scientific **data**

Check for updates

Ice phenology dataset reconstructed from remote sensing and modelling for lakes over the Tibetan Plateau

Yanhong Wu[1,2], Linan Guo[2,3], Bing Zhang [1,2,4 ✉], Hongxing Zheng [5], Lanxin Fan[1,2,4], Haojing Chi[1,2,4], Junsheng Li[1,2,4] & Shenglei Wang[1,2]

The Tibetan Plateau (TP) is a region sensitive to global climate change and has been experiencing substantial environmental changes in the past decades. Lake ice phenology (LIP) is a perceptible indicator reflecting changes of lake thermodynamics in response to global warming. Lake ice phenology over the Tibetan Plateau is however rarely observed and recorded. This research presents a dataset containing 39-year (1978–2016) lake ice phenology data of 132 lakes (each with area >40 km²) over the Tibetan Plateau by combining the strengths of both remote sensing (MOD11A2, MOD10A1) and numerical modelling (air2water). Data validation shows that the ice phenology data derived by our method is highly consistent with that based on existing approaches (with $R^2 > 0.75$ for all phenology index and RMSE < 5d). The dataset is valuable to investigate the lake-atmosphere interactions and long-term hydrothermal change of lakes across the Tibetan Plateau.

Background & Summary

The formation and duration of ice cover for lakes in the Cryosphere plays an important role in thermodynamics of lakes. It interacts with lake water temperature and shapes aquatic ecosystems as well. Lake ice phenology (LIP, e.g., dates of freezing-up and breaking-up) are closely related to temperature variation[1,2]. Change of lake ice phenology in the Cryosphere is considered as one of the most reliable pieces of evidence reflecting global warming[3–5]. It could subsequently lead to substantial alteration in biogeochemical processes (e.g., influencing growth of plankton and causing anoxia in deep water[6]). Research efforts on investigating changes of lake ice phenology are mainly based on ground observation, satellite observation or modelling[7–10]. Most of the research, however, are for lakes locating in North America or Northern Europe as there are relatively abundant in ground observations[10].

Lake ice phenology observations are not always widely available particularly for regions like the Tibetan Plateau (TP), where ground-based observation is challenging and costly. Observation from space therefore has been becoming more attractive in investigating the lake ice phenology. For instance, with stratified random sample from Sentinel-2 and PlanetScope, Pickens *et al.* (2021) reported that 41% of inland water area around the globe were seasonally frozen[11]. Though with uncertainties[12] from both data sources[13–15] and extraction methods, various satellite-based observations have shown their merits in tracking the freezing-thawing cycle of lake ice cover[16]. The most commonly used satellite data includes that from the optical sensors like Advanced Very High Resolution Radiometer (AVHRR)[17] and Moderate-resolution Imaging Spectroradiometer (MODIS)[18], the passive microwave sensors like Scanning Multichannel Microwave Radiometer (SMMR)[3], the Advanced Microwave Scanning Radiometer for EOS (AMSR-E)[19] and Special Sensor Microwave/Imager (SSM/I)[20]. Active microwave also shows capabilities in monitoring lake ice status based on sensors like European Remote Sensing Satellite (ERS)-1/2 synthetic aperture radar (SAR)[21] and Radarsat-1/2 SAR[22], but the narrow swath width and

[1]International Research Centre of Big Data for Sustainable Development Goals, Beijing, 100094, China. [2]Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, 100094, China. [3]Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, 100101, China. [4]University of the Chinese Academy of Sciences, Beijing, 100049, China. [5]CSIRO Land and Water, Canberra, ACT 2601, Australia. ✉e-mail: zb@radi.ac.cn

1
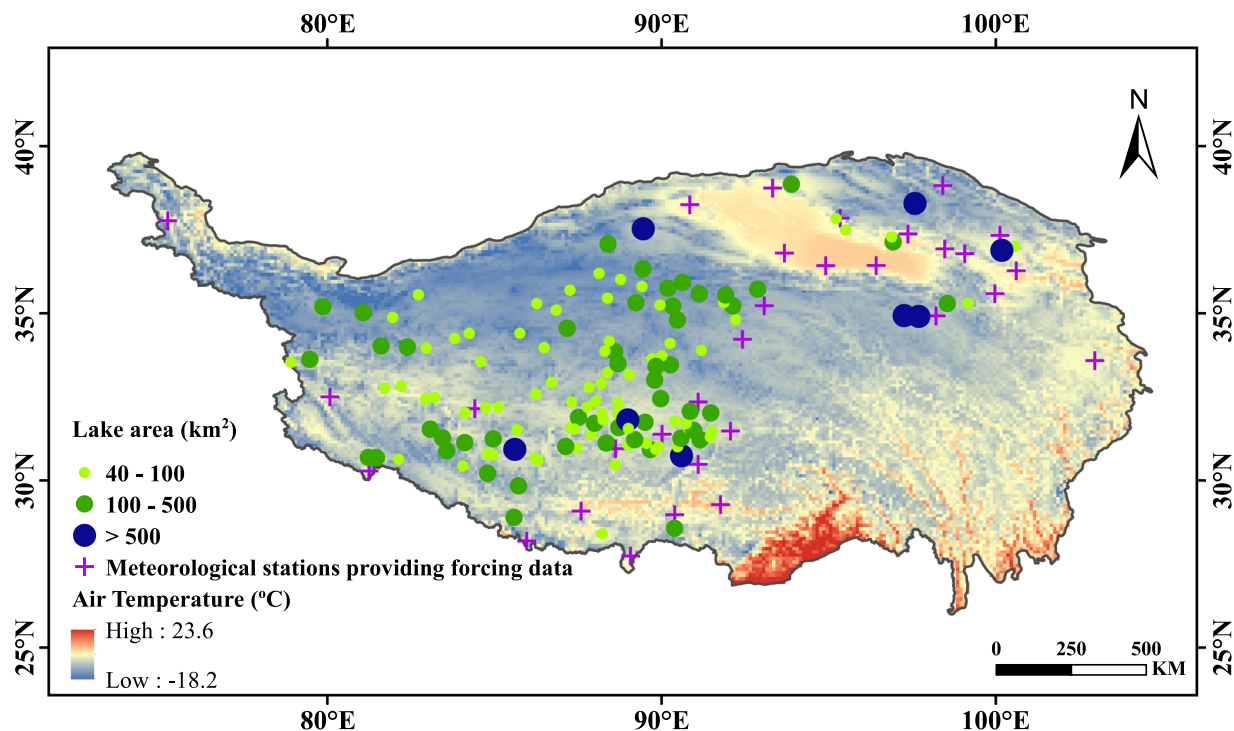
Fig. 1 Locations of studied lakes in the Tibetan Plateau.

the relatively low temporal resolution limited the application of the technology to monitor lake ice phenology at a daily scale and for a large area. It is also noted that the satellite-based observations of lake ice phenology mostly start from the late 1990s and could be with considerable data gaps. For example, the currently available dataset for lake ice phenology in the TP is that derived from microwave brightness temperature measured by AMSR-E, the Advanced Microwave Scanning Radiometer 2 (AMSR2) and Micro-Wave Radiation Imager (MWRI). The dataset however is limited to 51 lakes on the TP covering the period from 2002 to 2015[23].

In the literature of limnology, numerous research efforts also have been invested on developing mathematical models to reconstruct lake ice phenology for the historical period or to predict the response of lake ice phenology to a warming climate[5,24]. The model-based approach goes beyond reproducing data on lake ice phenology and can be used to quantify the response of lake ice phenology to climate change or climate variation. The models may cover different level of details on lake thermodynamics, subject to their research focuses[25–29]. It is recognized that a well simplified process-based model (like air2water[30,31]) could perform as good as a more complicated model (like MINLAKE[26], LIMNOS[27], HIGHTSI[28]) if the research interest is mainly on the timing of freezing and thawing of lake ice. The simplified model could be more competitive and applicable for regions with little *in-situ* observations (like the TP) as it requires much less data for running and calibration.

The dataset herein provides complete, consistent and continuous time series (1978–2017) of reconstructed ice phenology for 132 lakes in the TP (Fig. 1) by combing the strengths of remote sensing and mathematical modelling forced by meteorological data. The lake surface water temperature (LSWT) derived from MODIS products (MOD11A2)[32] is applied to calibrate and validate the numerical model in reproducing the consecutive daily surface water temperature for the studied period, based on which algorithm has been developed to determine the ice phenology indices (including dates of freezing-up and breaking-up and durations). Lake ice phenology derived directly from satellite observations in this research and by other researchers is used to validate the reconstructed LIP[33].

The reconstructed LIP represents a unique dataset to investigate the long-term trends of lake ice phenology in the TP under a warming climate. The dataset can also be used for a variety of applications related to climate change, limnology, hydrology, and aquatic ecology. It is conductive to assess the impacts of climate warming on dynamics of water/heat budget and aquatic biota in lakes across the TP.

Methods

Figure 2 shows the general framework of our study in determining the lake ice phenology for the 132 largest lakes across the Tibetan Plateau. The satellite-based lake surface temperature is used to calibrate the modified air2water model. The air temperature data from meteorological station is the solely required input data to drive the model. The temperature threshold of lake ice phenology is determined with the aid of MODIS snow cover product. Details of the procedures and methods are described in the following sections.

Data. The modified air2water model requires solely air temperature as input. For the TP, the best publicly accessible ground-based observation of daily air temperature is that from the dataset of daily climate
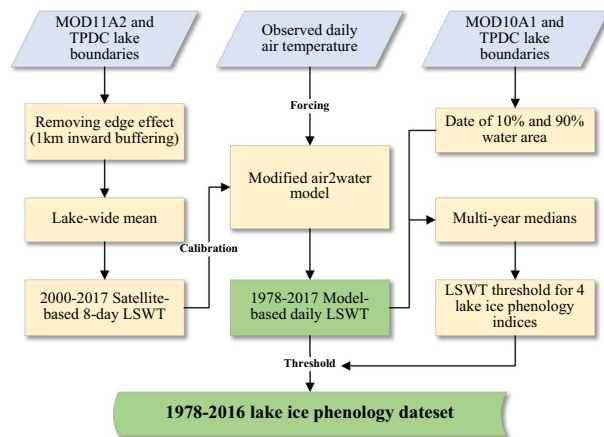
Fig. 2 Flowchart in producing the lake ice phenology (LIP) dataset.

data from Chinese surface stations (V3.0) provided by China Meteorological Data Service Centre of National Meteorological Information Centre (http://data.cma.cn/data/). The daily air temperature from the dataset is only available at 31 meteorological stations located in the TP. The spread and density of the meteorological stations are big constraints for earth system research in the region. We noticed that quite a few research efforts have been invested to provide reliable reanalysis climate datasets for the TP (e.g., China meteorological forcing dataset). The public available climate reanalysis datasets however are found not necessarily of high good quality for our simulation when we compare them against the observations from the meteorological stations. Hence, we use the air temperature interpolated from the nearest station with altitude adjustment (by a lapse rate of 0.65 °C 100 m^{-1})[34] to provide air temperature input to drive the modified air2water model.

The direct output of the model is lake surface water temperature. Hence, the model is firstly calibrated and validated against the satellite-based observation of lake surface water temperature deriving from the MOD11A2[32] (i.e., Terra product of Land Surface Temperature/Emissivity 8-day L3 Global 1 km). The surface water temperature of a specific lake is the lake-wide mean temperature based on all pixels within the lake. Boundaries of the lakes (shapefiles included in the dataset) are mainly from National Tibetan Plateau Data Center (TPDC, http://data.tpdc.ac.cn) and are cross-checked by maps from HydroLAKES (http://www.hydrosheds.org/)[35] and Google Earth. The lake-wide MODIS-based surface water temperature is the mean of daytime and night-time LST in the MOD11A2. All the data processing procedures are implemented via Google Earth Engine (GEE)[36].

Prior to determine the lake ice phenology given the modelled lake surface temperature, temperature thresholds should be defined, which is done herein with the aid of remotely sensed snow cover product, i.e., MOD10A1[37]. The MOD10A1 dataset is a snow cover product with a spatial resolution of 500 m containing daily, gridded snow cover and albedo derived from radiance data acquired by MODIS on board the Terra satellite. Snow cover is identified using the Normalized Difference Snow Index (NDSI) and a series of screens designed to alleviate errors and flag uncertain snow cover detections[37].

Lake surface temperature modeling. We used a slightly modified air2water model to simulate the daily lake surface temperature for the period 1978–2017. The air2water[27,28] is a semi-physical model designed to simulate lake surface temperature principally based on lake surface energy balance expressed as:

$$\rho c_p V_s \frac{dT_w}{dt} = A H_{net}$$

(1)

where T_w is water surface temperature at time t (day step in this study), H_{net} is the net heat flux per unit surface, A is the surface area of the lake, V_s is the volume of water involved in the heat exchange with the atmosphere, ρ and c_p is the density and specific heat capacity of water. The model considers all contributions in the energy balance (H_{net}) as a function of air temperature or is included in the parameters of the model. The original air2water model has simplified the lake energy balance by introducing eight calibratable parameters a_{1-8} with air temperature as the only required input, which can be expressed as:

$$\frac{dT_w}{dt} = \frac{1}{\delta}\left\{a_1 + a_2 T_a - a_3 T_w + a_5 \cos\left[2\pi\left(\frac{t}{t_y} - a_6\right)\right]\right\}$$

(2)
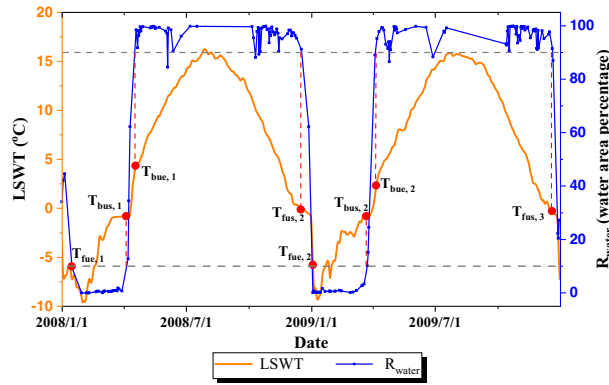
Fig. 3 Determination of lake surface water temperature (LSWT) thresholds for identifying the lake ice phenology indicators. $T_{FUS,i}$ $T_{FUE,i}$, $T_{BUS,i}$ and $T_{BUE,i}$ are the temperature threshold candidates derived for each specific year i ($=1,2,3$). R_{water} is the percentage of open water area within a lake. Data of this figure is from the Qinghai Lake for the period 2008–2009.

$$\delta = \begin{cases} exp\left(-\dfrac{T_w - T_h}{a_4}\right), & for \quad T_w \geq T_h \\ exp\left(-\dfrac{T_h - T_w}{a_7}\right) + exp\left(-\dfrac{T_w}{a_8}\right), & for \quad T_w \leq T_h \end{cases} \tag{3}$$

where, T_a is air temperature, t_y is days of a year. $\delta = D/D_r$ is normalized well-mixed depth, where D is the depth of well-mixed surface layer (the epilimnion thickness), D_r is the maximum thickness. T_h is the deep water temperature assumed to be 4 °C.

The model has been tested efficient in different regions[38,39]. However, the original air2water is limited to simulate the surface temperature of open water, which results in difficulties for applications in lakes with long ice cover duration. For this sake, we assume that when the lake is completely covered by ice, the heat exchange between air and water is blocked and surface energy balance would be replaced by land surface energy balance without mixing depth in water[36], which indicates $\delta = 1$ and Eq. (2) be rewritten as:

$$\frac{dT_i}{dt} = a_9 + a_{10}T_a - a_{11}T_w + a_{12}\cos\left[2\pi\left(\frac{t}{t_y} - a_{13}\right)\right] \tag{4}$$

where T_i is the ice surface temperature, a_{9-13} have similar physical significance to $a_{1,2,3,5,6}$. To represent the shifts of lake between open water and ice-covered, two additional parameters a_{14} and a_{15} are introduced to determine the states of the lake, which the lake surface water temperature (LSWT) can be expressed as:

$$LSWT = \begin{cases} T_w, & for \; T_L \geq a_{15} \\ T_i, & for \; T_L \leq a_{14} \\ (1 - K_{ice})T_w + K_{ice}T_i, & for \; a_{15} > T_L > a_{14} \end{cases} \tag{5}$$

where we set $K_{ice} = \sqrt{(a_{15} - LSWT)/(a_{15} - a_{14})}$ as the proportion of ice on the surface of the lake. Same as the original model, the model equations were solved numerically by using the Crank-Nicolson numerical scheme at daily time step.

The model is calibrated against the LSWT derived from MOD11A2 using the PSO (Particle Swarm Optimization) approach[31,40]. The objective function of model calibration is the root mean square error (RMSE). To ensure the rationality of the calibrated model parameters, a physically consistent priori range are assigned to each parameter[31] and D_r is bounded by the average depth of the lake obtained from Records of Lakes in China[41] and HydroLAKES[35]. The calibration period is 2000–2012, while 2013–2017 is for validation.

Temperature thresholds for ice phenology. The lake surface temperature output from air2water model is used to derive from lake ice phenology given prior defined temperature thresholds. The temperature thresholds are defined with the aid of the snow cover products (MOD10A1), where the MOD10A1 images with cloud proportion over the lake exceeding 20% are neglected. For each lake, the ratio of "inland water" (R_{water}) to total lake area is firstly calculated as: $R_{water} = N_{water}/(N_{water} + N_{ice})$, where N_{water} is the pixel number of "inland water", while N_{ice} is the pixel number of "ice". The lake is assumed to be open when $R_{water} > 90\%$ and frozen when $R_{water} < 10\%$[42,43]. As shown in Fig. 3 The temperature thresholds on determining lake ice phenology are then defined as the temperature of the date when water area percentage (R_{water}) drops down below 10% or rises above 90%. The multi-year medians of the temperature thresholds are finally considered as the overall thresholds to

File name	Description of each column
lake_info.csv (9 columns)	Lake ID, Name, Longitude, Latitude, Area, Elevation, Depth, Volume, Glacier.
TP_LIP_Lakes132.shp	Lake boundaries collected from National Tibetan Plateau Data Center.
XXXXX_LIP.csv (7 columns)	Year
	FUS (i.e., starting date of freezing-up).
	FUE (i.e., ending date of freezing-up).
	BUS (i.e., starting date of breaking-up).
	BUE (i.e., ending date of breaking-up).
	CID (i.e., fully ice-covering duration).
	IceD (i.e., ice-covered duration).

Table 1. Files and formats of the dataset.

determine the starting and ending dates of freezing-up (T_{FUS} and T_{FUE}) and breaking-up of lake ice phenology (T_{BUS} and T_{BUE}) separately for each lake.

Lake ice phenology determination. The long-term lake ice phenology for the period 1978–2016 is derived from the simulated lake surface temperature given the temperature thresholds. The lake ice phenology considered herein consists of 6 indicators reflecting the timing or duration of different freezing/thawing phases. In addition to starting and ending dates of freezing-up (FUS and FUE) and breaking-up (BUS and BUE), the ice-covered duration (IceD, in days) and the fully ice-covered duration (CID, in days) are also presented in the dataset. The indicators are defined as:

$$\begin{cases} FUS = min\,(t)\,s.\,t.\ \ LSWT_t < T_{FUS}\ and\ LSWT_{t-1} > T_{FUS} \\ FUE = max\,(t)\,s.\,t.\ \ LSWT_t < T_{FUE}\ and\ LSWT_{t-1} > T_{FUE} \\ BUS = min\,(t),\,s.\,t.\ \ LSWT_t > T_{BUS}\ and\ LSWT_{t-1} < T_{BUS} \\ BUE = max\,(t),\,s.\,t.\ \ LSWT_t > T_{BUE}\ and\ LSWT_{t-1} < T_{BUE} \\ \qquad\qquad IceD = BUE - FUS \\ \qquad\qquad CID = BUS - FUE \end{cases} \qquad (6)$$

Where $LSWT_t$ is the simulated lake surface temperature at date t. The date t is expressed as the day of a calendar year (starting at January 1st). Since the lake ice phenology could spans two consecutive years, we assign herein the value of t to range from 1 to 730, where $t>365$ represents the date in the next consecutive calendar year.

Data Records

The lake ice phenology data across Tibetan Plateau is archived and openly accessible via Figshare[44] via the link: https://doi.org/10.6084/m9.figshare.18852338.v2. Description about the 132 lakes is shown in the lake_info.csv file, including lake ID, name, latitude, longitude, lake area, water level, depth, water volume, freezing property, and glacial replenishment. The information about the lakes is mainly collected from Records of Lakes in China[41], HydroLAKES[35], and Map of Glaciers and Lakes on the Tibetan Plateau and Adjoining Regions. The Phenology folder in the dataset consists of 132 CSV files. Each CSV file is named according to the lake ID presented in lake_info.csv. The columns of each CSV file are listed in Table 1.

Technical Validation

Quality control. The reconstructed LIP dataset is produced with strict quality control throughout the procedures. The boundaries of the lakes are examined via professional GIS tools (both ArcGIS and Google Map) to ensure the consistency of their geographic locations, and topography. Lake-wide mean surface temperature instead of that of lake centroid is derived from MOD11A2 for model calibration. The 132 lakes presented in the dataset are all with good model performance (Nash-Sutcliffe efficiency coefficient[45] above 0.6) when the simulated lake surface temperature is compared against the MODIS-based one. Suspicious outliers (i.e., abnormal extremely high or low values) are eliminated from satellite-based lake surface temperature and snow product in providing reliable data for robust model calibration and validation. The MOD10A1 is used as an aid in determining the temperature thresholds of LIP. It is noted that the MOD10A1 is not always of good quality in continuously retrieving the ice-covered ratio of a lake due to accidental misclassification and effects of cloud cover. To alleviate the uncertainties induced by the optical RS data, we use open water ratio (see Methods) instead of directly using ice-covered ratio of a lake in determining the temperature threshold, and the medians of the temperature thresholds are finally accepted for robust lake ice phenology identification. The reconstructed LIP dataset is validated by comparing against the published LIP data as described below.

Validation of the reconstructed LIP dataset. There is little *in-situ* observation of lake ice phenology in the TP available for a complete one-by-one validation of our reconstructed LIP dataset. We herein validate our dataset in two different ways. Firstly, we validate the effectiveness of our approach by comparing the reconstructed LIP against *in-situ* observations of two lakes in the TP[46,47] and one lake in Canda[48]. The comparison is shown in Supplementary Fig. S1 suggesting that the estimated ice phenology overall is comparable with the *in-situ* observations, where the root mean squared error (RMSE) of the estimated FUE in Nam Co is around 7

5
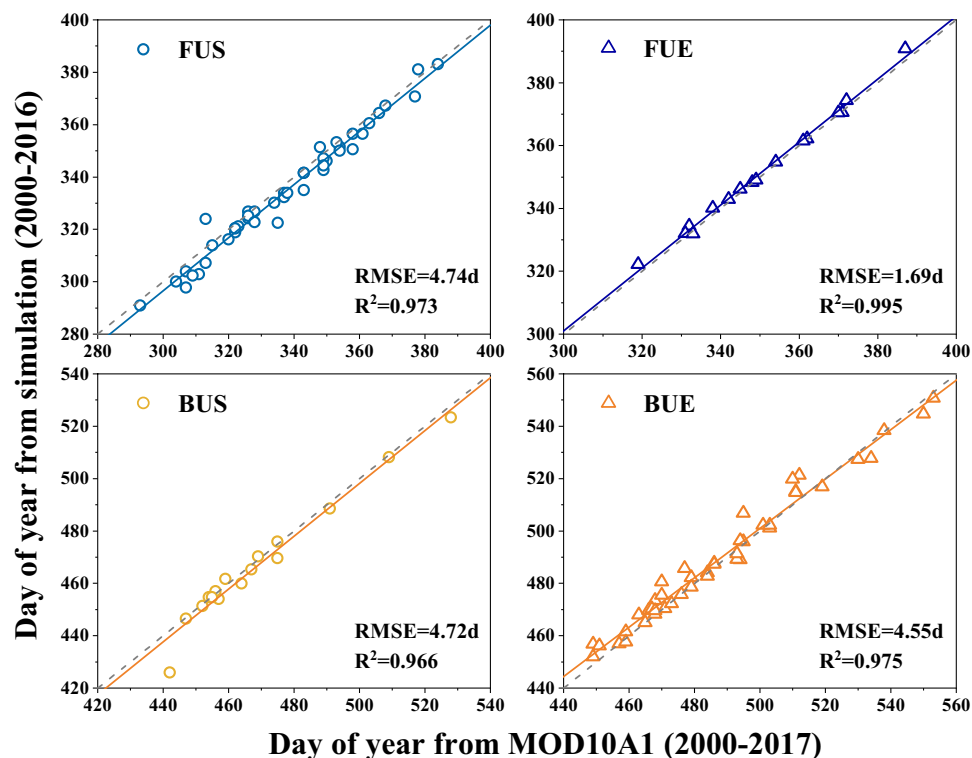
Fig. 4 Comparison of long-term mean lake ice phenology derived in this research against that from Cai *et al.*[32], where FUS, FUE, BUS, BUE represents starting date of freezing-up, ending date of freezing-up, starting date of breaking-up, and ending date of breaking-up respectively. The unit of FUS, FUE, BUS and BUE is day of year with the number larger than 365 indicates dates in the next calendar year.

days, RMSE of the estimated BUE in Lesser Slave Lake is near 9 days and RMSE of the estimated in Qinghai Lake is around 12 days. It should be noted that the bias could be due to both the uncertainties in the model and in the climate input to the model as well.

We further compare the reconstructed LIP against two existing datasets developed based on remotely sensed data, which were published by Cai *et al.*[32] and Qiu *et al.*[23] respectively. The dataset published by Cai *et al.*[32] is retrieved from MOD10A1 and Landsat data, providing mean lake ice phenology of 44 TP lakes for the period 2000–2017. Figure 4 shows the comparison of the mean annual LIP from our reconstructed dataset against that from Cai *et al.* for the 44 lakes. It can be seen that the LIP dataset produced by our approach is in high agreement with that solely based on MOD10A1, where the R^2 for all the four phenology indices (i.e., FUS, FUE, BUS and BUE) are higher than 0.96 and the root mean squared errors (i.e., RMSE) are all within 5 days.

The reconstructed LIP time series is also compared against the dataset published by Qiu *et al.*[23], which was derived from microwave brightness temperature measured by AMSR-E, AMSR2 and MWRI. The dataset from Qiu *et al.* provides LIP of 51 lakes in High Asia for the period from 2002 to 2016, among which there are 35 lakes identical to the lakes in our dataset. For the 35 lakes in the period 2002–2015 (sample size $= 35 \times 14$), the statistically significant correlations are found between the two datasets for all the six phenology indices with R^2 all above 0.75 (Fig. 5). The discrepancy between the two datasets is found to be the lowest in FUS (RMSE $= 8.6$ days) and to be the highest in BUS (RMSE $= 15$ days). The relative higher discrepancy of the two datasets in both BUS and FUE results at noticeable bias of the CID, which by definition equals to number of days between FUE and BUS. The uncertainties of both our reconstructed LIP dataset and those built on satellite-observations however need to be further quantified in the future particularly where *in-situ* observations becoming available.

Spatial distribution of LIP across the TP. Figure 6 shows the long-term mean spatial distribution of the LIP indices based on the reconstructed dataset. Overall, the results suggest that the phenology is close related to latitude (Fig. S2 in Supplementary). For lake in higher latitude, it tends to freeze-up earlier but break-up later hence with longer ice-covered duration. Overall, from south to north, FUS ranges from October 15th to January 15th while FUE varies from November 9th to January 21st. The gradients of FUS and FUE with respect to latitude are around 6.7 days/degree and 5.7 days/degree respectively (Fig. S2 in Supplementary). For ice-breaking dates, from south to north, BUS ranges between 27th January and 22nd June, while BUE varies between 2nd March and 5th July. The latitude gradients of BUS and BUE are about 11.8 days/degree and 8.5 days/degree respectively. From south to north, ice-covered duration ranges from 15 days to 215 days (CID) or from 85 to 255 days (IceD). The latitude gradients of CID and IceD are 17.5 days/degree and 15.2 days/degree respectively. Lake ice phenology is also affected by altitude (Fig. S3 in Supplementary). Statistically significant correlations are found between LIP and altitude particularly for lakes locating between 4500 to 5000 m (account for 70% of lakes in this dataset).
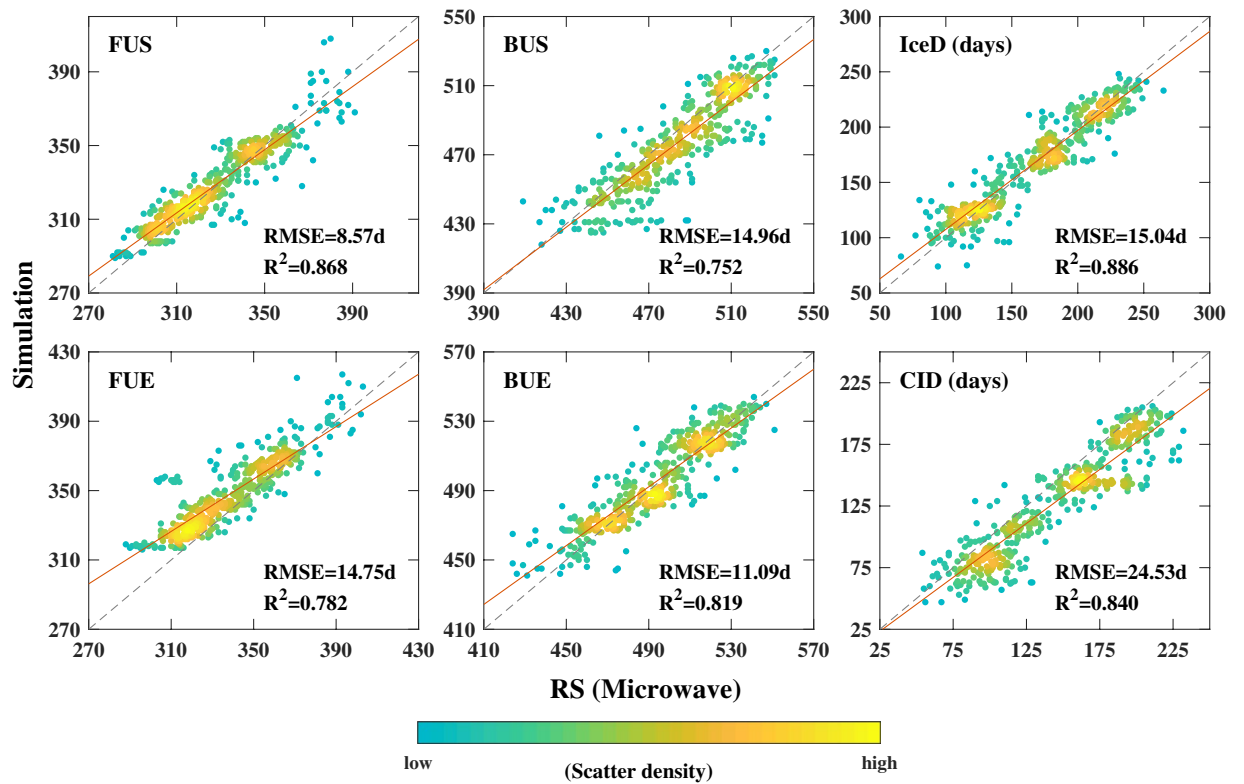
Fig. 5 Comparison of lake ice phenology derived in this research against that from Qiu *et al.*[23], where IceD and CID represents ice-covered duration and fully ice-covering duration respectively with unit day. The meaning and unit of FUS, FUE, BUS and BUE are the same to that in Fig. 4.
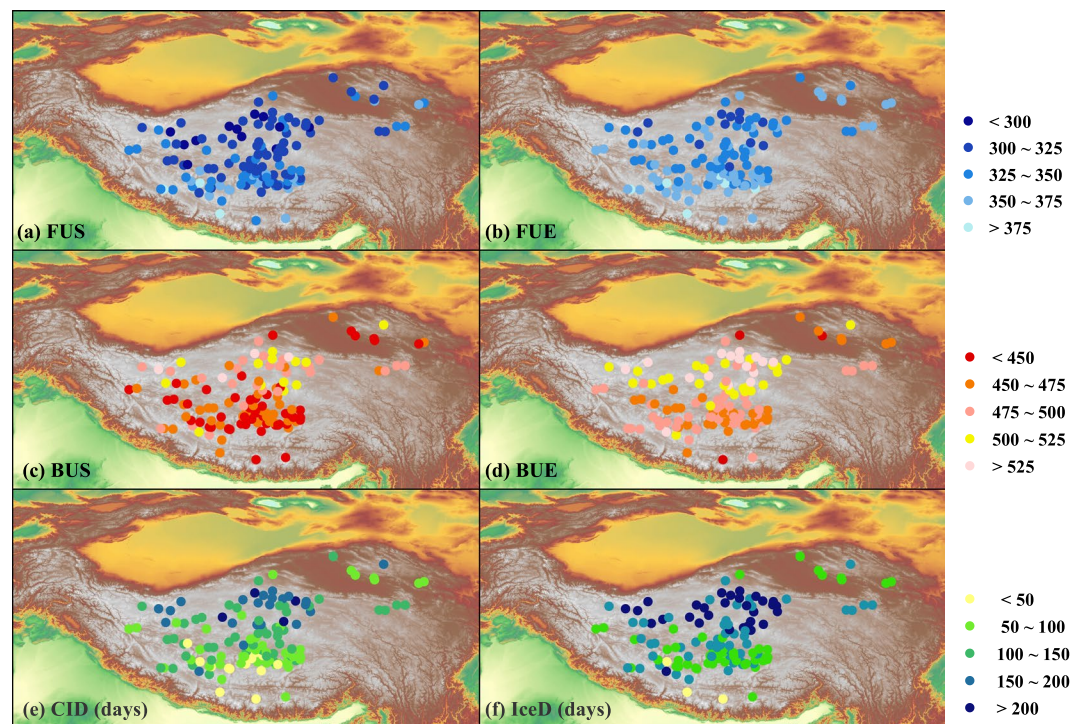
Fig. 6 Spatial distribution of lake ice phenology over the TP. The meaning of FUS, FUE, BUS, BUE, CID and IceD are the same to that in Fig. 4 and Fig. 5.

Overall, FUS and FUE are in negative correlations with altitude, while BUS, BUE, IceD and CID are in positive correlations with altitude. It suggests that for lakes with higher altitude, earlier the FUS and FUE, later BUS and BUE and longer IceD and CID can all be expected. The altitude gradient of FUS and FUE is 9.4 days/hm and 5.7 days/hm respectively, while it is around −13 days/hm and −11.7 days/hm for BUS and BUE respectively. The altitude gradient of CID and IceD is estimated to be 18.9 days/hm and 21.4 days/hm (Fig. S3 in Supplementary).

Code availability

Codes in developing the lake ice phenology dataset are freely available. Codes of the modified air2water model for lake surface water temperature simulation are available at: https://github.com/Siyu1993/ModifiedModel. Codes for lake ice phenology identification are written using Matlab and are available at: https://github.com/Siyu1993/MatlabCode.git.

References

1. Livingstone, D. M. Break-up dates of Alpine lakes as proxy data for local and regional mean surface air temperatures. *Clim. Change* **37**, 407–439, https://doi.org/10.1023/a:1005371925924 (1997).
2. Tian, B. S. *et al.* Characterizing C-band backscattering from thermokarst lake ice on the Qinghai-Tibet Plateau. *ISPRS-J. Photogramm. Remote Sens.* **104**, 63–76, https://doi.org/10.1016/j.isprsjprs.2015.02.014 (2015).
3. Kouraev, A. V. *et al.* Observations of Lake Baikal ice from satellite altimetry and radiometry. *Remote Sens. Environ.* **108**, 240–253, https://doi.org/10.1016/j.rse.2006.11.010 (2007).
4. Marszelewski, W. & Skowron, R. Ice cover as an indicator of winter air temperature changes: case study of the Polish Lowland lakes. *Hydrol. Sci. J.-J. Sci. Hydrol.* **51**, 336–349, https://doi.org/10.1623/hysj.51.2.336 (2006).
5. Walsh, S. E. *et al.* Global patterns of lake ice phenology and climate: Model simulations and observations. *J. Geophys. Res.-Atmos.* **103**, 28825–28837, https://doi.org/10.1029/98jd02275 (1998).
6. Caramatti, I., Peeters, F., Hamilton, D. & Hofmann, H. Modelling inter-annual and spatial variability of ice cover in a temperate lake with complex morphology. *Hydrol. Process.* **34**, 691–704, https://doi.org/10.1002/hyp.13618 (2020).
7. Bengtsson, L. Ice-covered lakes: environment and climate—required research. *Hydrol. Process.* **25**, 2767–2769 (2011).
8. Wang, X. *et al.* High-Resolution Mapping of ice cover changes in over 33,000 lakes across the north temperate zone. *Geophys. Res. Lett.* **48**, https://doi.org/10.1029/2021gl095614 (2021).
9. Zhang, S., Pavelsky, T. M., Arp, C. D. & Yang, X. Remote sensing of lake ice phenology in Alaska. *Environ. Res. Lett.* **16**, https://doi.org/10.1088/1748-9326/abf965 (2021).
10. Magnuson, J. J. *et al.* Historical trends in lake and river ice cover in the Northern Hemisphere. *Science* **289**, 1743–1746, https://doi.org/10.1126/science.289.5485.1743 (2000).
11. Pickens, A. H. *et al.* Global seasonal dynamics of inland open water and ice. *Remote Sens Environ.* **272**, https://doi.org/10.1016/j.rse.2022.112963 (2022).
12. Guo, L. *et al.* Uncertainty and variation of remotely sensed lake ice phenology across the Tibetan Plateau. *Remote Sens.* **10**, https://doi.org/10.3390/rs10101534 (2018).
13. Tolnaia, M., Nagy, J. G. & Bako, G. Spatiotemporal distribution of Landsat imagery of Europe using cloud cover-weighted metadata. *J. Maps* **12**, 1084–1088, https://doi.org/10.1080/17445647.2015.1125308 (2016).
14. Dörnhöfer, K. & Oppelt, N. Remote sensing for lake research and monitoring - Recent advances. *Ecol. Indic.* **64**, 105–122, https://doi.org/10.1016/j.ecolind.2015.12.009 (2016).
15. Song, C., Huang, B., Ke, L. & Richards, K. S. Remote sensing of alpine lake water environment changes on the Tibetan Plateau and surroundings: A review. *ISPRS-J. Photogramm. Remote Sens.* **92**, 26–37 (2014).
16. Nonaka, T., Matsunaga, T. & Hoyano, A. Estimating ice breakup dates on Eurasian lakes using water temperature trends and threshold surface temperatures derived from MODIS data. *Int. J. Remote Sens.* **28**, 2163–2179, https://doi.org/10.1080/01431160500391957 (2007).
17. Latifovic, R. & Pouliot, D. Analysis of climate change impacts on lake ice phenology in Canada using the historical satellite data record. *Remote Sens. Environ.* **106**, 492–507, https://doi.org/10.1016/j.rse.2006.09.015 (2007).
18. Cai, Y., Ke, C. Q. & Duan, Z. Monitoring ice variations in Qinghai Lake from 1979 to 2016 using passive microwave remote sensing data. *Sci. Total Environ.* **s 607–608**, 120–131 (2017).
19. Ruan, Y., Qiu, Y., Yu, X., Guo, H. & Cheng, B. Passive microwave remote sensing of lake freeze-thaw over High Mountain Asia. *IEEE international Geoscience and Remote Sensing Symposium.* 2818–2821, https://doi.org/10.1109/IGARSS.2016.7729728 (2016).
20. Ke, C. Q., Tao, A. Q. & Jin, X. Variability in the ice phenology of Nam Co Lake in central Tibet from scanning multichannel microwave radiometer and special sensor microwave/imager: 1978 to 2013. *J. Appl. Remote Sens.* **7**, 12, https://doi.org/10.1117/1.jrs.7.073477 (2013).
21. Jeffries, M. O., Morris, K., Weeks, W. F. & Wakabayashi, H. Structural and stratigraphic features and ers-1 synthetic-aperture radar backscatter characteristics of ice growing on shallow lakes in NW Alaska, winter 1991-1992. *J. Geophys. Res.-Oceans* **99**, 22459–22471, https://doi.org/10.1029/94jc01479 (1994).
22. Duguay, C. R., Pultz, T. J., Lafleur, P. M. & Drai, D. RADARSAT backscatter characteristics of ice growing on shallow sub-Arctic lakes, Churchill, Manitoba, Canada. *Hydrol. Process.* **16**, 1631–1644, https://doi.org/10.1002/hyp.1026 (2002).
23. Qiu, Y. *et al.* A dataset of microwave brightness temperature and freeze-thaw for medium-to-large lakes over the High Asia region (2002-2016). *Science Data Bank.* https://doi.org/10.11922/sciencedb.374 (2017).
24. Bernhardt, J., Engelhardt, C., Kirillin, G. & Matschullat, J. Lake ice phenology in Berlin-Brandenburg from 1947–2007: observations and model hindcasts. *Clim. Change* **112**, 791–817 (2012).
25. Duguay, C. R. *et al.* Ice-cover variability on shallow lakes at high latitudes: model simulations and observations. *Hydrol. Process.* **17**, 3465–3483, https://doi.org/10.1002/hyp.1394 (2003).
26. Gu, R. & Stefan, H. G. Year-round temperature simulation of cold climate lakes. *Cold Reg. Sci. Tech.* **18**, 147–160, https://doi.org/10.1016/0165-232x(90)90004-g (1990).
27. Vavrus, S. J., Wynne, R. H. & Foley, J. A. Measuring the sensitivity of southern Wisconsin lake ice to climate variations and lake depth using a numerical model. *Limnol. Oceanogr.* **41**, 822–831, https://doi.org/10.4319/lo.1996.41.5.0822 (1996).
28. Launiainen, J. & Cheng, B. Modelling of ice thermodynamics in natural water bodies. *Cold Reg. Sci. Tech.* **27**, 153–178, https://doi.org/10.1016/s0165-232x(98)00009-3 (1998).
29. Mironov, D. V. Parameterization of lakes in numerical weather prediction: description of a lake model. Report No. 11, (ARPA Piemonte, 2008).
30. Piccolroaz, S., Toffolon, M. & Majone, B. A simple lumped model to convert air temperature into surface water temperature in lakes. *Hydrology and Earth System Sciences* **17**, 3323–3338, https://doi.org/10.5194/hess-17-3323-2013 (2013).

31. Piccolroaz, S. Prediction of lake surface temperature using the air2water model: guidelines, challenges, and future perspectives. *Advances in Oceanography and Limnology* **7**, 36–50, https://doi.org/10.4081/aiol.2016.5791 (2016).
32. Wan, Z., Hook, S. & Hulley, G. MOD11A2 MODIS/Terra land surface temperature/emissivity 8-day L3 global 1km SIN grid V006. *NASA EOSDIS Land Processes DAAC.* https://doi.org/10.5067/MODIS/MOD11A2.006 (2015).
33. Cai, Y. *et al*. Variations of lake ice phenology on the Tibetan Plateau from 2001 to 2017 based on MODIS data. *J. Geophys. Res.-Atmos.* **124**, 825–843, https://doi.org/10.1029/2018jd028993 (2019).
34. Hu, Q., Jiang, D. & Fan, G. Evaluation of CMIP5 models over the Qinghai-Tibetan Plateau. *Chin. J. Atmos. Sci* **38**, 924–938, https://doi.org/10.3878/j.issn.1006-9895 (2014).
35. Messager, M. L., Lehner, B., Grill, G., Nedeva, I. & Schmitt, O. Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nat. Commun.* **7**, 11, https://doi.org/10.1038/ncomms13603 (2016).
36. Guo, L. A. *et al*. An integrated dataset of daily lake surface water temperature over the Tibetan Plateau. *Earth Syst. Sci. Data* **14**, 3411–3422, https://doi.org/10.5194/essd-14-3411-2022 (2022).
37. Hall, D. K. & Riggs, G. A. MODIS/Terra snow cover daily L3 global 500m SIN grid, version 61. *NASA National Snow and Ice Data Center Distributed Active Archive Center.* https://doi.org/10.5067/MODIS/MOD10A1.061 (2021).
38. Prats, J. & Danis, P. A. An epilimnion and hypolimnion temperature model based on air temperature and lake characteristics. *Knowl. Manag. Aquat. Ecosyst.*, 24, https://doi.org/10.1051/kmae/2019001 (2019).
39. Toffolon, M. *et al*. Prediction of surface temperature in lakes with different morphology using air temperature. *Limnol. Oceanogr.* **59**, 2185–2202, https://doi.org/10.4319/lo.2014.59.6.2185 (2014).
40. Kennedy, J. & Eberhart, R. Particle swarm optimization. *IEEE Proceedings of ICNN'95-International Conference on Neural Networks.* 1942–1948 (1995).
41. Wang, S. M. & Dou, H. S. *Records of lakes in China*. (Beijing: Science Press, 1998).
42. Reed, B., Budde, M., Spencer, P. & Miller, A. E. Integration of MODIS-derived metrics to assess interannual variability in snowpack, lake ice, and NDVI in southwest Alaska. *Remote Sensing of Environment* **113**, 1443–1452, https://doi.org/10.1016/j.rse.2008.07.020 (2009).
43. Yao, X. *et al*. Spatial-temporal variations of lake ice phenology in the Hoh Xil region from 2000 to 2011. *Journal of Geographical Sciences* **26**, 70–82 (2016).
44. Guo, L. *et al*. Lake ice Phenology across Tibetan Plateau (TPLIP). *Figshare.* https://doi.org/10.6084/m9.figshare.18852338.v2 (2022).
45. Nash, J. E. & Sutcliffe, J. V. River flow forecasting through conceptual models part I — A discussion of principles - ScienceDirect. *Journal of Hydrology* **10**, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6 (1970).
46. Gou, P. *et al*. Lake ice phenology of Nam Co, Central Tibetan Plateau, China, derived from multiple MODIS data products. *Journal of Great Lakes Research* **43** (2017).
47. Zhang, G. Q. & Duan, S. Q. Lakes as sentinels of climate change on the Tibetan Plateau. *All Earth* **33**, 161–165, https://doi.org/10.1080/27669645.2021.2015870 (2021).
48. Benson, B., Magnuson, J., & Sharma, S. Global lake and river ice phenology database, version 1. Boulder, Colorado USA NSIDC: National Snow and Ice Data Center. https://doi.org/10.7265/N5W66HP8 (2000, updated 2020).

Acknowledgements

Author contributions

Y.W. and L.G. contributed equally to this work. Y.W., H.Z. and L.G. conceived the research. L.G., Y.W. and H.Z. developed the approaches and datasets. L.F., H.C., J.L. and S.W. collected basic data of lakes. L.F., H.C. checked the results. Y.W., L.G. and H.Z. wrote the original draft. B.Z. and H.Z. revised the draft.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41597-022-01863-9.

Correspondence and requests for materials should be addressed to B.Z.

Reprints and permissions information is available at www.nature.com/reprints.

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