National Tibetan Plateau Data Center

Promoting Earth System Science on the Third Pole

Xiaoduo Pan, Xuejun Guo, Xin Li, Xiaolei Niu, Xiaojuan Yang, Min Feng, Tao Che, Rui Jin, Youhua Ran, Jianwen Guo, Xiaoli Hu, and Adan Wu

ABSTRACT: The Tibetan Plateau, known as the world’s “Third Pole” due to its high altitude, is experiencing rapid, intense climate change, similar to and even far more than that occurring in the Arctic and Antarctic. Scientific data sharing is very important to address the challenges of better understanding the unprecedented changes in the Third Pole and their impacts on the global environment and humans. The National Tibetan Plateau Data Center (TPDC, <http://data.tpdac.ac.cn>) is one of the first 20 national data centers endorsed by the Ministry of Science and Technology of China in 2019 and features the most complete scientific data for the Tibetan Plateau and surrounding regions, hosting more than 3,500 datasets in diverse disciplines. Fifty datasets featuring high-mountain observations, land surface parameters, near-surface atmospheric forcing, cryospheric variables, and high-profile article-associated data over the Tibetan Plateau, frequently being used to quantify the hydrological cycle and water security, early warning assessments of glacier avalanche disasters, and other geoscience studies on the Tibetan Plateau, are highlighted in this manuscript. The TPDC provides a cloud-based platform with integrated online data acquisition, quality control, analysis, and visualization capability to maximize the efficiency of data sharing. The TPDC shifts from the traditional centralized architecture to a decentralized deployment to effectively connect Third Pole–related data from other domestic and international data sources. As an embryo of data sharing and management over extreme environment in the upcoming “big data” era, the TPDC is dedicated to filling the gaps in data collection, discovery, and consumption in the Third Pole, facilitating scientific activities, particularly those featuring extensive interdisciplinary data use.

KEYWORDS: Atmosphere; Climate change; Data mining; Data processing; Databases; Data science

<https://doi.org/10.1175/BAMS-D-21-0004.1>

Corresponding author: Xin Li, xinli@itpcas.ac.cn

Supplemental material: <https://doi.org/10.1175/BAMS-D-21-0004.2>

In final form 4 June 2021

©2021 American Meteorological Society

For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#).

AFFILIATIONS: Pan, X. Guo, Niu, Yang, and Feng—National Tibetan Plateau Data Center, State Key Laboratory of Tibetan Plateau Earth System and Resources Environment, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China; Li—National Tibetan Plateau Data Center, State Key Laboratory of Tibetan Plateau Earth System and Resources Environment, Institute of Tibetan Plateau Research, and Center for Excellence in Tibetan Plateau Earth Sciences, Chinese Academy of Sciences, Beijing, China; Che and Jin—Center for Excellence in Tibetan Plateau Earth Sciences, Chinese Academy of Sciences, Beijing, and Heihe Remote Sensing Experimental Research Station, Key Laboratory of Remote Sensing of Gansu Province, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou, Gansu, China; Ran, J. Guo, Hu, and Wu—Heihe Remote Sensing Experimental Research Station, Key Laboratory of Remote Sensing of Gansu Province, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou, Gansu, China

Scientific data sharing benefits establishing an honest academic environment by increasing replicability (Carter et al. 2017; Nuijten 2019) and enhances the data value by reusing in further research (Piwowar et al. 2007; Li et al. 2020a, 2021). The concept that “science is driven by data, data is a mirror of science” (Hanson et al. 2011) has penetrated all aspects of scientific research. The essence of scientific data sharing is to provide scientific data to the public in an open and accessible manner to maximize the potential value of scientific data in wide applications, to enhance scientific and technological innovation and to promote scientific development. Fortunately, an increasing number of researchers have realized that “Data sharing can be complex for scientists to navigate, but the rewards are often career-enhancing” and that “Open science can lead to greater collaboration, increased confidence in findings and goodwill between researchers” (Popkin 2019). Well-documented, useful and preserved data can save researchers considerable time. It is estimated that PhD candidates in the sciences spend up to 80% of their time munging data before subjecting them to scientific analysis (Mons 2020).

The Tibetan Plateau (TP), being considered the world’s “Third Pole” due to its height (Qiu 2008) and as the “Water Tower of Asia” due to being the headwaters of Asia’s major rivers (Immerzeel et al. 2010; Immerzeel et al. 2020), is sensitive and vulnerable to global climate change, and along with Antarctic and Arctic, is experiencing a much higher rate of air temperature increase than other regions (Pepin et al. 2015; Liu et al. 2009). The impact of global warming on the Tibetan Plateau is of keen interest in the scientific community (Yao 2019; Yao et al. 2019; Chen et al. 2021). A series of observation and monitoring programs on the Tibetan Plateau have also been widely implemented, and various numerical simulation studies on exploring the mechanism of the interactions between Tibetan Plateau surface process and monsoons have been carried out (Yao et al. 2019).

Scientific data sharing is especially important for the Tibetan Plateau, where has strong multispherical interactions among the atmosphere, cryosphere, hydrosphere, and biosphere (Yao et al. 2015). However, scientific data on the Tibetan Plateau, including in situ observations, remote sensing observations, reanalysis data, and other data sources, are scattered among individuals or small groups and have not yet been integrated for comprehensive analysis of the Tibetan Plateau, thus hindering a better understanding of the unprecedented changes occurring on the Tibetan Plateau and their impacts on the global environment and humans. The collection, construction, publishing, and sharing of scientific data on the Tibetan Plateau are urgently needed to comprehensively understand the multilevel interactions, to provide insights into the ecological and environmental vulnerability associated with climate change and to institute corresponding countermeasures in response to global climate change.

To meet above challenges, the Tibetan Plateau Data Center (TPDC, <http://data.tpdc.ac.cn>) was built up in 2019. The missions of the TPDC are to 1) achieve extensive integration of scientific data resources over the Tibetan Plateau; 2) establish a comprehensive data management and sharing platform, and provide broad data access and services to the scientific research communities and the public; 3) facilitate the exploration of a new paradigm of “Big Data” to promote the Earth system science research and to support the sustainable development of the Tibetan Plateau and surrounding regions.

Overview of the TPDC

The TPDC is China’s most complete scientific data center on the Tibetan Plateau and its surrounding regions. The center was authorized as the National TPDC (one of the first 20 national data centers) in 2019 by the Ministry of Science and Technology of China. The goal of the TPDC is to facilitate the study of environmental changes in the Pan-Tibetan Plateau with improved accuracy and performance, as well as support decision-making for sustainable development of this region (Fig. 1). As of 15 April 2021, the TPDC has integrated 3,512 Tibetan Plateau–related datasets previously scattered on various platforms; has imposed measures to guarantee the intellectual property rights of scientific datasets and to promote the enthusiasm of scientific data sharing, such as data identification, Creative Commons (CC) attribution license, which is a public copyright license, data publishing, and data citation; and has provided preliminary services, including data curation, data quality control, data access, data analysis, and data visualization.

All data are sorted and integrated in strict accordance with the data standards specified by the TPDC and the relevant data acquisition specifications (Fig. 2). The datasets of the TPDC originate from a variety of sources using various methods, such as in situ observation, remote sensing, wireless sensor network, reanalysis and other value-added processes, and voluntary and mandatory sharing from projects and individual scientists. Then, these data are integrated at different levels: database integration, data conflation, and data fusion. Finally, they are preserved in a hybrid cloud environment that adheres to a standard system, thus embodying integrity, stewardship, and security and encouraging data publication. The TPDC hosts more than 3,500 datasets covering diverse disciplines, such as geography, atmospheric science, cryospheric science, hydrology, ecology, geology, geophysics, natural resource science, social economics, and other fields.

As shown in Fig. 3, these datasets are required to be shared under the Findable, Accessible, Interoperable, and Reusable (FAIR; Stall et al. 2019) data sharing principles in the TPDC. Thus, the scientific data and metadata are “findable” by anyone for exploration and use, “accessible” in that they can be examined by anyone, “interoperable” in that they can be analyzed and integrated with comparable data through the use of common vocabulary and formats, and “reusable” by the public as a result of robust metadata, provenance information

Fig. 1. Structure of the Tibetan Plateau Data Center.

Fig. 2. Data integration framework of the Tibetan Plateau Data Center.

and clear usage licenses. Under the guidance of the FAIR data sharing principles, the TPDC data platform provides open access for data users, supplemented by requestable access, with bilingual information in both Chinese and English. The requestable access data sharing is set in the TPDC according to the exclusive rights and interests of data generators. Open access data can be downloaded directly, requestable access data requires an approval process from the data generator, once the data applying has been approved, the downloading of the requestable access data are available and its procedure is same to that of the open access data. Access to requestable data in the TPDC can only be approved by the data provider, and the reasons for this accessible restriction are clarified in the “User Limits” term on the landing page. Meanwhile, the field work data should be submitted to an appropriate scientific data center every year in accordance with the project tasks according to the Notice of the General Office of the State Council (of China) on Regulations of Scientific Data Management [GBF (2018) No. 17]. To guarantee the data provider the priority of using these collecting data, the data protection period is set in the TPDC for them.

Datasets in the TPDC

More than 3,500 Tibetan Plateau-related datasets have been integrated into the TPDC from various data platforms. Among these datasets, five categories of datasets have been

Fig. 3. Data sharing principles of the Tibetan Plateau Data Center.

Fig. 5. Some examples of featured datasets. (a) Multiscale high-elevation river basin observation network (from Che et al. 2019; Li et al. 2013). (b) Water body distribution across the Tibetan Plateau (Zhang et al. 2013). (c) China meteorological forcing dataset (1979–2018) (He et al. 2020). (d) Plant functional type map of the Tibetan Plateau (Ran and Ma 2016). (e) A permafrost type map over the Tibetan Plateau in the past 50 years (from Ran et al. 2021). (f) A late Middle Pleistocene Denisovan mandible from the Tibetan Plateau (Chen et al. 2019). (g) Spatial and temporal patterns of glacier status in the Tibetan Plateau and surroundings (Yao et al. 2012).

Region Observation and Research Network for Land Surface Processes and Environment of China (HORN) (Peng and Zhu 2017). Additionally, comprehensive observation experiments have been conducted, such as the Heihe Watershed Allied Telemetry Experimental Research (HiWATER) project, an airborne-, satellite-, and ground-based integrated remote sensing experiment aiming to improve the observation ability of remote sensing techniques and the understanding and predictability of hydrological and related ecological processes on the catchment scale (Li et al. 2009, 2013).

The featured datasets of high-mountain observations on the Tibetan Plateau include datasets from the HORN, including the meteorological dataset, the hydrological dataset and the ecological dataset (Peng and Zhu 2017); a soil temperature and moisture observational dataset for the Tibetan Plateau (Su et al. 2011; Yang et al. 2013); multiscale observation datasets of the Heihe River basin (Che et al. 2019; Li et al. 2017, 2019; Liu et al. 2018) (Fig. 5a); and multiple datasets from the coordinated Asia–European long-term observation system for the Qinghai–Tibet Plateau, including hydrometeorological processes, the Asian monsoon system, satellite image data of the ground, and numerical simulations (Ma et al. 2009).

LAND SURFACE PARAMETER DATASETS. The parameters of the physical land surface are critical for Earth system models, and many of these parameters are dependent on the vegetation type and soil type index. Regional boundary maps are also needed for regional analysis and model comparison. These in-demand datasets are characterized as general geographic datasets by the TPD.

This category of featured datasets includes the boundary map of the Tibetan Plateau (Zhang et al. 2013), river basin map of the Tibetan Plateau (Zhang et al. 2013), administrative boundary map of the Tibetan Plateau, digital elevation model of the Tibetan Plateau, multisource integrated land cover map of the Tibetan Plateau (Ran et al. 2012a), multistage remote sensing monitoring datasets of land use/cover change over China (Liu et al. 2002), plant functional type map of the Tibetan Plateau (Ran and Ma 2016), soil particle-size distribution dataset for the Tibetan Plateau (Shangguan et al. 2012; Fig. 5c), soil properties for land surface modeling of the Tibetan Plateau (Shangguan et al. 2013), a long-term time series dataset of lake area on the Tibetan Plateau (1970–2013) (Zhang et al. 2013; Zhang et al. 2021), and water body distribution across the Tibetan Plateau (Zhang et al. 2013).

NEAR-SURFACE ATMOSPHERIC FORCING DATASETS. Among the elements in a surface Earth system model, hydrological, soil, ecological, and biogeochemical models all require the input of near-surface atmospheric conditions, including near-surface temperature, precipitation, pressure, water pressure, wind field, and shortwave and longwave radiation as boundary conditions, which are so-called forcing data (Li et al. 2011). Forcing data with high resolution (including both temporal and spatial resolutions) are the basis for running various models but are usually difficult to obtain. This challenge arises because the spatial distribution of station data are sparse, and the observation frequency of conventional stations is generally low. Therefore, interpolation or reanalysis of station data into a grid dataset usually cannot meet the quality and spatial–temporal resolution requirements of forcing data. The resolution of global reanalysis data are usually approximately 1° , although the spatial resolution of some regional reanalysis data can reach 0.25° , which is still relatively coarse for applications at regional/watershed scales. Therefore, it is urgent to develop regional forcing data products with resolutions at approximately 10 km or higher spatial resolution.

The China meteorological forcing dataset (1979–2018) (He et al. 2020), with a temporal resolution of three hours and a spatial resolution of 0.1° , is chosen as a featured dataset for near-surface atmosphere forcing data over the Tibetan Plateau due to its origin from

meteorological observation data, reanalysis data, and satellite remote sensing data, and its quality has been shown to be better than those of the reanalysis data for the Tibetan Plateau.

Additionally, global-scale forcing datasets are available at the TPDC, such as the dataset of high-resolution (3 h, 10 km) global surface solar radiation (1983–2017) (Tang et al. 2019; Fig. 5b), which was produced based on ISCCP-HXG cloud products, ERA5 reanalysis data, and MODIS aerosol and albedo products with an improved physical parameterization scheme. The quality of this dataset has proven superior to those of the ISCCP flux dataset (ISCCP-FD), the global energy and water cycle experiment surface radiation budget (GEWEX-SRB), and the Earth's Radiant Energy System (CERES) (Tang et al. 2019).

CRYOSPHERIC VARIABLES DATASETS. The cryosphere is a component of the Earth system, including solid precipitation, snow cover, glaciers, ice sheets, ice shelves, sea ice, lake and river ice, permafrost, and seasonal frozen ground. The cryosphere plays important roles in climate change, the water cycle, energy balance, ecosystems, and natural hazards at global and regional scales. With global warming, the accelerated retreat of the cryosphere has led to unprecedented impacts on our natural environment and human society. Some cryospheric components (e.g., ice cores) record historical signals of Earth's climate and environment, while others (e.g., sea ice) indicate current global changes. It is very important to understand cryosphere changes at different temporal and spatial scales for the assessment, mitigation, and adaptation of global change in the future.

The featured datasets of cryospheric variables in the TPDC include the first glacier inventory dataset for the Tibetan Plateau (Shi et al. 2009), the second glacier inventory dataset for the Tibetan Plateau (Guo et al. 2015), a permafrost temperature category map for the Tibetan Plateau (2021) (Ran et al. 2012b, 2018, 2021) (Fig. 5d), a new map of permafrost distribution on the Tibetan Plateau (Zou et al. 2017), a long-term land surface freeze–thaw dataset for the Tibetan Plateau (1979–2018) (Jin et al. 2009), a long-term snow depth dataset for the Tibetan Plateau (1979–2018) (Che et al. 2008), MODIS daily cloud-free snow cover products for the Tibetan Plateau (Zheng and Chu 2019), a glacial lake inventory for the Tibetan Plateau in 2015 (Yang et al. 2018), an active layer thickness dataset for the Tibetan Plateau (1981–2018) (Wu and Niu 2013), and an active layer temperature dataset for the Tibetan Plateau (1981–2018) (Xu et al. 2017).

HIGH-PROFILE ARTICLE-ASSOCIATED DATASETS. The purpose of the high-profile article-associated datasets over the Tibetan Plateau is to share the latest research progress on the Tibetan Plateau with researchers in a timely manner to contribute to the promotion of scientific research.

The high-profile article-associated datasets for the Tibetan Plateau and its surrounding areas include work on a late Middle Pleistocene Denisovan mandible from the Tibetan Plateau (Chen et al. 2019; Fig. 5e), differences in glacier status with atmospheric circulations on the Tibetan Plateau and its surroundings (Yao et al. 2012; Fig. 5f), agriculture-facilitated permanent human occupation of the Tibetan Plateau after 3,600 BP (Chen et al. 2015), seismic velocity reduction and accelerated recovery due to earthquakes on the Longmenshan fault (Pei et al. 2019), tree-ring-based winter temperature reconstruction for the southeastern Tibetan Plateau since 1,340 CE (Huang et al. 2019). These datasets are associated with research articles related to science on the Tibetan Plateau and are authorized by the authors, following the CC license and digital object identifier (DOI) numbers assigned by the journals.

The datasets hosted by TPDC have been used in lots of scientific publications, for example pertaining to identification of the change of glacier, permafrost, and snow cover over Tibetan Plateau under climate change, vulnerability assessment of Asian water tower, quantification of ecological change, risk assessment of frozen soil degradation, glacier melting, avalanche- and lake expansion–caused disasters, and calibration and validation of remote sensing products

over the Tibetan Plateau. A list of science highlights and references resulted from the TPDC datasets are compiled in Table ES1 in the online supplement.

New datasets from ongoing projects on the Tibetan Plateau. A series of major programs/projects related to the Earth sciences on the Tibetan Plateau are currently being carried out (Fig. 6), which will produce substantial refreshing and valuable observational datasets (including in situ and remote sensing data) and model outputs. The TPDC is focused on providing an operational supporting platform and database for these ongoing programs and on collecting, integrating, and sharing the data based on observational and research programs, enabling global scientists to explore the study of water resources, climate change adaptation, and disaster risk and resilience of the Tibetan Plateau.

The Second Tibetan Plateau Scientific Expedition and Research program (STEP) is a national key program initiated in August 2017 and led by the Chinese Academy of Sciences (CAS) (Yao 2019). The STEP program covers an area of more than 5 million square kilometers by involving more than 50 disciplines and will produce a series of massive scientific data involving cross-border, multiscale, multidisciplinary, and multitype research. The TPDC is taking the lead in effective management and sharing of these data, which is an important basis for achieving the goal of this scientific expedition, as well as supporting the study of regional and global environmental changes.

The CAS Strategic Priority Research Program entitled “The Pan-Third Pole Environment Study for a Green Silk Road (Pan-TPE)” was launched in 2018. The aim of this program is to explain environmental changes across the Third Pole and their implications, to provide solutions to environmental challenges in high-priority projects, and to explore pathways for sustainable development along the Silk Road. The TPDC has successfully completed the data collection, review, and publishing of the program outputs for two years.

The TPDC will also track the major projects related to the research on the Tibetan Plateau led by the National Natural Science Foundation of China (NSFC) as well as the basic research and development projects led by the Ministry of Science and Technology of China. Moreover, the TPDC has been strengthening cooperation with international programs and projects related to the Third Pole [e.g., Third Pole Environment (TPE), Alliance of International Science Organizations (ANSO) and the Global Energy and Water Cycle Exchanges (GEWEX)] to improve the collection, integration, and publication of data resources from these project outputs and to provide the relevant data support for them.

Fig. 6. Major programs/projects related to the Earth sciences on the Tibetan Plateau.

Data governance on crediting data contributors

Traditionally, the datasets were not cited in formal scientific publications, such as journal papers, which hindered scientists' willingness to share their data in data centers because it added little to advancing their academic careers (Parsons et al. 2010). To incentivize the sharing of scientific datasets, the TPDC has imposed the following measures.

Data identification. The DOI is independent of systems and languages to allow applications crossing disciplines, organizations, and countries. The DOI has been widely used for identifying academic publications, such as journal articles and research reports. In recent years, DOIs have started to be used for identifying datasets. The TPDC adopted the DOI system and created DOIs for every dataset to provide a permanent unique identifier for the dataset following the formula "10.11888/category.tpdc.metadataID," where "10.11888" is fixed as the DOI prefix, presenting the code of the TPDC. There are two variables in the DOI suffix: the item "category" indicates disciplines, and "metadataID" presents the serial number of datasets in the TPDC. For example, a DOI was provided for a long-term (2005–16) dataset of integrated land–atmosphere interaction observations on the Tibetan Plateau as "doi:10.11888/Meteoro.tpdc.270325."

The created DOI is embedded into the dataset metadata and is attached to the dataset during data downloading or accessing. The DOI created at the TPDC is registered with the Institute of Scientific and Technical Information of China, which is a DOI registration agency authorized by the International DOI Foundation, to embed it with the original dataset, which facilitates tracking and citing the dataset in publications or other datasets.

Additionally, Chinese Science and Technology Resource Identification (CSTR) guarantees the authenticity and scientificity of science and technology resources and is an important supplement to define the scientific attributes of resources with DOI identification. To make the identification of data resources concise and coordinate with the DOI, the following format has been adopted as the CSTR in the TPDC: "CSTR: 18046.11. category.tpdc.metadataID," where "CSTR: 18046" is fixed as the CSTR prefix, presenting the registration institution code of the TPDC in CSTR system; "11" means the attribution of scientific data resource, these two numbers are fixed in TPDC; the assignment of "category.tpdc.metadataID" is identical to that of the DOI introduced in last paragraph.

Creative Commons attribution license. The need to clarify the ownership and copyrights of datasets has been increasingly recognized as increasing amounts of data are shared across organizations. Data licensing, as a standard public legal approach, facilitates data sharing by strengthening copyright and removing restrictions that might otherwise limit the dissemination or reuse of data.

The TPDC adopted the Creative Commons 4.0 protocol, which allows the redistribution and reuse of licensed work on the condition that the data generator is appropriately credited. CC offers six options from among which data depositors can choose when they share data: 1) CC BY 4.0, 2) CC BY-SA 4.0, 3) CC BY-ND 4.0, 4) CC BY-NC 4.0, 5) CC BY-NC-SA 4.0, and 6) CC BY-NC-ND 4.0. Here, BY means attribution, AS means share-alike, NC means noncommercial, and ND means no derivative works. The default license in the online data submission system of the TPDC is CC BY 4.0, which means the datasets can be copied and redistributed in any medium or format with being given credit to the original author of the work and that any changes made be disclosed. The data providers can also review and choose other CC licenses to declare the proper copyright for accessing and using their dataset. The chosen CC license is attached to the dataset and will be shown along with the metadata when the dataset is provided or visualized.

Data publishing. Data publishing, emerging as a new form of scholarly publication and gradually being regarded as an important form of academic achievement, makes data usable,

citable, and accessible for long periods. Compared to conventional publications, data publishing makes it easier and more direct to credit data generators for data reuse (Pierce et al. 2019). Many data journals have been established that are dedicated to scientific data, such as *Earth System Science Data*, *Scientific Data*, *Data Science Journal*, *Geoscience Data Journal*, *Ecological Archives*, etc. The TPDC encourages data generators to share their datasets based on data publishing. Examples include Duan et al. (2018), He et al. (2019), Peng et al. (2019), K. Yang et al. (2020), and Ran et al. (2021). The TPDC also serves as a data repository for data publishing. Data should be shared openly before the publication of the data themselves or of corresponding articles, which is increasingly required by scientific data journals and conventional article journals, such as the American Geophysical Union (AGU), which requires that the data needed to understand and build upon the published research be available in public repositories following best practices and that the location where users can access or find the data for the paper be provided explicitly in the Acknowledgments section. Many datasets deposited in the TPDC have been published in scientific journals, such as “Dataset of high-resolution (3 h, 10 km) global surface solar radiation (1983–2017)” (Tang et al. 2019), “China lake dataset (1960s–2015)” (Zhang et al. 2019), “China meteorological forcing dataset (1979–2018)” (He et al. 2020), “The surface heterogeneity patterns and the flux imbalance under free convection based on the WRF LES” (Zhou et al. 2019), and “The 1-km Permafrost Zonation Index Map over the Tibetan Plateau (2019)” (Cao et al. 2019). The TPDC has been officially accepted to become a data repository in the broad scope Earth and environment sciences subsection in the Scientific Data and Springer Nature repository lists (www.nature.com/sdata/policies/repositories#broad-earth-env) and has also become a Trusted Digital Repository of AGU. The TPDC is also applying to become a recommended data repository for other international mainstream journals to incentivize data generators to share their well-documented and useful data by giving them credit and recognition.

Data citation. Data citation is a new concept raised by publishing agencies and data sharing communities to provide traceable information on data production and credit acknowledgment to data generators. The data reference information, particularly the names of the data generators and contributors, should be emphasized in both the metadata and the data documents. A reference to the data citation for each dataset in the TPDC, containing data generators, the dataset’s name, publication date, publisher, and a unique dataset DOI, is generated automatically in appropriate format by the data-sharing center and provided on the dataset-specific page of the TPDC. The data user is required to make the necessary references to the dataset he or she uses and is encouraged to acknowledge the TPDC as well.

Meanwhile, primary publications continue to be considered the main measure of the impact of research rather than the subsequent uses of the data (Pierce et al. 2019). In addition to data citation, in the TPDC, three types of literature related to data are listed on the landing page as the required or optional references to credit data generators: 1) data publications, as a first-hand scientific publication based on the dataset, that are closely related to the dataset’s research background, processing methods, quality evaluation, and application are typically provided by the data generators, generally as the required references; 2) articles that are loosely related to the data or that present analogical data, methods, or scientific topics are generally provided by the data generators as optional references; and 3) articles published by the data users are feedback by data users or collected by the data reuse metrics system as a data supplement.

Data services

Data curation. Scientific data curation is an active management of data interest and usefulness throughout the data life cycle and involves data authentication, archiving, management,

preservation retrieval, and representation. Following the certification criteria proposed by the CoreTrustSeal board—an organization of the World Data System of the International Science Council (WDS) and the Data Seal of Approval (DSA) (www.coretrustseal.org)—four data curation levels are available for the TPDC. The first is data distribution service as a data repository for data journals but provides a simple link with the corresponding paper. Second is data distribution but provides brief checking, addition of basic metadata or documentation. Third is the enhanced curation service through data format standardization and documentation enhancement. Finally, data-level curation provides additional data editing or integration to improve accuracy.

In practice, maintaining and managing the metadata is an important step to realize the curation of the dataset, which means that precise, rich and well-documented metadata are the premise for data curation. The quality of the metadata and the data in the TPDC are ensured by the online bilingual data submission system and the data semi-intellectual review system. The online bilingual data submission system, similar to the paper submission system, is characterized by flexibility and customization, including personalized data description and pop-up menu options. The data semi-intellectual review system is an interaction between data reviewers and data providers on data peer review, including a data expert library and a triggered email-sending function. The detailed workflow of the TPDC review system is shown in Fig. 7.

Data access. Data access is the means by which users can obtain data in an authenticated manner approved by the organization in possession of the data. As a data center dedicated to the Tibetan Plateau, the TPDC can provide a better role in helping data sharing and data use in the scientific research community by exchanging the capability of data and metadata through data services. The capability allows the TPDC to tightly integrate with other data centers to provide more complete and convenient data access to users as well as help promote its data resources to a wider range of user communities across the globe. The TPDC designed and implemented the data services to expose its metadata and datasets via the Internet. Interoperability is the greatest challenge for implementing such services because of the possible variety of implementations at each data center.

The TPDC attempts to reduce the barrier of interoperability by adopting standards and specifics that have been widely accepted by the community. The Open-Source Project for a Network Data Access Protocol (OPeNDAP) and Open Geospatial Consortium (OGC) standards are chosen for data exchange service protocols. Concerning interoperability at the data level,

Fig. 7. The semi-intellectual data review system of the TPDC.

we cannot request users to use a specific format but recommend encoding of data according to NetCDF following the Climate and Forecast convention wherever possible. There are several reasons for this, but the most important is that this is a community standard that covers a wide range of use cases, and it could support both real-time and archived data, support in a standardized manner many different data types [e.g., time series at stations, even moving stations, profiles, trajectories, various types of gridded data and, in the upcoming release, geometries (polygons, lines, etc.)]. It comes with a semantic framework in the form of standard names for variables, unit specification, missing values specification, aggregation levels in time and space, etc.

The TPDC also recommends adopting the structure of information of the published datasets to comply with schema.org and geoschemas.org, which is an emerging standard for describing datasets and data repositories across the geosciences to promote the data to be correctly searched and discovered in search engines, such as Google.

The TPDC website provides a user-friendly interface to regular users to obtain the data. However, it would be remarkably difficult or even impracticable for applications with complex processes to collect extensive datasets through the website. The data services also bypass user interferences to provide direct and continuous data access to such applications. One such example is Earth system modeling, which requires a large amount of data from a variety of sources and scales. These applications will be able to search and retrieve data from the TPDC data service directly in an automated way.

The TPDC also provides data services to support data access to support a variety of use scenarios beyond data downloading; for example, users are able to load and visualize geospatial datasets directly in different tools, such as visualizing the data maps in QGIS or ArcGIS through the OGC Web Map Services (WMS) protocol. Data services are also provided to support lightweight use environments, such as mobile phones, to facilitate a wide range of user communities and even the public with data needs.

Data analysis. Data analysis is the process of gleaning insights from data that are extracted, transformed, and centralized to analyze and discover hidden patterns, relationships, trends, correlations, and anomalies or to validate a theory or hypothesis. With the development of deep learning and machine learning, data can be processed to perform real-time analysis, spot emerging trends and uncover insights. Through incremental integration and independent research and development, the TPDC constructs a data analysis method and tool library of big data quality control, automatic modeling and analysis, data mining and interactive visualization using the Docker container environment and Jupyter + Python programming environment. The Common Software for Nonlinear and Non-Gaussian Land Data Assimilation (ComDA; Liu et al. 2020) is an example of online analysis in the TPDC. ComDA is an online analysis embryo of data assimilation for land surface, hydrological and other dynamic models based on long-term land surface data assimilation research. The online analysis of ComDA also supports users in introducing new dynamic models, observation operators, and data assimilation algorithms at the interface of the TPDC.

Data visualization. Scientific data visualization aims to graphically illustrate scientific data to enable scientists to understand, illustrate, and glean insight from the data (Morse et al. 2019). Geoscientific data visualization is comprehensive and helpful to develop human spatial thinking ability and reveal the relationship between things that may be ignored. In the TPDC, the live visualization of atmospheric/hydrological/ecological observational data are designed for the real-time data analysis and monitoring of the in situ instrumentations. A high-performance information service platform will be built using Web Service and Web Socket for establishing basic service layers, multidimensional maps of water resources, snow

cover, lake ice, observation stations and video surveillance are established by using GIS tools. With rapid technological development, 3D immersive visualization and interaction methods for multiscale geoscientific data based on virtual reality are proposed in the TPDC, and an early warning system of ice and snow disasters, ice lake outbursts and regional ecological monitoring are designed.

Strengthening international cooperation to promote third pole Earth system sciences

The TPDC is strengthening cooperation with international data centers for the sharing and application of third pole data at a global scale. These collaborations will enhance our understanding of climate and environmental changes through data sharing, exchange and interoperability. For example, the TPDC has joined the World Meteorological Organization (WMO) to promote the Integrated Global Cryosphere Information System (IGCryoIS) project and has officially signed a memorandum of collaboration with respect to comprehensive data sharing and research with the National Snow and Ice Data Center (NSIDC). The Third Pole region contains the largest store of ice and glacier mass outside the Arctic and Antarctic. Under global warming, glaciers, permafrost, and ice on the Third Pole are changing rapidly, resulting in a series of climate, ecological, environmental, and resource issues. Through cooperation with the WMO, NSIDC, and other international partners, the TPDC will extend to collect, integrate, and share data resources that are more systematic and relevant not only to the Third Pole but also to the three poles to provide strong data support for global climate and environmental research on extreme environments.

The TPDC is joining international data organizations [e.g., Committee on Data for Science and Technology (CODATA) and World Data System (WDS)] and providing data support for international science programs focused on the Tibetan Plateau and surrounding areas (e.g., TPE and ANSO), among which the TPE is an international program for interdisciplinary study of the relationships among water, ice, air, ecology and humankind in the Third Pole region and beyond (www.tpe.ac.cn/webindex/). It was initiated in 2009 by three world-renowned scientists, Professors Tandong Yao, Lonnie G. Thompson, and Volker Mosbrugger, and is endorsed by UNESCO (United Nations Educational, Scientific and Cultural Organization) as its flagship program and is in close partnership with UNEP (United Nations Environment Programme) and WMO. The TPE International Program Office resides at the Institute of Tibetan Plateau Research of CAS, where the TPDC is subordinate to. The TPDC is responsible for providing data and system support for TPE through developing data and information management mechanisms; storing, integrating, analyzing, excavating, and publishing scientific data; and developing online big data analysis for the Third Pole. High-quality data resources obtained from TPE programs are published on the TPDC platform, which not only enhances the international influence of these data resources but also makes full use of these data to provide support for research on the third pole environment.

Conclusions

The TPDC has recently been built to share scientific data over the Tibetan Plateau and its surrounding regions, and there are approximately 3,500 datasets covering multiple disciplines, such as geography, atmospheric science, cryospheric science, hydrology, ecology, geology, sociology, and economics. All the datasets were sorted and integrated in strict accordance with high-quality data standards, including accuracy, integrity, consistency, validity, uniqueness, and availability. Among these datasets, five categories of featured datasets have been highlighted, including high-mountain observations, land surface parameters, near-surface atmospheric forcing, cryospheric variables, and high-profile article-associated datasets over the Tibetan Plateau. These datasets are applied in Asian water tower investigations, early warning assessments of glacier avalanche disasters, and other geoscience studies on the

Tibetan Plateau. Each dataset in the TPDC is identified by a unique DOI and assigned the CC 4.0 license to guarantee the copyrights of the data generator and its redistributor in the Internet environment with multiple transfers, and the data citation and literature are provided to credit the acknowledgments to the data generators and contributors. The TPDC complies with the FAIR data sharing policy, providing open accessor users, supplemented by requestable access, with information presented in both Chinese and English.

With the rapid developments of the Internet of Things (IoT), artificial intelligence (AI), and machine learning, TPDC is breaking through the traditional concept of data sharing and constructing an online cloud platform integrating online data acquisition, quality control, analysis and visualization. For example, due to wireless transmission technology, the wireless sensor network (WSN), including the automatic collection, transmission and real-time processing of wireless sensor data, has been preliminarily implemented in the Heihe River basin and Qilian Mountain on the northeastern Tibetan Plateau and will be spread throughout the entire Tibetan Plateau and surrounding regions. With the successful application of WSNs, data from WSNs are becoming a live data source housed with TPDCs. In the online big data analysis aspect, based on the latest progress on data assimilation for Earth system science (He et al. 2019; Li et al. 2020b; Liu et al. 2020; Z. L. Yang et al. 2020), the effective integration of information from both model predictions and multisource observations is anticipated. Therefore, high-quality datasets of past, present and future Earth systems over the Tibetan Plateau are expected. The online big data analysis method library and comprehensive multisphere interaction model library for the TPDC are proposed in the next year, and online visualization will come after.

The TPDC has strengthened cooperation with international data organizations (e.g., CODATA, WDS) and provided data support for international science programs of the Tibetan Plateau (e.g., TPE, ANSO), has become a trusted data repository of Springer Nature and AGU and is striving to become a recommended data repository for other international mainstream journals either. The TPDC is shifting from monolithic centralized architectures to decentralized deployments by setting up data interoperability with national and international data centers relevant to the Third Pole Earth science system.

Acknowledgments. This work was supported by Basic Science Center for Tibetan Plateau Earth System (BCTPES, NSFC project 41988101) and the Strategic Priority Research Program of the Chinese Academy of Sciences under Grant XDA20060600. The authors thank the anonymous reviewers and the editor for their very helpful comments.

References

- Cao, B., T. Zhang, Q. Wu, Y. Sheng, L. Zhao, and D. Zou, 2019: Permafrost zonation index map and statistics over the Qinghai-Tibet Plateau based on field evidence. *Permafrost Periglacial Processes*, **30**, 178–194, <https://doi.org/10.1002/ppp.2006>.
- Carter, C. S., C. E. Bearden, E. T. Bullmore, D. H. Geschwind, D. C. Glahn, R. E. Gur, A. Meyer-Lindenberg, and D. R. Weinberger, 2017: Enhancing the informativeness and replicability of imaging genomics studies. *Biol. Psychiatry*, **82**, 157–164, <https://doi.org/10.1016/j.biopsych.2016.08.019>.
- Che, T., X. Li, R. Jin, R. Armstrong, and T. Zhang, 2008: Snow depth derived from passive microwave remote-sensing data in China. *Ann. Glaciol.*, **49**, 145–154, <https://doi.org/10.3189/172756408787814690>.
- , and Coauthors, 2019: Integrated hydrometeorological, snow and frozen-ground observations in the alpine region of the Heihe River Basin, China. *Earth Syst. Sci. Data*, **11**, 1483–1499, <https://doi.org/10.5194/essd-11-1483-2019>.
- Chen, F. H., and Coauthors, 2015: Agriculture facilitated permanent human occupation of the Tibetan Plateau after 3600 B.P. *Science*, **347**, 248–250, <https://doi.org/10.1126/science.1259172>.
- , and Coauthors, 2019: A late middle Pleistocene Denisovan mandible from the Tibetan Plateau. *Nature*, **569**, 409–412, <https://doi.org/10.1038/s41586-019-1139-x>.
- , and Coauthors, 2021: The Tibetan Plateau as the engine for Asian environmental change: the Tibetan Plateau Earth system research into a new era. *Sci. Bull.*, **66**, 1263–1266, <https://doi.org/10.1016/j.scib.2021.04.017>.
- Duan, A., S. Liu, Y. Zhao, K. Gao, and W. Hu, 2018: Atmospheric heat source/sink dataset over the Tibetan Plateau based on satellite and routine meteorological observations. *Big Earth Data*, **2**, 179–189, <https://doi.org/10.1080/20964471.2018.1514143>.
- Guo, W., and Coauthors, 2015: The second Chinese glacier inventory: Data, methods and results. *J. Glaciol.*, **61**, 357–372, <https://doi.org/10.3189/2015JoG14J209>.
- Hanson, B., A. Sugden, and B. Alberts, 2011: Making data maximally available. *Science*, **331**, 649, <https://doi.org/10.1126/science.1203354>.
- He, J., and Coauthors, 2019: Development and evaluation of an ensemble-based data assimilation system for regional reanalysis over the Tibetan Plateau and surrounding regions. *J. Adv. Model. Earth Syst.*, **11**, 2503–2522, <https://doi.org/10.1029/2019MS001665>.
- , K. Yang, W. Tang, H. Lu, J. Qin, Y. Chen, and X. Li, 2020: The first high-resolution meteorological forcing dataset for land process studies over China. *Sci. Data*, **7**, 25, <https://doi.org/10.1038/s41597-020-0369-y>.
- Huang, R., H. Zhu, E. Liang, B. Liu, J. Shi, R. Zhang, Y. Yuan, and J. Griebinger, 2019: A tree ring-based winter temperature reconstruction for the southeastern Tibetan Plateau since 1340 CE. *Climate Dyn.*, **53**, 3221–3233, <https://doi.org/10.1007/s00382-019-04695-3>.
- Immerzeel, W. W., L. P. H. van Beek, and M. F. P. Bierkens, 2010: Climate change will affect the Asian water towers. *Science*, **328**, 1382–1385, <https://doi.org/10.1126/science.1183188>.
- , and Coauthors, 2020: Importance and vulnerability of the world's water towers. *Nature*, **577**, 364–369, <https://doi.org/10.1038/s41586-019-1822-y>.
- Jin, R., X. Li, and T. Che, 2009: A decision tree algorithm for surface soil freeze/thaw classification over China using SSM/I brightness temperature. *Remote Sens. Environ.*, **113**, 2651–2660, <https://doi.org/10.1016/j.rse.2009.08.003>.
- Li, X., and Coauthors, 2009: Watershed allied telemetry experimental research. *J. Geophys. Res.*, **114**, D22103, <https://doi.org/10.1029/2008JD011590>.
- , and Coauthors, 2011: Toward an improved data stewardship and service for environmental and ecological science data in west China. *Int. J. Digital Earth*, **4**, 347–359, <https://doi.org/10.1080/17538947.2011.558123>.
- , and Coauthors, 2013: Heihe Watershed Allied Telemetry Experimental Research (HiWATER): Scientific objectives and experimental design. *Bull. Amer. Meteor. Soc.*, **94**, 1145–1160, <https://doi.org/10.1175/BAMS-D-12-00154.1>.
- , and Coauthors, 2017: A multiscale dataset for understanding complex eco-hydrological processes in a heterogeneous oasis system. *Sci. Data*, **4**, 170083, <https://doi.org/10.1038/sdata.2017.83>.
- , and Coauthors, 2019: Internet of things to network smart devices for ecosystem monitoring. *Sci. Bull.*, **64**, 1234–1245, <https://doi.org/10.1016/j.scib.2019.07.004>.
- , and Coauthors, 2020a: CASEarth poles: Big data for the three poles. *Bull. Amer. Meteor. Soc.*, **101**, E1475–E1491, <https://doi.org/10.1175/BAMS-D-19-0280.1>.
- , F. Liu, and M. Fang, 2020b: Harmonizing models and observations: Data assimilation in Earth system science. *Sci. China Earth Sci.*, **63**, 1059–1068, <https://doi.org/10.1007/s11430-019-9620-x>.
- , and Coauthors, 2021: Boosting geoscience data sharing in China. *Nat. Geosci.*, **14**, 541–542, <https://doi.org/10.1038/s41561-021-00808-y>.
- Liu, F., L. Wang, X. Li, and C. Huang, 2020: ComDA: A common software for non-linear and Non-Gaussian Land Data Assimilation. *Environ. Modell. Software*, **127**, 104638, <https://doi.org/10.1016/j.envsoft.2020.104638>.
- Liu, J. Y., M. L. Liu, X. Z. Deng, D. F. Zhuang, Z. X. Zhang, and D. Luo, 2002: The land use and land cover change database and its relative studies in China. *J. Geogr. Sci.*, **12**, 275–282, <https://doi.org/10.1007/BF02837545>.
- Liu, S., and Coauthors, 2018: The Heihe integrated observatory network: A basin-scale land surface processes observatory in China. *Vadose Zone J.*, **17**, 180072, <https://doi.org/10.2136/vzj2018.04.0072>.
- Liu, X., Z. Cheng, L. Yan, and Z. Y. Yin, 2009: Elevation dependency of recent and future minimum surface air temperature trends in the Tibetan Plateau and its surroundings. *Global Planet. Change*, **68**, 164–174, <https://doi.org/10.1016/j.gloplacha.2009.03.017>.
- Ma, Y., and Coauthors, 2009: Recent advances on the study of atmosphere-land interaction observations on the Tibetan Plateau. *Hydrol. Earth Syst. Sci.*, **13**, 1103–1111, <https://doi.org/10.5194/hess-13-1103-2009>.
- Mons, B., 2020: Invest 5% of research funds in ensuring data are reusable. *Nature*, **578**, 491, <https://doi.org/10.1038/d41586-020-00505-7>.
- Morse, P. E., A. M. Reading, and T. Stål, 2019: Well-posed geoscientific visualization through interactive color mapping. *Front. Earth Sci.*, **7**, 274, <https://doi.org/10.3389/feart.2019.00274>.
- Nuijten, M. B., 2019: Practical tools and strategies for researchers to increase replicability. *Dev. Med. Child Neurol.*, **61**, 535–539, <https://doi.org/10.1111/dmcn.14054>.
- Parsons, M. A., R. Duerr, and J. B. Minster, 2010: Data citation and peer review. *Eos, Trans. Amer. Geophys. Union*, **91**, 297–298, <https://doi.org/10.1029/2010EO340001>.
- Pei, S., F. Niu, Y. Ben-Zion, Q. Sun, Y. Liu, X. Xue, J. Su, and Z. Shao, 2019: Seismic velocity reduction and accelerated recovery due to earthquakes on the Longmenshan fault. *Nat. Geosci.*, **12**, 387–392, <https://doi.org/10.1038/s41561-019-0347-1>.
- Peng, P., and L. P. Zhu, 2017: Observations of land surface processes of the Tibetan Plateau based on the field stations network. *Sci. Technol. Rev.*, **35**, 97–102.
- Peng, S., Y. Ding, W. Liu, and Z. Li, 2019: 1 km monthly temperature and precipitation dataset for China from 1901 to 2017. *Earth Syst. Sci. Data*, **11**, 1931–1946, <https://doi.org/10.5194/essd-11-1931-2019>.
- Pepin, N., and Coauthors, 2015: Elevation-dependent warming in mountain regions of the world. *Nat. Climate Change*, **5**, 424–430, <https://doi.org/10.1038/nclimate2563>.
- Pierce, H. H., A. Dev, E. Statham, and B. E. Bierer, 2019: Credit data generators for data reuse. *Nature*, **570**, 30–32, <https://doi.org/10.1038/d41586-019-01715-4>.
- Piwowar, H. A., R. S. Day, and D. B. Fridsma, 2007: Sharing detailed research data is associated with increased citation rate. *PLOS ONE*, **2**, e308, <https://doi.org/10.1371/journal.pone.0000308>.
- Popkin, G., 2019: Data sharing and how it can benefit your scientific career. *Nature*, **569**, 445–447, <https://doi.org/10.1038/d41586-019-01506-x>.
- Qiu, J., 2008: China: The Third Pole. *Nature*, **454**, 393–396, <https://doi.org/10.1038/454393a>.
- Ran, Y. H., and H. Q. Ma, 2016: 1KM plant functional types map over China in 2000. *Remote Sens. Tech. Appl.*, **331**, 827–832.

- , X. Li, L. Lu, and Z. Y. Li, 2012a: Large-scale land cover mapping with the integration of multi-source information based on the Dempster–Shafer theory. *Int. J. Geogr. Inf. Sci.*, **26**, 169–191, <https://doi.org/10.1080/13658816.2011.577745>.
- , G. D. Cheng, T. J. Zhang, Q. B. Wu, H. J. Jin, and R. Jin, 2012b: Distribution of permafrost in China: An overview of existing permafrost maps. *Permafrost and Periglacial Processes*, **23**, 322–333, <https://doi.org/10.1002/ppp.1756>.
- , ———, and ———, 2018: Climate warming over the past half century has led to thermal degradation of permafrost on the Qinghai–Tibet Plateau. *Cryosphere*, **12**, 595–608, <https://doi.org/10.5194/tc-12-595-2018>.
- , and Coauthors, 2021: Mapping the permafrost stability on the Tibetan Plateau for 2005–2015. *Sci. China Earth Sci.*, **64**, 62–79, <https://doi.org/10.1007/s11430-020-9685-3>.
- Shangguan, W., Y. Dai, B. Liu, A. Ye, and H. Yuan, 2012: A soil particle-size distribution dataset for regional land and climate modelling in China. *Geoderma*, **171**–**172**, 85–91, <https://doi.org/10.1016/j.geoderma.2011.01.013>.
- , and Coauthors, 2013: A China data set of soil properties for land surface modeling. *J. Adv. Model. Earth Syst.*, **5**, 212–224, <https://doi.org/10.1002/jame.20026>.
- Shi, Y., C. Liu, and E. Kang, 2009: The glacier inventory of China. *Ann. Glaciol.*, **50**, 1–4, <https://doi.org/10.3189/172756410790595831>.
- Stall, S., and Coauthors, 2019: Make scientific data FAIR. *Nature*, **570**, 27–29, <https://doi.org/10.1038/d41586-019-01720-7>.
- Su, Z., J. Wen, L. Dente, R. van der Velde, L. Wang, Y. Ma, K. Yang, and Z. Hu, 2011: The Tibetan Plateau observatory of plateau scale soil moisture and soil temperature (Tibet-Obs) for quantifying uncertainties in coarse resolution satellite and model products. *Hydrol. Earth Syst. Sci.*, **15**, 2303–2316, <https://doi.org/10.5194/hess-15-2303-2011>.
- Tang, W., K. Yang, J. Qin, X. Li, and X. Niu, 2019: A 16-year dataset (2000–2015) of high-resolution (3 h, 10 km) global surface solar radiation. *Earth Syst. Sci. Data*, **11**, 1905–1915, <https://doi.org/10.5194/essd-11-1905-2019>.
- Wu, Q. B., and F. J. Niu, 2013: Permafrost changes and engineering stability in Qinghai–Xizang Plateau. *Chin. Sci. Bull.*, **58**, 1079–1094, <https://doi.org/10.1007/s11434-012-5587-z>.
- Xu, X., Q. Wu, and Z. Zhang, 2017: Responses of active layer thickness on the Qinghai–Tibet Plateau to climate change. *Bingchuan Dongtu*, **39**, 1–8.
- Yang, C. D., X. Wang, J. F. Wei, and S. Y. Liu, 2018: A dataset of glacial lake inventory of West China in 2015. *Chin. Sci. Data*, **3**, 9.
- Yang, K., and Coauthors, 2013: A multiscale soil moisture and freeze–thaw monitoring network on the Third Pole. *Bull. Amer. Meteor. Soc.*, **94**, 1907–1916, <https://doi.org/10.1175/BAMS-D-12-00203.1>.
- , Y. Chen, J. He, L. Zhao, H. Lu, J. Qin, D. Zheng, and X. Li, 2020: Development of a daily soil moisture product for the period of 2002–2011 in Chinese mainland. *Sci. China Earth Sci.*, **63**, 1113–1125, <https://doi.org/10.1007/s11430-019-9588-5>.
- Yang, Z. L., L. Zhao, Y. He, and B. Wang, 2020: Perspectives for Tibetan Plateau data assimilation. *Natl. Sci. Rev.*, **7**, 495–499, <https://doi.org/10.1093/nsr/nwaa014>.
- Yao, T., 2019: Tackling on environmental changes in Tibetan Plateau with focus on water, ecosystem and adaptation. *Sci. Bull.*, **64**, 417, <https://doi.org/10.1016/j.scib.2019.03.033>.
- , and Coauthors, 2012: Different glacier status with atmospheric circulations in Tibetan Plateau and surroundings. *Nat. Climate Change*, **2**, 663–667, <https://doi.org/10.1038/nclimate1580>.
- , and Coauthors, 2015: Multispherical interactions and their effects on the Tibetan Plateau’s Earth system: A review of the recent researches. *Natl. Sci. Rev.*, **2**, 468–488, <https://doi.org/10.1093/nsr/nwv070>.
- , and Coauthors, 2019: Recent Third Pole’s rapid warming accompanies cryospheric melt and water cycle intensification and interactions between monsoon and environment: Multidisciplinary approach with observations, modeling, and analysis. *Bull. Amer. Meteor. Soc.*, **100**, 423–444, <https://doi.org/10.1175/BAMS-D-17-0057.1>.
- Zhang, G., T. Yao, H. Xie, S. Kang, and Y. Lei, 2013: Increased mass over the Tibetan Plateau: From lakes or glaciers? *Geophys. Res. Lett.*, **40**, 2125–2130, <https://doi.org/10.1002/grl.50462>.
- , and Coauthors, 2019: Regional differences of lake evolution across China during 1960s–2015 and its natural and anthropogenic causes. *Remote Sens. Environ.*, **221**, 386–404, <https://doi.org/10.1016/j.rse.2018.11.038>.
- , and Coauthors, 2021: 100 years of lake evolution over the Qinghai–Tibet Plateau. *Earth Syst. Sci. Data*, **13**, 3951–3966, <https://doi.org/10.5194/essd-13-3951-2021>.
- Zheng, Z. J., and D. Chu, 2019: Snow cover dataset based on optical instrument remote sensing with 1 km spatial resolution on the Qinghai–Tibet Plateau (1989–2018). National Tibetan Plateau Data Center, accessed 19 April 2021, <https://doi.org/10.11888/Snow.tpd.270465>.
- Zhou, Y., D. Li, and X. Li, 2019: The effects of surface heterogeneity scale on the flux imbalance under free convection. *J. Geophys. Res. Atmos.*, **124**, 8424–8448, <https://doi.org/10.1029/2018JD029550>.
- Zou, D., and Coauthors, 2017: A new map of permafrost distribution on the Tibetan Plateau. *Cryosphere*, **11**, 2527–2542, <https://doi.org/10.5194/tc-11-2527-2017>.