nature geoscience

ARTICLES

https://doi.org/10.1038/s41561-021-00833-x

Check for updates

# Attribution of global lake systems change to anthropogenic forcing

Luke Grant[1]✉, Inne Vanderkelen[1], Lukas Gudmundsson[2], Zeli Tan[3], Marjorie Perroud[4], Victor M. Stepanenko[5,6], Andrey V. Debolskiy[5,6,7], Bram Droppers[8], Annette B. G. Janssen[8], R. Iestyn Woolway[9], Margarita Choulga[10], Gianpaolo Balsamo[10], Georgiy Kirillin[11], Jacob Schewe[12], Fang Zhao[12], Iliusi Vega del Valle[12], Malgorzata Golub[13], Don Pierson[13], Rafael Marcé[14,15], Sonia I. Seneviratne[2] and Wim Thiery[1,2]

**Lake ecosystems are jeopardized by the impacts of climate change on ice seasonality and water temperatures. Yet historical simulations have not been used to formally attribute changes in lake ice and temperature to anthropogenic drivers. In addition, future projections of these properties are limited to individual lakes or global simulations from single lake models. Here we uncover the human imprint on lakes worldwide using hindcasts and projections from five lake models. Reanalysed trends in lake temperature and ice cover in recent decades are extremely unlikely to be explained by pre-industrial climate variability alone. Ice-cover trends in reanalysis are consistent with lake model simulations under historical conditions, providing attribution of lake changes to anthropogenic climate change. Moreover, lake temperature, ice thickness and duration scale robustly with global mean air temperature across future climate scenarios (+0.9 °C $°C_{air}^{-1}$, −0.033 m $°C_{air}^{-1}$ and −9.7 d $°C_{air}^{-1}$, respectively). These impacts would profoundly alter the functioning of lake ecosystems and the services they provide.**

Lakes provide ecosystem services to local communities[1,2] and modulate local climates[3–7]. The seasonality of lake ice cover and lake temperatures are the foundations of the lake environment, controlling many lake processes[8,9]. In recent decades, lake temperatures have been rising, and seasonal ice cover has been declining on regional[10–12] and global scales[13,14]. Among other things, these changes alter lake stratification[15], impact lake ecosystem productivity[16] and disturb fisheries[17,18].

New historical simulations of lake ice cover and mixed-layer temperature from the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis v5 - Land (ERA5-Land) reanalysis[19]—lake model simulations forced by the observation-assimilated atmosphere of ERA5—provide a high-resolution outlook on lake changes in recent decades (Fig. 1a–c and Supplementary Fig. 1). From 1981–1990 until 2010–2019, these simulations reveal rapid changes; 130,472 lake grid cells worldwide have experienced two weeks of lake ice-cover loss, while on average lakes have lost nine days of ice cover. Likewise, global-scale reanalysed lake mixed-layer temperature shows substantial increases, with 64,382 lake grid cells warming more than 1.5 °C and a global annual average increase of 0.4 °C (Fig. 2e).

While observed and reanalysed changes in lake ice cover and lake temperatures are large, the possibility that they are due to natural climate variability has so far not been ruled out. They have also not been attributed to anthropogenic drivers using formal statistical approaches. Formally, 'detection'[20,21] of climate change impacts consists of showing that observed changes are inconsistent with natural variability by comparing them with simulated variability under human-free climate conditions. Upon successful detection, anthropogenic greenhouse gas emissions are a plausible candidate to explain ongoing changes in lakes, but this causal link must again be formally established. Such 'attribution'[20,21] to anthropogenic emissions is achieved by showing consistencies between observed changes and response patterns derived from historical climate impact simulations. Together, detection and attribution represent a cornerstone of assessments by the Intergovernmental Panel on Climate Change (IPCC)[22,23].

## Climate change detection and attribution

We investigate climate change detection and attribution in ERA5-Land reanalysed lake variables using two complementary approaches and simulations with five global-scale lake models forced by four global climate models (GCMs)[24] (Methods and Supplementary Note 1.5). The first approach[25–27] considers a distribution of rank correlations between the multimodel mean of lake simulations forced by GCMs under historical climate forcings (HIST) and a collection of individual pre-industrial control (PIC) lake simulations. This distribution of correlations, assumed to arise from pre-industrial climate variability, is compared with the single correlation between HIST and the reanalysed time series (ERA5L for reanalysis). Here, detection is inferred by rejecting the null hypothesis that reanalysed trends are consistent with the

[1]Department of Hydrology and Hydraulic Engineering, Vrije Universiteit Brussel, Brussels, Belgium. [2]ETH Zurich, Institute for Atmospheric and Climate Science, Zurich, Switzerland. [3]Pacific Northwest National Laboratory, Richland, WA, USA. [4]Institute for Environmental Sciences, University of Geneva, Geneva, Switzerland. [5]Lomonosov Moscow State University, Moscow, Russia. [6]Moscow Center for Fundamental and Applied Mathematics, Moscow, Russia. [7]Obukhov Institute for Atmospheric Physics, Russian Academy of Science, Moscow, Russia. [8]Water Systems and Global Change Group, Wageningen University & Research, Wageningen, the Netherlands. [9]European Space Agency Climate Office, ECSAT, Harwell Campus, Didcot, UK. [10]Research Department, European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, UK. [11]Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany. [12]Potsdam Institute for Climate Impact Research, Potsdam, Germany. [13]Department of Ecology and Genetics, Uppsala University, Uppsala, Sweden. [14]Catalan Institute for Water Research (ICRA), Girona, Spain. [15]University of Girona, Girona, Spain. ✉e-mail: luke.grant@vub.be

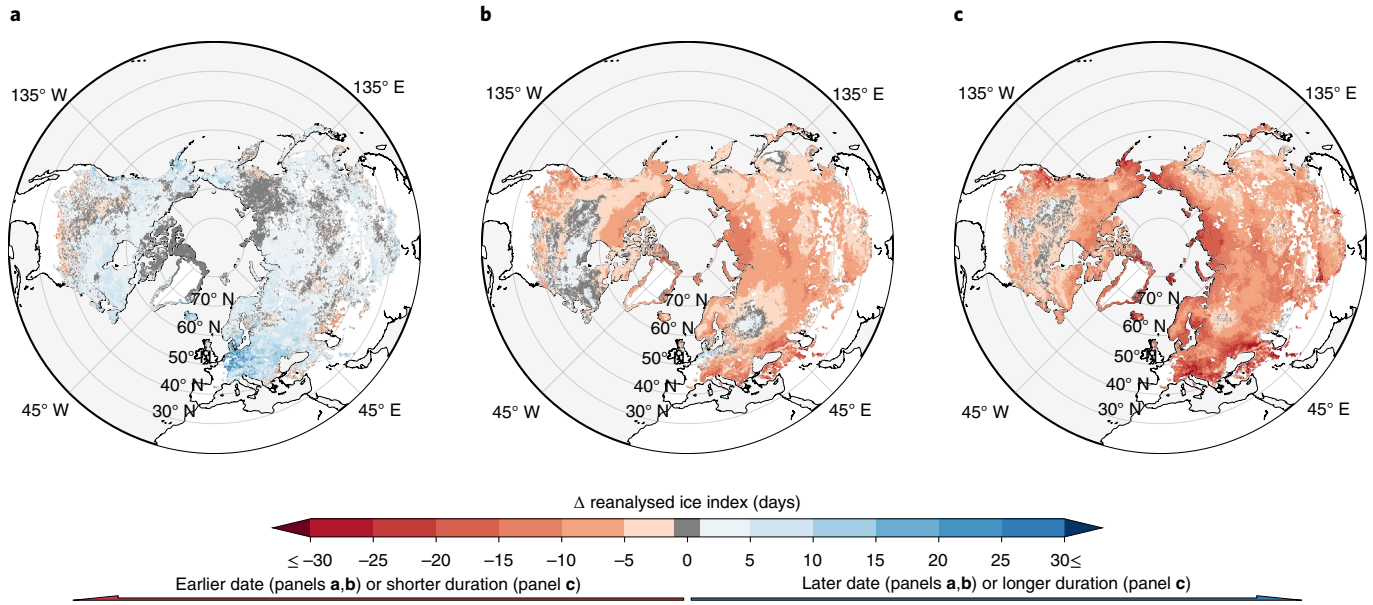


**Fig. 1 | Reanalysed historical lake ice changes. a–c**, Changes (Δ) in ice onset (**a**), ice break-up (**b**) and ice duration (**c**) in 40 years across baseline (1981–1990) and recent (2010–2019) periods as obtained from ERA5-Land.

distribution of correlations representative of pre-industrial climate variability (correlation-based approach[28]; Fig. 2a–d). The second approach employs regularized optimal fingerprinting (ROF)[20,29]. In this method, the slope parameters (henceforth referred to as scaling factors) that scale HIST to fit ERA5L in a total least squares (TLS) regression communicate detection when they are significantly different from 0 (that is, when the 95% confidence intervals of the scaling factors exclude 0). Attribution is achieved when scaling factors additionally overlap with unity (Methods).

Strict attribution to anthropogenic emissions requires both all-forcings historical and natural historical response patterns (including, for example, solar and volcanic influences but without anthropogenic emissions). Through rigorous statistics, the relative disparities or likeness of these experiments to a reference observed trend can confirm attribution. Our experimental framework does not include a natural historical climate scenario. This limits formal attribution to all combined historical forcings (Methods). However, in light of the dominant role of anthropogenic emissions relative to natural forcings in historical climate change[30], we argue that any attribution in this framework entails the imprint of human influence. In addition, out of necessity for a comprehensive spatial and temporal representation of lake trends, we substitute the use of observed records for lake ice cover and temperatures with observation-validated reanalysis in ERA5L (Supplementary Note 1.1).

For lake water temperature at 2 m depth (hereafter, lake temperature), the correlation-based approach shows a strong distinction between the correlation of ERA5L and HIST on the one hand and the distribution of correlation coefficients of PIC and HIST on the other hand (Fig. 2a). This implies that lake temperature reanalysis simulations for the recent past lie outside the typical variability of pre-industrial climate conditions and therefore cannot be explained by pre-industrial climate variability (99% confidence level). For ice onset, break-up and duration, correlations between ERA5L and HIST anomalies are again substantially larger than PIC versus HIST correlations (Fig. 2b–d) and significant (at confidence levels of 95%, 95% and 99%, respectively). This supports the detection of a climate change signal in lake temperature and all three lake ice indices.

Scaling factor confidence intervals for lake temperature and all three ice indices are significantly different from 0, confirming the detection of a climate change imprint in all four variables (Fig. 2e–h). For ice onset, break-up and duration, the HIST time series closely resembles ERA5L, and scaling factors overlap with unity (Fig. 2g–h), providing strong evidence to attribute changes in these variables to external forcings. Lake temperature changes cannot be likewise attributed, as the scaling factor interval for this variable remains below unity. On the whole, this formal statistical evidence confirms that external forcings—and by extension, anthropogenic emissions—can explain reanalysed changes in lake ice onset, break-up and duration.

## Future climate projections

Only a few recent studies[13,15,31] project end-of-century changes in lake temperature and ice cover over large areas under multiple GCM forcings and representative concentration pathways (RCPs), thereby accounting for uncertainties related to meteorological forcing and climate scenario. However, these studies so far disregard both lake model uncertainty and transient lake response to greenhouse gas forcing. Having demonstrated the foregone imprint of climate change on lakes, we project lake temperature and ice conditions across pre-industrial to future periods (1661–2099) under RCPs 2.6, 6.0 and 8.5 (Methods).

By the end of the century, annual mean lake temperatures increase and ice cover decreases unanimously under the high-emission scenario RCP 8.5 (Fig. 3a–e). Lakes warm the most (+4–5 °C by 2070–2099 relative to 1971–2000) in southern temperate latitudes in North America and in temperate latitudes across Eurasia (Fig. 3a and Supplementary Figs. 2–7). In many boreal zones, the June–July–August lake temperature warming exceeds global mean surface air temperature warming by a factor of 1.5–2.0 (Fig. 3b), indicating a high climate sensitivity for these lakes associated with the polar amplification of atmospheric warming[32] and local amplification due to decreased ice cover and local stratification. These spatial sensitivity patterns are consistent across RCPs for lake temperature (Supplementary Figs. 8–10), ice thickness (Supplementary Figs. 11–13) and ice-cover indices (Supplementary Figs. 14-–16). Ice duration decreases by 28–80 days (5th to 95th percentile), with the largest reductions occurring in coastal regions and Scandinavia (>45 days, Fig. 3e). Ice duration projections are

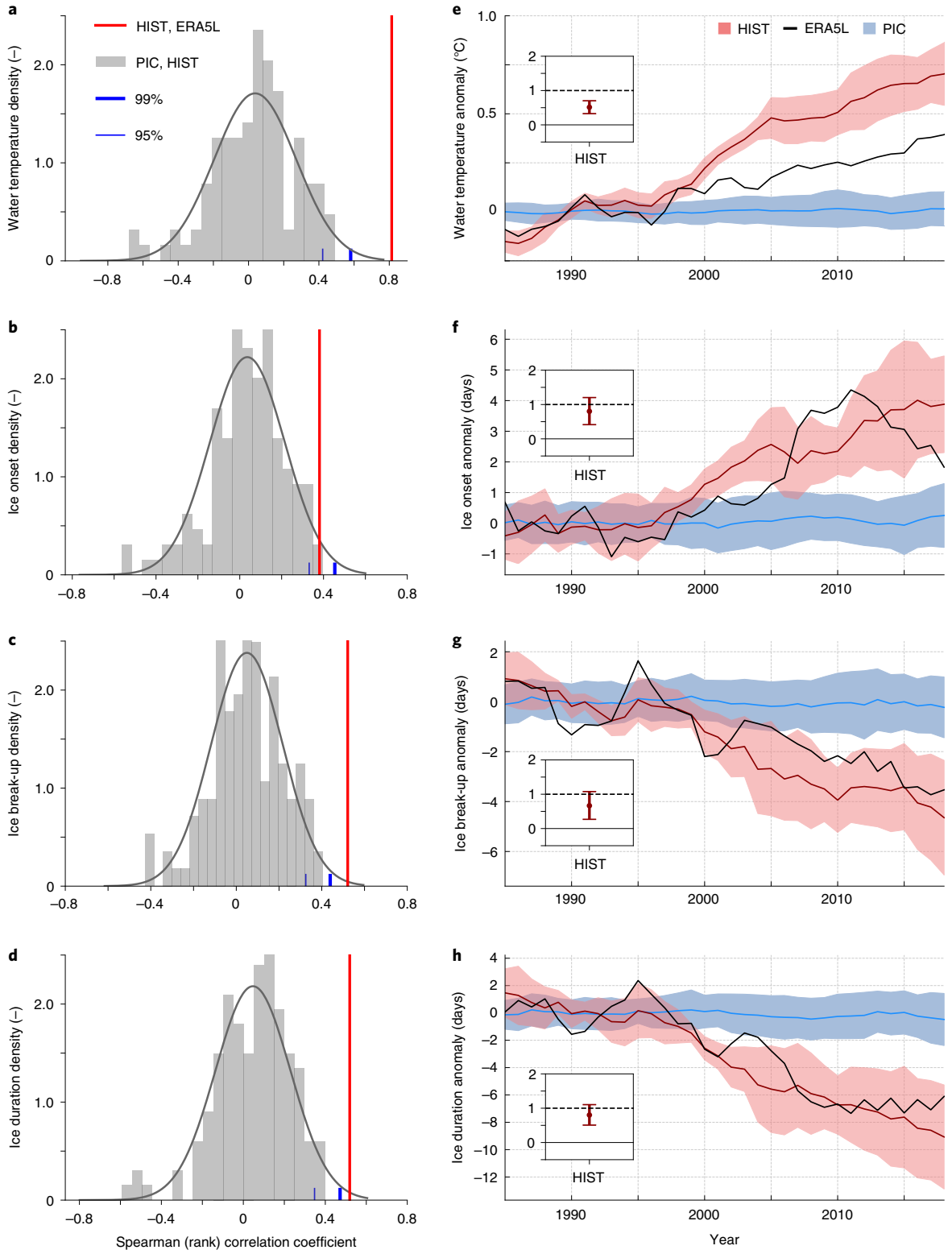


**Fig. 2 | Detection and attribution of the human imprint on lake variables. a–d**, Empirical distribution of correlation coefficients between PIC and HIST for lake temperature (**a**), ice onset (**b**), ice break-up (**c**) and ice duration (**d**). Red lines show the correlation coefficient between HIST and ERA5L. Vertical blue lines mark the 95% and 99% cumulative probability of an assumed normal distribution for the sample of PIC–HIST coefficients. **e–h**, Global multimodel mean time series and spread (ensemble standard deviation) for PIC and HIST forced response patterns and ERA5L smoothed by a 5 yr running mean for lake temperature (**e**), ice onset (**f**), ice break-up (**g**) and ice duration (**h**). Results of single-factor ROF output on HIST are displayed in insets. Scaling factor confidence intervals denote their 2.5–97.5% uncertainty range and infer detection when excluding the 0 line. Attribution is achieved when confidence intervals additionally include unity.

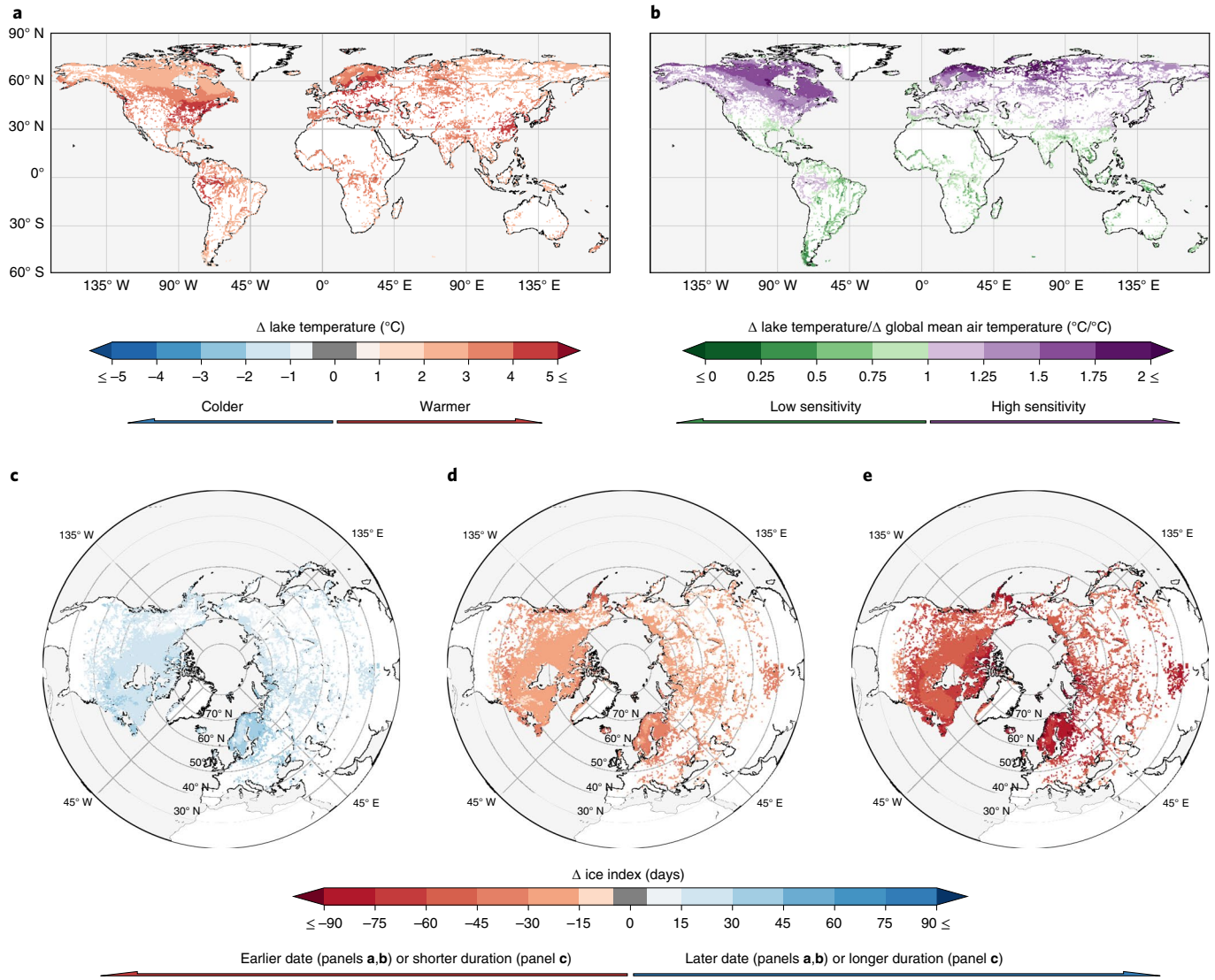


**Fig. 3 | End-of-century change in lake temperature and ice onset, break-up and duration according to RCP 8.5. a**, Multimodel mean change in annual lake temperatures at 2 m depth. **b**, The mean June–July–August lake temperature change at 2 m depth divided by the change in same-year, global-average, annual mean surface air temperature. **c**–**e**, Changes in ice onset, break-up and duration, respectively. All results compare end-of-century (2070–2099) with present-day (1971–2000) conditions.

driven mostly by changes in the timing of ice break-up, which happens consistently earlier in the year by the end of the century and agrees with the seasonality of ice thickness losses (Fig. 3c–e and Supplementary Figs. 17–22).

In all future scenarios, global mean lake temperatures increase while ice thickness and ice duration decrease (Fig. 4). Multimodel mean projections under RCPs 2.6, 6.0 and 8.5 diverge by 2050 at the latest, with only RCP 2.6 showing an end-of-century stabilization (Fig. 4a–c). Global mean projections show high inter-model consistency for all variables, except for ice thickness computed by Community Land Model version 4.5 (Supplementary Figs. 23–25). By 2100, the scenario spread exceeds the uncertainty originating from the lake models, GCMs and natural variability, underscoring the value of mitigation for avoiding severe lake system changes.

Across all future climate scenarios, multimodel mean lake temperature, ice thickness and ice cover scale robustly with air temperature at the global mean level (Fig. 4d–f). Projected global-average mean annual scalings with global mean air temperature for lake temperature, ice duration and ice thickness are +0.9 °C °$C_{air}^{-1}$, –9.7 d °$C_{air}^{-1}$ and –0.033 m °$C_{air}^{-1}$, respectively. RCP 8.5 projections indicate end-of-century global mean anomalies of +4.0 °C for lake temperature, –0.17 m for ice thickness and a 46-day decrease in ice duration relative to pre-industrial conditions. Compared with changes at the global scale, zonal mean annual projections reveal that impacts for lake ice scale the strongest with global mean air temperature anomalies in boreal latitudes (Supplementary Fig. 38b,c). By contrast, annual lake temperature scaling in the tropics exceeds its global mean rate of change (Supplementary Fig. 38g).

## Patterns of change

Our projections reveal coastal–inland gradients in ice duration projections around northern European and Scandinavian coasts and far eastern and western North America that agree with previous studies[33]. Large decreases in ice thickness projected in spring months relative to fall months (Supplementary Figs. 17–19) agree with observed changes in lake ice cover around the Northern Hemisphere[11,34,35]. This is also consistent with the dominant contribution of earlier ice break-up dates to ice duration changes relative to delayed ice onset (Supplementary Figs. 14–16 and 20–22), which has been ascribed to a stronger climate change impact on the spring

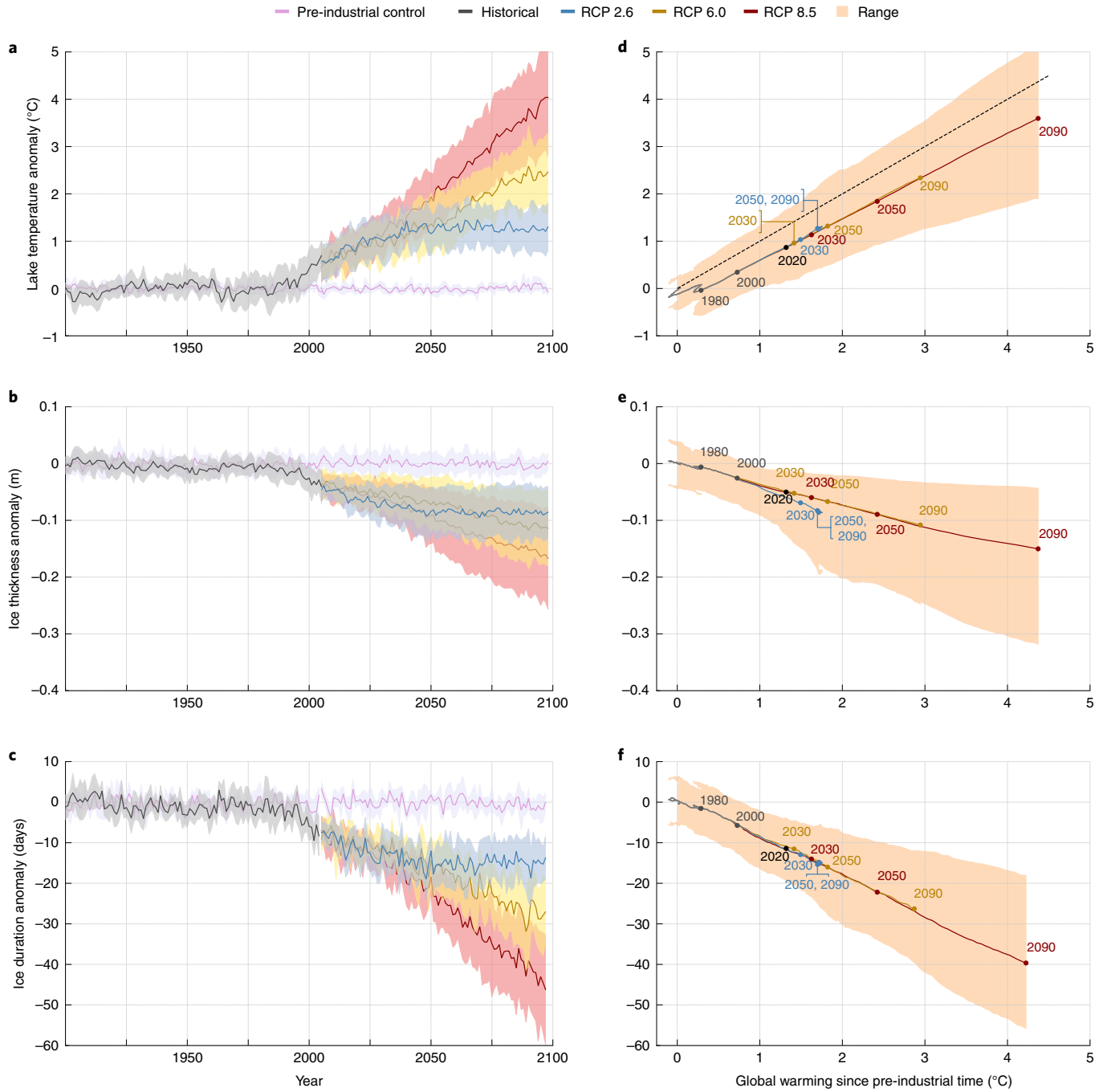


**Fig. 4 | Anomalies for lake temperature, ice thickness and ice cover. a–c**, Multimodel mean anomaly time series of global annual mean lake temperatures (**a**), ice thickness (**b**) and ice-cover duration (**c**). Uncertainty bands in **a**, **b** and **c** represent ±1 standard deviation in lake model ensemble projections. **d–f**, The same lake variable anomalies are scaled against global-average annual mean air temperature anomalies, with uncertainty bands representing the full range of scaled projections. The dashed line in **d** represents a 1:1 scale.

return of the 0 °C isotherm than on its fall timing[36]. At the global mean level, our lake temperature and ice-cover projections for 2100 (Supplementary Figs. 23 and 25) agree with RCP 2.6 and 6.0 projections from a single lake model study over a smaller set of lakes[15].

## Suitability of ERA5-Land

Challenges to global-scale lake modelling arise from parameter value selection, the spatio-temporal coverage and quality of reference products, and the selection of adequate impact variables. While anchored to reality through the step-wise bias correction of their boundary conditions[19] (Methods), the lake variables of ERA5-Land are diagnostics and not subject to direct assimilation with remote sensing or in situ data. Importantly, when compared with in situ and satellite-observed lake changes, ERA5-Land lake simulations perform satisfactorily in representing the global-scale time-series characteristics required as a substitute reference product to lake surface temperature and ice-cover detection and attribution (Supplementary Note 1.1 and Supplementary Figs. 29–36). In addition, ERA5-Land is the only available lake product with sufficient spatial and temporal extent necessary for detection and attribution purposes. In light of this validation, the notable warm bias that exists between ERA5L and HIST lake temperatures (Fig. 2e) could result

from differences in the HIST lake model depth fields, forcing and water-clarity parameterization (Supplementary Note 1.3). Despite this shortcoming and substantial mean biases for some models and variables (Supplementary Figs. 26–28 and Supplementary Note 1.2), the inter-model agreement both at the global scale (Supplementary Figs. 23–25) and with respect to latitudinal, coastal and seasonal characteristics (Supplementary Figs. 2–22) adds confidence to the quality of our projections. Future attribution studies may, however, benefit from the ongoing development of global-scale, multidecadal lake temperature and ice-cover datasets based on remote sensing[37]. As reference datasets and lake models are updated in the near future, optimal fingerprinting techniques may provide even more robust evidence of detection and attribution.

In summary, we showed increases in lake temperature and decreases in ice cover with strong inter-model consistency using an ensemble of five global-scale lake models. With detections achieved at the 95% confidence level, we demonstrate that reanalysed historical changes in lakes worldwide are extremely unlikely[38] to have occurred due to pre-industrial climate variability alone. Further, we attribute changes in all three ice-cover indices to anthropogenic emissions. Our ensemble framework encompasses climate model, lake model, natural variability and scenario uncertainties, which bolsters our projections and reduces sampling uncertainties in detecting and attributing the anthropogenic signal in historical lake variable changes. These projected changes could have manifold consequences for lake thermal regimes, lake ecological processes and provision of lake ecosystem services. The clear dependency of our projections on the radiative forcing scenario and the strong arguments we make for reanalysed changes being both unexplainable by pre-industrial climate variability alone and consistent with anthropogenic forcings underline the benefit of stabilizing lake systems through major societal adjustments towards mitigating climate change.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41561-021-00833-x.

Received: 19 August 2020; Accepted: 31 August 2021;
Published online: 18 October 2021

## References

1. Mueller, H., Hamilton, D. P. & Doole, G. J. Evaluating services and damage costs of degradation of a major lake ecosystem. *Ecosyst. Serv.* **22**, 370–380 (2016).
2. Rinke, K., Keller, P. S., Kong, X., Borchardt, D. & Weitere, M. *Ecosystem Services from Inland Waters and Their Aquatic Ecosystems* (Springer, 2019).
3. Bonan, G. B. Sensitivity of a GCM simulation to inclusion of inland water surfaces. *J. Clim.* **8**, 2691–2704 (1995).
4. Subin, Z. M., Murphy, L. N., Li, F., Bonfils, C. & Riley, W. J. Boreal lakes moderate seasonal and diurnal temperature variation and perturb atmospheric circulation: analyses in the Community Earth System Model 1 (CESM1). *Tellus A* **64**, 15639 (2012).
5. Thiery, W. et al. Understanding the performance of the FLake model over two African Great Lakes. *Geosci. Model Dev.* **7**, 317–337 (2014).
6. Thiery, W. et al. The impact of the African Great Lakes on the regional climate. *J. Clim.* **28**, 4061–4085 (2015).
7. Scott, R. W. & Huff, F. A. Impacts of the Great Lakes on regional climate conditions. *J. Great Lakes Res.* **22**, 845–863 (1996).
8. Griffiths, K., Michelutti, N., Sugar, M., Douglas, M. S. & Smol, J. P. Ice-cover is the principal driver of ecological change in High Arctic lakes and ponds. *PLoS ONE* **12**, e0172989 (2017).
9. Tan, Z., Yao, H. & Zhuang, Q. A small temperate lake in the 21st century: dynamics of water temperature, ice phenology, dissolved oxygen, and chlorophyll *a*. *Water Resour. Res.* **54**, 4681–4699 (2018).
10. Austin, J. A. & Colman, S. M. Lake Superior summer water temperatures are increasing more rapidly than regional temperatures: a positive ice-albedo feedback. *Geophys. Res. Lett.* **34**, L06604 (2007).
11. Ghanbari, R. N., Bravo, H. R., Magnuson, J. J., Hyzer, W. G. & Benson, B. J. Coherence between lake ice cover, local climate and teleconnections (Lake Mendota, Wisconsin). *J. Hydrol.* **374**, 282–293 (2009).
12. Duguay, C. R. et al. Recent trends in Canadian lake ice cover. *Hydrol. Process.* **20**, 781–801 (2006).
13. Sharma, S. et al. Widespread loss of lake ice around the Northern Hemisphere in a warming world. *Nat. Clim. Change* **9**, 227–231 (2019).
14. O'Reilly, C. M. et al. Rapid and highly variable warming of lake surface waters around the globe. *Geophys. Res. Lett.* **42**, 10773–10781 (2015).
15. Woolway, R. I. & Merchant, C. J. Worldwide alteration of lake mixing regimes in response to climate change. *Nat. Geosci.* **12**, 271–276 (2019).
16. O'Reilly, C. M., Alin, S. R., Piisnier, P. D., Cohen, A. S. & McKee, B. A. Climate change decreases aquatic ecosystem productivity of Lake Tanganyika, Africa. *Nature* **424**, 766–768 (2003).
17. Hansen, G. J., Read, J. S., Hansen, J. F. & Winslow, L. A. Projected shifts in fish species dominance in Wisconsin lakes under climate change. *Glob. Change Biol.* **23**, 1463–1476 (2017).
18. Lyons, J. et al. Trends in the reproductive phenology of two Great Lakes fishes. *Trans. Am. Fish. Soc.* **144**, 1263–1274 (2015).
19. Muñoz-Sabater, J. *ERA5-Land Hourly Data from 1981 to Present* (Copernicus, 2019).
20. Ribes, A., Azaís, J. M. & Planton, S. Adaptation of the optimal fingerprint method for climate change detection using a well-conditioned covariance matrix estimate. *Clim. Dyn.* **33**, 707–722 (2009).
21. Allen, M. R. & Stott, P. A. Estimating signal amplitudes in optimal fingerprinting, part I: theory. *Clim. Dyn.* **21**, 477–491 (2003).
22. Bindoff, N. L. et al. in *Climate Change 2013: The Physical Science Basis* (eds Stocker, T. F. et al.) 867–952 (Cambridge Univ. Press, 2013).
23. Gillett, N. P. et al. The Detection and Attribution Model Intercomparison Project (DAMIP v1.0) contribution to CMIP6. *Geosci. Model Dev.* **9**, 3685–3697 (2016).
24. Frieler, K. et al. Assessing the impacts of 1.5 °C global warming—simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b). *Geosci. Model Dev.* **10**, 4321–4345 (2017).
25. Wan, H., Zhang, X., Zwiers, F. & Min, S. K. Attributing northern high-latitude precipitation change over the period 1966–2005 to human influence. *Clim. Dyn.* **45**, 1713–1726 (2015).
26. Qian, C. & Zhang, X. Human influences on changes in the temperature seasonality in mid- to high-latitude land areas. *J. Clim.* **28**, 5908–5921 (2015).
27. Gudmundsson, L., Seneviratne, S. I. & Zhang, X. Anthropogenic climate change detected in European renewable freshwater resources. *Nat. Clim. Change* **7**, 813–816 (2017).
28. Padrón, R. S. et al. Observed changes in dry-season water availability attributed to human-induced climate change. *Nat. Geosci.* **13**, 477–481 (2020).
29. Ribes, A., Planton, S. & Terray, L. Application of regularised optimal fingerprinting to attribution. Part I: method, properties and idealised analysis. *Clim. Dyn.* **41**, 2817–2836 (2013).
30. Myhre, G. et al. in *Climate Change 2013: The Physical Science Basis* (eds Stocker, T. F. et al.) 659–740 (Cambridge Univ. Press, 2013).
31. Maberly, S. C. et al. Global lake thermal regions shift under climate change. *Nat. Commun.* **11**, 1232 (2020).
32. Ito, A. et al. Pronounced and unavoidable impacts of low-end global warming on northern high-latitude land ecosystems. *Environ. Res. Lett.* **15**, 044006 (2020).
33. Dibike, Y., Prowse, T., Saloranta, T. & Ahmed, R. Response of Northern Hemisphere lake-ice cover and lake-water thermal structure patterns to a changing climate. *Hydrol. Process.* **25**, 2942–2953 (2011).
34. Bonsal, B. R., Prowse, T. D., Duguay, C. R. & Lacroix, M. P. Impacts of large-scale teleconnections on freshwater-ice break/freeze-up dates over Canada. *J. Hydrol.* **330**, 340–353 (2006).
35. Korhonen, J. Long-term changes in lake ice cover in Finland. *Nord. Hydrol.* **37**, 347–363 (2006).
36. Bonsal, B. R. & Prowse, T. D. Trends and variability in spring and autumn 0 °C-isotherm dates over Canada. *Climatic Change* **57**, 341–358 (2003).
37. Giardino, C., Merchant, C. & Simis, S. Preparing for the first Lakes ECV climate data record. *Lakes Newsletter* (October 2019).
38. Mastrandrea, M. D. et al. *Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties* Technical Report (IPCC, 2010).

## Methods

**Inter-Sectoral Impact Model Intercomparison Project.** We perform global-scale simulations with five lake models as a part of phase 2b of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b). All simulations adhere to the lake sector protocol[24], which determines simulated periods and scenarios, lake model forcing datasets, the spatial and temporal resolutions of model outputs and lake locations and depths. Pre-industrial control simulations (1661–2099) assume a pre-industrial climate without anthropogenic greenhouse gas forcing[39]. Historical simulations (1861–2005) use a historical climate, whereas future projections (2006–2099) consider RCPs 2.6, 6.0 and 8.5. Four GCMs contributing to phase 5 of the Coupled Model Intercomparison Project (CMIP5)—GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-LR and MIROC5—are used as input to the lake models after bias adjustment to the EWEMBI reference dataset, which is compiled from ERA-Interim reanalysis data adjusted by WATCH forcing methodology, eartH2Observe forcing data and NASA/GEWEX Surface Radiation Budget data[39,40].

The lake models contributing to this study are the Community Land Model version 4.5[41] (CLM4.5), the Arctic Lake Biogeochemistry Model[42] (ALBM), SIMSTRAT-UoG[43], VIC-lake[44] and LAKE[45]. All lake models operate globally at 0.5°×0.5° horizontal resolution. The lake models simulate daily vertical temperature profiles per grid cell based on the mean depth and summed surface area of all lakes within that cell. Presence and grid-scale fractions of lakes within each 0.5° grid cell are given by the Global Lake Database (GLDB)[46–48], which is aggregated from the original 30 arcsec to the 0.5°×0.5° grid. GLDB also provides average lake depth per grid cell for all models except CLM4.5. In CLM4.5, every grid cell has a constant lake depth at 50 m. Individual model characteristics are provided in Supplementary Table 1.

**ERA5-Land.** We use ERA5-Land reanalysis lake ice depth and mixed-layer temperature datasets as reference for lake model evaluation and climate change detection and attribution[19]. The ERA5-Land product delivers lake variables at 0.1° horizontal and hourly temporal resolution computed by the Freshwater Lake model (FLake). ERA5-Land is a land-only re-run of ERA5 with a finer resolution for improved application as reference product for land-based energy and water-flux studies. The ERA5-Land lake reanalysis uses lower atmospheric forcing from the ERA5 reanalysis as boundary conditions and is therefore forced by observations through their assimilation in the atmosphere of ERA5. Lake model computations are embedded as a tile in the Tiled ECMWF Scheme for Surface Exchanges over Land incorporating land surface hydrology[49]. Here, lake variables are computed in each grid cell where inland water bodies cover at least 1% of the surface area of the cell. At the time of analysis, this dataset spans 1981 to 2019 (inclusive).

Using an ERA5-Land derived ice-cover duration series, we apply selection criteria on our projection maps for lake ice variables (Figs. 1 and 3 and Supplementary Figures) based on the first two decades of the series (1981–1999) to avoid presenting results for lakes with erratic or ephemeral ice cover. This baseline is chosen because it contains the reference period in Fig. 1 (1981–1990) and it overlaps with the 'present-day' period (1971–2000) used in results for Fig. 3 and projections in the Supplementary Information. Moreover, using two decades for the following criteria better accounts for inter-annual variability than does using a single decade (1981–1990). Our criteria involve selecting pixels from the native ERA5L ice-cover series where ice occurs for at least 20 days in 10 years of 1981–1999. This effectively limits our consideration of changes in ice cover to regions with more consistent ice according to our reference product.

**Data processing.** Post-processing of model ice thickness outputs was performed to attain homogenized ice onset, break-up and duration values. Ice-cover indices were calculated with hydrological years, defined as year-long periods that contain ice onset or break-up dates for lakes in the Northern Hemisphere. For ice onset calculations, we select the October to September hydrological year and convert each pixel value with ice cover to the day of the year of its time step. After this, we added 365 to periods between 1 January and 30 September so that the days of the year monotonically increase during one hydrological year. A temporal minimum was calculated across this adjusted 1 October (year $t$) to 30 September (year $t+1$ series). This was performed for all available October to September hydrological years in the series, resulting in annual maps of ice start dates. The same process with a temporal maximum calculation across its September to August hydrological year was done for ice break-up calculations, resulting in maps of annual ice end dates. Ice duration is computed as the sum of all 'ice-on' days across the October to September hydrological year. We analyse lake temperature at 2 m depth to enable comparison with ERA5-Land mixed-layer temperatures and to avoid an overly strong dependence on surface air temperature, which can be expected from lake surface temperature analyses. Global mean calculations on ice thickness datasets include all pixels without ice cover. Reanalysis data are coarsened to the 0.5°×0.5° ISIMIP grid. Before calculating spatial means, all datasets are masked for overlapping pixels between lake model simulations and reanalysis data. For all scaling plots, we use global mean air temperature anomalies to scale lake variable impacts to communicate results in the context of the IPCC trajectories and internationally agreed climate targets such as the Paris Agreement to limit global warming to well below 2 °C (ref. [39]).

**Detection and attribution.** We generate all-forcings response patterns (HIST) by concatenating each ISIMIP lake model's historical time series (1861–2005) with the RCP 8.5 (2006–2099) future simulations to sample forced response patterns for the same period as the ERA5L (1981–2019). Here, only the RCP 8.5 anomalies are used for the 2006–2019 period to avoid the artificial consistencies among HIST patterns that would occur if all future scenarios were added to the ensemble, which would replicate 1981–2005 anomalies. Next, global annual means are computed from these series, yielding a total of 40 HIST realizations (8 per lake model). For a forced response pattern without human influence (PIC), all available ISIMIP pre-industrial control simulations are concatenated for each lake model and cut into non-overlapping global mean 'chunks' matching the time span of ERA5L. This ideally provides 44 (11×4) chunks of pre-industrial climate-variability-driven simulations per lake model if pre-industrial control simulations span 1661–2099 for each GCM forcing. Reanalysis reference products and response patterns are then computed as anomalies through temporal centring (each series' temporal mean is subtracted) and applied to two detection and attribution approaches: a correlation-based view on detection and ROF to confirm detection and attribution.

The correlation approach (Fig. 2a–d) uses all available HIST and PIC anomalies without smoothing. For each lake variable, Spearman (rank) correlation coefficients are calculated between the global annual mean of all available historical simulations (HIST) and every available global annual mean PIC chunk. These correlation coefficients comprise the empirical distributions in Figure 2. A correlation coefficient is then computed between ERA5L and the mean of the HIST ensemble, plotted as a red vertical line. A normal distribution using the mean and standard deviation of PIC–HIST correlations is assumed for reporting the 95% and 99% confidence levels for comparison with ERA5L–HIST correlation. We use the Spearman correlation coefficient because of its resistance to outliers; however, results are consistent with a Pearson correlation.

We use ROF with a TLS regression to compute scaling factors that fit annual, multimodel mean HIST anomalies to ERA5L at the global mean level (Fig. 2e–h) for one lake variable at a time. This follows a generalized linear regression model of the form:

$$\mathbf{y} = X^{*}\boldsymbol{\beta} + \varepsilon$$

$$X = X^{*} + \nu$$

where $\mathbf{y}$ is a vector of $n$ observations (ERA5-Land lake reanalysis; ERA5L), $X$ is a matrix of $m$ columns of multimodel mean simulated response patterns containing noise $\nu$ (ISIMIP simulations; HIST), $\boldsymbol{\beta}$ is a vector of scaling factors and $\varepsilon$ is the regression residual, representing the internal variability in $\mathbf{y}$. We take a single-factor approach; the regression is fit for one response pattern generated with all external forcings (HIST) instead of regressing $\mathbf{y}$ onto a linear combination of response patterns to separate external forcings. The latter approach is often some collection of additive response patterns to natural and anthropogenic forcings (such as NAT representing volcanoes and aerosols, ANT for anthropogenic emissions and LU for land cover change). Therefore, in this study, $X$ contains only one column or response pattern ($m=1$) in each lake variable's analysis. The outcome of the analysis is therefore the single slope parameter of the regression, $\boldsymbol{\beta}$ and its confidence interval. If $\boldsymbol{\beta}$ does not overlap with 0, ROF communicates detection, revealing that an aspect of lake systems—the lake variable of interest—is statistically different from its natural or pre-industrial state. In addition, if $\boldsymbol{\beta}$ overlaps with 1, ROF attributes the changes in a given lake variable to all external forcings, revealing that a lake variable's trend can be explained by historical climate forcings dominated by anthropogenic climate change.

In a TLS framework, the regression is computed to minimize residuals perpendicular to the best fit line[21]. This addresses uncertainty in $X$, underlining the assumption in the TLS approach to optimal fingerprinting that model simulated response patterns are not perfectly known. In other words, TLS contends with the presence of noise in the observed $X$, represented by $\nu$, which affects the true deterministic $X^{*}$. Therefore, TLS is a strong choice for small-ensemble study cases with greater sampling uncertainty. This contrasts the ordinary least squares approach, which fits a regression by minimizing vertical residuals, thereby only accounting for noise in $\mathbf{y}$ and assuming that the response patterns in $X$ are perfectly known (as residuals are not accounted for along the $x$-axis). The TLS regression is achieved through a singular value decomposition (SVD) on $[\mathbf{y}, X]$.

Before the TLS fit, $\mathbf{y}$ and $X$ are converted to 5 yr block means, temporally centred (by subtracting their mean) and pre-whitened. Pre-whitening to achieve unit noise is the 'optimization' of signals in ROF. This is done with a regularized covariance matrix, $\widehat{C}_1$, which represents internal variability in our lake variables. $\widehat{C}_1$ is derived from one of two covariance estimates, $C_1$ and $C_2$, computed from non-overlapping, equal-sized samples (chunks) of all available PIC series. As we use a model-derived estimate of noise to pre-whiten $\mathbf{y}$, its compatibility with the noise in $\mathbf{y}$ is later validated second-hand by a residual consistency test (RCT). Key to ROF, regularization involves conforming $\widehat{C}_1$ to equal $\lambda C_1 + \rho I$. Here, $I$ is the identity matrix, and $\lambda$ and $\rho$ are coefficients whose estimators are provided by Ledoit and Wolf[50]. This avoids underestimating the lowest eigenvalues of $\widehat{C}_1$[29]. $C_2$ is used for calculating the confidence intervals on scaling factors and performing the RCT. Final computations of scaling factors, their confidence intervals and RCTs are

taken as the median of 1,000 realizations of ROF through shuffling the PIC chunks from which $C_1$ and $C_2$ are computed[51].

The RCT validates the residuals in the TLS regression against the assumed internal variability estimated using PIC chunks[29]. Here, $C_2$ and X are used in Monte Carlo simulations to bootstrap 1,000 samples of virtual reference series, fingerprints and covariance matrices assuming a perfect fit with $\boldsymbol{\beta}=1$. The smallest squared singular value (or eigenvalue, $\lambda$) of the SVD in the original TLS fit—representing the residuals in the regression—is then corrected and used as a test statistic against 1,000 virtual eigenvalues ($\lambda_{\mathrm{virt},i=1,\ldots1000}$) and their kernel density estimates ($\lambda$ is tested against 1,000 virtual, empirical distributions). The RCT is passed if $\lambda$ is consistent with these distributions, which is considered true if the average position of $\lambda$ in the virtual distributions yields a $P$ value greater than 0.10 (Supplementary Note 1.4).

**Future projections.** We calculate all maps as signals across 1971–2000 and 2070–2099 mean baseline and future periods. For scaling, each signal map is first divided by the change in global mean air temperature for the same period before calculating ensemble means. For each GCM–Lake model combination, we compute global mean anomalies relative to the global temporal average of the pre-industrial control simulation (Fig. 4). Global mean air temperature series from GCMs are treated the same. In Fig. 4d–f, series are smoothed with a 21 yr running mean to reduce natural variability effects.

## Data availability

The ISIMIP2b lake sector simulations presented in this study are available through the Earth System Grid Federation (ESGF, https://esgf-data.dkrz.de/). The ERA5-Land lake data used in this study are developed by the European Centre for Medium-Range Weather Forecasts (ECMWF) and are available through the Copernicus Climate Change Service's Climate Data Store (CDS, https://cds.climate.copernicus.eu/cdsapp#!/search?type=dataset). The Global Lake Temperature Collaboration Dataset lake surface temperatures used for evaluating ERA5-Land can be found here: https://portal.edirepository.org/nis/mapbrowse?packageid=knb-lter-ntl.10001.3. ESA CCI lake products can be found here: https://catalogue.ceda.ac.uk/uuid/3c324bb4ee394d0d876fe2e1db217378. The Global Lake and River Ice Phenology Database is available at https://nsidc.org/data/lake_river_ice/.

## Code availability

All code used to generate these analyses are available through the GitHub repository of the Department of Hydrology and Hydraulic Engineering at VUB (https://github.com/VUB-HYDR/2021_Grant_etal).

## References


39. Frieler, K. et al. Assessing the impacts of 1.5 °C global warming—simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b). *Geosci. Model Dev.* **10**, 4321–4345 (2017).
40. Lange, S. *EartH2Observe, WFDEI and ERA-Interim Data Merged and Bias-corrected for ISIMIP (EWEMBI)* (GFZ Data Services, 2016).
41. Lawrence, D. M. et al. Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model. *J. Adv. Model. Earth Syst.* **3**, M03001 (2011).
42. Tan, Z. et al. Modeling methane emissions from arctic lakes: Model development and site-level study. *J. Adv. Model. Earth Syst.* **6**, 513–526 (2015).
43. Goudsmit, G. H., Burchard, H., Peeters, F. & Wüest, A. Application of k-$\epsilon$ turbulence models to enclosed basins: the role of internal seiches. *J. Geophys. Res. Oceans* **107**, 23-1–23-13 (2002).
44. Bowling, L. C. & Lettenmaier, D. P. Modeling the effects of lakes and wetlands on the water balance of Arctic environments. *J. Hydrometeorol.* **11**, 276–295 (2010).
45. Stepanenko, V. et al. LAKE 2.0: a model for temperature, methane, carbon dioxide and oxygen dynamics in lakes. *Geosci. Model Dev.* **9**, 1977–2006 (2016).
46. Kourzeneva, E., Asensio, H., Martin, E. & Faroux, S. Global gridded dataset of lake coverage and lake depth for use in numerical weather prediction and climate modelling. *Tellus A* **64**, 15640 (2012).
47. Subin, Z. M., Riley, W. J. & Mironov, D. An improved lake model for climate simulations: model structure, evaluation, and sensitivity analyses in CESM1. *J. Adv. Model. Earth Syst.* **4**, M02001 (2012).
48. Choulga, M., Kourzeneva, E., Zakharova, E. & Doganovsky, A. Estimation of the mean depth of boreal lakes for use in numerical weather prediction and climate modelling. *Tellus A* **66**, 21295 (2014).
49. Balsamo, G., Dutra, E., Beljaars, A. & Viterbo, P. Evolution of land surface processes in the Integrated Forecast System. *ECMWF Newsl.* **127**, 17–22 (2011).
50. Ledoit, O. & Wolf, M. A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivar. Anal.* **88**, 365–411 (2004).
51. Gudmundsson, L., Seneviratne, S. I. & Zhang, X. Anthropogenic climate change detected in European renewable freshwater resources. *Nat. Clim. Change* **7**, 813–816 (2017).


## Acknowledgements


We are grateful to the Potsdam Institute for Climate Impact Research (PIK) for initiating and coordinating the ISIMIP initiative, with special thanks to M. Büchner for his oversight of ISIMIP data publishing, and to the modelling centres for making their impact simulations publicly available through ESGF. We acknowledge the European Centre for Medium-Range Weather Forecasts (ECMWF) and the Copernicus Climate Change Service for their provision of publicly available ERA5-Land lake data; this paper contains modified Copernicus Climate Change Information [2021]. Furthermore, L.Grant is funded by European Copernicus Climate Change Service (C3S) implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF) under the service contract Independent Assessment on ECVs led by National Research council of Italy (CNR) with the funding number ECMWF/Copernicus/2017/C3S_511_CNR. We owe many thanks to F. Fröb and A. Winkler for sharing their regularized optimal fingerprinting python code and to M. Schmid for the helpful discussions. We also thank the National Center for Atmospheric Research (NCAR) for maintaining CLM and making the source code publicly available. I.V. is a research fellow at the Research Foundation Flanders (FWO) (FWOTM920). W.T. acknowledges the Uniscientia Foundation and the ETH Zurich Foundation for their support to this research. Z.T. is supported by the US DOE's Earth System Modeling programme through the Energy Exascale Earth System Model (E3SM) project. The computational resources and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by the Research Foundation Flanders (FWO) and the Flemish Government, department EWI. R.M. participated through the project WATExR of the JPI Climate ERA4CS Program and acknowledges funding from the CERCA programme of the Generalitat de Catalunya. V.M.S. and A.V.D. used the HPC facilities of Lomonosov Moscow State University ('Lomonosov-2' supercomputer) and were supported by the Russian Ministry of Science and Higher Education, agreement no. 075-152019-1621. A.B.G.J acknowledges the Talent Programme Veni of the Netherlands Organisation for Scientific Research (NWO) (VI.Veni.194.002).


## Author contributions

L. Grant, I.V. and W.T. designed the study. L. Grant wrote the manuscript with support from all authors and performed all analyses under the supervision of I.V. and W.T. L. Gudmundsson provided guidance on the detection analysis. Z.T., M.P., V.M.S., A.V.D., B.D., A.B.G.J., S.I.S. and W.T. conducted the global lake model simulations. J.S., F.Z., M.G., D.P., R.M. and W.T. coordinated the ISIMIP lake sector activities. M.C. and G.B. helped validate ERA5-Land reanalysis data as reference products. I.V.d.V. provided oversight for data publishing. L. Grant and I.V. performed additional analyses in response to referee comments and together composed the referee response letter with the help of all authors.

## 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/s41561-021-00833-x.

**Correspondence and requests for materials** should be addressed to Luke Grant.

**Peer review information** *Nature Geoscience* thanks Peter Stott, Matthew Hipsey and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Thomas Richardson.

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