COMPASS

Exposure datasets at multiple scales

Deliverable 3.1

31st March 2025

Title	Exposure datasets at multiple scales
Lead Beneficiary	PIK
Lead Author(s)	Dominik Paprotny (PIK)
Contributors	
Deliverable number	3.1
Work Package	WP3: Enabling impact attribution by multi-scale modeling of exposure and vulnerability
Submission date	31/03/2025

Dissemination Level

PU	Public — fully open (automatically posted online)	X
SEN	Sensitive — limited under the conditions of the Grant Agreement	
CI	EU classified —Restreint-UE/EU-Restricted, Confidentiel-UE/EU-Confidential, Secret-UE/EU-Secret under Decision 2015/444	

Version History

Date	Version	Contributors	Comments
19/02/2025	0.1	Dominik Paprotny (PIK)	First version for internal review
12/03/2025	0.2	Doris Vertegaal, Anaïs Couasnon (Deltares)	Internal review
14/03/2025	1.0	Dominik Paprotny (PIK), Anaïs Couasnon (Deltares)	Final review and first finished version

Citation

Paprotny, D. (2025): Exposure datasets at multiple scales. Horizon Europe project COMPASS. Deliverable D3.1.

Funded by the European Union

The COMPASS project has received funding from the European Union's HORIZON Research and Innovation Actions Programme under Grant Agreement No. 101135481

Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or of the European Health and Digital Executive Agency (HADEA). Neither the European Union nor the granting authority HADEA can be held responsible for them.

Executive summary

Deliverable 3.1 "Exposure datasets at multiple scales" is a data deliverable, hence this report only documents the dataset without providing a detailed analysis. The data is a gridded dataset of global exposure (population, gross domestic product and net fixed asset value) at a spatial resolution of 30 arc seconds (≈0.93 km at the equator) and lower. It spans years 1850 to 2100, has an annual temporal resolution and includes five future trajectories consistent with the Shared Socio-economic Pathways (SSPs). Due to the large file sizes, only a selection of spatial resolutions and timesteps is provided in the repository, but the code and input data provided alongside the dataset enable users to generate different resolutions and timesteps as required by their research needs.

The exposure dataset and input data for its generation (currently v1.0 at the time of writing) are available on the COMPASS research repository Zenodo at the following DOI: 10.5281/zenodo.14892500

The code is hosted at the following repository: https://github.com/HORIZON-COMPASS/Exposure-and-vulnerability-modelling.git under the COMPASS Github repository. The version of the code using in the deliverable is in the 'D3_1_final' branch (https://github.com/HORIZON-COMPASS/Exposure-and-vulnerability-modelling/tree/D3_1_final) and under the release D3.1 (https://github.com/HORIZON-COMPASS/Exposure-and-vulnerability-modelling/releases/tag/D3.1).

Note: the use of term 'country' as well as the specific representation of national and subnational boundaries used in the data and documentation does not represent a particular point in time, but is done for practical and statistical purposes.

Table of Contents

1. Overview

Table 1. Main characteristics of the global exposure dataset.

Attribute	Value
Variables	Population Gross domestic product (2017 US dollar in Purchasing Power Parity) Net fixed asset value (2017 US dollar in Purchasing Power Parity)
Temporal extent	1850 - 2100
Spatial extent	Global
Temporal resolution	Annual
Spatial resolution	30 arc seconds, 30 arc minutes (with possibility of generating other resolutions that are multipliers of 30 arc seconds)
Format	GeoTIFF raster
Scenarios	Historical (1850 – 2023) Short-term projection (2024 – 2029) Five trajectories of Shared Socio-economic Pathways version 3.1, 2030 – 2100 (Van Vuuren 2017, IIASA 2024)
Timesteps	Benchmarks: 1850, 1900, 1950, 2100 COMPASS case study years: 2010, 2013, 2014, 2019, 2020, 2022, 2023

2. Data structure

The data is stored on the COMPASS project Zenodo repository (https://doi.org/10.5281/zenodo.14892500) and contains, apart from the exposure datasets, the documentation, input data and code for producing additional datasets at other spatial resolutions and/or specific timespan according to user needs. The input data is stored as multiple file parts of one zip file therefore they have to be unzipped together. **Note:** the input data mentioned in red are not provided, as they are publicly available elsewhere and at the same time have very large file sizes. The code uses the raw data as they are downloaded from the sources indicated and only needs to be inserted into the indicated folders.

- **code/** - contains Python code version release 1.0 for producing the global exposure dataset
    - **exposure/ -** contains scripts for generating the dataset
        - **combine_national_data.py –** script loads and harmonizes national timeseries
        - **disaggregation_exposure.py –** script generates the gridded exposure dataset at defined timesteps and resolutions.
        - **exposure_functions.py –** auxiliary functions for other scripts. The user needs to define the path to **data/** folder (variable MAIN_PATH) before being able to run the scripts.
    - **visualization/**
        - **country_graphs.py –** script for generating graphs of national timeseries of exposure
    - **LICENSE –** description of the code license
    - **README.md –** description of the code, including installation and running instructions
      requirements.txt – list of required Python packages for building the virtual environment
- **data/** - contains input and output datasets
    - **Inputs/** - contains input datasets
        - **Admin/ -** datasets defining national and subnational units
            - **Global_OSM_boundaries_2024_v4 –** vector dataset with national-level units
            - **OSM_country_map.tif –** raster with national-level units
            - **OSM_subnational_map.tif –** raster with subnational-level units
        - **GHSL/ -** selected Global Human Settlement Layer files (population and build-up surface at 30” resolution, 1975-2030), for build-up surface: total residential and non-residential (“Total RES+NRES”)[1]
        - **HYDE/** - selected HYDE 3.2 files
            - **popc_1975AD.tif –** interpolated gridded population in HYDE for 1975
            - **zip/** - HYDE 3.2 population count (variable ‘popc’, decennial 1850 AD-1980 AD) in the ‘baseline’ scenario[2]
        - **National_data/ -** input national and subnational exposure data
            - **National_exposure_all.xlsx –** national-level historical data
            - **Subnational_exposure_all.xlsx –** national-level historical data
            - **SSP_3_1_main_drivers.xlsx –** selection of SSP scenario data (‘SSP basic drivers 3.1 full’)[3]
            - **UN_PPP2024_Output_PopTot.xlsx –** United Nations probabilistic population projections, total population both sexes[4]
            - **WEOOct2024all.xlsx –** IMF World Economic Outlook Database, October 2024, By Countries[5]

[1] https://human-settlement.emergency.copernicus.eu/download.php
[2] https://geo.public.data.uu.nl/vault-hyde/HYDE%203.2%5B1710494848%5D/original/
[3] https://data.ece.iiasa.ac.at/ssp/#/downloads
[4] https://population.un.org/wpp/downloads?folder=Probabilistic%20Projections&group=Population

        - **Wang_SSP/ -** gridded population projections from Wang et al. (2022)[6]
    - **Outputs/** - contains output exposure datasets
        - **Figures/ -** graphs of national timeseries of exposure (1850-2100) providing an overview of historical changes and the different future SSP trajectories. Generated with 'country_graphs.py' script
        - **Fixed_asset_value/ -** raster files containing gridded net fixed asset value in $2017 Purchasing Power Parities (PPPs)
        - **GDP/ -** raster files containing gridded gross domestic product in $2017 PPPs
        - **National_timeseries/** - harmonized national timeseries of data that are used as basic inputs for disaggregation of exposure. Generated with 'combine_national_data.py' script.
        - **Population/** - raster files containing gridded mid-year population estimate
- **documentation/**
    - **Documentation.pdf –** this document, provided for methodological background

The gridded raster exposure datasets in **data/Outputs/** directory are named with the following convention:

VAR_YEAR_RESOLUTION.tif

Where VAR is the variable name (Pop – population, GDP – gross domestic product, FA – fixed asset value), YEAR is the timestep and RESOLUTION is the grid resolution in arc seconds.

For files between years 2021 and 2100, the name is expanded to indicate the selected scenario:

VAR_YEAR_RESOLUTION_SSPX_HARMONIZATION.tif

Where SSPX denotes SSP scenario (SSP1-SSP5) and HARMONIZATION is either 'harmonized' (harmonization procedure is applied, see section 3.3) or 'not_harm' (harmonization is not used). Here, we only provide outputs with harmonization, but the user can use the code to generate data without harmonization.

[5] https://www.imf.org/en/Publications/WEO/weo-database/2024/October/download-entire-database
[6] https://doi.org/10.6084/m9.figshare.19608594.v2

3. Methodology

3.1. General approach

The global exposure dataset combines multiple input datasets to produce homogenized and gridded datasets (Fig. 1). The dataset is generated by an algorithm and coded in Python that combines three layers of input data:

(1) A database of national-level historical and projected timeseries of population, gross domestic product (GDP) and fixed asset value for all countries of the world for 1850-2100;
(2) A database of subnational-level historical GDP per capita (for selected countries, where available);
(3) External gridded datasets of population and build-up area.

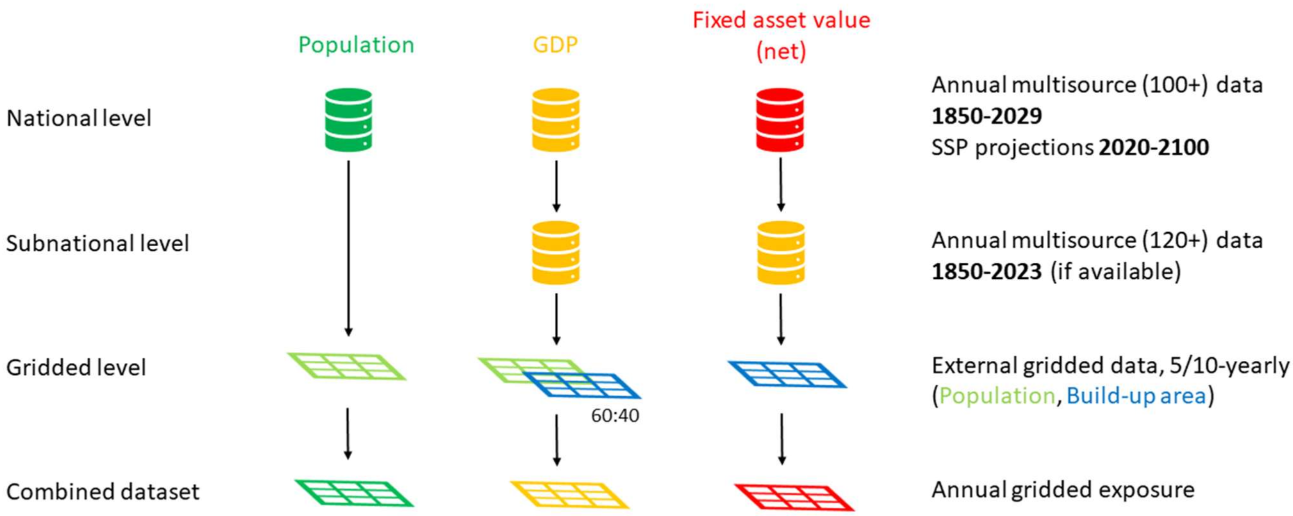

Figure 1. General approach of combining datasets from input database to final gridded product.

The database of national-level statistics was compiled for all countries of the world from 106 data sources for the historical period (1850-2023) and uses two datasets for harmonization with SSP projections (Table 2). The database is described in more detail in section 3.2.2 and the harmonization procedure in section 3.3. For selected countries (due to data availability), GDP is disaggregated according to a database of subnational GDP per capita compiled from 127 sources (see section 3.2.3). National population was disaggregated proportionally to interpolated and homogenized gridded data from other sources (see section 3.2.4). National or subnational GDP is disaggregated in 60% proportionally to gridded population and 40% proportionally to build-up surface, both from interpolated and homogenized external gridded data. The 60:40 ratio is intended to represent the typical ratio between the contributions of labour and capital inputs to GDP (see e.g. Arpaia et al. 2009, Guerriero 2019). Fixed asset value is distributed first proportionally to subnational GDP (if available) and then according to build-up surface. The result are three gridded datasets of exposure for any year between 1850 and 2100 (“combined dataset” in Figure 1) at any spatial resolution that are multipliers of 30 arc seconds.

Table 2. List of main datasets combined and harmonized in the multi-scale exposure dataset. Only the most important (by number of data points used) national-level datasets are listed (106 data sources in total).

Dataset	Version	Resolution used	Timespan used	Time step (yrs)	Variables used	Main role
Gridded						
HYDE	3.2.1	5′ (~9 km)	1850-1980	10	Population	Pre-1975 exposure disaggregation (combined with 1975 GHSL)
GHSL	R2023A	30′′ (~1 km)	1975-2030	5	Population, build-up surface	1975-2030 exposure disaggregation
Wang et al.	2022	30′′ (~1 km)	2020-2100	5	Population	Post-2030 exposure disaggregation (combined with 2030 GHSL)
National (main sources)						
MPD	2023	Country	1850-2010	1	GDP, population	Historical data source
UN NAMAD	2024	Country	1970-2022	1	GDP	Historical data source
PWT	10.01	Country	1950-2019	1	Fixed assets, GDP	Historical data source
UN WPP	2024 Rev.	Country	1950-2029	1	Population	Harmonization historical-projected data
IMF WEO	Oct 2024	Country	1980-2029	1	GDP	Harmonization historical-projected data
SSP	3.1.0	Country	2020-2100	5	GDP, population	Projected data source

3.2. Input datasets

3.2.1. Administrative divisions

Disaggregation of national and subnational statistics requires a high-resolution map of administrative units fully consistent with the databases of those statistics (sections 3.2.2 and 3.2.3). At national level, 248 geopolitical entities were defined based on ISO 3166-1 standard[7]. The numeric code of the ISO 3166-1 standard was used as a link between the database of national statistics in 3.2.2 and the map in vector and raster format[8]. The map was generated from OpenStreetMap national boundaries (including territorial waters) with some manual adjustments (Fig. 2). Similarly, OpenStreetMap was used to generate subnational boundaries, except European countries, which largely follow Eurostat's NUTS 2021 classification[9] using a map created for the HANZE database (Paprotny and Mengel 2023). In total, 2652 subnational units were defined for 82 countries (Fig. 3). They are assigned artificial numeric codes based on the ISO country code and serve as a link between the database of statistics and the raster map. Where available, the subnational units are also

[7] https://www.iso.org/obp/ui/#search
[8] ISO 3166-1 defines 249 'country codes', we exclude here Antarctica (uninhabited) and Åland Islands (part of Finland, defined in subnational statistics), but add Kosovo with codes used internationally (see https://population.un.org/wpp/downloads?folder=Documentation&group=Documentation)
[9] https://ec.europa.eu/eurostat/web/nuts/history

identifiable by ISO 3166-2 'country subdivision codes', Eurostat's NUTS codes or OECD 'territorial levels' codes[10].

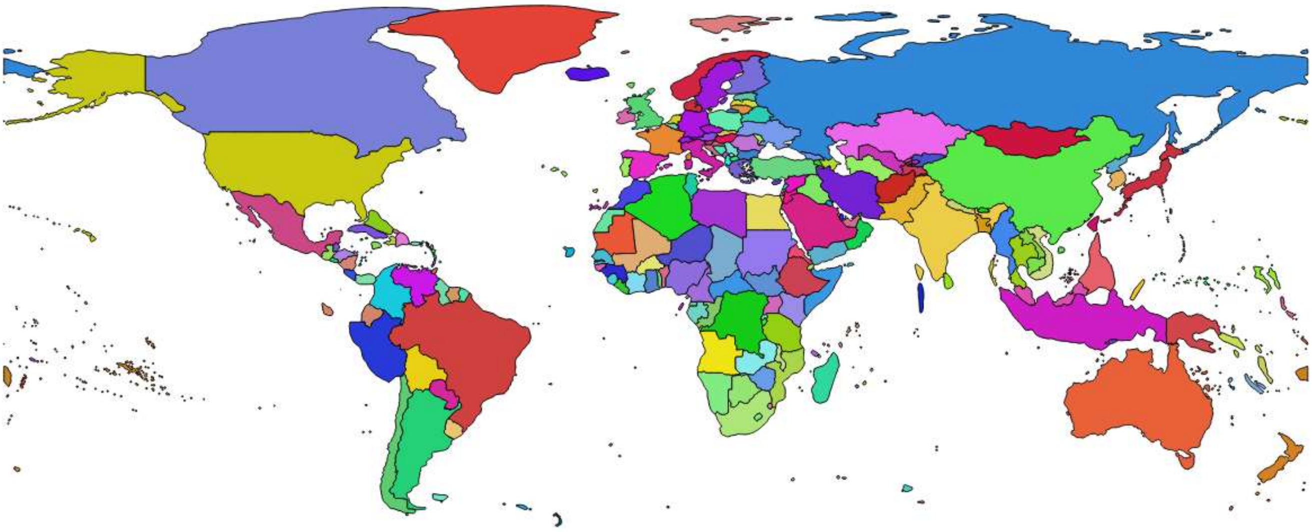

Figure 2. 248 national-level units defined in the database of national statistics (3.2.2)

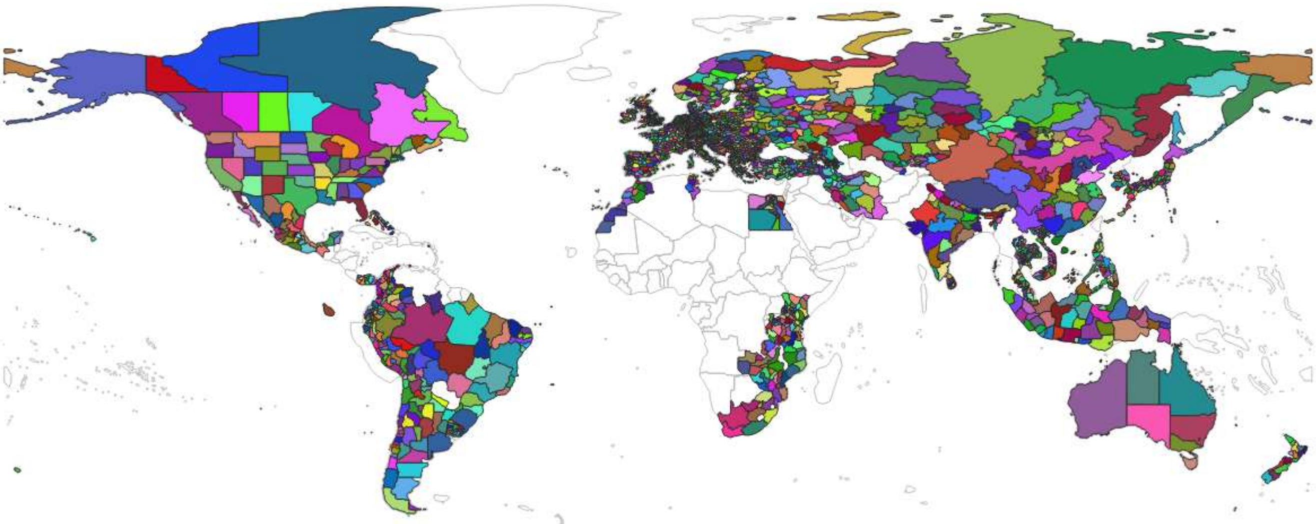

Figure 3. 2652 subnational-level units defined in the database of subnational GDP per capita (3.2.3).

3.2.2. National population, GDP and fixed asset value

The dataset of national-level statistics was compiled from 106 sources. Each country was researched individually to collect the best information publicly available in each case. For some countries, many interpolations, extrapolations or assumptions had to be made to compile complete population and GDP data. Further, all data had to be adjusted to a single country boundary definition for consistency of the disaggregation. Each case is documented individually in the file **National_exposure_all.xlsx** in the tab "Sources_by_country". If no source is provided for a particular country and years, it means that one of the default sources was used, as follows:

[10] https://www.oecd.org/en/data/datasets/oecd-geographical-definitions.html

- Population: 1950-2023 from United Nations (2024), *World Population Prospects 2024 Revision*;
- GDP per capita: 2011-2022 from United Nations (2025), *National Accounts - Analysis of Main Aggregates*;
- GDP per capita: 1850-2010 from Bolt and van Zanden (2025) *Maddison Project Database 2023*[11], interpolated where necessary.

Population is provided in the dataset as mid-year estimates in thousand persons. GDP per capita is provided in US dollars in constant 2017 prices and purchasing power parities (PPPs). PPPs are mostly revised estimates for 2017 from World Bank's 2021 International Comparison Programme database[12]. If 2017 PPPs were not available from that source, other PPP estimates were used as indicated in the dataset file (see tab 'PPP').

Projected population and GDP from SSP 3.1 database (IIASA 2024) were interpolated from 5-yearly timesteps provided in that dataset. Two GDP projections are available from the SSP database, with different spatial coverage. The OECD projection was used were available, and the IIASA projection were not. The IIASA GDP data start in 2025, therefore they were extrapolated back to 2020 using 2025-2030 trend. Where population or GDP projections are not available (for small countries), either appropriate regional aggregates were used or, in case of dependent territories, data for the territory's sovereign. The linking of country definitions between the historical dataset of statistics and the SSP dataset is done in **SSP_3_1_main_drivers.xlsx** file in the tab 'SSP_ISO_reference'.

Fixed asset value is defined as the stock of all produced non-financial assets that are used repeatedly or continuously in production processes for more than one year, both tangible (e.g. buildings, machinery, transport equipment) and intangible (e.g. software and patents)[13]. Inventories, valuables, non-produced assets (e.g. land) and consumer durables are excluded. The value of assets here is their net value, i.e. after deprecation, and inserted into the database as fixed asset-to-GDP ratio. Data on fixed assets are limited and were mostly obtained from Penn World Tables for 1950-2019 (Feenstra et al. 2015)[14] as well as Eurostat and OECD databases for some countries, supplemented by national statistics. Pre-1950 data are mostly from Goldsmith (1985), but as with population and GDP all individual data sources are identified on country basis in the database file. To gap-fill the fixed asset value, we leverage the observation that fixed asset-to-GDP ratio has a long-term upward trend (popularized by Piketty and Zucman 2014). This is observed in Figure 4 showing available data of GDP per capita and fixed asset-to-GDP ratio (1458 estimates)[15] that have been normalized using a standard normal distribution. A positive Spearman's correlation of 0.32 is measured from the data. Therefore, we use copulas, a statistical model, to model the relationship between GDP per capita and fixed asset-to-GDP ratio

[11] https://www.rug.nl/ggdc/historicaldevelopment/maddison/releases/maddison-project-database-2023
[12] https://www.worldbank.org/en/programs/icp/data
[13] In national accounts systems (SNA 2008, ESA 2010), this corresponds to it 'fixed assets' category (AN.11).
[14] https://www.rug.nl/ggdc/productivity/pwt/
[15] Only every 10 years of data was used to reduce autocorrelation resulting from very large inertia of fixed asset stock.

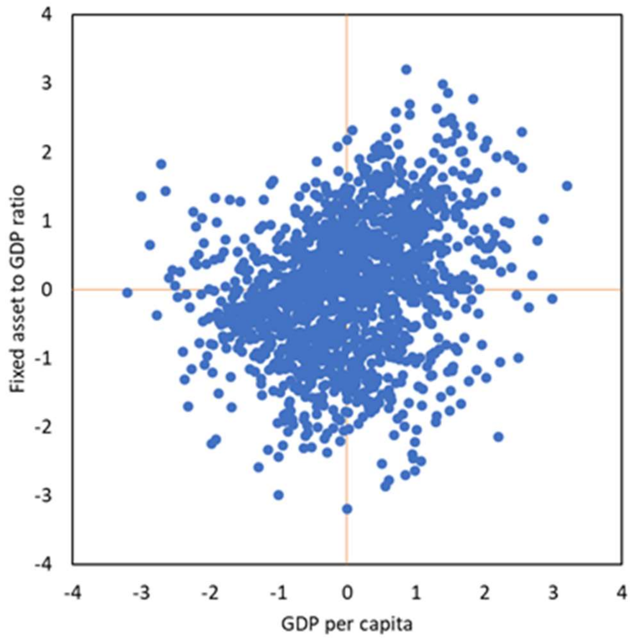

Figure 4. Normalised GDP per capita and fixed asset-to-GDP ratio using a standard normal distribution. The Spearman's correlation between the two datasets is 0.32.

The goodness–of-fit of different copula types was measured using a blanket test based on the Cramèr–von Mises statistic described by Genest et al. (2009). The Frank copula was shown to be the best fit. By randomly sampling the copula conditionalized on GDP per capita (which was compiled for all countries and years), the fixed asset-to-GDP ratio was computed for all missing years. If some fixed asset value data was already available for a given country, timeseries from gap-filling were used to extrapolate the earliest and latest available information. The results in the global gap-filled estimates (Fig. 5, light blue line) show an upward trend in fixed asset-to-GDP ratio, except for the period containing the Great Depression and World War II (c. 1930-1950).

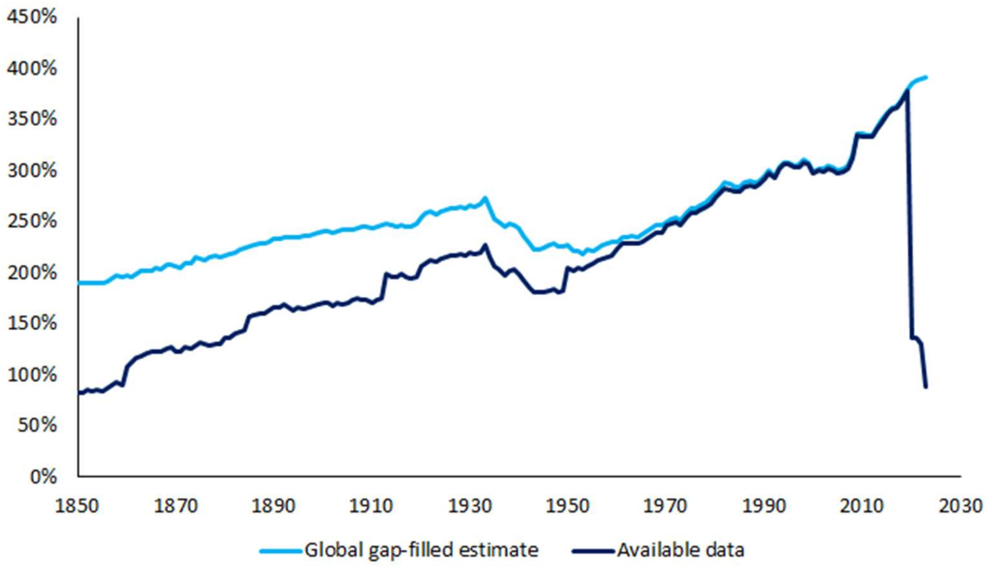

Figure 5. Net fixed asset value in the world relative to global GDP (%), summation of available data compared with summation of gap-filled estimates.

3.2.3. Subnational GDP per capita

Subnational GDP per capita was collected for 82 countries comprising 2652 subnational administrative divisions between 1850 and 2023, representing about 89% of global GDP. Data availability is largely limited to the 21st century, but some countries have estimates spanning many decades before, even for 1850 (Denmark and Sweden). In many cases, the data was interpolated between available benchmark estimates. All data were, to the extent possible, homogenized to the same administrative boundaries for all years. In many cases of European countries, data at a more aggregated level are available for earlier decades (typically Eurostat's NUTS 2 level), therefore the values at more detailed NUTS 3 level were extrapolated using the appropriate upper-level divisions. All data sources and transformations made to the data are documented for every country, in the file **Subnational_exposure_all.xlsx** in the tab "Sources_by_country". The data are presented as percentages relative to national GDP per capita. Source GDP and population data, where available, are also included in the dataset, but only the relative GDP per capita is used by the exposure disaggregation algorithm. For years where there is no data, the earliest timestep is used for all years prior and the latest data is used for all years afterwards.

3.2.4. External gridded data

The primary dataset used to disaggregate national and subnational data is the Global Human Settlement Layer (European Commission 2023). Among others, it provides gridded population and build-up surface in 30 arc second resolution between 1975 and 2030 with a 5-year timestep. This dataset was preferable to other available population grid due to a corresponding build-up area and also higher quality. The author's comparison of GHSL, Gridded Population of the World[17] and HYDE (Klein Goldewijk et al. 2017) has shown a much better match with georeferenced population for Europe from 2011 census round[18] for GHSL than for the other two population grids. Despite a higher spatial resolution, the validation results for Gridded Population of the World were much worse than for HYDE.

For years before 1975, the GHSL population and build-up surface in 1975 was extrapolated using changes in population at 5 arc minute resolution from HYDE 3.2.1, interpolated from 10-yearly timesteps[19]. HYDE dataset incorporates many subnational population statistics and an algorithm to represent changes in urbanization processes. It remains the best source of global gridded population data before 1975 despite limited documentation and quality issues with input subnational data.

For timesteps after 2030, the GHSL population and build-up surface in 2030 was extrapolated using changes in population at 30 arc second resolution from a dataset by Wang et al. (2022), interpolated from 5-yearly timesteps. The dataset from Wang et al. (2022), which used machine learning to compute spatial patterns of historical changes in population distribution and extrapolated the into the future (for different SSP scenarios) is preferable to other published datasets, which were found to strongly misrepresent observed patterns of population change in Europe (see Steinhausen et al. 2022).

3.3. Harmonization of historical and projected data

SSP projections were created with a 2020 base year and 5-yearly timestep. Therefore, they do not account for recent events affecting short-term changes especially in GDP. The algorithm for generating the global exposure dataset can harmonize (default option) or not harmonize the SSP projections with latest historical data and short-term projections (Fig. 6). Harmonization is done both on the level of national statistics and the gridded datasets. With harmonization, historical data are used until 2023, extrapolated with population from United Nations *World Population Prospects 2024 Revision* and GDP from International Monetary Fund's *World*

[17] https://www.earthdata.nasa.gov/data/projects/gpw

[18] https://ec.europa.eu/eurostat/web/gisco/geodata/population-distribution/geostat

[19] Newer HYDE 3.3 dataset is available, but it introduces some unexplained changes to population distribution in some countries, in the author's opinion reducing the quality of the data. Consequently, the previous version was used here.

Economic Outlook (IMF 2024). The latter is published every six months with projects up to five years ahead. Therefore, population and GDP are extrapolated under a single scenario until 2029. Then, population and GDP slowly converge with the SSP projections by the year 2100. The population and GDP in 2100 are assumed to be the same as if the SSP projection applied already all years from 2020 to 2100. The difference between the SSP trajectory starting in 2020 and harmonized projection for 2029 is linearly reduced to zero by 2100. In the variant without harmonization, SSP trajectory is applied from year 2020.

The harmonization on gridded level is done by using GHSL deterministic projections for years 2021-2030 and only extrapolating the population and build-up surface with estimates from Wang et al. (2022) after 2030. Without harmonization, estimates from Wang et al. (2022) are already used after 2020, using the 2020 GHSL grids as baseline.

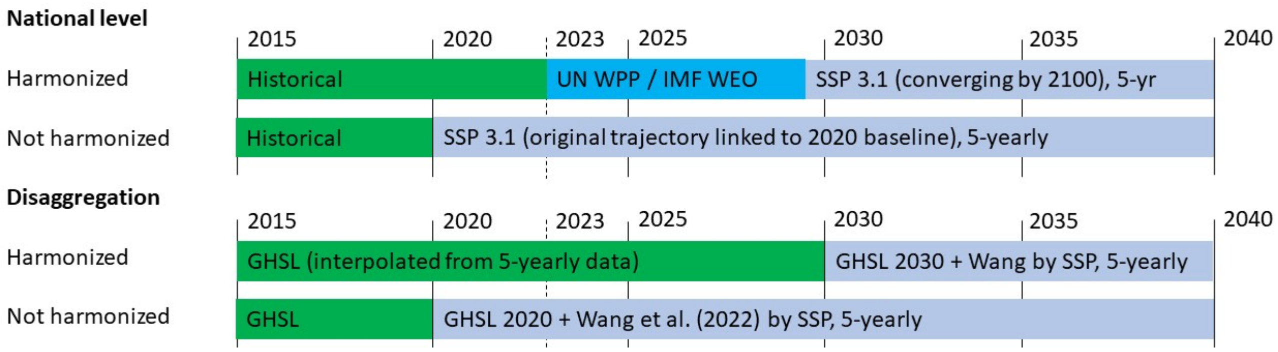

Figure 6. Harmonization of historical and projected data using WPP and WEO datasets. Without harmonization, SSP projections start directly after 2020 timestep.

An example harmonization for GDP per capita of Estonia is shown in Figs. 7-8. The SSP timeseries (dark red in Fig. 7) have a 5-year timestep, therefore do not capture the short-term economic variations (in this case, the COVID-19 pandemic and a recession resulting from the Russian invasion of Ukraine in 2022), which are included in our historical GDP dataset. Additionally, there is an apparent error in the valuation baseline of SSP data (Fig. 7a): the data are according to the metadata in 2017 prices and PPPs, but comparison with other datasets such as the World Economic Outlook indicates that they were in fact computed using the 2015 price level. Therefore, SSP data were always extrapolated from our historical timeseries starting with year 2020 to avoid a break in the series (Fig. 7b). In the default harmonized dataset, the GDP per capita timeseries for Estonia follow the historical database until 2023 and then follow short-term projections from the World Economic Outlook until 2029 (blue dashed line in Fig. 7b). Without harmonization, the interpolated SSP projections start already after 2020 (red solid line in Fig. 7b).

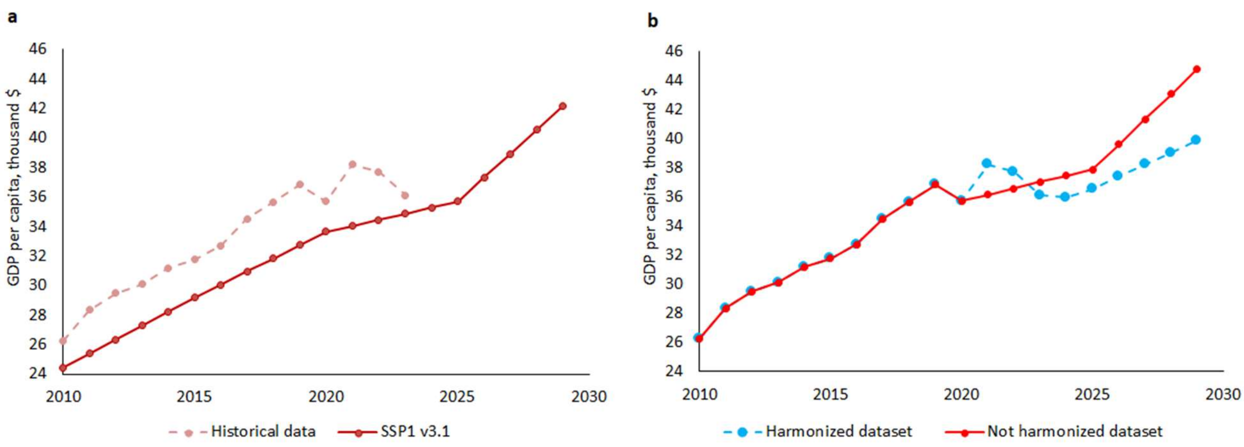

Figure 7. GDP per capita in Estonia from this study and from the SSP database, before (a) and after (b) correction.

In the harmonized dataset, GDP per capita of Estonia linearly converges with the SSP projection after 2029, merging with the value of the not harmonized dataset in year 2100 (light-colored dashed lines in Fig. 8). Without harmonization, the interpolated SSP projections is extrapolated from year 2020 all the way to 2100 (dark solid lines in Fig. 8).

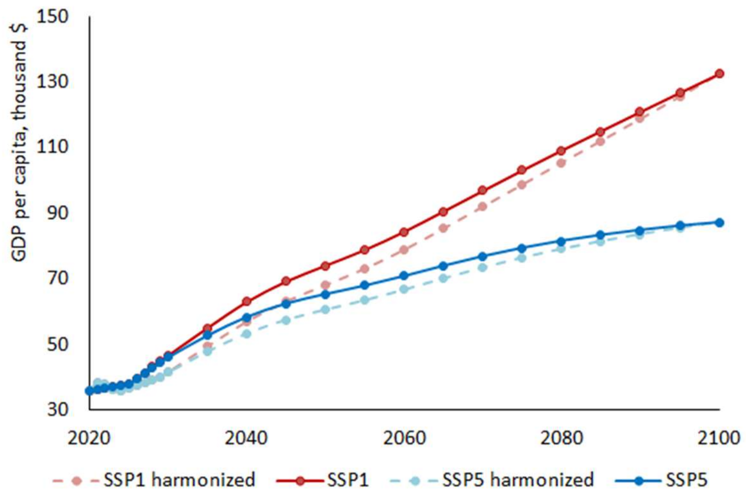

Figure 8. Projected GDP per capita in Estonia harmonized (light color, dashed) and not harmonized (dark color, solid).

3.4. Maintaining and updating the dataset

Dynamic economic and, to a lesser extent, demographic changes influence projected changes in the near-term perspective. Historical experience has shown that short-term swings do not invalidate the long-term trajectory but nonetheless affect quantification of exposure of recent events (e.g. within the framework of the COMPASS project, 2024-2026). The code was designed to facilitate relatively easy updating of the exposure dataset. While the historical statistics datasets (1850-2023) require considerable effort in updating, the harmonization with WPP and WEO datasets can be done relatively easily. The respective Excel files used as inputs can be downloaded directly from the sources and replace existing ones. WPP is updated every two years (typically in July), but WEO is updated every six months (April and October). This enables relatively frequent update of the harmonized data. The SSP projections were updated in 2024 after 11 years, but it is still uncertain whether they will form the basis for CMIP7 climate and impact modelling (van Vuuren et al., 2025). The SSP data is loaded into the code as a single file based on the raw dataset downloadable from the IIASA's SSP database.

GHSL, the main dataset used in the disaggregation, is in active research and updated roughly once a year. An improved future version could be used to replace rasters currently by the code. Raw data from GHSL is used, which should make updating easy.

We encourage users to report bugs in the code by opening an issue on github (https://github.com/HORIZON-COMPASS/Exposure-and-vulnerability-modelling). We also welcome information on any problems or errors in the data or documentation. Such issues, and other inquiries, can be reported directly to the main author (dominik.paprotny@pik-potsdam.de).

4. References

Arpaia, A., Pérez, E., and Pichelmann, K. (2009) Understanding labour income share dynamics in Europe. Economic Papers, 379, European Commission, Directorate-General for Economic and Financial Affairs, https://ec.europa.eu/economy_finance/publications/pages/publication15147_en.pdf

Bolt, J., van Zanden, J. L. (2025) Maddison style estimates of the evolution of the world economy: A new 2023 update. Journal of Economic Surveys, 39, 631–671. https://doi.org/10.1111/joes.12618

European Commission (2023) GHSL Data Package 2023. JRC133256, Publications Office of the European Union, Luxembourg, https://doi.org/10.2760/098587

Feenstra, R. C., Inklaar, R., and Timmer, M. P. (2015) The Next Generation of the Penn World Table. American Economic Review, 105(10), 3150-3182, https://doi.org/10.1257/aer.20130954

Genest, C., Rémillard, B., and Beaudoin, D. (2009) Goodness-of-fit tests for copulas: A review and a power study. Insurance Math. Econ., 44(2), 199–213.

Goldsmith, R. W. (1985) Comparative national balance sheets: a study of twenty countries, 1688-1978. University of Chicago Press, Chicago, 364 pp.

Guerriero, M. (2019) The Labor Share of Income around the World: Evidence from a Panel Dataset. ADBI Working Paper Series, 920, Asian Development Bank Institute, https://www.adb.org/sites/default/files/publication/484346/adbi-wp920.pdf

IIASA (2024) SSP Scenario Explorer 3.1.0, https://data.ece.iiasa.ac.at/ssp/#/workspaces

IMF (2024) World Economic Outlook Database, https://www.imf.org/en/Publications/WEO/weo-database/2024/October.

Klein Goldewijk, K., Beusen, A., Doelman, J., and Stehfest, E. (2017) Anthropogenic land use estimates for the Holocene – HYDE 3.2. Earth System Science Data, 9, 927–953, https://doi.org/10.5194/essd-9-927-2017

Paprotny, D., and Mengel, M. (2023) Population, land use and economic exposure estimates for Europe at 100 m resolution from 1870 to 2020. Scientific Data, 10, 372, https://doi.org/10.1038/s41597-023-02282-0

Piketty, T., and Zucman, G. (2014) Capital is Back: Wealth-Income Ratios in Rich Countries, 1700-2010. Quarterly Journal of Economics, 129(3), 1155-1210.

Steinhausen, M., Paprotny, D., Dottori, F., Sairam, N., Mentaschi, L., Alfieri, L., Lüdtke, S., Kreibich, H., and Schröter, K. (2022) Drivers of future fluvial flood risk change for residential buildings in Europe. Global Environmental Change, 76, 102559, https://doi.org/10.1016/j.gloenvcha.2022.102559

United Nations (2024) World Population Prospects, https://population.un.org/wpp/

United Nations (2025) National Accounts - Analysis of Main Aggregates (AMA), https://unstats.un.org/unsd/snaama/

Van Vuuren, D.P., Riahi, K., Calvin, K., Dellink, R., Emmerling, J., Fujimori, S., KC, S., Kriegler, E., O'Neill, B. (2017) The Shared Socio-economic Pathways: Trajectories for human development and global environmental change. Global Environmental Change, 42, 148-152, https://doi.org/10.1016/j.gloenvcha.2016.10.009

Van Vuuren, D., O'Neill, B., Tebaldi, C., Chini, L., Friedlingstein, P., Hasegawa, T., Riahi, K., Sanderson, B., Govindasamy, B., Bauer, N., Eyring, V., Fall, C., Frieler, K., Gidden, M., Gohar, L., Jones, A., King, A., Knutti, R., Kriegler, E., Lawrence, P., Lennard, C., Lowe, J., Mathison, C., Mehmood, S., Prado, L., Zhang, Q., Rose, S., Ruane, A., Schleussner, C.-F., Seferian, R., Sillmann, J., Smith, C., Sörensson, A., Panickal, S., Tachiiri, K., Vaughan, N., Vishwanathan, S., Yokohata, T., and Ziehn, T. (2025) The Scenario Model Intercomparison Project for CMIP7 (ScenarioMIP-CMIP7), EGUsphere [preprint], https://doi.org/10.5194/egusphere-2024-3765

Wang, X., Meng, X., and Long, Y. (2022) Projecting 1 km-grid population distributions from 2020 to 2100 globally under shared socioeconomic pathways. Scientific Data, 9, 563, https://doi.org/10.1038/s41597-022-01675-x