downscaling procedure and limitations, but then fail to provide any methodological advancement. I started collecting major and minor comments while going through the document until I lost motivation towards the end, see below. In summary, novelty is missing, the methodology is faulty and unconvincing, the description and discussion incomplete, uncertainty assessment is missing, English language needs thorough corrections, legibility of figures needs to be improved.

Major Comments

- Title: The authors name their data spatially explicit gross domestic product (GDP), which is very colloquial and misleading as GDP usually refers to the output per country and not per cell. Consider using gross cell product instead.
- Introduction: The introduction is not very well structured and needs revision: E.g. lines 84-88 appear very unconnected to the preceding paragraphs, as are lines 98-102.

- Line 74-75: The authors should cite and refer to this recent paper, which is a global extension of Zhao et al 2017 and covers large parts of the methodological work presented in the current manuscript in much more detail: Eberenz, S., Stocker, D., Rösli, T., and Bresch, D. N.: Asset exposure data for global physical risk assessment, *Earth Syst. Sci. Data*, 12, 817–833, <https://doi.org/10.5194/essd-12-817-2020>, 2020.
- Line 94-95, line 317-334: Murakami and Yamagata downscaled SSP1-3 only for a specific reason: They found their very elaborate methodology to be too uncertain to be applied to SSP4-5. In the current manuscript I am missing a thorough discussion of related uncertainties.
- Line 118-120: This statement about missing open access availability is not true. There are various freely available sources for gridded historical and/or future GDP data, e.g. Eberenz et al 2020, Geiger et al (2017), Kummu et al (2018).
- Section 2 (starting from line 131) is a presentation of various data sources without any introduction whatsoever. It would clearly be helpful for the reader to provide an initial paragraph motivating the description and the use of the different data sources in the context of the study.
- Section 2.2.1 (starting from line 205) is very confusing and it remains totally unclear why the concept of the ICP or PPP rates are introduced. This, however, is a very relevant point fundamental for the whole study! SSP data is provided in units of 2005 PPP USD (to allow for global intercomparison of national numbers) but most other data is not (or not anymore) (E.g., Worldbank GDP data in 2005 PPP USD was withdrawn and I don't expect sub-national GDP data to be available in PPP-adjusted values). Please see Geiger (2018) and references therein for a thorough description of this issue. I expect the authors to describe in clear language including references to data sources how they tackled the problem of data conversion step by step. It's not a simple issue and faulty data might be the result. This comment also applies to Section 3.1.
- Section 2.3.2 (starting from line 301): SEDAC for a couple of years already provides downscaled SSP population projections at 1km scale (<https://sedac.ciesin.columbia.edu/data/set/popdynamics-1-km-downscaled-pop-base-year-projection-ssp-2000-2100-rev01>). These data are not mentioned neither applied in the current manuscript and could significantly improve the results provided.
- Line 317-334: A revised version of the Murakami data was provided at 1/8° resolution (using the Jones and O'Neill 2016 population data as input) here: Geiger, Tobias; Daisuke, Murakami; Frieler, Katja; Yamagata, Yoshiki (2017): Spatially-explicit Gross Cell Product (GCP) time series: past observations (1850-2000) harmonized with future projections according to the Shared Socioeconomic Pathways (2010-2100). GFZ Data Services. <https://doi.org/10.5880/pik.2017.007>. This reference should be cited in the manuscript. RELATE TO THIS POINT LATER on doubts of data quality.
- Section 2.3.4 (starting from line 335): It's worthwhile to acknowledge how policy changes can affect the projections and I agree that China's policy change regarding the two-children policy will have significant impacts. But I doubt that deliberately replacing the GDP estimates for one country only provide an improvement for a globally-consistent and dynamically-modelled dataset. On the contrary, China's economy is globally interconnected and simply adjusting China's output by as much as 40% will misrepresent the interaction with all other countries. There are two solutions: 1) stick with the original data provided or 2) wait for a new update of globally-consistent GDP projections that do not only include policy changes in China but similar changes across the globe including past economic crises and the Covid-19 pandemic.
- Section 3.1.: I see many different data sources being used and reshuffled to obtain GDP per capita estimates. Relating to my major comment 8), I doubt that all GDP sources can be easily transferred between different units (in particular sub-national estimates used here). I further miss a thorough consistency check of GDP and population data from different sources, including strange and un-referenced sources like the Wire & Plastic Products Group. Differences between different sources can be quite significant (e.g. because of different and/or changing assumptions on country shapes, ...) and erroneous estimates of GDP per capita are the result. These issues are

discussed for gridded population (Section 3.2.1) but totally neglected for national estimates here.

- Line 412-414: I find it quite strange to use a county-level analysis for China only to judge on a global population dataset quality. When using Fig S1 one would reject LandScan based on its largest RMSE contribution.
- Eq 3: This approach is very general and does not refer to nor tests recent advancements in the field, see Eberenz et al (2020).
- Figure 2: I would be curious to see how the methodology applies to the U.S. or China alone instead of mixing up those countries. I would also strongly suggest to estimate the quality of the downscaling using log-transformed data. It seems in Fig 2c that the results are dominated by a few counties with largest GDP and not by the bulk of data.
- Section 3.4: The discussion is very general, not convincing and does not provide strong evidence to believe in the methodology presented. Again using only evidence from China and the U.S. for a global study provides very limited evidence, without discussing these limitations at all.
- Line 508: See major comment 9.
- Line 510: This assumption is very strong for a period of about 80 years and could be overcome by either using the available downscaled SSP population data on 1km or by using more advanced refinements, as e.g. presented by Murakami and Yamagata 2019.

Minor comments

- Line 24: Definition ScenarioMIP missing
- Line 28-31: This sentence is unclear and does not provide insights. What is not available open access?
- Line 66: Reference Tobias 2018 should be corrected to read Geiger 2018
- Line 78: What are DN values? Please specify!
- Line 95-98: I don't understand what the authors want to imply with this sentence.
- Line 114: What gives rise to the range of projected GDP changes in China? Please specify!
- Line 377: Please provide reference to Wire & Plastic Products Group data sources.
- Line 395: Please specify and reference which GIS-based boundaries were used in this study as this choice may have strong implications for the provided data.
- Eq 3: The subscript lit-pop and NTL-Pop are used simultaneously, please correct.

