

scientific data



OPEN

DATA DESCRIPTOR

Global gridded GDP data set consistent with the shared socioeconomic pathways

Tingting Wang & Fubao Sun^{1,2,3,4}

The vulnerability, exposure and resilience of socioeconomic activities to future climate extremes call for high-resolution gridded GDP in climate change adaptation and mitigation research. While global socioeconomic projections are provided mainly at the national level, and downscaling approaches using nighttime light (NTL) images or gridded population data can increase the uncertainty due to limitations. Therefore, we adopt an NTL-population-based approach, which exhibits higher accuracy in socioeconomic disaggregation. Gross regional product of over 800 provinces, which covering over 60% of the global land surface and accounted for more than 80% of GDP in 2005, were used as input. We present a first set of comparable spatially explicit global gridded GDP projections with fine spatial resolutions of 30 arc-seconds and 0.25 arc-degrees for the historical period of 2005 and for 2030–2100 at 10-year intervals under the five SSPs, accounting for the two-child policy in China. This gridded GDP projection dataset can broaden the applicability of GDP data, the availability of which is necessary for socioeconomic and climate change research.

Background & Summary

The shared socioeconomic pathways (SSPs), which qualitatively and quantitatively describe possible future patterns of global socioeconomic development under different challenges to climate change mitigation and adaptation¹, are of vital importance in global scientific assessment reports^{2,3} and the current literature^{4,5}. The widely used Scenario Model Intercomparison Project (ScenarioMIP) was formed based on different SSPs corresponding to specific representative concentration pathways (RCPs) within Phase Six of the Coupled Model Intercomparison Project (CMIP6)⁴. Future socioeconomic impacts on climate change are built upon socioeconomic projections and require gridded gross domestic product (GDP) data of a higher resolution for future scenarios^{4,5}.

GDP is the most commonly used indicator to assess and compare economic development within and across countries^{6–8}. Some global institutions (e.g., the World Bank, the Organization for Economic Co-operation and Development (OECD)), research groups (e.g., Penn World Tables (PWTs)), and researchers^{6,9} have collected or extended global GDP data for various applications. However, gross regional product (GRP) data at the provincial (state), city or county scales are often problematic, especially in many developing countries^{7,8}. Kummu *et al.*⁷ presented a set of global gridded GDP data in 2011 purchasing power parity (PPP) international dollars for 1990–2015 at 5 arc-minutes annually and at 30 arc-seconds for 1990, 2000, and 2015, making it easily integrated with data from other disciplines^{10,11}. While national and supranational GDP in the SSP database, which is the most widely used inputs in disaggregation, are mainly in 2005 PPP USD, resulting in biased data when compared with other historical periods directly. It is therefore crucial to spatialize GDP in 2005 PPP USD at a fine scale so that it can be easily analyzed and compared with GDP projections of future scenarios.

Gridded GDP estimates can be obtained using a spatially explicit gridded population dataset and/or satellite-derived nighttime light (NTL) images to better support current research^{8,12–15}. However, the GDP proportional to population estimates are based on the assumption that GDP per capita is uniformly distributed

¹Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China. ²State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, 830011, China.

³Akesu National Station of Observation and Research for Oasis Agro-ecosystem, Akesu, 843017, China. ⁴College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, 100049, China. e-mail: Sunfb@igsnrr.ac.cn

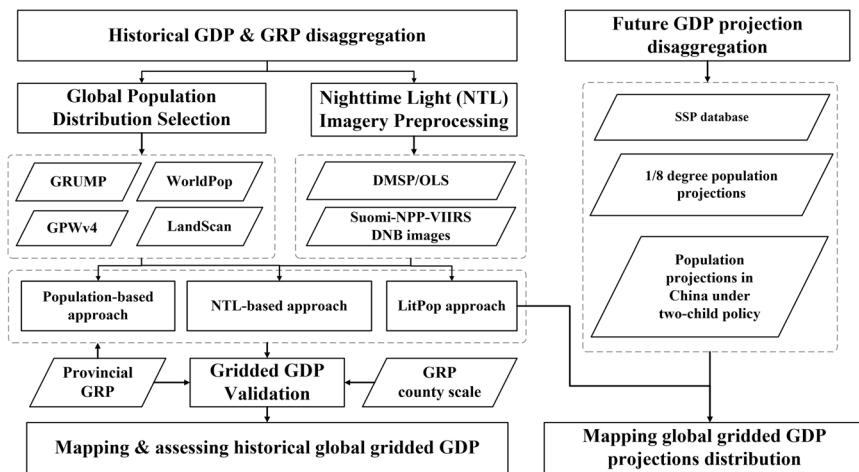


Fig. 1 Workflow for developing the spatially explicit GDP dataset consistent with the SSPs.

within an administrative boundary and are often biased at smaller scales⁹. The Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS) NTL images have been widely used in GDP downscaling for 1992–2013. However, these gridded GDP estimates are highly dependent on the NTL values and are often underestimated in urban centers and overestimated in rural regions due to the saturation problem¹⁶. Fortunately, this can be mitigated to a certain extent when incorporating gridded populations^{17,18}. In contrast, the global Soumi National Polar-Orbiting Partnership Visible Infrared Imaging Radiometer Suite (NPP-VIIRS) NTL images address this saturation problem and provide a more accurate proxy for GDP downscaling after 2012¹².

The increasing vulnerability, exposure, and resilience of socioeconomic activities to future climate extremes are promoting a move beyond administrative boundaries to enable flexible integration with the socioeconomic projections of the five SSPs^{5,19,20}. However, the widely used GDP projections in the SSP database are restricted at national and supranational scales and depict a wide range of uncertainties within organizations^{21,22}. The spatially explicit demographic projections at 0.125 arc-degrees⁴ have been further downscaled to 1-km resolution²³ and are publicly available, but the GDP projections are rather limited in their ability to facilitate global climate change research. Using gridded population estimates, Geiger *et al.*⁹ downscaled and provided a set of spatially explicit global GDP estimates at 5 arc-seconds for 1850–2100 under SSP2 scenario. Murakami and Yamagata²⁴ disaggregated the global population and GDP into 0.5 arc-degree grids for SSP1–3 scenarios, but SSP4–5 projections were unavailable to the public. Jiang *et al.*^{25,26} projected the Chinese population and GDP at the provincial level, and Huang *et al.*²⁷ later showed that Chinese GDP would increase by 38.1–43.9% by the late 21st century under the two-child policy.

To date, high-resolution global gridded GDP projections for all five SSPs are rather limited, and historical GDP has been bound to various data sources and to clearly limited downscaling approaches that use population estimates or NTL images. Hence, we downscale and present a set of comparable spatially explicit global GDP estimates for both the historical period and 2030–2100 at 10-year intervals. These estimates have high spatial resolutions of 1 km and 0.25 arc-degrees under all five SSPs, obtained by incorporating reliable data sources and methods to broaden the applicability of GDP data to scenario-based climate change research.

Data and Method

The NTL images and gridded population estimates are well correlated with GDP and have been widely used in global and regional GDP disaggregation at various spatial scales but have their own limitations, as described above. The pursuit of a reliable GDP projection dataset requires more accurate inputs, proper downscaling approaches, etc. In what follows, we describe the data, including the gridded population estimates and NTL images used to generate the base map, GDP and GRP, and the methods used for GDP downscaling. The data cleaning and estimation workflow is presented in Fig. 1.

Gridded population dataset. Gridded population data have been widely used to spatialize GDP (GRP) to a fine scale^{9,18}. To obtain a set of highly accurate global gridded population data, data from the Gridded Population of the World dataset, Version 4 (GPWv4), Revision 11 (<http://sedac.ciesin.columbia.edu/data/collection/gpw-v4>)²⁸, WorldPop (www.worldpop.org)^{29,30}, the Global Rural-Urban Mapping Project, Version 1 (GRUMPv1) dataset ([https://sedac.ciesin.columbia.edu/data/set/grump-v1-population-density](http://sedac.ciesin.columbia.edu/data/set/grump-v1-population-density)), and the LandScan Global Population database ([https://landscan.ornl.gov/](http://landscan.ornl.gov/))³¹ are compared at the national, provincial (state), and county scales first (Figures S1, S2). The LandScan Global Population database from the Urban Oak Ridge National Laboratory, USA, was chosen due to its relatively higher accuracy at the county scale (Figure S2), its best census and geographic data in terms of availability, and remote sensing imagery analysis techniques within a multivariate dasymetric modeling framework to disaggregate census counts within an administrative boundary³¹.

Nighttime light images. The NTL images are well correlated with global and regional socioeconomic activities and have also been widely used in GDP disaggregation^{8,14}. The widely used version 4 of the DMSP-OLS stable

NTL images in 2005 and 2012 were obtained from the National Oceanic and Atmospheric Administration's National Geophysical Data Center (NGDC) at <https://www.ngdc.noaa.gov/eog/dmsp/downloadV4composites.html>. These data have a spatial resolution of 30 arc-seconds and latitudinal and longitudinal ranges of 75°N to 65°S and 180°W to 180°E. The values for the DMSP-OLS stable NTL data range from 0 to 63, and the saturation problem (value of 63) is significant, mainly in large city centers and other brightly lit zones¹².

In addition, the Suomi-NPP-VIIRS Day/Night Band (DNB) images (https://eogdata.mines.edu/download_dnb_composites.html) from 2015 were obtained. These images have a higher resolution of 15 arc-seconds. To ensure the same cell size, NPP-VIIRS NTL images were resampled to a spatial resolution of 30 arc-seconds. In addition, the DMSP-OLS images in 2012 were used to eliminate the negative values in the NPP-VIIRS DNB images.

GDP and GRP. The World Development Indicators assembled by the World Bank (WB-WDI) provide a vast resource of relevant, high-quality, and internationally comparable socioeconomic statistics for 217 economies from 1961 to the present. For most economies, GDP in 2017 PPP international dollars, in current prices, and in other units can be easily obtained from <https://databank.worldbank.org/home.aspx>. However, GDP in 2005 PPP USD is no longer available from the WB-WDI, which makes GDP projections under the SSPs in 2005 PPP USD incomparable with historical data. We therefore obtained the GDP of 195 countries from Geiger *et al.*^{6,9}, which mainly comes from PWT8.1, in 2005 PPP USD, which is consistent with the SSP projections, and data for missing countries are taken from PWT9.0 and the WB-WDI after rescaling from 2011 to 2005 PPP USD.

GRP in current PPP USD for over 800 provinces (states) in 48 countries (35 OECD countries plus 13 non-OECD countries) were obtained from OECD data at https://stats.oecd.org/Index.aspx?DatasetCode=SNA_TABLE1. However, GRP data for Egypt (http://hdr.undp.org/sites/default/files/2008_egypt_nhdr_en.pdf) and Kenya (http://hdr.undp.org/sites/default/files/kenya_2006_en.pdf) were found in the Human Development Reports from the United Nations Development Programme since GRP (PPP) and GRP per capita (PPP) in 2005 were provided. These 50 countries cover over 60% of the global land surface and accounted for more than 80% of global GDP in 2005.

Moreover, GRP at a more explicit county-level was obtained for China (2160 counties with valid data in 2005 from the China County Statistical Yearbook 2006) and the U.S. (3071 out of 3145 counties from the U.S. Bureau of Economic Analysis (<https://www.bea.gov/data>)) for validation purposes. The Chinese figures were converted to USD using the conversion factors provided by the World Bank.

SSP projection data. Long-term demographic and GDP projections have been made available by different organizations to facilitate research on future impacts, adaptation, and vulnerability. There are several widely used demographic and GDP projections available to the public.

SSP database. Based on the assumptions of the SSP storylines in terms of the main drivers of economic growth, three sets of global GDP projections have been developed and are provided in the SSP database^{21,32–34}. As recommended by the SSP database, GDP projections in 2005 PPP USD were obtained from the OECD³³ for 184 countries up to the end of the 21st century under the five SSP scenarios at <https://tntcat.iiasa.ac.at/SspDb>. Then, national GDP growth rate projections relative to 2005 can be calculated using these figures. Projections in the remaining missing countries were filled in by using supranational GDP growth rate projections.

The national and supranational population projections in the SSP database fail to meet the increasing demand for spatially explicit demographic projections for these future scenarios. Hence, we obtained the gridded population projections developed by Jones and O'Neill¹⁴, who further extended these national totals and produced a scenario-based gridded population dataset that is qualitatively consistent with the assumptions in the SSP narratives regarding spatial development patterns at a spatial resolution of 1/8° (approximately 7.5 arc-minutes at the equator). The GRUMP population dataset from 2000 was used as the base map. Using a parameterized gravity-based downscaling model, the demographic drivers, namely, current fertility, income, urbanization, and international migration, are explicitly included in the population projections corresponding to each SSP to produce gridded population projections for all five SSPs that are quantitatively consistent with national population and urbanization projections.

Chinese population projections under the two-child policy. The implementation of the two-child policy in 2016 will no doubt have a substantial effect on the demographic and socioeconomic projections for China in the long run. In addition, the clear differences among the drivers (e.g., population distribution by age, sex, and educational enrollment in the historical period) between the census data in China and the United Nations (UN) estimations²⁷ require revisions and updates to the Chinese population and GDP projections for the five SSPs. Jiang *et al.*^{25,26} employed data from the China Statistical Yearbook and the Sixth National Population Census and made projections for 2020–2100 based on assumptions about future education, fertility, mortality, and migration for the five SSPs. Their projections account for the two-child policy. Using a parameterized multidimensional model, population projections by age, sex, education, and the economy are developed for China up to the year 2100 under the framework of the SSPs but are revised based on regional characteristics. The provincial demographic projections are further downscaled and provided at a 0.5-degree resolution for the five SSPs.

Population-based GDP disaggregation method. A broad literature has emphasized the key role of human capital as a driver of socioeconomic growth and population is the central link between human capital and economic growth in almost all econometric models^{32,33,35}. Shiogama *et al.*³⁶ suggested the robustness of an ensemble learning-based downscaling approach, which is defined by (baseline variable) × (control variable) in accordance with distribution weights. GDP can thus be proportionally disaggregated to match population counts

based on the assumption that GDP per capita is uniformly distributed within a given administrative boundary. This population-based GDP disaggregation (denoted GDP_{pop}) can be described as follows:

$$GDP_{pop} = Pop_{pixel} \times Pcap_i = Pop_{pixel} \times \frac{GDP_i}{Pop_i} \quad (1)$$

where Pop_{pixel} is the population count in each grid, which acts as the baseline variable, and $Pcap_i$ (GDP per capita within a given administrative boundary i) acts as the control variable.

NTL-based GDP disaggregation method. NTL images have served as a useful proxy for socioeconomic development and have been widely used to improve the quality of socioeconomic data, especially in many developing countries given their high spatial resolution^{8,14,16}. There are two separate annual stable light images derived from two satellites (F14 and F15, working simultaneously to collect global NTL information) in DMSP-OLS, and average visible, stable light, and cloud free coverage products are provided for 2005. The stable light product from F15, which contains a composite of all available cloud-free data with a spatial resolution of 1 km, was chosen since newer sensors have less degradation in their data quality and have been utilized to disaggregate global GDP into 1 km × 1 km grids. The NPP-VIIRS products have several demonstrated improvements over the DMSP-OLS NTL images, e.g., resolving the saturation problem and improving their calibration with a higher spatial resolution of 15 arc-seconds. Given their public availability and regular updates, global NTL images from two satellites have proven to be a useful source of information and are the most commonly used proxy in socioeconomic and climate change research. Based on the statistical linear relationship between GDP and the luminosity in NTL images, GDP is directly redistributed to each grid in proportion to the values within a given administrative boundary. Thus, the NTL-based GDP (denoted GDP_{Lit}) can be described as:

$$GDP_{Lit} = GDP_{per_light} \times Lit_{pixel} = \frac{GDP_i}{SL_i} \times Lit_{pixel} \quad (2)$$

where GDP_i is a GDP total, SL_i is the sum of the values from NTL images of administrative unit i, and GDP_{per_light} is a constant within this administrative unit i.

NTL-population-based approach. Studies have shown that the saturation problem in the DMSP-OLS NTL images has resulted in overestimation for large urban areas and underestimation for rural and distant regions^{16–18}. Zhao *et al.*¹⁸ improved the accuracy of GDP disaggregation by combining NTL images with gridded population data using a function in which GDP is proportional to NTL luminosity and population count and named it the LitPop approach. Eberenz *et al.*¹⁷ further improved this method by filling in the nonilluminated cells with population counts and the depopulated zones with NTL values, providing a better proxy for the spatial disaggregation of economic indicators. The LitPop images were first produced by multiplying the data from the NTL images with data from the Landscan population dataset (Eq. 3), assuming that the values in the nonilluminated grids are equal to 1 to ensure that nonilluminated but populated grid cells are not assigned a value of zero and that illuminated but unpopulated areas are not assigned a value other than zero. The LitPop images were then used (Eq. 4) for GDP downscaling (denoted GDP_{LitPop}) across the globe.

$$LitPop = \begin{cases} Pop & Lit = 0 \\ Lit \cdot Pop & Lit > 0 \& Pop > 0 \\ Lit & Pop = 0 \end{cases} \quad (3)$$

$$GDP_{LitPop} = \frac{GDP_i}{SL_i} \times LitPop_{pixel} \quad (4)$$

where $LitPop$ is the value of each pixel in the LitPop images, and SL_i is the sum of the LitPop data within administrative unit i.

Global GDP downscaling in 2005. Since GRP estimates in 2005 PPP USD are not available, we first made the following adjustments that are consistent with the SSP projections to ensure a more accurate GDP disaggregation. Using the ratio between GDP in 2005 PPP USD^{6,9} and GDP in current PPP USD (obtained from the WB-WDI) for each country, we first recalculated GRP in 2005 PPP USD based on a simple method (Eq. 5) under the assumption that the GDP in 2005 PPP USD is directly redistributed to each subnational area in proportion to the GRP in current prices within a given national administrative boundary i, overlooking regional differences in inflation.

$$GRP_{2005} = \frac{GDP_{i_2005}}{GDP_{i_current}} \times GRP_{current} \quad (5)$$

GDP in 2005 PPP USD for the 50 countries mentioned above was replaced by the GRP of over 800 provinces (states) after being converted to 2005 PPP USD. Following the LitPop procedure, the NTL-based approach, and the population-based disaggregation method, 2005 GDP and GRP (in 2005 PPP USD) for countries across the globe were spatially linked to the corresponding GIS-based administrative boundaries (<https://www.diva-gis.org/>) and then disaggregated into 1 km × 1 km grids using the DMSP-OLS NTL images and the Landscan population

dataset for 2005. After making the comparison reported in the Technical Validation Section, the LitPop approach to GDP downscaling was adopted for both 2005 and for the future scenarios.

GDP projection downscaling for the five SSPs. Using the LitPop approach, the GDP projections for the five SSPs were downscaled to a 1-km resolution. The GDP projections were first completed for all countries (regions) for the five SSPs. National projections for 177 countries and supranational projections for the world regions where the remaining countries are located were obtained from the OECD projections in the SSP database. GDP growth rates relative to 2005 (the figures in the SSP database) were calculated, and the supranational values were used to fill data for countries with missing values. GDP and GRP projections were then generated using these GDP growth rates and official figures from 2005 to complete the data series and ensure global consistency for all five SSPs.

Next, the population projections were downscaled to a 1-km resolution. Using the spatial distribution pattern in the LandScan population dataset for 2015, the 1/8-degree population projections consistent with the SSPs¹⁹ and the Chinese projections developed by Jiang *et al.*²⁵ in the context of China's two-child policy were downscaled to 1-km grids while still keeping the population totals at the original resolution. The two population projections were mosaiced together to create a new population raster for 2030–2100 at 10-year intervals for the five SSPs.

Finally, the global GDP projections were downscaled to a spatial resolution of 1 km using the LitPop approach. First, GDP and GRP projections were adopted and spatially linked to the corresponding administrative boundaries for 2030–2100 under the five SSPs. Second, the NPP-VIIRS DNB products for 2015 were used as the fixed NTL images, and the population projections that were downscaled to a 1-km resolution were used to generate the LitPop images for 2030–2100 under the five SSPs. Then, the GDP and GRP projections were disaggregated to a spatial resolution of 30 arc-seconds (~1 km) for all five SSPs.

In addition, the above spatially explicit GDP data for 2005 and 2030–2100 under the five SSPs were upscaled to 0.25 arc-degrees as an alternative choice.

Data Records

In this research, we produced a set of comparable spatially explicit global GDP estimates, which, to the best of our knowledge, is the first dataset to present substantial long-term changes in global GDP (in 2005 PPP USD) between the historical period (with 2005 as a representative) and 2030–2100 under all five SSPs while accounting for the two-child policy in China³⁷. This global gridded GDP dataset is provided in “GeoTIFF” format with two spatial resolutions: 30 arc-seconds and 0.25 arc-degrees. Global GDP is disaggregated within administrative boundaries, and therefore, Antarctica, oceans and some desert or wilderness areas (nonilluminated and depopulated zones) are assigned a value of 0. The spatial ranges are 60°S–90°N and 180°E–180°W in the standard WGS84 coordinate system. All files are available at <https://doi.org/10.5281/zenodo.5880037>³⁷.

The downscaled GDP for 2005 in 1-km and 0.25 arc-degree cells are shown in Fig. 2, and the GDP projections are shown in Fig. 3 for 2030 under SSP1-5 as a case study. This gridded GDP dataset can broaden the applicability of GDP data to allow for comparisons of global and regional socioeconomic changes between the historical period and future projections under the different SSPs. It can also increase research on scenario-based climate change and across disciplines.

Technical Validation

To examine the performance of the three GDP disaggregation approaches described above, namely, GDP_{Pop} , GDP_{Lit} , and GDP_{LitPop} , we first used national 2005 GDP for 203 countries only and spatialized the data into 1 km × 1 km grids based on each of the three approaches. The GRPs for over 800 provinces (states) in 50 countries and 5231 counties (in the USA and China) were used to evaluate these approaches by assessing their ability to disaggregate national GDP to the subnational level.

The national comparison yields a high degree of accuracy, with all R^2 values approaching 1.0 and very small RMSE values, which is to be expected since the national totals are inputs in each of these three methods. At the provincial (state) and county scales, the performance of GDP_{LitPop} is superior to that of GDP_{Pop} and GDP_{Lit} , with clear advantages in terms of slope, R^2 and RMSE. Specifically, GDP_{LitPop} is more accurate in identifying the spatial allocation of GDP at a fine scale, with R^2 values reaching 0.97, 0.90, and 0.80 and RMSE values of 24.28 billion, 4.67 billion, and 7.44 billion at the provincial (state) scale across the globe (Fig. 4a) and at the county scale for China (Fig. 4b) and the U.S. (Fig. 4c), respectively. However, the R^2 values are only 0.75, 0.62, and 0.46, and RMSE values reach as high as 69.55 billion, 7.12 billion, and 13.88 billion, respectively, when using the GDP_{Lit} approach, exhibiting relatively large bias due to the saturation problem in the DMSP-OLS NTL images. In addition, GDP_{Pop} also identifies GDP redistributions at a fine scale well, with R^2 and RMSE values close to but less accurate than those of GDP_{LitPop} . It should be mentioned that the GRP values are in current prices, which would lead to additional bias if used directly or converted roughly to 2005 PPP USD.

The NTL images from NPP-VIIRS instead of those from DMSP-OLS in 2015 and the LandScan population dataset are used for the disaggregation of future GDP projections. To assess the performance of each approach in GDP downscaling, GRP figures for 737 provinces (states) and 2270 counties in China and 2996 counties in the U.S. from 2015 were obtained and used as another case for technical validation. The national GDP in 2015 in 2011 PPP international dollars for 199 countries from the WB-WDI were used for the GDP downscaling and then reaggregated into gridded GDP estimates at the provincial and county scales.

Similar results to those for 2005 can be obtained for 2015 (Fig. 5). The comparisons show that GDP_{LitPop} outperforms the other methods with slightly higher R^2 values of 0.83, 0.85, and 0.75 at the provincial (Fig. 5a), and county scales in China (Fig. 5b) and the U.S. (Fig. 5c). The RMSE values are thus the smallest at 43.92 billion, 16.92 billion, and 9.49 billion when compared with those for GDP_{Lit} (58.00 billion, 24.46 billion, and 13.97

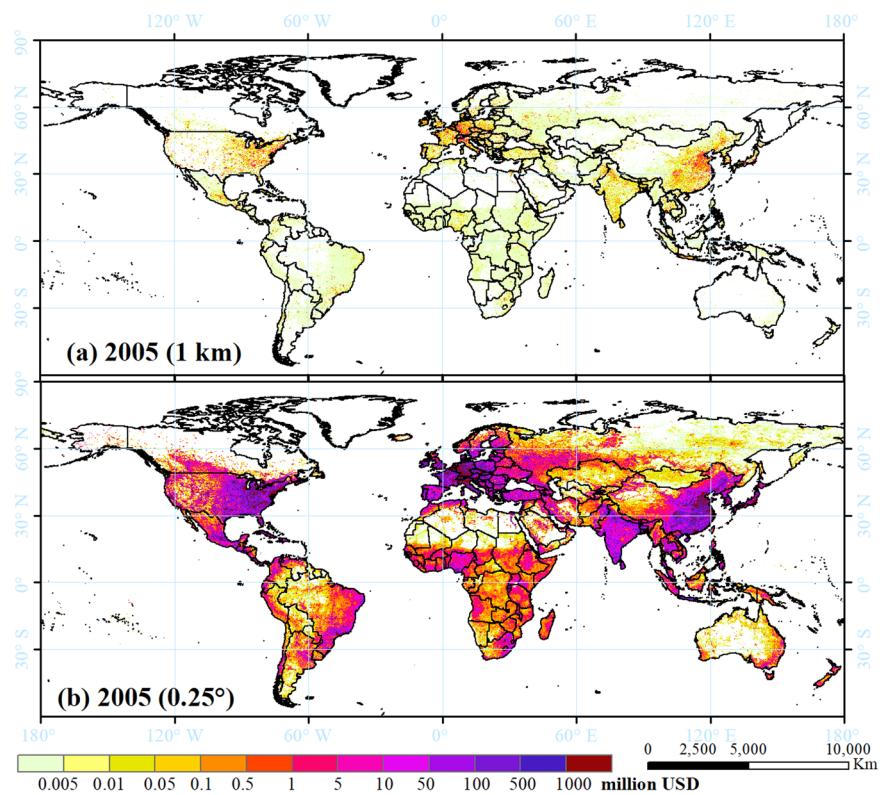


Fig. 2 Spatial distribution of GDP in 2005 at spatial resolutions of 1 km (a) and 0.25 arc-degrees (b). The units are 2005 PPP USD per cell.

billion) and GDP_{Pop} (47.12 billion, 28.76 billion, and 10.63 billion). This not only shows the superiority of the LitPop approach for GDP downscaling but also ensures the availability of gridded GDP projections by including the NPP-VIIRS NTL images from 2015 when building the LitPop base map for future SSPs.

We further assessed the performance of our gridded GDP estimates for 2005 by incorporating over 800 provincial (state) values into the downscaling procedure (Fig. 6). Estimates were obtained from the LitPop approach, the NTL-based approach, and the population-based disaggregation method. The performance of $\text{GDP}_{\text{LitPop}}$ exceeded that of other two, with higher R^2 values of 0.95 and 0.97 for China and the U.S., respectively. In addition, the RMSE values were smaller of 4.62 billion and 7.04 billion in China and the U.S., compared with the values of 5.41 billion and 12.40 billion for GDP_{Lit} and 4.98 billion and 7.48 billion for GDP_{Pop} (Fig. 6). This again shows that $\text{GDP}_{\text{LitPop}}$ better identifies GDP redistributions at a fine scale. Moreover, the gridded GDP estimates at the provincial (state) scales are clearly excellent, with all R^2 values approximately equal to 1.0, and the RMSE values decrease drastically to 0.14 billion, 0.32 billion, and 0.13 billion USD, respectively, indicating high accuracy when these provincial (state) GRP data are incorporated into the GDP downscaling. Additionally, the accuracy of the gridded GDP estimates at the county scale in China (Figs. 4b, 6a) and the U.S. (Figs. 4c, 6b) is also improved with much higher R^2 values and smaller RMSE values for all three downscaling methods. This shows the high accuracy and credibility of both the LitPop approach and our gridded GDP dataset.

We also compared and assessed the performance of the gridded GDP datasets from Murakami and Yamagata²⁴, which include GDP estimates for 1980–2100 at 10-year intervals, through two scenarios using the LitPop approach for 2010 (Fig. 7). National GDP data from the WB-WDI were obtained and downscaled into 1-km grids using the LitPop approach. LandScan population data for 2010 and NTL images for 2010 and 2005 were used to generate two LitPop images: one with 2010 population and NTL images as a historical case (denoted $\text{GDP}_{\text{L10P10}}$), and another with 2010 population and 2005 NTL images (denoted $\text{GDP}_{\text{L05P10}}$) as a specious future scenario since fixed NTL images were used for the SSPs. Comparisons between GRP and gridded GDP in 304 provinces (states) and 1996 counties with valid figures in each gridded dataset for the U.S. (Fig. 7b) show the superiority of our gridded GDP using the LitPop approach compared with that of Murakami and Yamagata²⁴. The R^2 values are much higher (over 0.95 vs. 0.56 at the county scale), and the RMSE values are approximately half as large (less than 50 billion vs. 106 billion USD at the provincial scale and approximately 8 billion vs. 16.2 billion at the county scale in the U.S.) (Fig. 7). This comparison also indicates that using NTL images from the historical period can induce additional bias in the future scenarios, and societal development could intensify the spatial differences across grids.

Above all, the LitPop approach has clear advantages for GDP disaggregation, especially at finer spatial scales, when compared with the population-based or NTL-based approaches. This was in line with many studies on overcoming the limitations of using either gridded population estimates based on the assumption of a uniform GDP per capita distribution across a given administrative unit or DMSP-OLS NTL images, which are subject

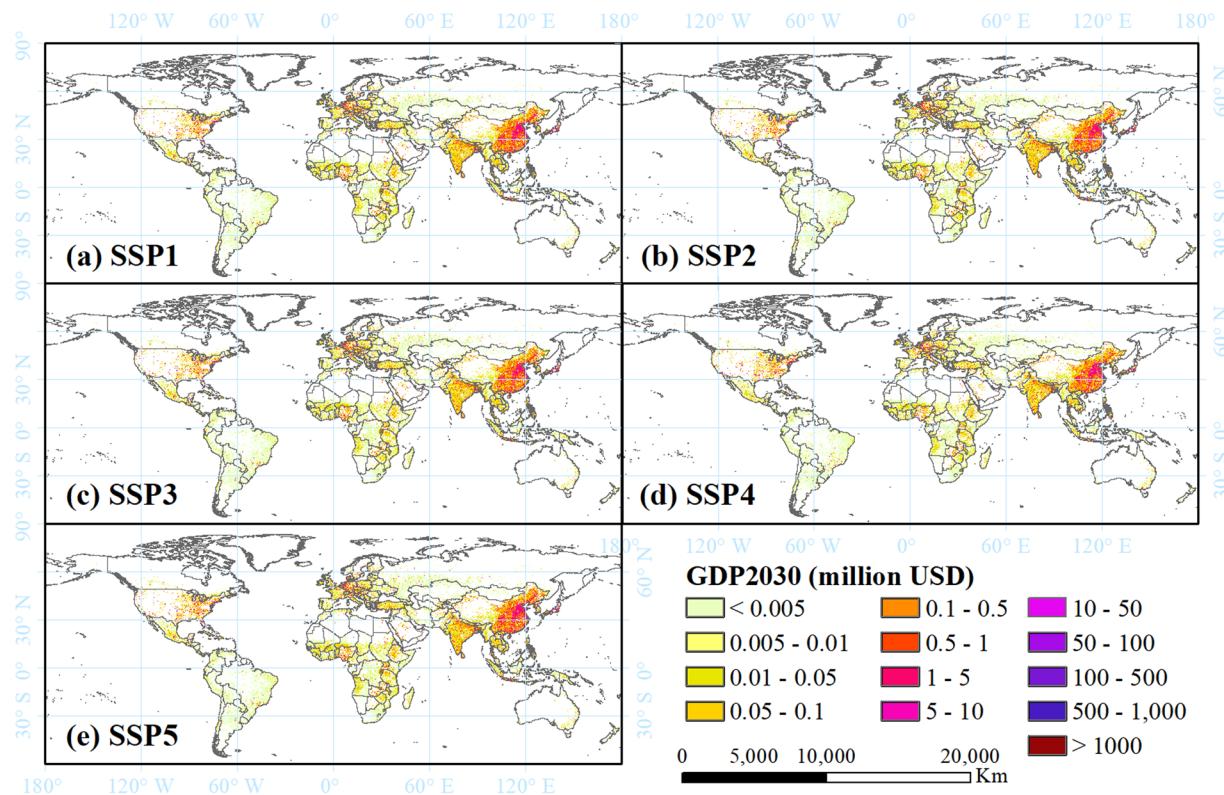


Fig. 3 Spatial distribution of global GDP in 2030 under scenarios SSP1–5 (a–e) at a spatial resolution of 1 km. The units are 2005 PPP USD per cell.

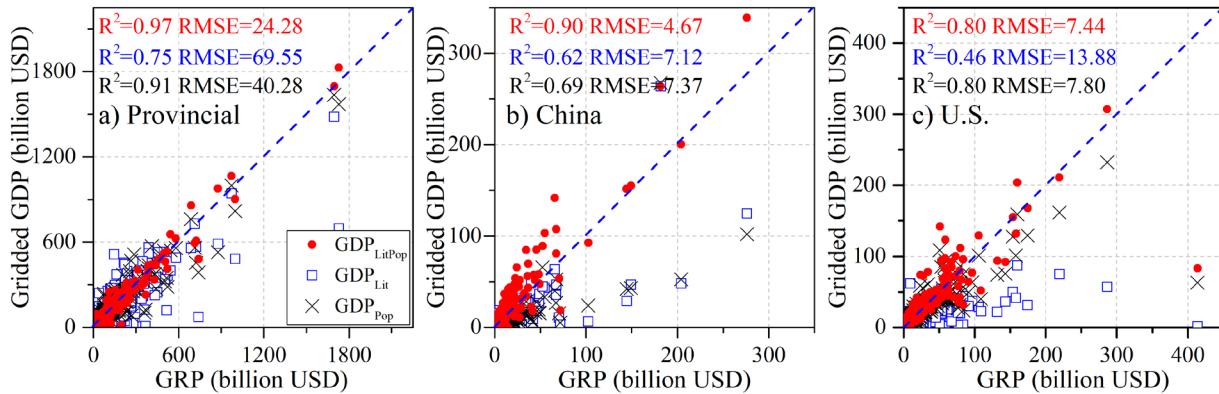


Fig. 4 Evaluation of GRP and gridded GDP in 2005 (national GDP-based downscaling) at the provincial (state) (a) and county scales in China (b) and the U.S. (c) using the population-based disaggregation method, the NTL-based approach, and the LitPop approach.

to saturation problems^{8,17,18,38}. Our gridded GDP dataset is recommended for use across disciplines due to its relatively high accuracy and credibility.

Usage Notes

We adopted the LitPop approach and presented a first set of comparable spatially explicit global gridded GDP estimates with a high spatial resolution of 30 arc-seconds (approximately 1 km at the equator) as well as of 0.25 arc-degrees for 2005 and for 2030–2100 (in 2005 PPP USD) at 10-year intervals for all five SSPs, while accounting for the two-child policy in China. This gridded GDP projection dataset can broaden the ability of GDP data to account for the reality of global and regional socioeconomic changes between the historical period and future projections under the different SSPs. It can also increase research into scenario-based climate change.

To be clear, this GDP dataset is bound to input data from various sources and to the approaches used, including the use of a uniform GDP growth rate within administrative boundaries, the use of NPP-VIIRS data from 2015 as a fixed NTL image, and the use of the population distribution patterns from LandScan as a base map

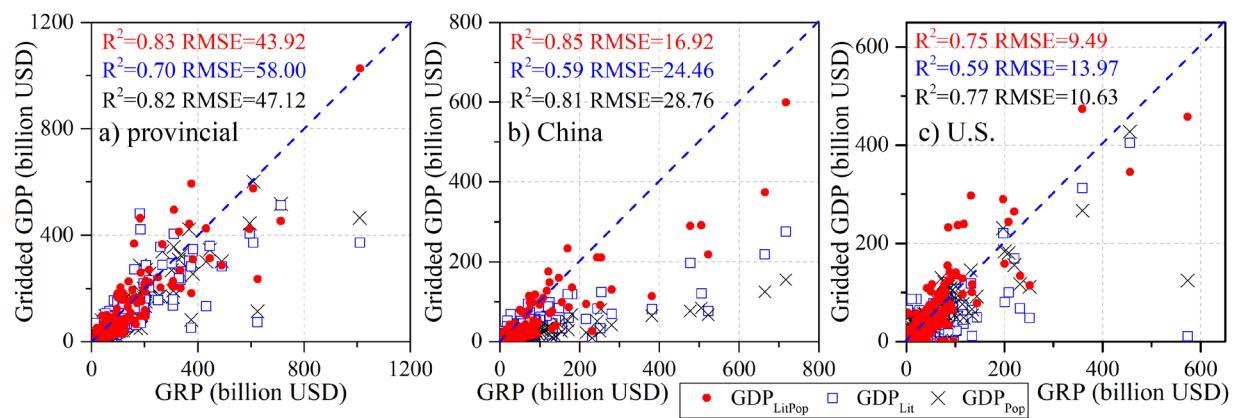


Fig. 5 Evaluation of GRP and gridded GDP in 2015 (national GDP-based downscaling with data from the WB-WDI) at the provincial (state) (**a**) and county scales in China (**b**) and the U.S. (**c**), using the population-based disaggregation method, the NTL-based approach, and the LitPop approach.

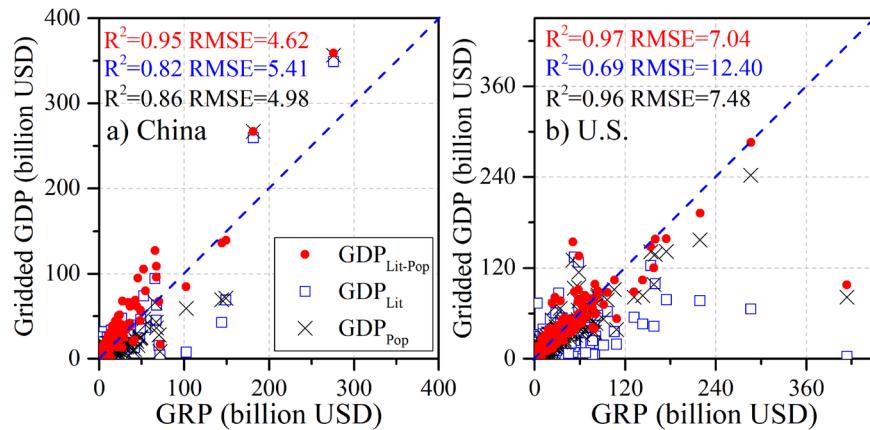


Fig. 6 Evaluation of GRP and gridded GDP in 2005 (provincial GRP-based downscaling) at the county scale in China (**a**) and the U.S. (**b**), using the population-based disaggregation method, the NTL-based approach, and the LitPop approach.

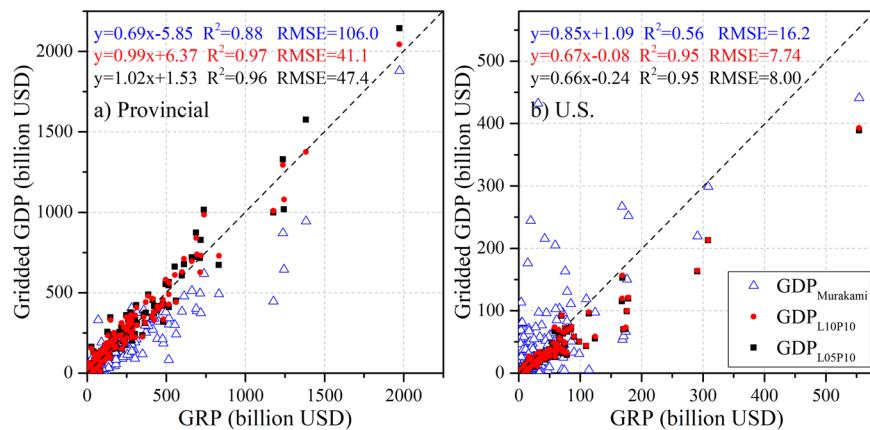


Fig. 7 Evaluation of the GRP and gridded GDP from Murakami and Yamagata²⁴ (denoted $GDP_{Murakami}$) and national GDP disaggregation using the LitPop approach (two LitPop images: the NTL image in 2005 and the Landscan population in 2010 (denoted GDP_{L05P10}) and the NTL image in 2010 and population in 2010 (denoted GDP_{L10P10}) at the provincial (state) (**a**) and at the county scale in the U.S. (**b**) in 2010.

for future population downscaling. Moreover, the disaggregated GDP values may be underestimated in some underdeveloped areas where nonilluminated cells or unpopulated zones are located when using the LitPop approach. Therefore, these GDP projections fail to capture the spatial variation caused by migration and urban development in the future scenarios.

In addition, the GDP projections with a 1-km resolution are produced based on the assumption that there is no population mobility across grids when disaggregating the population projections from 1/8 degree¹⁹ to 1-km grids and then using them to build a future LitPop base map. In addition, the national and supranational GDP projections from the SSP database contain large amounts of uncertainty, as explained in Riahi *et al.*²¹, and should be treated with caution. The effects of financial crises, policy changes, technological progress, political and societal factors, climate system feedback, etc., are also not considered in this gridded GDP dataset.

Code availability

No specific code was generated for analysis of these data.

Received: 12 August 2021; Accepted: 25 March 2022;

Published online: 19 May 2022

References

- O'Neill, B. C. *et al.* A new scenario framework for climate change research: the concept of shared socioeconomic pathways. *Climatic change* **122**, 387–400, <https://doi.org/10.1007/s10584-013-0905-2> (2014).
- IPCC. *Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change*. (Cambridge University Press, 2014).
- Mitchell, D. *et al.* Realizing the impacts of a 1.5 C warmer world. *Nature Climate Change* **6**, 735–737 (2016).
- O'Neill, B. C. *et al.* The scenario model intercomparison project (ScenarioMIP) for CMIP6. *Geoscientific Model Development* **9**, 3461–3482, <https://doi.org/10.5194/gmd-9-3461-2016> (2016).
- Wilbanks, T. J. & Ebi, K. L. SSPs from an impact and adaptation perspective. *Climatic change* **122**, 473–479, <https://doi.org/10.1007/s10584-013-0903-4> (2014).
- Geiger, T. Continuous national gross domestic product (GDP) time series for 195 countries: past observations (1850–2005) harmonized with future projections according to the Shared Socio-economic Pathways (2006–2100). *Earth System Science Data* **10**, 847–856, <https://doi.org/10.5194/essd-10-847-2018> (2018).
- Kummu, M., Taka, M. & Guillaume, J. H. Gridded global datasets for gross domestic product and Human Development Index over 1990–2015. *Scientific data* **5**, 180004, <https://doi.org/10.1038/sdata.2018.4> (2018).
- Nordhaus, C. W. D. Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 8589–8594, <https://doi.org/10.1073/pnas.1017031108> (2011).
- Geiger, T., Daisuke, M., Frieler, K. & Yamagata, Y. 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> (2017).
- Smits, J. & Permanyer, I. The Subnational Human Development Database. *Scientific Data* **6**, 190038, <https://doi.org/10.1038/sdata.2019.38> (2019).
- Chen, Q., Ye, T., Zhao, N., Ding, M. & Yang, X. Mapping China's regional economic activity by integrating points-of-interest and remote sensing data with random forest. *Environment and Planning B Urban Analytics and City science*, 1–19, <https://doi.org/10.1177/2399808320951580> (2020).
- Bennett, M. M. & Smith, L. C. Advances in using multitemporal night-time lights satellite imagery to detect, estimate, and monitor socioeconomic dynamics. *Remote Sensing of Environment* **192**, 176–197, <https://doi.org/10.1016/j.rse.2017.01.005> (2017).
- Doll, C. N. H., Muller, J. P. & Morley, J. G. Mapping regional economic activity from night-time light satellite imagery. *Ecological Economics* **57**, 75–92, <https://doi.org/10.1016/j.ecolecon.2005.03.007> (2006).
- Ghosh, T. *et al.* Shedding light on the global distribution of economic activity. *The Open Geography Journal* **3**, 147–160 (2010).
- Wang, X. *et al.* Estimation and mapping of sub-national GDP in Uganda using NPP-VIIRS imagery. *Remote Sensing* **11**, <https://doi.org/10.3390/rs11020163> (2019).
- Zhu, X., Ma, M., Yang, H. & Ge, W. Modeling the Spatiotemporal Dynamics of Gross Domestic Product in China Using Extended Temporal Coverage Nighttime Light Data. *Remote Sensing* **9**, 626, <https://doi.org/10.3390/rs9060626> (2017).
- Eberenz, S., Stocker, D., Rslı, T. & Bresch, D. N. Asset exposure data for global physical risk assessment. *Earth System Science Data* **12**, 817–833, <https://doi.org/10.5194/essd-12-817-2020> (2020).
- Zhao, N., Liu, Y., Cao, G., Samson, E. L. & Zhang, J. Forecasting China's GDP at the pixel level using nighttime lights time series and population images. *Mapping Environ & Remote Sensing* **54**, 407–425, <https://doi.org/10.1080/15481603.2016.1276705> (2017).
- Jones, B. & O'Neill, B. C. Spatially explicit global population scenarios consistent with the Shared Socioeconomic Pathways. *Environmental Research Letters* **11**, 084003, <https://doi.org/10.1088/1748-9326/11/8/084003> (2016).
- Su, B. *et al.* Drought losses in China might double between the 1.5 C and 2.0 C warming. *Proceedings of the National Academy of Sciences* **115**, 10600–10605, <https://doi.org/10.1073/pnas.1802129115> (2018).
- Riahi, K. *et al.* The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: an overview. *Global Environmental Change* **42**, 153–168, <https://doi.org/10.1016/j.gloenvcha.2016.05.009> (2017).
- Tacchella, A., Mazzilli, D. & Pietronero, L. A dynamical systems approach to gross domestic product forecasting. *Nature Physics* **14**, 861–865, <https://doi.org/10.1038/s41567-018-0204-y> (2018).
- Gao, J. & Pesaresi, M. Downscaling SSP-consistent global spatial urban land projections from 1/8-degree to 1-km resolution 2000–2100. *Scientific Data* **8**, 1–9, <https://doi.org/10.1038/s41597-021-01052-0> (2021).
- Murakami, D. & Yamagata, Y. Estimation of gridded population and GDP scenarios with spatially explicit statistical downscaling. *Sustainability* **11**, 1–18, <https://doi.org/10.3390/su11072106> (2019).
- Jiang, T. *et al.* National and Provincial Population Projected to 2100 Under the Shared Socioeconomic Pathways in China. *Climate Change Research* **13**, 128–137, <https://doi.org/10.12006/j.issn.1673-1719.2016.249> (2017).
- Jiang, T. *et al.* Projection of national and provincial economy under the shared socioeconomic pathways in China. *Climate Change Research* **14**, 50–58, <https://doi.org/10.12006/j.issn.1673-1719.2017.161> (2018).
- Huang, J. *et al.* Effect of Fertility Policy Changes on the Population Structure and Economy of China: from the Perspective of the Shared Socioeconomic Pathways. *Earth's Future* **7**, 250–265, <https://doi.org/10.1029/2018EF000964> (2019).
- Center for International Earth Science Information Network - CIESIN - Columbia University. (NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, NY, 2018).
- Stevens, F. R., Gaughan, A. E., Linard, C. & Tatem, A. J. Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. *PloS one* **10**, <https://doi.org/10.1371/journal.pone.0107042> (2015).

30. Tatem, A. J. WorldPop, open data for spatial demography. *Scientific data* **4**, 1–4, <https://doi.org/10.1038/sdata.2017.4> (2017).
31. Bhaduri, B., Bright, E., Coleman, P. & Urban, M. L. LandScan USA: a high-resolution geospatial and temporal modeling approach for population distribution and dynamics. *GeoJournal* **69**, 103–117, <https://doi.org/10.1007/s10708-007-9105-9> (2007).
32. Cuaresma, J. C. Income Projections for Climate Change Research: A Framework Based on Human Capital Dynamics. *Global Environmental Change* **42**, 226–236, <https://doi.org/10.1016/j.gloenvcha.2015.02.012> (2015).
33. Dellink, R., Chateau, J., Lanzi, E. & Magné, B. Long-term economic growth projections in the Shared Socioeconomic Pathways. *Global Environmental Change* **42**, 200–214, <https://doi.org/10.1016/j.gloenvcha.2015.06.004> (2017).
34. Leimbach, M., Kriegler, E., Roming, N. & Schwanitz, J. Future growth patterns of world regions – A GDP scenario approach. *Global Environmental Change* **42**, 215–225, <https://doi.org/10.1016/j.gloenvcha.2015.02.005> (2017).
35. KC, S. & Lutz, W. The human core of the shared socioeconomic pathways: Population scenarios by age, sex and level of education for all countries to 2100. *Global Environmental Change* **42**, 181–192, <https://doi.org/10.1016/j.gloenvcha.2014.06.004> (2017).
36. Shiogama, H. *et al.* Observational constraints indicate risk of drying in the Amazon basin. *Nature Communications* **2**, 1–7, <https://doi.org/10.1038/ncomms1252> (2011).
37. Wang, T. & Sun, F. Gross domestic product (GDP) downscaling: a global gridded dataset consistent with the Shared Socioeconomic Pathways. *Zenodo* <https://doi.org/10.5281/zenodo.5880037> (2022).
38. Dai, Z., Hu, Y. & Zhao, G. The suitability of different nighttime light data for GDP estimation at different spatial scales and regional levels. *Sustainability* **9**, 305, <https://doi.org/10.3390/su9020305> (2017).

Acknowledgements

This research was funded by the National Natural Sciences Foundation of China (42025104, 42001031), the National Key Research and Development Program of China (2019YFA0606903), the TopNotch Young Talents Program of China (Fubao Sun), China Postdoctoral Science Foundation funded project (2018M640173, 2020T130646).

Author contributions

T.W. and F.S. designed the research, and T.W. performed the analysis and drafted the manuscript; F.S. provided insights on data product characteristics and underlying procedures.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-022-01300-x>.

Correspondence and requests for materials should be addressed to F.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022