CAPACITY BUILDING FOR GIS-BASED SDG INDICATORS ANALYSIS WITH GLOBAL HIGH-RESOLUTION LAND COVER DATASETS

D. Oxoli[1,]*, S. R. Reyes[2], S. Peng[3], M. A. Brovelli[1], S. Coetzee[4], I. Ivanova[5], J. A. Leonardi[1], D. Rawal[6], G. Vacca[7], S. Zlatanova[8]

[1] Dept. of Civil and Environmental Engineering, Politecnico di Milano, P.zza Leonardo da Vinci 32, 20133 Milano, Italy - (daniele.oxoli, maria.brovelli)@polimi.it, juliaanna.leonardi@mail.polimi.it
[2] United Nations Satellite Centre, United Nations Institute for Training and Research, United Nations Building, Bangkok, Thailand – sheryl.reyes@unitar.org
[3] National Geomatics Center of China, Beijing, China – pengshu@ngcc.cn
[4] Dept. of Geography, Geoinformatics and Meteorology, University of Pretoria, Hatfield Campus, Lynnwood Road, Hatfield, 0083 Pretoria, South Africa – serena.coetzee@up.ac.za
[5] School of Earth and Planetary Sciences, Curtin University, Kent Street Bentley, Perth, Australia – ivana.ivanova@curtin.edu.au
[6] Dept. of Languages and Literatures, Communication, Training and Society, Laboratory of Geomatics Guido Barbina, Udine University, Via Palladio 8, 33100 Udine, Italy – rawalnet@yahoo.com
[7] Dept. of Civil and Environmental Engineering and Architecture, University of Cagliari, Piazza D'Armi, 09123 Cagliari, Italy – vaccag@unica.it
[8] School of Built Environment, University of New South Wales, 2052 Sydney, Australia – s.zlatanova@unsw.edu.au

KEY WORDS: SDGs, Capacity Building, Global Land Cover, Open Geospatial Data, GIS, Free and Open-Source Software

ABSTRACT:

The support of geospatial data and technologies for the United Nations Sustainable Development Goals (SDG) framework is critical for assessing and monitoring key indicators, revealing the planet's trajectory towards sustainability. The availability of global open geospatial datasets, especially high-resolution land cover datasets, provides significant opportunities for computing and comparing indicators across different regions and scales. However, barriers to their proficient use remain due to a lack of data awareness, management and processing capacities using geographic information systems software. To address this, the "Capacity Building for GIS-based SDG Indicator Analysis with Global High-resolution Land Cover Datasets" project created open training material on discovering, accessing, and manipulating global geospatial datasets for computing SDG indicators. The material focuses on water and terrestrial ecosystems, urban environments, and climate, by leveraging world-class global geospatial datasets and using the Free and Open Source Software QGIS. The training material is released under a Creative Commons Attribution 4.0 License, ensuring broad accessibility and facilitating continuous improvement.

1. INTRODUCTION

The support of geospatial data and technologies to the United Nations Sustainable Development Goals (UN SDGs) framework has turned out to be critical for both the assessment and the monitoring of key indicators, revealing the trajectory of our planet and society towards sustainability (Chen et al., 2020). The increasing availability of global coverage and openly licensed geospatial datasets unlocks noteworthy opportunities for the computation and comparison of these indicators across different geographical regions, as well as multiple spatial and temporal scales (UN-GGIM, 2021).

To that end, a notable class of global geospatial products is represented by High-Resolution Land Cover (HRLC) datasets which have a critical role in many scientific and policy-making applications, including climate modelling, natural resource management, landscape and biodiversity preservation, and urbanization monitoring (Bratic et al., 2021). Global HRLC data, coupled with complementary global open geospatial layers (such as spatial demographic datasets, satellite-derived soil and vegetation products, etc.) provide plentiful capabilities to investigate SDG indicators (Avtar et al., 2020).

* Corresponding author

Generally, national or local geospatial data provide (where available) a higher spatial and temporal resolution than coarser global products and - in turn - better capabilities for accurate computation of SDG indicators. However, the uneven availability of such data often prevents consistent assessment of SDG progress across countries worldwide. In this context, the use of open global geospatial datasets represents a quick, efficient, and cost-effective solution for providing a first comprehensive assessment before local data is produced and analysed (Carter, 2019). The tangible benefits of these datasets are particularly evident in developing countries, where the availability of such information from local authorities is often limited. Nevertheless, there are still several barriers to their proficient use due to the lack of data awareness, management and processing capacities using Geographic Information Systems (GIS) software tools (Brovelli et al., 2018). The presence of sufficient human expertise in GIS technology is recognized as a crucial foundation for effectively incorporating geospatial information into SDG implementation. Particularly in developing countries, this capacity is still in its early stages, and there is a strong need for initiatives that prioritize capacity building in GIS for SDGs (Sarvajayakesavalu, 2015).

Given the above, we present here the Capacity Building for GIS-based SDG Indicators Analysis with Global High-resolution Land Cover Datasets project, funded by the ISPRS Education and Capacity Building Initiatives 2022 (`https://www.isprs.org/society/ecbi`). The project addressed the creation of open training material covering the complete learning process of discovering, accessing, and manipulating global open geospatial datasets for computing SDG indicators, with a focus on those directly connected to water and terrestrial ecosystems, urban environment, and climate. The training material emphasizes the use of Free and Open Source Software (specifically QGIS) as well as of premier open global geospatial datasets, and it is made available under a Creative Commons Attribution 4.0 License, allowing widespread access and facilitating its maintenance and updates.

The remainder of this paper is as follows. Section 2 reports a focused review of global geospatial datasets suitable for SDG indicators computation. Sections 3 and 4 describes respectively the strategy used to collect training requirements and the development of the training material. Section 5 outlines dissemination and outreach activities carried out to promote the training material, while conclusions are reported in Section 6.

2. GLOBAL GEOSPATIAL DATASETS FOR SDG INDICATORS

Despite the development of the SDG indicators framework was originally based on traditional statistical data, many statistical offices and governmental agencies have recognised the need for geospatial data to augment non-spatial information and provide new and consistent data sources, which are critical to SDGs monitoring (UN-GGIM, 2021). In fact, It has been estimated that approximately 20% of the SDG indicators can be interpreted and measured either through the direct use of geospatial data itself or through integration with complementing statistical data (Allen et al., 2021). GIS, Earth observations and global-coverage geospatial data provide an important link to enable consistent comparison among countries, provide granularity and disaggregation of the indicators and communicate their geographic dimensions (Avtar et al., 2020).

Global geospatial data products that cover major thematic areas of the biosphere and society (e.g. land cover, vegetation productivity, forests, wetlands, surface water, human settlements, etc.) are strongly supporting the methodological development and measurement of several SDG indicators (European Space Agency, 2020). National or local geospatial data, when accessible, offer improved resolution for accurate measurement of SDG indicators compared to global data. However, the uneven availability of such data hinders consistent assessment of SDG progress worldwide. Open global geospatial datasets provide a quick, efficient, and cost-effective solution for conducting an initial comprehensive assessment before national data becomes available (Carter, 2019; Allen et al., 2021).

2.1 Global high-resolution global land cover datasets

A key class of global geospatial products, which plays a major role in many SDGs, is the class of high-resolution global land cover (HRLC) maps (Bratic et al., 2021). The land cover stands for all the geospatial information describing the physical and biological cover of the Earth's surface. Land cover data products at high spatial resolution have a critical role in many scientific and policy-making applications, including climate modelling, natural resources management, landscape and biodiversity preservation, urbanization monitoring, and spatial demography (Radwan et al., 2021). In the last two decades, the advancement of satellite Earth observation as well as the availability of global coverage satellite imagery with high temporal resolution - available as free and openly-licensed data - have significantly promoted international coordination in SDGs monitoring and favoured production, access and usability of global and multi-temporal HRLC data products (Avtar et al., 2020; Guo et al., 2023).

Accordingly, it has been assessed that currently available open and global HRLC data could potentially be exclusively used for measuring indicators from four of the SDGs (i.e. 6 Clean water and sanitation, 13 Climate action, 14 Life below water, 15 Life on land), and could be used to complement other data types for four other goals (2 Zero hunger, 9 Industry, innovation and infrastructure, 11 Sustainable cities and communities, 12 Responsible consumption and production) (UN-GGIM, 2021).

This increased global HRLC data availability has led to a wealth of studies, applications and methodology guidelines connected to the use of such data in SDG indicators computing and monitoring. Relevant examples in the literature include the analysis of land-cover efficiency (Estoque et al., 2021), land consumption rate (Wang et al., 2020), and forest cover (Sayer et al., 2019), which are key thematic variables e.g. for SDG 15 and 11. Open and global HRLC data provide different temporal coverages, spatial resolutions, thematic details (or classes), and accuracies. Therefore, the suitability of global HRLC data for each specific SDG indicator has to be assessed case by case. According to Carter (2019), fit-for-purpose global HRLC data should be characterized by a minimum spatial resolution of 1 km and a minimum temporal resolution of 3 to 5 years. The thematic detail is instead indicator-specific. Finally, the accuracy of global HRLC data generally varies across the globe and local validation is strongly suggested depending on study areas and the user-expected accuracy performances (Brovelli et al., 2018). Some of the most popular and mature, openly available, and fit-for-purpose global HRLC datasets are reported and detailed in Table 1.

2.2 Complementary global geospatial datasets

Global HRLC data were also used in combination with complementary global geospatial information demonstrating their capabilities of directly supporting the computation of indicators e.g. from SDG 2 (Radwan et al., 2021), 6 (Mulligan et al., 2020), and 9 (Ilie et al., 2019). Actually, land cover is only one of the Global Fundamental Geospatial Data Themes to SDGs identified by the UN (UN-GGIM, 2019). These fundamental themes include a variety of geospatial datasets which are necessary to spatially represent key phenomena and objects for the realisation of economic, social, and environmental targets identified by the SDGs. These datasets include - among others - demography, soil/vegetation parameters, administrative boundaries, infrastructures, transport networks, 2D/3D topographic maps, water bodies, etc. (UN-GGIM, 2021). The suitability of complementary geospatial datasets is challenging to assess globally for each specific SDG indicator or country. Spatial and temporal resolution, thematic details and accuracy should be evaluated case-by-case, using recommendations valid also for the global HRLC data, and their use should be functional to the provision of baseline data in support of national SDGs reporting (Arnold et al., 2016).

Table 2 contains information on popular global open geospatial datasets which might be employed in combination with global

HRLC Dataset	Temporal Coverage	Spatial Resolution / Thematic detail	Source
GlobeLand30	2000, 2010, 2020	30 m / 10 land cover classes	http://www.globeland30.org
Finer Resolution Observation and Monitoring of Global Land Cover (FROM-GLC)	2010, 2015, 2017	30 m / 10 land cover classes	http://data.ess.tsinghua.edu.cn
Copernicus Global Land Cover	2015-2019 (every year)	100 m / 23 land cover classes and 10 land cover fractions	https://land.copernicus.eu/global/products/lc
ESA-CCI Land Cover	1992-present (every year)	300 m / 22 land cover classes	https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-land-cover
Terra and Aqua combined Moderate Resolution Imaging Spectroradiometer (MODIS) Land Cover Type	2001-present (every year)	500 m / multiple land cover classes and classification schemas	https://lpdaac.usgs.gov/products/mcd12q1v006
Global Human Settlement built-up surface (GHS-BUILT-S)	1975 - 2030 (every 5 years)	Up to 100 m (3 arc seconds) / single land cover class (built-up)	https://data.jrc.ec.europa.eu/collection/ghsl
Freshwater Ecosystems Explorer	2000-present	Up to 300 m / permanent and seasonal surface water, reservoir, wetlands, mangroves	https://www.sdg661.app

Table 1. Selection of most popular and mature open global HRLC datasets suitable for SDG indicators computation.

HRLC data for the computation of land cover-connected SDG indicators.

3. TRAINING REQUIREMENTS

The identification of training requirements and the primary recipients of the training material played a crucial role in the project's development. The training requirements and the training material syllabus were determined through surveys and active discussions with experts and potential users who were involved in all phases of the project.

Dataset	Temporal Coverage	Spatial Resolution / Thematic detail	Source
World Bank Official Boundaries	Annual	vector map / administrative boundaries including international boundaries, disputed areas, coastlines, lakes	https://datacatalog.worldbank.org/search/dataset/0038272
WorldPop Population Counts	2000-2020 (every year)	Up to 100 m (3 arc seconds) / population	https://hub.worldpop.org/project/categories?id=3
Global Human Settlement - Population (GHS-POP)	1975-2030 (every 5 years)	Up to 100 m (3 arc seconds) / population	https://data.jrc.ec.europa.eu/collection/ghsl
GHS Urban Centre Database (GHS-UCDB)	2015, 2019	1 km (30 arc seconds) / urban centres boundaries	https://data.jrc.ec.europa.eu/collection/ghsl
OpenStreetMap	/	vector map / multiple, including transport networks	https://www.openstreetmap.org

Table 2. Selection of popular global open geospatial datasets, complementary to global HRLC data, relevant to SDG indicators computation.

To outline and assess the training requirements, the project team conducted an open survey. The survey was distributed among the project investigators and extended contacts, including the media channels of the ISPRS Student Consortium, in alignment with the project's educational objectives. The survey results helped identify the primary user group and their specific training needs.

Additionally, the survey collected valuable feedback on the initiative and generated an initial list of individuals interested in participating in the training. Survey respondents (as of May 2022) were 102 whose demographics are reported in Figure 1. Respondents' familiarity with SGDs, GIS and global geospatial data was also investigated, together with their interest in using the training material.

The respondents' group was composed of mainly graduate students and both young researchers and professionals who, accordingly, are considered the primary users' group for the project. Respondents declared a good knowledge of GIS while a generally lower familiarity with SDGs (see Figure 2) and with the use of global maps to compute SDG indicators. More than 70% of the respondents provided a personal email address to receive updates from the project and declared interest in open

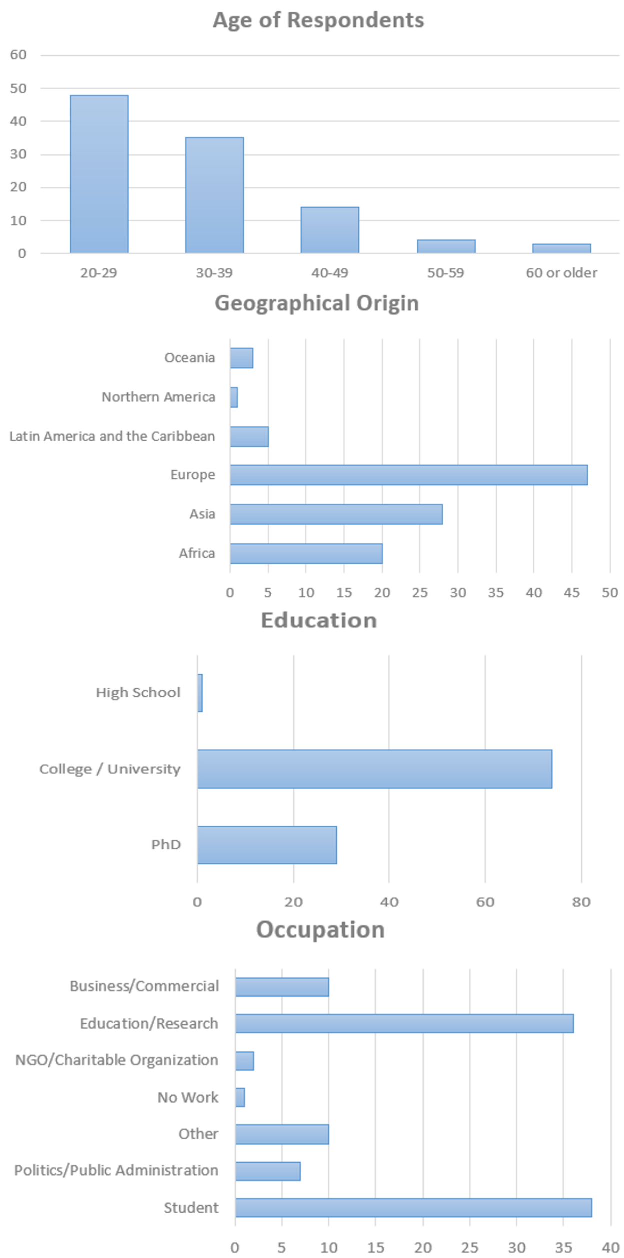


Figure 1. Demographics of survey respondents.

online training material as well as workshops on the project topics (see Section 5).

4. TRAINING MATERIAL DEVELOPMENT

The training material was designed as a web book and includes an introductory practical course on the combined use of open GIS and global HRLC and complementary datasets to compute SDG indicators. The web book sections are enriched with a theoretical/informative preamble to the SDGs and their connection to geospatial data and GIS, descriptive lists of relevant datasets, and hands-on exercises (see Figure 3a). Currently, the practices section of the web book (see Figure 3b) proposes computer exercises on the computation of indicators 15.1.1 (Forest area as a proportion of total land area), 11.3.1 (Ratio of land consumption rate to population growth rate), 9.1.1 (Proportion of the rural population who live within 2 km of an all-season

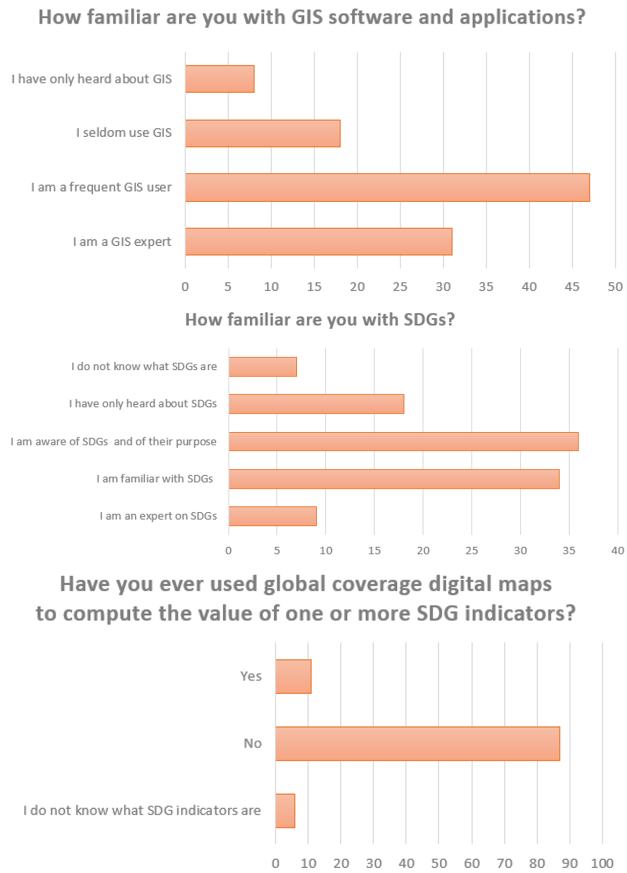


Figure 2. Familiarity of the survey respondents with SGDs and GIS.

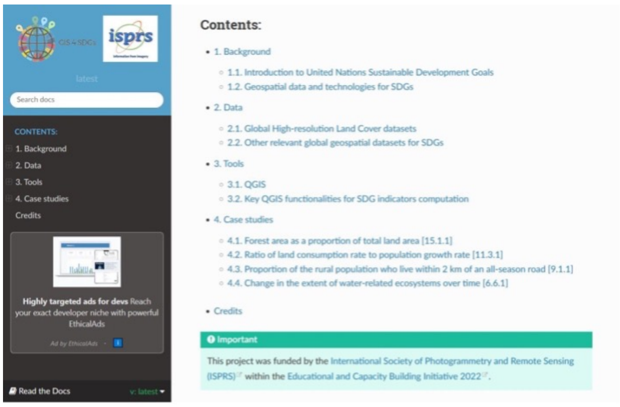
road, and 6.6.1 (Change in the extent of water-related ecosystems over time).

The datasets and bibliographic references, identified in Sections 2, represented both the theoretical and the material background on which the project is based. The openness of both data and software is considered key to the project's scope. To that end, the digital assets exploited in the development of material were selected as follows. To ensure the widest possible accessibility, the developed material leverages the Free and Open Source Software (FOSS) QGIS and it is released under a Creative Commons Attribution 4.0 License (CC BY 4.0). The web book was created using the Python-Sphinx documentation generator. The web book source code was versioned on GitHub in an online public repository at `https://github.com/opengeolab/ISPRS_GIS_SDG`, while the compiled version was published using the Read the Docs open hosting service at `https://isprs-gis-sdg.readthedocs.io`. Sample datasets for hands-on exercises were published on Zenodo in an open repository at `https://zenodo.org/record/8096448`.

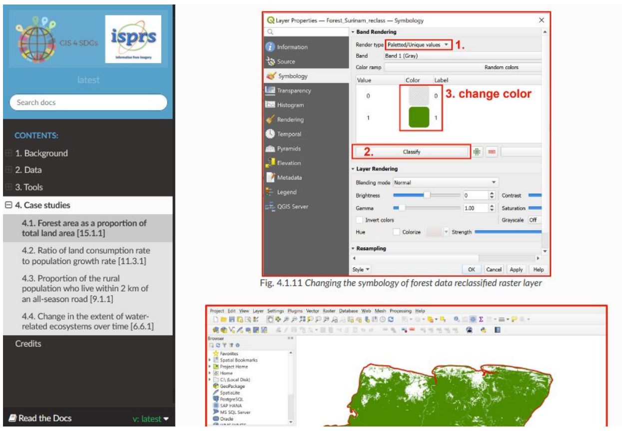
5. DISSEMINATION AND OUTREACH

Workshops connected to the training material were envisaged by the project proposal and further justified by the survey responses, as discussed in Section 3.

Three public workshops were organised during which both project development works and outputs were presented. Workshop 1 focused on the preliminary project presentation, includ-



(a)



(b)

Figure 3. (a) landing page and (b) sample of the practices section of the web book.

ing survey outcomes, by seeking feedback from both potential users and domain experts on the course syllabus. This activity was carried out through the participation of the project team at the Geospatial Information-enabled SDG Monitoring (GI4SDG) Forum Track of the XXIV ISPRS Congress held in Nice, France on June 7, 2022. Workshop 2 was dedicated to pilot testing of the draft web book content by collecting feedback which was used useful to adjust and complete the final version of the material (recording available at: `https://tinyurl.com/52xw95jd`). Workshop 3 presents the final version of the web book and it was provided as an open blended - in person and online - lecture (recording available at: `https://tinyurl.com/ms4b9e2m`).

Finally, the web book was listed on the UN Global Geospatial Information Management (UN-GGIM) Academic Network's website, within the educational material section (`http://unggim.academicnetwork.org/educational-material`) with the goal of promoting its diffusion.

6. CONCLUSIONS

The ultimate goal of the presented project is to enhance awareness of the connection between GIS and SDGs and engage, especially, GIS students and practitioners due to their potential key contribution to achieving Agenda 2030 in the next few years. To that end, advocating open data and software in this path is critical for maximising access to the required knowledge for students worldwide (without any economic barrier). The transformative potential of geospatial technologies in capacity-building for SDGs emphasizes the importance of collaboration, knowledge sharing, and inclusive approaches to tackle complex global challenges. The open availability of training material, analysis methods and tools promoted by the project represents a small but significant step forward in this direction.

Future activities of this work will include a continuous revision of the developed material. Integration with additional case studies on different SDG indicators is also envisaged.

ACKNOWLEDGEMENTS

This work has been funded by the Educational and Capacity Building Initiative 2022 (`https://www.isprs.org/society/ecbi`) of the International Society for Photogrammetry and Remote Sensing (ISPRS).

References

Allen, C., Smith, M., Rabiee, M., Dahmm, H., 2021. A review of scientific advancements in datasets derived from big data for monitoring the Sustainable Development Goals. *Sustainability Science*, 16(5), 1701–1716.

Arnold, S., Chen, J., Eggers, O., 2016. Global and complementary (non-authoritative) geospatial data for SDGs: Role and utilisation. *Working Group on Geospatial Information of the Inter-agency and Expert Group on Sustainable Development Goal Indicators (IAEG-SDGs)*.

Avtar, R., Aggarwal, R., Kharrazi, A., Kumar, P., Kurniawan, T. A., 2020. Utilizing geospatial information to implement SDGs and monitor their progress. *Environmental monitoring and assessment*, 192, 1–21.

Bratic, G., Vavassori, A., Brovelli, M. et al., 2021. Review of high-resolution global land cover. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43, 175–182.

Brovelli, M., Minghini, M., Molinari, M., Wu, H., Zheng, X., Chen, J. et al., 2018. Capacity building for high-resolution land cover intercomparison and validation: What is available and what is needed. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 15–22.

Carter, S. L., H. M., 2019. Specifications of land cover datasets for sdg indicator monitoring. https://tinyurl.com/mr44yj2h.

Chen, J., Peng, S., Chen, H., Zhao, X., Ge, Y., Li, Z., 2020. A comprehensive measurement of progress toward local SDGs with geospatial information: methodology and lessons learned. *ISPRS International Journal of Geo-Information*, 9(9), 522.

Estoque, R. C., Ooba, M., Togawa, T., Hijioka, Y., Murayama, Y., 2021. Monitoring global land-use efficiency in the context of the UN 2030 Agenda for Sustainable Development. *Habitat International*, 115, 102403.

European Space Agency, 2020. Compendium of earth observation contributions to the sdg targets and indicators. https://eo4society.esa.int/wp-content/uploads/2021/01/EO$_{C}ompendium - for - SDGs.pdf$.

Guo, H., Dou, C., Chen, H., Liu, J., Fu, B., Li, X., Zou, Z., Liang, D., 2023. SDGSAT-1: The world's first scientific satellite for Sustainable Development Goals. *Science Bulletin*, 68(1), 34–38.

Ilie, C. M., Brovelli, M. A., Coetzee, S., 2019. Monitoring SDG 9 with global open data and open software–A case study from rural Tanzania. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 1551–1558.

Mulligan, M., van Soesbergen, A., Hole, D. G., Brooks, T. M., Burke, S., Hutton, J., 2020. Mapping nature's contribution to SDG 6 and implications for other SDGs at policy relevant scales. *Remote Sensing of Environment*, 239, 111671.

Radwan, T. M., Blackburn, G. A., Whyatt, J. D., Atkinson, P. M., 2021. Global land cover trajectories and transitions. *Scientific Reports*, 11(1), 1–16.

Sarvajayakesavalu, S., 2015. Addressing challenges of developing countries in implementing five priorities for sustainable development goals. *Ecosystem Health and Sustainability*, 1(7), 1–4.

Sayer, J., Sheil, D., Galloway, G., Riggs, R. A., Mewett, G., MacDicken, K. G., Arts, B., Boedhihartono, A. K., Langston, J., Edwards, D. P., 2019. Sdg 15 life on land–the central role of forests in sustainable development. *Sustainable development goals: their impacts on forest and people*, Cambridge University Press, 482–509.

UN-GGIM, 2019. The global fundamental geospatial data themes. https://tinyurl.com/ystnzvsf.

UN-GGIM, 2021. SDGs geospatial roadmap. https://tinyurl.com/4z6wvbde.

Wang, Y., Huang, C., Feng, Y., Zhao, M., Gu, J., 2020. Using earth observation for monitoring SDG 11.3. 1-ratio of land consumption rate to population growth rate in Mainland China. *Remote Sensing*, 12(3), 357.

https://doi.org/10.5194/isprs-archives-XLVIII-1-W2-2023-559-2023