# A framework of freshwater and saline lake typology classification through leveraging hydroclimate, spectral, and literature evidence

Meng Ding[1], Jida Wang[1*], Chunqiao Song[2], Yongwei Sheng[3], J. M. Shawn Hutchinson[1], Abigail L. Langston[1], and Landon Marston[4]

[1]Department of Geography and Geospatial Sciences, Kansas State University, Manhattan, Kansas, United States

[2]Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, China

[3]Department of Geography, University of California, Los Angeles (UCLA), California, United States

[4]Department of Civil and Environmental Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, United States

*Corresponding to jidawang@ksu.edu

## Abstract

Separating saline and freshwater lakes is of immense importance for monitoring lentic ecosystems and assessing surface freshwater availability. However, conventional lake typology often relies on in situ salinity measurements, thus limiting the feasibility of a global application. To enable an effective workaround, we here propose and test a framework that classifies saline and freshwater lakes through systematic, integrated leveraging of hydrological, climatological, spectral, and literature evidence. In principle, saline lakes were sequentially identified if a lake is (1) located at the drainage terminal where salinity tends to build up as a result of lacking surface outflow (hydrological evidence), (2) distributed in an arid/semiarid climate with lacustrine evaporites visible from a high-resolution multispectral image (climatological and spectral evidence), or (3) documented to be (sub)saline or have measured salinity exceeding the maximum freshwater threshold (literature evidence). We applied the HydroSHEDS hydrography dataset, the Köppen-Geiger climate classification, and a large volume of spectral satellite images and literature to implement this framework. This framework was tested on the Tibetan Plateau using the Landsat-based, 30-m-resolution UCLA circa-2015 global lake mask, and the result was benchmarked against the Chinese national lake survey. Our collected literature shows a 95% agreement between the lake types classified using hydrological, climatological, and spectral evidence alone and those documented in the literature. After further corrections based on the literature evidence, our final classification reached an overall accuracy of nearly 70% in terms of lake count and 94% in lake area. Given such promising accuracies, we further applied our framework to Australia, showcasing its effectiveness for lake typology classification without additional in situ surveys across the world's most saline continent. While our method, in theory, results in a conservative recognition of saline lakes and should not be considered a complete substitute for in situ surveys, the proposed

framework demonstrates potential for performing global-scale freshwater and saline lake typology classification with reasonable accuracies, using only remote sensing images, open-source hydrography and climate data, and existing literature.

## 1. Introduction

Fresh surface water is one of the most important natural resources for ecosystems and human societies (Oki and Kanae, 2006; Schewe et al., 2014; Wada et al., 2011). There are more than one billion cubic kilometers of water on the Earth (Schlesinger and Bernhardt, 2013; Chahine, 1992; Jackson et al., 2001; Schmitt, 1995), but the storage of fresh water only accounts for about 3% of the water mass in the hydrosphere. However, only 0.5% of the freshwater supplies can be easily accessed by people, as over 80% of all available freshwater is stored in ice sheets and glaciers (Kashid and Pardeshi, 2014). Climate change and a growing and more affluent population are placing significant pressure on freshwater resources worldwide, which is likely to have increasingly negative consequences. The impacts of these challenges are numerous, including earlier and shorter melt seasons for snowpack and glaciers, reduced streamflow during summer and fall, and greater human consumptive water demands (Wada et al.; 2016; Barnett et al., 2005).

There are several major components of freshwater resources for human water consumption, including river flow, groundwater, and desalinated water. But the most readily accessible liquid water stock resides in lakes, ponds, and reservoirs (Oki and Kanae, 2006). River flow and lake storage are two main components of the surface water system. River flow includes discharge in the river/stream channels accumulated from surface runoff and base flow (Stricker, 1983), and lake storage contains water in two major types of lakes: freshwater lakes and saline lakes. While water in saline lakes may not be directly used for human consumption, saline lake systems have unique environmental and ecological functions, and they influence the biological dynamics of ecosystems that depend on the lakes (Jódar et al., 2020). Since freshwater and saline lakes offer different ecologic and societal benefits, it is important that we know where these types of lakes are located and their storage volume. The spatial distribution of freshwater and saline lakes, and their interactive dynamics in area and storage, are of essential significance for monitoring surface freshwater water availability and regional and global climate change (Alley et al., 1993). However, because of the extensive distribution and dynamic nature in the global lake systems (Adrian et al., 2009; Pekel et al., 2016; Smith et al., 2005; Song et al., 2016; Wang et al., 2012; Wang et al., 2014; Wang et al., 2017; Yamazaki et al., 2015; Yao et al., 2019), our knowledge of their accurate extents, quantity, and quality (including saline and freshwater typology) remains incomplete (Messager et al., 2016; Cael et al., 2017).

Existing studies on global-scale lake abundance focus on lake quantity, areas, and volumes (Messager et al., 2016; Cael et al., 2017). The Mullard Space Science Laboratory (MSSL) Global Lakes Database (MGLD) has the locations and types (closed lakes/open lakes/reservoirs) of 1415 inland natural and artificial water bodies (Birkett and Mason, 1995). The Shuttle Radar Topography Mission (SRTM) Water Body Data (SWBD) (SWBD, 2005; Slater et al., 2006) constitutes water bodies with a total area of 2.76 million km$^2$. However, this dataset only collected water bodies from 56°S to 60°N, limited by the spatial coverage of the SRTM. To fill in this incomplete coverage, Carroll et al. (2009) created a comprehensive global water mask at 250-m resolution by combining SWBD and MODIS imagery. About 254,000 water bodies larger than 0.1 km$^2$ were mapped in the Global Lakes and Wetlands Dataset (GLWD) (Lehner and Döll, 2004). The Global Inland surface Waterbody dataset (GIW) (Feng et al., 2016) mapped global water bodies from 30-m Landsat images, covering a total area of ~3.65 million km$^2$. However, some of the included water bodies are ephemeral or intermittent, while some others were mapped incompletely because of cloud contamination. In addition, there is no explicit distinction between rivers and lakes/reservoirs. The more recently published HydroLAKES dataset mapped 1.42 million global lakes and reservoirs larger than 0.1 km$^2$. By considering the variability of not only lake areas but also the terrain in the vicinity of the lakes, Messager et al. (2016) integrated a large number of documentation and geostatistical models to estimate the depth and volume of each inventoried lake/reservoir. One of the major limitations of HydroLAKES is the negligence of a great number of lakes smaller than 0.1 $km^2$.

Very recently, Birkett et al. (2022) produced a new database consisting of three individual lake and reservoir datasets with a total of 6,495 waterbody records: 1. Global Reservoir and Lakes Database (GREALD) with global water bodies larger than 100 $km^2$, 2. GREALD with water bodies (mid-low latitude open lakes and all-latitude reservoirs) with surface area 10-99 $km^2$, and 3. Global Ephemeral Lakes Database (GELD) with ephemeral lakes larger than 100 $km^2$. This database contributes to surface water level monitoring using satellite radar altimetry. All three datasets classified water bodies as reservoirs, open lakes, closed lakes, ephemeral lakes, and lagoons. However, these datasets do not classify the limited number of identified lakes as freshwater or saline lakes. Another recently published global lake dataset, GLAKES (Pi et al., 2022), contains 3.4 million global lakes and reservoirs and their area changes over the past four decades. It was produced using the Global Surface Water Occurrence (GSWO) dataset (Pekel et al., 2016) and deep learning methods. However, GLAKES also has no classification between freshwater and saline lakes, and the minimum lake size is 0.03 $km^2$. Among existing datasets, Sheng et al. (2016) produced a high-resolution global lake mask from cloud-free Landsat-8 images, which contains quality-controlled water extents of ~9 million perennial lakes and reservoirs larger than 0.004 $km^2$. The water extents correspond to representative inundation areas within the "stable lake season" during the initial 2.5 years (May 2013 to November 2015) of Landsat-8's operation (Lyons and Sheng, 2017; Sheng et al., 2016). This dataset, hereafter referred to as the "UCLA circa-2015 lake dataset" or "circa-2015 lake dataset", substantially improved our understanding of the spatial distribution of global lakes and reservoirs.

While these global waterbody datasets have advanced our understanding of the global location, storage, and number of different types of surface water bodies, we still miss critical information regarding the salinity of these water bodies. Distinguishing between freshwater and saline lakes provides a more accurate estimation of available freshwater supplies for human development (Abbott et al., 2019) and locates distinct ecosystems unique to lakes of different salinities. Therefore, clearly understanding the locations of freshwater and saline lakes has profound implications for water management and freshwater allocation. The most reliable way to differentiate freshwater and saline lakes is to use the salinity of the lake water, although the threshold of salinity that separates them varies depending on the purpose. In Hammer (1983) and Wang (1993), a salinity of 0.05% is considered as the salinity threshold between freshwater and saline lakes. However, due to the sheer number of global lakes and their diverse environmental settings, it is practically impossible to determine the exact salinity of each lake.

Two widely used methods for distinguishing freshwater and saline lakes are field survey and remote sensing (Harris, 1976; Ueda et al., 2000; Leidonald et al., 2019; Wang and Xu, 2008). Both methods have advantages and limitations. Lake salinity can be obtained by directly measuring the concentration of dissolved minerals in the water sample, or indirectly, it can be inferred by measuring the density or conductivity of the lake water (Williams and Sherwood, 1994). These field surveys provide high accuracy in salinity. However, field measurement prohibits large-scale applications and is relatively time-consuming and costly. Compared to field surveys, remote sensing provides a low-cost alternative to track the changes in salinity over time and space. For example, soil salinization around lakes could be distinguished by using the unique signatures of salts on multispectral or hyperspectral images (Safaee and Wang, 2020; Azarafza and Mokhtari, 2013; Wang and Xu, 2008; Abou et al., 2018). However, freshwater and saline lake classification that relies on remote sensing methods is also challenging because satellite images are often affected by changes in external environmental factors (Kim and Han, 2021; Huang et al., 2016), such as the percentage of cloud coverage, atmospheric effects, and the sensor response. In addition, saline lakes in different regions have various mineral constituents whose evaporites can exhibit different spectral signatures and textures, and there are similarities between the spectral signatures of salts and snow, ice, and impervious surface. Therefore, the accuracy of identifying saline lakes is reduced compared to field survey and the salinity of the lakes is hard to be remotely quantified. As a result, it is technically challenging to formulate a remote sensing algorithm that works equivalently well for different individual lakes.

To strike a balance between applicability and accuracy, we here aim at developing and validating a framework that leverages hydrological analysis, climatological evidence, spectral remote sensing, and existing references to classify lake freshwater and saline types. The organization of this paper is as follows. Section 2 introduces the testing and validation areas and the used datasets. Section 3 uses the Tibetan Plateau as a testing site to describe the logic of our lake type classification framework and showcase how it was implemented. The classification results and the accuracy validation for the Tibetan Plateau are shown in Section 4, and in this section, the classification approach is also applied in Australia as a case study. Section 5 summarizes our findings and discusses potential applications.

## 2. Study areas and datasets used

2.1 Study areas

We selected two study regions that are rich in saline lakes to demonstrate the efficacy of our freshwater and saline lake classification framework. This general framework can be applied globally. First, we applied our framework (see section 3 for more details) to the Tibetan Plateau. The Tibetan Plateau has ample verified data classifying lakes as either saline or freshwater that allowed us to extensively validate our approach. Once our model had been trained on the Tibetan Plateau, we applied our framework to the entire continent of Australia to showcase the generality and wide-applicability of our framework. Our rationale is that if our method can be effectively applied in these saline lake-rich hotspots, it may in principle be applied at a global scale.

### 2.1.1 Tibetan Plateau

The Tibetan Plateau has an altitude of more than 4,000 meters and a land area of 2.5 million $km^2$ (Zhang et al., 2011). One of the most well-known features is the large number and area of lakes (Fig. 1). Zhang et al. (2014) estimated more than 32,000 lakes with a total area of over 43,000 $km^2$ on the entire Tibetan Plateau based on GeoCover 2000 Landsat imagery. Historically, the Tibetan Plateau experienced substantive tectonic activities (Li and Fang, 1999) and glacial erosion, resulting in a large number of depressions for the formation of terminal or endorheic lakes. As rivers carry dissolved minerals into the terminal lakes but the lakes do not have a surface outlet, salt is accumulated in the lakes through time. Moreover, there is plenty of solar radiation but limited precipitation in the Tibetan Plateau, so the lake water becomes more saline as it evaporates.

The water cycle of the Tibetan Plateau, especially the dynamics in surface water storage, can directly reflect the impacts of large-scale regional climate change. The Tibetan Plateau has a complicated distribution of freshwater and saline lakes. Some of the largest saline lakes in China, such as Qinghai Lake, Nam Co, and Siling Co, are all located on the Tibetan Plateau. The lakes in the Tibetan Plateau are mainly terminal lakes, which are mostly saline, and are widely distributed in the Changtang Plateau which has a large number of landlocked and arid watersheds. In the context of climate warming and moistening in the Tibetan Plateau (Yang et al., 2014), terminal lakes are more subject to expansion, indicating increasing relocation of freshwater resources (mainly from increasing precipitation) to saline water resources (terminal lakes). In addition, authoritative information of lake salinity and types on the Tibetan Plateau are accessible through the Chinese national lake survey (CNLS) (see Section 3.7). For these reasons, the development and validation of our lake classification method are exemplified in the Tibetan Plateau.

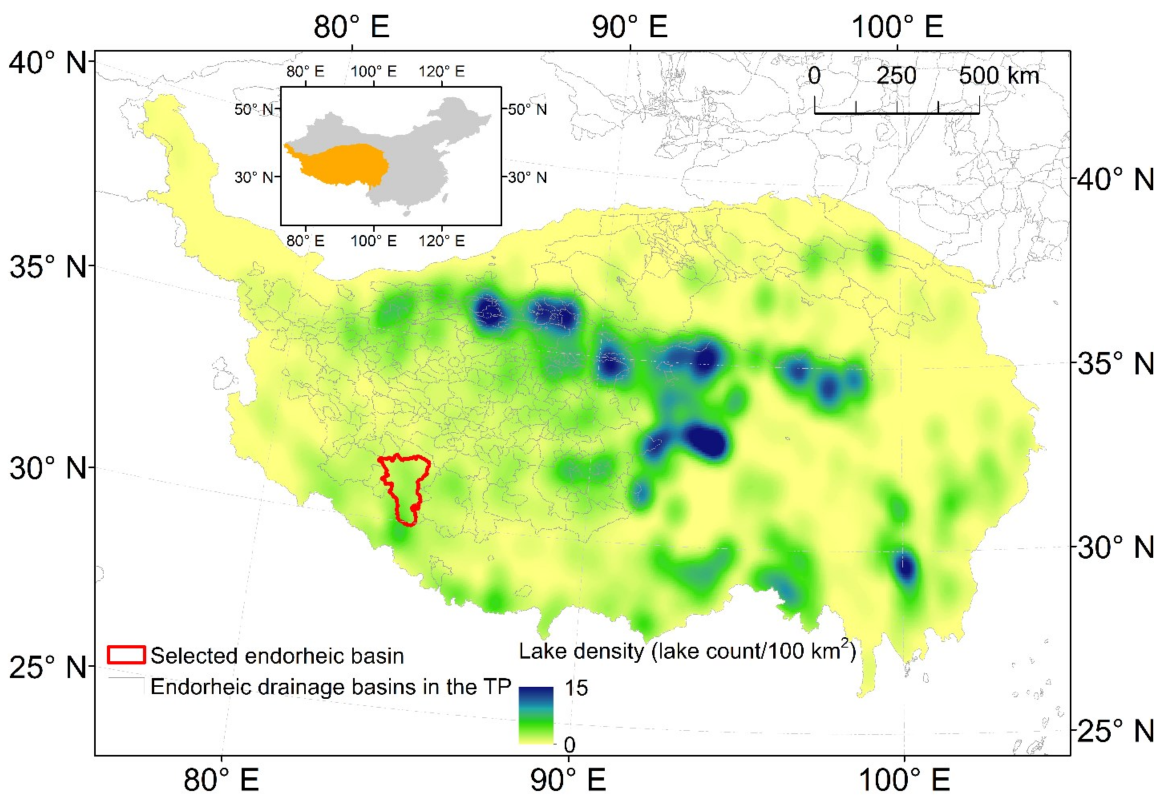


**Fig. 1.** The studied Tibetan Plateau. Drainage basins in the Tibetan Plateau (including both endorheic and exorheic basins) were derived from the HydroSHEDS 15-arc-second Drainage Basins (Lehner et al., 2008). The background is lake count per 100 km$^2$ (generated from the UCLA circa-2015 lake dataset) for the purpose of visualization. The inlet map shows the position of the Tibetan Plateau in China.

### 2.1.2 Australia

Most of Australia is located in low and middle latitudes (10°S−40°S) and contains a variety of climate zones (Sheng et al., 2016). The northeastern part of Australia has a tropical rainforest climate, and the eastern part has a subtropical humid climate. There is relatively more precipitation near coastal areas, and the southern part has a Mediterranean climate or a temperate maritime climate (Oliver, 2005; Sheng et al., 2016).

The majority of the landmass in Australia is, however, dominated by arid and semiarid climates (Byrne et al., 2008). The Tropic of Capricorn passes through the middle of Australia, thus the main climate categories in Australia are the tropical and subtropical climate types. In terms of the wind zone and the pressure zone in Australia, the area near the regression line (inside the continent) is mainly controlled by the subtropical high-pressure zone (Wandres et al., 2016). The suppressed precipitation forms a tropical desert climate in central Australia. Western Australia is affected by the cooling and dehumidifying effects of the Australian cold current, so the west coast also has a tropical desert climate (Beard, 1969). The extensive deserts are surrounded by savannas, with high temperatures throughout the year and low precipitation. As a result of climatic aridity, a large proportion of the landmass in Australia is landlocked or endorheic (Wang et al., 2018), making it another hotspot with highly concentrated saline lakes. For these reasons, we selected Australia as a showcase of our lake classification framework.

## 2.2 Datasets and pre-processing

### 2.2.1 Lake mask

Choosing an accurate and high-resolution lake mask plays an important role in classifying freshwater and saline lakes. Fortunately, due to the advances in spectral and microwave remote sensing on mapping surface hydrology, several datasets on global water bodies have been produced successively (refer to Section 1). Among them, the UCLA circa-2015 lake dataset was produced from about 60 thousand Landsat-8 scenes acquired during the initial 2.5 years of the sensor operation (2013-2015). Lake extents were mapped using a self-adaptive algorithm that took into account the unique spectral conditions of individual lakes to optimize land-water separation. The mapped lakes were mosaicked to represent an intermediate inundation extent to minimize seasonal and inter-annual variations. The automated mapping was then visually inspected and improved through a semi-automated quality assurance and quality control (QA/QC) procedure (Wang et al., 2014) to further improve the mapping quality.

The circa-2015 lake dataset contains more than 9 million perennial lakes and reservoirs with a total area of 2.97 million $km^2$. In comparison with HydroLAKES, the circa-2015 lake dataset has a higher spatial resolution, with a minimum lake size of 0.004 $km^2$ (~4 Landsat pixels) as opposed to 0.1 $km^2$ in HydroLAKES. The circa-2015 lake dataset also removed many commission errors (such as mountain and glacier shadows) as occasionally identified in HydroLAKES, owing to the rigorous QA/QC process. The improved spatial detail for smaller lakes in the circa-2015 dataset is illustrated in Fig. 2 for a focal region in the Canadian Shield (around 66°13’N and 116°40’W). As a partial result of the paleo-glaciation processes, the Canadian Shield possesses the highest lake density in the world today. Due to this improvement, we selected the circa-2015 lake dataset as the primary lake mask to test and perform our freshwater and saline lake classification. Our logic is that if we can validly apply our method to this higher-resolution lake mask, our method may be applicable to coarser-resolution lake masks as well.

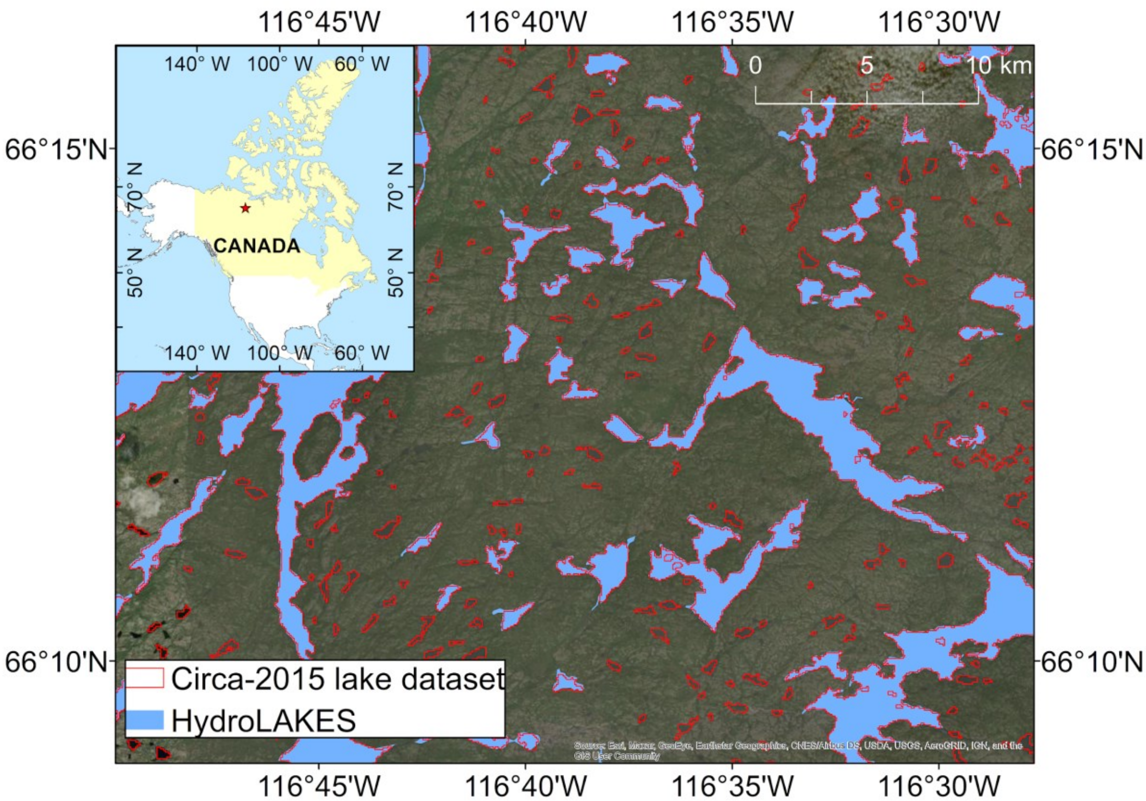


**Fig. 2.** Comparison between HydroLAKES and the UCLA circa-2015 lake dataset. The shown region is in the northern Canadian Shield (see inlet). Red (hollow) and blue (filled) polygons depict waterbody extents in the circa-2015 lake dataset and HydroLAKES, respectively.

### 2.2.2 Hydrography data

The HydroSHEDS dataset, which was developed from the SRTM Digital Elevation Model (DEM), depicts detailed hydrographic information, including river networks, drainage directions, and watershed boundaries (Lehner et al., 2008; https://www.hydrosheds.org/downloads). Multiple datasets were also assembled and compared during the development process. While hydrological adjustment was made through the use of automatic procedures, it also went through rigorous manual editing and quality control to remove spurious sinks and to delineate endorheic basins. The version we used here has a resolution of 15 arc seconds (about 450 m at the equator), which is the finest among the provided layers of river accumulation, river network, and drainage basins. We pre-processed the data to extract endorheic basins from the exorheic basins (Fig. 3), by first identifying inland sinks (terminal points) from the drainage direction layer, and then intersecting the sink points with the drainage basins. The derived terminal points, endorheic/exorheic basins, and the river networks within them were jointly used as the hydrological evidence for our lake type classification (see Section 3.2).

### 2.2.3 Climate data

The 0.5-degree world map of Köppen-Geiger climate classification by Rubel and Kottek (2010; accessed at http://koeppen-geiger.vu-wien.ac.at/shifts.htm) was used to help locate arid and semiarid regions, where saline lakes have a higher chance to occur due to limited precipitation but high potential evapotranspiration. This updated climate classification dataset includes a series of digital world maps from 1901 to 2100 to capture the shift of global climate zones due to climate change. The serial maps before 2000 were based on temperature and precipitation observations acquired from the Climatic Research Unit (CRU) of the University of East Anglia and the Global Precipitation Climatology Centre (GPCC) at the German Weather Service. We used the map for 1976-2000 to represent the most recent observation of global climate zones. Guided by our expert-based experimentation, we merged the classes of arid climate (B), dry summer (S), and polar tundra (T) to define the arid regions where saline lakes may likely occur. We included "polar tundra" mainly because a large area of the Tibetan Plateau, which is arid and home to many saline lakes (Section 2.1.1), was labeled as this class in the dataset. Merging these climate zones leads to about 40% of the global landmass (excluding Antarctica) and exceeds the maximum extent of the conventional arid climates. However, it ensures our purpose of identifying saline lakes as thoroughly as possible. As shown in Fig. 3, our processed global "arid zones" largely overlap with endorheic basins, but it also supplements the latter by encompassing more drylands peripheral to the endorheic regions.

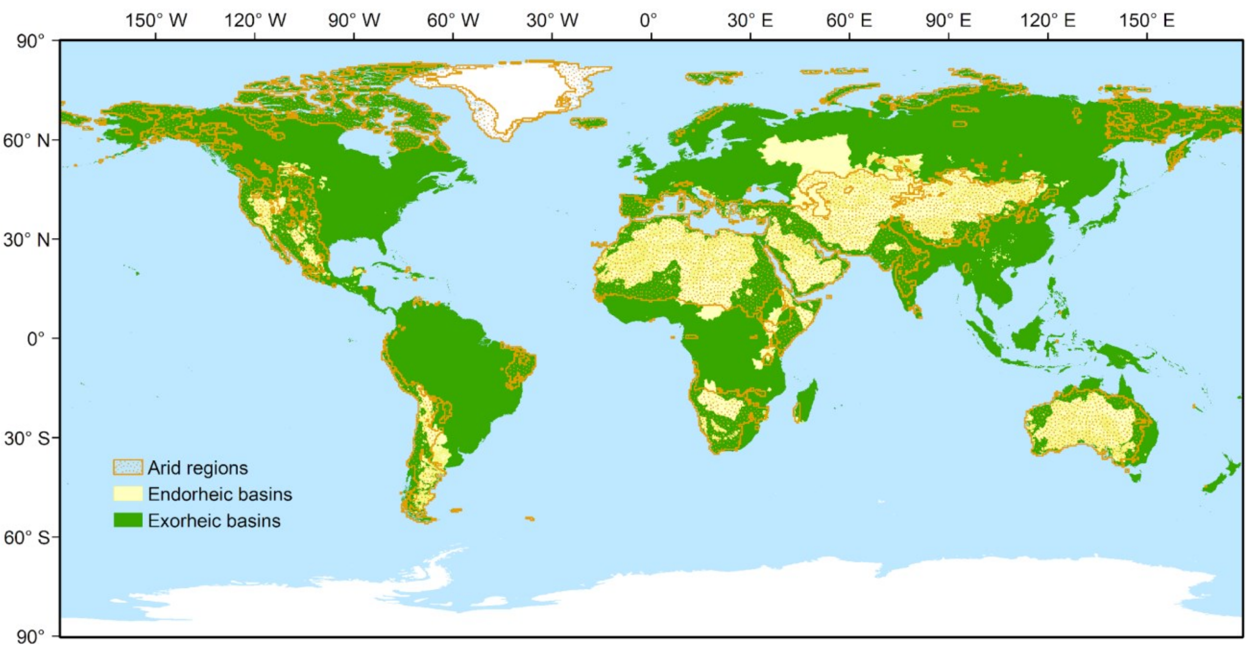


**Fig. 3.** Global distribution of arid regions, endorheic basins, and exorheic basins. Endorheic basins and exorheic basins were derived from the HydroSHEDS 15-second hydrography dataset (Lehner et al., 2008). The global spatial distribution of arid regions was based on the Köppen−Geiger Climate Classification (at 0.5-degree resolution) (Rubel and Kottek, 2010).

### 2.2.4 Satellite imagery

We referenced a large number of multi-source remote sensing images for visually inspecting and detecting the spectral evidence (e.g., deposited evaporites and salt crusts) of saline lacustrine environments. The primary reference is thousands of 30-m resolution Landsat-8 images, which served as the mapping source for the circa-2015 lake dataset. In addition, we also referred to higher-resolution images from Google Earth, Bing Maps, and the Esri World Imagery Map, to assist in our visual interpretation of the spectral evidence.

### 2.2.5 Other literature

A large volume of existing literature and documents record the surveyed salinity and water types of many lakes in the world. Although these records are mostly limited to large lakes, they are useful auxiliary information to complement the other evidence for a systematic lake type classification. Accordingly, we conducted an extensive literature review for the two study areas. Examples include some of the well-referred books such as "Saline Lake Ecosystems of the World" (Hammer, 1986), "Distribution of the World's Large Lakes" (Herdendorf, 1990), and the Chinese Lake Catalogue (Wang and Dou, 1998), among others. The literature evidence was also used to assess the effectiveness of our lake type classification using hydroclimate and spectral evidence alone (see Section 3.7).

### 2.2.6 Validation data

Authoritative survey data that document detailed lake type and/or salinity for a particular region are rare or poorly shared. For our study, we were able to access such a dataset based on CNLS in the year 2006. CNLS documents survey-based lake types for 2,735 lakes and reservoirs larger than 1 $km^2$ across China (Ma et al., 2010; Zhu et al., 2020), and applied a salinity threshold of 0.1% to distinguish freshwater lakes and saline lakes. Then CNLS used more specific salinity thresholds to classify saline lakes to three sub–categories: oligosaline lakes (salinity 0.1-2.4 %), polysaline lakes (2.4-3.5 %), and hypersaline lakes (>3.5 %). This lake survey was used to validate the accuracy of our typology results for the Tibetan Plateau.

## 3. Methods

3.1 Overview of the lake type classification framework

Our framework separated freshwater and saline lakes by following the decision tree in Fig. 4. In this decision tree, each of the 16 branches represents a unique combination of the binary conditions (existence or absence) for the four general evidence types, i.e., hydrological, climatological, spectral, and literature, which were used jointly to determine whether a lake is likely saline. For clarity, we also translated this decision tree into a tabular format in Table 1. The classification of lakes was based on four types of evidence, each with specific conditions that were categorized as “primary” or “secondary”. A “primary” condition or evidence means that the presence of this condition alone (such as being at the drainage terminal) could already lead to a high likelihood that a lake is saline unless there is literature evidence opposing this judgment. In comparison, a “secondary” condition (such as being under arid climate or on the drainage pathway, rather than at the terminal, of an endorheic basin) defines a favorable situation where primary conditions (such as exhibiting lacustrine evaporites) may possibly occur, but this condition alone is insufficient to conclude the lake type.

We aggregated the results of the 16 node scenarios to four general lake classes, depending on the classified freshwater or saline type and the drainage position of the lake, i.e., whether being on the drainage pathway or at the terminal. These four lake classes are terminal (closed) saline, open (flow-through) saline, terminal (closed) freshwater, and open (flow-through) freshwater.

Among them, open saline lakes may include saline lagoons if the lakes are distributed along the coast. We opted to inform the drainage positions (flow-through or terminal) along with the binary water types (saline and freshwater), because the former provides additional useful information indicative of lake retention time, water quality, and ecosystem functions. More details for each type of evidence are explained in the following Method sections.

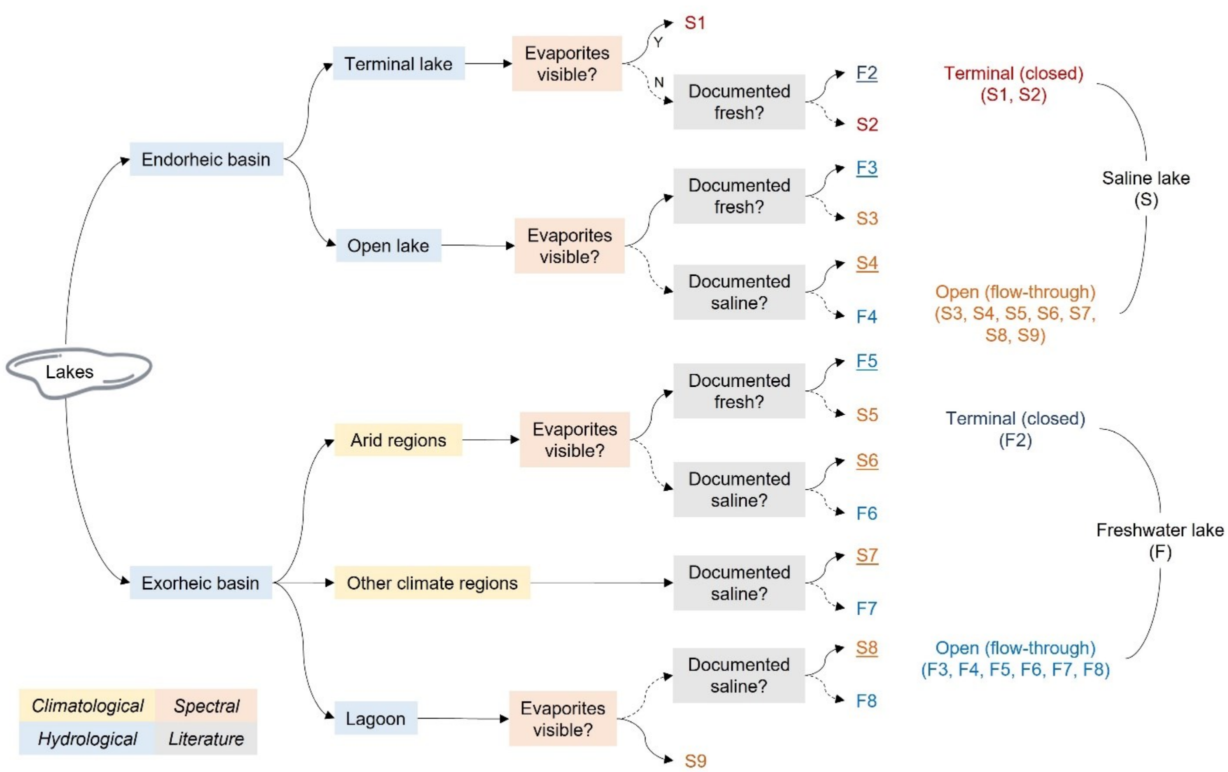


**Fig. 4.** Schematic flowchart illustrating the framework of classifying freshwater and saline lakes. Solid arrows after a question indicate “Yes” whereas dashed arrows indicate “No”. Classified lake types are annotated as “S1” to “S9” and “F2” to “F8”, with “S” indicating saline lakes and “F” freshwater. Underscored types (e.g., “F2” and “S4”) indicate results corrected by opposing literature evidence. Detailed lake types were further grouped into four general classes (as color-coded): terminal saline, open saline, terminal freshwater, and open freshwater. Here a “terminal”

or “open” lake means if the lake is located at the drainage terminal or pathway. Refer to Table 1 for more detail.

**Table 1.** Detailed rules for classifying freshwater and saline lake types. Typology notations are consistent with those in Fig. 4. Blank cells indicate this condition or evidence was less relevant and was not verified. “Yes” in “opposing literature” indicates the classified lake type would have been the opposite if using other evidence (hydroclimate and spectral) alone. Conditions that are not labeled “primary” are “secondary” conditions.

| ***Typology*** | | ***Hydrology*** | | | ***Climate*** | ***Spectral*** | ***Literature*** |
|---|---|---|---|---|---|---|---|
| *General* classes | Specific types | Endorheic basin | Terminal lake (*primary*) | Coastal lagoon | Aridity | Evaporite (*primary*) | Opposing literature (*primary*) |
| Terminal saline | S1 | Yes | Yes | | | Yes | No |
| | S2 | Yes | Yes | | | No | No |
| Open saline | S3 | Yes | No | | | Yes | No |
| | S4 | Yes | No | | | No | Yes |
| | S5 | No | No | No | Yes | Yes | No |
| | S6 | No | No | No | Yes | No | Yes |
| | S7 | No | No | No | No | | Yes |
| | S8 | No | No | Yes | | No | Yes |
| | S9 | No | No | Yes | | Yes | No |
| Terminal freshwater | F2 | Yes | Yes | | | No | Yes |
| Open freshwater | F3 | Yes | No | | | Yes | Yes |
| | F4 | Yes | No | | | No | No |
| | F5 | No | No | No | Yes | Yes | Yes |
| | F6 | No | No | No | Yes | No | No |
| | F7 | No | No | No | No | | No |
| | F8 | No | No | Yes | | No | No |

3.2 Paradox of lake water and solute balances

We theorize that a saline lake may form when its water balance tends to be maintained at the cost of a positive solute imbalance. We refer to this paradoxical condition as the "water-solute imbalance", which can be explained by the equations below. The balance of lake water mass or volume ($V$) through time ($t$) is governed by:

$$\frac{dV}{dt} = P + Q_{in} - Q_{out} - Q_g - E \quad (1)$$

where the main influxes include precipitation recharge on the lake surface ($P$) and surface inflow to the lake ($Q_{in}$), and the main outflux terms are surface outflow from the lake ($Q_{out}$), infiltration to the groundwater ($Q_g$), and lake surface evaporation ($E$). Among these fluxes, surface inflow, surface outflow, and groundwater outflow carry minerals dissolved along the drainage paths, but evaporation carries only water molecules. Assuming minerals in precipitation are limited, the balance of soluble salts ($S$) in a lake is governed by:

$$\frac{dS}{dt} = S_{Q_{in}} - S_{Q_{out}} - S_{Q_g} + S_e \quad (2)$$

where the only exogenous influx comes from the solute supply carried by lake surface inflow ($S_{Q_{in}}$), and the outfluxes are solute losses through surface outflow ($S_{Q_{out}}$) and seepage loss ($S_{Q_g}$). In addition, endogenous sources such as geochemical and microbial processes within the lake itself may also affect the lake's mineral balance, and we use $S_e$ to notate the net effect of these endogenous processes.

Assuming $S_e$ has a limited contribution, $dS$ accumulates if $S_{Q_{in}}$ exceeds the sum of $S_{Q_{out}}$ and $S_{Q_g}$. This happens when $Q_{out}$ and $Q_g$, which carry $S_{Q_{out}}$ and $S_{Q_g}$, are limited or negligible, leaving $E$ the predominant outflux to balance out $Q_{in}$. Despite a possible water mass balance under this condition, the lake experiences a positive imbalance (gain) of soluble salts and may eventually turn saline. For the same reason, saline lakes are often intermittent or ephemeral in arid climates, when $Q_{in}$ is so limited that $E$ can easily become disproportional, leading to a negative water mass imbalance. Although endogenous biochemical processes ($S_e$) also modulate lake salinity, our collected hydroclimate evidence below was guided based on this water-solute imbalance, which essentially defines an overarching condition that lake outflux is dominated by evaporation.

## 3.3 Hydrological evidence

Most saline lakes occur in endorheic basins, especially at the drainage terminals (Hammer, 1986; Wang et al., 2018; Wang, 2020; Wurtsbaugh et al., 2017; Yang et al., 2017), because endorheic hydrology well represents the above-described water-solute imbalance. Hydrologically, the global continental surface is divided into exorheic and endorheic basins, depending on whether surface flow from the basin can drain to the ocean (Fig. 3). Rivers in an endorheic basin do not flow into the ocean, but instead drain to an inland terminus or sink. Lakes located at the sinks are terminal or endorheic lakes, whereas lakes on the drainage paths or at the headwater are often referred to as flow-through or open lakes. Note open lakes can be located in either endorheic or exorheic basins. Since a terminal lake receives $Q_{in}$ but produces no $Q_{out}$, water storage in most terminal lakes can only be balanced out via $E$, leaving the dissolved minerals (solutes) in the lakes. Through time, the water salinity in a terminal lake accumulates and the lake becomes saline. Except significant infiltration and subsurface outflow ($Q_g$, such as Lake Chad), terminal lakes are usually saline due to this landlocked hydrology. For this reason, we considered being at the sink of an endorheic basin a primary condition, meaning that a terminal lake was usually assumed to be saline unless there is opposing literature suggesting otherwise. In contrast with terminal lakes, an open lake can be balanced out via both $E$ and $Q_{out}$, limiting the likelihood of solute accumulation in the lake. Therefore, we did not assume flow-through (open) lakes in an endorheic basin to be saline without other evidence.

Lakes in an exorheic basin are, by definition, in the pathway of the drainage networks and thus open to surface outflow. Since dissolved minerals are carried away by the surface outflow, the salinity of open lakes often remains low. Lakes in exorheic basins are often freshwater, except when located in very arid areas with high lake evaporation (see Section 3.3). Coastal lagoons are usually shallow water bodies that are separated from the ocean by barrier islands and are subject to tidal effects (Kjerfve et al., 1996). A lagoon can be both freshwater and saline or brackish depending on its hydrodynamics with the ocean (Joyce et al., 2005; Emmanuel and Chukwu, 2010), so literature evidence is required to identify their water types.

Here a selected endorheic basin in the southern Tibetan Plateau (Fig. 5) is used to illustrate our classification procedure throughout the methods. We first referred to HydroSHEDS to sort out the drainage position of each lake in this basin. A terminal lake named Chabyer Co (Fig. 5a and 5b) was identified at the drainage sink (green dots in Fig. 5a) of this basin. As being terminal is a primary condition, Chabyer Co is considered a saline lake for now. Other lakes in this endorheic basin, such as Qoi'nyi Co (Fig. 5c) draining from the northern shore and Taruo Co (Fig. 5d) draining eastward, are open or flow-through lakes, thus assumed to be freshwater lakes for now.

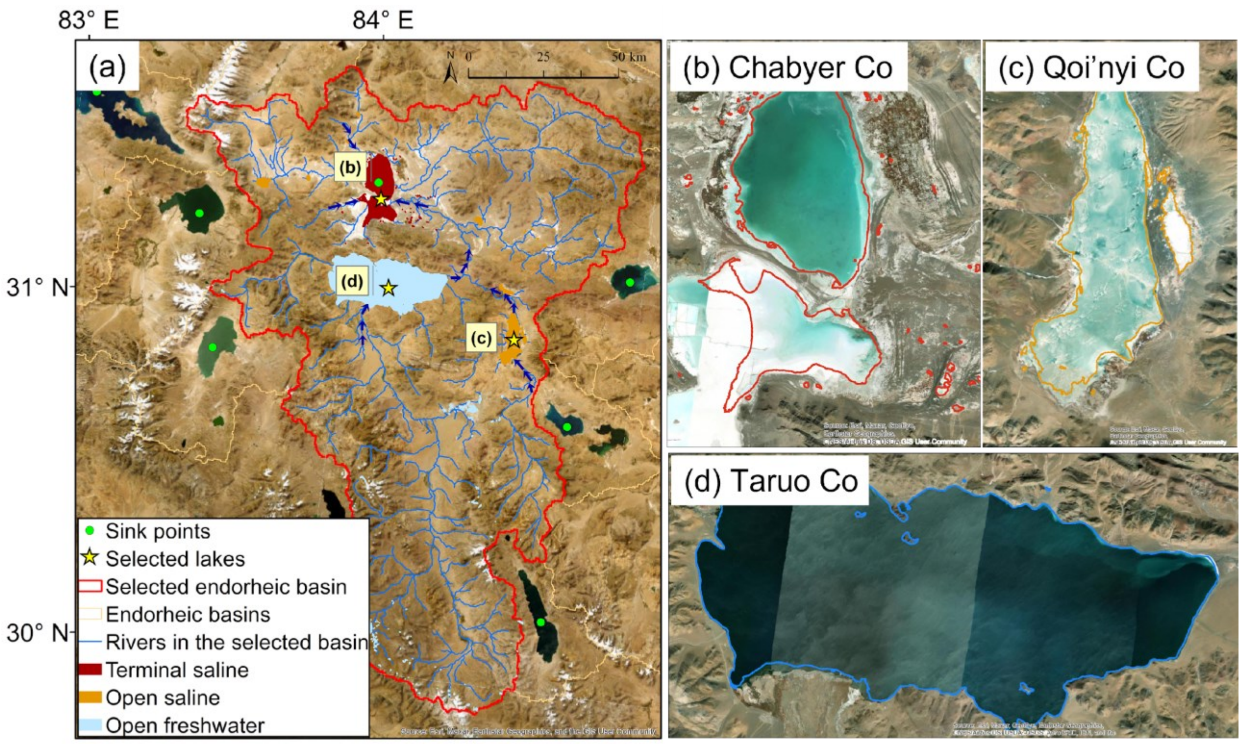


**Fig. 5.** Lake typology classification in a selected endorheic basin on the Tibetan Plateau. Background is the Esri satellite imagery.

3.4 Climatological evidence

Being terminal is not a necessary condition for the water-solute imbalance. As described in Section 3.2, salinity in a lake can also accumulate if its water loss via $Q_{out}$ and $Q_g$, both carrying soluble salts, is substantially less than the loss via $E$ carrying only water molecules. Such a disproportional surface evaporation tends to occur in arid climates, where long-term atmospheric demands for water vapor (i.e., potential evapotranspiration) exceed the atmospheric supplies $P$. Despite a limited recharge from local $P$, lake water supplies in arid climate can be compensated by more abundant $Q_{in}$ (e.g., glacial meltwater from headwaters). This may lead to lake $Q_{out}$ exceeding $E$, which is not conducive to solute accumulation. For this reason, we considered high climate aridity a secondary condition to locate the environments favored by saline lakes, which was then used together with additional evidence (Sections 3.5 and 3.6) to determine the lake type. The conventional arid climate zones (i.e., regions with potential evapotranspiration exceeding precipitation) are overall spatially consistent with endorheic basins (Wang et al., 2018; Wang 2020), but as described in Section 2.2.3, our defined global arid zones (Fig. 3) are more inclusive, in order to identify saline lakes as thoroughly as possible. These arid zones cover most part of the endorheic Tibetan Plateau, including the example in Fig. 5.

### 3.5 Spectral evidence

Under intense evaporation or saturated salinity, soluble minerals can precipitate from the water and crystallize on the lake beds or littoral zones, forming evaporites that are indicative of hypersaline environments. Evaporite deposits are usually highly reflective in the visible and infrared spectra, providing another type of primary evidence for saline lakes. In a recent study, Safaee and Wang (2020) applied machine learning on Landsat-8 images to extract salt pans and salt playas across the western United States. While this study demonstrated the potential of multispectral imagery for automated evaporite extraction, the method also implied the challenge of separating evaporites from other high-reflectance substances such as clays, sands, and manmade impervious surfaces. A more accurate classification may be achieved via spectroscopic analysis based on detailed spectral signatures, but this requires hyperspectral images which are not always freely accessible. Along with high spectral reflectance, crystallized evaporites tend to exhibit patchy and crusty textures, which are sometimes discernible in high-resolution multispectral images. Therefore, as a workaround, we combined this unique texture information and the high spectral reflectance of evaporites to define the “spectral evidence” of saline lakes.

Specifically, we inspected the existence of the above-defined spectral evidence in the immediate vicinity of each lake. Our focus was lakes in endorheic basins and/or arid climate zones (Section 2; Fig. 3) where evaporites are mostly likely to deposit. We leveraged a variety of image sources, including Landsat-8 source imagery of the circa-2015 lake dataset (see Section 1), Google Earth imagery, Bing Maps imagery, and the Esri World Imagery basemap. Given that the main purpose of this study was to demonstrate the classification framework, our inspection was largely based on expert-based visual interpretation, which was akin to how a human technician conducts an in situ survey based on the surface spectral reflectance and texture surrounding a lake. To assist in visual interpretation, we also followed Safaee and Wang (2020) to flag high-reflectance regions, defined empirically by Landsat-8 pixels where the mean of visible (red, green, and blue) reflectivity exceeds 0.2. The high-reflectance flags, in combination with our meticulous visual inspection, aimed at optimizing the accuracy of evaporite identification. However, our framework allows for the flexibility of potentially replacing this visual method by more automated algorithms which, for example, synthesize both spectral and textural information to effectively identify lacustrine evaporites.

In the example of the Tibetan Plateau (Fig. 5), highly reflective brines and evaporites are distributed across Lake Chabyer (Fig. 5b), Lake Qoi'nyi (Fig. 5c), and their littoral zones, showing clear evidence that these two lakes are saline or hypersaline. For Lake Chabyer, the spectral evidence is present on top of the hydrological evidence (i.e., being at the drainage terminal), further confirming this lake is indeed saline. For Lake Qoi'nyi, the spectral evidence complements the climatological evidence (i.e., being in arid climate), suggesting that this lake, although open to surface outflow, has a predominant outflux via evaporation and is therefore saline. In comparison, Lake Taruo (Fig. 5d), although exposed to arid climate, displays no hydrological or spectral evidence and was thus assumed to be freshwater. It is worth mentioning that the use of spectral evidence led to an overall conservative detection of saline lakes because not all saline lakes exhibit evident lacustrine evaporites, at least seen from multispectral imagery.

## 3.6 Literature evidence

While the lake type might be reasonably inferred from the combination of hydrological, climatological, and spectral evidence, the most authoritative evidence stemmed from literature documentation. Although rarely available and skewed towards larger lakes, literature evidence served an important dual purpose. As shown in Fig. 4 and Table 1, the lake types documented in the literature were used to correct misclassifications from the other set of evidence, therefore determining the final lake type. The literature also provided an independent source to validate the classification accuracy using the other evidence alone. This is especially valuable for benchmarking the fidelity of our lake type classification framework in literature-lacking regions.

We considered two types of literature evidence. One is documented nominal lake types (e.g., freshwater, subsaline, and saline) and the other is numeric salinity according to which lake type can be categorized. We followed Hammer (1983) and used a salinity of ~0.05% as the divide between freshwater and saline lakes. When possible, we further labeled the lakes whose salinities fall between 0.05% and 0.3% as subsaline (Hammer 1983; Wang 1993). This intermediate category allows for additional flexibility of lake type classification at the user's discretion. Unless otherwise specified, saline lakes in our study include those that are subsaline.

Through a deep search from the literature pool (2,261 lakes with literature evidence, with a total area of 29,703.7 km$^2$), we collected as much evidence as possible to supplement, confirm, and correct the classification results using hydroclimate and spectral evidence alone. This led to the final leaf nodes of the decision tree (Fig. 4), defining specific scenarios depending on whether the literature opposes or agrees with the classifications based on other sets of evidence. For instance, a terminal lake, which possesses the primary hydrological evidence, was assumed to be saline even though it lacks evident evaporites. However, if the literature opposes this assumption, the lake was reclassified to be freshwater with a scenario label of “F2”; otherwise, the lake type remained to be saline with a scenario label of “S2”. An example of such exceptions (“F2” lakes) is Lake Chad, where water remains fresh although the lake has a high evaporation level and is located at the surficial drainage terminal. This is because the local river inflow ($Q_{in}$) has low salinity and the lake has a high groundwater outflow ($Q_g$) allowing soluble salts to escape to the aquifers (Isiorho and Matisoff, 1990; Wang, 2020).

In the example of the Tibetan Plateau (Fig. 5), we referred to Wang and Dou (1998) to confirm that Chabyer Co is hypersaline with a high salinity of 43.4%. This also proves that our classification for Chabyer Co (label “S1”) given both hydroclimate and spectral evidence was reliable (Fig. 4; Table 1). Qoi’nyi Co was classified to be saline due to the presence of visible evaporites. However, we were unable to find any literature reference to verify this classification. As a result, Qoi’nyi Co was labeled “S3”, meaning a probable saline lake that has evident evaporite deposits but is open (flow-through). Guo et al (2016) documented a salinity of ~0.05% for Taruo Co, confirming our assumption of a freshwater type was correct. This rendered a label of “F4” for Taruo Co, meaning a literature-validated freshwater lake that is open (flow-through) and has no evident evaporites.

### 3.7 Validation

We validated our lake type classification framework in two ways. First, we applied available literature evidence to evaluate the reliability of our classification scheme (as described in Section 3.6) based on hydrological, climatological, and spectral evidence alone (thereafter "non-literature evidence"). This validation was performed for both Tibetan Plateau and Australia. Relative classification error was estimated as the count of lakes where the type determined by non-literature evidence opposes that by the literature, divided by the total count of lakes where literature evidence is available. Specifically, lake scenarios with contradictory literature and non-literature evidence include F2, F3, S4, F5, S6, S7, and S8 (Fig. 4 and Table 1), and these scenarios represent the known classification errors from non-literature evidence alone. It is worth noting that literature evidence was absorbed as part of the typology framework, so any error identified in this aspect has been corrected in the final classification result (see Section 4).

The second validation approach used the external data CNLS (Ma et al., 2010; Zhu et al., 2020) to independently validate the accuracy of the final lake types classified from both non-literature and literature evidence for the Tibetan Plateau. As described in Section 3.6, our classification was based on the lake typology system in Hammer (1983), which used a minimum salinity of 0.05% to define saline lakes, whereas CNLS used a minimal salinity of 0.1% (see Section 2.2.6). A more detailed discrepancy between the two typology systems was visualized in Fig. 6, where the range from 0.05% to 0.1% represents the uncertainty of lake types due to disagreeing salinity thresholds. To remove the impact of this discrepancy, we excluded any lake whose salinity falls in the uncertainty range from validation.

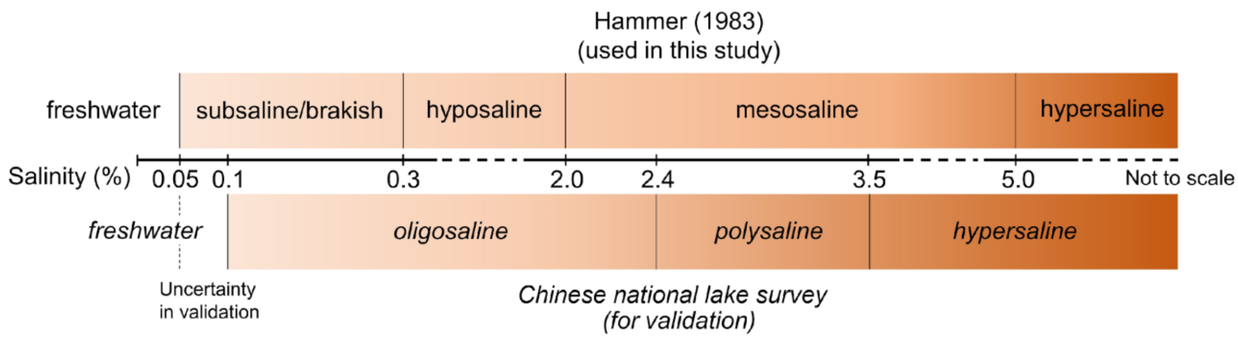


**Fig. 6.** Lake typology systems adopted in Hammer (1983) and the Chinese national lake survey (CNLS).

## 4. Results and discussion

4.1 Classified freshwater and saline lake typology on the Tibetan Plateau

Our classification framework comprehensively integrates hydroclimate analysis, spectral remote sensing, and literature evidence and was developed on the Tibetan Plateau using the circa-2015 lake dataset. Our classification approach identified the diversity of lakes on the Tibetan Plateau, including the number, type, size, and location of freshwater and saline lakes. Figure 7 shows the spatial distribution of classified freshwater and saline lakes on the Tibetan Plateau. Freshwater lakes are distributed almost ubiquitously across the entire Tibetan Plateau, whereas saline lakes are concentrated within the endorheic basins, mostly in the northern and northwestern Tibetan Plateau. Figure 8 shows the quantities of the four general lake types, i.e., terminal saline, open saline, terminal freshwater, and open freshwater. Among the 34,723 lakes on the Tibetan Plateau, 4,825 (13.9%) were classified as saline, and the remaining 29,898 (86.1%) were classified to be freshwater. While much fewer in number, the total surface area of saline lakes was much larger. The saline lakes have an aggregate area of 35,490.3 km$^2$ (82.4% of total lake area), whereas the freshwater lakes have a total area of 7,563.4 km$^2$ (17.6%). Freshwater lakes outnumber saline six to one, but saline lakes account for more than two-thirds of the total lake area, indicating that some of the largest lakes on the Tibetan Plateau are terminal saline lakes.

Focusing more narrowly on open lakes, we find that open freshwater lakes (86.0%) make up the largest share of all lakes in the Tibetan Plateau, whereas open saline lakes only account for a small fraction (4.8%) of total lakes. However, these two open lake categories are comparable in area. This is partially because open lakes were classified as saline only when they exhibit clear spectral evidence of surficial and/or littoral evaporites, and such spectral evidence was more likely to be identified for larger lakes given the limitation of image resolution. This contrast also implies that our saline lake detection is conservative. We found little evidence from spectral remote sensing or the literature to counter our assumption that terminal lakes are saline.

About 82% of the saline lakes are located on the Changtang Plateau (Fig. 7), the largest cluster of endorheic basins on the Tibetan Plateau (Yao et al., 2018). The Changtang Plateau has an average altitude of ~4800 meters and includes a wide variety of landscapes including plains, hügellands, and mountains, and it is surrounded by the Tanggula Mountains, the Nyainqêntanglha Mountains, and the Kailash Range (Xu et al., 2020). It is cold and dry on the Changtang Plateau because the mountains on the southern side block the warm and humid airflow from the Indian Ocean (Cuo et al., 2014; Tian et al., 2022). Therefore, the climate and hydrography of the Changtang Plateau both result in widespread saline lakes. About 8% of the saline lakes, including Qinghai Lake (37° N, 100.13° E), Hala Lake (38.3° N, 97.6° E), and Chaka Salt Lake (36.7° N, 99.11° E), are in the endorheic basins in the northeastern of the Tibetan Plateau. This region encompasses the Hexi basin, the headwater of the Yellow basin, and the Qaidam basin which is one of the three largest endorheic basins in China, and the climate is dry throughout the year and deserts are widespread (Zhang and Liu, 2020; Wang et al., 2022).

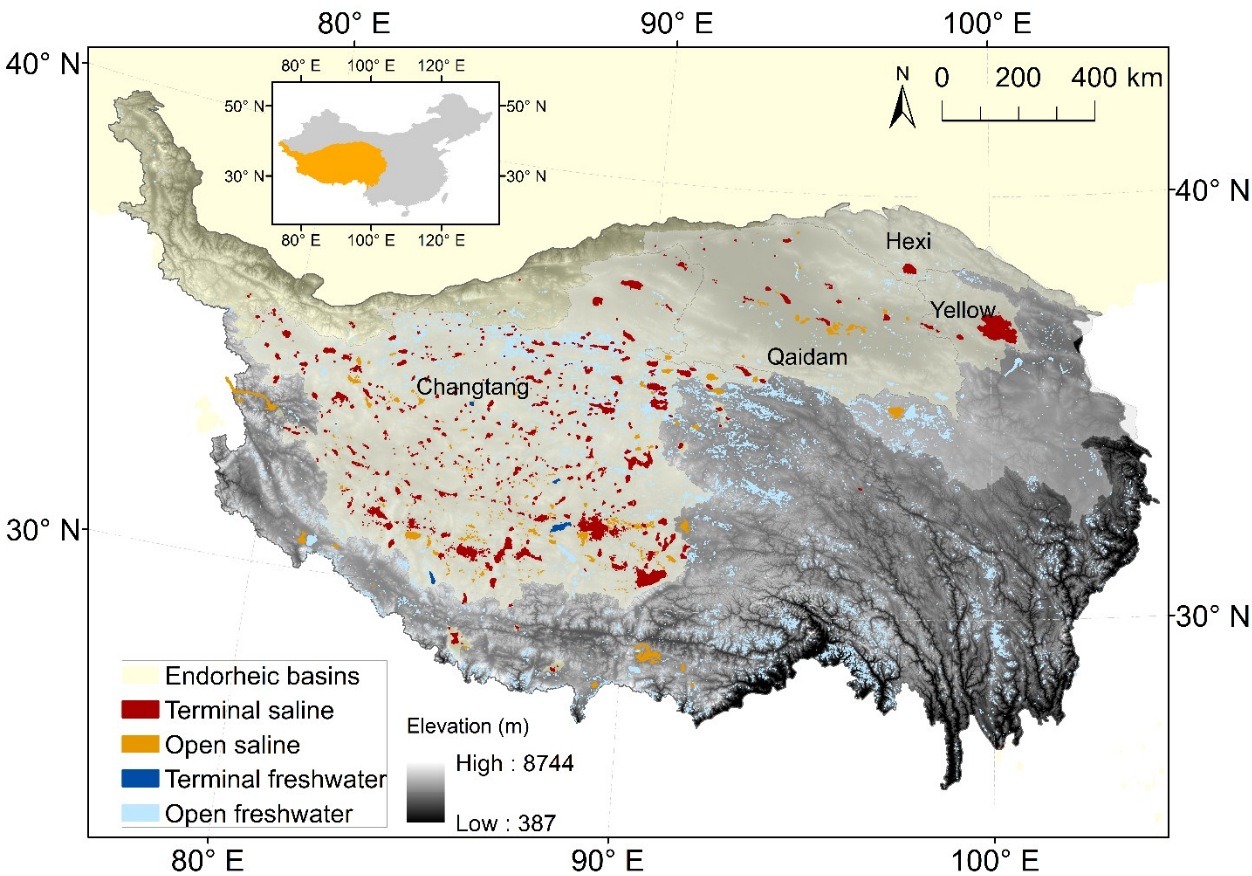


**Fig. 7.** Spatial distribution of classified terminal saline, open saline, terminal freshwater, and open freshwater lakes on the Tibetan Plateau. The lake mask circa-2015 lake dataset was used to identify lakes.

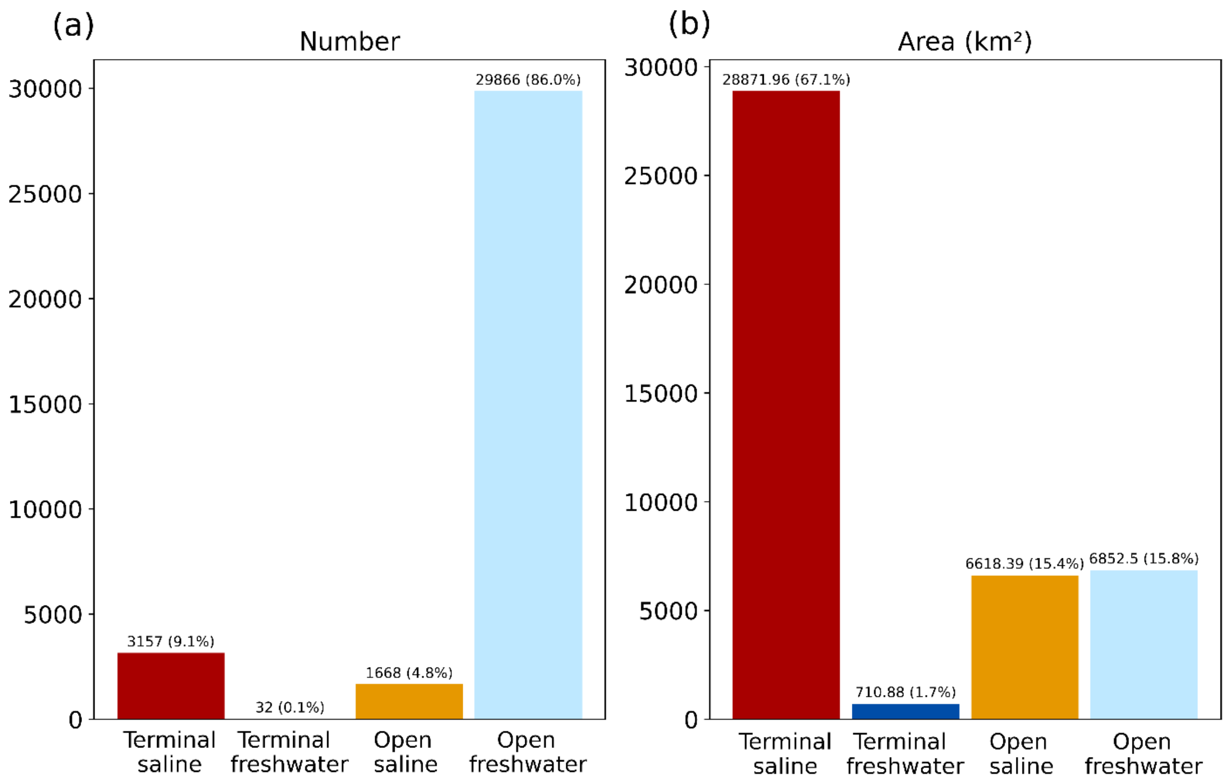


**Fig. 8.** The number of lakes (a) and size of lakes (b) by classified lake typology on the Tibetan Plateau.

Regardless of typology, lake count (abundance) and size follow a log-log linear relationship, also termed as Pareto or power-law distribution (Fig. 9). In general, the abundance of lakes decreases as the lake area increases, and the decay rate remains fairly constant on a logarithmic scale. There is also a significant difference between the abundance of freshwater lakes and saline lakes as the lake area increases. When the lake area is less than 1 $km^2$, the number of freshwater lakes is nearly an order of magnitude higher than that of saline lakes. However, the abundance of freshwater and saline lakes begins to converge for medium size lakes (1-10 $km^2$), and the count of saline lakes exceeds that of freshwater lakes when lake area is larger than 10 $km^2$ (Fig. 9). As a result, saline lakes dominate the total lake area (more than 80%) on the Tibetan

Plateau. As noted previously, our framework tends to be more conservative in identifying saline lakes as the lakes become smaller, while larger lakes are more likely to be appropriately classified. Nonetheless, the classified saline lakes span the entire range of lake size (Fig. 9), including very small lake sizes when supported by multiple lines of evidence. The Pareto relationship weakens when lakes become smaller than 0.015 $km^2$, probably because very small lakes in the circa-2015 lake dataset were mapped incompletely as they approach the minimum mapping unit in Landsat-8 images (0.004 $km^2$).

This explicit classification will facilitate a range of studies on the Tibetan Plateau, from understanding the spatial distribution of lake water resources to attributing lake changes (Zhu et al., 2020). Particularly, there has been a significant expansion of saline lakes across the Changtang Plateau over the past two to three decades, due to increased precipitation as well as warming-induced cryospheric contributions (e.g., glacier melting and permafrost thawing) (Song et al., 2014; Yao et al., 2018; Zhang et al., 2017). Therefore, understanding the time series of Tibetan lake dynamics with consideration of lake typology will be critical to monitor the changes in fresh and saline water budgets, lake salinity, and saline lake ecosystem functions.

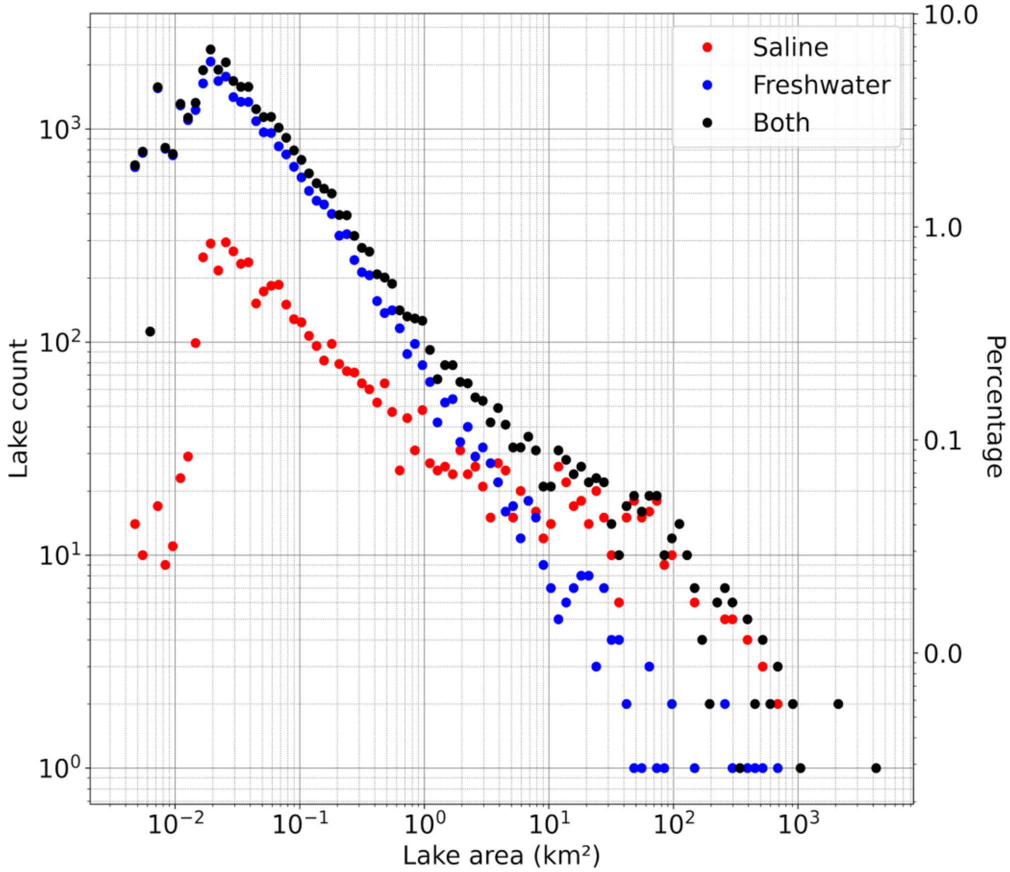


**Fig. 9.** Log-log relationship between lake area and count (abundance) in the Tibetan Plateau.

4.2 Validation

As described in Section 3.7, our validation first emphasizes the reliability of lake type classification based on non-literature (i.e., hydrological, climatological, and spectral) evidence only. We then benchmarked the accuracy of our final classification result (i.e., with inclusion of literature evidence) on the Tibetan Plateau against CNLS (Ma et al., 2010; Zhu et al., 2020).

4.2.1 Accuracy using hydroclimate and spectral evidence alone

We assessed the classification error by both lake count and lake area by calculating the number and the area of misclassified lakes (where literature evidence disagrees with hydroclimate/spectral evidence) as percentage of the total number and the total area of all classified lakes with literature evidence. Table 2 shows that the number of all classified lakes with literature evidence is 2,261 (94.5% of total lake count), which have a total area of 29,703.7 $km^2$ (93.2% of total area). Of these lakes, the number of inaccurately classified lakes is 125, with an area of 2,020.0 $km^2$. This means only using hydrological, climatological, and spectral evidence alone, our lake type classification may render an accuracy of up to 94.5% by lake count and 93.2% by lake area.

This accuracy assessment is dependent on the literature available for validation. Additional literature evidence would further strengthen validation of our framework. As described in Section 3.6, the identified lake type errors were labeled as S4, S6, S7, S8, F2, F3, and F5 depending on their evidence scenarios (Fig. 4) and have been corrected to the corresponding freshwater or saline types in the final classification result.

**Table 2.** Accuracy of lake type classification for the Tibetan Plateau using hydrological, climatological, and spectral evidence alone. Misclassified lakes fall into the following seven lake categories: i) terminal lakes without visible evaporites in endorheic basins but documented as freshwater lakes (F2); ii) open lakes with visible evaporites in endorheic basins but documented as freshwater lakes (F3); iii) open lakes without visible evaporites in endorheic basins but documented as saline lakes (S4); iv) lakes located in exorheic basins and arid regions and with visible evaporites but documented as freshwater lakes (F5); v) lakes located in exorheic basins and arid regions and without visible evaporites but documented as saline lakes (S6); vi) lakes located in exorheic basins and not in arid regions but documented as saline lakes (S7); vii) lagoons without visible evaporites but documented as saline lakes (S8)

| Category | Non-literature evidence | Literature evidence | Count | Area ($km^2$) |
|---|---|---|---|---|
| F2 | Saline | Freshwater | 32 | 710.9 |
| F3 | Saline | Freshwater | 1 | 57.7 |
| S4 | Freshwater | Saline | 91 | 1,251.4 |
| F5 | Saline | Freshwater | 1 | 0.03 |
| S6 | Freshwater | Saline | 0 | 0 |
| S7 | Freshwater | Saline | 0 | 0 |
| S8 | Freshwater | Saline | 0 | 0 |
| Total | - | - | 125 | 2,020.0 |
| All lakes with literature evidence | - | - | 2,261 | 29,703.7 |
| Accuracy | - | - | 94.5% | 93.2% |

4.2.2 Accuracy using all types of evidence

Our final classification results, which combined both literature and non-literature evidence, were benchmarked against the 1152 Tibetan Plateau lakes - including 270 freshwater lakes and 882 saline lakes - documented by CNLS. As described in Section 3.7, we excluded 19 lakes with salinity falling in the uncertainty range (i.e., salinity between 0.05% to 0.1%; Fig. 6) from validation. The results (Fig. 10) are grouped into four scenarios: correctly classified saline lakes, correctly classified freshwater lakes, saline lakes misclassified as freshwater, and freshwater lakes misclassified as saline. Their statistics are summarized in Table 3.

Our framework shows high levels of accuracy identifying saline lakes (95.0% by lake count and 98.1% by area) because we only classified a lake as saline if it possesses at least one type of primary evidence (hydrological, spectral, literature, or combined). This conservative strategy means some saline lakes lacking the considered evidence were misclassified to freshwater (i.e., omission errors), leading to a lower procedure's accuracy for saline lakes (63.3% by lake count). Even so, the procedure's accuracy for saline lakes remains high in terms of lake area (95.3%), suggesting such omission errors tend to occur only for small saline lakes. For the same reasons, problems in the classified freshwater lakes are dominated by commission errors, leading to lower user's accuracies for freshwater lakes (40.7% by lake count and 71.1% by area). The producer's accuracies for freshwater lakes are higher (88.2% by count and 86.4% by area), indicating rarer omission errors where freshwater lakes were misclassified as saline. Combining all scenarios, the overall accuracy of our classified lake typology reaches 68.9% by lake number and 94.3% by area. The classification accuracy increases as the lake size increases due to stronger classification evidence for larger lakes. In general, the conservative nature of our method suggests that our method should not be a complete substitute for in situ surveys, especially for small lakes; however, it does offer a cost-effective way to identify many of the world's saline lakes.

To further assess the reliability of using non-literature evidence only, we compared our classification results before literature correction with CNLS. The overall accuracies (67.9% by count and 92.2% by area) are comparable to those after literature correction. In other words, our classification framework based on non-literature evidence alone has correctly classified most of the large lakes, and the inclusion of literature evidence further increased the accuracy but by a limited extent (~1–2%). It is important to note that lake salinity is a dynamic parameter as it is affected by seasonal and interannual variability (Song et al., 2022). For example, the lake types in our literature evidence (mostly from Wang and Dou (1998)) were recorded decades ago whereas the information in CNLS was more recent (i.e., around 2006). Therefore, our mapping error, in theory, has included the uncertainties due to salinity changes between the period of the two datasets.

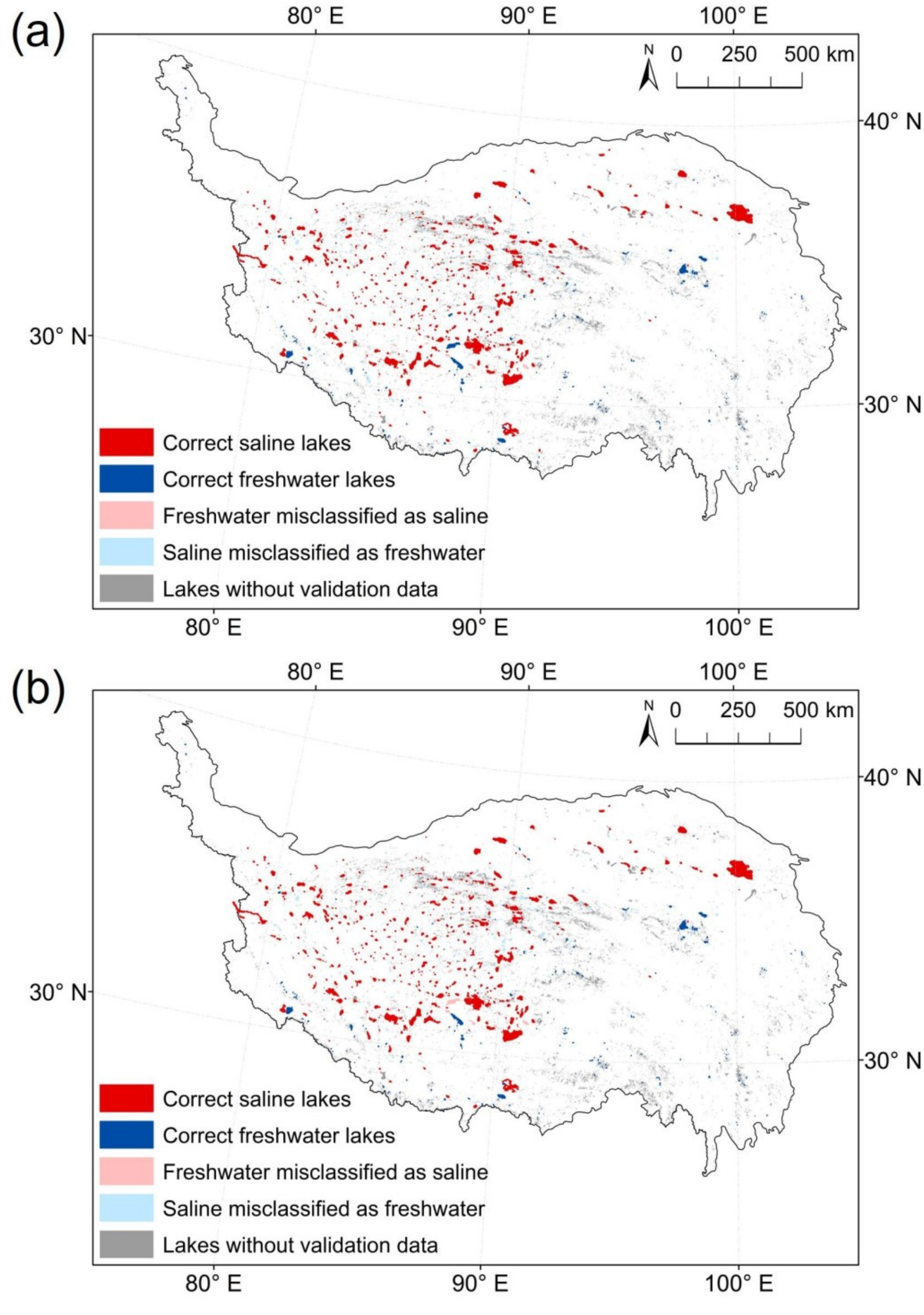


**Fig. 10.** Validation results for saline and freshwater lake type classification on the Tibetan Plateau against the Chinese national lake survey. Panel (a) shows the validation of classification results using all evidence, and (b) shows the validation of classification results using hydrological, climatological, and spectral (non-literature) evidence only.

**Table 3.** Comparison between lake typology statistics in this study and those in the Chinese national lake survey.

| Validation results | All evidence | | Non-literature evidence | |
|---|---|---|---|---|
| | Count | Area (km$^2$) | Count | Area (km$^2$) |
| Correct saline lakes | 506 (49.3%) | 30,799.2 (84.1%) | 500 (48.7%) | 30,635 (83.6%) |
| Correct freshwater lakes | 201 (19.6%) | 3,732.8 (10.2%) | 197 (19.2%) | 3,166.6 (8.6%) |
| Saline misclassified as freshwater | 293 (28.5%) | 1,513.6 (4.1%) | 299 (29.1%) | 1,677.8 (4.6%) |
| Freshwater misclassified as saline | 27 (2.6%) | 589.7 (1.6%) | 31 (3%) | 1,155.9 (3.2%) |
| Total | 1,027 (100%) | 36,635.3 (100%) | 1,027 (100%) | 36,635.3 (100%) |
| Accuracy summary | | | | |
| User's accuracy of saline lakes | 95.0% | 98.1% | 94.2% | 96.4% |
| User's accuracy of freshwater lakes | 40.7% | 71.1% | 39.7% | 65.4% |
| Producer's accuracy of saline lakes | 63.3% | 95.3% | 62.6% | 94.8% |
| Producer's accuracy of freshwater lakes | 88.2% | 86.4% | 86.4% | 73.3% |
| Overall accuracy | 68.9% | 94.3% | 67.9% | 92.2% |

Table 4 and Table 5 report more detailed accuracy assessments for our classified four lake types (i.e., terminal saline, open saline, terminal freshwater, and open freshwater). There are 104 open saline lakes misclassified as open freshwater (Table 4). These errors were attributed to the fact that the literature we collected disagrees with CNLS. We also found that two lakes, Laxiong (34.35° N, 85.23° E) and Jiesa (30.23° N, 84.79° E), were documented to be saline in CNLS but terminal freshwater in the Chinese Lake Catalogue (Wang and Dou, 1998), which we used as literature evidence. In the Chinese Lake Catalogue, Laxiong Lake was reported to have a salinity of less than 0.001%, and Jiesa Lake has no salinity record but is documented to be freshwater in Wikipedia. This situation indicates that discrepancies between the references can indeed occur, so future work may require additional collation among different literature. Some terminal lakes can be freshwater, without literature evidence, this type of lake is easily misclassified, leading to a low producer's accuracy (Table 5). Still, terminal freshwater lakes are rare, accounting for only 0.5% of all terminal lakes in the Tibetan Plateau and about 0.4% of all validated lakes. Improvements also need to be made in finding the "hidden" saline lakes without any of the three kinds of evidence. In the Tibetan Plateau, where saline lakes are very abundant, these hidden saline lakes account for about 14.8% of all validated lakes (104 open freshwater in this study). However, we expect this proportion to reduce in other regions where saline lakes prevail to lesser extents.

**Table 4.** Confusion matrix for the classified terminal saline lakes, open saline lakes, terminal freshwater lakes, and open freshwater lakes. Numbers in the upper triangle represent the user's accuracies and the numbers in the lower triangle represent the producer's accuracies.

| | | Chinese national lake survey (CNLS) | | | | |
|---|---|---|---|---|---|---|
| | Count (Area in km$^2$) | Terminal saline | Open saline | Terminal freshwater | Open freshwater | Total |
| This study | Terminal saline | 411 (27,155.8) | 0 | 7 (59.7) | 0 | 418 (27,215.5) |
| | Open saline | 0 | 18 (1,929.2) | 0 | 0 | 18 (1,929.2) |
| | Terminal freshwater | 2 (199.1) | 0 | 3 (508.5) | 0 | 5 (707.6) |
| | Open freshwater | 0 | 104 (430.2) | 0 | 155 (2,099.6) | 259 (2,529.8) |
| | Total | 413 (27,354.9) | 122 (2,359.4) | 10 (568.2) | 155 (2,099.6) | 700 (32,382.1) |

**Table 5.** Summary of classification errors and accuracies for each lake type.

| Count (Area) | Commission error | Omission error | Producer's accuracy | User's accuracy |
|---|---|---|---|---|
| Terminal saline | 1.7% (0.2%) | 0.5% (0.7%) | 99.5% (99.3%) | 98.3% (99.8%) |
| Open saline | 0% (0%) | 85.2% (18.2%) | 14.8% (81.8%) | 100% (100%) |
| Terminal freshwater | 40.0% (28.1%) | 70% (10.5%) | 30% (89.5%) | 60.0% (71.9%) |
| Open freshwater | 40.2% (17.0%) | 0% (0%) | 100% (100%) | 59.8% (83%) |

## 4.3 Application of framework to Australia

Encouraged by the promising accuracy of our classification on the Tibetan Plateau, we then applied it to the continent of Australia using the same circa-2015 lake mask. Figure 11 shows the spatial distribution of our classified terminal saline, open saline, terminal freshwater, and open freshwater lakes in Australia. Most saline lakes are found in endorheic and/or arid regions that account for a large proportion (79%) of the Australian landmass, whereas freshwater lakes are mainly located in the peripheral exorheic basins. This result echoes one of the major climatic characteristics in Australia: a vast arid interior partially caused by the subtropical high-pressure zone. As shown in the latitudinal and longitudinal distributions (Fig. 11), it is clear that saline lakes dominate the continent in terms of lake area. Three large saline lakes, Lake Mackay (22.5° S, 128.58° E) in the Great Sandy Desert, Lake Eyre (28.37° S, 137.37° E) in the Simpson Desert, and Lake Torrens (31.04° S, 137.86° E), contribute a significant proportion (12%) of the total saline lake area (Fig. 11a). Based on the literature evidence, several terminal lakes in Australia are freshwater. Examples include Lake Woods (17.84° S, 133.45° E) and Lake Goyder (26.78° S, 139.13° E), both of which are ephemeral with low salinity (Shipton et al., 2021; Badman and May, 1983). As summarized in Fig. 12, saline and freshwater lakes account for 44.15% and 55.85% of the total lake number, respectively, and the aggregated area of saline lakes is about three times larger than that of freshwater lakes.

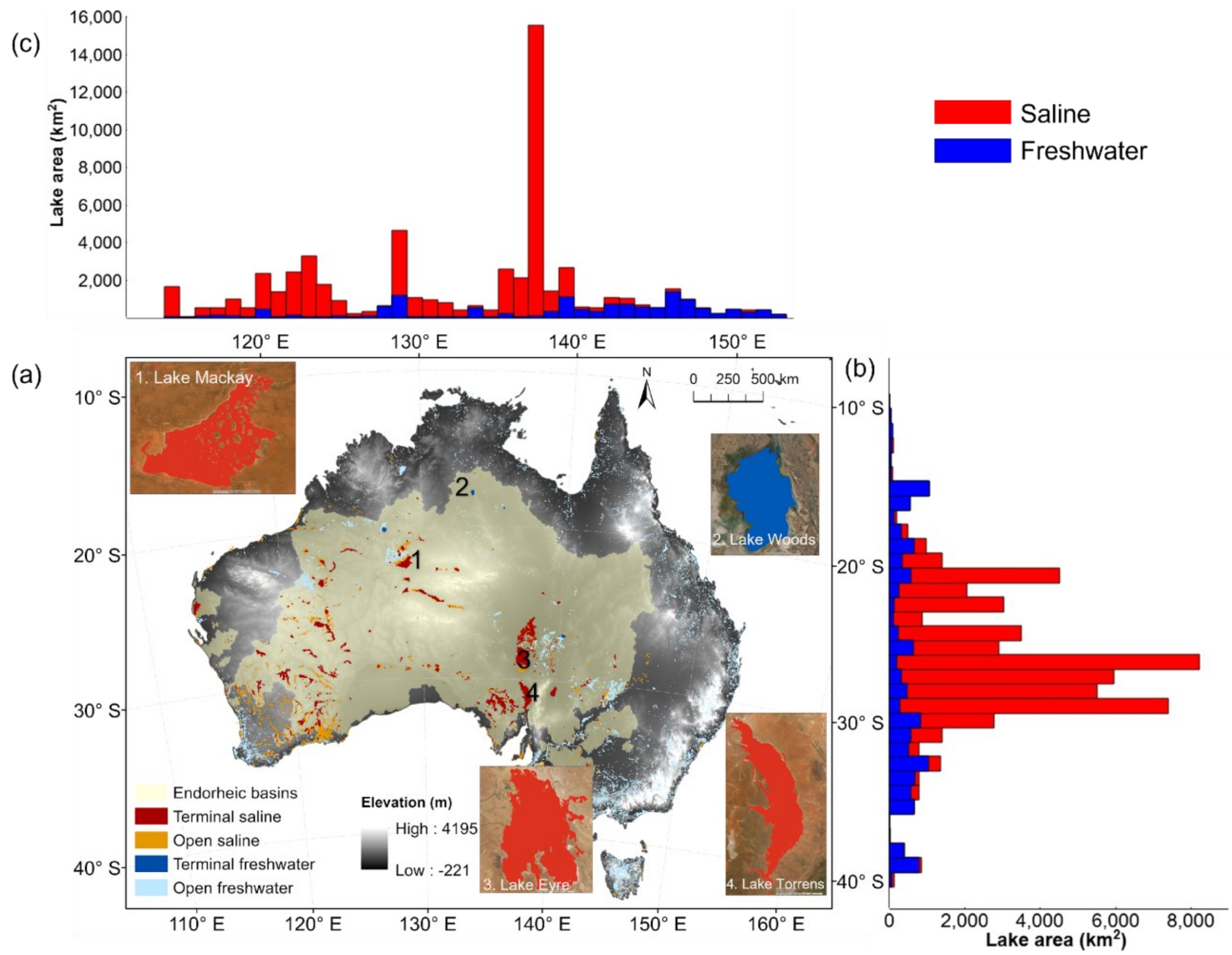


**Fig. 11.** Classified freshwater and saline lake typology in Australia using the circa-2015 lake dataset. Panel (a): Spatial distribution of terminal saline lakes, open saline lakes, terminal freshwater lakes, and open freshwater lakes. Panels (b) and (c) are latitudinal and longitudinal distributions of the total freshwater lake area and saline lake area, respectively.

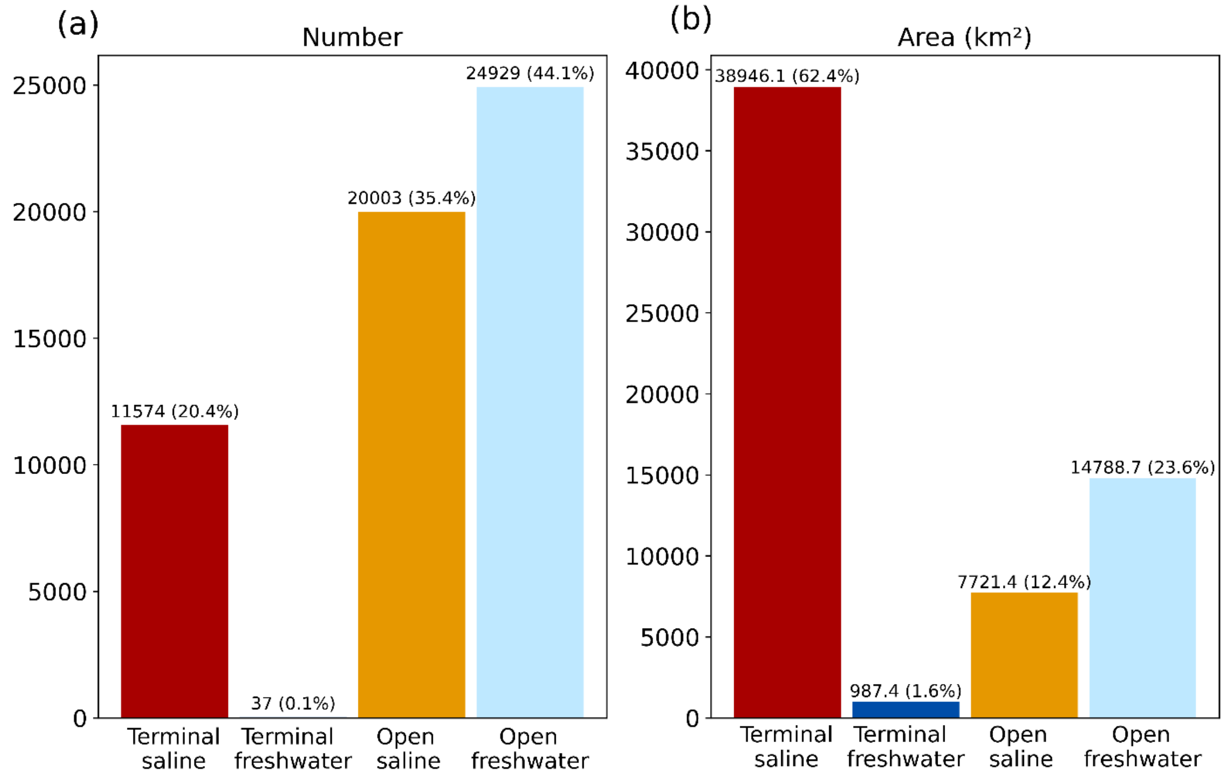


**Fig. 12.** Statistics of the classified lake typology in Australia. Panel (a) and panel (b) show lake numbers and lake areas, respectively.

Table 6 reports the accuracy of lake type classification based on hydroclimate and spectral evidence alone, validated with reference to our collected literature evidence. The overall accuracies are 99.1% by lake count and 96.9% by lake area, which are generally consistent with, albeit slightly higher than, the accuracies for the Tibetan Plateau. The category F2, terminal freshwater lake, accounts for the largest proportion (66.1% by count and 86.7% by area) of classification errors. Several factors may contribute to terminal lakes being freshwater, including salt decreases due to lake water seepage to the aquifer (Pham-Duc et al., 2020), ephemeral lakes caused by short-term flooding (Whitford and Duval., 2019), and local government protection and management of the lake water environment. However, this type of error constitutes only 0.6% (or 2.7% by area) of all lakes with literature evidence.

**Table 6.** Accuracy of lake type classification for Australia using hydrological, climatological, and spectral evidence alone. Misclassified lakes fall into the following seven lake categories: i) terminal lakes without visible evaporites in endorheic basins but documented as freshwater lakes (F2); ii) open lakes with visible evaporites in endorheic basins but documented as freshwater lakes (F3); iii) open lakes without visible evaporites in endorheic basins but documented as saline lakes (S4); iv) lakes located in exorheic basins and arid regions and with visible evaporites but documented as freshwater lakes (F5); v) lakes located in exorheic basins and arid regions and without visible evaporites but documented as saline lakes (S6); vi) lakes located in exorheic basins and not in arid regions but documented as saline lakes (S7); vii) lagoons without visible evaporites but documented as saline lakes (S8)

| Category | Non-literature evidence | Literature evidence | Count | Area ($km^2$) |
|---|---|---|---|---|
| F2 | Saline | Freshwater | 37 | 987.4 |
| F3 | Saline | Freshwater | 0 | 0 |
| S4 | Freshwater | Saline | 13 | 22.8 |
| F5 | Saline | Freshwater | 6 | 129.7 |
| S6 | Freshwater | Saline | 0 | 0 |
| S7 | Freshwater | Saline | 0 | 0 |
| S8 | Freshwater | Saline | 0 | 0 |
| Total | - | - | 56 | 1,139.9 |
| All lakes with literature evidence | - | - | 6,145 | 36,649.4 |
| Accuracy | - | - | 99.1% | 96.9% |

The relationship between lake count (abundance) and lake area in Australia follows a Pareto distribution (Fig.13), as was the case in the Tibetan Plateau. However, the abundance of saline lakes, as well as the slope of the count-area curve, is comparable to that of freshwater lakes (also shown in Fig. 12), which contrast with what was observed in the Tibetan Plateau. Our classification also captures many small saline lakes, as the power-law relationship still holds when saline lake size decreases. This is partially because under extreme aridity, many small saline lakes in Australia still exhibit clear spectral evidence from high-resolution images, and some of them are clustered as patchy water residuals within large saline lake basins or salt pans. Some of the classified small saline lakes, particularly in the periphery of major fluvial systems, can be highly ephemeral or transitory. During the flood season, some of them can be replenished by freshwater flows, so their salinity may vary significantly within the year. However, we opted to treat them as saline lakes to indicate that these water bodies are located in visibly saline environments and tend to become saltier as they are disconnected from the rivers and their water evaporates during the drier seasons. It is also worth mentioning that saline lagoons also contribute to the saline lake area in Australia. About 3% of the total saline lake area in Australia comes from lagoons, and they were separated from freshwater lagoons using spectral and/or literature evidence (Fig. 4; Table 1).

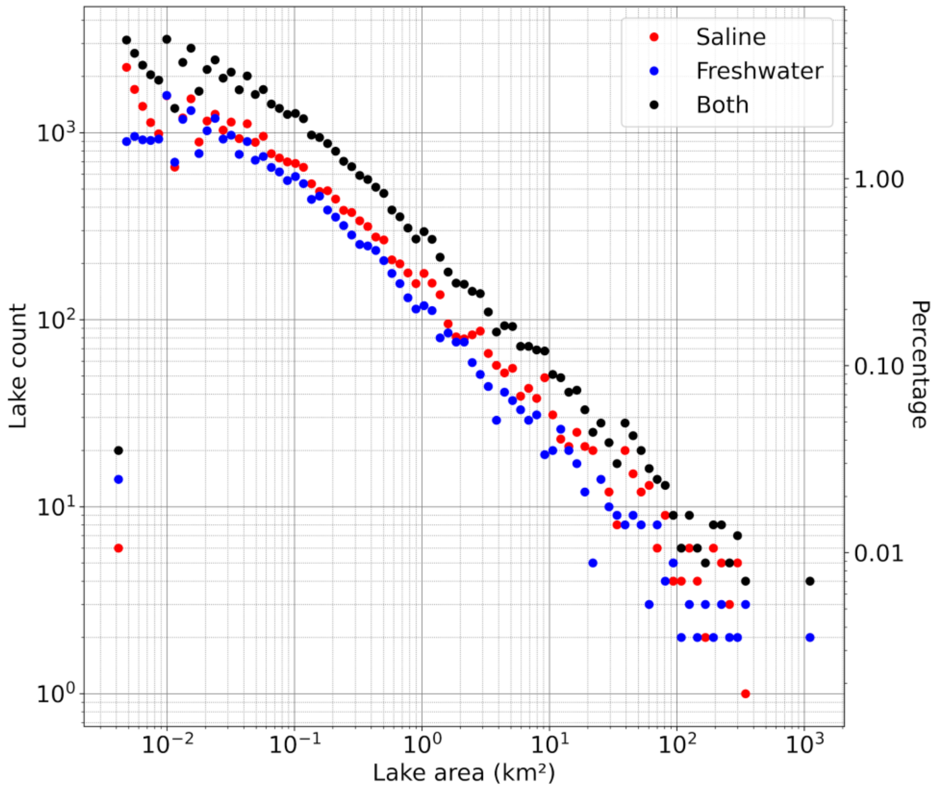


**Fig. 13.** Log-log relationship between lake area and count (abundance) in Australia.

## 4.4 Limitations and discussion

Our classification method is overall conservative in extracting saline lakes. As our schematic flowchart shows, only lakes that exhibit sufficient evidence were classified as saline lakes. In this case, our method provides feasibility to identify some of the most evident saline lakes with reasonable accuracies. However, since the evidence for saline lakes tends to weaken as the lake size decreases, the omission error generally increases for smaller saline lakes, leading to an overall conservative recognition of the saline lake number. This was reflected in our validation for the Tibetan Plateau. Although the quantity of saline lakes is underestimated, our method is more successful in capturing the total area of saline lakes, with both user's and producer's accuracies exceeding 95% for the Tibetan Plateau. For this reason, it is important to reiterate that our classification method, with its effectiveness dependent on the presence of remotely sensed evidence, should not be considered a complete substitute for in situ measurement. Furthermore, we also acknowledge that human activities also affect lake salinity (Tang et al. 2022; Williams, 2001). For example, agricultural water use can lead to overexploitation of rivers and lakes and thus increase in salinity. The influence of human activities can be further considered when studying changes in lake typology and salinity. The focus of this study is to provide a framework to differentiate between freshwater and saline lakes, and therefore the discussion of the effects of human activities is not involved.

Our current identification of spectral evidence relies on visual interpretation. Existing studies related to salt-affected surface extraction applied more automated methods. For example, Safaee and Wang (2020) implemented machine learning to delineate salt pans and playas, but the accuracy of the results is limited to the image spectral and spatial resolutions (Landsat-8 multispectral images) and the ability to use multispectral information alone to separate salts from other high-reflectance substance. Wang et al (2020) used satellite hyperspectral images (acquired from China's Huanjing Satellite 1) and the fractional derivatives method to estimate soil salt contents. In another study, Liu et al (2018) analyzed the surface reflectance spectrum from airborne hyperspectral images to construct a salt index and label the salinized soil. Hyperspectral information, although more effective in separating surface substances, is not always freely available. In addition, a common challenge of these studies is that the validity of the methods may be limited by the characteristics of the study area and sample collection, and the accuracy varies from place to place. Therefore, these methods may not always be effective on a global scale. For the purpose of "protocol and framework development", we replaced the automated methods by an expert-based visual interpretation based on high-resolution multispectral imagery. This visual interpretation allowed us to comparatively analyze textural and contextual information so as to optimize the accuracy across heterogeneous surface conditions. However, our framework remains flexible and does not inhibit future inclusion of automated methods.

In addition, our current lake type classification framework outputs a binary typology (i.e., freshwater and saline), although many of the saline lakes we identified were also provided with a numeric salinity value collected from the literature (~7% in Tibetan Plateau, and ~5% in Australia). These records were used to binarize the lake classes, and they were also used (together with other nominal literature records) to verify and assess our classification accuracy based on hydroclimate and spectral evidence alone (see details in Section 4.2). While we made every endeavor to collect as much documentation in the form of published papers and reports, lakes with numeric salinity documentation are still relatively rare given the sheer number of lakes in the study areas. We could have applied the limited numeric salinities to enable a more detailed typology (e.g., including subsaline, polysaline, and hypersaline), but some of the literature only reported nominal lake types which were not always based on a consistent typology system or did not inform specific salinity thresholds. Given these uncertainties, we decided to only keep binary categories (freshwater and saline) for the final lake typology.

## 5. Summary and concluding remarks

Freshwater resources are finite and vital for human activities. As one of the most readily accessible water resources, freshwater lake storage can contribute to human domestic, energy, industrial, and agricultural uses. In addition, both freshwater lakes and saline lakes play important roles in ecological services and climate change adaptation (Vareschi, 1987; Jódar et al., 2020).

Therefore, understanding the spatial distribution of freshwater and saline lakes has important ramifications for limnology, ecology, and water security and management. However, classifying freshwater and saline lake typology across large areas is challenging considering the high cost and low efficiency of in situ surveys. Lake classification is further challenged by remote sensing algorithms that are not always transferable among different saline environments. To address these challenges, we composed a systematic lake type classification framework, which followed the "water-solute imbalance" to theorize the conditions where lake salinity tends to accumulate. The conditions were then identified as thoroughly as possible, by leveraging multi-dimensional evidence including drainage terminals (hydrological), aridity (climatological), evaporite deposits (spectral), and existing documentations (literature), which required no additional in situ measurements.

When our framework was applied to the Tibetan Plateau, we found that saline lakes, despite constituting less than 15% of the lake population, cover 82.4% of the total lake area across the Tibetan Plateau. Nearly all of the classified saline lakes are distributed in endorheic basins. Freshwater lakes dominate the lake population (86.1%) despite a smaller total area, and most of the freshwater lakes (64.16%) are located in exorheic basins and/or humid regions. Validated by CNLS, our overall classification accuracy reaches nearly 70% in terms of lake count and more than 90% in terms of lake area. Given such promising accuracies, the classification method was replicated in Australia. Our result shows that freshwater and saline lakes in Australia are comparable in count (55.85% for saline lakes and 44.15% for freshwater lakes) but the total area of saline lakes is about three times that of freshwater lakes. The case studies in these two saline lake dense regions suggest the potential of applying our method in classifying lake typology in other regions of the world.

We would like to reiterate that our lake type classification was achieved completely based on existing literature and open-source remote sensing, climate, and hydrography datasets. Although the method was proven to be reasonably accurate, there are cases where a saline lake possesses none of the considered evidence, leading to an overall conservative saline lake identification. For this reason, our method should not be considered a substitute for field surveys. Classifying global lakes into freshwater and saline categories on a high-resolution global lake inventory is a difficult task due to the sheer number of lakes and the complex hydroclimate settings. Despite these limitations, particularly a conservative accuracy for identifying saline lakes, our method provides a cost-effective way to provide a first-order separation between freshwater and saline lake systems at broad geographic scales.

## Authorship contribution statement

**Meng Ding:** Formal analysis, Investigation, Validation, Visualization, Writing - original draft, Writing - review & editing. **Jida Wang:** Conceptualization, Methodology, Formal analysis, Investigation, Validation, Visualization, Resources, Supervision, Project administration, Writing – original draft, Writing - review & editing. **Chunqiao Song:** Methodology, Resources, Writing - review & editing. **Yongwei Sheng:** Resources, Writing - review & editing. **J.M. Shawn Hutchinson:** Writing - review & editing. **Abigail L. Langston:** Writing - review & editing. **Landon Marston:** Writing - review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

This research was in part supported by the faculty start-up fund to JW and NASA Surface Water and Ocean Topography (SWOT) Grant (#80NSSC20K1143). The authors gratefully acknowledge Takuto Urano, Pedro Satori, Blake A. Walter, and Colin Bailey at Kansas State University, who made critical contributions to data preparations and part of the lake classification at an earlier stage of the research.

## References


Abbott, B. W., Bishop, K., Zarnetske, J. P., Minaudo, C., Chapin, F. S., Krause, S., Hannah, D. M., Conner, L., Ellison, D., Godsey, S. E., Plont, S., Marçais, J., Kolbe, T., Huebner, A., Frei, R. J., Hampton, T., Gu, S., Buhman, M., Sayedi, S., Ursache, O., Chapin, M., Henderson K. D., & Pinay G. (2019). Human domination of the global water cycle absent from depictions and perceptions. *Nature Geoscience*, 12, 533–540.

Adrian, R., O'Reilly, C. M., Zagarese, H., Baines, S. B., Hessen, D. O., Keller, W., Livingstone, D. M., Sommaruga, R., Straile, D., Donk, E., Weyhenmeyer, G. A., & Winder, M. (2009). Lakes as sentinels of climate change. *Limnology and oceanography*, 54(6part2), 2283–2297.

Alley, R. B., Meese, D. A., Shuman, C. A., Gow, A. J., Taylor, K. C., Grootes, P. M., White, J. W. C., Ram, M., Waddington, E. D., Mayewski, P. A., & Zielinski, G. A. (1993). Abrupt increase in Greenland snow accumulation at the end of the Younger Dryas event. *Nature*, *362*(6420), 527–529.

Azarafza, M., & Mokhtari, M. (2013). Evaluation of drought effect on Urmia Lake salinity changes using remote sensing techniques. *Journal of Arid Biome*, 3(2), 1–14.

Badman, F. J., & May, I. A. (1983). Waders in northern South Australia. *South Australian Ornithologist*, 29, 29–39.

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438(7066), 303–309.

Beard, J. S. (1969). The natural regions of the deserts of Western Australia. *The Journal of Ecology*, 677-711.

Birkett, C. M., & Mason, I. M. (1995). A new global lakes database for a remote sensing program studying climatically sensitive large lakes. *Journal of Great Lakes Research*, 21(3), 307–318.

Birkett, C. M., O'Brien, K., Kinsey, S., Ricko, M., & Li, Y. (2022). Enhancement of a global lake and reservoir database to aid climate studies and resource monitoring utilizing satellite radar altimetry. *Journal of Great Lakes Research*, *48*(1), 37-51.

Byrne, M., Yeates, D. K., Joseph, L., Kearney, M., Bowler, J., Williams, M. A. J., Cooper, S., Donnellan, S.C., Keogh, J.S., Leys, R., Melville, J., Murphy, D.J., Porch, N., & Wyrwoll, K. H. (2008). Birth of a biome: insights into the assembly and maintenance of the Australian arid zone biota. *Molecular ecology*, *17*(20), 4398.

Cael, B. B., Heathcote, A. J., & Seekell, D. A. (2017). The volume and mean depth of Earth's lakes. *Geophysical Research Letters*, 44(1), 209–218.

Carroll, M. L., Townshend, J. R., DiMiceli, C. M., Noojipady, P., & Sohlberg, R. A. (2009). A new global raster water mask at 250 m resolution. *International Journal of Digital Earth*, 2(4), 291–308.

Chahine, M. T. (1992). The hydrological cycle and its influence on climate. *Nature*, 359(6394), 373–380.

Cuo, L., Zhang, Y., Zhu, F., & Liang, L. (2014). Characteristics and changes of streamflow on the Tibetan Plateau: A review. *Journal of Hydrology: Regional Studies*, *2*, 49-68.

Emmanuel, B. E., & Chukwu, L. O. (2010). Spatial distribution of saline water and possible sources of intrusion into a tropical freshwater lagoon and the transitional effects on the lacustrine ichthyofaunal diversity. *African Journal of Environmental Science and Technology*, *4*(7), 480-491.

Feng, M., Sexton, J. O., Channan, S., & Townshend, J. R. (2016). A global, high-resolution (30-m) inland water body dataset for 2000: First results of a topographic–spectral classification algorithm. *International Journal of Digital Earth*, 9(2), 113–133.

Guo, Y., Zhu, L., Frenzel, P., Ma, Q., Ju, J., Peng, P., Wang, J., & Daut, G. (2016). Holocene lake level fluctuations and environmental changes at Taro Co, southwestern Tibet, based on ostracod-inferred water depth reconstruction. *The Holocene*, *26*(1), 29–43.

Haddeland, I., Heinke, J., Biemans, H., Eisner, S., Flörke, M., Hanasaki, N., Konzmann, M., Ludwig, F., Masaki, Y., Schewe, J., Stacke, T., Tessler, Z. D., Wada, Y., & Wisser, D. (2014). Global water resources affected by human interventions and climate change. *Proceedings of the National Academy of Sciences*, *111*(9), 3251–3256.

Hammer, U. T. (1983). *Saline Lakes (Developments in Hydrology)*, Junk, 264.

Hammer, U. T. (1986). *Saline lake ecosystems of the world* (Vol. 59). Springer Science & Business Media.

Huang, Q., Qin, Z., & Zeng, Z. (2016). A study on the relative radiometric normalization of multi-sources and multi-temporal remote sensing data. *J. Geo-Inf. Sci.*, *18*, 606-14.

Harris, P. W. V. (1976). The seasonal temperature-salinity structure of a glacial lake: Jökulsarlon, south-east Iceland. *Geografiska Annaler: Series A, Physical Geography*, *58*(4), 329–336.

Herdendorf, C. E. (1990). Distribution of the world's large lakes. *Large Lakes*, 3-38. Springer, Berlin, Heidelberg.

Isiorho, S.A., & Matisoff, G. (1990). Groundwater recharge from Lake Chad. *Limnology and Oceanography*, *35*(4): 931–938.

Jackson, R. B., Carpenter, S. R., Dahm, C. N., McKnight, D. M., Naiman, R. J., Postel, S. L., & Running, S. W. (2001). Water in a changing world. *Ecological Applications*, 11(4), 1027–1045.

Jódar, J., Rubio, F. M., Custodio, E., Martos-Rosillo, S., Pey, J., Herrera, C., Turu, V., Pérez-Bielsa, C., Ibarra, P. & Lambán, L. J. (2020). Hydrogeochemical, isotopic and geophysical characterization of saline lake systems in semiarid regions: The Salada de Chiprana Lake, Northeastern Spain. *Science of The Total Environment*, *728*, 138848.

Joyce, C. B., Vina-Herbon, C., & Metcalfe, D. J. (2005). Biotic variation in coastal water bodies in Sussex, England: implications for saline lagoons. *Estuarine, Coastal and Shelf Science*, *65*(4), 633-644.

Kashid, S. G., & Pardeshi, S. A. (2014). A survey of water distribution system and new approach to intelligent water distribution system. In *2014 First International Conference on Networks & Soft Computing (ICNSC2014)*, 339–344.

Kim, T., & Han, Y. (2021). Integrated Preprocessing of Multitemporal Very-High-Resolution Satellite Images via Conjugate Points-Based Pseudo-Invariant Feature Extraction. *Remote Sensing*, *13*(19), 3990.

Kjerfve, B., Schettini, C. A. F., Knoppers, B., Lessa, G., & Ferreira, H. O. (1996). Hydrology and salt balance in a large, hypersaline coastal lagoon: Lagoa de Araruama, Brazil. *Estuarine, Coastal and Shelf Science*, 42(6), 701–725.

Lehner, B., & Döll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology*, 296(1–4), 1–22.

Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos, Transactions American Geophysical Union*, 89(10), 93–94.

Leidonald, R., Muhtadi, A., Lesmana, I., Harahap, Z. A., & Rahmadya, A. (2019). Profiles of temperature, salinity, dissolved oxygen, and pH in Tidal Lakes. In *IOP Conference Series: Earth and Environmental Science* (Vol. 260, No. 1, p. 012075). IOP Publishing.

Li, J., & Fang, X. (1999). Uplift of the Tibetan Plateau and environmental changes. *Chinese Science Bulletin*, *44*(23), 2117-2124.

Liu, D., Qiu, J., Tong, Q., Xie, T., Yang, Y., & Wang, Z. (2018). Airborne hyperspectral study of spatial relationship between oil and gas leakage and salinization in Jimusar area. *Geology in China*, *45*(5), 1062–1073.

Lyons, E. A., & Sheng, Y. (2017). LakeTime: Automated seasonal scene selection for global lake mapping using Landsat ETM+ and OLI. *Remote Sensing*, *10*(1), 54.

Ma, R., Duan, H., Hu, C., Feng, X., Li, A., Ju, W., Jiang J., & Yang, G. (2010). A half - century of changes in China's lakes: Global warming or human influence? *Geophysical Research Letters*, 37(24).

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., & Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications*, *7*, 13603, DOI:10.1038/ncomms13603.

Oki, T., & Kanae, S. (2006). Global hydrological cycles and world water resources. *Sicence*, 313(5790), 1068–1072.

Oliver J.E. (2005) Maritime Climate. In: Oliver J.E. (eds) *Encyclopedia of World Climatology*. *Encyclopedia of Earth Sciences Series*. Springer, Dordrecht. https://doi.org/10.1007/1-4020-3266-8_131

Pham-Duc, B., Sylvestre, F., Papa, F., Frappart, F., Bouchez, C., & Crétaux, J. F. (2020). The Lake Chad hydrology under current climate change. *Scientific reports*, *10*(1), 5498.

Pi, X., Luo, Q., Feng, L., Xu, Y., Tang, J., Liang, X., Ma, E., Cheng, R., Fensholt, R., Brandt, M., Cai, X., Gibson, L., Liu, J., Zheng, C., Li, W., & Bryan, B. A. (2022). Mapping global lake dynamics reveals the emerging roles of small lakes. *Nature Communications*, *13*(1), 5777.

Rubel, F., & Kottek, M. (2010). Observed and projected climate shifts 1901-2100 depicted by world maps of the Köppen-Geiger climate classification. *Meteorologische Zeitschrift*, *19*(2), 135.

Safaee, S., & Wang, J. (2020). Towards global mapping of salt pans and salt playas using Landsat imagery: a case study of western United States. *International Journal of Remote Sensing*, *41*(22), 8693–8716.

Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N. W., Clark, D. B., Dankers, R., Eisner, S., Fekete, B. M., Colón-González, F. J., Gosling, S. N., Kim, H., Liu, X., Masaki, Y., Portmann, F. T., Satoh, Y., Stacke, T., Tang, Q., Wada, Y., Wisser, D., Albrecht, T., Frieler, K., Piontek, F., Warszawski, L., & Kabat, P. (2014). Multimodel assessment of water scarcity under climate change. *Proceedings of the National Academy of Sciences*, 111(9), 3245–3250.

Schlesinger, W. H., & Bernhardt, E. S. (2013). The global cycles of nitrogen and phosphorus. *Biogeochemistry: an analysis of global change*, 445–467.

Schmitt, R. W. (1995). The ocean component of the global water cycle. *Reviews of Geophysics*, 33(S2), 1395–1409.

Sheng, Y., Song, C., Wang, J., Lyons, E. A., Knox, B. R., Cox, J. S., & Gao, F. (2016). Representative lake water extent mapping at continental scales using multi-temporal Landsat-8 imagery. *Remote Sensing of Environment*, 185, 129–141.

Shipton, C., Cohen, T., Forbes, M., Boesl, F., Jacobs, Z., Dixon, R. D., Kingston, S., Albert, C., & O'Connor, S. (2021). Diverse stone artefacts around Lake Woods, Central Northern Territory, Australia. *Australian Archaeology*, 87(2), 156–178.

Slater, J. A., Garvey, G., Johnston, C., Haase, J., Heady, B., Kroenung, G., & Little, J. (2006). The SRTM data "finishing" process and products. *Photogrammetric Engineering & Remote Sensing*, 72(3), 237–247.

Smith, L. C., Sheng, Y., MacDonald, G. M., & Hinzman, L. D. (2005). Disappearing arctic lakes. *Science*, 308(5727), 1429–1429.

Song, C., Huang, B., Richards, K., Ke, L., & Hien Phan, V. (2014). Accelerated lake expansion on the Tibetan Plateau in the 2000s: Induced by glacial melting or other processes?. *Water Resources Research*, 50(4), 3170–3186.

Song, C., Luo, S., Liu, K., Chen, T., Zhang, P., & Fan, C. (2022). Widespread declines in water salinity of the endorheic Tibetan Plateau lakes. *Environmental Research Communications*, *4*(9), 091002.

Song, C., Sheng, Y., Ke, L., Nie, Y., & Wang, J. (2016). Glacial lake evolution in the southeastern Tibetan Plateau and the cause of rapid expansion of proglacial lakes linked to glacial-hydrogeomorphic processes. *Journal of hydrology*, 540, 504–514.

Stricker, V. (1983). *Base flow of streams in the outcrop area of southeastern sand aquifer, South Carolina, Georgia, Alabama, and Mississippi*. Atlanta: US Geological Survey.

SWBD, 2005. Shuttle Radar Topography Mission Water Body Data set (SWBD). Accessed from www2.jpl.nasa.gov/srtm.

Tang, X., Xie, G., Deng, J., Shao, K., Hu, Y., He, J., ... & Gao, G. (2022). Effects of climate change and anthropogenic activities on lake environmental dynamics: A case study in Lake Bosten Catchment, NW China. *Journal of Environmental Management*, *319*, 115764.

Tian, Q., Zengrang, X., & Ziqian, W. (2022). Estimation of the Tibetan Wild Ass Population in Gaize County of Chang Tang Plateau based on the Belt Transect Method and Random Forest Model. *Journal of Resources and Ecology*, *13*(5), 860-869.

Pekel, J. F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633), 418–422.

Ueda, S., Kawabata, H., Hasegawa, H., & Kondo, K. (2000). Characteristics of fluctuations in salinity and water quality in brackish Lake Obuchi. *Limnology*, *1*(1), 57–62.

Vareschi, E. (1987). Saline lake ecosystems. *Potentials and limitations of ecosystem analysis*, 347-364.

Vörösmarty, C. J., Green, P., Salisbury, J., & Lammers, R. B. (2000). Global water resources: vulnerability from climate change and population growth. *Science*, 289(5477), 284–288.

Wada, Y., de Graaf, I. E., & van Beek, L. P. (2016). High‐resolution modeling of human and climate impacts on global water resources. *Journal of Advances in Modeling Earth Systems*, 8(2), 735–763.

Wada, Y., Van Beek, L. P. H., Viviroli, D., Dürr, H. H., Weingartner, R., & Bierkens, M. F. (2011). Global monthly water stress: 2. Water demand and severity of water stress. *Water Resources Research*, 47(7), W07518, DOI: 10.1029/2010WR009792.

Wandres, M., Pattiaratchi, C., Wijeratne, E. S., & Hetzel, Y. (2016). The influence of the subtropical high-pressure ridge on the western Australian wave climate. *Journal of Coastal Research*, (75 (10075)), 567-571.

Wang, F., & Xu, Y. J. (2008). Development and application of a remote sensing-based salinity prediction model for a large estuarine lake in the US Gulf of Mexico coast. *Journal of hydrology*, *360*(1-4), 184–194.

Wang, J. (2020). Endorheic water. *International Encyclopedia of Geography: People, the Earth, Environment and Technology*. The American Association of Geographers; John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118786352.wbieg20

Wang, J., Sheng, Y., Hinkel, K. M., & Lyons, E. A. (2012). Drained thaw lake basin recovery on the western Arctic Coastal Plain of Alaska using high-resolution digital elevation models and remote sensing imagery. *Remote Sensing of Environment*, 119, 325–336.

Wang, J., Sheng, Y., & Tong, T. S. D. (2014). Monitoring decadal lake dynamics across the Yangtze Basin downstream of Three Gorges Dam. *Remote Sensing of Environment*, 152, 251–269.

Wang, J., Sheng, Y., & Wada, Y. (2017). Little impact of the Three Gorges Dam on recent decadal lake decline across China's Yangtze Plain. *Water Resources Research*, 53(5), 3854–3877.

Wang, J., Song, C., Reager, J. T., Yao, F., Famiglietti, J. S., Sheng, Y., MacDonald, G. M., Brun, F., Schmied, H. M., Marston R. A., & Wada. Y., (2018). Recent global decline in endorheic basin water storages. *Nature Geoscience*, 11(12), 926–932.

Wang, S., and H. Dou (1998). *Chinese Lake Catalogue*, Science Press, Beijing.

Wang, S (1993). Physical and chemical characteristics of saline lakes in the world. *Lake Science*, 5(3), 278-286.

Wang, Z., Wang, J., & Han, J. (2022). Spatial prediction of groundwater potential and driving factor analysis based on deep learning and geographical detector in an arid endorheic basin. *Ecological Indicators*, *142*, 109256.

Wang, Z., Zhang, X., Zhang, F., Weng Chan, N., Liu, S., & Deng, L. (2020). Estimation of soil salt content using machine learning techniques based on remote-sensing fractional derivatives, a case study in the Ebinur Lake Wetland National Nature Reserve, Northwest China. *Ecological Indicators*, *119*, 106869.

Williams, W. D. (2001). Anthropogenic salinisation of inland waters. In *Saline Lakes: Publications from the 7th International Conference on Salt Lakes, held in Death Valley National Park, California, USA, September 1999* (pp. 329-337). Springer Netherlands.

Williams, W. D., & Sherwood, J. E. (1994). Definition and measurement of salinity in salt lakes. *International Journal of Salt Lake Research*, *3*(1), 53–63.

Whitford, W. G., & Duval, B. D. (2019). *Ecology of desert systems*. Academic Press.

Wurtsbaugh, W. A., Miller, C., Null, S. E., DeRose, R. J., Wilcock, P., Hahnenberger, M., Howe, F., & Moore, J. (2017). Decline of the world's saline lakes. *Nature Geoscience*, 10(11), 816–821.

Xu, Z., Wei, Z., & Jin, M. (2020). Causes of domestic livestock–wild herbivore conflicts in the alpine ecosystem of the Chang Tang Plateau. *Environmental Development*, *34*, 100495.

Yamazaki, D., Trigg, M. A., & Ikeshima, D. (2015). Development of a global~ 90 m water body map using multi-temporal Landsat images. *Remote Sensing of Environment*, 171, 337–351.

Yang, K., Wu, H., Qin, J., Lin, C., Tang, W., & Chen, Y. (2014). Recent climate changes over the Tibetan Plateau and their impacts on energy and water cycle: A review. *Global and Planetary Change*, *112*, 79-91.

Yang, K., Yao, F., Wang, J., Luo, J., Shen, Z., Wang, C., & Song, C. (2017). Recent dynamics of alpine lakes on the endorheic Changtang Plateau from multi-mission satellite data. *Journal of Hydrology*, 552, 633–645.

Yao, F., Wang, J., Wang, C., & Crétaux, J. F. (2019). Constructing long-term high-frequency time series of global lake and reservoir areas using Landsat imagery. *Remote Sensing of Environment*, 232, 111210.

Yao, F., Wang, J., Yang, K., Wang, C., Walter, B. A., & Crétaux, J. F. (2018). Lake storage variation on the endorheic Tibetan Plateau and its attribution to climate change since the new millennium. *Environmental Research Letters*, 13(6), 064011.

Zhang, G., Xie, H., Kang, S., Yi, D., & Ackley, S. F. (2011). Monitoring lake level changes on the Tibetan Plateau using ICESat altimetry data (2003–2009). *Remote Sensing of Environment*, *115*(7), 1733–1742.

Zhang, G., Yao, T., Shum, C. K., Yi, S., Yang, K., Xie, H., Feng, W., Bolch, T., Wang, L., Behrangi, A., Zhang, H., Wang, W., Xiang, Y., & Yu, J. (2017). Lake volume and groundwater storage variations in Tibetan Plateau's endorheic basin. *Geophysical Research Letters*, 44(11), 5550–5560.

Zhang, G., Yao, T., Xie, H., Zhang, K., & Zhu, F. (2014). Lakes' state and abundance across the Tibetan Plateau. *Chinese Science Bulletin*, *59*(24), 3010–3021.

Zhang, M., & Liu, X. (2020). Climate changes in the Qaidam Basin in NW China over the past 40 kyr. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *551*, 109679.

Zhu, J., Song, C., Wang, J., & Ke, L. (2020). China's inland water dynamics: The significance of water body types. *Proceedings of the National Academy of Sciences*, 117(25), 13876–13878.