

# Making China's water data accessible, usable and shareable

Received: 27 July 2022

Accepted: 2 February 2023

Published online: 6 April 2023

 Check for updates

Jingyu Lin<sup>1,2,3,34</sup>, Brett A. Bryan<sup>3,34</sup>, Xudong Zhou<sup>4</sup>, Peirong Lin<sup>5,6</sup>, Hong Xuan Do<sup>7,8</sup>, Lei Gao<sup>9</sup>, Xinchen Gu<sup>10</sup>, Zhifeng Liu<sup>11</sup>, Luwen Wan<sup>12</sup>, Shanlin Tong<sup>13</sup>, Jiacong Huang<sup>14</sup>, Qian Wang<sup>15</sup>, Yuan Zhang<sup>1,2</sup>, Hongkai Gao<sup>16</sup>, Jiabo Yin<sup>13</sup>, Zilong Chen<sup>17</sup>, Weili Duan<sup>18</sup>, Zheyu Xie<sup>17</sup>, Tong Cui<sup>19</sup>, Junzhi Liu<sup>20</sup>, Mingqian Li<sup>21</sup>, Xiaodong Li<sup>22</sup>, Zhenwu Xu<sup>23,24</sup>, Fei Guo<sup>3,25</sup>, Lele Shu<sup>26</sup>, Bin Li<sup>27</sup>, Jingwen Zhang<sup>28</sup>, Pingping Zhang<sup>29</sup>, Bo Fan<sup>30</sup>, Yanwen Wang<sup>31</sup>, Yifan Zhang<sup>32</sup>, Jinliang Huang<sup>17</sup>, Xin Li<sup>33</sup>, Yanpeng Cai<sup>1,2</sup>✉ & Zhifeng Yang<sup>1,2</sup>

Water data are essential for monitoring, managing, modelling and projecting water resources. Yet despite such data—including water quantity, quality, demand and ecology—being extensively collected in China, it remains difficult to access, use and share them. These challenges have led to poor data quality, duplication of effort and wasting of resources, limiting their utility for supporting decision-making in water resources policy and management. In this Perspective we discuss the current state of China's water data collection, governance and sharing, the barriers to open-access water data and its impacts, and outline a path to establishing a national water data infrastructure to reform water resource management in China and support global water-data sharing initiatives.

Water has been central to the development of China's ancient civilization and to its contemporary development and industrialization, and will shape its path to an ecological civilization—an aspirational, more sustainable mode of human interaction with the environment<sup>1</sup>. Chinese understanding of water and its management stems from the Neolithic period, supporting agriculture, irrigation and intensive cultivation. This has also involved extensive engineering and hydrology, as well as drought and flood management (for example, Yu the Great had a legendary ability to control floods around 4,100 yr BP) and defined Chinese culture (the Liangzhu culture developed in the context of extreme hydrological dynamics in the Yangtze delta around 5,100 yr BP; ref. 2). Today, and in the years to come, China's burgeoning economy, cities and population desperately need access to scarce freshwater resources to support industry and ensure food and energy security<sup>3–5</sup>, to manage the impacts of increasingly frequent and severe droughts and floods, and to improve the quality of water resources and water-dependent ecosystems given the impacts of industrialization and urbanization. Ready access to comprehensive, high-quality data on multiple aspects of surface, ground and coastal water resources (including quantity, quality, ecology, demand and infrastructure<sup>6,7</sup>;

Table 1) is fundamental to understanding, managing and monitoring China's water resources and meeting the nation's commitment to the UN Sustainable Development Goals (SDGs), particularly SDG 6 (clean water and sanitation; refs. 8–10). However, although water data are extensively collected in China, governance and sharing lag well behind the international community.

In this Perspective we use the term 'data governance' to refer to the principles, practices, roles, standards and metrics necessary to ensure the security, availability, usability and integrity of data. By adopting the terminology of Cantor et al.<sup>6</sup>, we use the term water data to describe a broad range of data and information that support decision-making and research on water-related topics, including data to characterize and monitor water-dependent ecosystems. Given that water data are produced by many different entities, including government agencies, academia, industry and non-governmental organizations, water data can be grouped into authoritative (that is, certified and provided by an authoritative source such as government) and non-authoritative data<sup>7</sup>. We focus here on government provision of authoritative data from the first five water categories in Table 1 (that is, W1–W5), which are directly related to water resources.

Here we assess the state of China's water data, discuss the barriers to water data sharing and provide a practical framework for modernizing its water data infrastructure. The assessment is based on our collective experience and a review of the literature and available data sources. In addition, to support our assessment, we conducted an anonymous online survey published on the Tencent Questionnaire Platform and promoted via WeChat from 20 February 2022 to 6 March 2022 (see Word File 1 at <https://doi.org/10.6084/m9.figshare.21532908>). The questionnaire was posted on the Hydro90 WeChat channel to target the community of water scientists actively using water data in analyses and modelling for water resource management in China. The survey was also promoted via our personal WeChat connections. The survey was read by 845 people, and completed by 305 people (36%). We asked participants about the types of water data they lacked, where their data were sourced, difficulties in sourcing data, their willingness to share water data (used in published papers), reasons for the lack of data availability and recommendations for improving data access and sharing. Most respondents (88%) were early-career researchers (for example, research students, young academics and young professionals; Fig. 1a).

## State of water data collection, governance and sharing

Water data are extensively collected in China under the responsibilities of multiple government agencies (Table 1). Most hydrometric stations for collecting water resource data are managed by the Ministry of Water Resources (MWR), and these have increased in number from 353 in the 1950s to ~120,000 (~80,000 of which are automated) in 2020<sup>11</sup>. The MWR also supervises nearly 100,000 dams and reservoirs<sup>12</sup>, following growing investment in water conservancy projects from \$0.01 billion (in 2021 US\$) in 1950 to a peak of ~\$127 billion in 2020 (see Data File 1 at <https://doi.org/10.6084/m9.figshare.21532908>). Beyond that, many stations have been established and are overseen by other agencies. For example, the Ministry of Ecology and Environment (MEE) supervises 3,641 surface water quality monitoring stations (2,024 automated)<sup>13</sup> and maintains approximately 10,826 municipal wastewater treatment plants (capacity of over 500 m<sup>3</sup> of sewage discharge per day). The Ministry of Natural Resources also built and manages 10,171 groundwater monitoring stations (10,168 automated), which generate over 90 million data points per year<sup>14,15</sup>, and manages 1,503 stations for collecting coastal water quality data<sup>16</sup> (<http://ep.nmemc.org.cn:8888/Water/>).

However, these data are not widely shared and obtaining water data is extremely difficult (Fig. 1d). For example, most time series of streamflow data have been collected by local Hydrology Bureau branches (managed by the MWR) and published in hard-copy annual yearbooks, with no digital versions available online<sup>17</sup>. Scientists who want to work on China's water resource issues must first source the relevant hard-copy yearbook from a local library, university library or bureau (Fig. 1b) and then digitize the streamflow data themselves. Surface water quality data collected by MWR is not published even in hard copy and hence are difficult to find and verify. Most water demand and use data are published in the Water Resources Bulletin and formatted as a report each year for each city separately, which is also time-consuming to source, digitize, curate and use. While researchers often share digitized water data among their networks, these actions lack coordination and result in unnecessary duplication of effort.

Some water data are made available digitally via national data platforms (for example, see the list of streamflow data in Data File 2 at <https://doi.org/10.6084/m9.figshare.21532908>), but the quality is often poor with temporal discontinuities and limited sites. Data can also be secondary, descriptive, summarized (for example, see water quality reports in Data File 2 at <https://doi.org/10.6084/m9.figshare.21532908>), modelled (for example, estimated and reanalysis data products; see streamflow data in Data File 2 at <https://doi.org/10.6084/m9.figshare.21532908>) or presented in a user-unfriendly

way (for example, data in the format of Microsoft Word files, see weekly water quality data in Data File 2 at <https://doi.org/10.6084/m9.figshare.21532908>). Essential metadata information (for example, locations or attribute descriptions) describing hydrometric sites available on common data platforms is frequently missing or unclear (see Data File 2 at <https://doi.org/10.6084/m9.figshare.21532908> and Fig. 1c) and there are no publicly accessible standards for this information. Most water datasets on these platforms are only available on an ad hoc basis and access to each item is subject to approval.

Water data for China available from online international repositories, particularly streamflow and water quality data, are few and outdated and China lags most other nations of comparable land area in making this data available (Table 2). For example, the Global Runoff Data Centre (GRDC), which supports large-scale hydrological studies, provides hydrological data for 10,702 stations (as of 29 April 2022) worldwide. However, only 39 of these stations are in China and these records have not been updated since 2004. Similarly, the Global Streamflow Indices and Metadata (GSIM) project was initiated to collate publicly available data and promote the widespread use of streamflow data, with 30,959 stations in total<sup>18</sup>. However, Chinese streamflow data are only available for rivers in Yunnan Province and the Tibet Autonomous Region maintained by the China Hydrology Data Project<sup>19</sup>. This includes just 163 stations with records spanning 1947 to 1987 that have not been updated since the establishment of GSIM<sup>18</sup>. Aggregating five large water-quality datasets, the Global River Water Quality Archive (GRQA) comprehensively reports 42 indicators for 93,057 sites with records from 1898 to 2020. There are 244 of these stations located in China with only 3,595 observations in total, and daily records are only available from 1980 to 2009<sup>20</sup>. However, water infrastructure and utilities data (for example, wastewater treatment plants, dams and reservoirs) for China is much better represented in global databases (Table 2).

Some live, direct, real-time water quality, streamflow and water level data for China's surface water resources are temporarily viewable online at the website of the China National Environmental Monitoring Centre, which provides 2,024 automated surface water quality stations and the website of National Hydraulics and Hydroinformatics, which provides data for 1,145 hydrological stations. This has led some researchers to resort to web scraping and other automated means of assembling and cleaning real-time water data. However, historical data are not accessible to the public<sup>21</sup>.

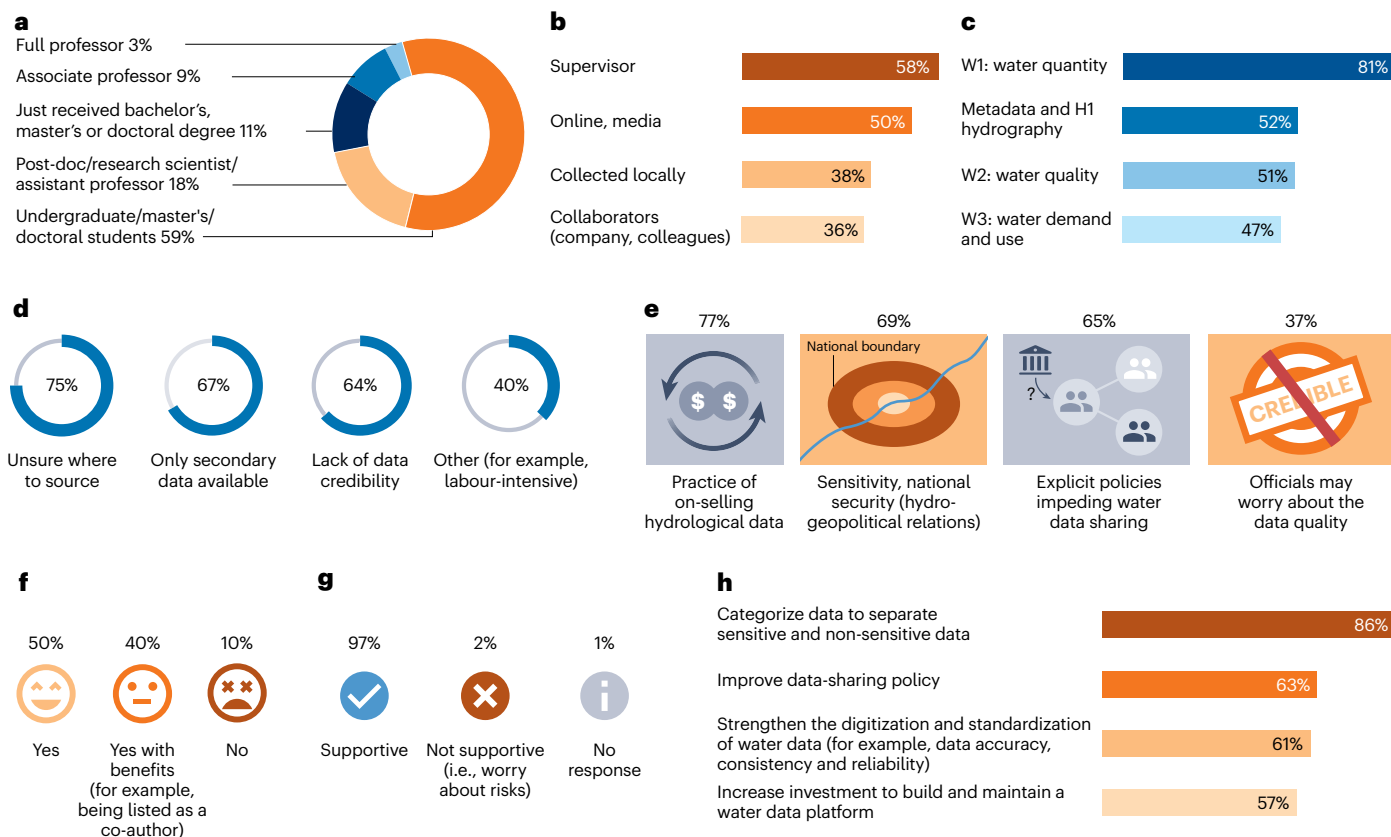
## Barriers and impacts

There have been many barriers to sharing China's water data. Of our survey respondents, 77% believed that the greatest barrier to the accessibility of water data is the widespread practice of onselling hydrological data and the high costs involved for users. For example, purchasing a time series of daily streamflow data for a single year at just one station costs around US\$1,023. Although a cost-recovery model may have been reasonable in the early stages of data collection due to the labour-intensive nature of manual water resource monitoring, ongoing high data costs are difficult to justify with the widespread adoption of automated monitoring equipment and hydroinformatic techniques. Concerns around national security and hydro-geopolitical relations as a barrier to water data sharing were expressed by 69% of respondents. Water plays an important role in national security and regional geopolitics for China, which shares 42 major transboundary rivers with neighbouring countries<sup>22</sup>. For example, countries surrounding the Tibetan Plateau, which share multiple major transboundary rivers including the Mekong, Brahmaputra and Indus, have already experienced conflicts around water resource management, the impacts of global warming, atmospheric circulation changes and increased water demand<sup>23</sup>. With water a key resource and foundation of national security and development, water data in China have historically been considered to be highly sensitive. Sixty-five percent of respondents also

**Table 1 | Summary and descriptions of water data categories under the SDGs and the Chinese government agencies responsible**

Category	Summary of category	Agencies	SDG indicators	Water resource category	
				Cantor et al. <sup>6</sup>	Larson et al. <sup>7</sup>
W1 Water quality	Spatial, quantitative and metadata information about the quantity and timing of water, including surface water (for example, stream, lake, snow, glacier) and groundwater (for example, well, soil moisture) monitoring and modelled data, water storage and height, and aquifers.	Ministry of Water Resources, Ministry of Natural Resources	6.1.1	Water supply, groundwater, water storage	Hydrology, flood control
W2 Water quality	Spatial, quantitative and metadata information about drinking water, surface water, groundwater and coastal/ocean water quality (for example, pH, desalination, temperature, dissolved oxygen, water chemistry, sediments, contaminants and so on).	Ministry of Water Resources, Ministry of Ecology and Environment, Ministry of Natural Resources	6.3.2, 14.1	Water quality, groundwater	Water quality
W3 Water demand and use	Spatial, quantitative and metadata information about water demand/ use for public supply, domestic use, irrigation, thermoelectric power, industry, mining, livestock, aquaculture and recreation; water rights; and water use efficiency, recycling and reuse technologies.	Ministry of Natural Resources, Ministry of Housing and Urban-Rural Development, Ministry of Agriculture and Rural Affairs, National Bureau of Statistics	6.2.1, 6.4.1, 6.4.2, 6.a	Agriculture, water demand and use, land use	Water availability and use, recreation and aesthetic importance
W4 Water pollutant loadings	Spatial, quantitative and metadata information about water pollutant loadings or sewage discharge from different users (for example, domestic, industry, livestock manure).	Ministry of Ecology and Environment, Ministry of Natural Resources, Ministry of Agriculture and Rural Affairs, National Bureau of Statistics	6.3.1	–	–
W5 Water infrastructure and utilities	Spatial, quantitative and metadata information about hydroelectric dams, reservoirs, water pipes, water deliveries, pumps, wastewater treatment plants, flood control infrastructure, stormwater infrastructure, sanitation services and equipment.	Ministry of Water Resources, Ministry of Ecology and Environment, Ministry of Agriculture and Rural Affairs, National Development and Reform Commission, National Health Commission	6.3.1	Water demand and use, infrastructure and utilities, water storage	Hydropower, dams, flood control
W6 Water policies	Historical, current and future water-related policies, strategies, law, actions, plans, projects, programmes and management.	Ministry of Water Resources, Ministry of Ecology and Environment, Ministry of Natural Resources, Ministry of Agriculture and Rural Affairs, National Development and Reform Commission	6.5.2, 6.6.1, 6.a, 6.b, 13.2	–	–
H1 Hydrography	Spatial and quantitative data about the physical characteristics of basins, sub-basins, watersheds, streams, stream order, canals, pipelines and so on.	Ministry of Water Resources, Ministry of Natural Resources, Ministry of Housing and Urban-Rural Development	6.5.1, 6.5.2	Mapping and modelling	Hydrography
H2 Hydroecology	Spatial and quantitative data about the presence of sensitive aquatic species, habitats, biodiversity, bacteria, algal blooms and biological indicators (for example, fish counts, invasive species, estimates of wildlife population and extent, fish migration).	Ministry of Ecology and Environment, Ministry of Natural Resources, National Forestry and Grassland Administration	6.6.1, 14.6, 15.8	Ecology	Ecology, migratory barriers
H3 Hydroclimatology	Spatial and quantitative data depicting current or future regional climate (temperature and precipitation (rainfall, snowfall), atmospheric water) and climate-informed hydrology.	China Meteorological Administration	13.1	Weather and climate	Hydroclimatology

Six water-specific categories are identified and labelled W1–W6, and three hydrogeographic categories are denoted H1–H3, referring to the categories of Cantor et al.<sup>6</sup> and Larson et al.<sup>7</sup>



**Fig. 1 | China's water data survey results. a**, Respondent profile. **b**, Most common sources of data. **c**, Top four water data aspects lacking. **d**, Top four difficulties encountered when sourcing data. **e**, Top four reasons for data being unavailable. **f**, Willingness to share data from respondents' published papers. **g**, Opinions on the call for better water data sharing. **h**, Top four ways to effectively boost data sharing. Data are available in Data File 3 at <https://doi.org/10.6084/m9.figshare.21532908>.

identified the lack of explicit policies as a barrier for promoting water data sharing (Fig. 1e) including data privacy laws, security standards and data governance protocols for data providers; restricted data use policies and data publishing strategies for end users; standards for metadata and data interoperability; policies for promoting supporting data infrastructure and technology; and policies for connecting data providers and users (for example, incentive policies, engagement, user networks)<sup>6,24</sup>. Although China has existing legal structures administering scientific data<sup>25</sup>, stronger policies (acts, laws or executive orders) that are targeted towards water data and specify clear roles for different agencies are required.

Inadequate data sharing and overlapping responsibilities has resulted in competition between multiple agencies and stakeholders from local to national levels in establishing, managing and monitoring gauging stations. For instance, the MWR established the Zhimenda station at the outlet of the source of the Yangtze River (33.013° N, 97.238° E) which has collected water data (for example, streamflow, water quality, and sediment) since 1956. This station is very close to another station also called Zhimenda (33.022° N, 97.248° E) established by the MEE to collect water quality data since 1999<sup>26</sup>. Water quality sampling programmes have also been implemented for hundreds of lakes and rivers of high concern by different stakeholders. For example, multiple water sampling programmes for Lake Taihu are conducted by water management agencies, institutes and universities, leading to a considerable waste of resources and time, inconsistent data quality, varied results of hydrological analyses and divergent policy recommendations<sup>27</sup>.

The lack of an openly available dense network of digital, high-quality, up-to-date, continuous, standardized, long-term daily or

hourly data capturing key water indicators has limited the credibility of China's water information (Fig. 1e). More importantly, it has hampered the ability of Chinese and international scientists to contribute to the sustainable management of water resources in China and globally, impeded the development of China's next generation of water scientists and stymied the growth of water resource science, management and policy<sup>28</sup>. Although some studies have used original national-scale streamflow data in China (see Data File 1 at <https://doi.org/10.6084/m9.figshare.21532908>), the datasets underlying these studies are not publicly available due to licensing restrictions, which limits the reproducibility of the work.

### An emerging willingness to act but more is needed

Efforts have been made towards improving the quality and availability of water data in China. More than 20 years ago, the MWR launched a notice on the disclosure of public-welfare water data. In 2007, the Deputy Minister of the MWR announced that China would strive to set up a national water database to boost data sharing<sup>29</sup>. Just 2 years ago, the MWR announced measures for assembling and managing water resource information<sup>30</sup>. China's government also recently pledged to share year-round Mekong River data to downstream Asian countries for better monitoring and forecasting of floods and droughts through the Lancang-Mekong Cooperation Mechanism<sup>31</sup>, and published a detailed report proposing the construction of digital twin basins with big water data to enhance smart water conservancy in early 2022<sup>32</sup>. Despite these initiatives, the systematic collection, governance and sharing of high-quality water data in China has not gained momentum.

**Table 2 | Data available in global water datasets from the seven largest nations by area**

Country	Number of sites				
	Water quantity (streamflow)		Water quality	Water infrastructure and utilities	
	GRDC (2022)	GSIM (streamflow, 2018)	GRQA (2021)	HydroWASTE (wastewater treatment plant, 2021 <sup>50</sup> )	GeoDAR V11 (dams and reservoirs, 2020 <sup>51</sup> )
Russia	1,491	480	233	1,269	53
Canada	1,153	6,325	1,737	2,064	562
China	39	189*	244	2,486	5,282
United States	1,021	9,404	74,917	14,819	7,958
Brazil	487	3,313	2,822	2,820	767
Australia	701	2,941	900	1,234	351
India	44	318	254	816	1,291

\*Comprises 163 sites from the China Hydrology Data Project and 26 sites from the GRDC recorded by 2018. Note that water quantity data in the GRDC is actively updated and the data counts were accurate as of 29 April 2022. Date of last data upload is as indicated for other datasets.

As the next generation of water scientists working to understand and manage China’s water resources, we watch with envy the ever-expanding open global provision of data for supporting water resource monitoring, modelling and prediction. For example, real-time streamflow records for approximately 8,500 gauging stations and water quality for about 2.7 million sites (covering both surface and groundwater resources in inland and coastal areas) in the United States are available on the National Water Information System and Water Quality Portal<sup>33</sup> (<https://waterdata.usgs.gov/nwis>). This data resource is supporting the US National Water Model for real-time flood forecasting and will support the Surface Water and Ocean Topography mission, which will further enhance estimates of Earth’s surface water from 2022<sup>34</sup>. Based on the GRDC database, the European Union has built continental-scale hydrological models to assess the impacts of climate change on water scarcity<sup>35</sup>. This information has enabled France to develop the Observatories National des Étiages network for in-depth modelling of headwater hydrological processes. Although some other countries do not have a network of hydrologic stations as dense as China’s, they have pushed forward with the publication of non-sensitive data online<sup>36,37</sup>. These efforts are pertinent examples of the important role of national water agencies in generating major new advances in water data as a catalyst for promoting evidence-based water resource management policy.

Given the need to accelerate progress towards SDG 6 and to manage the looming impacts of climate change, the availability of water data is becoming even more important and urgent to underpin hydrological and interdisciplinary water research. Water data are a crucial asset in developing early warning systems for flooding; providing information to support decision-making in drought planning and management, agricultural and irrigation management, and integrated river and basin management; and developing a modelling and strategic foresighting capacity for managing water resources under future uncertainty<sup>10</sup>. These toolsets are urgently needed as more and more research points to the accelerating impacts of compounding extremes of weather and natural hazards<sup>38</sup>, as well as surging demand for clean water—especially during crises such as the COVID-19 pandemic<sup>39</sup>. With the prevalence of artificial intelligence, data-driven modelling is increasingly complementing process-based modelling, making it a fast-moving field connecting water scientists from all countries in countering the adverse impacts of climate change. These issues are likely to be exacerbated over the next few decades due to intensified international conflicts over water.

### A path forward

We call upon the Chinese government to reform and provide open access to its water data, and to coordinate its water-related administrative

sectors and transform water data governance and sharing via integrating the now fragmented components into a connected, specialized, national water data infrastructure to advance water resource management (Fig. 1g,h). Below we propose the priority, e-technology, archetypes, cooperation and engagement (PEACE) framework to illustrate the major components of such a modern data infrastructure (Fig. 2).

#### Priority

The digitization, governance, internationalization and open provision of the full historical record of non-sensitive gauging station water data within China needs to be prioritized according to national rules and scientific data strategies<sup>25,40</sup>. As recommended by the Jiangsu Committee of the Chinese People’s Political Consultative Conference<sup>41</sup>, stronger policies and laws are also urgently required to prevent onselling of basic and non-sensitive water data.

#### E-technology

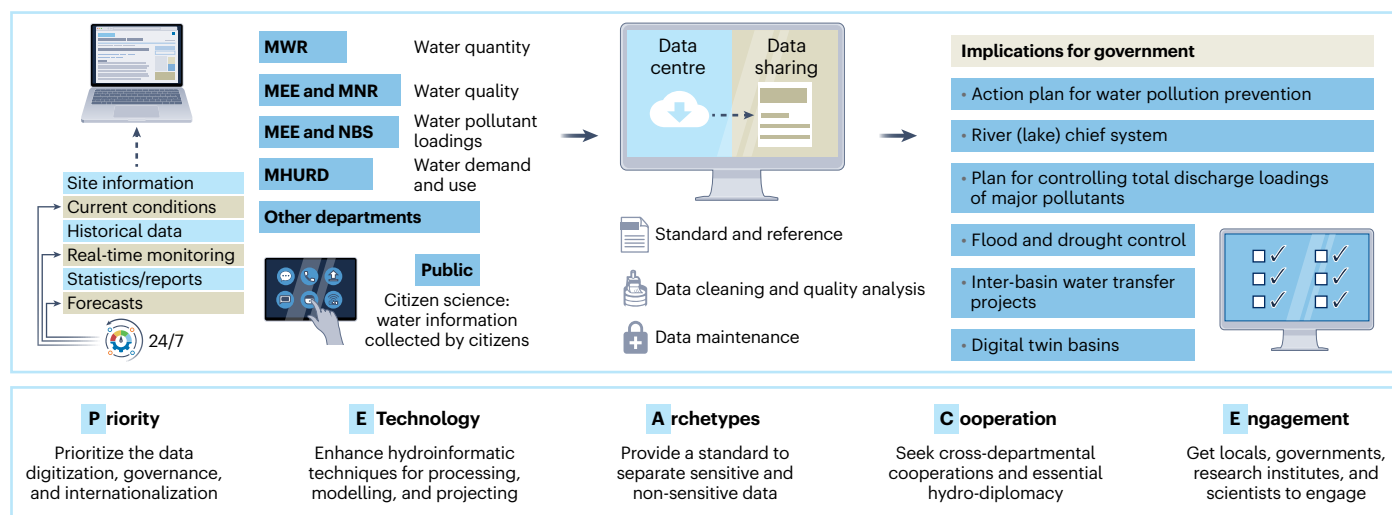
Next-generation automatic water resource monitoring, modelling, projection, standards, services and wireless sensing are required, with a focus on data transparency and openness. Digital technology will play a key role in the usability of water data via integrating multi-source and heterogeneous data into interoperable data. For example, the great success of the Google Earth Engine platform has transformed big data into information that meets diverse user needs. Developed by a group at Stanford University, the CEDAR Workbench illustrates a potential solution for creating national standardized metadata<sup>42</sup>.

#### Archetypes

We urge the establishment of archetypes for specific standards, rules and regulations to identify water information (for example, sensitive data with regards to national security) for restricted data sharing following the Water Law of the People’s Republic of China 2017 Edition<sup>40</sup> and the Measures for the Administration of Scientific Data in 2018<sup>25</sup>. Water data are inherently sensitive, especially in transboundary river systems and areas subject to geopolitical and resource conflicts. It is therefore important for the related agencies to classify the regions and associated data clearly to avoid misuse<sup>24</sup> and allay national security concerns.

#### Cooperation

We appeal for cross-departmental cooperation to reduce duplication of effort and expand transboundary cooperation to seek basin-level integrated water resource management. The exchange of water information between different departments can improve the value discovery and credibility of water data across interdisciplinary producers<sup>7</sup>. Hydro-diplomacy is also essential to bring different stakeholders



**Fig. 2 | Anticipating water data sharing challenges with PEACE.** The PEACE framework offers a connected, specialized infrastructure for advancing water resources management. MHURD, Ministry of Housing and Urban-Rural Development; MNR, Ministry of Natural Resources; NBS, National Bureau of

Statistics. River (lake) chief system is a policy to address water governance via assigning each part of a river/lake to a certain official and associating the status of a river/lake (i.e., water quality) with the overall assessment of officials.

and countries in these conflict-prone river basins together to mediate hydro-geopolitical conflict via dialogue and cooperation<sup>43</sup>.

### Engagement

We encourage research institutes, scientists, industry, local communities and non-governmental organizations to engage, and governments (for example, different administrative divisions from regional to local level) to improve corporate data governance following the principles of transparency, responsibility, user focus, sustainability and technology (TRUST)<sup>44</sup> and findable, accessible, interoperable and reusable data (FAIR)<sup>45</sup>. In particular, empowering the participation of women would promote gender equality and diversity in water resource management<sup>46</sup>. Ongoing stakeholder engagement helps to ensure that investment will ultimately be worthwhile for different data users. Lessons from California's efforts to promote a data-driven water platform indicated that stakeholder engagement helped to expand data system usability and awareness, even though it requires time, resources, and commitment<sup>6</sup>.

The above transformation will set China on a path to more effective and efficient water data governance characterized by increased transparency and accountability, and promote trust and credibility in institutions and government. Efficiencies gained via reduced duplication of effort will save money and resources. Sharing high-quality data will not only increase the value of data to stakeholders, but also create a return on investment of 10–20 times for data users, as well as a 20–50 times greater spill-over value for the wider economy<sup>47</sup>. One US dollar invested in data services alone has been found to provide returns of up to US\$8.30 (ref. 48). In addition, the community can connect and engage in citizen science to improve awareness of water-related issues such as scarcity and pollution<sup>49</sup>. Open science and data sharing will promote collaboration between Chinese and international scientists, and boost understanding of and improve water resource management in China and globally. It will also enable China to make progress in water data governance and analysis, and facilitate its young scientists to develop the skills required to manage China's water resources sustainably<sup>28</sup> on the path towards an ecological civilization and the SDGs. An optimized water data infrastructure is essential to underpin China's contribution to international scientific programmes and to building a regional and global community with a shared aspiration to enhance our ability to manage rapidly changing water resources.

### Data availability

All Supplementary Materials are available via Figshare at <https://doi.org/10.6084/m9.figshare.21532908>.

### References

- Frazier, A. E. et al. Ecological civilization: perspectives from landscape ecology and landscape sustainability science. *Landscape Ecol.* **34**, 1–8 (2019).
- Liu, B. et al. Earliest hydraulic enterprise in China, 5,100 years ago. *Proc. Natl Acad. Sci. USA* **114**, 13637–13642 (2017).
- He, C. Y. et al. Future global urban water scarcity and potential solutions. *Nat. Commun.* **12**, 4667 (2021).
- Ur Rehman, S. A., Cai, Y. P., Mirjat, N. H., Walasai, G. D. & Nafees, M. Energy-environment-economy nexus in Pakistan: lessons from a PAK-TIMES model. *Energy Policy* **126**, 200–211 (2019).
- Cai, Y. P., Huang, G. H., Yang, Z. F. & Tan, Q. Identification of optimal strategies for energy management systems planning under multiple uncertainties. *Appl. Energy* **86**, 480–495 (2009).
- Cantor, A. et al. *Data for Water Decision Making: Informing the Implementation of California's Open and Transparent Water Data Act through Research and Engagement* (Center for Law, Energy & the Environment, UC Berkeley School of Law, 2018).
- Larson, K. B. *Improving Discovery, Sharing, and Use of Water Data* (Pacific Northwest National Laboratory, 2021).
- Yin, J. B. et al. Does the hook structure constrain future flood intensification under anthropogenic climate warming? *Water Resour. Res.* **57**, e2020WR028491 (2021).
- Yu, C. Q. et al. Managing nitrogen to restore water quality in China. *Nature* **567**, 516–520 (2019).
- Sadoff, C. W., Borgomeo, E. & Uhlenbrook, S. Rethinking water for SDG 6. *Nat. Sustain.* **3**, 346–347 (2020).
- 2020 Statistics Bulletin on China Water Activities (MWR, 2022).
- Song, C. et al. A comprehensive geospatial database of nearly 100,000 reservoirs in China. *Earth Syst. Sci. Data* **14**, 4017–4034 (2022).
- Report of the National Surface Water Quality (from January to June) (in Chinese) (MEE, 2022); [https://www.mee.gov.cn/ywdt/xwfb/202207/t20220718\\_989007.shtml](https://www.mee.gov.cn/ywdt/xwfb/202207/t20220718_989007.shtml)

14. National Groundwater Monitoring Project Completed. China Geological Survey [https://en.cgs.gov.cn/news1/201904/t20190409\\_478767.html](https://en.cgs.gov.cn/news1/201904/t20190409_478767.html) (30 January 2019).
15. China Improves Groundwater Monitoring System (State Council of the People's Republic of China, 2020); [http://english.www.gov.cn/statecouncil/ministries/202001/05/content\\_WS5e11e49ec6d0cee5d284b2ff.html](http://english.www.gov.cn/statecouncil/ministries/202001/05/content_WS5e11e49ec6d0cee5d284b2ff.html)
16. Site Information for China's Coastal Water Quality Monitoring in 2016. Ministry of Natural Resources of the People's Republic of China (MNR) [http://gc.mnr.gov.cn/201806/t20180614\\_1794783.html](http://gc.mnr.gov.cn/201806/t20180614_1794783.html) (2017).
17. Henck, A. C. et al. Spatial controls on erosion in the Three Rivers Region, southeastern Tibet and southwestern China. *Earth Planet. Sci. Lett.* **303**, 71–83 (2011).
18. Do, H. X., Gudmundsson, L., Leonard, M. & Westra, S. The Global Streamflow Indices and Metadata Archive (GSIM) - Part 1: the production of a daily streamflow archive and metadata. *Earth Syst. Sci. Data*, **10**, 765–785 (2018).  
**The most comprehensive and publicly available global streamflow data collection.**
19. Henck, A. C. *China Hydrology Data Project* (Oberlin College and Conservatory, 2012); <https://www2.oberlin.edu/faculty/aschmidt/chdp/index.html>
20. Virro, H., Amatulli, G., Kmoch, A., Shen, L. & Uuemaa, E. GRQA: global river water quality archive. *Earth Syst. Sci. Data* **13**, 5483–5507 (2021).
21. Huang, J. C. et al. Characterizing the river water quality in China: recent progress and on-going challenges. *Water Res.* **201**, 117309 (2021).
22. Xie, L. & Jia, S. *China's International Transboundary Rivers: Politics, Security and Diplomacy of Shared Water Resources* (Routledge, 2019).
23. Yao, T. D. et al. The imbalance of the Asian water tower. *Nat. Rev. Earth. Environ.* **3**, 618–632 (2022).
24. Li, X. et al. Boosting geoscience data sharing in China. *Nat. Geosci.* **14**, 541–542 (2021).  
**A Comment summarizing the progress that China has made in promoting geoscientific data sharing.**
25. Notice on the Measures for the Administration of Scientific Data (in Chinese). State Council of the People's Republic of China [http://www.gov.cn/zhengce/content/2018-04/02/content\\_5279272.htm](http://www.gov.cn/zhengce/content/2018-04/02/content_5279272.htm) (2018).
26. Water Data Sharing in Yangtze River Encounters the Embarrassment: Two Monitoring Sites are Separated By only a Wall, But They don't Share the Data (in Chinese). *Outlook Weekly* [http://lw.xinhuanet.com/2021-01/25/c\\_139694943.htm](http://lw.xinhuanet.com/2021-01/25/c_139694943.htm) (25 January 2021).
27. Qin, B. Q. et al. Are nitrogen-to-phosphorus ratios of Chinese lakes actually increasing? *Proc. Natl Acad. Sci. USA* **117**, 21000–21002 (2020).
28. Jain, S. et al. Are we at risk of losing the current generation of climate researchers to data science? *AGU Adv.* **3**, e2022AV000676 (2022).
29. China Will Set up a National Hydrological Database for Sharing Publicly Water Data (in Chinese). *China Xinhua News* [http://www.gov.cn/jrzq/2007-05/30/content\\_630913.htm](http://www.gov.cn/jrzq/2007-05/30/content_630913.htm) (30 May 2007).
30. *Measures for Assembling and Managing Water Resources Information* (MWR, 2020); [http://www.gov.cn/gongbao/content/2020/content\\_5570073.htm](http://www.gov.cn/gongbao/content/2020/content_5570073.htm)
31. Finney, R. China agrees to provide year-round Mekong River data to downstream countries. *Reuters* <https://www.rfa.org/english/news/china/agrees-10222020160430.html> (22 October 2020).
32. Li, D. R., Yu, W. B. & Shao, Z. F. Smart city based on digital twins. *Comput. Urban Sci.* **1**, 4 (2021).
33. Read, E. K. et al. Water quality data for national-scale aquatic research: the water quality portal. *Water Resour. Res.* **53**, 1735–1745 (2017).
34. Altenau, E. H. et al. The surface water and ocean topography (SWOT) mission river database (SWORD): a global river network for satellite data products. *Water Resour. Res.* **57**, e2020GL088946 (2021).
35. Abbaspour, K. C. et al. A continental-scale hydrology and water quality model for Europe: calibration and uncertainty of a high-resolution large-scale SWAT model. *J. Hydrol.* **524**, 733–752 (2015).  
**An empirical study using global monitoring water datasets and a distributed model to simulate large-scale hydrology and water quality.**
36. Addor, N. et al. Large-sample hydrology: recent progress, guidelines for new datasets and grand challenges. *Hydrol. Sci. J.* **65**, 712–725 (2020).
37. Trambly, Y. et al. ADHI: the African Database of Hydrometric Indices (1950-2018). *Earth Syst. Sci. Data* **13**, 1547–1560 (2021).
38. Zscheischler, J. et al. Future climate risk from compound events. *Nat. Clim. Change* **8**, 469–477 (2018).
39. van der Voorn, T., van den Berg, C., Bhattacharya, P. & Quist, J. Never waste a crisis: drawing first lessons from the COVID-19 pandemic to tackle the water crisis. *ACS EST Water* **1**, 8–10 (2020).
40. *Water Law of the People's Republic of China 2017 Edition* (State Council of the People's Republic of China, 2017); [http://www.gov.cn/gongbao/content/2017/content\\_5219177.htm](http://www.gov.cn/gongbao/content/2017/content_5219177.htm)
41. *Open Hydrological Data to Realize Public Resource Sharing* (Jiangsu Committee of the Chinese People's Political Consultative Conference, 2012); [http://www.jszx.gov.cn/zxta/2012ta/201202/t20120211\\_13617.html](http://www.jszx.gov.cn/zxta/2012ta/201202/t20120211_13617.html)
42. Musen, M. A. Without appropriate metadata, data-sharing mandates are pointless. *Nature* **609**, 222–222 (2022).
43. Zhang, H. Z. & Li, M. J. A process-based framework to examine China's approach to transboundary water management. *Int. J. Water Resour. Dev.* **34**, 705–731 (2018).
44. Lin, D. W. et al. The TRUST Principles for digital repositories. *Sci. Data* **7**, 144 (2020).  
**An empirical study proposing guiding principles to demonstrate the trustworthiness of digital repositories and sharing.**
45. Wilkinson, M. D. et al. The FAIR guiding principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016).
46. Coles, A. & Wallace, T. *Gender, Water, and Development* (Routledge, 2021).
47. *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use Across Societies* (OECD, 2019).
48. Beagrie, N. & Houghton, J. *The Value and Impact of the Archaeology Data Service: A Study and Methods for Enhancing Sustainability* (Charles Beagrie Ltd and Victoria University, 2013).
49. Jollymore, A., Haines, M. J., Satterfield, T. & Johnson, M. S. Citizen science for water quality monitoring: data implications of citizen perspectives. *J. Environ. Manage.* **200**, 456–467 (2017).
50. Ehalt, H. M. et al. Distribution and characteristics of wastewater treatment plants within the global river network. *Earth Syst. Sci. Data* **14**, 559–577 (2022).
51. Wang, J. et al. GeoDAR: georeferenced global dams and reservoirs dataset for bridging attributes and geolocations. *Earth Syst. Sci. Data* **14**, 1869–1899 (2022).

## Acknowledgements

We thank the Hydro90 Research community in China for supporting this research. We thank Q. Tang, P. Yang, F. Li and Y. Xu for their constructive comments. This work was supported by the Program for Guangdong Introducing Innovative and Entrepreneurial Teams (grant

number 2019ZT08L213), the National Natural Science Foundation of China (grant numbers 52200213, 52239005 and 42222104) and the Key-Area Research and Development Program of Guangdong Province (grant number 2020B1111380003).

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** should be addressed to Yanpeng Cai.

**Peer review information** *Nature Water* thanks Xiaowei Jin and Kyle Larson for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© Springer Nature Limited 2023

<sup>1</sup>Guangdong Provincial Key Laboratory of Water Quality Improvement and Ecological Restoration for Watersheds, School of Ecology, Environment and Resources, Guangdong University of Technology, Guangzhou, China. <sup>2</sup>Southern Marine Science and Engineering Guangdong Laboratory, Guangzhou, China. <sup>3</sup>School of Life and Environmental Sciences, Deakin University, Burwood, Victoria, Australia. <sup>4</sup>Global Hydrological Prediction Center, Institute of Industrial Science, University of Tokyo, Tokyo, Japan. <sup>5</sup>Institute of Remote Sensing and GIS, School of Earth and Space Sciences, Peking University, Beijing, China. <sup>6</sup>International Research Center for Big Data for Sustainable Development Goals, Beijing, China. <sup>7</sup>Faculty of Environment and Natural Resources, Nong Lam University - Ho Chi Minh City, Ho Chi Minh City, Vietnam. <sup>8</sup>Center for Technology Business Incubation, Nong Lam University - Ho Chi Minh City, Ho Chi Minh City, Vietnam. <sup>9</sup>Commonwealth Scientific and Industrial Research Organisation (CSIRO), Urrbrae, South Australia, Australia. <sup>10</sup>State Key Laboratory of Hydraulic Engineering Simulation and Safety, School of Civil Engineering, Tianjin University, Tianjin, China. <sup>11</sup>State Key Laboratory of Earth Surface Processes and Resource Ecology (ESPRE) and School of Natural Resources, Faculty of Geographical Science, Beijing Normal University, Beijing, China. <sup>12</sup>Department of Earth and Environmental Sciences, Michigan State University, East Lansing, MI, USA. <sup>13</sup>State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan, China. <sup>14</sup>Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, China. <sup>15</sup>Institute of Water Science and Technology, Hohai University, Nanjing, China. <sup>16</sup>Key Laboratory of Geographic Information Science (Ministry of Education of China), East China Normal University, Shanghai, China. <sup>17</sup>College of the Environment and Ecology, Xiamen University, Xiamen, China. <sup>18</sup>Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, China. <sup>19</sup>Department of Hydraulic Engineering, State Key Laboratory of Hydrosience and Engineering, Tsinghua University, Beijing, China. <sup>20</sup>Center for the Pan-Third Pole Environment, Lanzhou University, Lanzhou, China. <sup>21</sup>Key Laboratory of Groundwater Resources and Environment, Ministry of Education, Jilin University, Changchun, China. <sup>22</sup>College of Hydrology and Water Resources, Hohai University, Nanjing, China. <sup>23</sup>Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China. <sup>24</sup>University of Chinese Academy of Sciences, Beijing, China. <sup>25</sup>State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing, China. <sup>26</sup>Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou, China. <sup>27</sup>School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, China. <sup>28</sup>School of Civil Engineering, Center of Water Resources and Environment, Sun Yat-sen University, Zhuhai, China. <sup>29</sup>College of Water Conservancy and Civil Engineering, South China Agricultural University, Guangzhou, China. <sup>30</sup>State Key Laboratory of Pollution Control and Resource Reuse, School of the Environment, Nanjing University, Nanjing, China. <sup>31</sup>School of Economics and Management, China University of Geosciences, Wuhan, China. <sup>32</sup>College of Soil and Water Conservation, Beijing Forestry University, Beijing, China. <sup>33</sup>National Tibetan Plateau Data Center, State Key Laboratory of Tibetan Plateau Earth System, Environment and Resources, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China. <sup>34</sup>These authors contributed equally: Jingyu Lin, Brett A. Bryan. ✉e-mail: [yanpeng.cai@gdut.edu.cn](mailto:yanpeng.cai@gdut.edu.cn)