# *Widespread deoxygenation of temperate lakes*

Article

Accepted Version

**Title:**

Widespread deoxygenation of temperate lakes

**Authors:**

Stephen F. Jane[1,2], Gretchen J. A. Hansen[3], Benjamin M. Kraemer[4], Peter R. Leavitt[5,6], Joshua L. Mincer[1], Rebecca L. North[7], Rachel M. Pilla[8], Jonathan T. Stetler[1], Craig E. Williamson[8], R. Iestyn Woolway[9], Lauri Arvola[10], Sudeep Chandra[11], Curtis L. DeGasperi[12], Laura Diemer[13], Julita Dunalska[14], Oxana Erina[15], Giovanna Flaim[16], Hans-Peter Grossart[17, 18], K. David Hambright[19], Catherine Hein[20], Josef Hejzlar[21], Lorraine L. Janus[22], Jean-Philippe Jenny[23], John R. Jones[7], Lesley B. Knoll[24], Barbara Leoni[25], Eleanor Mackay[26], Shin-Ichiro S. Matsuzaki[27], Chris McBride[28], Dörthe C. Müller-Navarra[29], Andrew M. Paterson[30], Don Pierson[2], Michela Rogora[31], James A. Rusak[29], Steven Sadro[32], Emilie Saulnier-Talbot[33], Martin Schmid[34], Ruben Sommaruga[35], Wim Thiery[36, 37], Piet Verburg[38], Kathleen C. Weathers[39], Gesa A. Weyhenmeyer[2], Kiyoko Yokota[40], and Kevin C. Rose[1]

1. Department of Biological Sciences, Rensselaer Polytechnic Institute, Troy, NY, USA
2. Department of Ecology and Genetics/Limnology, Uppsala University, Uppsala, Sweden
3. Department of Fisheries, Wildlife and Conservation Biology, University of Minnesota, St. Paul, MN, USA
4. Department of Ecosystem Research, IGB Leibniz institute for freshwater ecology and inland fisheries, Berlin, Germany

5. Institute of Environmental Change and Society, University of Regina, Regina, Saskatchewan, Canada

6. Institute for Global Food Security, Queen's University Belfast, Belfast, County Antrim, United Kingdom

7. School of Natural Resources, University of Missouri, Columbia, MO, USA

8. Department of Biology, Miami University, Oxford, OH, USA

9. Centre for Freshwater and Environmental Studies, Dundalk Institute of Technology, Dundalk, Ireland

10. Lammi Biological Station, University of Helsinki, Pääjärventie 320, Lammi, Finland

11. Biology Department and Global Water Center, University of Nevada, Reno, NV, USA

12. King County Water and Land Resources Division, Seattle, WA, USA

13. FB Environmental Associates, Portsmouth, NH, USA

14. Department of Water Protection Engineering, University of Warmia and Mazury in Olsztyn, Prawocheńskiego str. 1, Olsztyn, Poland

15. Department of Hydrology, Lomonosov Moscow State University, Leninskiye Gory 1, Moscow, Russia

16. Department of Sustainable Agro-ecosystems and Bioresources, Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige, Italy

17. Department of Experimental Limnology, Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Alte Fischerhuette 2, Stechlin, Germany

18. Institute of Biochemistry and Biology, Potsdam University, Maulbeerallee 2, Potsdam, Germany

19. Plankton Ecology and Limnology Laboratory, Geographical Ecology Group, and Program in Ecology and Evolutionary Biology, Department of Biology, The University of Oklahoma, Norman, Oklahoma, USA

20. Wisconsin Department of Natural Resources, Madison, WI, USA

21. Institute of Hydrobiology, Biology Centre CAS, Na Sádkách 7, České Budějovice, Czech Republic

22. Bureau of Water Supply, New York City Department of Environmental Protection, 465 Columbus Ave., Valhalla, NY, USA

23. CARRTEL Limnology Center, Institut National de la Recherche Agronomique (INRA), Université Savoie Mont Blanc, 73000 Chambéry, France

24. Itasca Biological Station and Laboratories, University of Minnesota, Lake Itasca, MN, USA

25. Department of Earth and Environmental Sciences, University of Milan-Bicocca, Piazza della Scienza 1, Milan, Italy

26. Lake Ecosystems Group, Centre for Ecology & Hydrology, Bailrigg, Lancaster, United Kingdom

27. Center for Environmental Biology and Ecosystem Studies, National Institute for Environmental Studies, 16-2 Tsukuba, Ibaraki, Japan

28. Environmental Research Institute, Hamilton, New Zealand

29. Department of Biology, University of Hamburg, Hamburg, Germany

30. Ontario Ministry of the Environment, Conservation and Parks, Dorset Environmental Science Centre, Dorset, ON, Canada

31. CNR Water Research Institute, L. go Tonolli 50, Verbania Pallanza, Italy

32. Department of Environmental Science and Policy, University of California, One Shields Ave., Davis, CA, USA

33. Centre d'études nordiques (CEN) and Faculty of Sciences and Engineering, Université Laval, Québec, Canada

34. Eawag, Swiss Federal Institute of Aquatic Science and Technology, Surface Waters – Research and Management, Kastanienbaum, Switzerland

35. Department of Ecology, University of Innsbruck, Innsbruck, Austria

36. Vrije Universiteit Brussel, Department of Hydrology and Hydraulic Engineering, Pleinlaan 2, Brussels, Belgium

37. ETH Zurich, Institute for Atmospheric and Climate Science, Universitaetstrasse 16, Zurich, Switzerland

38. National Institute of Water and Atmospheric Research Ltd (NIWA), Gate 10 Silverdale Rd., Hillcrest, Hamilton, New Zealand

39. Cary Institute of Ecosystem Studies, Box AB, Millbrook, New York, USA

40. Biology Department, State University of New York College at Oneonta (SUNY Oneonta), Oneonta, New York, USA

**Summary paragraph:**

The concentration of dissolved oxygen in aquatic systems helps regulate biodiveristy[1, 2], nutrient biogeochemistry[3], greenhouse gas emissions[4], and drinking water quality[5]. The long-term declines in dissolved oxygen concentrations in coastal and ocean waters have been linked to climate warming and human activity[6, 7], but little is known about changes in dissolved oxygen concentrations in lakes. While dissolved oxygen solubility decreases with increasing water temperatures, long-term lake trajectories are not necessarily predictable. Oxygen losses in warming lakes may be amplified by enhanced decomposition and stronger thermal stratification[8, 9] or they may increase as a result of enhanced primary production[10]. Here we analyse 45,148 dissolved oxygen and temperature profiles from 393 temperate lakes spanning 1941-2017. We find that a decline in dissolved oxygen is widespread in surface and deep-water habitats. The decline in surface waters is primarily associated with reduced solubility under warmer water temperatures, although surface dissolved oxygen increased in a subset of highly-productive warming lakes, likely due to increasing phytoplankton production. In contrast, the decline in deep waters is associated with stronger thermal stratification and water clarity losses, but not with changes in gas solubility. Our results suggest that climate change and declining water clarity have altered the physical and chemical environment of lakes. Freshwater dissolved oxygen losses are 2.5-10 times greater than observed in the world's oceans[6, 7] and could threaten essential lake ecosystem services[2, 3, 5, 11].

**Main text:**

The concentration of dissolved oxygen (DO) in aquatic systems influences biodiversity[1,2], nutrient biogeochemistry[3], greenhouse gas emissions[4], drinking water quality[5], and, ultimately, human health[12]. Many aquatic species require well-oxygenated habitat[11, 13] and cool water to survive warm summers[2, 11]. Loss of deep-water DO degrades water quality by promoting the release of accumulated nutrients from sediments into water[1, 3], which can increase phytoplankton biomass. This process can also facilitate harmful algal blooms[5], which can compromise water supplies and harm human health[12]. Despite clear evidence of large-scale deoxygenation in ocean waters[6, 7], there are no systematic large-scale studies of this phenomenon in lakes[3].

DO concentrations should decline with increasing water temperature due to reduced gas solubility. However, other mechanisms can alter DO, potentially amplifying or counteracting losses predicted from solubility changes alone. For example, rates of heterotrophic respiration increase with temperature faster than primary production[9], and surface-temperature warming can increase the strength and duration of thermal stratification, reducing water circulation, and preventing deep-water DO replenishment[8, 14, 15]. Studies of individual lakes demonstrate deep-water DO concentrations can decrease with lake warming[3, 8, 15, 16], reducing access to cold-water habitat essential to many organisms[11]. However, given the many feedbacks and processes regulating DO, overall trajectories currently defy *a priori* prediction.

We addressed this critical issue by compiling and analyzing an extensive database of lake temperature and DO profiles to characterize widespread and long-term changes in DO concentration and its causes. We used data from 393 temperate lake and reservoir basins, each with a minimum of 15 years of observation (median: 24 years), and report population medians from long-term surface- (epilimnion) and deep-water (hypolimnion) trends in temperature, DO concentration, and DO saturation during the late summer period when seasonal DO depletion is expected to be pronounced[17]. Our analyses revealed that lake DO concentrations have declined in both surface and deep waters from 1980 to 2017 by 0.45 and 0.42 mg $L^{-1}$, respectively (Fig. 1). These rates represent losses of 5.1 and 20.2% for surface and deep waters, respectively, and were substantially greater than those observed for the oceans, where total water-column DO has declined about 2% since 1960[6].

While deep-water temperatures have been virtually stable since observations began (Fig. 1a; –0.01°C decade$^{-1}$), both deep-water DO concentration and percent saturation declined through time (–0.12 mg $L^{-1}$ decade$^{-1}$ and –1.2% decade$^{-1}$; respectively, Fig. 1b, c). Declines were unrelated to solubility as predicted changes based on solubility (slight increase of 0.01 mg $L^{-1}$) were negligible compared with observed losses (median -0.23 mg $L^{-1}$ based on last five years relative to first five years of each time series, Fig. 2b) Declining DO, despite essentially unchanging solubility, implies deep-water habitats have become increasingly inhospitable for organisms with aerobic metabolism, including fishes. We quantified potential impacts of such declines on habitat availability by calculating trends in $T_{DO3}$, the minimum water column temperature where DO was at least 3 mg $L^{-1}$. This metric was developed to quantify oxy-thermal habitats for cold-water fisheries[11]. In lakes where DO was below 3 mg $L^{-1}$ anywhere in the water column at least once in the time series (n = 369), $T_{DO3}$ increased by 0.17°C decade$^{-1}$, with 68.0% of lakes having positive trends and declining habitat for many cold-water species.

In contrast to trends observed for deep waters, variation in surface-water DO concentrations was well explained by changes in gas solubility. Consistent with other global-scale lake studies[18], median air temperatures warmed at 0.30°C decade$^{-1}$ and median lake surface waters warmed at 0.39°C decade$^{-1}$. Additionally, median wind speed and precipitation declined (trends of -0.04 m $s^{-1}$ $decade^{-1}$ and -4.23 mm $decade^{-1}$, respectively), while shortwave radiation increased (1.88 W $m^2$ $decade^{-1}$; Table S1). Surface-water temperature increases were best explained by spring and summer air temperature increases and by summer wind speed declines (Table S2). Surface-water DO concentrations declined at -0.11 mg $L^{-1}$ $decade^{-1}$ (Fig. 1b). Comparing the last five years relative to first five years of each time series revealed that the median change predicted due to solubility loss was ~63% of the median observed decline in DO concentration, with solubility-predicted loss of 0.12 versus observed losses of 0.19 mg $L^{-1}$ (Fig. 2a).

Despite a strong influence of water temperature on DO concentration in surface-waters, there was substantial variability among lakes (Fig. 2a), and a large subset of lakes exhibited increases in both water temperature and DO concentration (n=87; Fig. 3d). Analysis of the interaction between DO concentration, surface temperature, and water clarity (measured as Secchi depth, a proxy for trophic status[19]) showed that DO concentration generally decreased with increasing temperature. However, in lakes with low water clarity (< 2 m), DO concentration increased when average mean summer surface-water temperatures exceeded ~24°C (Fig. 3c). Similarly, in a subset of lakes with chlorophyll data (a proxy for phytoplankton biomass; n = 162), positive DO trends were observed when chlorophyll was high and surface temperatures exceeded ~25°C, (Fig. 3b; $P < 0.001$). Thus, we suggest that eutrophication and warming interact to increase surface-water DO concentration despite reduced gas solubility.

Lakes with increasing DO concentration in warming surface waters had significantly higher surface-water temperatures (Fig. 3a; $P = 0.016$) and their watersheds contained a significantly higher proportion of agriculture ($P = 0.046$) and developed land cover ($P < 0.001$) compared with other lakes. When developed land exceeded ~50% of a watershed and surface water temperature exceeded ~25°C, the probability of a warming lake having an increasing DO trend was >50%. Combined, these analyses highlight a potential threshold above which water temperatures and lake productivity interact to elevate DO concentration in surface waters despite declining gas solubility. While we lack data on phytoplankton taxonomic composition, evidence indicates that phytoplankton blooms are increasing globally[20], in particular due to cyanobacteria[21]. High temperatures and elevated nutrient loading can promote surface cyanobacteria blooms whose photosynthesis leads to DO supersaturation, particularly in eutrophic lakes as temperatures exceed ~23-25°C[10, 21]. Consistent with this inferred mechanism, we note these same lakes exhibited consistently low deep-water DO concentration (median: 0.64 mg $L^{-1}$) relative to other lakes (median: 3.42 mg $L^{-1}$), as is expected when a large phytoplankton biomass sinks and is decomposed in deep-water habitats[22]. Deep water DO changes are described in more detail below.

Decadal-scale trends in DO were associated with non-linear changes in surface-water temperature (Fig. 2c-f; Fig. S1). For example, although surface-water temperatures generally increased from 1980 onwards, there was a period of accelerated increase during 1990-2000, with slower warming thereafter (Fig. 2c), consistent with the "warming hiatus" observed during 1998-2012[23]. This trend occurs across the population of all lakes, as well as the subset of lakes sampled continuously throughout this period. Similarly, surface-water DO exhibited periodic deviations from an overarching trend of declining DO concentration (Fig. 2d), mainly due to the productive lakes exhibiting increasing DO levels in surface waters (Fig. 2d, blue line). Excluding these lakes, analysis of the remaining sites showed a consistent long-term decline in surface-water DO (Fig. 2d, red line). Deep-water temperatures exhibited a pronounced multi-decadal oscillation since 1980 (Fig. 2e) as has been observed in some lakes previously[24], whereas deep-water DO concentration declined consistently through time (Fig. 2f).

While surface-water DO concentration changes were generally well predicted by solubility changes, deep-water DO changes were more strongly associated with changes in water clarity and water-column density differences (Figs. 4 and S2). For example, water clarity losses exceeding 1 m were associated with substantial reductions in deep-water DO saturation (Fig. S2). Mechanistically, increases in phytoplankton biomass or dissolved organic matter (DOM) reduce water clarity while increasing oxygen-consuming respiration[19, 22, 25]. Increases in phytoplankton biomass and DOM are often caused by land use change and recovery from acid deposition, respectively[26]. However, there was no overarching decline in water clarity across study lakes. Indeed, 51% of lakes had clarity increases and 49% had decreases, and only 39% of lakes exhibited both water clarity loss and DO saturation loss (Fig. 4a).

Deep-water DO decreased substantially in lakes where the water column density difference between surface and deep waters increased by more than ~0.5 kg $m^{-3}$ (Fig 4b; Fig. S2b). Strong increases in the density difference indicate intensified stratification that reduces vertical mixing and replenishment of deep-water DO from the atmosphere, and may reduce nutrient upwelling to surface waters[3, 15]. Water column density differences increase due to water clarity losses as well as other factors that increase heat gain in near-surface waters, including climate warming[26] and atmospheric stilling[27]. Increased water column density differences may also be associated with earlier onset of seasonal stratification and thus more time for oxygen consumption before the summer sampling period[22]. We found that changes in water-column density differences were best explained by changes in deep water temperature and climate characteristics (Fig. S3). Despite no overarching among-lake trend in water clarity or deep-water temperature, stratification strength increased in 84% of lakes that stratified, with 61% of basins exhibiting both increased density difference and DO saturation loss (Fig 4b). Warming surface-water temperatures combined with unchanging deep-water temperatures (Fig. 1a) increases the density difference in lake water columns (median rate: 0.10 kg $m^{-3}$ $decade^{-1}$). We observed unchanging deep-water DO in lakes where both clarity and stratification were unchanged (Fig. 4c, d). Therefore, we anticipate further DO losses in deep waters of lakes where water clarity continues to decline or thermal stratification intensifies, whether due to atmospheric warming, stilling, or both [26, 27].

Despite a wide range of lake and catchment characteristics, the overall trend of temperate lake deoxygenation is clear, with climate changes and water clarity losses contributing to declines in lake DO concentration at rates ~2.5-10 times greater than those observed in the global oceans[6, 7]. We find deep-water lake habitats are especially threatened, and deep-water DO trends may portend future losses of cold-water and oxygen-sensitive species[2], increased internal nutrient loading which exacerbates eutrophication[3] and the formation of harmful algal blooms[5], and potentially increased outgassing of stored methane[4]. While already rapid, future losses in lake DO may accelerate due to continued anthropogenic modifications of the environment, including eutrophication[22], salinization[28], and hydrological management[28]. While many lakes have undergone active management to reduce nutrient loads, in part to mitigate phytoplankton growth and deep-water oxygen loss[28], our findings suggest such actions will likely require more rigorous efforts in the future to counter the effects of climate and land use change.

**References:**


1. Wetzel, R. G. 2001. Chap. 9 oxygen. In: *Limnology*, 3rd edn (ed Wetzel R. G.), pp 151-168, Academic Press, San Diego.

2. Schindler, D. Warmer climate squeezes aquatic predators out of their preferred habitat. *Proc. Natl. Acad. Sci.*, **114**, 9764-9765 (2017).

3. North, R. P., North, R. L., Livingstone, D. M., Köster, O., & Kipfer, R. Long-term changes in hypoxia and soluble reactive phosphorus in the hypolimnion of a large temperate lake: consequences of a climate regime shift. *Glob. Change Biol.,* **20**, 811-823 (2014).

4. Fernández, J. E., Peeters, F., & Hofmann, H. Importance of the autumn overturn and anoxic conditions in the hypolimnion for the annual methane emissions from a temperate lake. *Environ. Sci. Technol.*, **48**, 7297-7304 (2014).

5. Michalak, A. M., et al. Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. *Proc. Natl. Acad. Sci.*, **110**, 6448-6452 (2013).

6. Schmidtko, S., Stramma, L., & Visbeck, M. Decline in global oceanic oxygen content during the past five decades. *Nature*, **542**, 335-339 (2017).

7. Breitburg, D., et al. Declining oxygen in the global ocean and coastal waters. *Science*, **359**, DOI: 10.1126/science.aam7240 (2018).

8. Jankowski, J., Livingstone, D. M., Bührer, H., Forster, R., & Niederhauser, P. Consequences of the 2003 European heat wave for lake temperature profiles, thermal stability, and

hypolimnetic oxygen depletion: Implications for a warmer world. *Limnol. Oceanogr.,* **51**, 815-819 (2006).

9. Yvon-Durocher, G., Jones, J. I., Trimmer, M., Woodward, G., & Montoya, J. M. Warming alters the metabolic balance of ecosystems. *Philos. T. R. Soc. B*, **365**, 2117-2126 (2010).

10. Seki, H., Takahashi, Y., Hara, Y., & Ichimura, S. Dynamics of dissolved oxygen during algal bloom in Lake Kasumigaura, Japan. *Water Res.*, **14**, 179-183 (1980).

11. Jacobson, P. C., Stefan, H. G., & Pereira, D. L. Coldwater fish oxythermal habitat in Minnesota lakes: influence of total phosphorus, July air temperature, and relative depth. *Can. J. Fish. Aquat. Sci.*, **67**, 2002-2013 (2010).

12. Harke, M. J., et al. A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, *Microcystis* spp. *Harmful Algae*, **54**, 4-20 (2016).

13. Vaquer-Sunyer, R., & Duarte, C. M. Thresholds of hypoxia for marine biodiversity. *Proc. Natl. Acad. Sci.*, **105**, 15452-15457 (2008).

14. Woolway, R. I., & Merchant, C. J. Worldwide alteration of lake mixing regimes in response to climate change. *Nat. Geosci.*, **12**, 271-276 (2019).

15. Livingstone, D. M. Impact of secular climate change on the thermal structure of a large temperate central European lake. *Clim. Change*, **57**, 205-225 (2003).

16. Zhang, Y., et al. Dissolved oxygen stratification and response to thermal structure and long-term climate change in a large and deep subtropical reservoir (Lake Qiandaohu, China). *Water Res.*, **75**, 249-258 (2015).

17. Bouffard, D., Ackerman, J. D., & Boegman, L. Factors affecting the development and dynamics of hypoxia in a large shallow stratified lake: hourly to seasonal patterns. *Water Resour. Res.*, **49**, 2380-2394 (2013).

18. 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).

19. Nürnberg, G. K. Trophic state of clear and colored, soft- and hardwater lakes with special consideration of nutrients, anoxia, phytoplankton and fish. *Lake Reserv. Manag.*, **12**, 432-447 (1996).

20. Ho, J. C., Michalak, A. M., & Pahlevan, N. Widespread global increase in intense lake phytoplankton blooms since the 1980s. *Nature*, **574**, 667-670 (2019).

21. Kosten, S., et al. Warmer climates boost cyanobacterial dominance in shallow lakes. *Glob. Change Biol.*, **18**, 118-126 (2012).

22. Müller, B., Bryant, L. D., Matzinger, A., & Wüest, A. Hypolimnetic oxygen depletion in eutrophic lakes. *Environ. Sci. Technol.*, **46**, 9964-9971 (2012).

23. Winslow, L. A., Leach, T. A., & Rose, K. C. Global lake response to the recent warming hiatus. *Environ. Res. Lett.*, **13**, 054005 (2018).

24. Livingstone, D. M. An example of the simultaneous occurrence of climate-driven "sawtooth" deep-water warming/cooling episodes in several Swiss lakes, *Verh. Int. Ver. Limnol.*, **26**, 822-828 (1997).

25. Williamson, C. E., et al. Ecological consequences of long-term browning in lakes. *Sci. Rep.*, **5**, DOI:10.1038/srep18666, (2015).

26. Rose, K. C., Winslow, L. A., Read, J. S., & Hansen, G. J. A. Climate-induced warming of lakes can be either amplified or suppressed by trends in water clarity. *Limnol. Oceanogr. Letters.*, **1**, 44-53 (2016).

27. Woolway, R. I., Merchant, C. J., Van Den Hoek, J., Azorin-Molina, C., Nõges, P., Laas, A., Mackay, E. B., and Jones, I. D. Northern hemisphere atmospheric stilling accelerates lake thermal responses to a warming world. *Geophys. Res. Lett.*, **46**, 11983-11992 (2019).

28. Carpenter, S. R., Stanley, E. H., & Vander Zander, M. J. State of the world's freshwater ecosystems: physical, chemical, and biological changes. *Annu. Rev. Env. Resour.*, **36**, 75-99 (2011).

**Acknowledgments:** This manuscript benefited from conversations at meetings of the Global Lake Ecological Observatory Network (GLEON; supported by funding from US NSF grants 1137327 and 1702991). SFJ and KCR acknowledge support from US NSF grants 1638704, 1754265, and 1761805 and SFJ was supported by a US Fulbright Student grant to Uppsala University, Sweden. GH acknowledges the many employees of the Minnesota Department of Natural Resources, the Minnesota Pollution Control agency, and citizen volunteers for data collection and collation. PRL acknowledges support from a NSERC Discovery Grant, the Canada Research Chair Program, Canada Foundation for Innovation, the Province of Saskatchewan, University of Regina, and Queen's University Belfast. RLN & JJ acknowledge support from the Missouri Department of Natural Resources and the Missouri Agricultural Experiment Station and many students that collected and processed reservoir samples under the leadership of Daniel V. Obrecht and Anthony P. Thorpe. RMP and CEW acknowledge support from NSF grants 1754276 and 1950170, Miami University Eminent Scholar Fund, and the Lacawac Sanctuary and Biological Field Station for access to Lake Lacawac and use of research facilities. RIW acknowledges support from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 791812. SC acknowledges support of the Castle Lake Research Program through the University of Nevada and UC Davis via Charles R. Goldman. CLD acknowledges the King County Environmental Laboratory for the long-term monitoring data for Lake Washington and Lake Sammamish. JD acknowledges support from the University of Warmia and Mazury in Olsztyn (Grant under the Senate Committee for International Cooperation financing) and staff at Department of Water Protection Engineering for long-term data collection and analysis. OE acknowledges support from the Russian Scientific Foundation (grant 19-77-30004) for Mozhaysk Reservoir. GF acknowledges support for long-term sampling of Lake Caldonazzo by the Fondazione Edmund Mach. HPG acknowledges funding for long-term sampling of Lake Stechlin by the Leibniz association and assistance by members of the IGB team. KDH acknowledges the Oklahoma Department of Wildlife Conservation, the Oklahoma Water Resources Board, the Grand River Dam Authority, the US Army Corps of Engineers, the City of Tulsa, W.M. Matthews, T. Clyde, R.M Zamor, P. Koenig, and R. West for support, assistance, and data for Lakes Texoma, Thunderbird, Grand, Eucha, and Spavinaw. JH acknowledges support from the ERDF/ESF project Biomanipulation as a tool for improving water quality of dam reservoirs (No CZ.02.1.01/0.0/0.0/16_025/0007417). BL acknowledges support from the FA-UNIMIB for long-term monitoring of Lake Iseo. EBM acknowledges support from the UK Natural Environment Research Council funding for the long-term monitoring on Blelham Tarn and the staff of the Freshwater Biological Association and UK Centre for Ecology and Hydrology for carrying out the work. AP and JAR acknowledge support from the Ontario Ministry of the Environment, Conservation and Parks for providing data from south-central Ontario lakes ("Dorset lakes") and staff and students at Ontario's Dorset Environmental Science Centre for data collection and analysis. MR acknowledges the International Commission for the Protection of Italian-Swiss Waters (CIPAIS) for funding long-term research on Lake Maggiore. EST acknowledges Prof. Lauren Chapman (McGill University) and her team for the long-term data collection in Lake Nkuruba. MS acknowledges the City of Zurich Water Supply and the cantonal agencies of the cantons of Bern (AWA, Gewässer- und Bodenschutzlabor), Zurich (AWEL), St. Gallen (AFU), and Neuchatel for providing data for the Swiss lakes, CIPEL and INRA for data from Lake Geneva, and IGKB for data from Lake Constance. WT acknowledges support from the Belgian Science Policy Office through the research project EAGLES (CD/AR/02A) on Lake Kivu. PV acknowledges support from the councils of the regions of Waikato, West Coast, and Bay of Plenty for long-term sampling of lakes Taupo, Brunner, and Tarawera. KY acknowledges support from the Clark Foundation for long-term monitoring of Otsego Lake and past and current members of SUNY Oneonta BFS for sampling. KS and JS acknowledge the National Park Service, W Gawley, Acadia National Park for providing data for Jordan Pond, Bubble Pond, and Eagle Lake in Maine. The views expressed in this article are those of the authors and do not necessarily reflect views or policies of funding agencies.

**Authorship contributions:** SFJ and KCR designed the study, compiled the data, conducted analyses, and drafted the manuscript. GJAH, BMK, PRL, JLM, RLN, RMP, JTS, CEW and RIW helped design the study and conduct analyses, contributed data, and edited the manuscript. All other authors contributed data, edited the manuscript, or both.

**Author information:** Derived statistics used in our analyses are publicly available via the Environmental Data Initiative (EDI) repository at: https://doi.org/10.6073/pasta/ac8b05bb0da19032b3df3efc21f83874.

Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing interests.

Correspondence and requests for materials should be addressed to KCR (rosek4@rpi.edu).

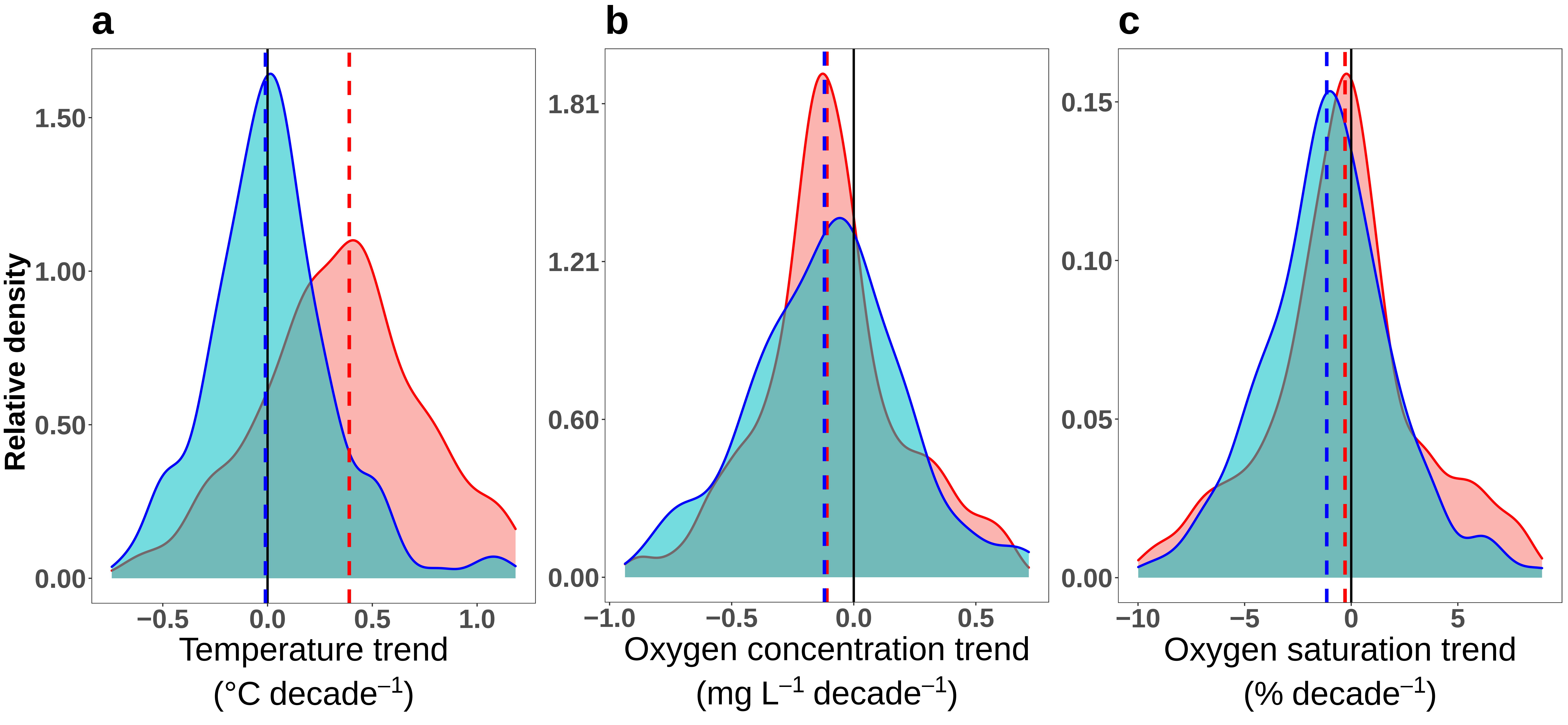
## Figures and Figure Captions:

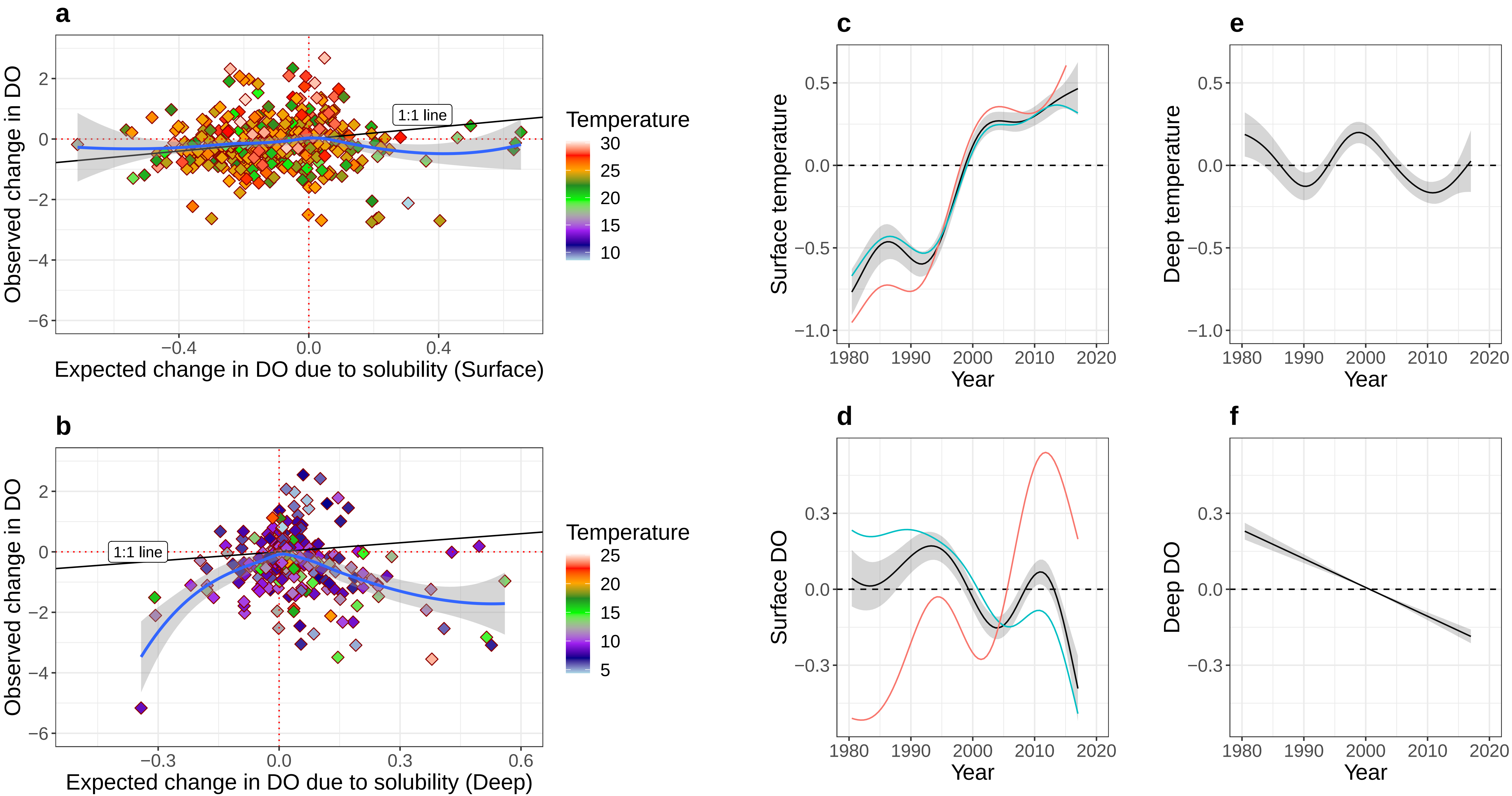
**Fig. 1 | Trends in dissolved oxygen and temperature. a-c**, Density plots of trend magnitudes for **a** temperature (°C decade$^{-1}$), **b** DO concentration (mg L$^{-1}$ decade$^{-1}$) and **c** DO percent saturation (% decade$^{-1}$). Red distribution indicates surface water trends and blue indicates deep-water trends. The x-axis range for each plot covers two standard deviations from the median, or approximately 95% of data. Vertical dashed lines indicate median trends, and the zero trend is highlighted with a black vertical line.

**Fig. 2 | Solubility effects and changes in temperature and DO concentration through time. a, b**, Observed vs. predicted change in DO concentration (mg L$^{-1}$) due to solubility for surface (**a**) and deep (**b**) waters. Solid black line is the 1:1 line and the blue line is loess smoothed, while the gray regions are 95% confidence intervals. **c-f**, Smoothed curves of GAMM models, showing deviation from the mean model predictions for selected response variables with year as the predictor variable. Gray regions represent one standard error from the predicted line for **c**, temperature (°C) and **d**, DO (mg L$^{-1}$) through time for surface waters. The red line represents lakes where both surface temperature and DO were increasing (n = 87) and the blue line is all other lakes (n = 332). **e**, Temperature and **f**, DO for deep waters.

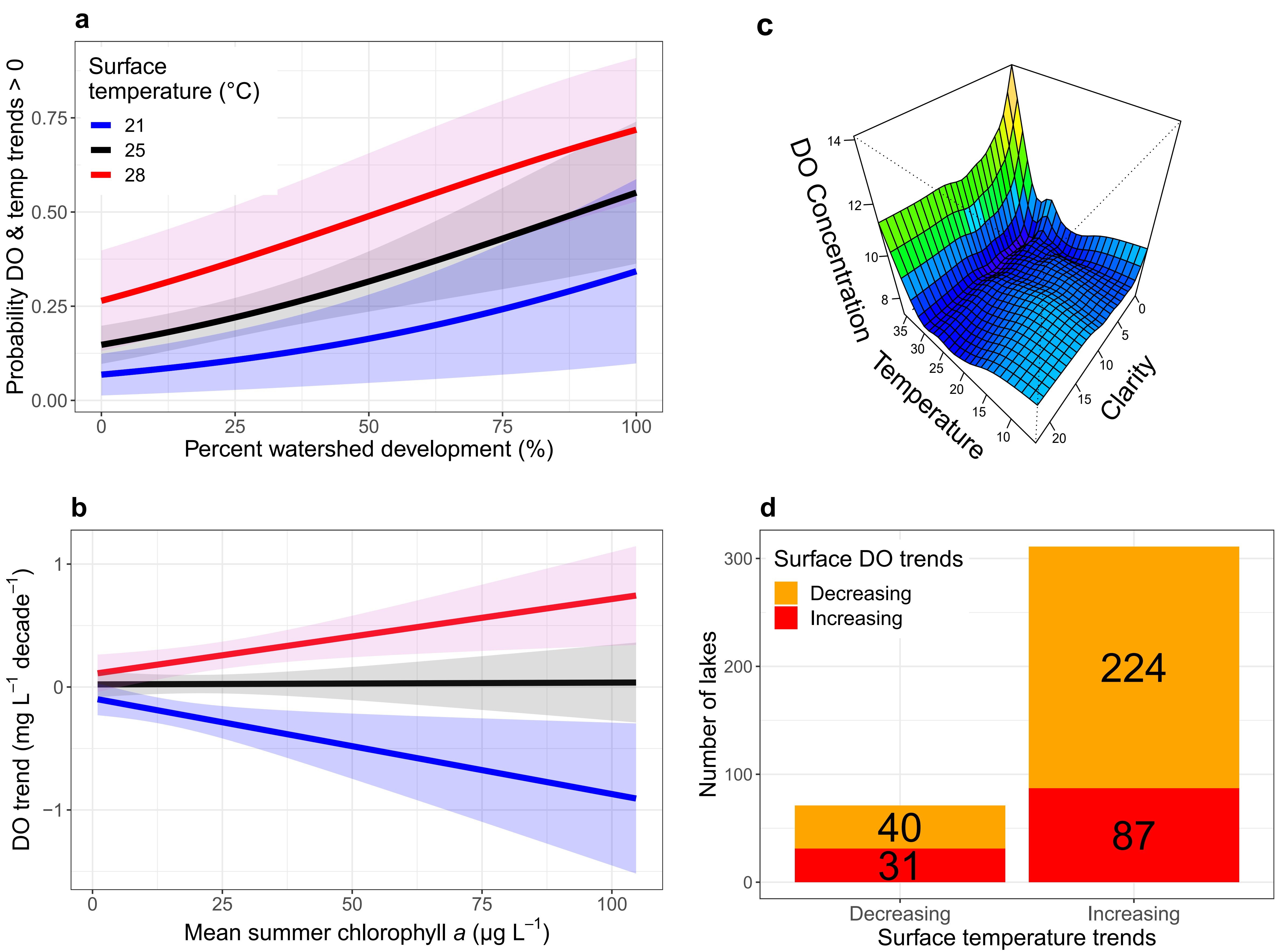
**Fig. 3 | Interaction of productivity and temperature in surface waters. a**, Predicted probability of a lake having both increasing surface temperature and DO concentration from a fitted logistic regression model at three different mean surface water temperatures: 21°C (blue), 25°C (black), 28°C (red) **b**, Predictions of DO trends from a fitted multiple regression model for chlorophyll (used as a surrogate for primary productivity) at these same temperatures (legend same as **a**) **c**, The interaction of water clarity (measured as Secchi depth in m) and surface-water temperature (°C) and their effects on surface DO (mg $L^{-1}$) from fitted generalized additive mixed models (GAMM) **d**, Most lakes exhibited increasing surface temperatures and decreasing DO concentration consistent with solubility effects, but a subset of lakes (n = 87) have both increasing surface temperature and DO concentration.

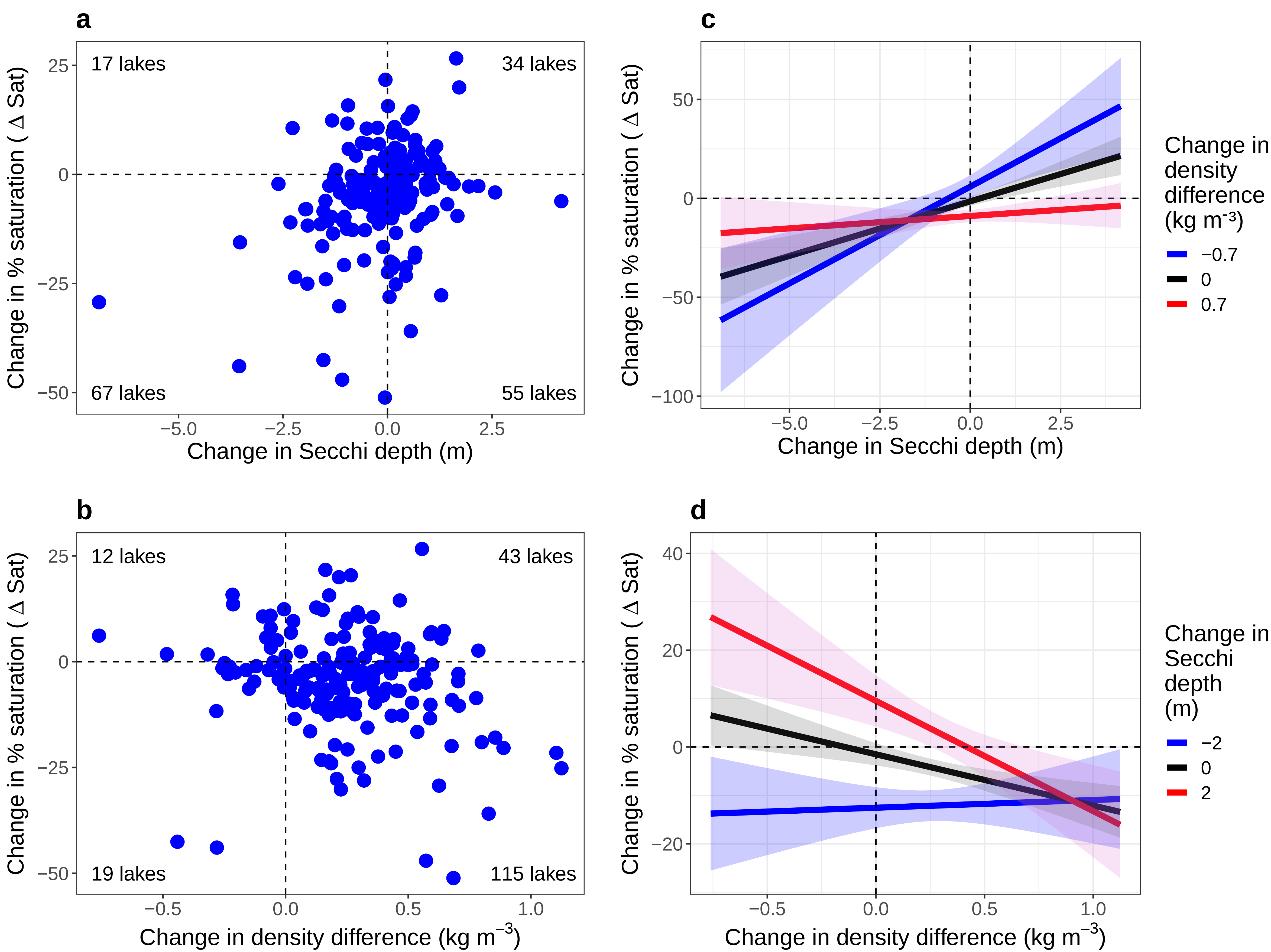
**Fig. 4 | Effect of changes in water clarity and density difference on deep-water DO saturation change.** **a**, Change in % saturation versus change in water clarity (Secchi depth). **b**, Change in % saturation versus change in water column density difference between surface and deep waters. The number of lakes in each quadrant in a and b are indicated by text. **c**, Predictions of change in % saturation from a fitted multiple regression model for change in water clarity at three density changes. **d**, Predictions of change in % saturation from a fitted multiple regression model for change in density difference at three clarity changes. Note that for both c and d the origin sits at no change in either predictor.


a
b
c
Relative density
1.50
1.00
0.50
0.00
−0.5
0.0
0.5
1.0
Temperature trend
(°C decade⁻¹)
1.81
1.21
0.60
0.00
−1.0
−0.5
0.0
0.5
Oxygen concentration trend
(mg L⁻¹ decade⁻¹)
0.15
0.10
0.05
0.00
−10
−5
0
5
Oxygen saturation trend
(% decade⁻¹)



a
Observed change in DO
Expected change in DO due to solubility (Surface)
1:1 line
Temperature
30
25
20
15
10
b
Observed change in DO
Expected change in DO due to solubility (Deep)
1:1 line
Temperature
25
20
15
10
5
c
Surface temperature
Year
d
Surface DO
Year
e
Deep temperature
Year
f
Deep DO
Year


a
Surface temperature (°C)
21
25
28
Probability DO & temp trends > 0
0.00
0.25
0.50
0.75
0
25
50
75
100
Percent watershed development (%)
c
DO Concentration
Temperature
Clarity
b
DO trend (mg L$^{-1}$ decade$^{-1}$)
1
0
−1
Mean summer chlorophyll $a$ (μg L$^{-1}$)
d
Surface DO trends
Decreasing
Increasing
Number of lakes
0
100
200
300
224
87
40
31
Decreasing
Increasing
Surface temperature trends



a
17 lakes
34 lakes
67 lakes
55 lakes
Change in % saturation ( Δ Sat)
Change in Secchi depth (m)
b
12 lakes
43 lakes
19 lakes
115 lakes
Change in % saturation ( Δ Sat)
Change in density difference (kg m−3)
c
Change in % saturation ( Δ Sat)
Change in Secchi depth (m)
Change in density difference (kg m-3)
−0.7
0
0.7
d
Change in % saturation ( Δ Sat)
Change in density difference (kg m−3)
Change in Secchi depth (m)
−2
0
2

**Supplemental information**

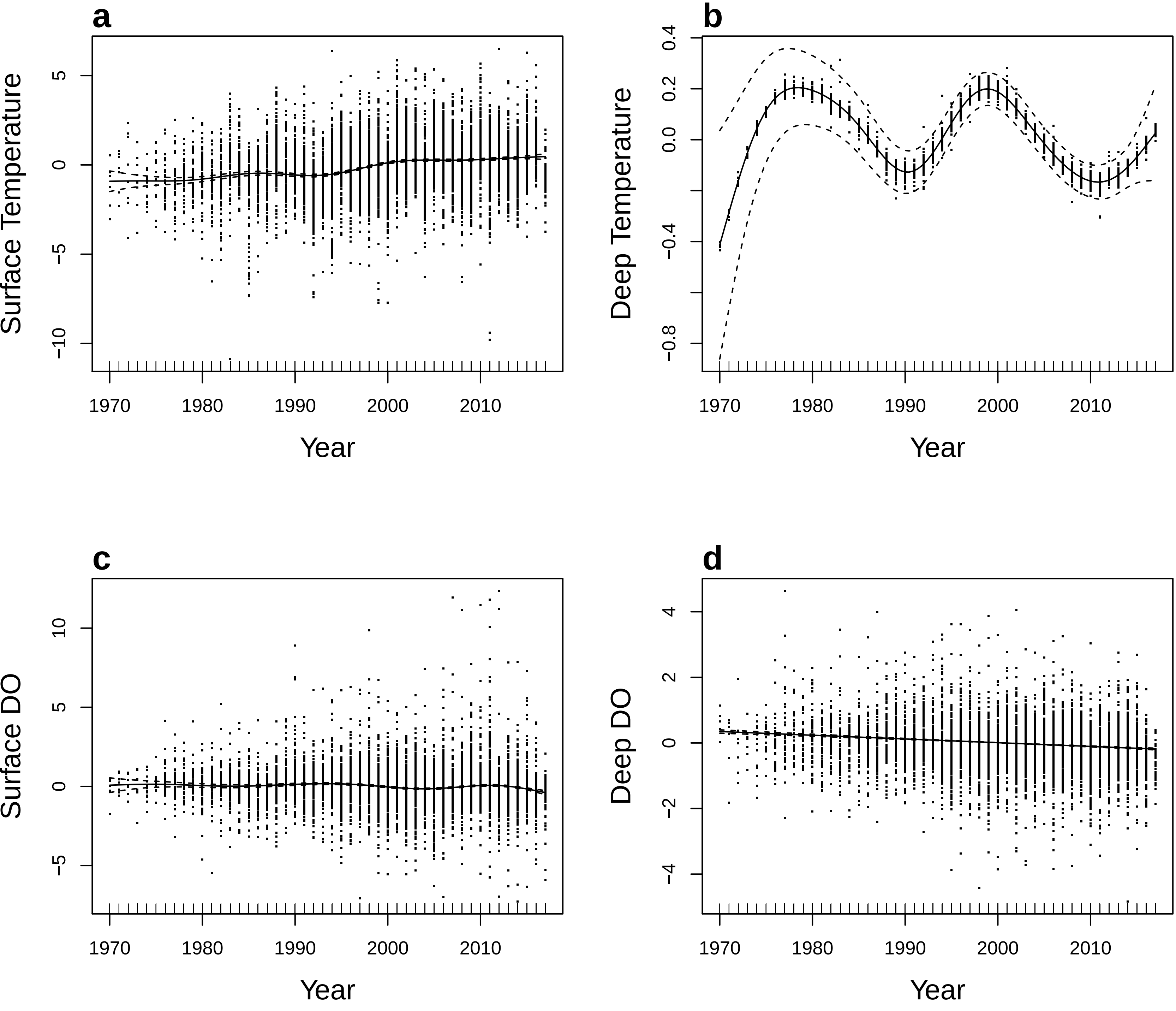
There are seven supplemental information tables and four supplemental information figures. Tables S1 and S2 are referenced in text. Table S3 describes data contributors for this project and Table S4 provides location and trend information for each lake. Trend data were not reported for a) two lakes where providers did not provide permission to publish data but that were included in trend analyses (Annecy and Geneva; 'NP' in table S4), b) lakes had less than 15 years of data at a given depth (not shown in table), or c) deep-water trends in lakes that did not thermally stratify ('NA' in table S4). In one lake (T Bird), epilimnetic water was artificially aerated and this depth layer was excluded from analysis. Table S5 presents statistics associated with spatial autocorrelation analyses. Table S6 describes trends over the entire population of lakes versus a sub-sample of lakes after accounting for the large numbers of samples obtained in lake-rich regions. Table S7 describes trends and uncertainty in trends over two time periods for subsets of lakes having data for at least 80% of years: 1980-2017 and 1990-2017. Fig. S1 presents the results of GAMM analysis of trends zoomed out to visualize distribution of residuals for surface and deep-water temperature and dissolved oxygen trends. Fig. S2 presents the partial dependency plots for the top predictors of changes in deep-water DO percent saturation as determined by a random forest analysis. Fig. S3 presents partial dependency plots for the top predictors of changes in water column density difference between surface and deep waters as determined by a random forest analysis. Fig. S4 presents the locations of lakes used in this study (n=393).

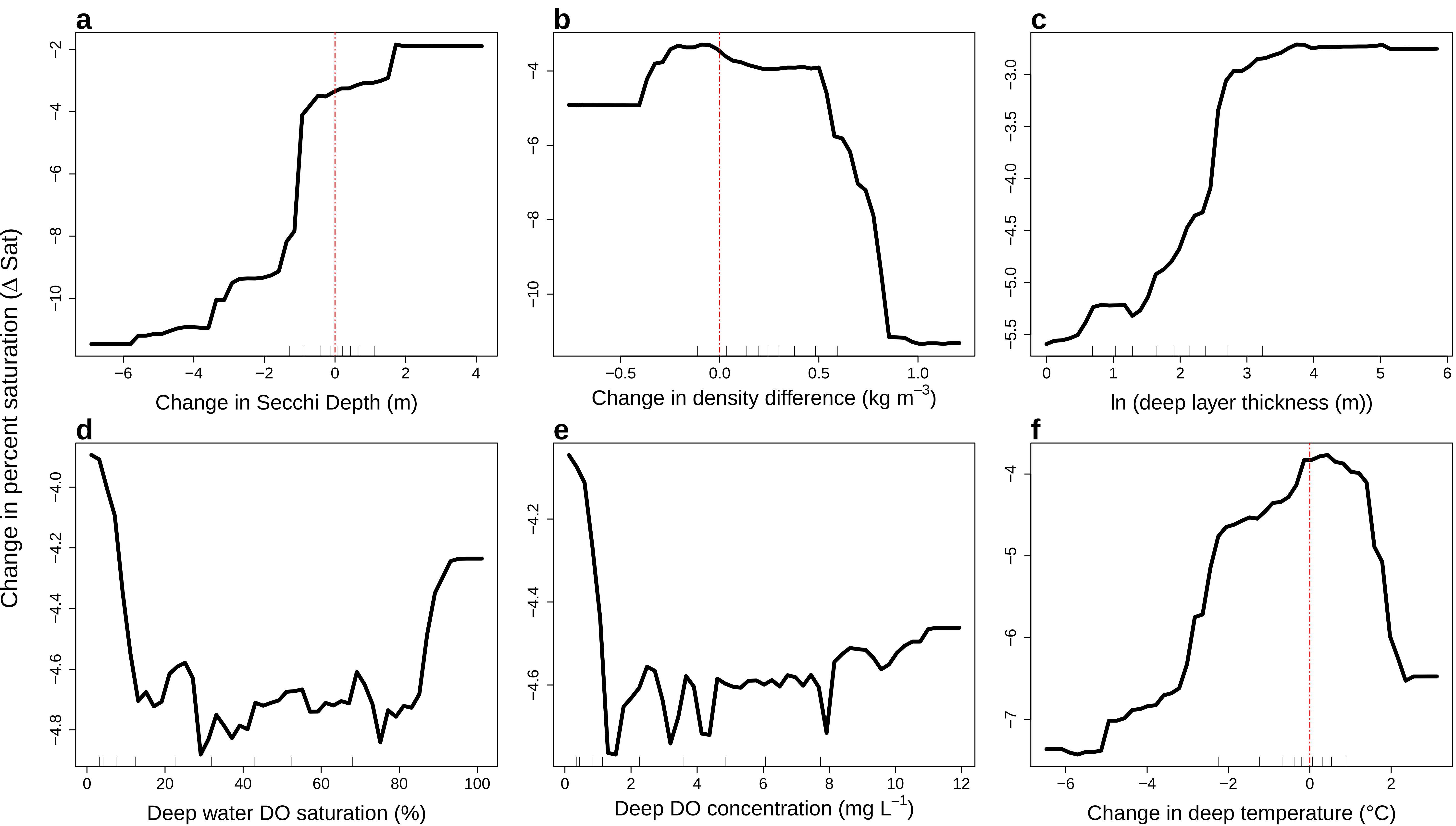
**Figure S1 |** Results of GAMM analysis of trends zoomed out to visualize distribution of residuals. **a,** Surface-water temperature (°C) **b,** Deep-water temperature (°C) **c,** Surface-water DO (mg $L^{-1}$) and **d,** Deep-water DO concentration (mg $L^{-1}$).

**Figure S2 | a-f**, Partial dependency plots from a random forest algorithm of deep-water change in % dissolved oxygen saturation (Δ Sat) in the last five years of record relative to the first five years of record for each lake. Plots are ordered by predictor variable importance, decreasing in importance from the upper left to lower right (a to f). Vertical red lines indicate zero change in predictor variable and hash marks on the x-axis indicate lake distribution deciles. Partial dependencies indicate the relationship between predictor and response variables when holding other variables at their mean value. Lakes that experienced no change in either water clarity or density difference between surface and deep waters exhibited little change in deep-water saturation (see also, Fig. 4).

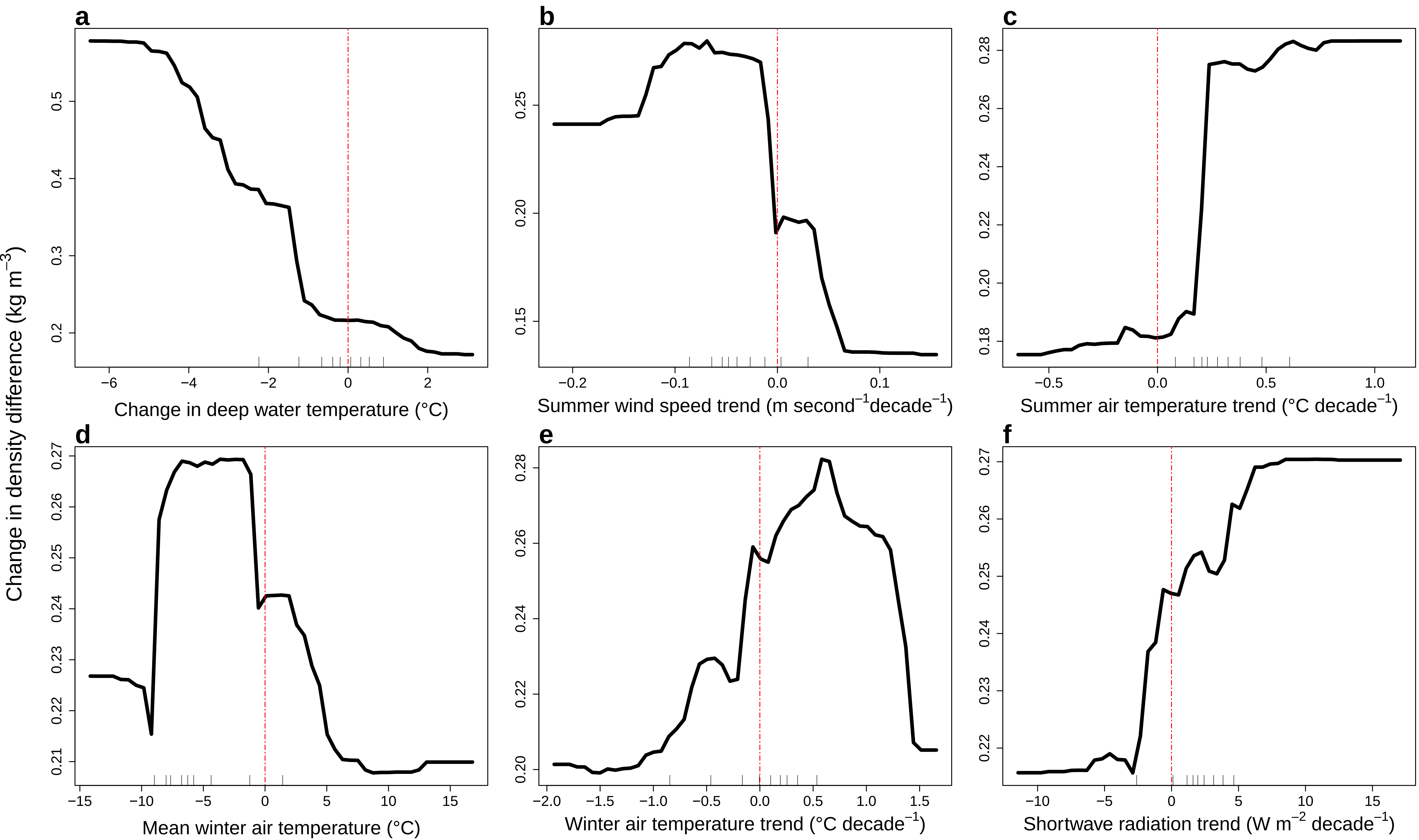
**Figure S3 |** Drivers of the change in density difference between surface and deep waters. **a-f**, Partial dependency plots from a random forest algorithm of deep-water change in water column density difference in the last five years of record relative to the first five years of record for each lake. Plots are ordered by predictor variable importance, decreasing in importance from the upper left to lower right (a to f). Vertical red lines indicate zero values for predictor variable and hash marks on the x-axis indicate lake distribution deciles. Partial dependencies indicate the relationship between predictor and response variables when holding other variables at their mean value.

**Fig. S4 |** Locations of lakes used in this study (n=393).

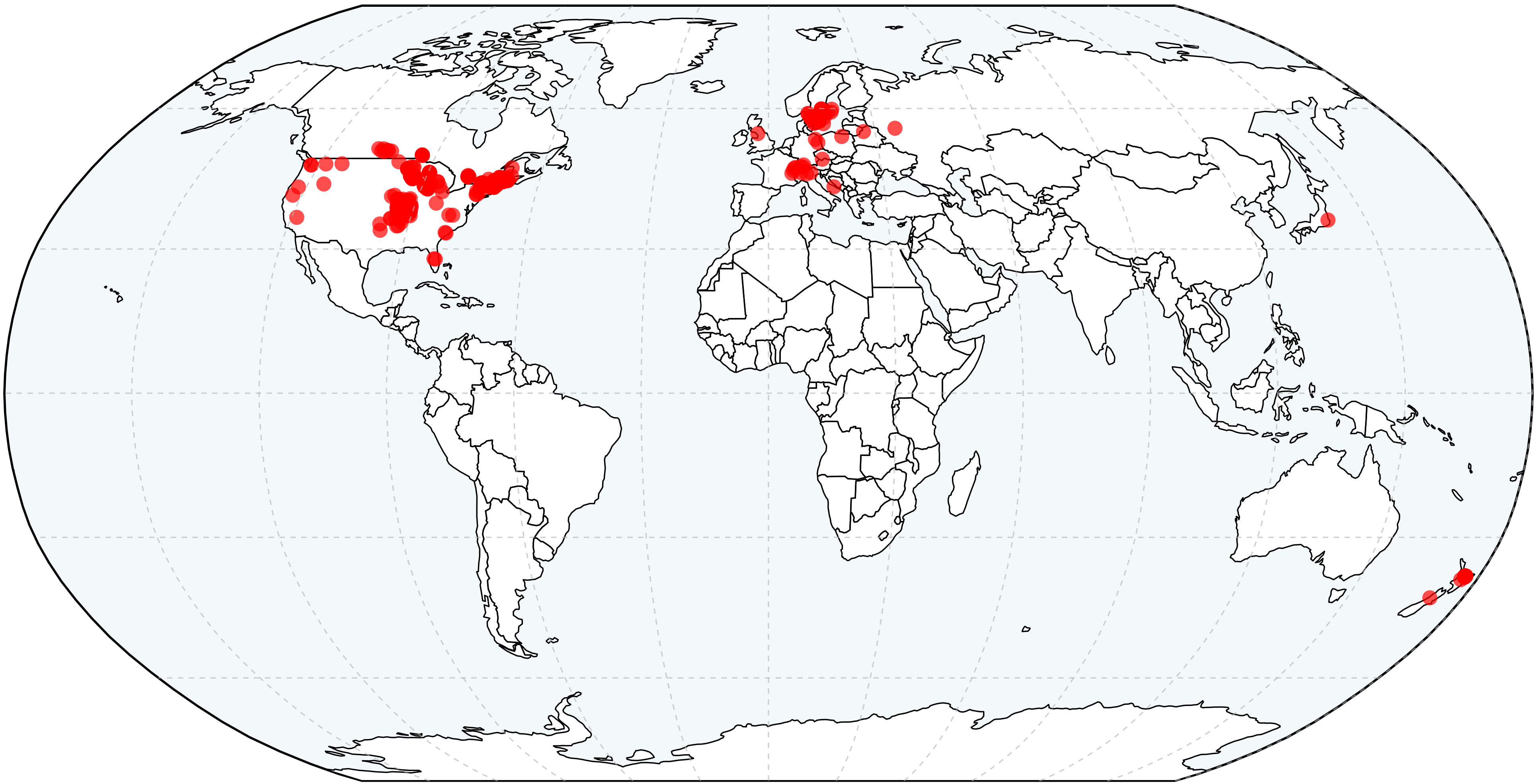

a
Surface Temperature
Year
b
Deep Temperature
Year
c
Surface DO
Year
d
Deep DO
Year
1970
1980
1990
2000
2010



Change in percent saturation (Δ Sat)
a
Change in Secchi Depth (m)
b
Change in density difference (kg m−3)
c
ln (deep layer thickness (m))
d
Deep water DO saturation (%)
e
Deep DO concentration (mg L−1)
f
Change in deep temperature (°C)



a
b
c
d
e
f
Change in density difference (kg m−3)
Change in deep water temperature (°C)
Summer wind speed trend (m second−1decade−1)
Summer air temperature trend (°C decade−1)
Mean winter air temperature (°C)
Winter air temperature trend (°C decade−1)
Shortwave radiation trend (W m−2 decade−1)



## Methods:

### Overview

Our methods here describe how we 1) compiled and quality-checked data, 2) interpolated and delineated water layer strata, and 3) statistically analyzed these data. Our statistical analyses focused on characterizing long-term trends in climate characteristics (air temperature, wind speed, precipitation, and short-wave radiation), DO concentration and saturation, water temperature, and deep-water habitat quality; identifying and characterizing potential non-linearity in DO concentration and water temperature through time; characterizing the relationship between DO concentration changes and solubility, chlorophyll, and land use; identifying the predictors of changes in deep-water DO saturation, and characterizing meteorological drivers of surface temperature trends. These methods are described in detail in the sections below.

### Data compilation and quality control

We compiled lake temperature and DO concentration water column measurements from a wide range of government, university, and not-for-profit sources (Fig. S4 and Tables S3 and S4). To assess long-term trends in temperature and DO concentration, we required profiles be made at least once annually during the peak summertime stratification (defined as the late summer period, July 15 - August 31 for northern hemisphere lakes and January 15 - February 28 for southern hemisphere lakes) offshore (e.g., nearest the deepest location in each lake) for at least 15 years. In some larger lakes (n = 6 lakes), we used profiles from two separate locations if the lake had more than one distinct basin and treated these as separate waterbodies. For some analyses other than long-term trend analyses we included lake time series data less than 15 years long, but always at least 10 years in duration (described below).

We conducted quality control on the compiled data as follows. We first removed impossible values, defined as those outside the range 0-40 for both temperature (units: °C) and DO concentration (units: mg $L^{-1}$). We then removed profiles from consideration if our initial quality control step process removed greater than 95% of the profile or if the profile had less than three distinct depth points. To reduce the potential impacts of DO measurements made when sensors sat on or in sediments, we removed the deepest measurement for individual profiles if the maximum depth for that profile exceeded the maximum depth of 90% of the remaining profiles for a given lake.

Not all profiles surveyed the entire water column. Some lakes had some profiles where the shallowest depth was greater than 0 (meaning near-surface measurements were not made), yet temperature measurements showed the nearest surface measurements were within the epilimnion. In these cases, we made the assumption of uniform DO and temperature from the surface to the shallowest measurement and added a 0 m depth point. We did this by either 1) changing the minimum depth in the profile to 0 if it was less than 0.5 m, 2) adding a 0 depth point and assigning temperature and DO values equal to that of the minimum depth point if the minimum depth point was greater than or equal to 0.5 m but less than or equal to 3 m. If the minimum depth was greater than 3 m, we excluded the profile from analyses. If there were multiple values of either temperature or DO for a given depth, the mean value at this depth was used. These operations and all further analyses were conducted in R version 3.4.2[29].

In total, the above QA steps removed 2,040 profiles out of a total of 25,023 (8.2%). After processing and removing eight non-temperate lakes, we had 22,574 DO profiles with corresponding temperature profiles. There was a median of 2.1 profiles per year (range: 1-38) and 23 years of data per lake (see also, Table S4).

**Profile interpolation and strata delineation**

In order to generate a dataset with consistent depth resolution within and among lakes, we interpolated each temperature and DO profile from depth 0 m to the deepest depth of each profile at intervals of 0.5 m using the pchip function of the R package pracma[30]. This interpolation procedure preserves the overall shape of the profile by preventing overshooting of data values[30]. Following interpolation, we calculated temperature and stability characteristics using the R package rLakeAnalyzer[31]. We delineated the epilimnion and hypolimnion using the meta.depths function (slope = 0.1, seasonal = FALSE), which calculates the top and bottom depths of the metalimnion[31]. If the range of temperatures through the profile is less than 1°C, the meta.depths function does not return values for the metalimnion (i.e., the profile is not considered stratified).

Many lakes did not have a well-defined hypolimnion. To identify those with a hypolimnion, we first removed lakes where the meta.depths function failed to calculate a bottom metalimnion depth for more than 10% of profiles. We then calculated the mean of the maximum profile depths across all profiles for each lake, to get a mean profile depth for the lake. If the mean value of the bottom of the metalimnion for a lake was shallower than the calculated mean profile depth for that lake, it was considered to have a hypolimnion. We defined “surface waters” as all depths shallower than or equal to the top metalimnetic depth and “deep waters” as all depths deeper than the bottom depth of the metalimnion.

**Characterizing trends in dissolved oxygen and temperature**

We calculated the mean surface- or deep-water temperature and DO concentration and percent saturation. For each lake, we calculated the mean of surface- or deep-water DO concentration or temperature for all profiles in a given year (in our defined late-summer period) to obtain a mean annual value. We then calculated the percent DO saturation from temperature, DO concentration, and lake elevation data[32]. Mean annual surface- and deep-water temperature and DO concentration measurements were then used to calculate long-term trends for surface waters (n = 393 lakes; median number of years per lake: 24) and deep waters (n = 260; median number of years: 24). All trends were calculated using the nonparametric Sen's slope in the R package openair[33]. For trend analysis, we only used lakes with at least 15 years of data.

For deep-water trends, lakes that were essentially anoxic (average hypolimnetic DO < 0.5 mg $L^{-1}$) had trend magnitudes that clustered near 0 relative to other lakes. This was not unexpected as lakes with essentially no hypolimnetic DO have little potential to lose additional DO. When calculating median trends and for graphical depiction of trends (Fig. 1), we removed these lakes (n = 69; difference = 191).

We conducted several analyses to examine the potential of variability in lake data through time (i.e., not all lakes sampled all years of observation) or variability in space (i.e., some regions sampled much more heavily than others) to influence overall population level trends (see following sections and Tables S5-S6).

**Spatial autocorrelation and effects of lake clusters**

Because the lakes included in this study were not uniformly dispersed over all continental land masses, we examined the potential of large numbers of lakes in relatively concentrated regions to drive overall patterns. To do this, we first examined spatial autocorrelation in trends in lake temperature and dissolved oxygen concentration using Moran's I in the R package lctools[34, 35]. This statistic ranges from -1 for data that are perfectly dispersed to +1 for data that are perfectly autocorrelated. Values near zero suggest randomly distributed data. We observed weak but significant spatial autocorrelation in some variables (Table S5; Moran's I values ranging 0.02 to 0.27).

Following this analysis, we examined the potential for the large numbers of lakes in some regions to dominate overall trends we reported. We tested for potential bias by examining trends for a subset of lakes. We identified four regions in the US with high numbers of lakes (Maine = 113 lakes, New Hampshire = 38 lakes, Missouri = 41 lakes, and Minnesota = 84 lakes). For each of these clustered regions, we randomly subsampled 10% of the lakes. After this random subsetting, we found that the overall trends are similar to the trends obtained from all lakes (see Table S6). These results demonstrate that our observed population-level trends are not driven solely by trends observed in our lake-rich regions. While our analysis focuses on temperate lakes, we obtained data from a small number of non-temperate lakes (n=8). Including these non-temperate lakes in our analysis (Table S6) did not change our overall results.

**Uncertainty estimates and temporal variation in trends**

We conducted an analysis to compare trends, confidence intervals, and significance of trends over two time periods: 1980-2017 (n = 80) and 1990-2017 (n = 197) to assess whether different lake observation years influenced the overall trends in DO concentration and temperature we observed. For each time period, we used a subset of lakes that had data for at least 80% of years within the defined time period. Following established methods[18], we calculated a yearly anomaly in temperature and dissolved oxygen for each lake as the difference between each year's observation and the long-term mean. We then averaged these anomalies across all lakes and used linear regression to calculate the slope, significance, and confidence intervals of these averaged anomalies (Table S7).

## Characterizing trends in climate characteristics

We examined trends in air temperature, total precipitation, wind speed, and shortwave radiation using the ERA-5 reanalysis from the European Centre for Medium-Range Weather Forecasts (ECMWF)[36]. This data set provides a single gridded global product with a resolution of 0.25° latitude by 0.25° longitude over the period 1979-2019 available as monthly averages (air temperature, wind speed, and shortwave radiation) or totals (precipitation). We used ECMWF time-series data from the gridded location closest to each lake and over the two-month period around when lakes were sampled (July-August for Northern hemisphere lakes, January-February for Southern hemisphere lakes). We calculated temporal trends in mean summer air temperature, mean summer wind speed, summer total precipitation, mean summer shortwave radiation, mean winter air temperature, mean spring air temperature, mean fall air temperature using the same methods we used to examine lake temperature and DO trends (see above). We then conducted a multiple regression analysis to assess which of these predictor variables (trends in air temperature, total precipitation, wind speed, or shortwave radiation) best explained surface-water temperature trends.

## Trends in climatic variables over the temperate zone

We delineated gridded latitude and longitudes at 2° intervals across the entire temperate zone over land masses only as well as over large regions, including Asia (defined by longitude ≥ 29.3°; latitude 23.5° to 60°) Europe and North America (longitude < 29.3°; latitude 23.5° to 60°), South America and western Africa (longitude < 0°; latitude ≤ -23.5° to -60°); and southern Africa, Australia, and Oceania (longitude ≥ 0°; latitude -23.5° to -60°). We then used data from the ERA-5 reanalysis (see 'Characterizing trends in climate characteristics' in Methods for details) to calculate trends in climate variables over each of these regions (Table S1).

**Multiple regression analysis of drivers of surface water temperature trends**

We conducted a multiple regression analysis of the meteorological drivers of observed surface water temperature trends. Predictors in the analysis included: summer air temperature trend, summer total precipitation trend, summer wind speed trend, summer shortwave radiation trend, winter air temperature trend, spring air temperature trend, fall air temperature trend, and mean winter temperature (as a proxy for ice cover[18]). We z-score standardized all variables to facilitate comparison of model coefficients across variables having different units[37]. We verified that multicollinearity was not a problem by checking that the variance inflation factor was well below ten for all variables[38]. We used the leaps R package to select subset models including all predictors and two-way interactions, and selected the fitted model having the lowest AIC[39]. Coefficients and p-values for the selected model appear in Table S2.

**Characterizing trends in deep-water habitat quality**

We used $T_{DO3}$[11] to quantify trends in oxythermal habitat relevant for cold-water organisms. $T_{DO3}$ represents the minimum temperature in the water column where DO concentration was greater than or equal to 3 mg $L^{-1}$ and has been used to describe habitat availability for cold-water fisheries[11]. To calculate trends in $T_{DO3}$ we excluded lakes where the DO concentration was higher than 3 mg $L^{-1}$ across all depths in all profiles. For the remaining lakes, we calculated $T_{DO3}$ for each profile. If a given profile did not have DO below 3 mg $L^{-1}$, we assigned it the minimum temperature in the profile. We then calculated an annual mean $T_{DO3}$ for the late summer period and excluded lakes that had ≤ 15 years of data. This left 369 lakes where DO went below 3 mg $L^{-1}$ at least once.

**Non-linearity in DO and temperature through time**

We conducted a generalized additive mixed model (GAMM) analysis to characterize overall response of lake temperature and DO concentration through time and to identify any non-linearity. GAMMs fit a smooth function of the predictor variables showing the relationship of the predictors to the response variable[40]. We conducted separate analyses for four response variables, surface-water temperature, surface-water DO concentration, deep-water temperature, and deep-water DO concentration. For each GAMM, our only predictor variable was the year, resulting in models that show the change in the response variable through time. We used the gamm4 function of the gamm4 package to fit these models using the default thin plate spline for smooth terms[41]. Gamm4 uses penalized regression splines of moderate rank for the smooth function. For two of these models we used a normal error distribution. Because residuals for the deep-water temperature analysis were skewed, we used a gamma distribution. Residuals in the deep-water DO analysis were also skewed, but because there were a large number of 0 values we used a Tweedie distribution instead of a gamma distribution. We limited this analysis to data from 1970 and later and included all lakes with data in the specified time period (total lake n = 419). To account for the non-independent nature of the repeated measurements through time within each individual lake, the slope and intercept were allowed to vary randomly by lake[42].

We next conducted a GAMM to understand how surface water DO concentration responded to temperature and productivity (n = 419 lakes). We used Secchi disk depth as a surrogate for productivity[19]. We included fixed effects of mean summer surface water temperature, mean Secchi depth, and the interaction of these two terms in the model. We included a random intercept and slope by year within each lake and included a corresponding year fixed effect.

**Relationship between dissolved oxygen concentration changes and solubility**

To determine the relative importance of solubility in explaining changes in DO concentration, we calculated the expected change in DO concentration due to solubility alone and compared this amount to the observed DO change. To do this, we first calculated the difference between the observed mean DO concentration across the last five years and the first five years of record for each lake, requiring a minimum of ten years of data per lake (n = 415 lakes for surface (Fig. 2a); n = 259 lakes for deep (Fig 2b)). We then calculated the expected change due solely to solubility and compared observed to expected DO changes. Specifically, we calculated the mean percent saturation in the first five years by first calculating the mean DO saturation for each water column layer (surface or deep waters) and then calculated the mean of all of these values. We then used an analogous approach to calculate mean temperature, DO concentration, and mean DO concentration at 100% saturation in the last five years of record for each lake. Once we calculated these values, we multiplied the mean DO concentration at 100% saturation by the decimal value of percent saturation in the first five years of record. This product represents the expected DO concentration if the percent saturation in the last five years of record remained the same as it was in the first five years of record. In other words, we removed the effect of temperature so that if all changes were due solely to solubility, observed changes in DO concentration would be identical to this value.

**Relationship between dissolved oxygen trends and chlorophyll**

We used multiple regression to test if chlorophyll concentration and surface-water temperature were predictors of lakes having both increasing surface DO concentration and temperature trends. We first calculated the long-term mean late-summer surface-water (epilimnetic) chlorophyll concentration, which was available for 162 lakes having at least ten years of chlorophyll measurements. We next predicted DO concentration trends using chlorophyll and mean surface-water temperature as independent variables. We first fit the linear regression models, starting with a full model that included the interaction of chlorophyll and temperature. We then fit all subset models and selected the model with the lowest AIC value[43]. Using this selected model, we predicted DO concentration trends at three different mean epilimnetic temperatures (21, 25, and 28°C) across the observed values for chlorophyll.

**Relationship between dissolved oxygen trends and land use**

We used logistic regression to better understand the drivers of increasing DO concentration in lakes with increasing surface-water temperatures, using land use/land cover data to model the probability of this phenomenon[44]. Logistic regression predicts the probability of a binary response outcome for different values of predictor variables. Predictors in our logistic regression included the percent of agriculture and developed land cover in the watershed and the mean surface-water temperature over the last ten years of record because these land use characteristics have been associated with increased growth of some phytoplankton taxa in warmer lakes[5, 21]. Our binary response was: either a lake had both increasing surface temperature and DO concentration (1) or it did not (0). We tested for all two-way interactions and all main effects. We used the National Land Cover Database 2011 to derive land cover metrics for US lakes[45]. We considered any land falling into any of the developed classes as developed (Developed – Open Space, Developed – Low Intensity, Developed – Medium Intensity, Developed – High Intensity). We tested the goodness of fit of the final model using the Hosmer-Lemeshow test, available in the ResourceSelection R package (function hoslem.test)[46]. This test showed an acceptable goodness of fit ($P = 0.166$). The final number of lakes for analysis that had both land cover data and sufficient data to calculate trends was 326.

**Identifying the predictors of changes in deep-water DO saturation**

We first used a random forest algorithm to obtain predictors of the observed change in percent saturation (i.e., drivers beyond pure solubility effects) in deep waters[47]. We used the percent increase in mean squared error as a measure of predictor variable importance. We conducted the random forest algorithm analysis using the randomForest package[48]. For each analysis, we only used lakes that had no missing values for any of the predictor variables (n = 224 lakes).

For the random forest algorithm, the response variable was the change in mean DO percent saturation in the last five years of record relative to the first five years of record for each lake (Δ Sat). A positive Δ Sat indicated an increase in percent saturation while a negative Δ Sat indicated a decrease in percent saturation. Predictor variables included mean hypolimnetic DO percent saturation, DO concentration, temperature, and thickness of the hypolimnion (ln transformed), mean Secchi depth, ln of mean lake depth, log10 of residence time, change in hypolimnetic thickness, change in hypolimnetic temperature, change in Secchi depth, and change in the density difference between surface and deep waters. Mean lake depth and residence time were obtained from the HydroLakes Database[49]. We calculated the density difference across the water column using rLakeAnalyzer to calculate densities for each interpolated depth point in each water column profile[31]. If a given profile was stratified, we then used the mean epilimnetic density and the mean hypolimnetic density and calculated the difference between these densities. If a given profile was not stratified, we took the mean density across the top two meters and the mean density across the bottom two meters and calculated the difference between these densities. We also included trends in the following ERA-5 meteorological variables: summer, fall, and winter air temperature, summer shortwave radiation, and summer wind speed. Finally we included mean winter air temperature as a proxy for ice cover[18].

Following the above analysis, change in the density difference between surface and deep waters came out as an important predictor. Although this could be explained by increased surface water temperatures driven by meteorological variables, it is possible that other changes, such as water clarity[25], could also explain changes in density difference. To disentangle the drivers of changes in water column density differences, we conducted another RF using the same predictor variables as the above analysis but changing the response variable to the change in the density difference. We did not include the response variable from the first analysis (Δ Sat). The six most important variables are presented in Fig. S3.

Based on results of the RF analysis, we conducted a multiple regression analysis to predict change in percent saturation (Δ Sat) for different levels of predictor variables (ln of mean lake depth, change in the density difference across the water column, and change in Secchi depth). We used a subset of lakes where mean deep-water DO concentration exceeded 0.5 mg/L to avoid lakes with little potential to lose DO. Predictor variables were selected because they were the three most important variables identified by RF, except we substituted ln mean lake depth for ln deep layer thickness. This substitution was made because models using ln of deep layer thickness demonstrated substantial non-linearity in plots of residuals against fitted values. Models built with ln mean lake depth greatly improved these patterns and these two variables were correlated ($r = 0.51$). We first fit the multiple regression models starting with a full model that included all predictors and two-way interaction terms. We then fit all subset models and selected the model with the lowest AIC value[43]. Using this selected model, we predicted Δ Sat at three different values of each of the two predictors change in Secchi depth ($P < 0.001$) and change in water column density difference ($P < 0.001$), with ln mean lake depth held at the median value.

**Data Availability:**

Many of the datasets analyzed during this study are publicly available on-line and associated links can be found in supplementary Table S3. Derived statistics are publicly available via the Environmental Data Initiative (EDI) repository at https://doi.org/10.6073/pasta/ac8b05bb0da19032b3df3efc21f83874. Most lakes are included here, but we note that due to the collaborative nature of this project and a wide range of data provenance, it was not possible to include every lake in this repository. Data not otherwise already publicly available are available upon request from the corresponding author pending permission from the appropriate data provider.

**References:**


29. R Core Team. R: a language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria (2017).

30. Borchers, H. W. pracma: Practical Numerical Math Functions. R package version 2.1.5 https://CRAN.R-project.org/package=pracma (2018).

31. Winslow, L. A., et al. rLakeAnalyzer: Lake Physics Tools. R package version 1.11.4. https://CRAN.R-project.org/package=rLakeAnalyzer (2017).

32. Winslow, L. A., et al. LakeMetabolizer: An R package for estimating lake metabolism from free-water oxygen using diverse statistical models. *Inland Waters*, **6**, 622-636 (2016).

33. Carslaw, D. C., & Ropkins, K. Openair – An R package for air quality data analysis. *Environ. Model. Softw.*, 27-28, 52-61 (2012).

34. Moran, P. A. P. The interpretation of statistical maps. *J. Roy. Stat. Soc. B Met.*, **10**, 243-251 (1948).

35. Kalogirou, S. lctools: Local Correlation, Spatial Inequalities, Geographically Weighted Regression and Other Tools. R package version 0.2-7. https://CRAN.R-project.org/package=lctools (2019).

36. Copernicus Climate Change Service (C3S). ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climage Change Service Climate Data Store (CDS), Accessed 10/1/2019. https://cds.climate.copernicus.eu/cdsapp#!/home

37. Gelman, G., & Hill, J. *Data Analysis Using Regression and Multilevel/Hierarchical Models.* Cambridge University Press, New York (2007).

38. Quinn, G. P., & Keough, M. J. *Experimental Design and Data Analysis for Biologists.* Cambridge University Press, Cambridge, U. K. (2002).

39. Lumley, T. leaps: Regression Subset Selection (based on Fortran code by Alan Miller). R package version 3.1. https://CRAN.R-project.org/package=leaps (2020).

40. Wood, S. N. *Generalized Additive Models: An Introduction with R* (2nd edition). CRC Press. Boca Raton, FL (2017).

41. Wood, S., & Scheipl, F. gamm4: Generalized Additive Mixed Models using 'mgcv' and 'lme4'. R package version 0.2-5. https://CRAN.R-project.org/package=gamm4 (2017).

42. Pinheiro, J. C., & Bates, D. M. *Mixed Effects Models in S and S-Plus*. Springer, New York (2000).

43. Burnham, K. P., Anderson, D. R., & Huyvaert, K. P. AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behav. Ecol. Sociobiol.*, **65**, 23-35 (2011).

44. Hosmer, D. W., & Lemeshow, S. *Applied Logistic Regression* ($2^{nd}$ edition). John Wiley and Sons, Inc., New York (2000).

45. Homer, C. G., et al. Completion of the 2011 National Land Cover Database for the conterminous United States – Representing a decade of land cover change information. *Photogramm. Eng. Remote Sensing*, **81**, 345-354 (2015).

46. Lele, S. R., Keim, J. L., & Solymos, P. ResourceSelection: Resource Selection (Probability) Functions for Use-Availability Data. R package version 0.3-2. https://CRAN.R-project.org/package=ResourceSelection (2017).

47. Cutler, D. R., et al. Random forests for classification in ecology. *Ecology*, **88**, 2783-2792, (2007).

48. Liaw, A., & Wiener, M. Classification and regression by randomForest. *R News*, **2**, 18-22 (2002).

49. 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**, 13603, doi:10.1038/ncomms1360 (2016).