Journal of Global Change Data & Discovery. 2024, 8(1): 90-97
www.geodoi.ac.cn

Global Change Research Data Publishing & Repository

Global Urban Expansion Simulation Dataset (1992–2050)

Liu, Z. F.[1, 2] Ying, J. H.[1, 2] He, C. Y.[1,3,4*] Huang, Q. X.[1, 2] Bai, Q. X.[1, 2] Pan. X. H.[1, 2]

1. State Key Laboratory of Earth Surface Processes and Resource Ecology (ESPRE), Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China;
2. School of Natural Resources, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China;
3. Key Laboratory of Environmental Change and Natural Disasters of Chinese Ministry of Education, Beijing Normal University, Beijing 100875, China;
4. Academy of Disaster Reduction and Emergency Management, Ministry of Emergency Management and Ministry of Education, Beijing 100875, China

Abstract: Global urban expansion data is the foundation for understanding the process and effects of global urban expansion and optimizing it accordingly, but the existing data do not effectively distinguish between urban and rural construction land, which leads to greater uncertainty in the results of the relevant research. In this paper, we reconstructed the global urban expansion year by year from 1992 to 2020 using the global built-up area data and the global urban center location data, and simulated the global urban expansion from 2020 to 2050 under the SSPs using Land Use Scenario Dynamics-urban (LUSD-urban) model, so as to develop a dataset of global urban expansion that effectively distinguishes between urban and rural construction land and is continuously comparable. The accuracy evaluation shows that the dataset is accurate and reliable, with a Kappa coefficient of 0.88 and a FoM of 0.23. This dataset consists of the annual global historical urban built-up area from 1992 to 2020 and the global future urban built-up area for every five years from 2021 to 2050. The spatial resolution of the dataset is 1 km, and it is archived in the format of .tif. The data size is 498 MB, and 23.7 MB after compression.

Keywords: urban expansion; global; 1992–2050; SSPs (shared socio-economic pathways); urbanization

DOI: https://doi.org/10.3974/geodp.2024.01.11

CSTR: https://cstr.escience.org.cn/CSTR:20146.14.2024.01.11

Dataset Availability Statement:

The dataset supporting this paper was published and is accessible through the *Digital Journal of Global Change Data Repository* at: https://doi.org/10.3974/geodb.2024.06.05.V1 or https://cstr.escience.org.cn/CSTR:20146.11.2024.06.05.V1.

Received: 15-01-2024; **Accepted:** 10-03-2024; **Published:** 25-03-2024

Foundation: Ministry of Science and Technology of P. R. China (2019YFA0607203)

***Corresponding Author:** He,,C. Y., Ministry of Emergency Management-Ministry of Education Institute of Disaster-Reduction and Emergency Management, Beijing Normal University, hcy@bnu.edu.cn

Data Citation: [1] Liu, Z. F., Ying, J. H., He, C. Y., *et al*. Global urban expansion simulation dataset (1992–2050) [J]. *Journal of Global Change Data & Discovery*, 2024, 8(1): 90–97. https://doi.org/10.3974/geodp.2024.01.11. https://cstr.escience.org.cn/CSTR:20146.14.2024.01.11.

[2] Liu, Z. F., Ying, J. H., He, C. Y., *et al*. Global urban expansion simulation dataset (1992–2050, V1.0) [J/DB/OL]. *Digital Journal of Global Change Data Repository*, 2024. https://doi.org/10.3974/geodb.2024.06.05.V1. https://cstr.escience.org.cn/CSTR:20146.11.2024.06.05.V1.

1 Introduction

The global urban expansion profoundly impacts the environment, society, and economy[1,2]. Global urban expansion datasets are crucial foundation for revealing the spatial and temporal patterns and the driving mechanisms of global urban expansion, evaluating the impacts and risks of global urban expansion, and optimizing this process. Currently, researchers have constructed several datasets of global urban expansion, providing rich data support for relevant studies[3–5]. However, most existing datasets do not effectively distinguish between urban built-up land and rural construction land, which leads to large uncertainty in the results of related research. Although He *et al*. have utilized deep learning methods to composite nighttime lighting data, vegetation index data, and surface temperature data, creating a global historical urban expansion dataset that can effectively distinguish between urban and rural construction land, and further produced continuously comparable global future urban expansion data[6,7], this dataset lacks historical year-by-year urban expansion information and urban expansion information post-2016. Additionally, they calculated the suitability of urban expansion based on global-scale data in the process of simulating future urban expansion, which does not fully account for inter-regional differences. To address these issues, we reconstructed the global urban expansion data on an annual basis from 1992 to 2020 by compositing multi-source data, and simulated the global urban expansion from 2020 to 2050 under the five Shared Socioeconomic Pathways (SSPs), producing a dataset of global urban expansion that effectively differentiates between urban and rural construction land and provides continuous comparability.

2 Metadata of the Dataset

The metadata of the global urban expansion simulation dataset (1992–2050, V1.0)[8] is summarized in Table 1. It includes the dataset full name, short name, authors, year of the dataset, temporal resolution, spatial resolution, data format, data size, data files, data publisher, and data sharing policy, etc.

3 Methods

3.1 Data Sources

The global built-up area data and urban center location data used to obtain the global historical urban expansion data were obtained from the Climate Change Initiative (CCI) land cover products published by the European Space Agency (ESA)[10] and the Global Human Settlement (GHS) Urban Center Database (UCDB)[11] published by the European Commission Joint Research Center (JRC)[11], respectively. Historical and future urban population data for projecting the global future urban built-up area were obtained from the World Urbanization Prospects (WUP) dataset published by the United Nations[12] and the WUP-SSPs Population and Urbanization Rates dataset[13]. We obtained digital elevation model (DEM) from the global multi-resolution topographic elevation dataset[14], soil data from the Harmonized World Soil Database (HWSD)[15], geographic information auxiliary data including coastline data from the National Oceanic and Atmospheric Administration of the United States[16], river data from the Global Rivers Network dataset[17], road and railroad data from the Resource and Environmental Data Cloud Platform (REDCP)[1], and meteorological data from the National Climate Information Center (NCIC)[18]. All the raster data were resampled to 1-km spatial resolution.

Table 1 Metadata summary of global urban expansion simulation dataset (1992–2050)

Items	Description
Dataset full name	Global urban expansion simulation dataset (1992–2050, V1.0)
Dataset short name	GlobalUrbanExpansion1992-2050_1.0
Authors	Liu, Z. F., Beijing Normal University, zhifeng.liu@bnu.edu.cn Ying, J. H., Beijing Normal University, jiahe.ying@mail.bnu.edu.cn He, C. Y., Beijing Normal University, hcy@bnu.edu.cn Huang, Q. X., Beijing Normal University, qxhuang@bnu.edu.cn Bai, Q. X., Beijing Normal University, qx_bai@163.com Pan, X. H., Beijing Normal University, xinhao.pan@mail.bnu.edu.cn
Geographical region	Global
Year	1992–2050
Temporal resolution	1-year (1992–2020), 5-year (2020–2050)
Spatial resolution	1 km
Data format	.tif
Data size	23.7 MB (compressed)
Data files	A total of 59 raster data files containing year-by-year global historical built-up area for 1992–2020, and five-year-by-five-year global future built-up area for 2021–2050 under SSP1–SSP5 scenarios
Foundation	Ministry of Science and Technology of P. R. China (2019YFA0607203)
Data publisher	Global Change Research Data Publishing & Repository, http://www.geodoi.ac.cn
Address	No. 11A, Datun Road, Chaoyang District, Beijing 100101, China
Data sharing policy	(1) *Data* are openly available and can be free downloaded via the Internet; (2) End users are encouraged to use *Data* subject to citation; (3) Users, who are by definition also value-added service providers, are welcome to redistribute *Data* subject to written permission from the GCdataPR Editorial Office and the issuance of a *Data* redistribution license; and (4) If *Data* are used to compile new datasets, the 'ten per cent principal' should be followed such that *Data* records utilized should not surpass 10% of the new dataset contents, while sources should be clearly noted in suitable places in the new dataset[9]
Communication and searchable system	DOI, CSTR, Crossref, DCI, CSCD, CNKI, SciEngine, WDS/ISC, GEOSS

3.2 Algorithm

(1) Reconstruction of historical global urban expansion

In order to identify the urban built-up area from the global built-up area data, the raster data of global built-up area was first converted into vector data, and the adjacent built-up patches were merged based on the Moore's Neighborhood (8-neighborhood) rule to obtain the Moore's Neighborhood-based global built-up patches. Then, spatial analysis was performed on the adjusted global built-up patches and urban center point data, and the built-up patches intersecting with the urban center points were screened out and treated as urban built-up patches. Finally, the urban built-up patch data was converted to raster data to obtain urban built-up area with a spatial resolution of 1 km.

(2) Simulation of future global urban expansion

We adopted the Land Use Scenario Dynamics-urban (LUSD-urban) model developed by He *et al*. to simulate the global future urban expansion based on the zonal simulation approach[7, 19]. Firstly, taking continent as the basic unit, we utilized the adaptive Monte Carlo method in the LUSD-urban model to obtain the weights of each suitability and

[1] Resource and Environmental Data Cloud Platform (REDCP). Global road database [OL/DB]. http://www.resdc.cn/data.aspx?DATAID=207.

restriction layer and the probability of each non-urban pixel transitioning to an urban pixel. Then, taking the country as the basic unit, we constructed linear regression equations based on the historical urban built-up area and urban population, and revised the equations to ensure that the built-up area obtained from the regression based on the urban population data is consistent with the actual built-up area in 2020[20], and then predicted the future urban built-up area of each country by using the revised equations and the future urban population data. Finally, urban pixels were spatially allocated based on the probability of each non-urban pixel transforming into an urban pixel, to meet the future demand for urban built-up area in each country.

3.3 Technical Approach

Based on the above methods, the global urban built-up area from 1992 to 2020 was first identified year by year using the global built-up area data and the global urban center location data (Figure 1). Then, by further combining the global urban population data of each country from 1992 to 2020 and the global urban population data of each country under the SSPs from 2020 to 2050, the global built-up area of each country under the SSPs was projected (Figure 1). Finally, the LUSD-urban model was calibrated and validated using the global urban expansion data from 1992 to 2020, and the global urban expansion under the SSPs was simulated using the calibrated model (Figure 1).

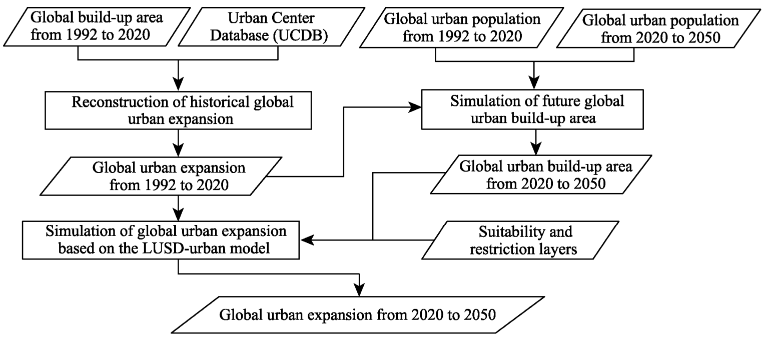

Figure 1 Flow chart of the dataset development

4 Data Results and Validation

4.1 Data Composition

This dataset consists of the annual global historical urban built-up area from 1992 to 2020 and the global future urban built-up area for every five years from 2021 to 2050 under the SSP1–SSP 5. The spatial resolution of the dataset is 1 km. The data size is 498 MB, and 23.8 MB after compression. The name of the data is explained as follows: GlobalUrban**** represents the global historical built-up area of the year ****; GlobalUrban20**_SSP* represents the global built-up area of the year 20** under the SSP*, SSP1 is the sustainability, SSP2 is the middle of the road, SSP3 is the regional rivalry, SSP4 is the inequality, and SSP5 is the fossil-fueled development.

4.2 Data Analysis

The world experienced a rapid process of urban expansion from 1992 to 2020, and this process is projected to continue from 2020 to 2050 (Figure 2). The global urban area increased from 229.8 thousand km^2 in 1992 to 486.7 thousand km^2 in 2020, a growth of 1.12

times. By 2050, the global urban area is projected to increase to 633.9–762.6 thousand km^2, which is 30.24%–56.69% higher than in 2020. Under the SSP5, the global urban expansion area is the largest, reaching 275.9 thousand km^2; under the SSP3, the global urban expansion area is the smallest, at 147.2 thousand km^2.

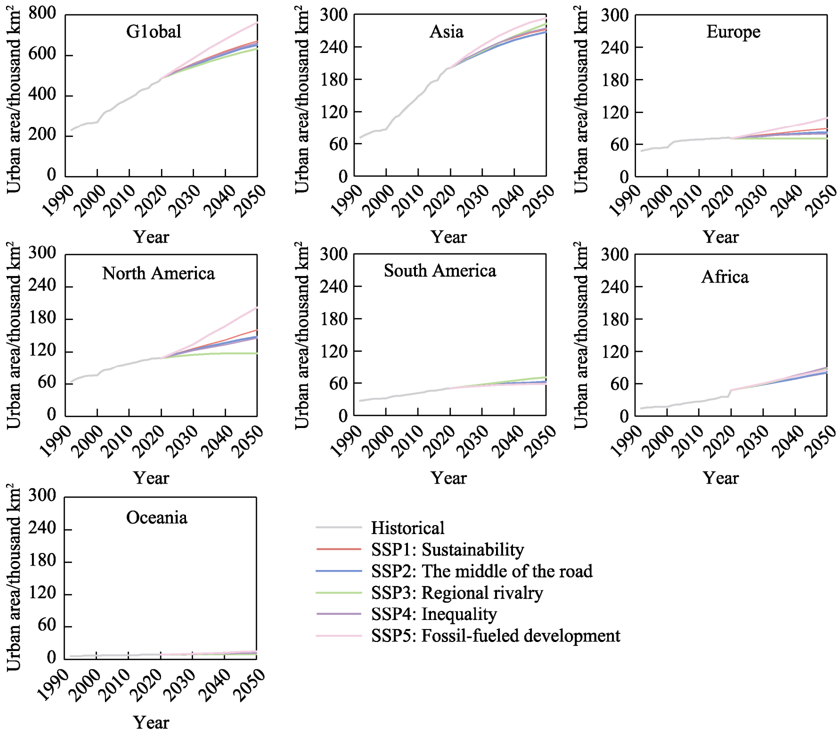

Figure 2 Maps of global and continental urban expansion

There are obvious differences in the process of urban expansion in different continents (Figure 2). From 1992 to 2020, Asia and North America had larger urban expansion area, with 127.6 thousand km^2 and 44.0 thousand km^2, respectively; Europe, South America, and Africa had the next largest urban expansion area, with 24.8 thousand km^2, 22.8 thousand km^2, and 20.6 thousand km^2, respectively; and Oceania had a smaller urban expansion area, with 2.6 thousand km^2. From 2020 to 2050, Asia and Africa are projected to have larger area of urban expansion, ranging from 65.9 thousand km^2 to 91.8 thousand km^2 and from 32.0 thousand km^2 to 42.8 thousand km^2, respectively. North America, Europe, and South America are projected to have large differences in urban expansion area among scenarios. The urban area in North America is projected to expand by 93.1 thousand km^2 under the SSP5, but only by 8.1 thousand km^2 under the SSP3. Europe is projected to have a slight increase in urban area under SSP1, SSP2, and SSP4, more urban expansion area under the SSP5, and will face urban shrinking pressures under SSP3. South America's urban area is projected to expand by 20.7 thousand km^2 under SSP3 and only 8.3 thousand km^2 under SSP5. Oceania is projected to have a smaller urban expansion area of 1.4–6.4 thousand km^2.

Differences in the process of urban expansion between countries are more obvious (Figure 3). From 1992 to 2020, China, United States, and India had larger urban expansion area of 67.7 thousand km^2, 41.5 thousand km^2, and 15.6 thousand km^2, respectively; and

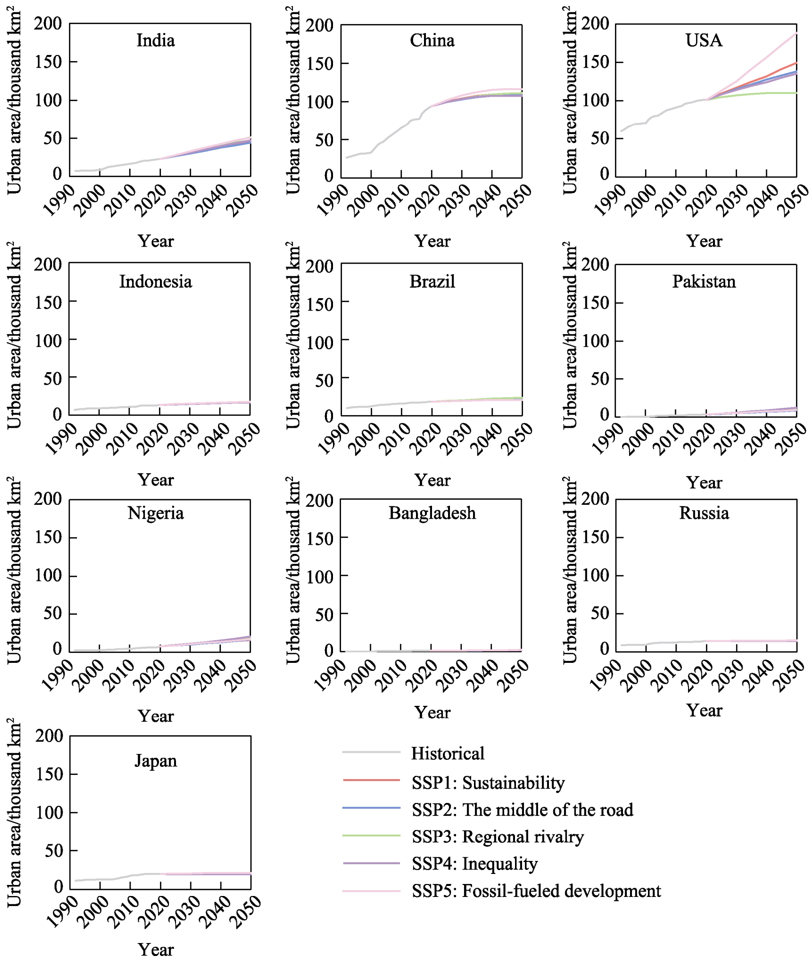

Figure 3 Urban expansion in the top ten most populous countries

Japan, Brazil, and Indonesia had the next largest urban expansion area of 8.8 thousand km^2, 7.8 thousand km^2, and 6.2 thousand km^2, respectively. From 2020 to 2050, India and Nigeria are projected to continue to expand urban area by 21.4–28.2 thousand km^2 and 8.3–12.8 thousand km^2, respectively. In the United States, the trend of urban expansion under different scenarios is projected to be inconsistent, with the maximum urban expansion area reaching 86.9 thousand km^2 and the minimum only 7.6 thousand km^2. China's urban area is projected to increase slightly, with an expansion area of 11.7–21.9 thousand km^2. Japan, Brazil, and Russia are projected to have smaller urban expansion area. The process of urban expansion also varies across regions (Figure 4).

4.3 Data Validation

To validate the LUSD-urban model, we calibrated the model using global urban expansion data from 1992 to 2010 and simulated global urban area in 2020 using the calibrated model. By comparing the simulated result in 2020 and the actual global urban area in 2020, we found that the Kappa coefficient is 0.88 and the Figure of Merit (FoM) is 0.23 at the global scale. The simulation results of the urban area of the top ten most populous countries worldwide have Kappa coefficients ranging from 0.78 to 0.93, and the FoM ranging from 0.13 to 0.34, which indicates that the LUSD-urban model can accurately simulate the global urban expansion (Table 2).

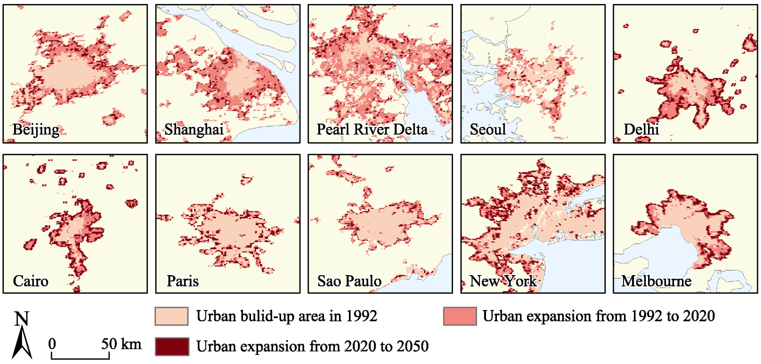

Figure 4 Maps of spatial and temporal patterns of the urban expansion in representative regions (using the SSP2 as an example)

Table 2 Accuracy of simulation results of urban expansion

Country	Kappa	FoM	Country	Kappa	FoM
India	0.82	0.22	Nigeria	0.81	0.34
China	0.81	0.24	Bangladesh	0.78	0.28
USA	0.93	0.15	Russia	0.92	0.13
Indonesia	0.87	0.22	Japan	0.91	0.20
Brazil	0.91	0.13	Global	0.88	0.23
Pakistan	0.79	0.23			

5 Discussion and Conclusion

We constructed a global urban expansion dataset in the period of 1992–2050 by combining the global built-up area data with the global urban center location data and performing the LUSD-urban model. This dataset can effectively distinguish between urban and rural built-up area and is continuously comparable. In addition, this dataset is accurate and reliable, and the accuracy evaluation shows that the Kappa coefficient of the global urban built-up area simulated by the LUSD-urban model is 0.88, and the FoM is 0.23.

Based on this dataset, we found that the global urban built-up area increased by 1.12 times from 1992 to 2020, and is projected to increase by 30.24% to 56.69% from 2020 to 2050. There are obvious differences in the process of urban expansion in different continents and countries. At the continental scale, Asia had the largest urban expansion area in the period of 1992–2020, and is projected to have the largest urban expansion area in the period of 2020–2050. At the national scale, China had the largest urban expansion area in the period of 1992–2020; and the United States is projected to have the largest urban expansion area in the period of 2020–2050. This dataset provides basic data support for understanding the spatial and temporal patterns, driving mechanisms, and environmental, social and economic effects of urban expansion at various scales worldwide.

Author Contributions

He, C. Y., Liu, Z. F., and Huang, Q. X. did the overall design for the development of the dataset. Liu, Z. F., Ying, J. H., and Bai, Q. X. collected and processed the data. He, C. Y., Liu, Z. F., and Pan, X. H. designed the models and algorithms. Ying, J. H. did the data validation. Liu, Z. F., Ying, J. H., and He, C. Y., wrote the data paper.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Seto, K. C., Güneralp, B., Hutyra, L. R. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools [J]. *PNAS*, 2012, 109: 16083–16088.

[2] Wu, J. Urban ecology and sustainability: the state-of-the-science and future directions [J]. *Landscape Urban Planning*, 2014, 125: 209–221.

[3] Klein, G. K., Beusen, A., Doelman, J., *et al*. Anthropogenic land use estimates for the holocene—HYDE 3.2 [J]. *Earth System Scientific Data*, 2017, 9(2): 927–953.

[4] Chen, G. Z., Li, X., Liu, X. P., *et al*. Global projections of future urban land expansion under shared socioeconomic pathways [J]. *Nature Communication*, 2020, 11(1): 12.

[5] Gao, J., O'Neill, B, C. Mapping global urban land for the 21st century with data-driven simulations and Shared Socioeconomic Pathways [J]. *Nature Communication*, 2020, 11: 2302.

[6] He, C., Liu, Z., Gou, S., *et al*. Detecting global urban expansion over the last three decades using a fully convolutional network [J]. *Environmental Research Letters*, 2019, 14: 034008.

[7] He, C., Liu, Z., Wu, J., *et al.* Future global urban water scarcity and potential solutions [J]. *Nature Communications,* 2021, 12: 4667.

[8] Liu, Z., Ying, J., He, C., *et al*. Global Urban Expansion Simulation Dataset (1992–2050, V1.0) [J/DB/OL]. *Digital Journal of Global Change Data Repository*, 2024. https://doi.org/10.3974/geodb.2024.06.05.V1.

[9] GCdataPR Editorial Office. GCdataPR data sharing policy [OL]. https://doi.org/10.3974/dp.policy.2014.05 (Updated 2017).

[10] ESA. Land cover CCI product user guide version 2.0 [OL/DB]. http://maps.elie.ucl.ac.be/CCI/viewer/index.php.

[11] Florczyk, A., Corbane C., Schiavina, M., *et al*. GHS-UCDB R2019A—GHS Urban Centre Database 2015, multitemporal and multidimensional attributes [Z]. European Commission, Joint Research Centre (JRC), 2019. https://doi.org/10.2905/53473144-b88c-44bc-b4a3-4583ed1f547e.

[12] Department of Economic and Social Affairs Population Dynamics. 2018 revision of world urbanization prospects [Z]. United Nations, 2018. https://population.un.org/wup.

[13] Chen, S., Huang, Q., Muttarak, R. *et al*. Updating global urbanization projections under the Shared Socioeconomic Pathways [J]. *Scientific Data*, 2022, 9: 137. https://doi.org/10.1038/s41597-022-01209-5.

[14] Jeffrey, J. D., Dean, B. G., Global multi-resolution terrain elevation data 2010 [Z]. Earth Resources Observation and Science (EROS) Center, 2018. https://lta.cr.usgs.gov/GMTED2010.

[15] Wieder, W. R., Boehnert, J., Bonan, G. B., *et al*. Regridded harmonized world soil database v1.2 [Z]. ORNL DAAC, Oak Ridge, Tennessee, USA.

[16] Global Self-consistent, Hierarchical, High-resolution Geography Database [Z]. National Centers for Environmental Information, 2017. https://www.ngdc.noaa.gov/mgg/shorelines/

[17] Schneider, A., Jost, A., Coulon, C., *et al*. Global scale river network extraction based on high-resolution topography and constrained by lithology, climate, slope, and observed drainage density [J]. *Geophysical Research Letter*, 2017, 44(6): 2773–2781. DOI: 10.1002/2016GL071844.

[18] National Climate Information Center (NCIC). Global surface meteorological observations historical database [OL/DB]. http://data.cma.cn/data/detail/dataCode.

[19] He, C., Okada, N., Zhang, Q., *et al*. Modelling dynamic urban expansion processes incorporating a potential model with cellular automata [J]. *Landscape and Urban Planning*, 2020, 86: 79–91.

[20] Huang, M., Wang, Z. C., Pan, X. H., *et al*. Delimiting China's urban growth boundaries under localized shared socioeconomic pathways and various urban expansion modes [J]. *Earth's Future,* 2022, 10: e2021EF002572.