



OPEN

DATA DESCRIPTOR

A synthesis of hydroclimatic, ecological, and socioeconomic data for transdisciplinary research in the Mekong

Amar Deep Tiwari¹ , Yadu Pokhrel¹ ✉, Daniel Kramer², Tanjila Akhter¹, Qihong Tang³ , Junguo Liu⁴ , Jiaguo Qi⁵, Ho Huu Loc⁶ & Venkataraman Lakshmi⁷

The Mekong River basin (MRB) is a transboundary basin that supports livelihoods of over 70 million inhabitants and diverse terrestrial-aquatic ecosystems. This critical lifeline for people and ecosystems is under transformation due to climatic stressors and human activities (e.g., land use change and dam construction). Thus, there is an urgent need to better understand the changing hydrological and ecological systems in the MRB and develop improved adaptation strategies. This, however, is hampered partly by lack of sufficient, reliable, and accessible observational data across the basin. Here, we fill this long-standing gap for MRB by synthesizing climate, hydrological, ecological, and socioeconomic data from various disparate sources. The data—including groundwater records digitized from the literature—provide crucial insights into surface water systems, groundwater dynamics, land use patterns, and socioeconomic changes. The analyses presented also shed light on uncertainties associated with various datasets and the most appropriate choices. These datasets are expected to advance socio-hydrological research and inform science-based management decisions and policymaking for sustainable food-water, livelihood, and ecological systems in the MRB.

Background & Summary

Features of the MRB. The Mekong is an important transboundary river that supports water supplies, fisheries, irrigated agriculture, energy generation, riverine ecosystems, navigation, and recreation in the six countries that it flows across: China, Myanmar, Laos, Thailand, Cambodia, and Vietnam^{1–3}. An important ecological phenomenon characterized by a strong seasonal flow pattern that supports periodic inundation, as well as the exchange of water and nutrients between rivers and neighboring floodplains, known as the flood pulse⁴, is a lifeline to the numerous ecosystem services that the Mekong River basin (MRB) provides. In particular, the Tonle Sap Lake (TSL) and Mekong Delta are two important socio-ecological systems that rely on the Mekong flood pulse and support critical river-floodplain ecosystems as well as agriculture—especially rice farming—and fisheries^{5–7}. The lake located in the Cambodian portion of the basin is home to one of the world's most diverse inland fishery systems (e.g., Ziv *et al.*⁸). The fisheries and flood-recession agriculture in and around the lake are supported by the two-way flow in the Tonle Sap River⁹, driven by the Mekong flood pulse⁵. Furthermore, the Mekong Delta in Vietnam is a densely populated region and the third largest Delta in the world¹⁰. It is highly fertile and intensively cultivated, supporting the production of 7–10% of all rice traded worldwide¹¹.

¹Department of Civil and Environmental Engineering, Michigan State University, East Lansing, Michigan, USA.

²Department of Fisheries and Wildlife, Michigan State University, East Lansing, Michigan, USA. ³Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China. ⁴School of Water Conservancy, North China University of Water Resources and Electric Power, Zhengzhou, China. ⁵Center for Global Change and Earth Observations, Michigan State University, East Lansing, Michigan, USA. ⁶Water Engineering and Management, Asian Institute of Technology, Khlong Nueng, Pathum Thani, Thailand. ⁷Engineering Systems and Environment, University of Virginia, Charlottesville, Virginia, USA. ✉e-mail: ypokhrel@msu.edu

Shifting hydrological and ecological systems. The natural synchronization among hydrology, flood pulse, fishery, agriculture, and riverine ecosystems has remained relatively stable for generations, but these tightly connected systems are being transformed in recent times due to climate change and growing human interventions^{12–16}. In particular, climatic shifts, land use land cover changes, and a boom in hydropower dams have altered the river's hydrology and ecology^{9,17–19}, with profound implications on ecosystem services^{20,21}, nutrient and sediment budgets²², and biodiversity and productivity^{23,24}. Most notably, the construction of a series of dams along the mainstream Mekong is the dominant factor leading to a rapid transformation in the hydrological regime^{25,26}. The effects of new dams also extend beyond hydrology and ecology, causing widespread changes in greenhouse gas (GHG) emissions, nutrient dynamics, and local livelihoods^{24,27–31}.

As dam construction continues to accelerate across the MRB with increasing downstream impacts on food-energy-water (FEW) and ecological systems, there has been a growing interest in studying these consequences and developing improved adaptation strategies^{32–36}. Data collected either on the ground or remotely, and models are indispensable tools for examining the nexus of these human activities, impacts and social implications, and providing reliable future scenarios. As such, numerous studies in the MRB have attempted to investigate these ongoing changes using observed data^{3,37–40}, remote sensing products^{1,16,41–43}, and process-based models^{41,44–53}.

Studies based on observed datasets. Using observed hydrometeorological and climate datasets in the MRB, several studies have investigated the role and implications of long-term climate change, short-term climate variability, and human activities on the FEW nexus. For example, Fan and He³⁷ quantified the trends in precipitation and temperature to examine their effect on streamflow in the MRB. Mohammed *et al.*⁴¹ explored recent extreme flow events at various locations in the MRB. Other studies have examined the changes in groundwater, especially over the Mekong Delta or other small regions within the MRB, including those^{46,54–56} that have focused on groundwater analysis in the different parts of the basin. Numerous other studies have examined various aspects of hydrology and climatology over some parts of the basin^{13,39,41,42,47,57,58} which provide a small-scale or sub-basin level understanding of spatial and temporal variability. Associated nutrient dynamics have also been explored in a number of studies. Binh *et al.*³⁴, for example, identified the decreasing trend in sediment in the Mekong Delta due to upstream dam construction and Whitehead *et al.*⁵⁹ quantified the changes in nitrogen and phosphorus concentrations and fluxes. These studies have provided insights into many emerging issues but have also underscored data limitations that hindered more comprehensive, and basin-scale studies on the rapidly changing hydrological, agricultural, and ecological systems. Indeed, for the MRB, observational datasets are very limited, non-continuous, sparse, or not easily accessible.

Studies based on remote sensing datasets. Some of the gaps and limitations in the ground-based observational data have been addressed using emerging data from satellite remote sensing. These data have been specifically used to characterize the changes in land use^{1,43,60–63}, cropping patterns^{64–66}, map open surface water areas^{53,58,67–69}, develop dam attributes^{67,70}, and landslide estimation⁷¹ among others. Further, remotely sensed precipitation data have been used as inputs in modeling studies^{40,44,72}. Data from the Gravity Recovery and Climate Experiment (GRACE) satellite mission have been used to evaluate simulated surface storage, filling the gap in publicly available groundwater data for the MRB^{9,44,73}. Remote sensing products have also been used to evaluate model results, filling the gaps in observational datasets, or providing added opportunities for model development and validation^{45,50}. Despite these applications, remote sensing products suffer from inherent uncertainties resulting from numerous technical and methodological issues such as cloud contamination^{67,74}. Further, while remote sensing products are spatially extensive, their temporal coverage and frequency for many products are limited.

Gaps and Opportunities. The studies based on observed and remote sensing datasets have advanced the understanding of the processes, impacts, and drivers of hydrological and ecosystem shifts, especially in relation to dam construction^{9,25,45,47,75–78}. However, they have also revealed critical data gaps that have hindered a more complete understanding of the many rapidly emerging issues. Ground-based observational data are either limited or not easily accessible for the MRB⁷⁹. Satellite products can fill certain gaps but are not always reliable in terms of data quality and temporal gaps. Further, models can provide spatially-complete and temporally-continuous information, but the lack of observational data is a critical hurdle in constraining and validating process-based models. Compared to other global regions such as North America and Europe, data availability for the MRB is very limited⁸⁰. Some data are available but there is no formal mechanism to share, are not easily accessible, or sharing is constrained due to institutional restrictions. Therefore, there are opportunities to bring light to the available data and their utility for increasingly important transdisciplinary research and collaboration^{81,82}.

This paper fills a long-discussed and critical knowledge gap in data availability, accessibility, and utilization for the MRB. In this study, a synthesis of climate, hydrological, ecological, and socioeconomic datasets is provided, focusing on accuracy and utility, to guide the research community to the most suitable and reliable datasets for different purposes. Specifically, we synthesize and archive publicly available data using online repositories. For the datasets that are not publicly available (e.g., those available for purchase or through formal agreements), we identified the sources and presented the findings in graphical form, where possible. The key contribution of the paper lies in that it brings together many disparate datasets for hydrological, ecological, agricultural, and socioeconomic studies in the MRB, which are not readily available, not easily accessible, or are only available in graphical form in the published literature. It is expected that the paper could serve as a one-stop shop for certain key datasets for the aforementioned studies, which are increasingly essential to address growing issues on food, energy, water, and environmental systems across the MRB.

The rest of this paper is organized as follows. Methods section provides a detailed description of the numerous datasets, their sources, and methods to collect and produce the data. We discuss how to make use of disparate datasets that are freely available from their providers. Datasets that are not publicly available but are obtainable

S. No.	Data	Source	Native resolution	Data type	Remarks or weblinks*
1	Precipitation	EM-Earth	0.1°	Satellite	https://www.frdr-dfdr.ca/repo/dataset/8d30ab02-f2bd-4d05-ae43-11f4a387e5ad
2	Temperature	EM-Earth	0.1°	Satellite	https://www.frdr-dfdr.ca/repo/dataset/8d30ab02-f2bd-4d05-ae43-11f4a387e5ad
3	Wind Speed	MRC	Point data	<i>In Situ</i>	https://portal.mrcmekong.org
4	Relative Humidity	MRC	Point data	<i>In Situ</i>	https://portal.mrcmekong.org
5	Sunshine Hours	MRC	Point data	<i>In Situ</i>	https://portal.mrcmekong.org
6	Streamflow	MRC	Point data	<i>In Situ</i>	https://portal.mrcmekong.org
		Published literature	Point data	<i>In Situ</i>	Multiple sources
7	Water level	MRC	Point data	<i>In Situ</i>	https://portal.mrcmekong.org
8	Evapotranspiration	GLEAM	0.25°	Hybrid	https://www.gleam.eu/
		MODIS	500 m	Satellite	https://modis.gsfc.nasa.gov/data/dataproduct/mod16.php
9	Surface water	Landsat based	30 m	Satellite	https://global-surface-water.appspot.com/download
		MODIS based	500 m	Satellite	http://data.starcloud.pcl.ac.cn/resource/9
10	Soil Moisture	SMAP	1 km	Satellite	https://nsidc.org/data/nsidc-0779/versions/1133
		Ground-based	Point data	<i>In Situ</i>	Collected by an individual scholar and his team
11	Groundwater	Published literature	Point data	<i>In Situ</i>	Multiple sources
12	Dam data	WLE	Point data	<i>In Situ</i>	https://wle-mekong.cgiar.org/
13	Nutrients	MRC	Point data	<i>In Situ</i>	https://portal.mrcmekong.org
14	Sediment Concentration	MRC	Point data	<i>In Situ</i>	https://portal.mrcmekong.org
15	Land use	ESA-CCI	300 m	Satellite	https://www.esa-landcover-cci.org/
16	Crop yield	FAOSTAT	Country-wise	<i>In Situ</i>	https://www.fao.org/faostat/en/#data/QCL
17	Crop calendar	GGCMI	0.5°	Hybrid	https://zenodo.org/record/5062513#.ZBHs5nrMfuU
18	Irrigated area and Irrigation water use	FAO	0.0833°	Satellite	https://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas/latest-version
19	Irrigation water use (Cambodia)	Census	Point data	<i>In Situ</i>	Available for Cambodia
20	Wetlands	SWAMP	123 m	Satellite	https://www2.cifor.org/swamp/
		Tootchi <i>et al.</i> ¹⁹²	500 m	Satellite	https://doi.pangaea.de/10.1594/PANGAEA.892657
21	Inundation maps	Fluet-Chouinard <i>et al.</i> ¹⁹³	500 m	Satellite	www.estellus.fr/index.php?static13/giems-d15
22	Greenhouse gases	(EDGAR) v4.3.2	0.1°	Satellite	https://edgar.jrc.ec.europa.eu/
23	Socio-economic	Census	-	Hybrid	Multiple sources
24	Population	SEDAC	30-arc second	Hybrid	https://sedac.ciesin.columbia.edu/data/collection/gpw-v4
25	Population projections	Jones & O'Neill ²¹²	1/8°	Hybrid	http://sedac.ciesin.columbia.edu/data/set/popdynamics-1-8th-pop-base-year-projection-ssp-2000-2100-rev01/data-download
26	GDP	DRYAD	30-arc second	<i>In Situ</i>	https://datadryad.org/stash/dataset/doi:10.5061/dryad.dk1j0
27	GDP projections	Wang & Sun ²⁰⁷	0.25°	Hybrid	https://doi.org/10.5281/zenodo.5880037

Table 1. Key variables and their primary sources. *These are the weblinks of the raw datasets.

from the authors or agencies are also summarized. A list of data records is mentioned in the following section. In the technical validation section, we compare various datasets for each data type and provide insights on quality and validity. Based on the findings from this section, we provide usage notes for the available datasets in the last section.

Methods

Meteorological data. Meteorological datasets (e.g., precipitation, temperature, wind speed, relative humidity, sunshine hours; Table 1) are important elements of climate, hydrological, and ecosystems studies. These datasets constitute key inputs in hydrological modeling. Among various meteorological variables, precipitation serves as an entry point for a variety of applications relating to climate variability, hydrological modeling, agriculture, and ecosystems, among others^{83,84}. In the MRB, precipitation measurements—along with other meteorological variables—are available at certain gauging stations but the placement of these stations is rather sparse and non-homogenous^{85–87}. Various other meteorological products, especially for precipitation, are available that help fill observational data gaps and provide spatially continuous data for the entire MRB. For example, the Asian Precipitation-Highly Resolved Observational Data Integration Towards Evaluation (APHRODITE) product^{88,89}, the only available long-term regional gauge-based daily gridded precipitation dataset for Asia⁹⁰, is a key reference data used in many MRB studies^{90–92}. Further, the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS; Funk *et al.*⁹³), a 35 + year quasi-global rainfall dataset derived from gauge observations and satellite-based products and have recently been used by some MRB studies (e.g., Luo *et al.*⁹⁴; Guo *et al.*⁹⁵). Moreover, in the past three decades, earth observations by many satellite missions have provided spatially-complete and relatively high-resolution precipitation products, which have been increasingly used for many applications in the MRB^{41,96–98}. Recent studies have compared these various precipitation datasets^{40,44,72}, some suggesting that APHRODITE is amongst the most reliable precipitation products for the MRB^{40,72}.

S. No.	Data	Period	Spatial resolution	Temporal resolution	Data type	Reference
1	Mekong River Commission (MRC)	Station-dependent	Point data	Daily	Observed	—
2	Asian Precipitation-Highly Resolved Observational Data Integration Towards Evaluation (APHRODITE)	1951–2015	0.25°	Daily	Satellite	Yatagai <i>et al.</i> ^{88,89}
3	Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS)	1981–present	0.05°	Daily	Satellite	Funk <i>et al.</i> ⁹³
4	Tropical Rainfall Measuring Mission (TRMM)	1998–2019	0.25°	3 hourly	Satellite	Simpson <i>et al.</i> ²²⁶
5	Global Precipitation Measurement (GPM)	2014–present	0.1°	3 hourly	Satellite	Hou <i>et al.</i> ^{227,228}
6	Integrated Multi-satellitE Retrievals for GPM (IMERG)	2000–present	0.1°	30 minutes	Satellite	Huffman <i>et al.</i> ²²⁹
7	TRMM Multi-satellite Precipitation Analysis (TMPA)	1998–2019	0.25°	3 hourly	Satellite	Huffman <i>et al.</i> ²³⁰
8	Climate Prediction Center Morphing (CMORPH)	1998–present	8 km	30 minutes	Satellite	Joyce <i>et al.</i> ²³¹
9	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN)	1983–present	0.25°	Daily	Satellite	Hsu <i>et al.</i> ²³² ; Sorooshian <i>et al.</i> ²³³
10	Climate Prediction Center (CPC)	1979–present	0.5°	Daily	Satellite	Xie <i>et al.</i> ²³⁴
11	Climate Research Unit (CRU)	1901–2018	0.5°	Monthly	Satellite	Harris <i>et al.</i> ²³⁵
12	Global Precipitation Climatology Center (GPCC)	1891–present	0.5°	Monthly	Satellite	Schneider <i>et al.</i> ²³⁶
13	Precipitation Reconstruction over Land (PREC/L)	1948–present	0.5°	Monthly	Satellite	Chen <i>et al.</i> ²³⁷
14	Multi-Source Weighted-Ensemble Precipitation v1.0 (MSWEP)	1979–present	0.25°	Daily	Satellite	Beck <i>et al.</i> ²³⁸
15	ECMWF Reanalysis v5 (ERA5)	1959–present	0.25°	Hourly	Satellite	Hersbch <i>et al.</i> ⁹⁹
16	Princeton data	1948–2016	0.25°	Daily	Hybrid	He <i>et al.</i> ¹⁰⁰
17	TerraClimate	1958–2015	1/24°	Monthly	Hybrid	Abatzoglou <i>et al.</i> ²³⁹
18	EM-Earth	1950–2019	0.1°	Daily	Hybrid	Tang <i>et al.</i> ¹¹⁰

Table 2. Summary of various precipitation products useful for MRB studies.

As such, precipitation products are seemingly many; however, their spatial resolution and temporal availability period limit their utility for many applications; for example, process-based hydrological modeling often requires sub-daily data (e.g., Kabir *et al.*⁴⁴), but many products noted above include only daily datasets. Table 2 summarizes, to our knowledge, the existing precipitation products, mostly global, with details on their resolution and availability period.

Many studies, especially on hydrological modeling, require meteorological input other than precipitation, which includes temperature, solar radiation, humidity, surface pressure, and wind speed. Such data are largely lacking for the MRB, except for the sparse gauge-based data from the MRC (Fig. 1). Therefore, modeling studies generally employ data from global products, which are primarily based on atmospheric reanalysis such as the ECMWF Reanalysis v5 (ERA5; Hersbch *et al.*⁹⁹). There are numerous other global products that could be used for basin-scale modeling, which are derived from different reanalysis datasets. These include the Princeton Global Forcing data^{100,101}, WATCH Forcing methodology applied to ERA-Interim reanalysis data (WFDEI^{102,103}), meteorological forcing data of the third Global Soil Wetness Projects (GSWP3; Kim¹⁰⁴), and WFDEI5 over land merged with ERA5 over the ocean (W5E5; Lange *et al.*¹⁰⁵). Brocca *et al.*¹⁰⁶ proposed an algorithm to estimate the effective rainfall data from *in-situ* soil moisture data (SM2RAIN). Later, this algorithm has been applied to various satellite-based soil moisture datasets to estimate the global effective rainfall (e.g., SM2RAIN-CCI (Ciabatta *et al.*¹⁰⁷), SM2RAIN-ASCAT (Brocca *et al.*¹⁰⁸), and GMP + SM2RAIN (Massari *et al.*¹⁰⁹)). One common limitation in many of these products is the coarse spatial resolution (typically 0.5° ~50 km at the equator), which limits the application to only basin-scale modeling studies^{19,44}.

To overcome the limitations related to spatial-temporal resolution and inherent biases, recent efforts have led to the development of higher resolution products such as the Ensemble Meteorological Dataset for Planet Earth (EM-Earth) data at 0.1° (~10 km at the equator) spatial resolution over global land areas from 1950 to 2019¹¹⁰. These data have 25 ensemble members enabling uncertainty analyses and sensitivity test in hydrological modeling. Another such recent product is the Climatologies at high resolution for the earth's land surface areas (CHELSA) data (Karger *et al.*¹¹¹; <https://chelsa-climate.org/>), also available at 30 arc seconds (~1 km) globally. However, both of these products are available at a daily time step, limiting the utility to models that only resolve water balance; land surface models that resolve energy balance typically require sub-daily datasets¹¹². Nevertheless, the EM-Earth ensemble datasets have the potential to be useful for probabilistic climate and hydrological modeling. We have synthesized these datasets or have noted relevant sources where data are readily accessible.

Hydrological data. *Streamflow and water level.* The primary source of the observed hydrological data in the MRB is the Mekong River Commission (MRC), which provides gauge-based data on river discharge and water levels at over 29 and 47 stations for streamflow and water level, respectively, across the basin (available through formal agreement or for purchase). Water level observations are also available from other sources such as the Cambodia Ministry of Water Resources and have been presented in the published literature (e.g., Arias *et al.*¹¹³). Observed data for the Chinese portion of the basin (i.e., Upper MRB (UMRB)) are generally not available to the international community but have been presented in some journal articles^{3,26,37,114,115}. Some disparate streamflow data also exists for selected stations including Yun Jing Hong/ Yunjinghong^{114,116} (Chinese border), Jiuzhou^{37,116},

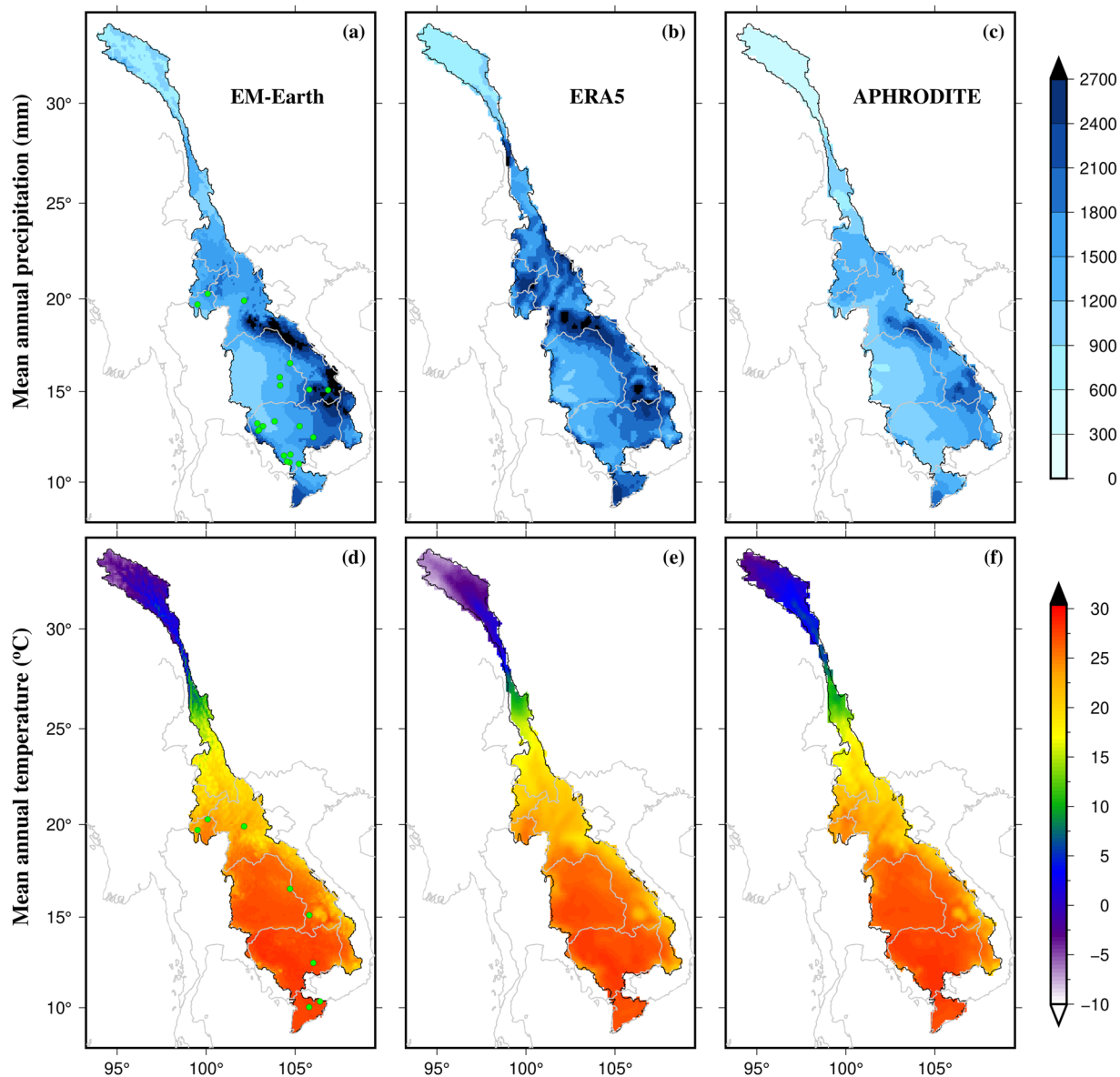


Fig. 1 Spatial coverage of mean annual precipitation (mm) (1971–2000) for (a) EM-Earth, (b) ERA5, and (c) APHRODITE data and mean annual temperature (°C) (1971–2000) for (a) EM-Earth, (b) ERA5, and (c) APHRODITE data. Green circles in precipitation and temperature data panels (a and d) show gauging stations for observed datasets obtained from the MRC.

Gajiu¹¹⁶, and Changdu¹¹⁷ stations in China; we have digitized these from the published literatures. The Global Runoff Data Center (GRDC) provides some streamflow data for the MRB within its global database but only for a small number of stations, which are included in the MRB database. Here, we present the complete information on available data from the MRB and other sources, along with some infographics.

Evapotranspiration (ET). Similarly, ET is typically not measured *in-situ* due to the difficulty of deploying a network of ground-based networks over the entire area in the MRB. Therefore, satellite-based ET products, which provide a continuous record of ET at a global scale with a relatively high temporal resolution, are often used as an alternative. Some global ET products that have been used in the MRB include the water balance (WB; Zeng *et al.*¹¹⁸) based ET, GLEAM product (Martens *et al.*¹¹⁹), Penman-Monteith-Leuning data (PMLv2; Zhang *et al.*¹²⁰) (500 m), Moderate-resolution Imaging Spectroradiometer global terrestrial evapotranspiration product (MOD16; Mu *et al.*¹²¹) (500 m), and Global LAnd Surface Satellite datasets (GLASS; Xie *et al.*¹²²) (1 km). Studies such as Hu and Mo¹²³ and Chen *et al.*¹²⁴ have evaluated the performance of these products in the MRB and found that their accuracy can vary depending on specific conditions and characteristics of the region. Chen *et al.*¹²⁴ concluded that Moderate Resolution Imaging Spectroradiometer (MODIS) ET underperforms in the MRB compared to other selected datasets. While Hu and Mo¹²³ compared model simulated ET with satellite datasets and suggested that, in the MRB, GLEAM ET performs comparatively better than other products. Further,

MODIS ET does not include data for land cover types specified as unclassified, urban, wetlands, perennial snow/ice, and permanent water bodies. Here, given certain uncertainties in both GLEAM (version 3.6b) and MODIS (version 6.1, gap filled) ET datasets and lack of observational data, we compare the two to demonstrate how they differ spatially and temporally.

Surface water. Monitoring surface water volume is a crucial aspect of water resource management, as it helps understand the availability and dynamics of water in a region. Surface water can be monitored using surface water area and water level^{125,126}. Satellite altimetry datasets, such as those provided by TOPEX/Poseidon, Jason-1, Jason-2, Jason-3, and Envisat, use radar measurements to determine the height of the water surface and satellite imagery, such as that provided by MODIS, LANDSAT, and Sentinel, can be used to measure surface water area. For example, European Commission's Joint Research Centre (JRC), developed by Pekel *et al.*¹²⁷ used LANDSAT data at a spatial resolution of 30 meters to monitor surface water extent from 1984 to 2015. However, the temporal resolution of these datasets is relatively coarse, and they are available only in the form of percentage water occurrence at the monthly scale or as yearly classification. Moreover, there are limited images only for MRB which are cloud-free⁹. The recently launched Surface Water and Ocean Topography (SWOT) mission is expected to enable us address some of these limitations and greatly improve our ability to monitor surface water volume, especially by providing high-resolution data on surface water area and water level. Further, there are other satellite-based surface water products such as those generated by Ji *et al.*¹²⁸ using MODIS data, which are available at the daily interval at the spatial resolution of 500 m and for the 2001–2016 period (data source: <http://data.starcloud.pcl.ac.cn/resource/9>). Here, we have processed and compared the two remotely sensed surface water products by Pekel *et al.*¹²⁷ and Ji *et al.*¹²⁸ for MRB, and present the processed surface water datasets to the community.

Soil moisture. *In-situ* soil moisture data for the MRB are limitedly available, if not non-existent at the basin-scale. As a result, the only choice is to use globally available remote sensing-based soil moisture products. For example, soil moisture data are available from the i) Soil Moisture Active Passive (SMAP; Entekhabi *et al.*¹²⁹) at 9 km spatial resolution, ii) Soil Moisture and Ocean Salinity Level 3 (SMOS L3; Jacques *et al.*¹³⁰) at 25 km, iii) European Space Agency Climate Change Initiative (ESA-CCI SM v2.7; Liu *et al.*¹³¹; Wagner *et al.*¹³²) at 25 km, and iv) Global Land Evaporation Amsterdam Model (GLEAM: Martens *et al.*¹¹⁹) at 25 km. Recently, the SMAP soil moisture data have been downscaled to a finer spatial resolution of 1 km globally^{133,134} as well as locally¹³⁵. In this study, we focus on the downscaled 1 km SMAP product by Fang *et al.*¹³³ while also noting the utility of the other products. Among such limitedly available and disparate observed data are the observations at five locations (Chaiyabhum, Srisaket, Amnatcharoen, Sakonnakhon, and Bungkan) in Thailand, available at 5-minute intervals from 14th December 2017 to 12th February 2019 and provided by an individual scholar (see Acknowledgment section).

Groundwater. Groundwater data in the MRB are collected by respective government agencies in each member country. For example, the National Centre for Water Resources Planning and Investigation (NAWAPI) in Vietnam, the Department of Groundwater Resources (DGR) in Thailand, the Ministry of Water Resources and Meteorology (MOWRAM) in Cambodia, and the Department of Water Resources under the Minister of Natural Resources and Environment (DWR- MONRE) in Laos conduct groundwater monitoring. However, these datasets are generally not available to the public, nor included within the MRC's database. Some of the datasets (e.g., from NAWAPI) are available for scientific research conducted with an in-country team but are restricted from broader sharing. Further, numerous previous studies have collected groundwater data on an individual basis or obtained from certain partner agencies in the region. Yet, the data have not been shared beyond certain graphics in journal articles. Here, we have digitized all published data, obtainable through our best efforts from published sources^{46,54–56,136–156}, and identified various other sources via which groundwater data can be obtained, for example, through formal agreements with respective agencies. Details are provided in Table S1.

Dam data. Recently, over 100 hydropower dams have been constructed across the MRB, dramatically increasing reservoir storage capacity from ~5 to ~70 km³ during 2010–2020²⁵. Therefore, dams and their operation have become crucial aspects of hydrologic and ecosystem studies in the MRB, which demand reliable data on the attributes of existing and planned dams as well as on the way reservoirs are operated. Globally, data on large dams are available through the database of the World Register of Dams (WRD), maintained by the International Commission of Large Dams (ICOLD). These data have been synthesized, for example producing the Global Reservoir and Dam (GRanD) data¹⁵⁷ and used in many global studies^{112,158–160}. However, these global data include only a few large dams in the MRB, leaving a major information gap regarding the smaller or recently built dams or those that are planned. The Global geOreferenced Database of Dams (GOODD; Mulligan *et al.*¹⁶¹) includes larger number of dams compared to GRanD and the georeferenced global dams and reservoirs (GeoDAR; Wang *et al.*¹⁶²) and provides richer information on global dams. Yet, the necessary dam attributes (e.g., dam height and reservoir storage capacity) are not comprehensively included in most of these datasets. Recently, Zhang and Gu¹⁶³ developed Global Dam Tracker (GDAT), a comprehensive dam database which includes more than 35,000 global dams with their location, catchment area, and other attributes. The GDAT dataset includes attributes of 466 dams in the MRB. Further, there are notable discrepancies or missing attributes in many of these products (e.g., Shin *et al.*⁴⁵).

In this study, we present the data from the Research Program on Water, Land, and Ecosystems (WLE Mekong; <https://wle-mekong.cgiar.org/>) as the base product and enhance the database by using information from various other sources. Note that WLE is the primary data source for GDAT for the MRB region. Specifically, building on the efforts of Shin *et al.*⁴⁵, we conducted a thorough inspection of the existing database, made manual corrections using various independent sources (e.g., Google Earth, internet resources on individual dams, published literature and

S. No.	Land Use	Satellite	Resolution	Source/ Agency
1	ESRI Land cover	Sentinel- 2	10 m	https://gisgeography.com/living-atlas-of-the-world/
2	Global Land Survey (GLS)	Landsat 7 ETM+	30 m	The University of Maryland with the USGS
3	European Space Agency- Climate Change Initiative (ESA-CCI)	ENVISAT	300 m	http://maps.elie.ucl.ac.be/CCI/viewer/index.php
4	Moderate Resolution Imaging Spectroradiometer (MODIS)	MODIS	500 m	https://modis.gsfc.nasa.gov/data/dataproduct/mod12.php ; Friedl <i>et al.</i> ²⁴⁰
5	Global Land Cover Characterization (GLCC)	Advanced Very High-Resolution Radiometer	1000 m	Loveland <i>et al.</i> ²⁴¹
6	GlobeLand30	Landsat	30 m	https://gisgeography.com/usgs-earth-explorer-download-free-landsat-imagery/
7	National Land Cover Database (NLCD)	NOAA- USGS	30 m	https://www.usgs.gov/node/279743
8	SPOT 5	SPOT	20 m	Bartholomé and Belward ¹⁷¹
9	Synthetic Aperture Radar	SAR	25 m	Balzter <i>et al.</i> ¹⁷²
10	Copernicus Global Land Service	PROBA- V	100 m	https://land.copernicus.eu/global/lcviewer
11	NASA SERVIR Mekong	Landsat 4–7	30 m	https://landcovermapping.org/en/landcover/

Table 3. Details of land use/land cover datasets.

reports), and further verified with credible sources (e.g., Yigzaw *et al.*¹⁶⁴; Yun *et al.*¹⁶⁵; Galelli *et al.*⁷⁵; Schmitt *et al.*¹⁶⁶). We also selected dams that have either or both the dam height and reservoir capacity as these are the two basic attributes for dam impact studies. Finally, we have selected large dams, satisfying on one of the following criteria: (1) dam height ≥ 15 m, (2) storage capacity > 1 million m^3 (Mm^3), and (3) installed hydropower capacity > 100 Mega Watts (MW); these large dams are of particular important for basin-wide modeling studies on the impacts of upstream dams on large-scale downstream impacts (e.g., Räsänen *et al.*¹⁶⁷; Pokhrel *et al.*⁹; Shin *et al.*⁴⁵; Dang *et al.*⁴⁷; Chaudhari and Pokhrel¹⁶⁸).

Land use and crop data. Remotely sensed land use and crop data are critical for studies on land use transitions, crop patterns, and hydrologic change. NASA SERVIR Mekong provides land use data for Mekong countries. In particular, the incorporated land use mapping tool within NASA SERVIR Mekong is useful in quantifying the spatial impacts of dams on land use and land cover disturbances^{62,169,170}. However, some global products such as Satellite Pour l'Observation de la Terre (SPOT 1–5¹⁷¹), Synthetic Aperture Radar (SAR; Balzter *et al.*¹⁷²), Landsat 4–7, Environmental Satellite (ENVISAT) Advanced Synthetic Aperture Radar (ASAR), MODIS, Advanced Very High-Resolution Radiometer (AVHRR), and Sentinel 1–2 satellites are routinely used products for land use change detection in local and global scale studies. Moreover, the European Space Agency- Climate Change Initiative (ESA-CCI) land use data based on the ENVISAT satellite provides continuous annual global data from 1981. Comprehensive information about these land use datasets and satellites is provided in Table 3. Among these land use datasets, we selected ESA-CCI (<https://www.esa-landcover-cci.org/>) land use data as demonstration in this study owing to its relatively longer temporal coverage and high spatial resolution.

Crop datasets are crucial for accurately modeling hydrological and agricultural processes but datasets on crop types and cropping patterns are not specifically available for the MRB. Thus, as for many other regions, the alternative is to use crop types from Remote sensing. The commonly used Leaf Area Index (LAI) data, an important modelling attribute, in many MRB studies are based on MODIS products (e.g., Son *et al.*⁶⁵; Hu and Mo¹⁷³). Another critical parameter for understanding food security and agricultural productivity is crop yield, which is not available at a basin-wide scale. Therefore, studies in the MRB use global annual data on crop yield such as Food and Agriculture Organization Corporate Statistical Database (FAOSTAT)^{76,174,175}. For this study, we obtained country-based annual crop yield data for the period of 1961–2021 for the Lower MRB (LMRB) countries (Cambodia, Laos, Thailand, and Vietnam) from FAOSTAT. The datasets include annual crop yield for crops such as rice, maize, banana, and sugarcane, etc.

Similar to crop yield, crop calendar datasets are available at the global scale. Crop calendar datasets are necessary inputs in hydrological-agricultural modeling, and crucial products for broader agricultural and food security studies. The International Production Assessment Division (IPAD) of the U.S. Department of Agriculture (USDA) Foreign Agricultural Service (FAS) provides global crop calendar data for planting, mid-season, and harvesting periods for grains, oilseeds, and cotton. In addition to IPAD, the Group on Earth Observations Global Agricultural Monitoring (GEOGLAM; Whitcraft *et al.*¹⁷⁶) has developed crop calendar data using MODIS products for several countries at the national and sub-national scales. Furthermore, Jägermeyr *et al.*¹⁷⁷ created a gridded dataset of crop calendars for the Global Gridded Crop Model Intercomparison (GGCMI) at a 0.5° spatial resolution. The datasets were generated by combining information from nine observational sources at 0.5° land grid cells for 18 different crops, distinguishing between rainfed and irrigated systems. The dataset includes information on planting day, maturity day, growing season length, primary data source, and the fraction of harvested area. The GGCMI datasets are produced and validated using multiple sources and are gridded products that can be readily used for modeling purposes. We utilized the GGCMI for our study by extracting the MRB region from the global database.

Irrigated area and Irrigation water use. Irrigation consumes a significant portion of global water withdrawals, accounting for $\sim 70\%$ of total human water use¹⁷⁸. This is particularly relevant for food security, as a significant portion of global food production (33–40%) is derived from irrigated cropland¹⁷⁹. Therefore,

understanding the spatial distribution of irrigation is essential for managing water resources and ensuring food security; this is crucial in the MRB in light of the growing impacts of climate change and dams on fishery systems (e.g., Sabo *et al.*⁷⁷; Ziv *et al.*⁸; Veldkamp *et al.*¹⁸⁰) and the potential need for irrigation expansion^{76,181}.

However, there are no specific datasets on irrigated areas and irrigation water use for the MRB. As a result, studies on MRB have relied on globally available datasets. The latest version of global maps of irrigated areas provided by Food and Agriculture Organization (FAO) and developed by Siebert *et al.*¹⁸² are available at 5 arc-minutes which has been widely used globally to identify the irrigated area and irrigation water use. Moreover, several studies combined various datasets to generate global maps on irrigated areas (e. g. Zabel *et al.*¹⁸³; Salmon *et al.*¹⁸⁴; Meier *et al.*¹⁸⁵). Among different global products, FAO¹⁸² based irrigated area and irrigation water use data are commonly used in the Mekong. Therefore, we have selected FAO data for this study. However, datasets for Cambodia are missing in global databases. Thus, we acquired the Cambodian census data and subsequently processed these. As a result, the gap in global datasets is filled by processed census data for Cambodia. Though, similar processing can be performed for other countries, however census data is not easily accessible for those countries. Furthermore, ongoing agricultural census surveys in other parts of the basin will be extremely valuable for the research and policy makers.

Ecological data. *Nutrients and sediment data.* The MRC provides some data on nutrients and sediment, but these datasets are even more sparse than streamflow data and are not freely available. Specifically, the data include Nitrite-Nitrate (NO₃-N), Total Phosphorous (TP), and Dissolved Oxygen (DO). The MRC Discharge Sediment Monitoring Project (DSMP; Koehnken¹⁸⁶) that started in 2009 monitors sediment data at certain locations in the downstream regions of the MRB¹⁸⁷. Sediment concentration estimates are also available from satellite remote sensing, developed by using empirical or physics-based approaches^{187,188}. Here, we present and examine the data from the MRC and identify various other data sources.

Wetland and inundation data. Accurate wetland datasets are crucial for research on climate change, biodiversity preservation, and the implementation of effective land use policies and wetland conservation strategies. Wetland related studies in the Mekong have primarily used global datasets that are based on satellite observations due to lack of basin wide in-situ data availability. For example, Cho and Qi⁷⁰ used multi-sensor approach to overcome limitations in detecting wetland inundations from 2014 to 2021 in Southeast Asia. Several studies have also identified wetlands in the MRB; however, these are limited to the Mekong Delta^{189,190}. On a global scale, Sustainable Wetlands Adaptation and Mitigation Program (SWAMP) wetland maps were produced by Gumbricht *et al.*¹⁹¹ which include the wetland categories identified by Ramsar (2013). Furthermore, Tootchi *et al.*¹⁹² identified global wetlands based on surface water imagery and groundwater constraints. In this study, we provide the comparative evaluation of wetland based on Gumbricht *et al.*¹⁹¹ and Tootchi *et al.*¹⁹² for MRB.

Several studies have used satellite products to generate inundation datasets globally^{193–195}. Here, we use the GIEMS-D15 (Global Inundation Extent from Multi-Satellites – Downscaled to 15 arc-seconds; Fluet-Chouinard *et al.*¹⁹³) dataset for inundation maps as the dataset were made available by the authors. Based on topographic indices, the GIEMS-D15 dataset was created by downscaling monthly inundation observations from multiple satellites over a 12-year period from 1993 to 2004^{194,195} to a finer grid resolution of 15 arc-second pixels (~500 m at the equator). However, inundation in the MRB—especially in its downstream regions—is strongly related to precipitation seasonality and flow regulations by dams rather than topography, therefore other methods such as normalized difference vegetation index based flood inundation¹⁹⁶ than downscaling the data to higher resolution could be more reliable in MRB. Nonetheless, in this study we present the GIEMS-D15 based inundation datasets for the MRB region.

GHG emission data. Studies on GHG emission in the MRB have focused primarily on emissions from rice cultivation in the Mekong Delta^{197,198}. Some studies have investigated alternate farming methods to reduce the GHG emissions in the Delta, but these are rather limited^{199–201}. Moreover, a handful of studies have also estimated GHG emissions from hydropower dams in MRB (e.g., Räsänen *et al.*³⁰; Shi *et al.*²⁰²; Wang *et al.*³⁶). These studies have produced certain GHG datasets, but a complete timeseries and for the entire MRB is lacking. Therefore, for basin wide studies, global GHG datasets have been used. Global Emissions Database for Global Atmospheric Research (EDGAR; Crippa *et al.*²⁰³) v4.3.2 is the primary and most reliable source among gridded GHG datasets. The EDGAR dataset compiles anthropogenic emissions data for CO₂, CH₄, and N₂O based on international statistics and emission factors. Moreover, country specific annual GHG datasets for CO₂, CH₄, and N₂O are also available from Ritchie *et al.*²⁰⁴ (OURWORLDINDATA: <https://ourworldindata.org/greenhouse-gas-emissions>). Here, we employed the EDGAR datasets to infer insights on GHG emissions, which is available at 0.1° (~10 km) spatial resolution and is comprehensive in terms of covering GHG emissions from local and global scales; we consider this dataset as a reliable alternative in the absence of local datasets^{205,206}.

Socio-economic data. Helping advance scientific research and inform science-based management decisions and policymaking for sustainable transboundary basin management requires not only biophysical data (e.g., water, climate, and nutrients), but also socioeconomic data. These data are crucial, for example, to better understand the interactions among climate, water, and societies and ensure food, energy, livelihoods, and water securities under climate change and growing human influence on water systems^{1,3,11}. In this study, we synthesize socio-economic data for the four LMRB countries (i.e., Cambodia, Laos, Thailand, and Vietnam), which are obtained from various sources including government websites, the National Institute of Statistics for Cambodia, and the Lao Statistics Bureau for Laos. We further combined these datasets with those available from public repositories such as the OpenDevelopment Mekong (<https://opendevlopmentmekong.net/>) and the Socioeconomic Data and Applications Center (SEDAC: <https://sedac.ciesin.columbia.edu/data/sets/browse>). These datasets cover

S. No.	Countries »	Cambodia	Laos	Thailand	Vietnam	All
1	National sources »	National Institute of Statistic ^a	Lao Statistics Bureau ^b	National Statistics Office (NSO) ^c	General Statistics Office ^d	—
2	Demographics	General Population Census: 1998, 2008, 2019	Lao Population & Housing Census: 2015 ^e	See link to NSO	Population: 2011 to 2021	—
		Demographic & Health Survey: 2010, 2021			Population & Housing Census: 2011 to 2021	
		Intercensal Population Survey: 2004, 2013			Health, Culture, Sport, Living standards, Social order, Safety and Environment: 1995 to 2017	
3	Agriculture/Fishing	Census of Agriculture: 2013, 2019	—	See link to NSO	Agriculture, Forestry, Fishery: 2000 to 2021	—
		Intercensal Agriculture Survey: 2019	Rural, Agriculture, and Fishery Census			
4	Environment	—	SDG ^f	See link to NSO	Land & Climate ^g	—
5	Economics	Economic Census: 2011	—	See link to NSO	Employment: 2000 – 2021	—
		Socioeconomic Survey: 2004, 2007, 2009, 2010, 2012, 2013, 2014, 2015, 2016, 2017, 2019				
		Labor Force Survey: 2019				
6	Satellite Imagery Based	—	—	See link to NSO	—	Nighttime Lights ^h
						Global Roads Data ⁱ
						GRIP Global Roads Data ^j
7	Other	OpenDevelopment Cambodia ^k	OpenDevelopment Laos ^l	OpenDevelopment Thailand ^m	OpenDevelopment Vietnam ⁿ	SEDAC ^o
						ASEAN Stats ^p
						APEC Key Economic Indicators ^q
						Global Internal Displacement Database ^r
						SDG Gateway Data ^s

Table 4. Details of socio-economic datasets synthesized. ^a<https://www.nis.gov.kh/index.php/km/> ^b<https://www.lsb.gov.la/> ^c<http://statbbi.nso.go.th/staticreport/page/sector/en/index.aspx> ^d<https://www.gso.gov.vn/en/homepage/> ^e<https://laosis.lsb.gov.la/tblInfo/TblInfoList.do?rootId=2101000&menuId=2101101&lang=en&keyword=&searchType=undefined> ^f<https://www.lsb.gov.la/sdg/en/> ^g<https://www.gso.gov.vn/en/administrative-unit-land-and-climate/> ^h<https://www.earthdata.nasa.gov/learn/backgrounders/nighttime-lights> ⁱ<https://sedac.ciesin.columbia.edu/data/set/groads-global-roads-open-access-v1> ^j<https://www.globio.info/download-grip-dataset> ^k<https://opendevolucioncambodia.net/> ^l<https://laos.opendevolucionmekong.net/> ^m<https://thailand.opendevolucionmekong.net/> ⁿ<https://vietnam.opendevolucionmekong.net/> ^o<https://sedac.ciesin.columbia.edu/data/sets/browse> ^p<https://data.aseanstats.org/> ^qhttps://statistics.apec.org/index.php/key_indicator/index ^r<https://www.internal-displacement.org/database/displacement-data> ^s<https://dataexplorer.unescap.org/>.

a range of attributes including population demographics, agriculture, gross domestic product, housing, forestry, fishery, road networks, and internal displacement. However, these data are often limited in terms of spatial and temporal coverage, as detailed in Table 4.

High resolution gridded population and Gross Domestic Product (GDP) data are key to understanding and better predicting exposure and vulnerability of socioeconomic activities to future climate extremes and developing improved adaptation and mitigation strategies²⁰⁷. Gridded population of the World (GPWv4; Doxsey-Whitfield *et al.*²⁰⁸) datasets have been extensively used in socioeconomic and environmental studies, such as vulnerability mapping, disaster impacts, and health implications of environmental change^{209–211}. However, for our study we used GPWv4 population datasets at 30-arc second (~1 km) spatial resolution and projected population datasets from SEDAC²¹² at 1/8th degree spatial resolution. Furthermore, we utilized gridded GDP data from Kumm *et al.*²⁴ at 10-year interval and gridded GDP projections datasets from Wang and Sun²⁰⁷ which are consistent with the shared socioeconomic pathways (SSPs). We further provide a comparison between population and GDP projections for all six Mekong countries.

Data Records

The synthesized datasets are available in the Zenodo repository²¹³ (<https://zenodo.org/record/7803254>). The uploaded datasets are optimized considering user convenience and data size reduction. For example, EM-Earth precipitation and temperature, GLEAM ET, GHG emissions, digitized groundwater, population projections, GDP projections, ground observations of soil moisture, and digitized streamflow datasets are provided in text format. The EM-Earth precipitation and temperature, GLEAM ET, and GHG emissions are gridded datasets with first two rows as locations (longitude and latitude), initial columns as time series (e.g., year, year-month, year-month-day), and rest of the columns as data time series. Moreover, first two columns of populations and GDP projection datasets are gridded locations (longitude and latitude) and the rest of the columns show data for

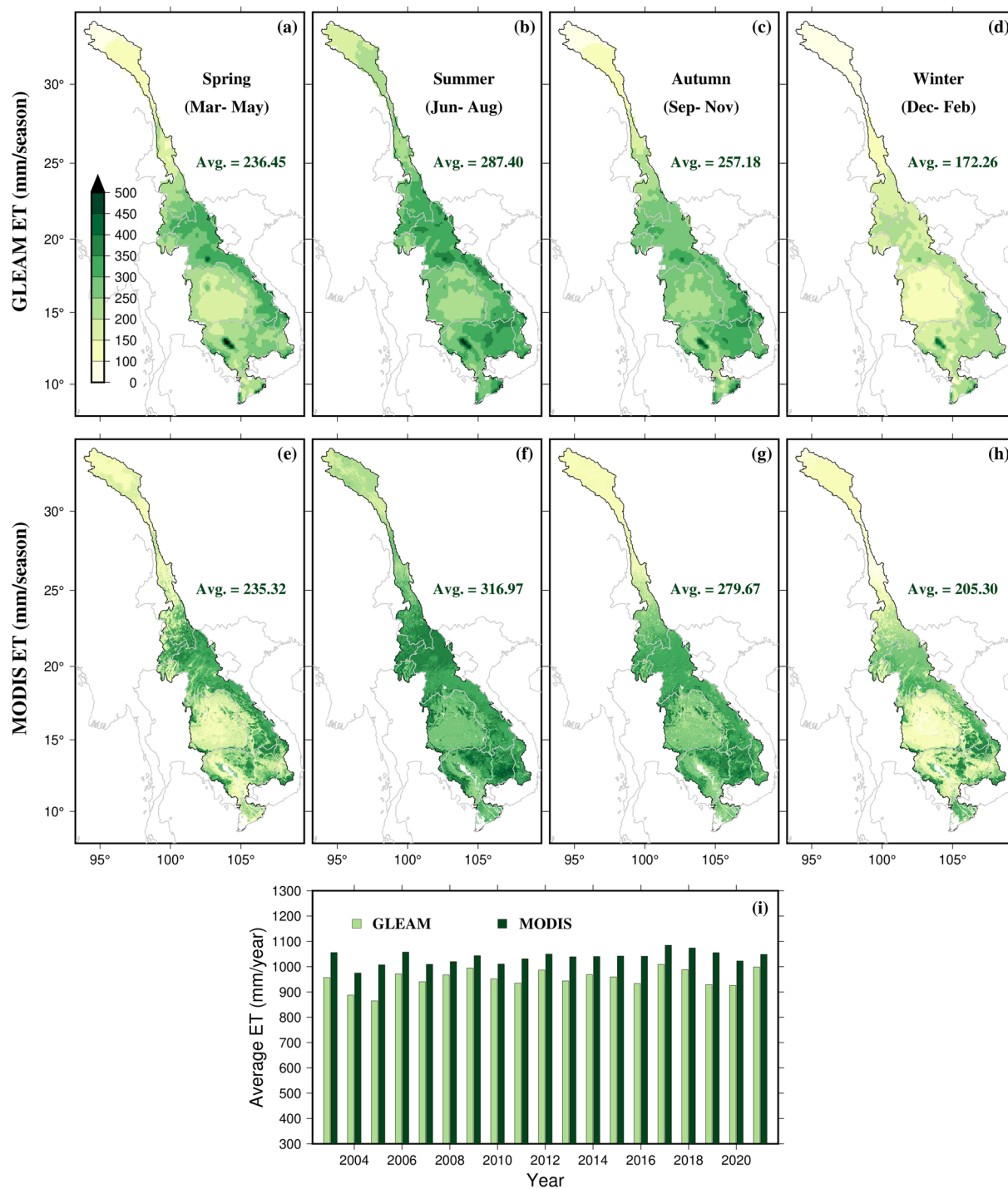


Fig. 2 (a–d) Spatial distribution of seasonal ET from GLEAM data, (e–h) same as (a–d) but for MODIS ET, “Ave.” indicates the basin-averaged ET (mm/season), and (i) comparison of basin-averaged annual GLEAM and MODIS ET (mm/year).

base year or projected years. Digitized groundwater, ground observations of soil moisture, and digitized stream-flow data files contain time series in the initial columns and followed by corresponding data in the last column. Data on crop yield, which is country-level annual data, is presented with year in the first column and crop types in the first row; dam attributes (first row) are stored in excel files. GeoTIFF image format is utilized for MODIS ET, irrigated area and irrigation water use, LULC, population, GDP, surface water, and wetland datasets. Soil moisture datasets are stored in MATLAB (.mat) files. Each data folder includes a “Readme” file that provides detailed data description, including the original source, where relevant.

The publicly available datasets such as, satellite precipitation and temperature, ET, surface water, satellite soil moisture, LULC, crop yield, crop calendar, wetlands, GHG, and socio-economic datasets are freely available for download

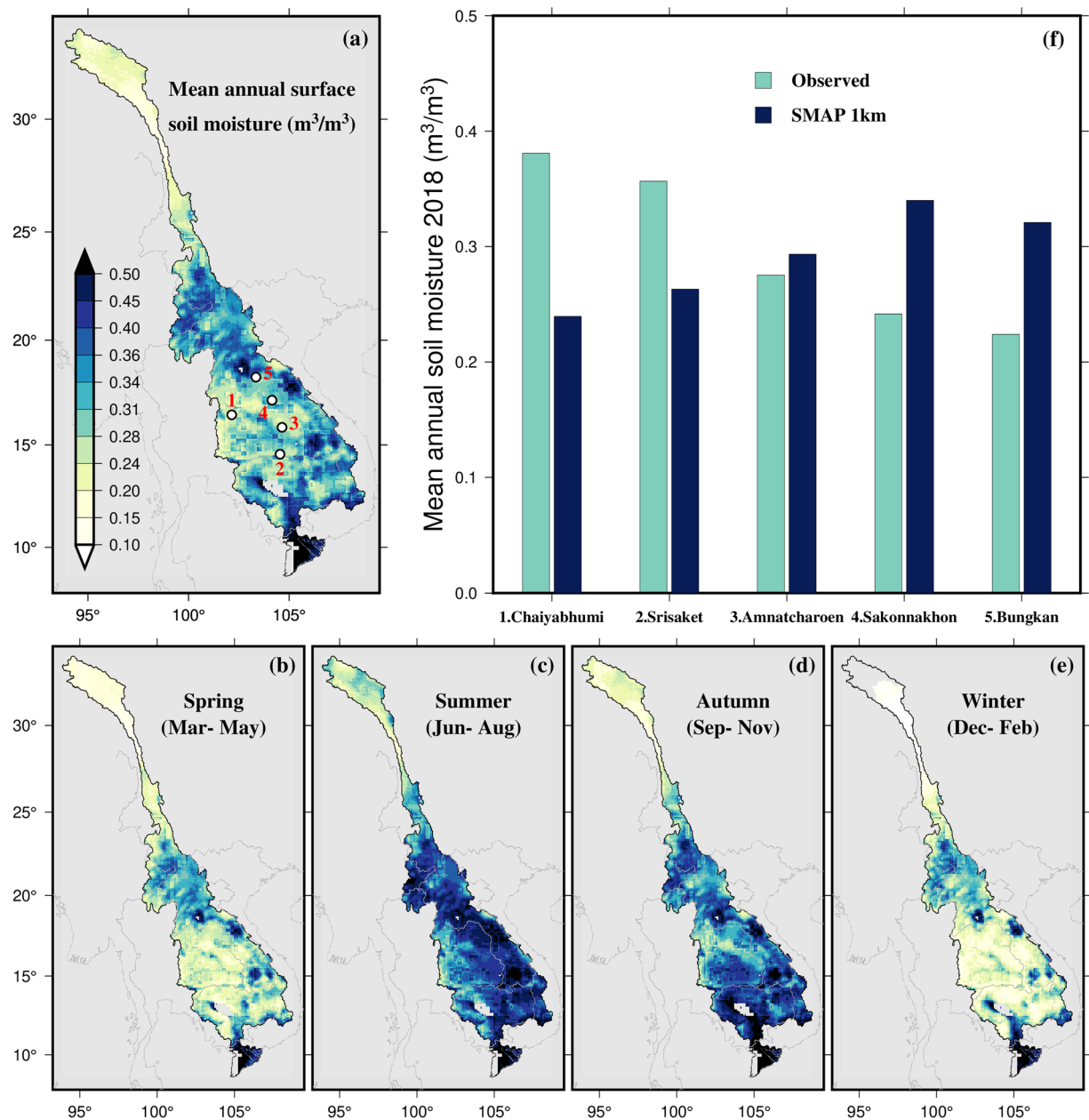


Fig. 3 (a) Spatial distribution of downscaled SMAP soil moisture data at 1 km spatial resolution for 2016–2021 period. (b–e) Seasonal variation in the spatial distribution of SMAP data. The grey color indicates ‘no data’. (f) Comparison of SMAP data with *in-situ* observation at five locations in Thailand (white circle in panel a).

from the original sources. Additionally, *in-situ* datasets from the MRC, including precipitation, temperature, wind speed, sunshine hours, specific humidity, streamflow, water level, nutrients, and sediment can be obtained through formal Procedure for Data and Information Exchange and Sharing (PDIES); these data are open to member countries of LMRB and to certain extent the MRC stakeholders²¹³.

Technical validation

Meteorological data. Among various hydrometeorological datasets identified in methods section, we find the EM-Earth data to be i) relatively inclusive of most climate variables required for analyses and modeling and ii) of reasonable spatial resolution. The dataset also includes multiple ensemble members useful for uncertainty quantification. Thus, we present an analysis of this product, focusing on precipitation and temperature (Fig. 1 and Fig. S1), the two variables of primary interest in many hydrological and ecological studies. Among EM-Earth, APHRODITE, TRMM, IMERGE, Princeton (He *et al.*¹⁰⁰), and ERA5 precipitation datasets, EM-Earth data show better results when compared against gauge-based data from the MRC at selected locations, except at Kratie (Fig. S1), indicating high accuracy of the ensemble-mean EM-Earth data. Substantial spatial heterogeneity can be seen in precipitation (Fig. 1a) and temperature exhibits a strong north-south gradient (Fig. 1d). In Laos, Vietnam, and the eastern half of Cambodia, annual precipitation is higher compared to other regions in the

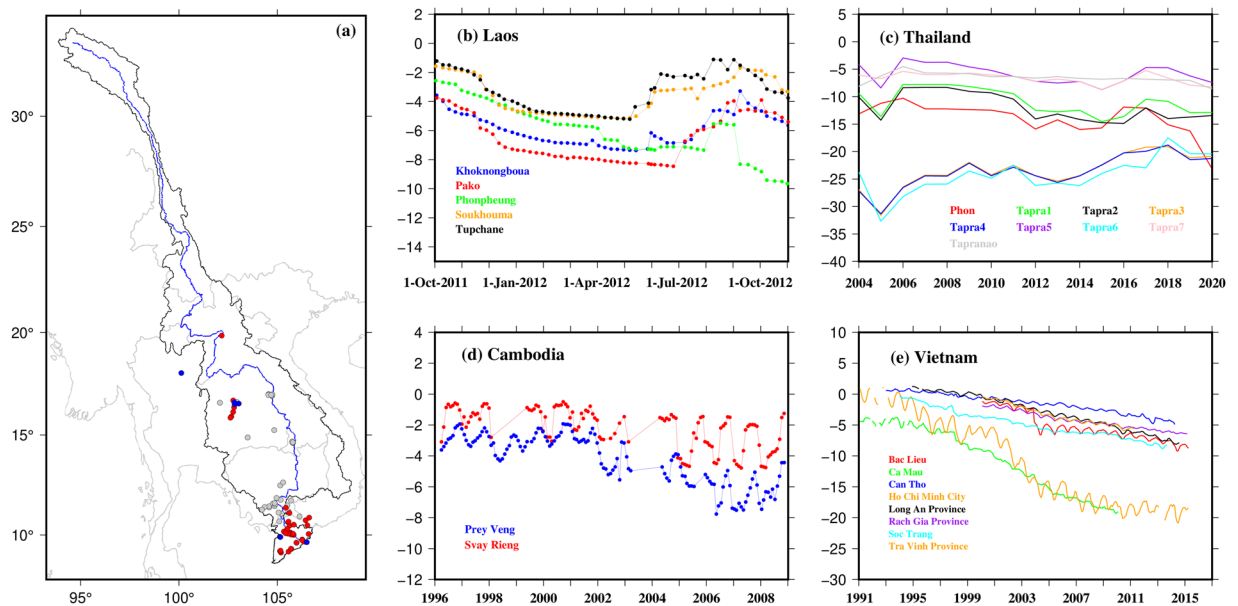


Fig. 4 (a) Locations of groundwater observation wells at which data are published in the literature. Red and blue colors indicate declining and increasing trends, respectively, in annual groundwater, whereas the grey color indicates either no significant change or groundwater data are available for less than 5 years. We used Mann-Kendall with Sen Slope method to calculate the trend. The time period of available datasets along with trend significance is provided in Table S1, (b) daily discrete groundwater depth at 5 locations in Laos from 01-Oct-2011 to 31-Oct-2012, digitized from Vote *et al.*¹⁵⁵, (c) yearly groundwater depth at 7 locations in Thailand from 2004 to 2020, digitized from Muenratch *et al.*⁵⁵, (d) daily discrete GW depth from 1996 to 2008 at 2 locations in Cambodia, digitized from Johnston *et al.*¹⁴², and (e) monthly GW depth from 1991–2015 at 7 location in Mekong Delta, Vietnam, digitized from Minderhoud *et al.*⁴⁶.

MRB (Fig. 1a). A higher mean annual temperature in Cambodia, Thailand, and the Mekong Delta was found compared to other parts of the basin (Fig. 1d).

Additionally, we compare the spatial patterns of precipitation and mean temperature for three different datasets: EM-Earth ensemble mean, ERA5, and APHRODITE (Fig. 1a–f), revealing interesting patterns and tendencies. The APHRODITE precipitation was comparatively lower than the other two products in Laos, but the three temperature products display similar spatial patterns across the basin. This suggests that while there may be some differences in the precipitation data, temperature data are more consistent across different sources.

Hydrological data. *Streamflow and water level.* We evaluate the availability and trends in digitized and MRC-based streamflow and water level at various locations in the MRB. In terms of streamflow and water level data, there are more stations with positive trends than with negative trends (Fig. S2a, b). The alternate positive and negative trends in the streamflow and water level data could be due to seasonal shift in water availability in the streams and different time-period considered to evaluate the trend based on data availability (Fig. S2a, b; Table S2, S3). Moreover, we present the seasonal cycle of streamflow and water level at 8 selected stations across the basin. We find that at all the locations streamflow and water level start increasing from May and peak in August or September (except for Changdu which is peaking in July), following the monsoonal rainfall patterns.

ET. We compare two ET datasets based on GLEAMv3.6b and MOD16A2GF for four seasons, finding that both datasets show similar spatial pattern across the basin (Fig. 2). Upon analyzing the seasonal pattern in both datasets, it is found that summer season had the highest ET, which is consistent with the seasonal precipitation patterns (Fig. 2a–d). Additionally, we observe that the spatial patterns of ET vary depending on the location within the MRB basin. However, with the exception of the spring season, the seasonal and annual MODIS ET is comparatively higher than GLEAM ET (Fig. 2). Finally, our investigation reveals that the mean annual ET for both datasets for the entire MRB basin exhibit similar increasing trend over time (Fig. 2i). Increased ET in the basin may change the percentage of precipitation that becomes surface water runoff or subsurface recharge which may affect the groundwater levels, groundwater surface water interactions, and soil moisture^{214,215}.

Surface water. Fig. S3a shows the long-term occurrence of surface water in the MRB based on the data from Ji *et al.*¹²⁸. We highlight two regions of particular interest: one mainly featuring multiple reservoirs, and the other featuring the TSL and Mekong Delta (Fig. S3). For these two regions we compare the surface water based on JRC and Ji *et al.*¹²⁸ (Fig. S3b–e), finding that the data from Ji *et al.*¹²⁸ show lesser extent, and also lower occurrence especially in the Mekong Delta compared to JRC data. Such surface water datasets are crucial for a wide range of studies in the MRB, including for model evaluation and studies on ecological, agricultural, fisheries, and livelihood changes, especially in relation to upstream dam construction. Many recent studies have used these datasets

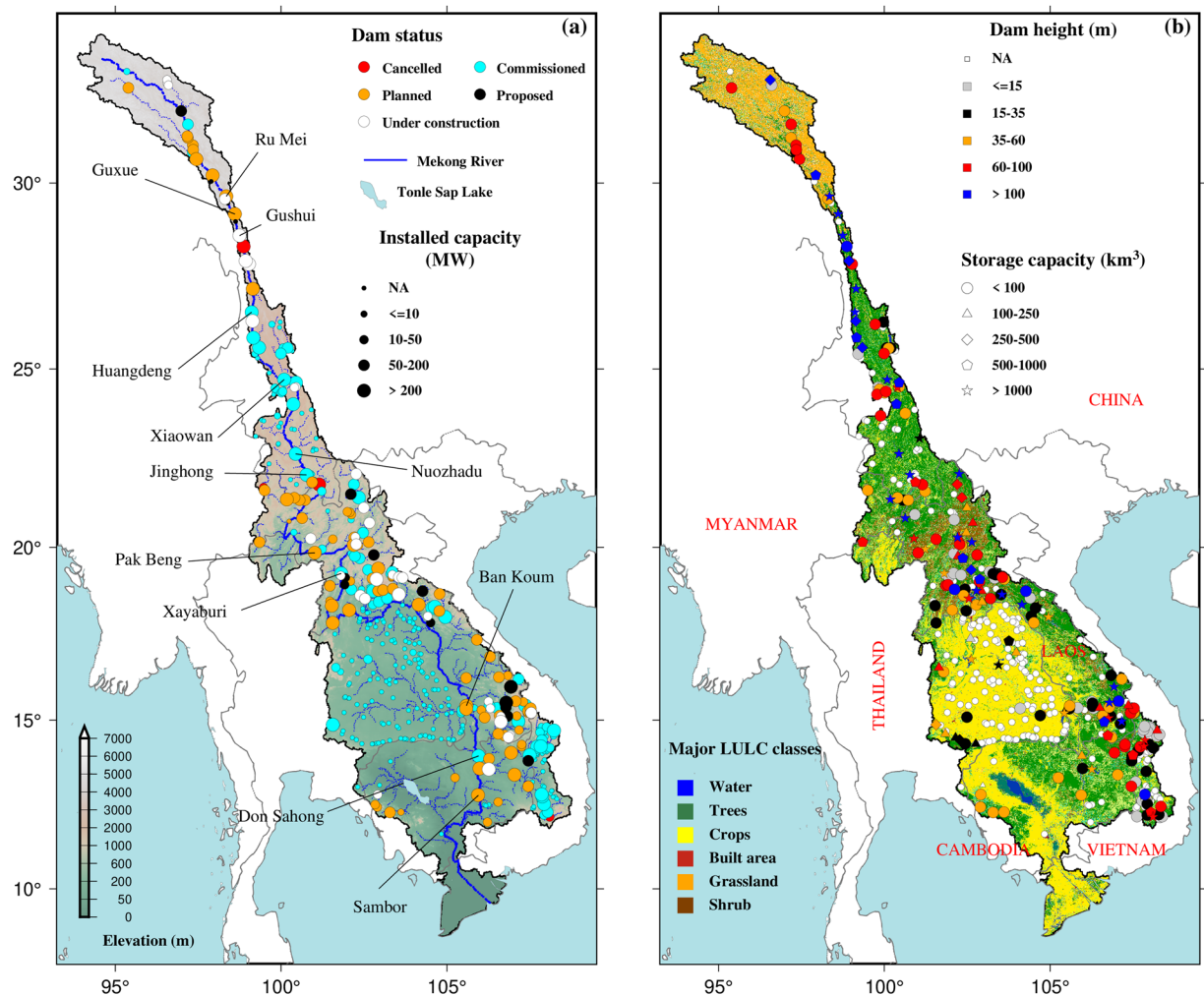


Fig. 5 (a) Locations of dams in the Mekong River basin. The background image shows elevation (source: Shuttle Radar Topography Mission (SRTM)). Selected large dams—commissioned, under construction, or planned—are labeled. Dam status is color-coded whereas the installed capacity (megawatts) is marked by circle size; NA denotes “data not available”. (b) Selected dams, categorized based on dam height and storage capacity; the background image shows land use and land cover (LULC) classes for the year 2020 obtained from ESA-CCI.

to examine the changing inundation patterns around TSL and Mekong Delta due to climate variability and dam construction^{9,29,47}. However, these datasets—mostly satellite based—provide limited information on the changes in surface water such as long-term occurrence or changes in permanent water bodies. Therefore, there are opportunities to develop more accurate and reliable surface water datasets, for example by using information from future satellite missions or improved modeling approaches.

Soil moisture. We chose the soil moisture data from Fang *et al.*¹³³ for this study due to its good accuracy and relatively high resolution after reviewing numerous soil moisture products available for the MRB region (discussed in methods section). We analyze the spatial variation of mean annual soil moisture across the MRB and compare downscaled soil moisture data with ground data at 5 locations in Thailand. The mean annual surface (up to 5 cm) soil moisture content in Laos and Vietnam is higher than in other parts of the MRB (Fig. 3a). Mekong Delta in Vietnam and the flood plains in Cambodia show higher soil moisture content. Similarly, the southern parts in the Chinese portion of the MRB, northern Laos, and the subsequent Thailand portions show higher soil moisture. Soil moisture levels are lower in Thailand and some areas of Cambodia that are primarily agricultural (Fig. 3a). A comparison with observed soil moisture at five locations in Thailand suggests that SMAP captures soil moisture content reasonably well (Fig. 3b).

We also analyze the seasonal patterns in the soil moisture data, which reveals that soil moisture is generally higher in the summer and autumn seasons compared to spring and winter (Fig. 3c–f). This pattern is consistent with the typical rainy season in the MRB region, which occurs during the summer and autumn months and results in increased soil moisture levels. We, however, note that this pattern could vary in different regions with different climate patterns that govern seasonal rainfall.

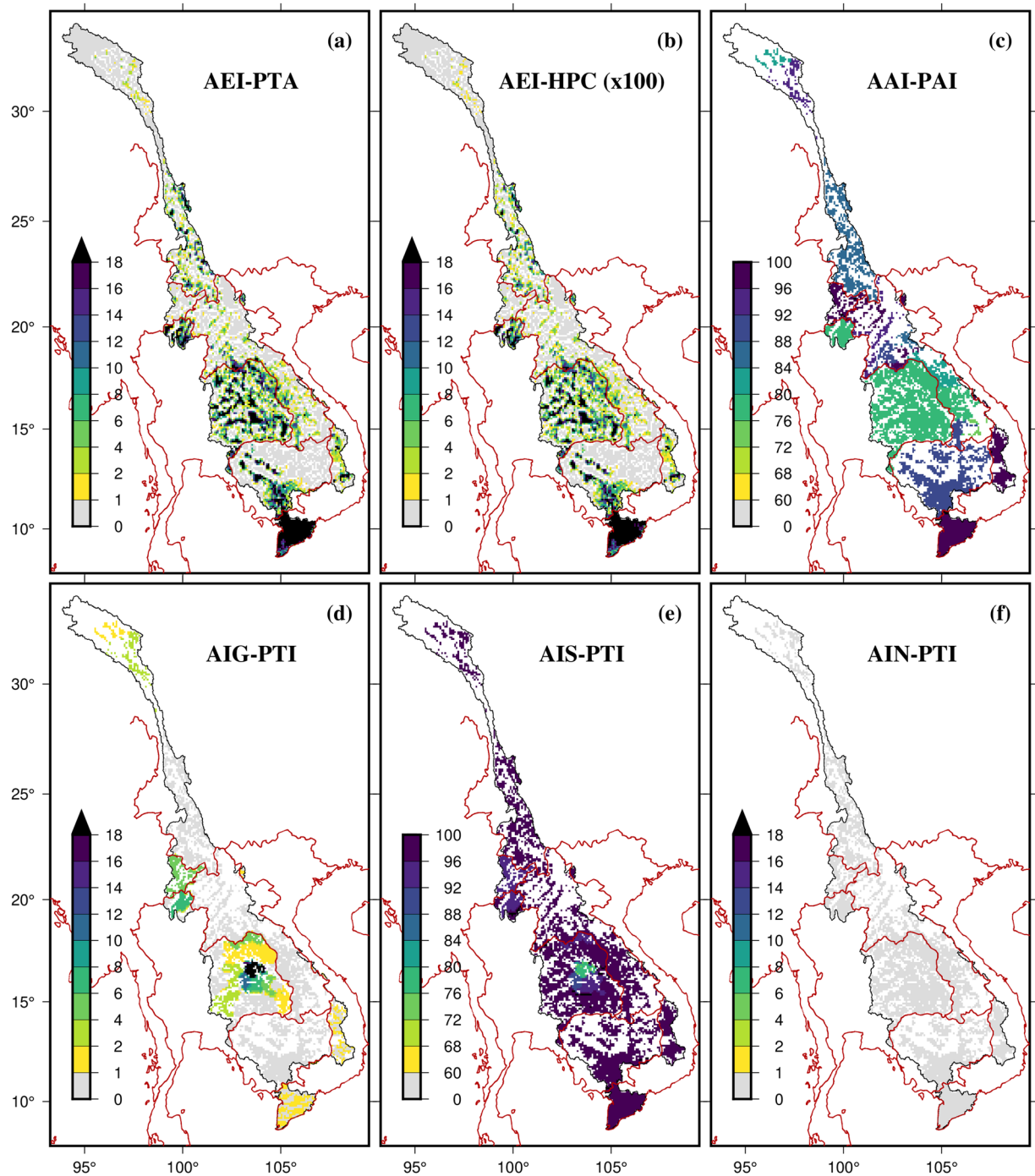


Fig. 6 Spatial distribution of (a) area equipped for irrigation expressed as a percentage of total area (grid resolution ~ 10 km) (AEI-PTA), (b) area equipped for irrigation expressed in hectares per cell (AEI-HPC), (c) area actually irrigated expressed as a percentage of area equipped for irrigation (AAI-PAI), (d) area irrigated with groundwater expressed as a percentage of total area equipped for irrigation (AIG-PTI), (e) area irrigated with surface water expressed as a percentage of total area equipped for irrigation (AIS-PTI), (f) area irrigated with water from non-conventional sources expressed as a percentage of total area equipped for irrigation (AIN-PTI).

Groundwater. Groundwater anomalies can be estimated by subtracting the modeled surface water anomalies (e.g., obtained from Global Land Data Assimilation System (GLDAS)) from terrestrial water storage (TWS) anomalies derived from GRACE satellite observations^{216,217}. However, the spatial resolution of GRACE data is low, and surface water from GLDAS contains high uncertainty, for example because of missing human interventions. Moreover, observed groundwater datasets for the MRB are not publicly available. Therefore, we present the digitized groundwater data from a series of published literature (Fig. 4). The data consists of temporal measurements of groundwater levels at various locations within the MRB, including daily, monthly, yearly, and

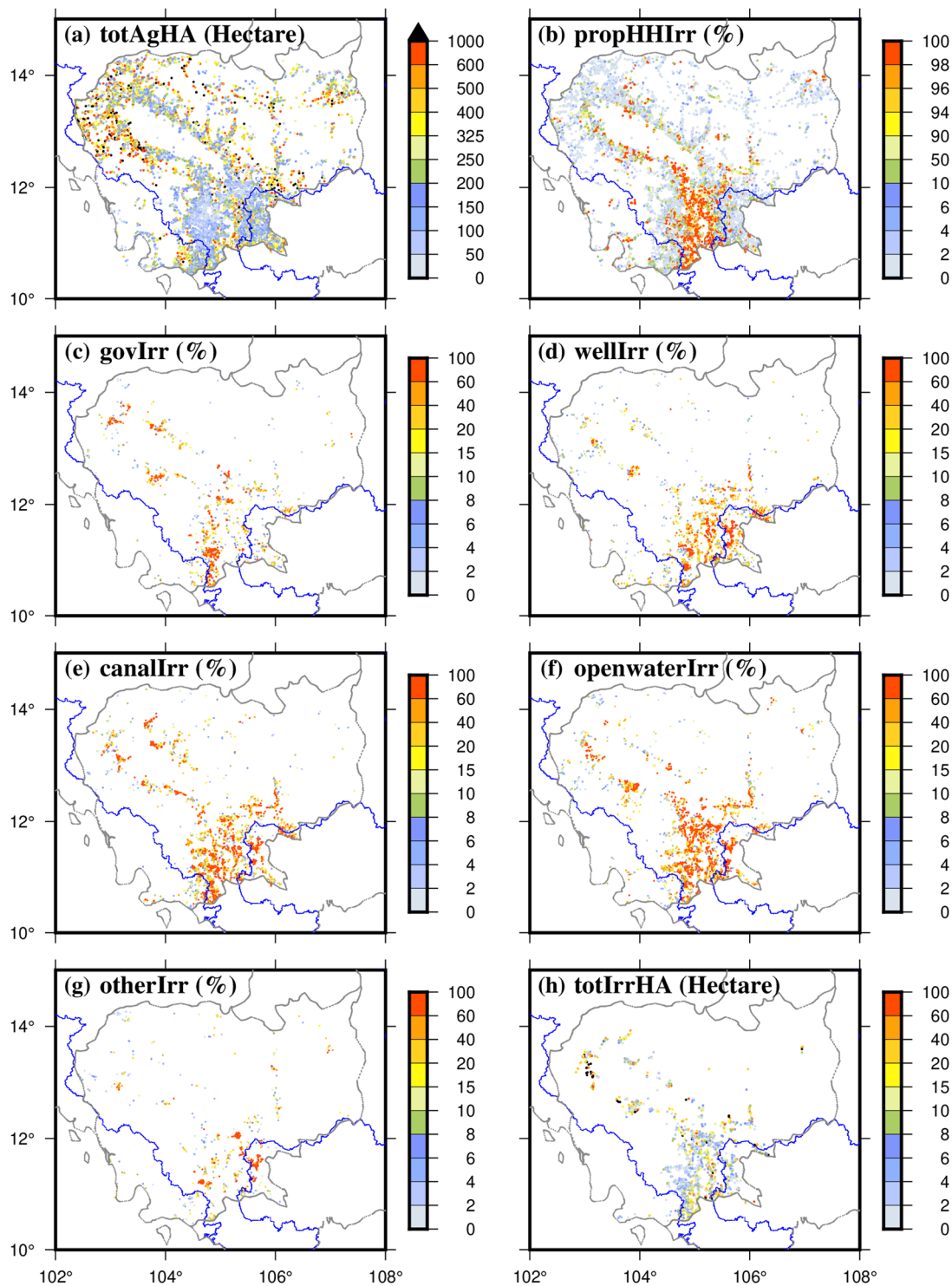


Fig. 7 Spatial distribution of irrigated area and irrigation water use datasets collected for Cambodia, (a) total hectares in agricultural use (temporary, permanent, meadow/pasture) by village (village size varies) (totAgHA), (b) proportion of households by village indicating some use of irrigation (propHHIrr), (c) proportion of parcels by village using government owned irrigation (govIrr), (d) proportion of parcels by village using well irrigation (wellIrr), (e) proportion of parcels by village using canal irrigation (canallIrr), (f) proportion of parcels by village using open water irrigation (openwaterIrr), (g) proportion of parcels by village using other types of irrigation (otherIrr), (h) total acreage of irrigated land by the village for all type of irrigation (totIrrHA).

seasonal cycles. The highest density of data points was found in the Mekong Delta region, encompassing parts of Vietnam and Cambodia. General examination of the digitized data reveals declining groundwater levels within the MRB, with the most pronounced decreases occurring in the Mekong Delta (Fig. 4). This decreasing

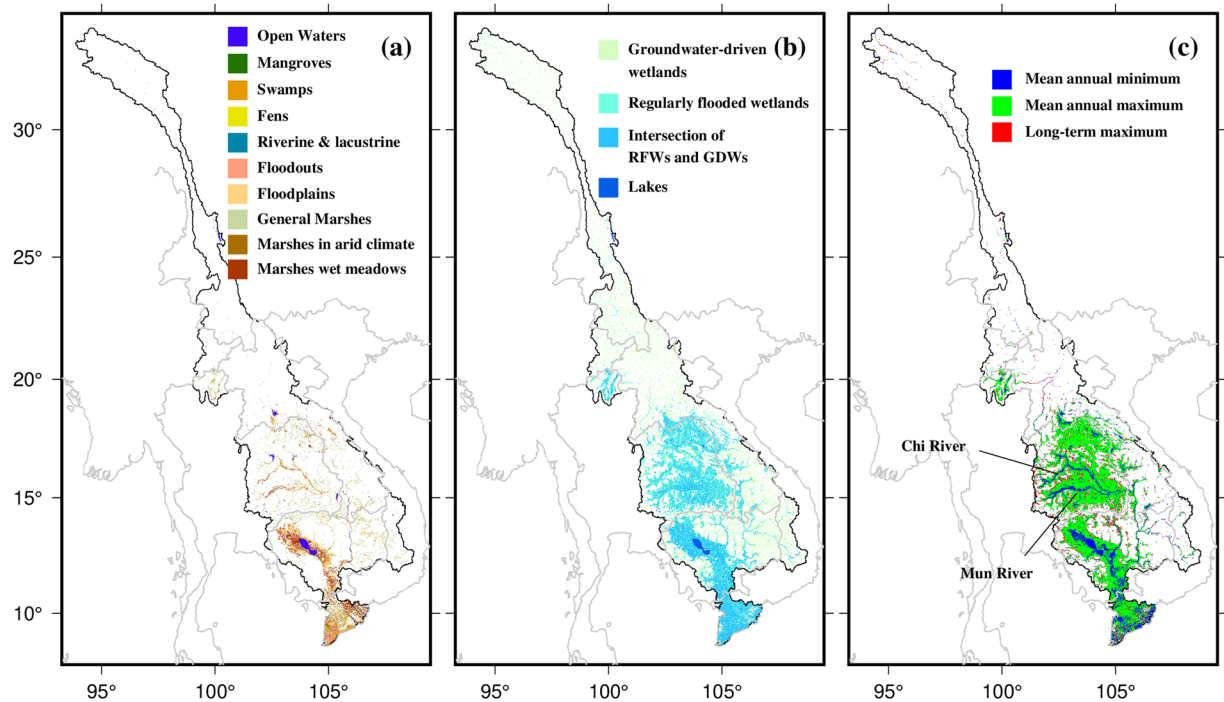


Fig. 8 Wetland maps based on (a) Gumbricht *et al.*¹⁹¹ and (b) Tootchi *et al.*¹⁹², (c) GIEMS-D15 map (Fluet-Chouinard *et al.*¹⁹³) inundation extent at 15 arc-second spatial resolution over the basin, with mean annual minimum, mean annual maximum, and long-term maximum inundation area.

trend in groundwater is likely influenced by the high level of groundwater pumping for agricultural purposes in both Vietnam and Cambodia^{46,137,146,218}. In addition, the extraction of groundwater for agricultural and domestic use has been linked to subsidence in the Mekong Delta^{11,46}. Given these findings, improved management and conservation efforts will be necessary to ensure the sustainable use of groundwater resources in the MRB, particularly in the Mekong Delta region.

Dam data. Figure 5 depicts the synthesized and corrected dam datasets (see methods section) for the MRB. A more detailed information for the dams selected from the database for hydrological modeling purpose is created, which includes information on dam height, reservoir storage capacity, and reservoir purpose, among others (Table S4). Here we present selected dam attributes such as dam status, installed capacity, dam height, and reservoir storage capacity (Fig. 5). In the past decade, ~100 dams have been constructed in the MRB²⁵, with several more currently being planned or under construction, particularly in China, Laos, and Cambodia. The construction of large dams such as Ru Mei, Guxue, Gushui, and Huangdeng in the UMRB (Fig. 5a) has sparked environmental concerns such as the decline in the flood season river flow and annual sediment flux, and water quality deterioration in reservoirs within China²¹⁹. Additionally, the construction of large dams such as Xayaburi, Nuozechadu, and Don Sahong has led to the trapping of sediment flow and disruption of fisheries, raising significant ecological concerns^{1,77,220}. Therefore, the dam database is expected to be useful in hydrological, ecological, and socio-economic modelling and consequently future planning and management.

Land use and crop data. The land use data for the MRB obtained from ESA-CCI were analyzed from 1992 to 2020. Out of eleven land use classes, cropland, tree cover, mosaic tree and shrub, shrubland, and grassland are dominant (Fig. S4). We select two regions in the upper and lower basin for a more in-depth examination (Fig. S4). We found that the LMRB is experiencing a significant increase in crop coverage. In contrast, in the UMRB, there is a substantial increase in tree cover (Fig. S4). In the upper region (region b), there is a slight increase in cropland but a significant increase in tree cover. Tree cover increased primarily from 1996 to 2000, compensating for the loss in shrubland. However, in the lower region (region c), cropland increased substantially with a corresponding decline in tree cover area. Urban areas in region “c” also increased considerably compared to region “b”. Overall, cropland areas in the basin rose steadily until 2012, but started declining since then (Fig. S4c).

To gain further insight into crop dynamics in the LMRB, we analyze data on crop yields for rice, maize, bananas, and sugarcane—the major crops grown in the region—for four LMRB countries (Fig. S5). Our analysis revealed that rice yield has been increasing in all countries, particularly after 1990. Vietnam exhibited the highest rice yield among the four countries. Similarly, maize yield has been increasing across all countries, with Laos exhibiting the highest yield. While banana and sugarcane yields have been decreasing in Cambodia, they have been increasing in the remaining three countries. These data, and the interesting patterns therein, could be useful for studies on water and food security issues in the MRB; however, these datasets are available only at the country level, hence cannot be used for basin-scale analyses or modeling. Nevertheless, the datasets

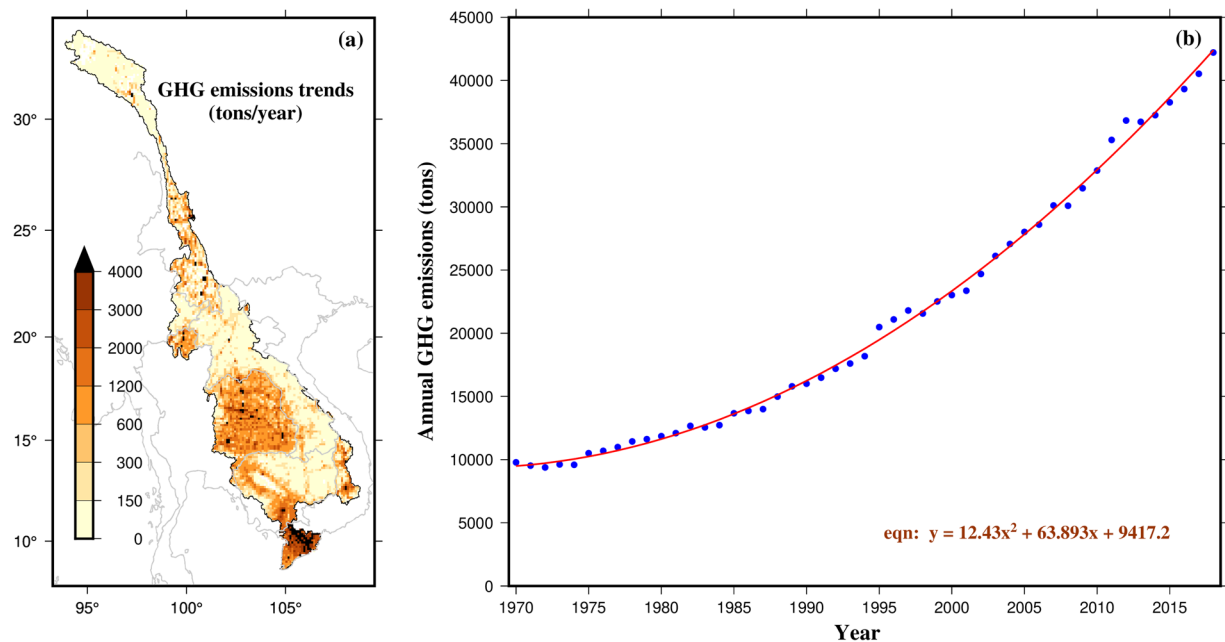


Fig. 9 (a) Spatial distribution of GHG (CH_4 , CO_2 , and N_2O) emission rate in tons/year at 0.1° resolution from year 1970–2018, (b) timeseries GHG emissions per year from 1970–2018 and trendline fitted with quadratic equation.

could be used to derive grid-based products through combination with other datasets such as on cropland areas (e.g., Burbano *et al.*⁷⁶).

Irrigated area and irrigation water use. The areas equipped for irrigation in the MRB mainly ranges from 0 to 20% of the total grid cell area at 0.083333° (~ 10 km) spatial resolution (Fig. 6a). The Mekong Delta, the flood plains in Cambodia, Thailand, southern part of UMRB, and some portions in Laos are the main areas that are intensively irrigated (Fig. 6b). Results suggest that Vietnam and North Laos portions are more irrigated compared to Cambodia, Thailand, and Southern portion of Laos (Fig. 6c). Except for some portions in Thailand and the Mekong Delta, which are irrigated with groundwater, rest of the basin is irrigated heavily by surface water (Fig. 6d, e). In addition to the conventional sources of irrigation, the use of non-conventional methods for irrigation is extremely limited, as demonstrated in the data presented in Fig. 6f. As the demand for agricultural products from the LMRB is projected to rise by 20–50% in the coming 30 years due to the growing global population²²¹, there is a growing risk of food and water insecurity in the basin. To address this issue, it is important to better understand where irrigation is currently happening, what the implications on water and food systems are, and how future irrigation expansion could affect sustainable water use.

To fill irrigation data gaps for Cambodia, we obtained and examined the data from the Cambodian census database, especially focusing on the spatial patterns of irrigation practices (Fig. 7). Results indicate that the highest agricultural land utilization in Cambodia is located in the proximity of the TSL and in flood plain zones (Fig. 7a). Moreover, a higher density of irrigation infrastructure was observed in the flood plain zones when compared to other regions, which was also found by Park *et al.*¹⁴⁶ (Fig. 7b). Furthermore, the census data also provided insights on the irrigation systems owned and operated by the government. It can be observed that government-owned irrigation systems were less prevalent in comparison to other irrigation practices such as wells, canals, and open water (Fig. 7c–e). Other than well, canal, and open water irrigation, rest of the irrigation sources are insignificant in MRB (Fig. 7g). The total irrigated area data looks incomplete since it is difficult to reconcile responses here with propHHIrr (many NaN values here) (Fig. 7h). Similar census data for irrigation purposes for the other LMRB countries would help in better understanding and modeling the changes in irrigation water use, however these datasets are currently inaccessible.

Ecological data. Nutrients and sediment data. We obtained nutrient datasets (DO, $\text{NO}_3\text{-N}$, and TP) from MRC for our study in MRB as mentioned in methods section. The locations of stations for which data on nutrients (specifically DO, $\text{NO}_3\text{-N}$, and TP) are obtained from the MRC are shown in Fig. S6. We only selected locations for DO, $\text{NO}_3\text{-N}$, and TP where the data was available from 1996–2021. Results indicate that except in Cambodia (near the TSL), DO exhibits a declining trend (Fig. S6a). Further, in the TSL region and Mekong Delta the $\text{NO}_3\text{-N}$ concentration has been increasing over time (Fig. S6b). However, TP is increasing across the entire basin except in the Mekong Delta (Fig. S6c). Contrary to the common belief, the construction of multiple dams in the upstream of MRB has increased nutrient concentration downstream^{52,222}, especially in Cambodia. Wang *et al.*²²³ also showed an increasing trend in total suspended solids in Cambodia between 2000 to 2018. However, DO and TP show a negative trend in the Mekong Delta. Moreover, nutrients in terms of DO and TP tend to have been discharging to inland water bodies (e.g., lakes) whereas the delivery of these nutrients to the Mekong Delta is declining (Fig. S6).

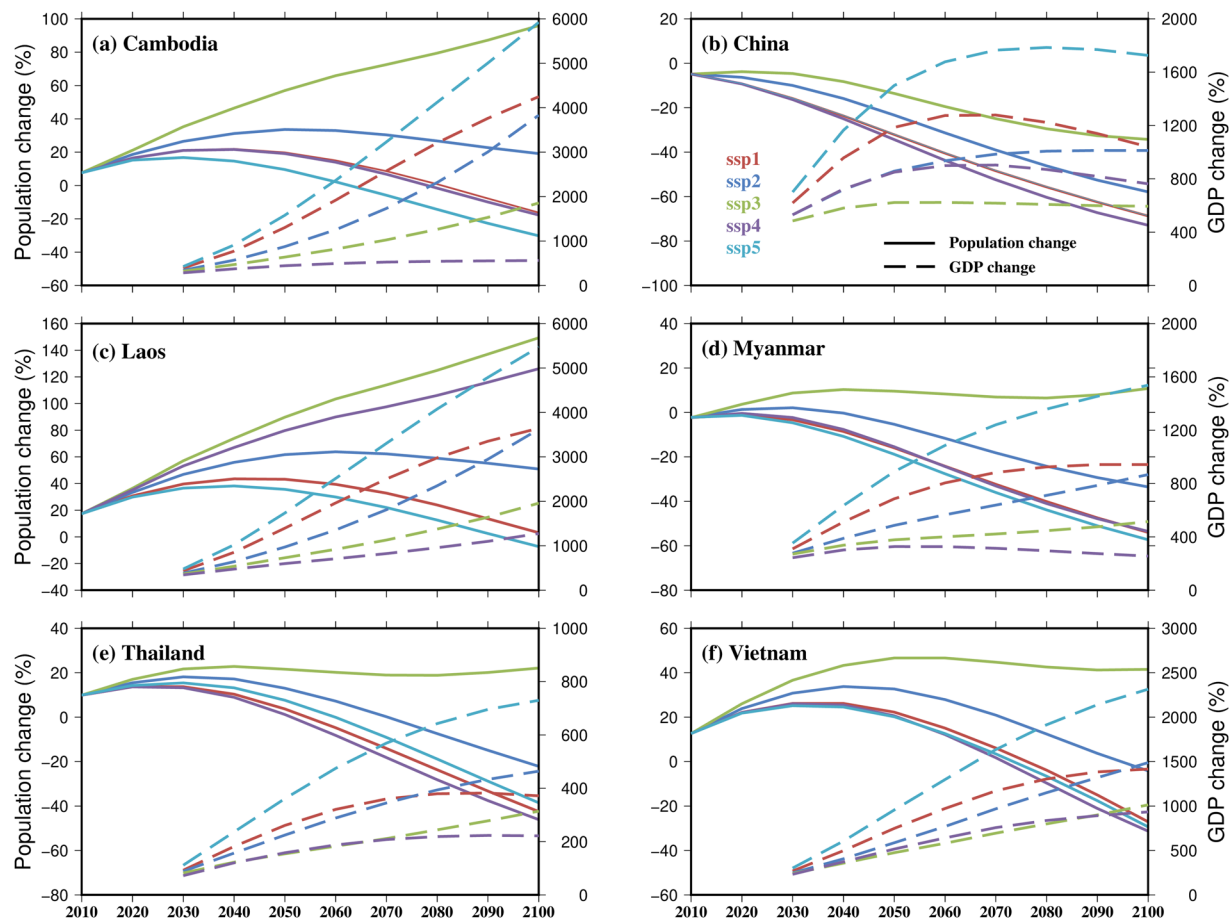


Fig. 10 Comparison of projections of population and GDP percentage change for the regions of Cambodia, China, Laos, Myanmar, Thailand, and Vietnam, which come under the MRB from 2010 to 2100. Population change is shown in left side of y-axis and GDP in terms of PPP (USD) change is shown in right side of y-axis with dashed lines.

Similarly, sediment concentration datasets were obtained from the MRC for 18 locations within the MRB (Table S5). However, the datasets are not continuous nor complete for all locations, hence a statistical analysis was not possible. Therefore, we conducted a visual inspection of the data at 10 locations, finding a decline in sediment concentration, particularly at locations within the mainstream Mekong (Figure S7). Only the Mae Suai dam site and Rasi Salai stations, which are not in the mainstream Mekong, show an increase in sediment concentration; the rest of the stations show a decline (Fig. S7). The reduction in sediment load implies downstream impacts including coastal erosion, reduced nutrient supply for aquatic species and agriculture, and land subsidence^{22,224}.

Wetland and inundation data. We provide a detailed comparison between the two selected wetland products: one based on the maps produced by Gumbricht *et al.*¹⁹¹ and the other from Tootchi *et al.*¹⁹² (Fig. 8a,b). Gumbricht *et al.*¹⁹¹ classified the wetlands mainly as open water, mangrove, swamps, fens, riverine and lacustrine, floodplains, and marshes based on geomorphology, moisture condition, and vegetation and soil condition. These wetlands are located mostly in the LMRB with the Mekong Delta housing many swamps, mangroves, and floodouts (Fig. 8a). On the other hand, floodplains of Cambodia and around the lake include open water, marshes, and meadows, etc. (Fig. 8a). However, Tootchi *et al.*¹⁹² classified the wetlands in two parts which are, i) regularly flooded wetlands (RFW) and ii) groundwater-driven wetlands (GDW). Regularly flooded wetlands were produced by taking the combination of three inundation datasets (ESA-CCI, GIEMS-D15, and JRC surface water). However, groundwater-driven wetlands were derived based on Fan *et al.*²²⁵ groundwater simulations, considering only pixels with water table depth less than 20 cm. Finally, Tootchi *et al.*¹⁹² proposed composite wetlands (CW), which are the combinations of RFWs and GDWs (Fig. 8b). However, due to difficulty in down-scaling the flooding to higher resolution in MRB, maps based on Tootchi *et al.*¹⁹² contain higher uncertainty. Therefore, SWAMP data developed by Gumbricht *et al.*¹⁹¹ could be considered better product for wetland identification in MRB.

Similarly, the GIEMS-D15 (Fluet-Chouinard *et al.*¹⁹³) dataset were used to identify the annual inundation areas in MRB (Fig. 8c). We find that, the portions of Mekong Delta and TSL basin and the main streamline of Chi and Mun Rivers, tributaries of the Ngun River, were inundated at least every year (Fig. 8c). However, the entire Mekong Delta, a larger portion of the TSL basin, the basins of the Chi and Mun rivers, and areas

around Vientiane station were all inundated, considering the annual mean of maximum extent of the flood. Moreover, long-term maximum inundation is almost similar to the mean annual maximum inundation in the MRB. Furthermore, the large uncertainties in the flood inundation data in GIEMS-D15 data propagated from downscaling using DEM data, results in overestimation of the inundation area in MRB (Fig. 8c). Therefore, a basin-wide study at a fine temporal and spatial resolution is essential for the management and conservation of biodiversity and other ecosystem services associated with freshwater.

GHG emission data. The EDGAR based GHG emission datasets cover a substantially long period, enabling a detailed analysis of the spatial patterns of GHG emissions in the MRB. Here, we specifically focus on the trends in GHG emissions over time, with the aim of understanding how these emissions have changed (Fig. 9a). Results indicate a rising trend in GHG emissions throughout the basin. We find a high rate of increase in mean annual GHG emissions in Thailand, the western part of Cambodia (around the TSL), and the Mekong Delta as compared to other regions in the MRB (Fig. 9a). These regions with high GHG emissions having high human population are intensive agricultural regions. Moreover, there is an alarming increase in terms of annual mean GHGs from 1970 to 2018 (Fig. 9b) considering the entire MRB.

Socio-economic data. We compare the projected increase in population and GDP under shared socioeconomic pathways (Fig. 10). Population (2010–2100) and GDP (2030–2100) projections are shown in terms of percentage of the base year as 2000 and 2005, respectively. Projections for the regions of Cambodia, China, Laos, Myanmar, Thailand, and Vietnam, which come within the MRB under shared socioeconomic pathways, were calculated for 5 SSP scenarios which are SSP1 (Sustainability), SSP2 (Middle of the road), SSP3 (Regional rivalry), SSP4 (Inequality), and SSP5 (Fossil-fueled development). In almost all countries, all scenarios show a decrease in the population at the end of the century, except the SSP3 scenario, which shows increasing population in most countries within MRB. Also, the SSP4 and SSP5 scenarios show the most decreasing trend among other. Similarly, there is an increase in the GDP for each country under all five SSP scenarios. Where SSP5 scenario which is projected to have one of the lowest populations shows the highest GDP growth in all countries within MRB. Therefore, population, and GDP are inversely projected in each country of MRB (Fig. 10).

Usage Notes

The datasets synthesized in this study could form the basis for a range of hydrological, agricultural, ecological, and socioeconomic studies in the MRB. For example, the meteorological datasets from EM-Earth provide a probabilistic approach to meet the diverse requirements of hydrometeorological and ecological applications. The observed climate and streamflow datasets can be used in hydrological models to constrain the streamflow data at sparse locations in the basin. Further, groundwater datasets we digitized from the published literature partly fill the complete vacuum in groundwater data for the MRB. Such data are crucial for groundwater modeling in the MRB, which is indispensable to better understand the rapidly evolving groundwater dynamics across the basin. Indeed, groundwater in the MRB remains relatively poorly studied and needs increased attention. The nutrient datasets at various locations in the MRB could be used to improve the understanding of the changes in water quality as well as to constrain and validate model simulations on riverine nutrient budgets, another research direction that has received very little attention, owing primarily to critical data gaps. Spatial and temporal changes in land use land cover are directly linked to the changes in hydrological, agricultural, and ecological systems across the basin. Thus, the land use data could be of use for a range of hydrological, agricultural, and ecological studies. Moreover, population projections can be used in determining the exposure and vulnerability to future hazards. The synthesized gridded GDP projections will help in identifying the vulnerability, exposure, and resilience of socioeconomic activities under future climate extremes. In summary, the datasets synthesized here are expected to fill the widely acknowledged and long-debated data gap for the MRB, which has hindered socio-hydrological studies—including modeling and analysis—toward improving the understanding of rapidly emerging hydrological, agricultural, and ecological systems within the basin, and providing improved future projections for transboundary water management and sustainability.

Code availability

The codes used for data processing, analysis, and generating results are available at GitHub repository [www.github.com/yadupokhrel/Tiwari_Mekong_DataSynthesis_SciData](https://github.com/yadupokhrel/Tiwari_Mekong_DataSynthesis_SciData).

Received: 9 February 2023; Accepted: 27 April 2023;

Published online: 15 May 2023

References

1. Pokhrel, Y. *et al.* A review of the integrated effects of changing climate, land use, and dams on Mekong river hydrology. *Water (Basel)* **10**, 266 (2018).
2. Mekong River Commission (MRC). *Overview of the Hydrology of the Mekong Basin*. (2005).
3. Liu, J., Chen, D., Mao, G., Irannezhad, M. & Pokhrel, Y. Past and Future Changes in Climate and Water Resources in the Lancang–Mekong River Basin: Current Understanding and Future Research Directions. *Engineering* **13**, 144–152 (2022).
4. Pokhrel, Y. & Tiwari, A. D. Re-operating dams in the Mekong. *Nat Sustain* <https://doi.org/10.1038/s41893-022-00998-2> (2022).
5. Chen, A. *et al.* Multidecadal variability of the Tonle Sap Lake flood pulse regime. *Hydrol Process* **35** (2021).
6. Uk, S. *et al.* Tonle Sap Lake: Current status and important research directions for environmental management. *Lakes Reserv* **23**, 177–189 (2018).
7. Triet, N. V. K. *et al.* Future projections of flood dynamics in the Vietnamese Mekong Delta. *Science of the Total Environment* **742**, (2020).

8. Ziv, G., Baran, E., Nam, S., Rodríguez-Iturbe, I. & Levin, S. A. Trading-off fish biodiversity, food security, and hydropower in the Mekong River Basin. *Proc Natl Acad Sci USA* **109**, 5609–5614 (2012).
9. Pokhrel, Y., Shin, S., Lin, Z., Yamazaki, D. & Qi, J. Potential Disruption of Flood Dynamics in the Lower Mekong River Basin Due to Upstream Flow Regulation. *Sci Rep* **8**, (2018).
10. Coleman, J. M. & Roberts, H. H. Deltaic coastal wetlands. in *Coastal Lowlands* 1–24 (Springer, Dordrecht, 1989).
11. Kondolf, G. M. *et al.* Save the Mekong Delta from drowning. *Science* (1979) **376**, 583–585 (2022).
12. Mekong River Commission (MRC). *State of the Basin Report 2010*. (2010).
13. Lauri, H. *et al.* Future changes in Mekong River hydrology: Impact of climate change and reservoir operation on discharge. *Hydrol Earth Syst Sci* **16**, 4603–4619 (2012).
14. Johnston, R. & Kummu, M. Water Resource Models in the Mekong Basin: A Review. *Water Resources Management* **26**, 429–455 (2012).
15. Nesbitt, H., Johnston, R. & Soliend, M. Mekong river water: will river flows meet future agriculture needs in Lower Mekong Basin? *Water in agriculture* **116**, (2004).
16. Wang, Y., Zang, S. & Tian, Y. Mapping paddy rice with the random forest algorithm using MODIS and SMAP time series. *Chaos Solitons Fractals* **140**, (2020).
17. Yang, J., Yang, Y. C. E., Chang, J., Zhang, J. & Yao, J. Impact of dam development and climate change on hydroecological conditions and natural hazard risk in the Mekong River Basin. *J Hydrol (Amst)* **579**, (2019).
18. Shrestha, B., Maskey, S., Babel, M. S., van Griensven, A. & Uhlenbrook, S. Sediment related impacts of climate change and reservoir development in the Lower Mekong River Basin: a case study of the Nam Ou Basin, Lao PDR. *Clim Change* **149**, 13–27 (2018).
19. Sridhar, V., Kang, H. & Ali, S. A. Human-induced alterations to land use and climate and their responses for hydrology and water management in the Mekong River Basin. *Water (Switzerland)* **11**, (2019).
20. Dugan, P. J. *et al.* Fish migration, dams, and loss of ecosystem services in the mekong basin. *Ambio* **39**, 344–348 (2010).
21. Intralawan, A., Wood, D., Frankel, R., Costanza, R. & Kubiszewski, I. Tradeoff analysis between electricity generation and ecosystem services in the Lower Mekong Basin. *Ecosyst Serv* **30**, 27–35 (2018).
22. Kondolf, G. M. *et al.* Changing sediment budget of the Mekong: Cumulative threats and management strategies for a large river basin. *Science of the Total Environment* **625**, 114–134 (2018).
23. Xu, B. & He, X. A Physics-Informed Bayesian Storyline Approach to Assess Sediment Transport in the Mekong. *Water Resour Res* **58**, (2022).
24. Kummu, M. & Varis, O. Sediment-related impacts due to upstream reservoir trapping, the Lower Mekong River. *Geomorphology* **85**, 275–293 (2007).
25. Hecht, J. S., Lacombe, G., Arias, M. E., Dang, T. D. & Piman, T. Hydropower dams of the Mekong River basin: A review of their hydrological impacts. *J Hydrol (Amst)* **568**, 285–300 (2019).
26. Yun, X. *et al.* Impacts of climate change and reservoir operation on streamflow and flood characteristics in the Lancang-Mekong River Basin. *J Hydrol (Amst)* **590**, (2020).
27. Maavara, T. *et al.* River dam impacts on biogeochemical cycling. *Nat Rev Earth Environ* **1**, 103–116 (2020).
28. Deemer, B. R. *et al.* Greenhouse gas emissions from reservoir water surfaces: A new global synthesis. *Bioscience* **66**, 949–964 (2016).
29. Wang, F., Maberly, S. C., Wang, B. & Liang, X. Effects of dams on riverine biogeochemical cycling and ecology. *Inland Waters* **8**, 130–140 (2018).
30. Rasanen, T. A., Varis, O., Scherer, L. & Kummu, M. Greenhouse gas emissions of hydropower in the Mekong River Basin. *Environmental Research Letters* **13**, (2018).
31. Soukhaphon, A., Baird, I. G. & Hogan, Z. S. The impacts of hydropower dams in the mekong river basin: A review. *Water (Switzerland)* **13**, (2021).
32. Liu, L. *et al.* Spatial and temporal variability of methane emissions from cascading reservoirs in the Upper Mekong River. *Water Res* **186**, (2020).
33. Gao, J., Zhao, J. & Wang, H. Dam-Impacted Water–Energy–Food Nexus in Lancang-Mekong River Basin. *J Water Resour Plan Manag* **147**, (2021).
34. Binh, D. van, Kantoush, S. & Sumi, T. Changes to long-term discharge and sediment loads in the Vietnamese Mekong Delta caused by upstream dams. *Geomorphology* **353**, (2020).
35. Binh, D. van *et al.* Long-term alterations of flow regimes of the Mekong River and adaptation strategies for the Vietnamese Mekong Delta. *J Hydrol Reg Stud* **32**, (2020).
36. Wang, J., Wu, W., Zhou, X., Huang, Y. & Guo, M. Nitrous oxide (N₂O) emissions from the high dam reservoir in longitudinal range-gorge regions on the Lancang-Mekong River, southwest China. *J Environ Manage* **295**, (2021).
37. Fan, H. & He, D. Temperature and precipitation variability and its effects on streamflow in the upstream regions of the Lancang-Mekong and Nu-Salween Rivers. *J Hydrometeorol* **16**, 2248–2263 (2015).
38. Li, D., Long, D., Zhao, J., Lu, H. & Hong, Y. Observed changes in flow regimes in the Mekong River basin. *J Hydrol (Amst)* **551**, 217–232 (2017).
39. Lu, X. X. & Chua, S. D. X. River Discharge and Water Level Changes in the Mekong River: Droughts in an Era of Mega-Dams. *Hydrol Process* **35**, (2021).
40. Irannezhad, M. & Liu, J. Evaluation of six gauge-based gridded climate products for analyzing long-term historical precipitation patterns across the Lancang-Mekong River Basin. *Geography and Sustainability* **3**, 85–103 (2022).
41. Mohammed, I. N., Bolten, J. D., Srinivasan, R. & Lakshmi, V. Satellite observations and modeling to understand the Lower Mekong River Basin streamflow variability. *J Hydrol (Amst)* **564**, 559–573 (2018).
42. Mohammed, I. N., Bolten, J. D., Srinivasan, R. & Lakshmi, V. Improved hydrological decision support system for the Lower Mekong River Basin using satellite-based earth observations. *Remote Sens (Basel)* **10**, (2018).
43. Vu, H. T. D. *et al.* Land use change in the Vietnamese Mekong Delta: New evidence from remote sensing. *Science of the Total Environment* **813**, (2022).
44. Kabir, T., Pokhrel, Y. & Felfelani, F. On the Precipitation-Induced Uncertainties in Process-Based Hydrological Modeling in the Mekong River Basin. *Water Resour Res* **58**, (2022).
45. Shin, S. *et al.* High Resolution Modeling of River-Floodplain-Reservoir Inundation Dynamics in the Mekong River Basin. *Water Resour Res* **56**, (2020).
46. Minderhoud, P. S. J. *et al.* Impacts of 25 years of groundwater extraction on subsidence in the Mekong delta, Vietnam. *Environmental Research Letters* **12**, (2017).
47. Dang, H. *et al.* Hydrologic balance and inundation dynamics of Southeast Asia's largest inland lake altered by hydropower dams in the Mekong River basin. *Science of the Total Environment* **831**, (2022).
48. Lu, Y. *et al.* Socio-hydrologic modeling of the dynamics of cooperation in the transboundary Lancang-Mekong River. *Hydrol Earth Syst Sci* **25**, 1883–1903 (2021).
49. Try, S., Lee, G., Yu, W., Oeurng, C. & Jang, C. Large-Scale Flood-Inundation Modeling in the Mekong River Basin. *J Hydrol Eng* **23**, (2018).
50. Wang, J. *et al.* Modeling Daily Floods in the Lancang-Mekong River Basin Using an Improved Hydrological-Hydrodynamic Model. *Water Resour Res* **57**, (2021).
51. Mondal, A., Le, M. H. & Lakshmi, V. Land use, climate, and water change in the Vietnamese Mekong Delta (VMD) using earth observation and hydrological modeling. *J Hydrol Reg Stud* **42**, (2022).

52. Li, S. & Bush, R. T. Rising flux of nutrients (C, N, P and Si) in the lower Mekong River. *J Hydrol (Amst)* **530**, 447–461 (2015).
53. Tran, T. N. D. *et al.* Quantification of global Digital Elevation Model (DEM) – A case study of the newly released NASADEM for a river basin in Central Vietnam. *J Hydrol Reg Stud* **45**, (2023).
54. Kabeya, N. *et al.* Long-term hydrological observations in a lowland dry evergreen forest catchment area of the lower Mekong River, Cambodia. *Japan Agricultural Research Quarterly: JARQ* **55**, 177–190 (2021).
55. Muenratch, P., Nguyen, T. P. L., Shrestha, S., Chatterjee, J. S. & Virdis, S. G. P. Governance and policy responses to anthropogenic and climate pressures on groundwater resources in the Greater Mekong Subregion urbanizing cities. *Groundw Sustain Dev* **18**, (2022).
56. Ribolzi, O. *et al.* Interacting land use and soil surface dynamics control groundwater outflow in a montane catchment of the lower Mekong basin. *Agric Ecosyst Environ* **268**, 90–102 (2018).
57. Arias, M. E., Piman, T., Lauri, H., Cochrane, T. A. & Kumm, M. Dams on Mekong tributaries as significant contributors of hydrological alterations to the Tonle Sap Floodplain in Cambodia. *Hydrol Earth Syst Sci* **18**, 5303–5315 (2014).
58. Bonnema, M. & Hossain, F. Assessing the Potential of the Surface Water and Ocean Topography Mission for Reservoir Monitoring in the Mekong River Basin. *Water Resour Res* **55**, 444–461 (2019).
59. Whitehead, P. G. *et al.* Water quality modelling of the Mekong River basin: Climate change and socioeconomics drive flow and nutrient flux changes to the Mekong Delta. *Science of the Total Environment* **673**, 218–229 (2019).
60. Vu, Q. M., Lakshmi, V. & Bolten, J. Assessment of the biomass productivity decline in the lower Mekong Basin. *Remote Sens (Basel)* **11**, (2019).
61. Spruce, J., Bolten, J., Srinivasan, R. & Lakshmi, V. Developing land use land cover maps for the lower mekong basin to aid hydrologic modeling and basin planning. *Remote Sens (Basel)* **10**, (2018).
62. Spruce, J., Bolten, J., Mohammed, I. N., Srinivasan, R. & Lakshmi, V. Mapping Land Use Land Cover Change in the Lower Mekong Basin From 1997 to 2010. *Front Environ Sci* **8**, (2020).
63. Poortinga, A. *et al.* Predictive analytics for identifying land cover change hotspots in the mekong region. *Remote Sens (Basel)* **12**, (2020).
64. Hui, T. R., Park, E., Loc, H. H. & Tien, P. D. Long-term hydrological alterations and the agricultural landscapes in the Mekong Delta: Insights from remote sensing and national statistics. *Environmental Challenges* **7**, (2022).
65. Son, N. T., Chen, C. F., Chen, C. R., Duc, H. N. & Chang, L. Y. A phenology-based classification of time-series MODIS data for rice crop monitoring in Mekong Delta, Vietnam. *Remote Sens (Basel)* **6**, 135–156 (2013).
66. Nguyen, T. T. H., de Bie, C. A. J. M., Ali, A., Smaling, E. M. A. & Chu, T. H. Mapping the irrigated rice cropping patterns of the Mekong delta, Vietnam, through hyper-temporal spot NDVI image analysis. *Int J Remote Sens* **33**, 415–434 (2012).
67. Lin, Z. & Qi, J. A new remote sensing approach to enrich hydropower dams' information and assess their impact distances: A case study in the Mekong River Basin. *Remote Sens (Basel)* **11**, (2019).
68. Aires, F. *et al.* Surface water evolution (2001–2017) at the Cambodia/Vietnam border in the upper mekong delta using satellite MODIS observations. *Remote Sens (Basel)* **12**, (2020).
69. Prigent, C., Matthews, E., Aires, F. & Rossow, W. B. Remote sensing of global wetland dynamics with multiple satellite data sets. *Geophys Res Lett* **28**, 4631–4634 (2001).
70. Cho, M. S. & Qi, J. Characterization of the impacts of hydro-dams on wetland inundations in Southeast Asia. *Science of the Total Environment* **864**, (2023).
71. Dandridge, C., Stanley, T., Amatya, P. & Lakshmi, V. The Influence of Land Use and Land Cover Change on Landslide Susceptibility in the Lower Mekong River Basin, <https://doi.org/10.21203/rs.3.rs-1595283/v1> (2022).
72. Tian, W. *et al.* Evaluation of six precipitation products in the Mekong River Basin. *Atmos Res* **255**, (2021).
73. Jing, W. *et al.* Variations in terrestrial water storage in the Lancang-Mekong river basin from GRACE solutions and land surface model. *J Hydrol (Amst)* **580**, (2020).
74. Li, Z. *et al.* Uncertainties in satellite remote sensing of aerosols and impact on monitoring its long-term trend: a review and perspective. *Ann Geophys* **27**, 2755–2770 (2009).
75. Galelli, S., Dang, T. D., Ng, J. Y., Chowdhury, A. F. M. K. & Arias, M. E. Opportunities to curb hydrological alterations via dam re-operation in the Mekong. *Nat Sustain* <https://doi.org/10.1038/s41893-022-00971-z> (2022).
76. Burbano, M., Shin, S., Nguyen, K. & Pokhrel, Y. Hydrologic changes, dam construction, and the shift in dietary protein in the Lower Mekong River Basin. *J Hydrol (Amst)* **581**, (2020).
77. Sabo, J. L. *et al.* Designing river flows to improve food security futures in the Lower Mekong Basin. *Science (1979)* **358**, (2017).
78. Lu, X. X., Li, S., Kumm, M., Padawangi, R. & Wang, J. J. Observed changes in the water flow at Chiang Saen in the lower Mekong: Impacts of Chinese dams? *Quaternary International* **336**, 145–157 (2014).
79. Kazama, S., Hagiwara, T., Ranjan, P. & Sawamoto, M. Evaluation of groundwater resources in wide inundation areas of the Mekong River basin. *J Hydrol (Amst)* **340**, 233–243 (2007).
80. Balica, S., Dinh, Q., Popescu, I., Vo, T. Q. & Pham, D. Q. Flood impact in the Mekong Delta. *Vietnam. J Maps* **10**, 257–268 (2014).
81. Endo, A. *et al.* Dynamics of water–energy–food nexus methodology, methods and tools. *Curr Opin Environ Sci Health* **13**, 46–60 (2020).
82. Gallagher, L. *et al.* Supporting stakeholders to anticipate and respond to risks in a Mekong river water-energy-food nexus. *Ecology and Society* **25**, 1–16 (2020).
83. Krzysztofowicz, R. Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Water Resour Res* **35**, 2739–2750 (1999).
84. Oki, T. & Kanae, S. Global hydrological cycles and world water resources. *Science (1979)* **313**, 1068–1072 (2006).
85. Wang, W. *et al.* Modelling hydrologic processes in the Mekong River basin using a distributed model driven by satellite precipitation and rain gauge observations. *PLoS One* **11**, (2016).
86. Le, H. M. *et al.* J. R. P. D. J. D. V. Comparison and bias correction of TMPA precipitation products over the lower part of Red-Thai Binh River Basin of Vietnam. *Remote Sens (Basel)* **10**, (2018).
87. Dandridge, C., Lakshmi, V., Bolten, J. & Srinivasan, R. Evaluation of satellite-based rainfall estimates in the Lower Mekong River Basin (Southeast Asia). *Remote Sens (Basel)* **11**, (2019).
88. Yatagai, A. *et al.* A 44-Year Daily Gridded Precipitation Dataset for Asia Based on a Dense Network of Rain Gauges. *Sola* **5**, 137–140 (2009).
89. Yatagai, A. *et al.* Aphrodite constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bull Am Meteorol Soc* **93**, 1401–1415 (2012).
90. Tanarhte, M., Hadjinicolaou, P. & Lelieveld, J. Intercomparison of temperature and precipitation data sets based on observations in the Mediterranean and the Middle East. *Journal of Geophysical Research Atmospheres* **117**, (2012).
91. Sidike, A., Chen, X., Liu, T., Durdiev, K. & Huang, Y. Investigating alternative climate data sources for hydrological simulations in the upstream of the Amu Darya river. *Water (Switzerland)* **8**, (2016).
92. Tan, M. L., Gassman, P. W. & Cracknell, A. P. Assessment of three long-term gridded climate products for hydro-climatic simulations in tropical river basins. *Water (Switzerland)* **9**, (2017).
93. Funk, C. *et al.* The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Sci Data* **2**, 150066 (2015).
94. Luo, X., Wu, W., He, D., Li, Y. & Ji, X. Hydrological Simulation Using TRMM and CHIRPS Precipitation Estimates in the Lower Lancang-Mekong River Basin. *Chin Geogr Sci* **29**, 13–25 (2019).

95. Guo, H. *et al.* Meteorological drought analysis in the Lower Mekong Basin using satellite-based long-term CHIRPS product. *Sustainability (Switzerland)* **9**, (2017).
96. Try, S. *et al.* Comparison of gridded precipitation datasets for rainfall-runoff and inundation modeling in the Mekong River Basin. *PLoS One* **15**, (2020).
97. Lauri, H., Räsänen, T. A. & Kumm, M. Using Reanalysis and Remotely Sensed Temperature and Precipitation Data for Hydrological Modeling in Monsoon Climate: Mekong River Case Study. *J Hydrometeorol* **15**, 1532–1545 (2014).
98. Tang, X., Zhang, J., Gao, C., Ruben, G. B. & Wang, G. Assessing the uncertainties of four precipitation products for SWAT modeling in Mekong River Basin. *Remote Sens (Basel)* **11**, (2019).
99. Hersbach, H. *et al.* The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* **146**, 1999–2049 (2020).
100. He, X., Pan, M., Wei, Z., Wood, E. F. & Sheffield, J. A global drought and flood catalogue from 1950 to 2016. *Bull Am Meteorol Soc* **101**, E508–E535 (2020).
101. Sheffield, J., Goteti, G. & Wood, E. F. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *J Clim* **19**, 3088–3111 (2006).
102. Weedon, G. P. *et al.* Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. *J Hydrometeorol* **12**, 823–848 (2011).
103. Weedon, G. P. *et al.* The WFDEI meteorological forcing data set: WATCH Forcing data methodology applied to ERA-Interim reanalysis data. *Water Resour Res* **50**, 7505–7514 (2014).
104. Kim, H. Global Soil Wetness Project Phase 3 Atmospheric Boundary Conditions (Experiment 1) [Data set]. *Data Integration and Analysis System (DIAS)* <https://doi.org/10.20783/DIAS.501> (2017).
105. Lange, S. WFDE5 over land merged with ERA5 over the ocean (W5E5) V. 1.0. *GFZ Data Service* (2019).
106. Brocca, L. *et al.* Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data. *J Geophys Res* **119**, 5128–5141 (2014).
107. Ciabatta, L. *et al.* SM2RAIN-CCI: A new global long-term rainfall data set derived from ESA CCI soil moisture. *Earth Syst Sci Data* **10**, 267–280 (2018).
108. Brocca, L. *et al.* SM2RAIN-ASCAT (2007–2018): Global daily satellite rainfall data from ASCAT soil moisture observations. *Earth Syst Sci Data* **11**, 1583–1601 (2019).
109. Massari, C. *et al.* A daily 25 km short-latency rainfall product for data-scarce regions based on the integration of the Global Precipitation Measurement mission rainfall and multiple-satellite soil moisture products. *Hydrol Earth Syst Sci* **24**, 2687–2710 (2020).
110. Tang, G., Clark, M. P. & Papalexiou, S. M. EM-Earth: The Ensemble Meteorological Dataset for Planet Earth. *Bull Am Meteorol Soc* **103**, E996–E1018 (2022).
111. Karger, D. N., Wilson, A. M., Mahony, C., Zimmermann, N. E. & Jetz, W. Global daily 1 km land surface precipitation based on cloud cover-informed downscaling. *Sci Data* **8**, (2021).
112. Pokhrel, Y. N., Hanasaki, N., Wada, Y. & Kim, H. Recent progresses in incorporating human land–water management into global land surface models toward their integration into Earth system models. *Wiley Interdisciplinary Reviews: Water* **3**, 548–574 (2016).
113. Arias, M. E. *et al.* Quantifying changes in flooding and habitats in the Tonle Sap Lake (Cambodia) caused by water infrastructure development and climate change in the Mekong Basin. *J Environ Manage* **112**, 53–66 (2012).
114. Mohammed, I. N. *et al.* Ground and satellite based observation datasets for the Lower Mekong River Basin. *Data Brief* **21**, 2020–2027 (2018).
115. Tang, J., Yin, X. A., Yang, P. & Yang, Z. F. Assessment of Contributions of Climatic Variation and Human Activities to Streamflow Changes in the Lancang River, China. *Water Resources Management* **28**, 2953–2966 (2014).
116. Tang, J., Yin, X. A., Yang, P. & Yang, Z. F. Climate-Induced Flow Regime Alterations and their Implications for the Lancang River, China. *River Res Appl* **31**, 422–432 (2015).
117. Zhang, X. *et al.* Satellite-based operational real-time drought monitoring in the transboundary lancang-mekong river basin. *Remote Sens (Basel)* **12**, (2020).
118. Zeng, Z. *et al.* Global evapotranspiration over the past three decades: Estimation based on the water balance equation combined with empirical models. *Environmental Research Letters* **7**, (2012).
119. Martens, B. *et al.* GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geosci Model Dev* **10**, 1903–1925 (2017).
120. Zhang, Y. *et al.* Coupled estimation of 500 m and 8-day resolution global evapotranspiration and gross primary production in 2002–2017. *Remote Sens Environ* **222**, 165–182 (2019).
121. Mu, Q., Zhao, M. & Running, S. W. Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sens Environ* **115**, 1781–1800 (2011).
122. Xie, Z. *et al.* The Global Land Surface Satellite (GLASS) evapotranspiration product Version 5.0: Algorithm development and preliminary validation. *J Hydrol (Amst)* **610**, (2022).
123. Hu, S. & Mo, X. Diversified evapotranspiration responses to climatic change and vegetation greening in eight global great river basins. *J Hydrol (Amst)* **613**, (2022).
124. Chen, H. *et al.* Assessment and Inter-Comparison of Multi-Source High Spatial Resolution Evapotranspiration Products over Lancang–Mekong River Basin, Southeast Asia. *Remote Sens (Basel)* **14**, 479 (2022).
125. Gao, H., Birkett, C. & Lettenmaier, D. P. Global monitoring of large reservoir storage from satellite remote sensing. *Water Resour Res* **48**, 1–12 (2012).
126. Zhang, S., Gao, H. & Naz, B. S. Monitoring reservoir storage in South Asia from multisatellite remote sensing. *Water Resource Research* **50**, 8927–8943 (2014).
127. Pekel, J. F., Cottam, A., Gorelick, N. & Belward, A. S. High-resolution mapping of global surface water and its long-term changes. *Nature* **540**, 418–422 (2016).
128. Ji, L., Gong, P., Wang, J., Shi, J. & Zhu, Z. Construction of the 500-m Resolution Daily Global Surface Water Change Database (2001–2016). *Water Resour Res* **54**, 10,270–10,292 (2018).
129. Entekhabi, D. *et al.* The soil moisture active passive (SMAP) mission. *Proceedings of the IEEE* **98**, 704–716 (2010).
130. Jacqueline, E. *et al.* SMOS CATDS level 3 global products over land. in *Remote Sensing for Agriculture, Ecosystems, and Hydrology XII* vol. **7824** 78240K (SPIE, 2010).
131. Liu, Y. Y. *et al.* Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sens Environ* **123**, 280–297 (2012).
132. Wagner, W. *et al.* Fusion of active and passive microwave observations to create an essential climate variable data record on soil moisture. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS Annals)* **7**, 315–321 (2012).
133. Fang, B. *et al.* A global 1-km downscaled SMAP soil moisture product based on thermal inertia theory. *Vadose Zone Journal* **21**, (2022).
134. Das, N. N. *et al.* The SMAP and Copernicus Sentinel 1 A/B microwave active-passive high resolution surface soil moisture product. *Remote Sens Environ* **233**, (2019).
135. Dandridge, C., Fang, B. & Lakshmi, V. Downscaling of SMAP Soil Moisture in the Lower Mekong River Basin. *Water (Switzerland)* **12**, (2020).
136. Duy, N. le *et al.* Groundwater dynamics in the Vietnamese Mekong Delta: Trends, memory effects, and response times. *J Hydrol Reg Stud* **33**, (2021).

137. Erban, L. E. & Gorelick, S. M. Closing the irrigation deficit in Cambodia: Implications for transboundary impacts on groundwater and Mekong River flow. *J Hydrol (Amst)* **535**, 85–92 (2016).
138. Fryar, A. E., Schreiber, M. E., Pholkern, K., Srisuk, K. & Ziegler, B. A. Variability in groundwater flow and chemistry in the Mekong River alluvial aquifer (Thailand): implications for arsenic and manganese occurrence. *Environ Earth Sci* **80**, (2021).
139. Xiao, H. *et al.* Saltwater intrusion into groundwater systems in the Mekong Delta and links to global change. *Advances in Climate Change Research* **12**, 342–352 (2021).
140. Hoan, T. V. *et al.* An Improved Groundwater Model Framework for Aquifer Structures of the Quaternary-Formed Sediment Body in the Southernmost Parts of the Mekong Delta, Vietnam. *Hydrology* **9**, (2022).
141. IUCN. *Mekong Water Dialogues: groundwater in the Mekong Delta*. International Union for Conservation of Nature http://scholar.google.com/scholar_lookup?&title=Mekong%20Water%20Dialogues%3A%20groundwater%20in%20the%20Mekong%20Delta&publication_year=2011 (2011).
142. Johnston, R., Roberts, M., Try, T. & de Silva, S. *Groundwater for irrigation in Cambodia*. (2013).
143. MRD, J. *The study on groundwater development in southern Cambodia: final report*. (2002).
144. Thu, N. T. Groundwater and Surface Water Cycle System in Mekong Delta, Vietnam. *Life and Environmental Sciences* **171** (2017).
145. Nobuhiro, T. *et al.* Evapotranspiration during the late rainy season and middle of the dry season in the watershed of an evergreen forest area, central Cambodia. *Hydrol Process* **22**, 1281–1289 (2008).
146. Park, E. *et al.* Impacts of agricultural expansion on floodplain water and sediment budgets in the Mekong River. *J Hydrol (Amst)* **605**, (2022).
147. Petpongpan, C., Ekkawatpanit, C. & Kositgittiwong, D. Climate change impact on surface water and groundwater recharge in northern Thailand. *Water (Switzerland)* **12**, (2020).
148. Le, P. V. V. *et al.* Responses of groundwater to precipitation variability and ENSO in the Vietnamese Mekong Delta. *Hydrology Research* **52**, 1280–1293 (2021).
149. Seeboonruang, U. An application of time-lag regression technique for assessment of groundwater fluctuations in a regulated river basin: a case study in Northeastern Thailand. *Environ Earth Sci* **73**, 6511–6523 (2015).
150. Shrestha, S., Bach, T. V. & Pandey, V. P. Climate change impacts on groundwater resources in Mekong Delta under representative concentration pathways (RCPs) scenarios. *Environ Sci Policy* **61**, 1–13 (2016).
151. Tu, T. A. *et al.* Localized recharge processes in the NE Mekong Delta and implications for groundwater quality. *Science of the Total Environment* **845**, (2022).
152. Tsubo, M. *et al.* Effects of soil clay content on water balance and productivity in rainfed lowland rice ecosystem in Northeast Thailand. *Plant Prod Sci* **10**, 232–241 (2007).
153. van Ty, T. *et al.* Spatiotemporal variations in groundwater levels and the impact on land subsidence in CanTho, Vietnam. *Groundw Sustain Dev* **15**, (2021).
154. Van, T. P. & Koontanakulvong, S. Estimation of groundwater use pattern and distribution in the coastal mekong delta, Vietnam via socio-economical survey and groundwater modelling. *Engineering Journal* **23**, 487–499 (2019).
155. Vote, C. *et al.* The use of groundwater as an alternative water source for agricultural production in southern Lao PDR and the implications for policymakers. in *A policy dialogue on rice futures: rice-based farming systems research in the Mekong region* 103–115 (2014).
156. Wagner, F., Tran, V. B. & Renaud, F. G. Groundwater Resources in the Mekong Delta: Availability, Utilization and Risks. in *The Mekong Delta System* 201–220 (Springer, 2012).
157. Lehner, B. *et al.* High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. *Frontiers in Ecology and the Environment* **9**, 494–502 (2011).
158. Pokhrel, Y. *et al.* Global terrestrial water storage and drought severity under climate change. *Nat Clim Chang* **11**, 226–233 (2021).
159. Wada, Y. & Bierkens, M. F. P. Sustainability of global water use: Past reconstruction and future projections. *Environmental Research Letters* **9**, (2014).
160. Hanasaki, N., Yoshikawa, S., Pokhrel, Y. & Kanae, S. A global hydrological simulation to specify the sources of water used by humans. *Hydrol Earth Syst Sci* **22**, 789–817 (2018).
161. Mulligan, M., van Soesbergen, A. & Sáenz, L. GOODD, a global dataset of more than 38,000 georeferenced dams. *Sci Data* **7**, (2020).
162. Wang, J. *et al.* GeoDAR: georeferenced global dams and reservoirs dataset for bridging attributes and geolocations. *Earth Syst Sci Data* **14**, 1869–1899 (2022).
163. Zhang, A. T. & Gu, V. X. Global Dam Tracker: A database of more than 35,000 dams with location, catchment, and attribute information. *Sci Data* **10**, (2023).
164. Yizgaw, W. *et al.* A New Global Storage-Area-Depth Data Set for Modeling Reservoirs in Land Surface and Earth System Models. *Water Resour Res* **54**, 10,372–10,386 (2018).
165. Yun, X., Tang, Q., Sun, S. & Wang, J. Reducing Climate Change Induced Flood at the Cost of Hydropower in the Lancang-Mekong River Basin. *Geophys Res Lett* **48**, (2021).
166. Schmitt, R. J. P., Bizzi, S., Castelletti, A., Opperman, J. J. & Kondolf, G. M. Planning dam portfolios for low sediment trapping shows limits for sustainable hydropower in the Mekong. *Sci Adv* **5**, eaaw2175 (2019).
167. Räsänen, T. A. *et al.* Observed river discharge changes due to hydropower operations in the Upper Mekong Basin. *J Hydrol (Amst)* **545**, 28–41 (2017).
168. Chaudhari, S. & Pokhrel, Y. Alteration of River Flow and Flood Dynamics by Existing and Planned Hydropower Dams in the Amazon River Basin. *Water Resour Res* **58**, (2022).
169. Cho, M. S. & Qi, J. Quantifying spatiotemporal impacts of hydro-dams on land use/land cover changes in the Lower Mekong River Basin. *Applied Geography* **136**, (2021).
170. Zhao, Q., Liu, S. & Dong, S. Effect of dam construction on spatial-temporal change of land use: A case study of Manwan, Lancang River, Yunnan, China. in *Procedia Environmental Sciences* **2**, 852–858 (2010).
171. Bartholomé, E. & Belward, A. S. GLC2000: A new approach to global land cover mapping from earth observation data. *Int J Remote Sens* **26**, 1959–1977 (2005).
172. Balzter, H., Cole, B., Thiel, C. & Schmillius, C. Mapping CORINE land cover from Sentinel-1A SAR and SRTM digital elevation model data using random forests. *Remote Sens (Basel)* **7**, 14876–14898 (2015).
173. Hu, S. & Mo, X. Attribution of long-term evapotranspiration trends in the mekong river basin with a remote sensing-based process model. *Remote Sens (Basel)* **13**, 1–18 (2021).
174. Do, P. *et al.* Exploring synergies in the water-food-energy nexus by using an integrated hydro-economic optimization model for the Lancang-Mekong River basin. *Science of the Total Environment* **728**, (2020).
175. Abhishek, A. *et al.* Evaluating the impacts of drought on rice productivity over Cambodia in the Lower Mekong Basin. *J Hydrol (Amst)* **599**, (2021).
176. Whitcraft, A. K., Becker-Reshef, I. & Justice, C. O. A framework for defining spatially explicit earth observation requirements for a global agricultural monitoring initiative (GEOGLAM). *Remote Sens (Basel)* **7**, 1461–1481 (2015).
177. Jägermeyr, J. *et al.* Climate impacts on global agriculture emerge earlier in new generation of climate and crop models. *Nat Food* **2**, 873–885 (2021).
178. Siebert, S. *et al.* Groundwater use for irrigation - A global inventory. *Hydrol Earth Syst Sci* **14**, 1863–1880 (2010).
179. FAO. *The state of world fisheries and aquaculture*. vol. **2014** (2014).

180. Veldkamp, T. I. E. *et al.* Water scarcity hotspots travel downstream due to human interventions in the 20th and 21st century. *Nat Commun* **8**, (2017).
181. Orr, S., Pittock, J., Chapagain, A. & Dumaresq, D. Dams on the Mekong River: Lost fish protein and the implications for land and water resources. *Global Environmental Change* **22**, 925–932 (2012).
182. Siebert, S., Henrich, V., Frenken, K. & Burke, J. Update of the digital global map of irrigation areas to version 5. in *Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany and Food and Agriculture Organization of the United Nations*, <https://doi.org/10.13140/2.1.2660.6728> (2013).
183. Zabel, F., Putzenlechner, B. & Mauser, W. Global agricultural land resources - A high resolution suitability evaluation and its perspectives until 2100 under climate change conditions. *PLoS One* **9**, (2014).
184. Salmon, J. M., Friedl, M. A., Frohling, S., Wisser, D. & Douglas, E. M. Global rain-fed, irrigated, and paddy croplands: A new high resolution map derived from remote sensing, crop inventories and climate data. *International Journal of Applied Earth Observation and Geoinformation* **38**, 321–334 (2015).
185. Meier, J., Zabel, F. & Mauser, W. A global approach to estimate irrigated areas - A comparison between different data and statistics. *Hydro Earth Syst Sci* **22**, 1119–1133 (2018).
186. Koehnken, L. *Discharge sediment monitoring project (DSMP) 2009–2013 summary and analysis of results (final report) MRC/GIZ*. (2014).
187. Beveridge, C., Hossain, F. & Bonnema, M. Estimating Impacts of Dam Development and Landscape Changes on Suspended Sediment Concentrations in the Mekong River Basin's 3 S Tributaries. *J Hydrol Eng* **25**, (2020).
188. Wackerman, C., Hayden, A. & Jonik, J. Deriving spatial and temporal context for point measurements of suspended-sediment concentration using remote-sensing imagery in the Mekong Delta. *Cont Shelf Res* **147**, 231–245 (2017).
189. Huu Nguyen, H., Dargusch, P., Moss, P. & Tran, D. B. A review of the drivers of 200 years of wetland degradation in the Mekong Delta of Vietnam. *Reg Environ Change* **16**, 2303–2315 (2016).
190. Dang, A. T. N., Kumar, L., Reid, M. & Nguyen, H. Remote sensing approach for monitoring coastal wetland in the mekong delta, vietnam: Change trends and their driving forces. *Remote Sens (Basel)* **13**, (2021).
191. Gumbrecht, T. *et al.* An expert system model for mapping tropical wetlands and peatlands reveals South America as the largest contributor. *Glob Chang Biol* **23**, 3581–3599 (2017).
192. Tootchi, A., Jost, A. & Ducharne, A. Multi-source global wetland maps combining surface water imagery and groundwater constraints. *Earth Syst Sci Data* **11**, 189–220 (2019).
193. Fluet-Chouinard, E., Lehner, B., Rebelo, L. M., Papa, F. & Hamilton, S. K. Development of a global inundation map at high spatial resolution from topographic downscaling of coarse-scale remote sensing data. *Remote Sens Environ* **158**, 348–361 (2015).
194. Prigent, C., Papa, F., Aires, F., Rossow, W. B. & Matthews, E. Global inundation dynamics inferred from multiple satellite observations, 1993–2000. *Journal of Geophysical Research Atmospheres* **112**, (2007).
195. Papa, F. *et al.* Interannual variability of surface water extent at the global scale, 1993–2004. *Journal of Geophysical Research Atmospheres* **115**, (2010).
196. Payne, J. V. *et al.* Flood mapping in the lower Mekong River Basin using daily MODIS observations. *Int J Remote Sens* **38**, 1737–1757 (2017).
197. Arai, H. *et al.* Greenhouse gas emissions from rice straw burning and straw-mushroom cultivation in a triple rice cropping system in the Mekong Delta. *Soil Sci Plant Nutr* **61**, 719–735 (2015).
198. Nguyen, D. B., Gruber, A. & Wagner, W. Mapping rice extent and cropping scheme in the Mekong Delta using Sentinel-1A data. *Remote Sensing Letters* **7**, 1209–1218 (2016).
199. Arai, H. Increased rice yield and reduced greenhouse gas emissions through alternate wetting and drying in a triple-cropped rice field in the Mekong Delta. *Science of the Total Environment* **842**, (2022).
200. Luu, Q., le, Nguyen, N. H., Halog, A. & Bui, H. van. GHG emission reduction in energy sector and its abatement cost: Case study of five provinces in Mekong delta region, Vietnam. *Int J Green Energy* **15**, 715–723 (2018).
201. Khai, N. H., Tinh, T. K., Tin, H. Q. & Sanh, N. V. Reducing Greenhouse Gas Emissions in Rice Grown in the Mekong Delta of Vietnam. *Environment Pollution and Climate Change* **02**, 158 (2018).
202. Shi, W. *et al.* Nitrous oxide emissions from cascade hydropower reservoirs in the upper Mekong River. *Water Res* **173**, (2020).
203. Crippa, M. *et al.* Fossil CO₂ and GHG emissions of all world countries. *Publication Office of the European Union: Luxemburg* <https://doi.org/10.2760/687800> (2019).
204. Ritchie, H., Roser, M. & Rosado, P. CO₂ and Greenhouse Gas Emissions. *Our World in Data* (2020).
205. Kim Oanh, N. T., Permadi, D. A., Dong, N. P. & Nguyet, D. A. Emission of Toxic Air Pollutants and Greenhouse Gases from Crop Residue Open Burning in Southeast Asia. in *Land-atmospheric research applications in South and Southeast Asia* 47–66, https://doi.org/10.1007/978-3-319-67474-2_3 (Springer, 2018).
206. Dasgupta, S., Lall, S. V & Wheeler, D. Tracking Methane Emissions by Satellite: A New World Bank Database and Case Study for Irrigated Rice Production. *Policy Research Working Paper;10224* (2022).
207. Wang, T. & Sun, F. Global gridded GDP data set consistent with the shared socioeconomic pathways. *Sci Data* **9**, (2022).
208. Doxsey-Whitfield, E. *et al.* Taking Advantage of the Improved Availability of Census Data: A First Look at the Gridded Population of the World, Version 4. *Papers in Applied Geography* **1**, 226–234 (2015).
209. Döll, P. Vulnerability to the impact of climate change on renewable groundwater resources: A global-scale assessment. *Environmental Research Letters* **4**, (2009).
210. Guha-Sapir, D., Rodriguez-Llanes, J. M. & Jakubicka, T. Using disaster footprints, population databases and GIS to overcome persistent problems for human impact assessment in flood events. *Natural Hazards* **58**, 845–852 (2011).
211. Caminade, C. *et al.* Impact of climate change on global malaria distribution. *Proc Natl Acad Sci USA* **111**, 3286–3291 (2014).
212. Jones, B. & O'Neill, B. C. Spatially explicit global population scenarios consistent with the Shared Socioeconomic Pathways. *Environmental Research Letters* **11**, (2016).
213. Tiwari, A. D. *et al.* A synthesis of hydroclimatic, ecological, and socioeconomic data for transdisciplinary research in the Mekong. *Zenodo* <https://doi.org/10.5281/zenodo.7803254> (2023).
214. Condon, L. E., Atchley, A. L. & Maxwell, R. M. Evapotranspiration depletes groundwater under warming over the contiguous United States. *Nat Commun* **11**, (2020).
215. Meixner, T. *et al.* Implications of projected climate change for groundwater recharge in the western United States. *J Hydrol (Amst)* **534**, 124–138 (2016).
216. Asoka, A., Gleeson, T., Wada, Y. & Mishra, V. Relative contribution of monsoon precipitation and pumping to changes in groundwater storage in India. *Nat Geosci* **10**, 109–117 (2017).
217. Bhanja, S. N., Mukherjee, A., Saha, D., Velicogna, I. & Famiglietti, J. S. Validation of GRACE based groundwater storage anomaly using *in-situ* groundwater level measurements in India. *J Hydrol (Amst)* **543**, 729–738 (2016).
218. Erban, L. E., Gorelick, S. M. & Zebker, H. A. Groundwater extraction, land subsidence, and sea-level rise in the Mekong Delta, Vietnam. *Environmental Research Letters* **9**, (2014).

219. Fan, H., He, D. & Wang, H. Environmental consequences of damming the mainstream Lancang-Mekong river: A review. *Earth Sci Rev* **146**, 77–91 (2015).
220. Stone, R. Mayhem on the Mekong. *Science* (1979) **333**, 814–818 (2011).
221. Trisurat, Y., Aekakkararungroj, A. & Ma, H. ok & Johnston, J. M. Basin-wide impacts of climate change on ecosystem services in the Lower Mekong Basin. *Ecol Res* **33**, 73–86 (2018).
222. Chen, Q. *et al.* Hydropower reservoirs on the upper Mekong River modify nutrient bioavailability downstream. *Natl Sci Rev* **7**, 1449–1457 (2020).
223. Wang, Y., Feng, L., Liu, J., Hou, X. & Chen, D. Changes of inundation area and water turbidity of Tonle Sap Lake: Responses to climate changes or upstream dam construction? *Environmental Research Letters* **15**, (2020).
224. Schmitt, R. J. P., Rubin, Z. & Kondolf, G. M. Losing ground - scenarios of land loss as consequence of shifting sediment budgets in the Mekong Delta. *Geomorphology* **294**, 58–69 (2017).
225. Fan, Y., Li, H. & Miguez-Macho, G. Global patterns of groundwater table depth. *Science* (1979) **339**, 940–943 (2013).
226. Simpson, J., Kummerow, C., Tao, W.-K. & Adler, R. F. On the tropical rainfall measuring mission (TRMM). *Meteorology and Atmospheric physics* **60**, 19–36 (1996).
227. Hou, A. Y., Skofronick-Jackson, G., Kummerow, C. D. & Shepherd, J. M. Global precipitation measurement. in *Precipitation: Advances in Measurement, Estimation and Prediction* vol. Springer 131–169 (Springer, 2008).
228. Hou, A. Y. *et al.* The global precipitation measurement mission. *Bull Am Meteorol Soc* **95**, 701–722 (2014).
229. Huffman, G. J. *et al.* NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). *Algorithm Theoretical Basis Document (ATBD) Version 4*, (2015).
230. Huffman, G. J. *et al.* The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J Hydrometeorol* **8**, 38–55 (2007).
231. Joyce, R. J., Janowiak, J. E., Arkin, P. A. & Xie, P. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J Hydrometeorol* **5**, 487–503 (2004).
232. Hsu, K.-L., Gao, X., Sorooshian, S. & Gupta, H. V. Precipitation estimation from remotely sensed information using artificial neural networks. *Journal of Applied Meteorology* **36**, 1176–1190 (1997).
233. Sorooshian, S., Hsu, K.-L., Gao, X. & Gupta, H. v & Braithwaite, D. Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull Am Meteorol Soc* **81**, 2035–2046 (2000).
234. Xie, P., Chen, M. & Shi, W. CPC unified gauge-based analysis of global daily precipitation. *Preprints, 24th Conf. on Hydrology, Atlanta, GA, Amer. Meteor. Soc.* **2**, (2010).
235. Harris, I., Osborn, T. J., Jones, P. & Lister, D. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci Data* **7**, (2020).
236. Schneider, U. *et al.* GPCP's new land surface precipitation climatology based on quality-controlled *in situ* data and its role in quantifying the global water cycle. *Theor Appl Climatol* **115**, 15–40 (2014).
237. Chen, M., Xie, P., Janowiak, J. E. & Arkin, P. A. Global Land Precipitation: A 50-yr Monthly Analysis Based on Gauge Observations. *Journal of Hydrometeorology* **3**, 249–266 (2002).
238. Beck, H. E. *et al.* MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrol Earth Syst Sci* **21**, 589–615 (2017).
239. Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A. & Hegewisch, K. C. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Sci Data* **5**, (2018).
240. Friedl, M. A. *et al.* Global land cover mapping from MODIS: algorithms and early results. *Remote Sens Environ* **83**, 287–302 (2002).
241. Loveland, T. R. *et al.* Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *Int J Remote Sens* **21**, 1303–1330 (2000).

Acknowledgements

This study was funded by the National Science Foundation (Belmont Forum Project, grant No.- 2127643 and CAREER, grant 1752729) and Henry Luce Foundation. We acknowledge that various datasets synthesized here were collected as a part of two previous NASA projects (Awards: 2127643 and 1752729). We acknowledge the Mekong River Commission for providing various *in-situ* datasets. We thank Nathan Torbick for providing ground-based soil moisture data for Thailand. J.L. acknowledges the support by the Strategic Priority Research Program of the Chinese Academy of Sciences (grant no. XDA20060402), and the Henan Provincial Key Laboratory of Hydrosphere and Watershed Water Security.

Author contributions

A.D.T. and Y.P. designed the study. A.D.T. drafted the manuscript with contribution from Y.P. A.D.T. performed data search, acquisition, and analyses. D.K. acquired the census-based irrigation attributes for Cambodia and helped to prepare socioeconomic dataset. All authors performed quality evaluation and revision of the manuscript.

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/s41597-023-02193-0>.

Correspondence and requests for materials should be addressed to Y.P.

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) 2023