

**Thank you very much to all three reviewers for their thoughtful review and commentary. It has lead to several substantial improvements to the manuscript. We have responded to each of your comments in bold below.**

Reviewer 1:

This study reveals the influence of background climate oscillations such as the El Nino Southern Oscillation, and the North Atlantic Oscillation on worldwide lake level changes (117 large lakes). The authors compared the lake level trends before and after removing of background climate variation. The idea of this study is novel. However, this manuscript still needs improve its quality for publications. In addition, I am a remote sensing hydrology scientist, but not a hydroclimatology scientist. The most content of this manuscript is about climatic effects on lake level changes. I suggest to add a review from hydroclimatology science for objective evaluation of this manuscript, and for some comments to improve the quality of this study.

My very limited comments are as follows:

1) The method of this study can be displayed by a tree-structure such as Figure 3a in <http://dx.doi.org/10.1002/2015GL066235> ?

**We welcomed the suggestion of including a regression tree in our analysis. We opted for a boosted regression tree analysis because it tends to outperform simple regression trees according to a variety of performance metrics. We decided that this additional analysis constituted a substantial diversion from the main aim of our manuscript so we have decided to cover it a separate follow up manuscript rather than include it here.**

2) Figures 1 and 2 can be combined together, as Figure 1 shows the lake level characters and Figure 2 shows the corresponding p-value?

**We are grateful to the reviewer for this suggestion. We have combined Figures 1 and 2 into the same figure. We agree that the content of Figures 1 and 2 were related enough to justify their merging.**

3) Figure color: it is usually a warm color (red) for water loss and cold color (blue) for water gain. In addition, a title in head of each figure could be removed? You can use such as "a) PCENSO" for Figure 4a in upper left, and others are similar.

**Thanks to the reviewer for these suggestions. We reversed the colour schemes so that red stands for water loss and blue stands for water gain. This improved colour scheme is more intuitive and more in agreement with norms in the literature. We also removed the title heads on each of the figure panels and replaced them with over-plot panel labels as suggested. This change reduces the clutter in the figures.**

4) L20: "The water level of a lake is an integrative indicator of local and regional hydrology." Two related citations are suggested here:

<http://dx.doi.org/10.1016/j.rse.2011.03.005>; <http://dx.doi.org/10.1007/s10712-016-9362-6>

**Thanks for these suggested citations, we enjoyed reading them and have added citations to them in the manuscript.**

5) L25: 1 meter to 1 m.

**Thanks for this suggestion. We have corrected it.**

6) L30: “On average, water levels in the world’s large lakes are increasing but are highly variable from lake to lake.” This conclusion is not consistent with a recent paper “Recent global decline in endorheic basin water storages” in Nature, <https://doi.org/10.1038/s41561-018-0265-7>. Could you explain which is right and reason?

**The difference can be reconciled by differences in the methods used. The recent article in Nature Geoscience used shorter time series (only 14 years), analysed terrestrial water storage generally as opposed to lake water storage specifically, and are based on water volumes not water levels. We have added a brief discussion of this topic to the manuscript and cited the recent paper that was mentioned.**

Reviewer 2:

Page 2, Ln 25-27: Please double check this calculation. The total area of the eight East Africa Great Lakes is  $\sim 152410 \text{ km}^2$ . The water volume increase associated with 1 meter of water level increase is  $\sim 152 \text{ km}^3$ .

**We have double checked this calculation and found it to be correct. The line states that “water levels in the eight East African Great Lakes went up by more than 1 m.” The actual increases are listed below:**

**Lake Tanganyika: Surface area (32000 km<sup>2</sup>) \* Water level increase (1.7 m) = 54.4 km<sup>3</sup>**  
**Lake Victoria: Surface area (69000 km<sup>2</sup>) \* Water level increase (1.4 m) = 96.6 km<sup>3</sup>**  
**Lake Mweru: Surface area (4500 km<sup>2</sup>) \* Water level increase (2.6 m) = 11.7 km<sup>3</sup>**  
**Lake Kyoga: Surface area (1700 km<sup>2</sup>) \* Water level increase (1.9 m) = 3.23 km<sup>3</sup>**  
**Lake Turkana: Surface area (6750 km<sup>2</sup>) \* Water level increase (3.8 m) = 25.65 km<sup>3</sup>**  
**Lake Rukwa: Surface area (3050 km<sup>2</sup>) \* Water level increase (2.4 m) = 7.32 km<sup>3</sup>**  
**Lake Malawi: Surface area (30000 km<sup>2</sup>) \* Water level increase (1.2 m) = 36.0 km<sup>3</sup>**  
**Lake Kariba: Surface area (5200 km<sup>2</sup>) \* Water level increase (5.9 m) = 30.68 km<sup>3</sup>**

**Total = 265.58 km<sup>3</sup>**

**This total is in agreement with the total offered in the manuscript.**

Page 4, Ln 34: “Collinear” is a too strong word here. I would use “correlated”.

**Thanks for this clarifying suggestion. We have replaced the word “collinear” with “correlated.”**

Page 5, Ln 30-32: Altimeter observed water level changes for global major lakes are also available from the CNES Hydroweb (<http://hydroweb.theia-land.fr/>). There are notably large differences between the G-REALM and Hydroweb solutions. The G-REALM solutions appear to show larger biases (based on preliminary comparisons in Caspian Sea and Lake Victoria). This is not to ask the authors to redo the analysis using the Hydroweb solutions, but to remind them the potential issues with the G-REALM solutions.

**Thanks for this reminder of the potential issues with the G-REALM solutions. Despite the trade-offs as mentioned by the reviewer, we decided to use G-REALM data because it covers more lakes over longer time periods.**

It will be helpful to show a distribution map of the 117 lakes. The current Figure 3 does not really suit the purpose.

**Thanks to the reviewer for this suggestion. We are not sure what a distribution map would show which is not already shown in Figure 3 but would be happy to include one regardless in the supplement to our revision.**

Page 6, Ln 6: water level altimetry -> altimetry water level observations

**Thanks for this suggestion. It has been changed as requested by the reviewer.**

Page 6, Ln 29-33: The authors claimed that they calculated a complete correlation matrix between each PC and all of the 37 major climate indices recognized by NOAA's Earth System Research Laboratory. Some of the indices (e.g., EA/WR and TPI/IPO) from the ESRL (<https://www.esrl.noaa.gov/psd/data/climateindices/list/>) do not cover the recent time spans. How did the authors deal with those indices?

**For the indices that were not calculated continuously to the present, we calculated correlation coefficients over the longest timespan that was available for each PC and its paired index. This has now been clarified in the methods section with the following text:**

**“For indices that aren't updated to the present, we calculated the correlation over the longest time period over which each major climate index was available.”**

Reviewer 3:

The manuscript analyses the water level variations for 117 globally-distributed lakes and surface air temperature in order to disentangle the effects of climate variation and anthropogenic activities on lake variations. The authors use boosted regression trees (BRTs) to model water level as a function of time and the PCs of temperature. Then those most influential PCs are correlated with climate indices. The topic is highly relevant and will provide important contributions to current climate and anthropogenic impacts on inland water bodies, e.g. lakes.

**We are grateful for these comments summarizing the manuscript.**

General Comments

In general, the manuscript is generally well organized and easy to read. But the method section is a bit redundant and needs to be clarified. The main assumption is that the pattern of global lake changes is somehow driven by global patterns of earth surface air temperature. But the results (27% of water level variation was associated with background climate variation) do not support this assumption very well. In my opinion, a better re-designation of the study objectives is needed.

**Thanks to the reviewer for this suggestion on how to clarify and improve the methods section. 27% of the variability was, in fact, attributable to the the PCs representing de-seasonalized and de-trended global variation in earth surface temperatures. However, the seasonal component of the model explained an additional 5%, and the long-term trends component explained an additional 13% of the variation. So, in sum, we were able to explain 45% of the variation in water levels on average across lakes. In this context, the importance of the PCs is more apparent because they explained more than 5 times the amount of lake level variation than background seasonality (27% > 5 x 5%). This has now been clarified in the results section and in the abstract with the following text:**

**In the abstract: “On average, 27% of water level variation in individual lakes was associated with background climate variation with an additional 18% explained by seasonal variation and the long-term trend.”**

**In the results section: “Together, the PCs selected for each lake explained an average of 27% of the variance in water levels (interquartile range: 4-44%) compared to 5% explained by seasonality and 13% explained by the long-term trend.”**

The authors stressed that a novel statistic method BRT is used, but in my opinion, more details are needed. On top of that, some justifications and clarifications are needed as pointed out below. I hope my comments are useful for authors to improve the manuscript.

**Thanks very much for the comments below on how to further clarify the methods section. Our responses are detailed below.**

Specific comments

P4L6: “. . . based on average annual water level. . .” why and how is annual water level used?

**Thanks for pointing out the need to clarify here. We have added the following text to the manuscript: methods section:**

**“Annual water levels were used in the trend analysis instead of raw water levels so that the trend residuals would not be serially autocorrelated.”**

P4L13-21: Does this consider the nonstationary of time series induced by climate?

**Yes it does. All effects in the model including the effects of the PCs and the effects of season are allowed to change from year to year. For example, the effect of PCenso on water level variation might be very important in driving water levels in some years for a specific lake but not in others. We have added the following text to the manuscript: methods section:**

**“The main advantages of BRTs over other statistical models is that they have higher predictive performance, do not require data transformation or outlier elimination, automatically handle complex nonlinear relationships and interactions, and allow for many types of predictor variables and partial missing data. Through these interactions, the BRT allows for non-stationarity of the timeseries (e.g. the effect of each PC is allowed to change over time).”**

P4L15: Here the climate indices should be renamed or clarified to be distinguishable from those called climate index, e.g. ENSO, IOD, NAO, etc. It is a bit confusing.

**Thanks for pointing out the need to clarify here. We have added the following text to the methods section:**

**“In cases where a PC is highly correlated to one of the major climate indices (PCs), we renamed it with a subscript (e.g., PC<sub>ENSO</sub>) to facilitate interpretation. In all remaining cases, PCs were named with a numeric subscript which matched their order in the PCA (e.g. “PC1”)**

P4L22-31: Why are the PCs detrended instead of temperature detrended? How can detrended PCs remove anthropogenic effect?

**Thanks for this insightful comment. In our revision, we detrended the temperature data prior to the PCA as suggested by the reviewer. Detrending the temperature data also removed the long-term trends from the PCs. We would also add that detrending in this way as suggested by the reviewer predominantly removed the effect of anthropogenic climate change. However, the broader “anthropogenic effect” which includes water extractions and reservoir construction have not been removed from the PCs. We have now clarified this throughout our revision of the methods section and the discussion section.**

P4-P5: I think the overview of method should be simplified and move some of the description to specific methodologies.

**Thanks for this suggestion. In our revision, we have moved much of the text from the “overview” section to the subsequent relevant topic sections.**

P6, 2.3: Regarding PCA, a bit more details maybe help to reproduce the results presented in the manuscript. Why 100 PCs are used? Personally, I think those beyond 20 are very weak and probably mainly noises. Is a constant number of PCs used in the BRT for all lakes? I think that not all PCs are statistically significant in the regression. Does BRT take this into consideration?

**We used an excess number of PCs in the model fitting to ensure that we captured all of the important background climate drivers. But with that said, in our revision, we removed all PCs which were indistinguishable from random noise using a Box-Ljung test ( $\alpha = 0.01$ ). Based on the Box-Ljung test, we eliminated 5 PCs from our BRT model selection procedure. We also include a rigorous evaluation of the importance of each of the PCs in the model through a model selection process. Through the model selection process, if the inclusion of a PC in the BRT did not lead to an improvement in the model’s performance during cross validation, it was removed from the model. This is a more rigorous way to exclude variables than exclusion based on “statistical significance.”**

P6L3: “the longest and highest resolution time series”. How long are they? Do they vary for different lakes? What does it mean when you say “highest resolution”? Be specific.

**We are grateful to the reviewer for bringing up these needed clarifications. In our revision, we have updated the time series for each lake so that it begins in 1992 and ends in 2020. We have included the exact number of observations for each lake in the supplementary Table S3. By “longest” we mean greater than 27 years, and by “highest resolution,” we mean, “the greatest number of observations per unit time.” This has been clarified in the line referenced by the reviewer. In our revision, the sentence reads,**

**“The lakes in this study comprise the 117 lakes with the longest (>28 years) and highest temporal resolution time series (greatest number of samples per year).”**

P6L5: “. . . is typically ~3 cm for large lakes. . .”. This is a bit overstated. The best case is around 3 cm. Please double check and revise.

**Thanks for this comment. This has been updated to “~5 cm for large lakes” based on the USDA G-REALM website.**

P6L9: “. . . linearly-interpolated each lake’s time series. . .”. Why not use a model considering annual and semi-annual variations given that you have quite long time series? I suggest using the approach used in previous studies, e.g. (Kleinherenbrink et al., 2015; Villadsen et al., 2014).

**Thanks for this suggestion. In our revision, we have interpolated the time series using ARMA time series modelling with a Kalman filter. The reviewer is correct that this is a much more robust and accepted way to interpolate the time series.**

P7 2.4: The description of statistical method, BRT should be expanded instead of just saying some R packages are used given that you state this method is novel in your study.

**We are grateful to the reviewer for pointing out this key area for improvement. In our revision we have included the following description of BRTs:**

**“A BRT is an ensemble machine learning approach that differs from conventional statistical techniques which use a single parsimonious model. Instead, BRTs combine the strengths of standard regression trees and boosting—a method for aggregating many models to improve the predictive capacity. The main advantages of BRTs over other statistical models is that they have higher predictive performance, do not require data transformation or outlier elimination, automatically handle complex nonlinear relationships and interactions, and allow for many types of predictor variables and partial missing data.”**

P9L26: Lake water level definitely is an integrative metrics of regional water budget. I think this sentence should be rephrased.

**This sentence has been rephrased. Our revised sentence states, “This contrast highlights how global analyses of waterbody surface level variation can enhance our ability to detect hydrological changes.”**

P9L26: “significant effects . . . in most lakes”, is this overstated?

**Thanks for pointing out our misuse of the term “significant.” We have replaced the phrase, “had significant effects on ” with the phrase “substantially influenced.”**

P11L28-29: As pointed above, I doubt that detrending the PCs helps to separate background climate indices from ongoing climate change. Instead, I would first detrend the time series of temperature, and then perform PCA.

**In our revised manuscript, we have detrended the temperature data prior to the PCA as suggested by the reviewer.**

#### References

Kleinherenbrink, M., Lindenbergh, R. C. and Ditmar, P. G.: Monitoring of lake level changes on the Tibetan Plateau and Tian Shan by retracking Cryosat SARIn waveforms, *J. Hydrol.*, 521, 119–131, doi:10.1016/j.jhydrol.2014.11.063, 2015.



Villadsen, H., Andersen, O. B. and Stenseng, L.: Annual cycle in lakes and rivers from CryoSat-2 altimetry - The Brahmaputra river, *Int. Geosci. Remote Sens. Symp.*, 894–897, doi:10.1109/IGARSS.2014.6946569, 2014.

## **Worldwide lake level trends and responses to background climate variation**

**Benjamin M. Kraemer<sup>1,3</sup>, Anton Seimon<sup>2</sup>, Rita Adrian<sup>1</sup>, Peter B. McIntyre<sup>3,4</sup>**

<sup>1</sup>Ecosystem Research Department, IGB Leibniz Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany

<sup>2</sup>Department of Geography and Planning, Appalachian State University, Boone, NC, USA

<sup>3</sup>Center for Limnology, University of Wisconsin–Madison, Madison, WI, USA

<sup>4</sup>Department of Natural Resources, Cornell University, Ithaca, NY, USA

*Correspondence to:* Benjamin M. Kraemer (ben.m.kraemer@gmail.com)

**Abstract.** Lakes provide many important benefits to society including drinking water, flood attenuation, nutrition, and recreation. Anthropogenic environmental changes may affect these benefits by altering lake water levels. However, background climate oscillations such as the El Niño Southern Oscillation, and the North Atlantic Oscillation can obscure long-term trends in water levels, creating uncertainty over the strength and ubiquity of anthropogenic effects on lakes. Here we account for the effects of background climate variation and test for long-term (1992-2019) trends in water levels in 17200 globally-distributed large lakes using satellite altimetry data. On average, 27% of the median percentage of water level variation in individual lakes was associated with background climate variation was 58% with an additional 10% explained by seasonal variation and 25% by the long-term trend. The relative influence of specific axes of background climate variation on water levels varied substantially across and within regions. After removing the effects of background climate variation on water levels, long-term water level trend estimates were lower ( $+1.08$  cm year<sup>-1</sup>) than calculated from raw water level data ( $+1.42$  cm year<sup>-1</sup>). However, the trends became more statistically significant in 76.86% of lakes after removing the effects of background climate variation (the median p-value of trends changed from 0.1216 to 0.02). Thus, robust tests for long-term trends in lake water levels which may or may not be anthropogenic will require prior isolation and removal of the effects of background climate variation. Our findings suggest that background climate variation often masks long-term trends in environmental variables, but can be accounted for through more comprehensive statistical analyses.

## 1 Introduction:

The water level of a lake is an integrative indicator of local and regional hydrology. By extension, variation in lake water levels through time captures the dynamic nature of the water cycle, particularly when coherent patterns are observed across many lakes (Molinos et al., 2015; Crétaux et al., 2016; Molinos et al., 2015; Zhang et al., 2011). Water level variation is often associated with oscillatory dynamics in Earth's hydroclimate such as the El Niño Southern Oscillation (ENSO) (Ghanbari and Bravo, 2008; Stager et al., 2007), the Pacific Decadal Oscillation (PDO) (Benson et al., 2003; Wang et al., 2010), North Atlantic Oscillation (NAO) (Benson et al., 1998), and the Indian Ocean Dipole (IOD) (Marchant et al., 2007). For instance, water levels in multiple lakes in Poland can increase by 20 cm or more during the positive phase of the NAO (Wrzesiński et al., 2018). When multiple axes of background climate variation overlap, the effects can be even more intense. For instance, due to strong oceanic temperature anomalies in the late 1990's associated with both the ENSO and the IOD, water levels in eight East African Great Lakes went up by more than 1 meter in less than a year (Mercier et al., 2002). This constituted a combined increase in water storage of more than 266 km<sup>3</sup>—more than half the volume of Lake Erie.

Human activities can also directly affect lake water levels, creating variation that is independent of background climate dynamics (Aladin et al., 2009; Pekel et al., 2016; Rodell et al., 2018; Tao et al., 2015) (Aladin et al., 2009; Pekel et al., 2016; Rodell et al., 2018; Tao et al., 2015). For instance, in the 1970's, the two main inflowing rivers to the Aral Sea were diverted in an attempt to irrigate cotton plantations in Central Asian deserts (Aladin et al., 2009; Micklin, 1988). As a result, the water level in the Aral Sea dropped by 2 m in the first decade following the onset of irrigation, and continued to decline as water use in the watershed intensified. In this case, attribution of water level variation to human activity was robust because an abrupt change

in surface water management coincided with a comparably abrupt change in the water level of downstream lakes. However, attributing water level variation to human activity is more difficult when anthropogenic effects are subtle relative to the effects of background climate variation (Corti et al., 1999). ~~This challenge is especially salient when attributing water level variation to anthropogenic climate change (Hassanzadeh et al., 2012; Rodell et al., 2018)~~ This challenge is especially salient when attributing water level variation to anthropogenic climate change (Hassanzadeh et al., 2012; Rodell et al., 2018), especially at the global scale (~~Rodell et al., 2018~~) (Rodell et al., 2018). Indeed, there is an ongoing debate about the global extent and strength of climate change effects on water level variation in lakes (Muller, 2018).

Climate change can affect water levels through a complex web of forces linking surface temperature with hydrology (Block and Strzepek, 2012; Ramanathan et al., 2001; Rodell et al., 2018). While warming has been observed across the Earth's surface, hydrological responses to warming are highly variable with some areas becoming wetter and others becoming dryer (Greve et al., 2014; Rodell et al., 2018; Wang et al., 2012). However, more than three-quarters of Earth's land mass has seen no substantial change in total wetness or dryness in response to recent climate change (Greve et al., 2014; Greve and Seneviratne, 2015). Thus, the effects of climate change on water levels may be subtle relative to background climate variation (Jöhnk et al., 2004), calling for statistical approaches that can effectively account for the effects of background climate variation. Analyzing water levels from lakes worldwide may be especially helpful in reducing uncertainty about the potential contribution of background climatic variation to long-term trends.

Here, we build off previous studies focused on specific lake regions (~~Clites et al., 2014; Molinos and Donohue, 2014; Molinos et al., 2015; Pasquini et al., 2008~~) and attempt to disentangle the effects of background climate variation from other drivers of water levels in 117 globally distributed lakes using time series of remotely sensed water levels from 1992 to 2019. We investigate two key areas of uncertainty: (1) whether apparent anthropogenic water level trends in specific lakes can be explained by background climate variation, and (2) whether trends in water levels can be detected in specific lakes only after accounting for and removing the influence of background climate variation. We use boosted regression trees (BRTs) as a means of removing the effects of background climate variation on water levels in each lake, enabling us to achieve more robust quantification of the multi-decadal trends which may or may not be anthropogenic. This approach differs from other recently published approaches (~~Chanut et al., 1988; Molinos et al., 2015~~) in that it allows for nonlinear relationships, high levels of interactions between axes of climate variation, nonstationarity, and missing data. (~~Clites et al., 2014; Molinos and Donohue, 2014; Molinos et al., 2015; Pasquini et al., 2008~~) and attempt to disentangle the effects of background climate variation from other drivers of water levels in 200 globally distributed lakes using time series of remotely sensed water levels from 1992 to 2019. We investigate two key areas of uncertainty: (1) whether apparent anthropogenic water level trends in specific lakes can be explained by background climate variation, and (2) whether trends in water levels can be detected in specific lakes only after accounting for and removing the influence of background climate variation. We use boosted regression trees (BRTs) as a means of removing the effects of background climate variation on water levels in each lake, enabling us to achieve more robust quantification of the multi-decadal trends which may or may not be anthropogenic. This approach differs from other recently published approaches (~~Chanut et al., 1988; Molinos et al., 2015~~) in that it allows for nonlinear relationships, high levels of interactions between axes of climate variation, non-stationarity, and missing data. Finally, we assess patterns across lakes to generalize about which regions are most strongly influenced by background

climate variation. Our overall goal is to better understand the impact of large-scale background climate forcing on lakes in ways that will help communities manage the benefits derived from lakes in the face of global climate dynamics and anthropogenic influences.

## 2 Methods

### 2.1 Overview

The ~~417200~~ lakes included in our study contain much of the Earth's liquid surface freshwater and a large proportion of freshwater biodiversity (Vadeboncoeur et al., 2011). They span a wide range of lake characteristics including surface area (~~79.523~~ to ~~378,149.3377.002~~ km<sup>2</sup>), catchment area (~~20093~~ to ~~3,174,0002.764,126~~ km<sup>2</sup>), perimeter (~~49.462~~ to ~~11,478.915.829~~ km), latitude (~~-33.1650.22~~ to ~~66.4214~~ °N), and elevation (~~-7471~~ to ~~5,216194~~ m above sea level) (Supplementary Material). All continents were represented except for Antarctica and Australia. According to the HydroLakes database (Messenger et al., 2016), at least ~~2478~~ of these lakes are reservoirs, and ~~1418~~ of them ~~are natural lakes with~~ have some water level regulatory structure such as a dam. The global scope of this analysis builds off similar analyses which aimed to disentangle background climate variation's influence on lake water levels using smaller numbers of lakes from specific regions (~~Mercier et al., 2002; Molinos et al., 2015~~)(~~Mercier et al., 2002; Molinos et al., 2015~~) or specific lakes (Cohn and Robinson, 1975; Stager et al., 2005; Tomasion and Valle, 2000).

Our first objective was to estimate trends in water levels in this global sample of lakes based on average annual water level. ~~Annual water levels were used in the trend analysis instead of raw water levels so that trend residuals would not be serially autocorrelated.~~ We calculated trends using Thiel-Sen non-parametric regression using the 'zyp' package in R. Thiel-Sen slopes represent the median of slopes derived from all pairwise combinations of points in a time series. The statistical significance of each trend (p-value) was calculated using a bootstrapped one sample Wilcoxon signed-rank test with 1,000 repetitions where the input data for the test was the complete list of all slopes derived from all pairwise combinations of points in the time series. The number of pairwise slopes used in each repetition of the Wilcoxon signed-rank test was equal to the number of years of water level data for each lake.

Our second objective was to characterize and account for the effects of background global ~~climate~~ variation on ~~lake~~ water ~~level~~ ~~variation~~ ~~levels~~. We did this by using boosted regression trees (BRTs) to model water level variation in each lake as a function of the year, the month of the year, and a large set of global climate indices. We calculated the relative importance of each global climate index in the models for each lake to assess its sensitivity to different axes of background climate variation. We use the partial dependence of water level variation on the year term in the model to reflect the long-term variation in water levels that is not attributable to background climate variation. We repeat the Thiel-Sen slope calculation and p-value calculation based on the partial dependence data for the year term as an estimate of the long-term trend that is not attributable to background climate variation. This remaining variation could be attributable to human activity, though we cannot draw causal conclusions or distinguish between the various aspects of human activity which can affect water levels.

We derived the background global climate indices for each lake's BRT using principal components analysis (PCA) applied to global variation in monthly earth surface air temperature data through time. This approach is widely applied in the climate sciences to global grids of temperature, pressure, or rainfall data, and is analogous to Empirical Orthogonal Function (EOF) analysis (Dommenges and Latif, 2008; Hannachi et al., 2009; Kim and Wu, 1999). Temperature

time series at each pixel were included as separate variables in the PCA with each time step as a separate observation of those variables. We linearly detrended ~~temperature variation at each principal component (PC) resulting from the PCA using Thiel-Sen non-parametric regression; pixel~~ such that the ~~PC temperature~~ values were equivalent to the residuals from ~~the~~ Thiel-Sen regression. ~~PCs through time. Temperatures~~ were detrended because long-term ~~temperature trends in the PCs~~ could be considered potentially related to anthropogenic climate change ~~which we wanted to separate from the PCs representing background climate~~ (Stenseth et al., 2003). Thus, the effect of all PCs on water levels in the BRTs were interpreted as the collective effects of ~~background surface air temperature; temperature variation and their~~ associated hydrological effects on lake levels.

Using the ~~PCs as predictors in each lake's BRT~~ was preferred over using the commonly recognized climate indexes (NAO, ENSO, IOD, etc.) directly in ~~the each lake's~~ BRTs for several reasons. First, many commonly recognized global climate indices are ~~collinear correlated~~ which makes them problematic for simultaneous use as predictors, whereas the PCs used here are uncorrelated by definition. Second, using the commonly recognized global climate indices alone would miss a substantial amount of variation in air temperature that ~~is not periodic but~~ may still drive climate and hydrological variation in lakes: ~~but is not yet well-recognized~~. Third, many of the commonly recognized global climate indices are defined subjectively; ~~(e.g. average temperature difference between two subjectively defined areas of the ocean)~~, whereas the PCs used here are ~~identified empirically less subjective~~.

Our modelling approach is based on the recognition that much of the variation in water levels is directly or indirectly driven by global patterns in earth's surface air temperature via the effects of global temperatures on hydrological fluxes (Dommenges and Latif, 2008). This relationship is well-supported because earth's surface air temperature is a key control on earth's hydrological cycle through the Clausius Clapeyron relation which, in turn, drives lake water budgets and water levels (Christensen et al., 2004; Fowler et al., 2007; Nijssen et al., 2001; Tierney et al., 2008). Earth's surface air temperature may also be a key control on the positioning of atmospheric rivers which can also drive lake water budgets and water levels (Gimeno et al., 2014; Lorenzo et al., 2008). This assumption is empirically well-supported by studies showing earth surface air temperature variation described by ENSO, PDO and IOD is strongly associated with water level variation across the globe (Stager et al., 2007; Tierney et al., 2013; Wang et al., 2010).

Disentangling the direct and indirect effects of earth surface air temperature variation on water level variation could be done using a reductionist approach, i.e. by constructing lake hydrological budgets with all of the water inputs and outputs and modelling the effect of human activity on each of those fluxes. But, for most lakes in our dataset, we lack the field measurements required to model the forces linking earth surface air temperature to water level variation for the entire ~~2728~~-year water level time series (~~Stenseth et al., 2003~~); (~~Hegerl et al., 2015~~; ~~Stenseth et al., 2003~~). So instead, we use BRTs which mimic the complex web of forces linking earth surface air temperature variation to water level variation. We use BRTs for this purpose because the model structure accommodates high levels of interactions among predictor variables and mimics the interactive and indirect effects we tried to capture. This approach also differs from other recently published approaches (~~Chanut et al., 1988~~; ~~Molinos et al., 2015~~); (~~Chanut et al., 1988~~; ~~Molinos et al., 2015~~) in that it allows for nonlinear relationships, non-stationarity, and missing data. We fit BRTs separately for each lake because the influence of a particular climate oscillation could differ across lakes due to geographic forcing, orographic

forcing or other local factors (Stenseth et al., 2003). -We used backward elimination variable selection techniques to identify the PCs for each specific lake that, when fit against a training dataset, performed the best when predicting water level variation in a test dataset with which the model had not been fit. Then, we used the resulting BRTs to determine the relative importance of different axes of background climate variation separately for each lake.

## 2.2 Data

Water level data were acquired from the NASA/CNES Topex/Poseidon and Jason satellite missions via the Global Reservoir and Lake Monitoring (G-REALM) project version 2.3 (Crétaux and Birkett, 2006) and can be obtained from:

[http://www.pecad.fas.usda.gov/cropexplorer/global\\_reservoir](http://www.pecad.fas.usda.gov/cropexplorer/global_reservoir). Although these altimeters were developed to map ocean surface height, they have also been used to detect water level changes in lakes (Birkett, 1995). Only a small subset of the world's lakes can be monitored in this way because the space-borne sensors must pass directly over the lake with sufficient regularity to produce accurate and complete time series. The US Department of Agriculture (USDA) uses these data to monitor water level variation for many inland water bodies globally. The lakes in this study comprise the ~~117200~~ lakes with the longest (~~>28 years~~) and highest ~~temporal resolution time series- (greatest number of samples per year)~~. Validation of satellite altimeter data over inland water bodies is typically performed by comparing satellite altimeter measurements and in situ measurements. The root mean squared error of elevation variations derived from the NASA/CNES Topex/Poseidon and Jason-1 satellite missions is typically ~~~3 cm for large lakes (Birkett, 1995; Crétaux and Birkett, 2006). Thus, it is justifiable to use water level altimetry~~ ~~5 cm for large lakes based on the USDA G-REALM website (https://ipad.fas.usda.gov/cropexplorer/global\_reservoir/validation.aspx). Thus, it is justifiable to use altimetry water level observations~~ in place of in situ gauge measurements (Birkett et al., 2011).

Water levels are typically measured every 10 days, but the exact dates on which water levels are measured vary from lake to lake. To make water level data temporally consistent, we linearly-interpolated each lake's time series to the daily scale using the 'deseasonalize' and 'zoo,' packages in R (R Core Team, 2017). Monthly averages were calculated so that all lakes had time series of the same interval that also matches the temporal resolution of surface air temperature data used in the PCA. ~~Nine70~~ of the ~~117200~~ water level time series had a data gap from late 2002 through the middle of 2008. The missing data were not estimated, ~~but instead~~, our analyses treated these lakes in the same way as ~~ones lakes~~ with complete data.

Monthly average land and ocean surface air temperature anomaly data were acquired from the Goddard Institute for Space Studies (GISS) Land Ocean Temperature Index analysis for a 2 x 2 degree grid with 1200 km smoothing (Hansen et al., 2010). These temperature data are derived from meteorological station observations distributed across the globe, and processed according to methods developed at the National Aeronautics and Space Administration (NASA) (Hansen et al., 2010). They are publicly available at <https://data.giss.nasa.gov/gistemp/>.

## 2.3 Principal components analysis

We used PCA to distill the spatial complexity of surface air temperatures for inclusion in each lake's BRT. PCA is an ordination-based statistical tool that converts potentially correlated variables into a set of orthogonal vectors that capture the variation across locations. PCA uses

orthogonal linear transformation to identify vectors that account for as much of the total variation in a set of variables as possible. The first PC ( $PC_1$ ) explains the largest percentage of the variation in the underlying set of variables followed by the second ( $PC_2$ ), third ( $PC_3$ ), and so on. Each succeeding PC is linearly uncorrelated to the others and accounts for as much of the remaining variation as possible. PCA can, therefore, be used to summarize the consistent aspects of time dynamics across space and reduce redundant spatial variation (stemming from spatial autocorrelation and teleconnections) in temperature.

To identify which known oscillations in surface air temperature are related to individual PCs, we calculated a complete correlation matrix between each PC and all of the [3743](#) major climate indices recognized by NOAA's Earth System Research Laboratory. [For indices that aren't updated to the present, we calculated the correlation over the longest time period over which each major climate index was available.](#) Monthly time series of major climate indexes were sourced from <https://www.esrl.noaa.gov/psd/data/climateindices/list/>. In cases where a PC is highly correlated to one of the major climate indices (PCs), we renamed it with a subscript (e.g.,  $PC_{\text{ENSO}}$ ) to facilitate interpretation. [In all remaining cases, PCs were named with a numeric subscript which matched their order in the PCA \(e.g., "PC<sub>18</sub>"\)](#)

#### 2.4 Boosted regression trees (BRTs) and model selection

BRTs were used to model mean monthly water levels in the lakes as a function of year, the month of the year, and a large set of PCs. [BRTs were fitA BRT is an ensemble machine learning approach that differs from conventional statistical techniques which use a single parsimonious model. Instead, BRTs combine the strengths of standard regression trees and boosting—a method for aggregating many models to improve the predictive capacity. The main advantages of BRTs over other statistical models is that they have higher predictive performance, do not require data transformation or outlier elimination, automatically handle complex nonlinear relationships and interactions, and allow for many types of predictor variables and partial missing data. Through these interactions, the BRT allows for non-stationarity of the timeseries \(e.g. the effect of each PC is allowed to change over time\). We fit BRTs using the "dismo" and "gbm" packages \(Hijmans et al., 2017\) in R \(R Core Team, 2017\).](#) To cross validate the BRTs for each lake, we fit the model using a training dataset and then used the fit BRT to predict monthly water levels using the PC values from a test dataset. For each lake, we fit six starting models using six different training and test dataset combinations. To get these dataset combinations, we first split the lake level time series for each lake into training and test datasets along its time series using 40-60, 50-50, and 60-40 splits. We split the data [serially](#) along the time series into training and test datasets instead of by randomly selecting observations for the training and test datasets because the data are temporally autocorrelated and we wanted to ensure that the training and test datasets were independent. For each of the three splits, the starting model was fit twice, once using the first part of the split as a training dataset and the second part as a test dataset, and once using the second part of the split as a training dataset and the first part as a test dataset, resulting in a total of six train-test dataset combinations. For each lake's six train-test dataset combinations, we repeatedly refit the starting model after dropping the PC with the lowest relative importance averaged across all six train-test dataset combinations (see explanation of relative importance below) until the starting model had only 2 predictors—the minimum number of predictors allowed in a BRT. Each time a variable was dropped from one of the 6 starting models, we calculated the average change in predicted residual error sum of



squares (PRESS; the sum of squares of the prediction residuals calculated using the test data) which resulted from dropping it.

Variables which, when dropped from the model, resulted in an average increase in PRESS across the six starting models were selected in that lake's "best BRT." We combined information across models so that the selection of a variable did not depend on the arbitrary decisions of where to split the time series and whether to use the first or second part as the training dataset. ~~Thus, the best BRT included the year, the month, and the combination of PCs from PC<sub>1</sub> to PC<sub>100</sub> which consistently improved its performance in cross validation. The first 100 PCs included all PCs which explained at least 0.1% of the variability in earth surface air temperatures. The first 100 PCs were used in the BRTs in recognition that NOAA distinguishes 37 similar axes of variation and many more potentially go undescribed. This best BRT All 336 PCs were used as predictors at the beginning of the model selection process because even very high order PCs can be important variables in a model (Phinyomark et al., 2015), and many of the high order PCs calculated here were statistically distinguishable from random noise (Supplementary Material).~~ Thus, the best BRT included the year, the month, and the combination of PCs from PC<sub>1</sub> to PC<sub>336</sub> which consistently improved its performance in cross validation. 336 PCs was the maximum number of PCs which could be calculated from 336 monthly observations of temperature over the 28-year span of the lake level time series. This best combination of PCs, specific to each lake, was used for determining the relative importance of the variables selected in the best model. We refit the best BRT for each lake to the full time series to calculate the final relative importance values of each PC. The relative importance of each predictor variable in the model is a function of the frequency with which it was included in the BRT's individual regression trees and the improvement to the model that resulted from its inclusion (Elith et al., 2008). The relative importance of variables that were not selected in each lake's best BRT was ~~considered set~~ to be zero.

### 3 Results

We observed considerable variation in water levels within the ~~117200~~ lakes in our analyses (Fig 1). Prior to accounting for the effects of background climate variation, water levels increased at a median rate of 1.42 cm year<sup>-1</sup> (interquartile range: -0.53 cm year<sup>-1</sup> to 34.1 cm year<sup>-1</sup>) (Fig 1). Water levels were decreasing in ~~3560~~ lakes (~~29.930~~ %), of which 14 were statistically significant ( $p < 0.05$  level), and increasing in ~~82140~~ lakes (~~70.1~~ %), of which ~~3051~~ were statistically significant ( $p < 0.05$  level). In total, ~~4465~~ lakes (~~37.632.5~~ %) in our analyses had significant trends in water level (Fig 21). Given a significance level of  $\alpha = 0.05$ , we would expect only ~~610~~ of our ~~117200~~ lakes to show "significant" trends by chance. ~~However; thus,~~ we observed ~~far more~~ significant long-term trends ~~in water levels in 44 lakes, far more than predicted based on~~ chance alone (Fig. 21). A comparable disparity was observed across a range of different arbitrary thresholds for statistical significance (i.e., 0.01, 0.05, and 0.1).

Changes in water levels from 1992 to 2019 displayed a moderate level of regional consistency in the direction and magnitude of water level trends (Fig 32). In particular, lakes in the Middle East and the Southwest United States tended to have decreasing water levels (Fig 32). Lakes in Canada and Europe tended to have weak or increasing water level trends. Water level trends in East Asia and Africa were highly variable from lake to lake (Fig 32).

BRTs performed well for most lakes in our cross-validation and model optimization procedure; the median PRESS of the best model was 57.0 cm (interquartile range: ~~2.6-17.23.5-~~ 85.7 cm) across lakes. ~~However, our BRTs failed for two~~ Two lakes with known anthropogenic

water level dynamics, both of which are reservoirs on the Mekong River in China where water levels increased performed poorly in cross validation because they had water level increases >50 m as a result of dam construction in the middle of the time series (Nuozhadu and Xiaowan Dams). Across lakes, the best model included a median of 511 of the 100336 PCs (interquartile range: 2103-19 PCs) that were fed into our model selection procedure. Remarkably, only 4154 out of 100336 PCs were never selected in any lake, and the overall frequency of inclusion across lakes decreased with PC order (Kendall's tau = -0.54,  $p < 0.01$ ). Together, the PCs selected for each lake explained a median of 27% of the variance in water levels (interquartile range: 4-44%)-75,  $p < 0.01$ ). The median percentage of water level variation associated with background climate variation was 58% (inner quartile range: 33-74%) with an additional 10% (inner quartile range: 4-22%) explained by seasonal variation and 25% (inner quartile range: 13-50%) explained by the long-term trend.

The relative importance of each specific predictor variable varied substantially from lake to lake. On average based on the median relative importance values across lakes, the year and the month were most influential across lakes. On a scale from 0-100 with higher values connoting higher relative importance, year and month had mean relative importance values of 29.1 and 11.1, respectively, strongly influenced water level variability. PC<sub>1</sub>, PC<sub>4</sub>, PC<sub>2</sub>, and PC<sub>5</sub>, and PC<sub>6</sub> were the predictor variables with the third (4.32.6), fourth (4.2), and 1.6), fifth (3.21.5), and sixth (1.4) highest mean median relative importance in the best performing BRTs (Fig 4). PC<sub>4</sub> was moderately correlated with the North Pacific Oscillation (NPO; Kendall's tau = 0.30,  $p < 0.01$ ); PC<sub>2</sub> was highly correlated with the Multivariate El Niño Index (MEI; Kendall's tau = 0.7268,  $p < 0.01$ ); PC<sub>4</sub> was correlated with the Atlantic Multidecadal Oscillation (AMO; Kendall's tau = 0.26,  $p < 0.01$ ); and PC<sub>5</sub> was moderately correlated with the North Atlantic Arctic Oscillation (NAOAO; Kendall's tau = 0.3730,  $p < 0.01$ ); PC<sub>6</sub> was correlated with the Indian Ocean Dipole (IOD; Kendall's tau = 0.21,  $p < 0.01$ ); so we renamed them here as PC<sub>NPO</sub>, PC<sub>ENSO</sub>, PC<sub>AMO</sub>, PC<sub>AO</sub>, and PC<sub>NAO</sub>PC<sub>IOD</sub> (Fig 53-4). These threefour PCs together encompass 18.222.0% of the variation in surface air temperature anomalies (PC<sub>NPO</sub> according to the eigenvalues from the PCA (PC<sub>ENSO</sub> = 9.2%, PC<sub>AMO</sub> = 5.5%, PC<sub>ENSO</sub> = 7.8%, PC<sub>AO</sub> = 4.0%, and PC<sub>NAO</sub> = 4.8%)-PC<sub>NPO</sub>; PC<sub>IOD</sub> = 3.3%). PC<sub>ENSO</sub>, PC<sub>AMO</sub>, PC<sub>AO</sub>, and PC<sub>NAO</sub>PC<sub>IOD</sub> were selected in the best models for 42, 43107, 78, 69, and 3270 lakes, respectively, but the direction of their effects differed among waterbodies lakes (Fig 53). Many of the remaining PCs of high mean relative importance across waterbodies (mean relative influence > 1) were only moderately correlated to indices from NOAA (Kendall's correlation coefficient range < 0.22-0.342). For instance, PC<sub>15</sub>PC<sub>18</sub> was not substantially correlated with any climate index recognized by NOAA (maximum Kendall's tau = 0.1714,  $p < 0.01$ ), yet exhibited the 9<sup>th</sup>11<sup>th</sup> highest mean median relative importance in explaining water levels across lakes (Fig 65).

PC<sub>NPO</sub>, PC<sub>ENSO</sub>, PC<sub>AMO</sub>, PC<sub>AO</sub>, and PC<sub>NAO</sub>PC<sub>IOD</sub> were strongly related to water level variation in lakes around the world but the strength and directionality of those effects were regionally concentrated (Fig 7-9)-6). The strongest effects of PC<sub>ENSO</sub> were concentrated in the tropics where it had positive effects on water levels (Fig 6). PC<sub>AMO</sub> was more positively associated with water levels in Central Canada and Northeast the Midwest United States, and negatively associated with water levels in Sub-Saharan Northern Canada and East Africa, Northern Europe, South Central Canada, the Equatorial Americas and Central Asia (Fig 7). The strongest effects of PC<sub>NAO</sub> were concentrated in North America and East Asia where it had uniformly positive effects on water levels (Fig 8). PC<sub>NPO</sub>, PC<sub>AO</sub> was also selected in the best models of lakes that were regionally concentrated; it was positively associated with water levels

in East Africa, Northern Europe, and South East Canada/Central Asia while it was negatively associated with water levels in parts of Canada, Alaska, Northern Canada/Europe, Brazil, and Northern East Africa (Fig 9). Not all effects of the PCs on water levels were monotonic—PC<sub>NPO</sub> had a U-shaped relationship with water levels in Southeast North America such that both very low and very high values of PC<sub>NPO</sub> were 6). PC<sub>IOD</sub> was positively associated with relatively high water levels in Europe (Fig 9)-6).

After removing the effects of background climate variation on water levels using the fitted BRTs, water level trend estimates were shallower compared to estimates from the original time series (Fig 1). The median water level trend across lakes dropped from +1.42 to +1.0.8 cm year<sup>-1</sup> after correcting for background climate variation (Fig 1). Even though they were lower on average, the trends became more statistically significant in 76.86% of lakes (Fig 21). Indeed, the median p-value of water level trends across lakes changed from 0.1216 to 0.02 after removing the effects of background climate variation (Fig 21). For instance, prior to removing the effects of background climate variation, Becharof Lake Mweru in central Africa/Alaska had an increasing trend (+2.620.40 cm year<sup>-1</sup>) with a relatively high-low statistical significance (p-value  $\leq$  0.3015). However, after accounting for the effects of background climate variation, the trend was not substantially affected (2.52 from +0.40 cm year<sup>-1</sup> to +0.51 cm year<sup>-1</sup>) but was became substantially more statistically significant (p-value: from 0.00915 to <0.01). Based on inspection of the time series of water levels in Lake Mweru/Becharof, we suspect that the strong peak in lake levels in the late 1990's due to climate oscillations affecting lake levels throughout the time series, in particular the water level local minima in 1998 and 2017 as well as the local maxima in 2003 and 2013 masked the long-term trend (Fig 10). In contrast, prior to removing the effects of background climate variation, Kainji/Gods Lake in Nigeria/Canada had an increasing trend (+4.321.26 cm year<sup>-1</sup>) with a relatively low statistical significance (p-value  $\leq$  0.05235). However, after accounting for the effects of background climate variation, the trend flipped sign and was moderately lower (+3.52 much weaker (from +1.26 cm year<sup>-1</sup>) but to -0.10 cm year<sup>-1</sup>) and less statistically significant (p-value: from 0.13535 to 0.58). Based on inspection of the time series, we suspect that synergies among background climate variables a series of three local water level minima early in the lake level time series in 1996, 2000, and 2004 created a specious appearance of a long-term trend in Kainji/Gods Lake (Fig 10)-7) that went away when those minima were attributed to background climate.

#### 4 Discussion:

We detected long-term trends in water levels before and after accounting for background climate variation in most lakes in our analyses. The evidence of trends in the majority of lakes belies reports that most of the Earth has experienced no consistent changes in annual wetness and dryness (Greve et al., 2014; Greve and Seneviratne, 2015). This contrast highlights the potential for waterbody surface levels to serve as integrative metrics of regional water budgets, thereby enhancing our ability to detect hydrological changes.

Background climate variation had significant effects on water levels in most large lakes between 1992 and 2019. Infrequently, the effects of multiple axes of background climate variation gave rise to the appearance of long-term trends which became less significant once background climate variation was factored out. But more often, background climate variation masked underlying trends in water levels which were detected when the effects of background climate were factored out (Table S2). Thus, attempting to detect anthropogenic effects on water

levels using water level time series without accounting for background climate variation may over- or under- estimate the multi-decadal water level trends in lakes.

The trends in water levels estimated here differed widely among lakes, presumably reflecting the heterogeneity of underlying changes in regional hydrological fluxes. Rising water levels in the majority of lakes may be attributable to increases in precipitation within their watersheds (Bintanja and Selten, 2014; Chadwick et al., 2013; IPCC, 2014; O’Gorman et al., 2012). However, even in watersheds which have experienced increased precipitation, greater inputs of water may be offset or even exceeded by increases in evapotranspiration (Dorigo et al., 2012; Vinukollu et al., 2011; Vörösmarty et al., 2000; Vörösmarty and Sahagian, 2006) that yield net decreases in water levels.

Not all of the lakes with significant trends in water levels followed the “wet gets wetter and dry gets dryer” pattern that is often predicted to occur with climate change (Wang et al., 2012). According to such predictions, surface water storage would be expected to decrease in many dry mid-latitude and subtropical regions, and to increase at high latitudes and in humid mid-latitude regions (IPCC, 2014). Several lakes with the strongest decreases in water levels (Aral, Mosul, Powell, Rakshastal, Salton, Tharthar, and Urmia) are indeed located in relatively dry regions (average long-term discharge / watershed area  $< 5000 \text{ cm}^3 \text{ s}^{-1} \text{ km}^{-2}$ ). Furthermore, several lakes with the strongest increases in water levels (Zeyaskoye and Atitlan) are located in relatively wet regions (average long-term discharge / watershed area  $> 5000 \text{ cm}^3 \text{ s}^{-1} \text{ km}^{-2}$ ). However, the intensification of wet-dry contrasts was violated in many places as some lakes in wet regions got dryer (Vermelha, Winnebago, and Woods) and some lakes in dry regions got wetter (Balkash, Cabora Bassa, Kapachagayskoye, Kariba, and Ulungar). This finding adds to others showing that a range of hydrological fluxes contradict the “wet gets wetter and dry gets dryer” pattern over land (Byrne et al., 2015). As a specific example from our study, Lake Turkana had a significant increase in water level from 1992 to 2019 despite being in a very dry region. Thus, intensification of contrasts in precipitation may be a useful heuristic for predicting water level trends in some regions, but is clearly inapplicable at the global scale (Greve et al., 2014; Greve and Seneviratne, 2015).

The changes in water levels in response to the threefour most important PCs ( $PC_{NPO}$ ,  $PC_{ENSO}$ ,  $PC_{AMO}$ ,  $PC_{AO}$ , and  $PC_{NAO}$   $PC_{IOD}$ ) often matched the direction of change predicted from the hydrological changes associated with these particular climate oscillations. For instance, we observed strong negative relationships between  $PC_{ENSO}$  and water levels for many waterbodies in Sub-Saharan Africa, the Equatorial Americas, and South Central Canada, as would be predicted based on studies of the global effects of ENSO on precipitation (Dai and Wigley, 2000; Ropelewski and Halpert, 1987). We also observed water level changes in North America as a function of  $PC_{NAO}$  that are consistent with predictions from observed regional changes in rainfall associated with NAO (Dai et al., 1997). We also observed water level changes in North America as a function of  $PC_{AMO}$  that are consistent with predictions from observed regional changes in rainfall associated with AMO (Wang et al., 2017). However, we also detected relationships between water levels and PCs that were not consistent with the known hydrological effects of  $NPO$ , the ENSO, AO, AMO, and NAOIOD. For example, we found that  $PC_{NPO}PC_{AO}$  had a strong effect on water levels in Northern Europe East Africa, where  $NPO$  the AO is not typically considered to be a major driver of hydrological fluxes. However, the correlation between  $PC_{NPO}PC_{AO}$  and the  $NPOAO$  was relatively weak, so the apparent influence of  $PC_{NPO}PC_{AO}$  could be driven by climate variation that was captured by  $PC_{NPO}PC_{AO}$  and not the  $NPO$  actual AO index.

—~~Twenty-seven percent~~ 58% of the explained variation in water levels (median across lakes) could be attributed to large-scale climate drivers as represented by the ~~first 100~~ 336 PCs derived from air temperature records. We included higher order PCs because they might have accounted for additional variation in water levels. But we note that higher order PCs encompassed far less of the variation in surface air temperatures and they were included in far fewer of the lakes' best water level models, so their implications for water levels are far weaker. For example, PC<sub>100</sub> explained only 0.1% of variation in global surface air temperatures, and was included in our best performing model for 24 lakes albeit with very low explanatory contribution in those cases. Thus, excluding high-order PCs is unlikely to lead to substantial improvements or changes to our conclusions. The performance of the BRTs might be improved by including lag effects (Hansen et al., 1998; Hidalgo and Dracup, 2003). However, the computation time required to include ~~more PCs or~~ lag effects was prohibitive.

Our interpretations of the statistical patterns reported herein require several caveats. First, there is substantial debate over which aspect of human activity (e.g. climate change, land use change, dam construction/management) is most important for driving water levels (Gyau-Boakye, 2001; Lenters, 2001; Mercier et al., 2002; Ricko et al., 2011; Wurtsbaugh et al., 2017). Our modelling approach does not discern whether the trends we calculated are anthropogenic or which aspects of human activity are driving water level trends. Hence, future work is needed to disentangle the various anthropogenic forces which may influence water levels. Detecting anthropogenic water level trends and distinguishing between the effects of different aspects of human activity on water levels could be achieved by including water level dynamics in Earth system models. To date, lake ecosystems are generally oversimplified in such models in which lakes are often assumed to be relatively static, inert bodies on the landscape.

Second, we interpreted the trend in the partial dependence values for the year term in each lake's BRT as being potentially anthropogenic. However, we cannot exclude the possibility that anthropogenic water level variation was attributed to background climate PCs themselves in cases where PCs are highly correlated with human activity. Human water extraction may consistently be a larger proportion of total water fluxes during certain phases of the ENSO cycle, for example. In ~~our~~ the BRTs, this aspect of anthropogenic water level variation would have been captured by PC<sub>ENSO</sub> rather than the year term, and would have been erroneously attributed to background climate variation. However, if the correlation between human water extraction and ENSO changed over time, it would be correctly attributed to the year variable in the BRTs, ~~as we expected in general accordance with our simplified interpretations.~~

Third, we distinguish between anthropogenic climate change and background climate variation in our interpretations because background climate indices like those used here are generally considered to be modes of natural variation. However, human activity may influence background climate variation as well, perhaps making certain phases of various climate oscillations more likely (Cai et al., 2015; Capotondi and Sardeshmukh, 2017; Timmermann et al., 1999). We partially accounted for this by removing any linear trend through time ~~for each PC in temperature~~ prior to fitting the BRTs PCA. Detrending the ~~PC temperature data~~ in this way helped to separate ~~our~~ the background climate indices from any ongoing climate change. However, more complex interactions between climate change and background climate variation would not be removed by this approach, as in a scenario where climate change enhances both positive and negative phases of background climate oscillations yet has no impact on trends ~~in~~ of the mean. Research on such complex relationships is still inconclusive (Allen and Ingram, 2002; Cane, 2005; Collins, 2000; Guilyardi et al., 2009; van Oldenborgh et al., 2005), so we performed

only linear detrending of PCs. We also recognize that our 2728-year time series could also reflect longer-period climate oscillations such as the ~~Multi-Pacific Decadal-Atlantic~~ Oscillation, but the limited duration of altimeter-based satellite monitoring of lake levels precludes testing for such influences.

The lakes represented in our study comprise a substantial portion of the global liquid surface freshwater on the planet. Our study includes the ten most voluminous freshwater lakes on Earth's surface (Baikal, Tanganyika, Superior, Michigan, Huron, Malawi, Victoria, Great Bear, Ontario, Great Slave), which collectively contain more freshwater (total 80,241 km<sup>3</sup>) than has been withdrawn from the environment by humans at the global scale over the last 20 years (Food and Agriculture Organization of the United Nations, 2016). But the lakes in this study are not representative of all lakes, which tend to have smaller surface area and shallower maximum depths on average. Thus, the relative importance of background climate oscillations in the remainder of lakes other than the ~~117200~~ large lakes studied here over the last 2728 years remains uncertain and should be investigated further.

Our modelling approach could be widely applied to disentangle the effects of background climate on other hydrological changes including streamflow and pan evaporation rates. The novel statistical method presented here using BRTs could be used to describe or factor out the effects of long-term variation in background climate variation on a wide variety of environmental variables including fires, floods, heatwaves, and droughts—all of which have been shown to be sensitive to climate teleconnections (Chen et al., 2016; Lau and Kim, 2012; Stenseth et al., 2003). Instead of using common climate indices, we encourage the use of the complete set of PCs calculated here, because important climate oscillations ~~would~~ could otherwise be missed. To illustrate this point, the PC with the ~~9th~~ 11th highest ~~mean~~ median relative importance in explaining water level variation was not strongly correlated with any of the reported NOAA climate indices (Fig 75).

Wherever water levels are affected by background climate and human activity, there is the potential to affect lake ecosystems and the benefits that humans derive from them (Clites et al., 2014). More than two billion people live in water stressed regions of the world where human demand for surface freshwater exceeds the available supply (Mekonnen and Hoekstra, 2016; Vörösmarty et al., 2000, 2010). Meeting the competing demands for surface freshwater, especially in water scarce regions and in the face of anthropogenic environmental change, is a key challenge for society. Our capacity to disentangle the effects of background climate oscillations on water levels is key to sustaining our freshwater resources, especially in the face of climate change (~~Clites et al., 2014; Gronewold et al., 2013~~). ~~By applying this BRT statistical approach, we partially disentangled the effects of background global climate indices on water levels. Many of the large lakes in our analyses were remarkably resilient to long-term changes from 1992 to 2019. Thus large lakes may be an increasingly important resource in the face of water scarcity in the future. Abrupt changes in water levels in large lakes remain possible due to human activities and climate change, but our analyses suggest that we have not yet crossed such thresholds~~(Clites et al., 2014; Gronewold et al., 2013). ~~By applying this BRT statistical approach, we partially disentangled the effects of background global climate indices on water levels. Many of the large lakes in our analyses were remarkably resilient to long-term changes from 1992 to 2019. Thus, large lakes may be an increasingly important resource as water scarcity intensifies in the future. Abrupt changes in water levels in large lakes remain possible due to human activities and climate change, but our analyses suggest that we have not yet observed such changes~~ in many of Earth's largest lakes.



## **5 Conclusions:**

On average, water levels in the world's large lakes are increasing but are highly variable from lake to lake. Background climate variability often masks these long-term trends in water levels and occasionally gives rise to the appearance of false trends that wane after background climate variation is factored out. Background climate variation alone can explain a large proportion of water level variability in lakes worldwide due to the strong influence of earth surface air temperatures on lake levels via climate-lake level teleconnections. These findings highlight further opportunities to investigate the specific mechanisms that couple climate and lake levels. The novel statistical method presented here using BRTs could be used to describe or factor out the effects of long-term variation in background climate variation on a wide variety of environmental variables including fires, floods, heatwaves, and droughts. Tests for long-term trends in environmental variables which may or may not be anthropogenic will likely benefit from prior isolation and removal of the effects of background climate variation using this method.

## **Acknowledgements:**

The lead author is grateful for support from the IGB Leibniz Institute for Freshwater Ecology and Inland Fisheries through their international visiting scholars program and from the German Research Foundation within the LimnoScenES project (AD 91/22-1). We are also grateful for financial support from the John D. and Catherine T. MacArthur Foundation grants G-108015-0 and G-1609-151200 to Appalachian State University, and for further support from the Packard Fellowship in Science and Engineering.

## **Data Availability Statement:**

Water level data products are courtesy of the USDA/NASA G-REALM program which can be found at [https://www.pecad.fas.usda.gov/cropexplorer/global\\_reservoir/](https://www.pecad.fas.usda.gov/cropexplorer/global_reservoir/). All GISS data can be found at <https://data.giss.nasa.gov/gistemp/>.

## **Code Availability Statement:**

All code used in the production of this manuscript including data analysis and figures are published in the Zenodo online repository with DOI: 10.5281/zenodo.3363187

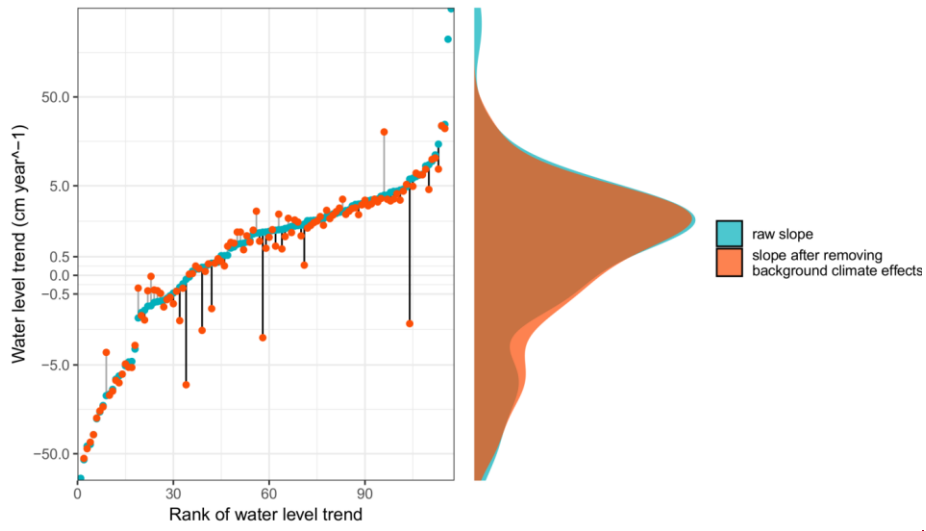
## **Author Contributions:**

BK designed the study, developed the model code, and performed the analyses. BK prepared the initial manuscript. BK, RA, AS, and PM contributed to the manuscript's revision and editing.

## **Competing Interests:**

The authors declare that they have no conflict of interest.

Figures:





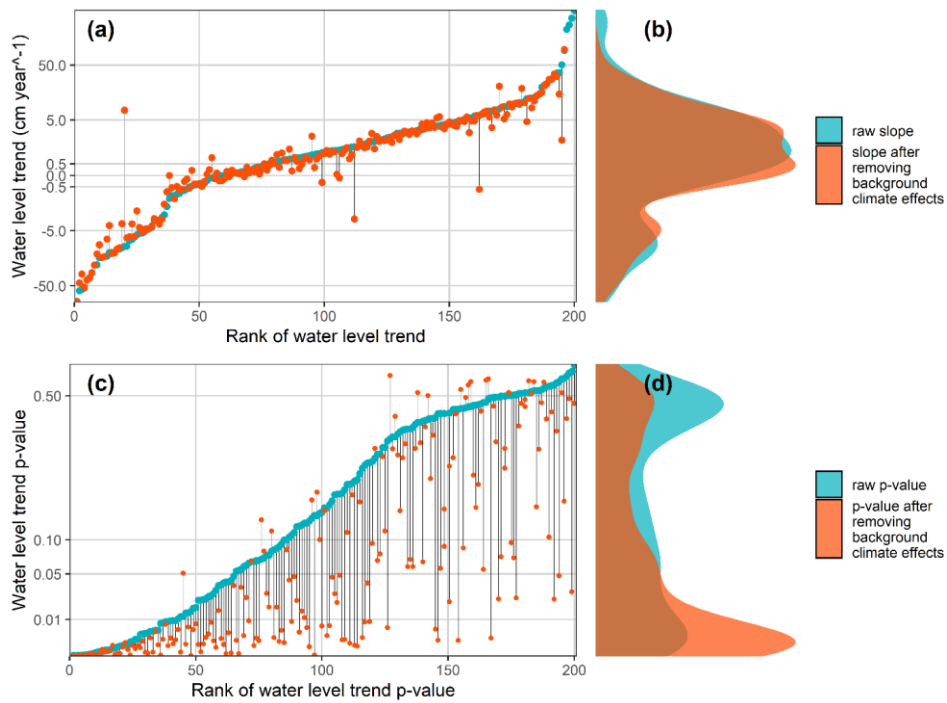


Figure 1: Water level trends in 117 globally distributed large lakes and theThe change in each lake's water level trend (a, b) and their trend p-value (c, d) which resulted from removing the effects of background climate variation on water level variation in 200 globally distributed large lakes. After removing the effects of background climate variation on water levels, long-term water level trend estimates (grey) were slightly more conservative overall compared to rates calculated from the raw water level data (black). The median value of water level trends across lakes changed from +1.42 cm year<sup>-1</sup> to +1.0.8 cm year<sup>-1</sup> after removing the effects of background climate variation.

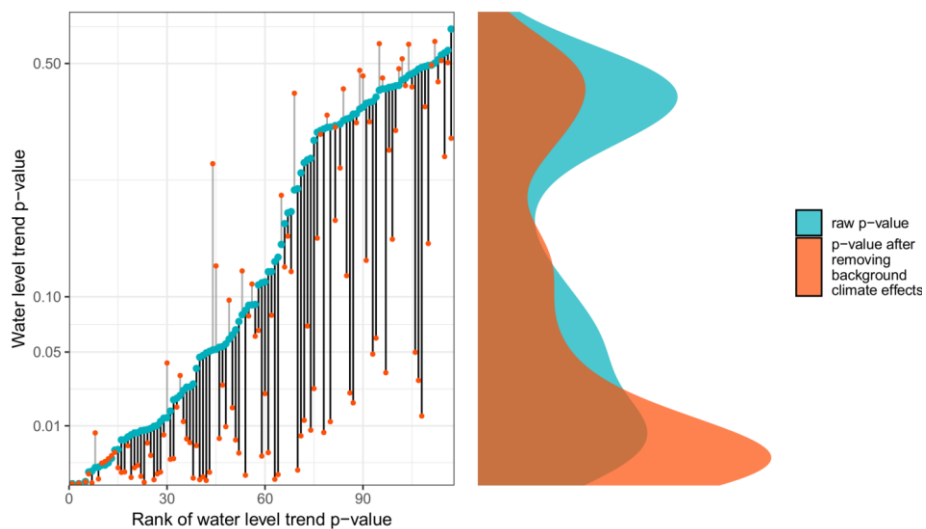


Figure 2: The change in each lake's water level trend p-value which resulted from removing the effects of background climate variation. The trends became more statistically significant in most (76%, bluegrey lines above the curve) but not all lakes (24%, redblack lines below the curve) after accounting for the effects of background climate variation. The median p-value of water level trends across lakes changed from 0.4216 to 0.02 after removing the effects of background climate variation. Lakes are ranked independently in panels (a) and (c).

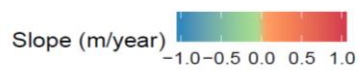
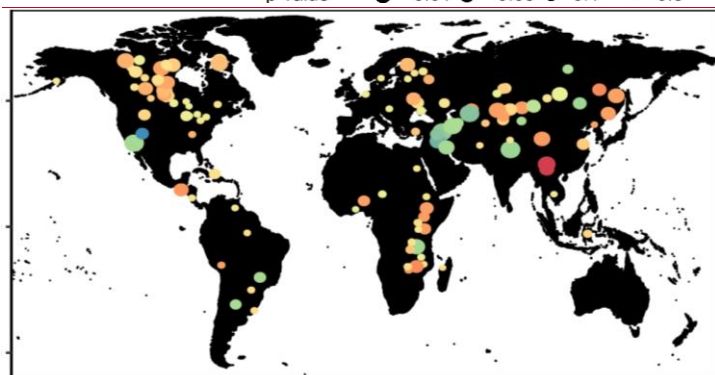
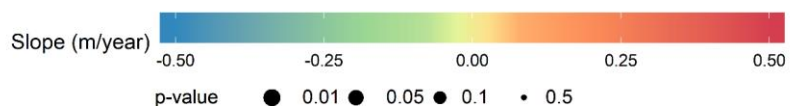
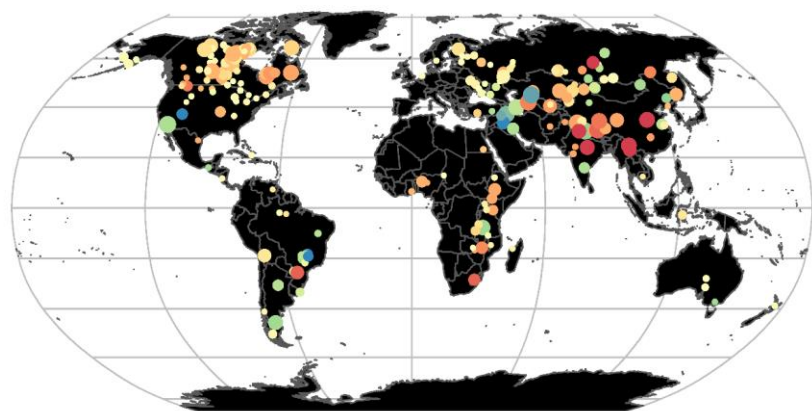
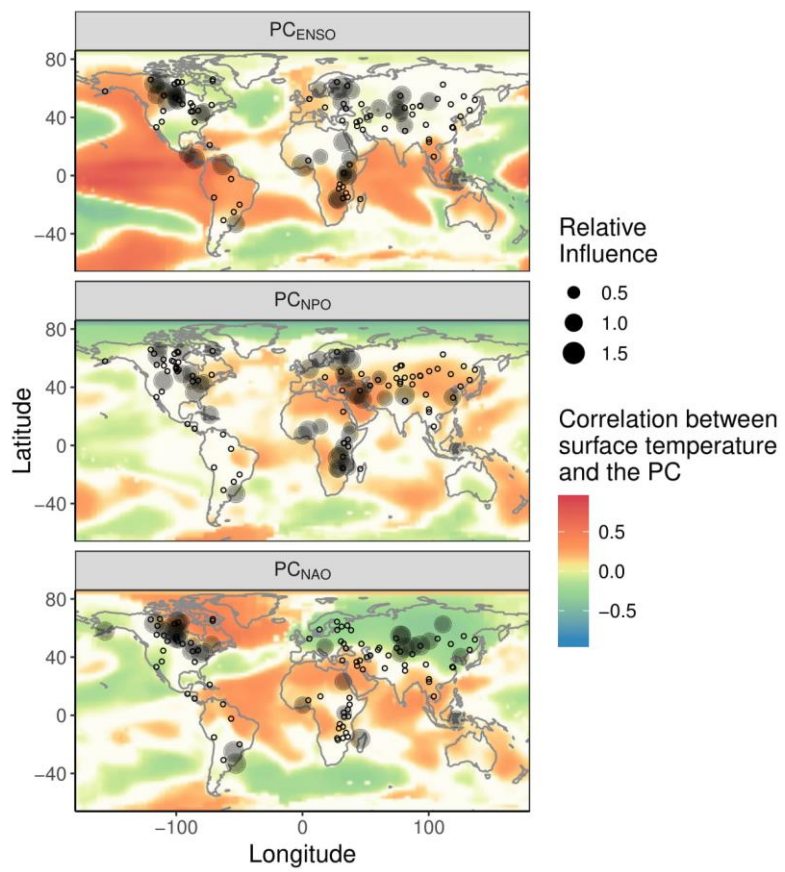


Figure 32: Long-term trends in water levels prior to removing the effects of background climate variation. Some lakes including those in the southwest United States, parts of Africa, and the Middle-East show regionally-consistent patterns in water levels.



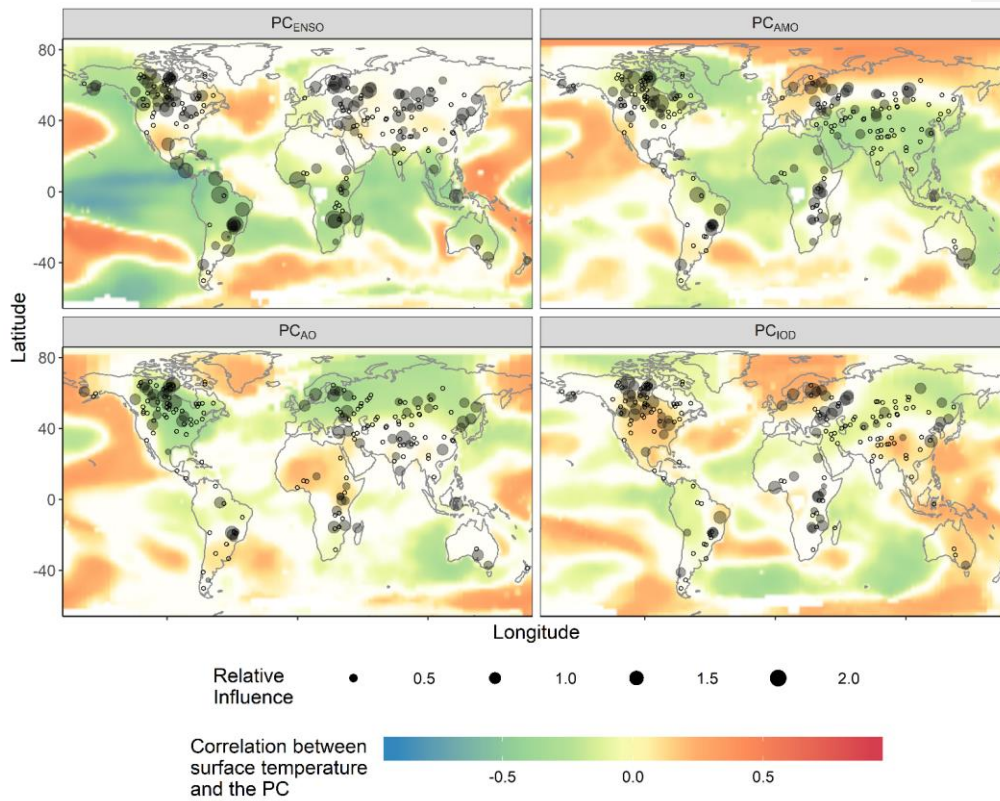
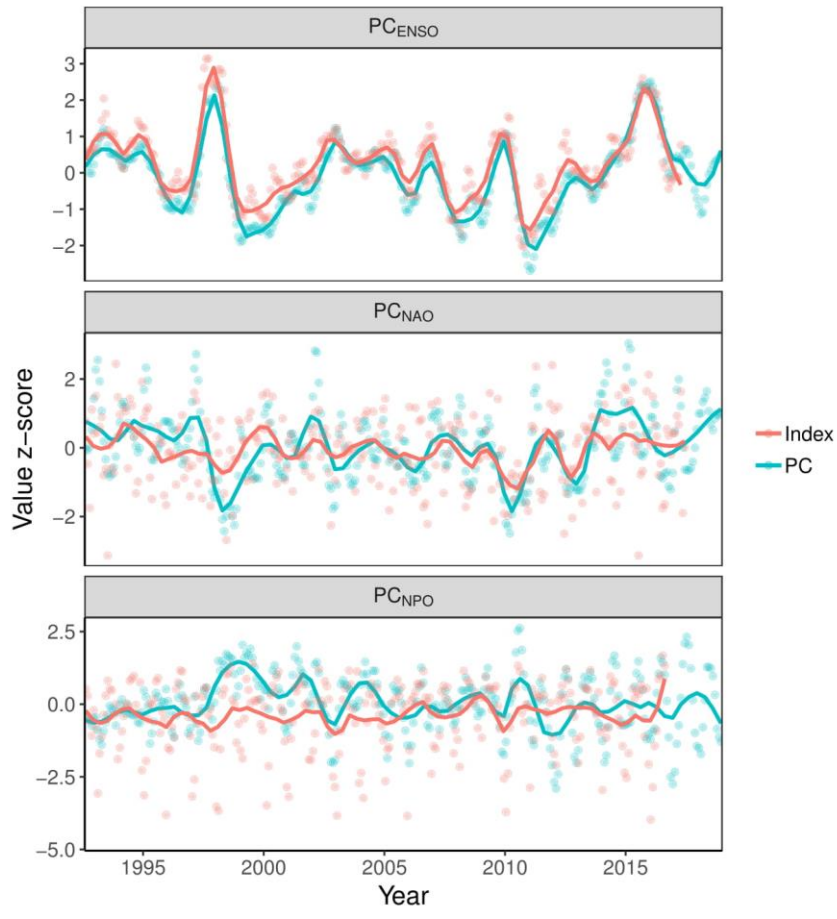


Figure 43: The relative influence of different axes of background climate variation (PCs) on water level variation. Empty dots represent lakes for which the PC was not selected in its “best model.” The opacity of each colored pixel in each map is related to the significance of the correlation between the PC and temperature at that pixel with less significant correlations appearing white.



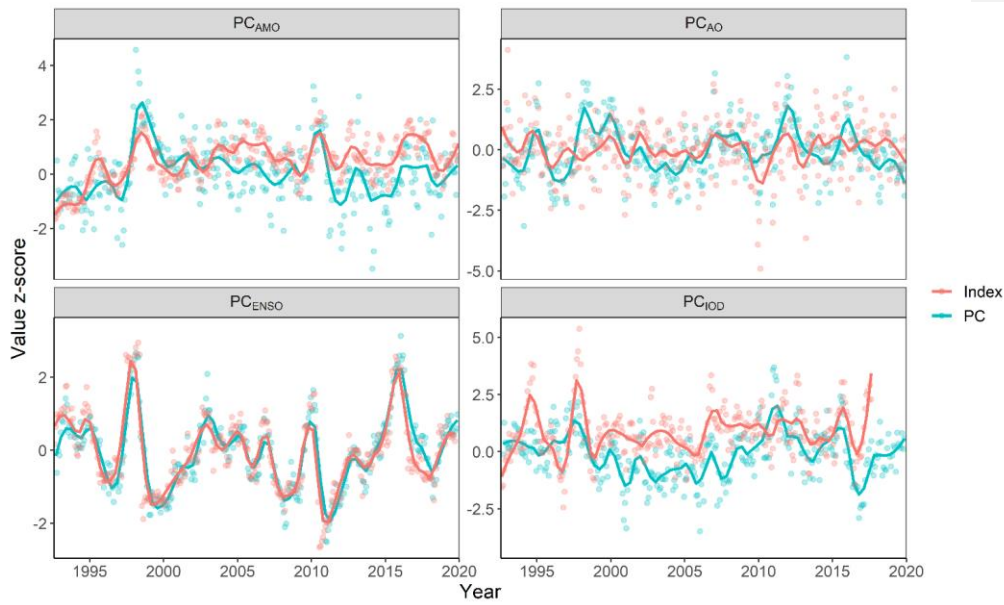
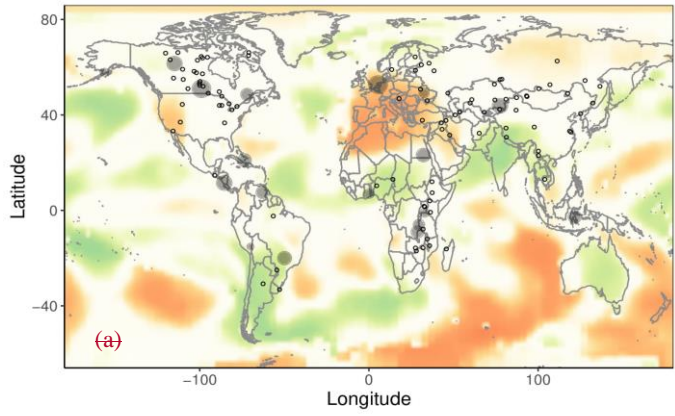
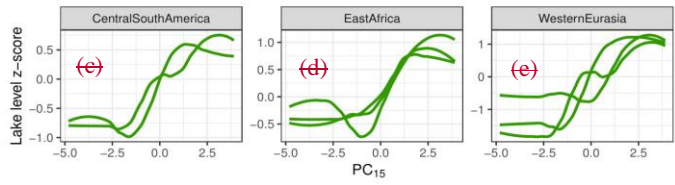
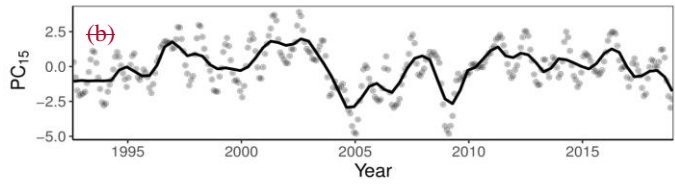


Figure 54: Time series of background climate variation for the three PCs that were most influential in the best models of water level variation on average across lakes. PCs ([PC<sub>AMO</sub>](#), [PC<sub>AO</sub>](#), [PC<sub>ENSO</sub>](#), [PC<sub>NAO</sub>](#), and [PC<sub>NPO</sub>](#)/[PC<sub>IOD</sub>](#)) and their corresponding climate indexes ([AMO](#), [AO](#), [ENSO](#), [NAO](#), and [NPO/IOD](#)) are transformed to their z-scores so that they can be more easily compared on the same, unit-less scale. The dots represent raw values and the lines represent locally-weight scatter plot (LOWESS)-smoothed time series.



Correlation between surface temperature and PC<sub>15</sub> -0.5 0.0 0.5

Relative Influence ● 0.7 ● 0.9 ● 1.1 ● 1.3  
● 0.8 ● 1.0 ● 1.2





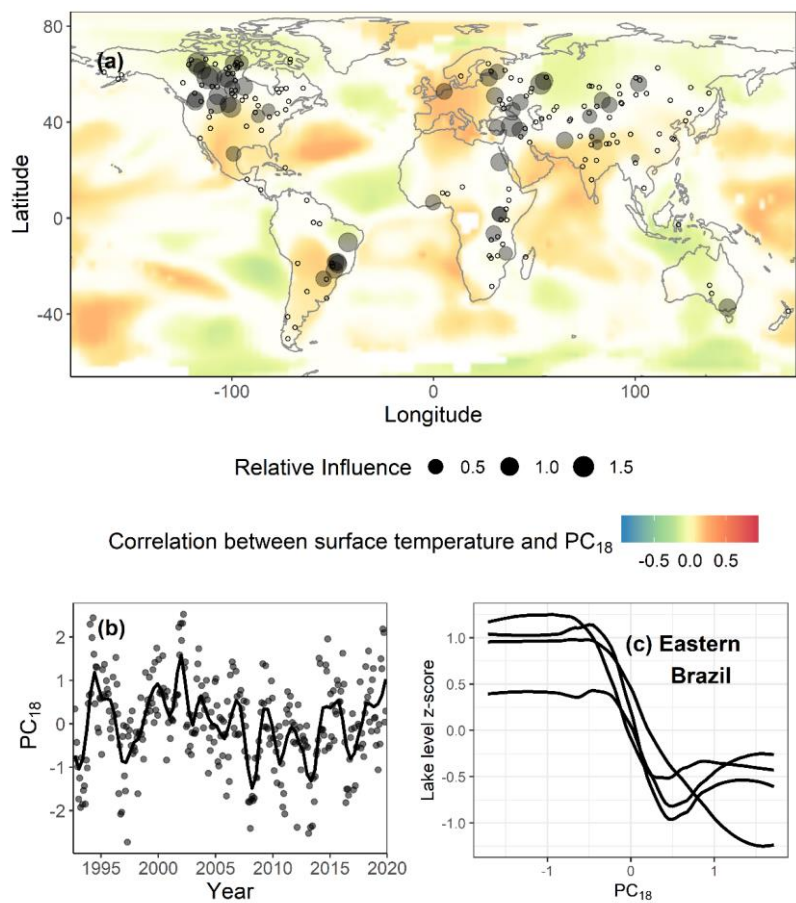
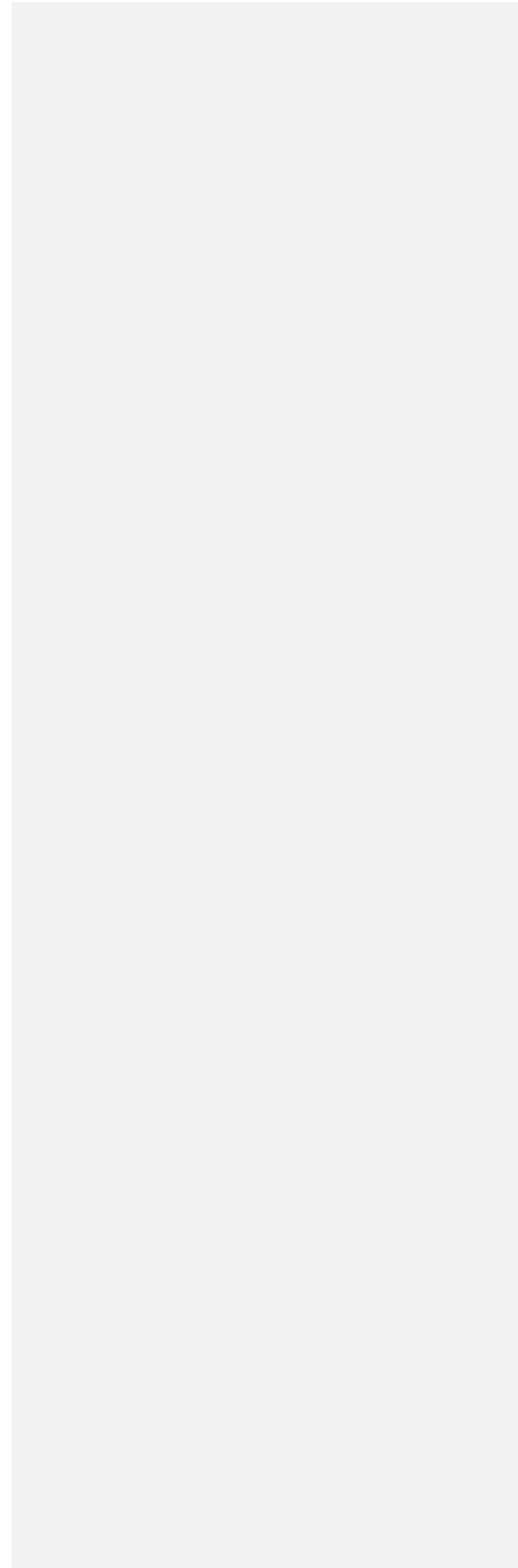
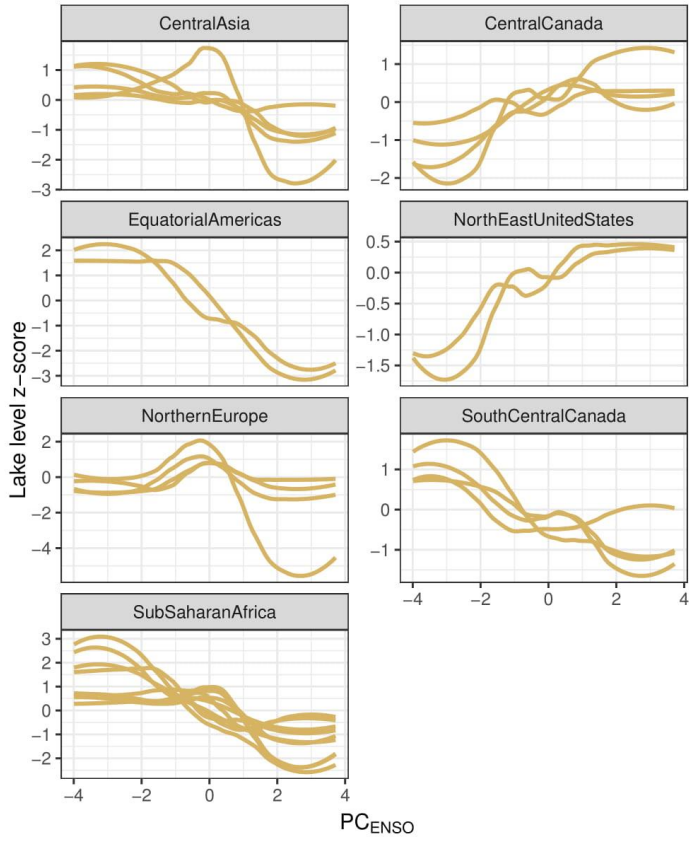


Figure 6: PC<sub>18</sub> is an example of a PC which was important for explaining water levels but was not strongly correlated with any of the climate indices recognized by NOAA. In panel (a), Empty dots represent lakes for which the PC<sub>18</sub> was not selected in its “best model.” The opacity of each colored pixel in each map is related to the significance of the correlation between the PC and temperature at that pixel with less significant correlations appearing white. In panel (b), the dots represent raw values of PC<sub>18</sub> and in panels (b-c) and (c), the line represents a locally-weight scatter plot (lines represent LOWESS)-smoothed values.



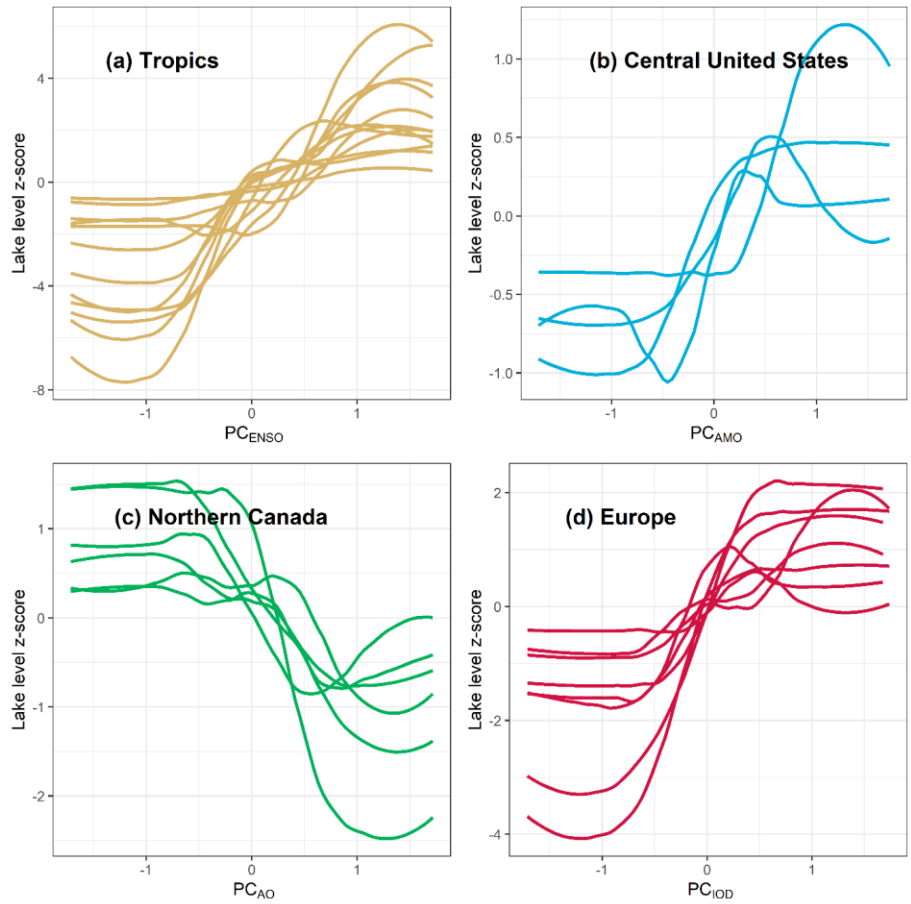
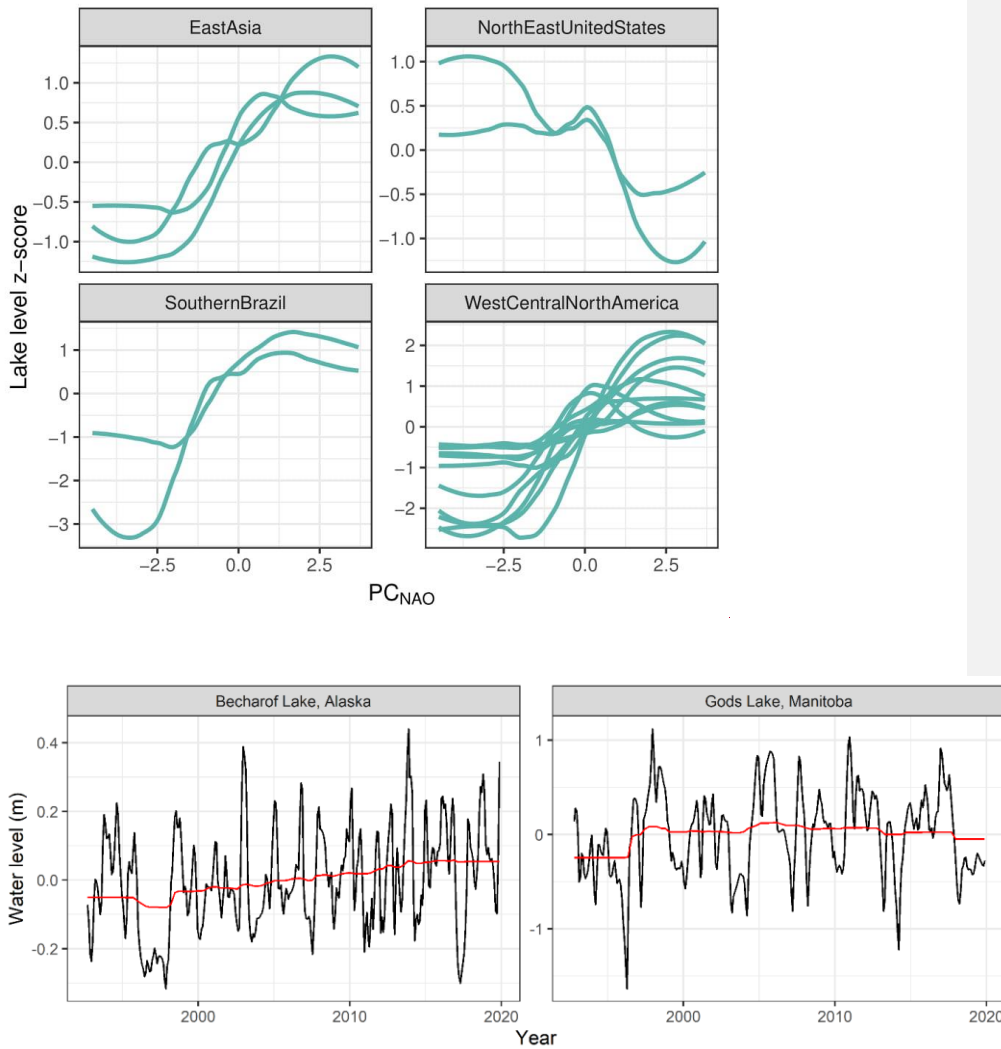


Figure 76: Regional consistently in the directionality of  $PC_{ENSO}$ ,  $PC_{AMO}$ ,  $PC_{AO}$ , and  $PC_{IOD}$  effects on water level variation. Each line represents a LOWESS-smoothed relationship between water level variation in a specific lake and  $PC_{ENSO}$ . Water levels have been transformed into z-scores so that they may all be plotted on the same axis.



Formatted: Add space between paragraphs of the same style

Figure 8: Regional consistency in the directionality of PC<sub>NAO</sub> effects on water level variation. Each line represents a LOWESS-smoothed relationship between water level variation in a specific lake and PC<sub>NAO</sub>. Water levels have been transformed into z-scores so that they may all be plotted on the same axis.

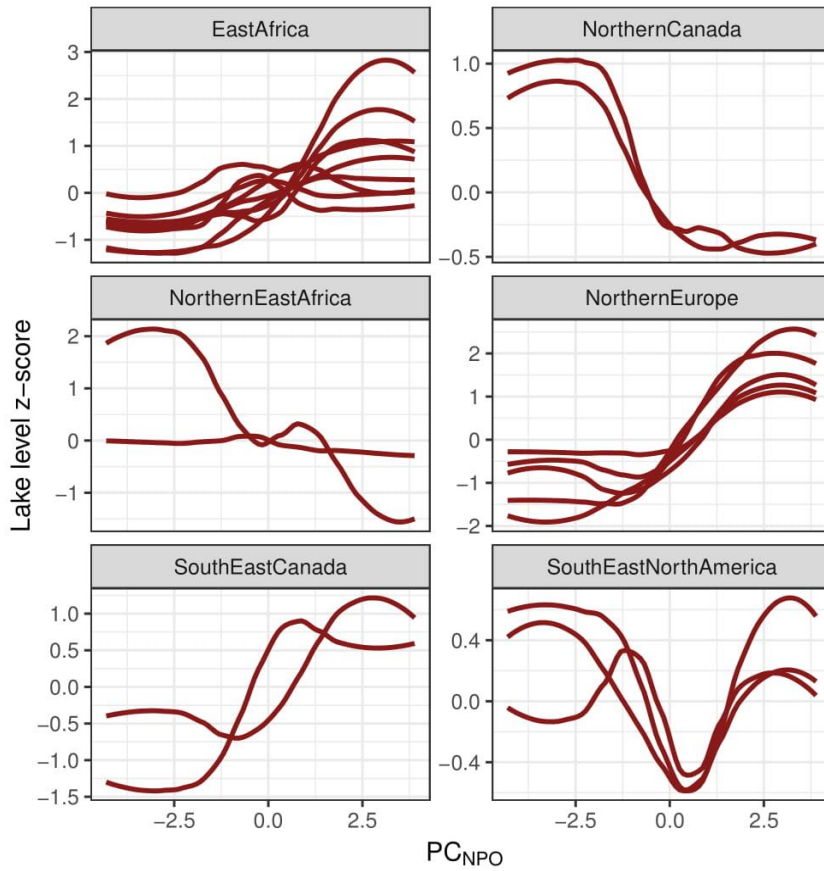


Figure 9: Regional consistency in the directionality of PC<sub>NPO</sub> effects on water level variation. Each line represents a LOWESS-smoothed relationship between water level variation in a specific lake and PC<sub>NPO</sub>. Water levels have been transformed into z-scores so that they may all be plotted on the same axis.

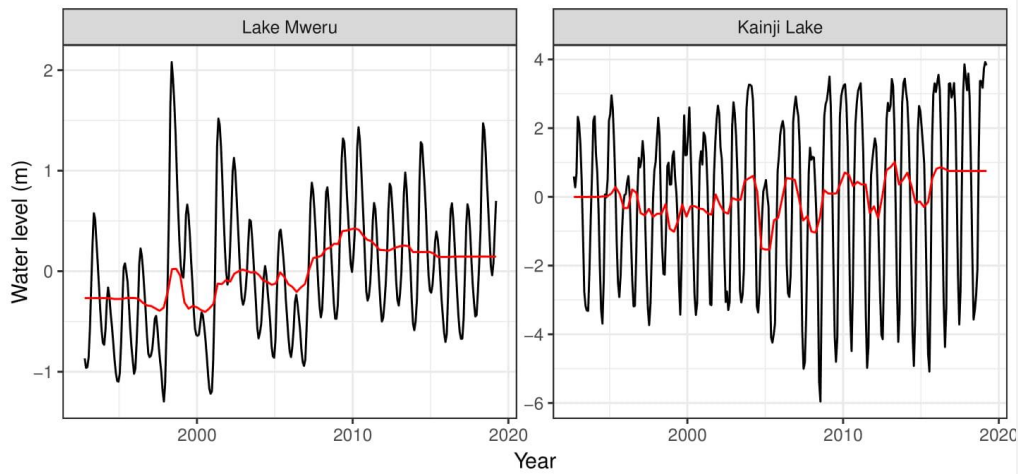


Figure 107: Water level time series for two tropical African lakes which show the potential for (a) background climate variation to mask long-term trends (Becharof Lake) and (b) the potential for climate variation to give rise to the false appearance of long-term trends (Kainji Lake). The water level trend in Lake Mweru on the Zambia-Democratic Republic of Congo border became more significant after accounting for background climate variation and the water level trend in Kainji Lake in western Nigeria became less significant after accounting for background climate variation. Raw data are shown here but all trends were calculated based on annual averages. The black line is the raw water level data and the red line is the partial dependence of water level on the year term in the boosted regression trees.

## References:

- Aladin, N. V., Plotnikov, I. S., Micklin, P. and Ballatore, T.: Natural Resources and Environmental Issues Aral Sea: Water level, salinity and long-term changes in biological communities of an endangered ecosystem-past, present and future, *Nat. Resour. Environ. Issues*, 15(36), 177–183 [online] Available from: <https://digitalcommons.usu.edu/nrei> (Accessed 7 April 2016), 2009.
- Allen, M. R. and Ingram, W. J.: Constraints on future changes in climate and the hydrologic cycle, *Nature*, 419(6903), 224–232, doi:10.1038/nature01092, 2002.
- Benson, L., Linsley, B., Smoot, J., Mensing, S., Lund, S., Stine, S. and Sarna-Wojcicki, A.: Influence of the Pacific decadal oscillation on the climate of the Sierra Nevada, California and Nevada, *Quat. Res.*, 59(2), 151–159, doi:10.1016/S0033-5894(03)00007-3, 2003.
- Benson, L. V., Lund, S. P., Burdett, J. W., Kashgarian, M., Rose, T. P., Smoot, J. P. and Schwartz, M.: Correlation of Late-Pleistocene Lake-Level Oscillations in Mono Lake, California, with North Atlantic Climate Events, *Quat. Res.*, 49(1), 1–10, doi:10.1006/qres.1997.1940, 1998.
- Bintanja, R. and Selten, F. M.: Future increases in Arctic precipitation linked to local evaporation and sea-ice retreat, *Nature*, 509(7501), 479–482, doi:10.1038/nature13259, 2014.
- Birkett, C., Reynolds, C., Beckley, B. and Doorn, B.: From research to operations: The USDA global reservoir and lake monitor, *Coast. Altimetry*, 19–50, doi:10.1007/978-3-642-12796-0\_2, 2011.
- Birkett, C. M. C.: The contribution of TOPEX/POSEIDON to the global monitoring of climatically sensitive lakes, *J. Geophys. Res.*, 100204(15), 179–25, doi:10.1029/95JC02125, 1995.
- Block, P. and Strzepek, K.: Power Ahead: Meeting Ethiopia’s Energy Needs Under a Changing Climate, *Rev. Dev. Econ.*, 16(3), 476–488, doi:10.1111/j.1467-9361.2012.00675.x, 2012.
- Byrne, M. P., O’Gorman, P. A. and O’Gorman, P. A.: The response of precipitation minus evapotranspiration to climate warming: Why the “Wet-get-wetter, dry-get-drier” scaling does not hold over land, *J. Clim.*, 28(20), 8078–8092, doi:10.1175/JCLI-D-15-0369.1, 2015.
- Cai, W., Santoso, A., Wang, G., Yeh, S. W., An, S. Il, Cobb, K. M., Collins, M., Guilyardi, E., Jin, F. F., Kug, J. S., Lengaigne, M., Mcphaden, M. J., Takahashi, K., Timmermann, A., Vecchi, G., Watanabe, M. and Wu, L.: ENSO and greenhouse warming, *Nat. Clim. Chang.*, 5(9), 849–859, doi:10.1038/nclimate2743, 2015.
- Cane, M. A.: The evolution of El Niño, past and future, *Earth Planet. Sci. Lett.*, 230(3–4), 227–240, doi:10.1016/j.epsl.2004.12.003, 2005.
- Capotondi, A. and Sardeshmukh, P. D.: Is El Niño really changing?, *Geophys. Res. Lett.*, 44(16), 8548–8556, doi:10.1002/2017GL074515, 2017.
- Chadwick, R., Boutle, I. and Martin, G.: Spatial patterns of precipitation change in CMIP5: Why the rich do not get richer in the tropics, *J. Clim.*, 26(11), 3803–3822, doi:10.1175/JCLI-D-12-00543.1, 2013.
- Chanut, J. P., D’astous, D. and El-Sabh, M. I.: Modelling the Natural and Anthropogenic Variations of the St. Lawrence Water Level, in *Natural and Man-Made Hazards*, pp. 377–394, Springer Netherlands, Dordrecht., 1988.
- Chen, Y., Morton, D. C., Andela, N., Giglio, L. and Randerson, J. T.: How much global burned area can be forecast on seasonal time scales using sea surface temperatures?, *Environ. Res. Lett.*, 11(4), 045001, doi:10.1088/1748-9326/11/4/045001, 2016.
- Christensen, N. S., Wood, A. W., Voisin, N., Lettenmaier, D. P. and Palmer, R. N.: The Effects

of Climate Change on the Hydrology and Water Resources of the Colorado River Basin, *Clim. Change*, 62(1–3), 337–363, doi:10.1023/B:CLIM.0000013684.13621.1f, 2004.

Clites, A. H., Smith, J. P., Hunter, T. S. and Gronewold, A. D.: Visualizing relationships between hydrology, climate, and water level fluctuations on Earth’s largest system of lakes, *J. Great Lakes Res.*, 40, 807–811, doi:10.1016/j.jglr.2014.05.014, 2014.

Cohn, B. P. and Robinson, J. E.: Cyclic fluctuations of water levels in Lake Ontario, *Comput. Geosci.*, 1(1–2), 105–108, doi:10.1016/0098-3004(75)90010-2, 1975.

Collins, M.: Understanding uncertainties in the response of ENSO to greenhouse warming, *Geophys. Res. Lett.*, 27(21), 3509–3512, doi:10.1029/2000GL011747, 2000.

Corti, S., Molteni, F., Palmer, T. N., Corti, S. and Molteni, F.: Signature of recent climate change in frequencies of natural atmospheric circulation regimes, *Nature*, 398(6730), 799–802, doi:10.1038/19745, 1999.

Crétau, J.-F. J.-F. and Birkett, C.: [Lake studies from satellite radar altimetry, \*Comptes Rendus - Geosci.\*, 338\(14–15\), 1098–1112, doi:10.1016/j.crte.2006.08.002, 2006.](#) [F., Abarca-del-Río, R., Bergé-Nguyen, M., Arsen, A., Drolon, V., Clos, G. and Maisongrande, P.: \*Lake Volume Monitoring from Space, \*Surv. Geophys.\*, 37\(2\), 269–305, doi:10.1007/s10712-016-9362-6, 2016.\*](#) [Crétau, J. F. J.-F. and Birkett, C.: \*Lake studies from satellite radar altimetry, \*Comptes Rendus - Geosci.\*, 338\(14–15\), 1098–1112, doi:10.1016/j.crte.2006.08.002, 2006.\*](#)

Dai, A. and Wigley, T. M. L.: Global patterns of ENSO-induced precipitation, *Geophys. Res. Lett.*, 27(9), 1283–1286, doi:10.1029/1999GL011140, 2000.

~~Dai, A., Fung, I. Y. and Del Genio, A. D.: *Surface Observed Global Land Precipitation Variations during 1900–88, *J. Clim.*, 10(11), 2943–2962, doi:10.1175/1520-0442(1997)010<2943:SOGLPV>2.0.CO;2, 1997.*~~

Dommenget, D. and Latif, M.: Generation of hyper climate modes, *Geophys. Res. Lett.*, 35(2), L02706, doi:10.1029/2007GL031087, 2008.

Dorigo, W., de Jeu, R., Chung, D., Parinussa, R., Liu, Y., Wagner, W. and Fernández-Prieto, D.: Evaluating global trends (1988–2010) in harmonized multi-satellite surface soil moisture, *Geophys. Res. Lett.*, 39(18), doi:10.1029/2012GL052988, 2012.

Elith, J., Leathwick, J. and Hastie, T.: A working guide to boosted regression trees, *J. Anim. Ecol.*, 77(4), 802–813, 2008.

Food and Agriculture Organization of the United Nations: AQUASTAT, [online] Available from: <http://www.fao.org/nr/water/aquastat/main/index.stm> (Accessed 11 July 2018), 2016.

Fowler, H. J., Blenkinsop, S. and Tebaldi, C.: Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling, *Int. J. Climatol.*, 27(12), 1547–1578, doi:10.1002/joc.1556, 2007.

Ghanbari, R. N. and Bravo, H. R.: Coherence between atmospheric teleconnections, Great Lakes water levels, and regional climate, *Adv. Water Resour.*, 31(10), 1284–1298, doi:10.1016/j.advwatres.2008.05.002, 2008.

Gimeno, L., Nieto, R., Vázquez, M. and Lavers, D. A.: Atmospheric rivers: a mini-review, *Front. Earth Sci.*, 2, doi:10.3389/feart.2014.00002, 2014.

Greve, P. and Seneviratne, S. I.: Assessment of future changes in water availability and aridity, *Geophys. Res. Lett.*, 42(13), 5493–5499, doi:10.1002/2015GL064127, 2015.

Greve, P., Orłowsky, B., Mueller, B., Sheffield, J., Reichstein, M. and Seneviratne, S. I.: Global assessment of trends in wetting and drying over land, *Nat. Geosci.*, 7(10), 716–721, doi:10.1038/NGEO2247, 2014.

Gronewold, A. D., Fortin, V., Lofgren, B., Clites, A., Stow, C. A. and Quinn, F.: Coasts, water



levels, and climate change: A Great Lakes perspective, *Clim. Change*, 120(4), 697–711, doi:10.1007/s10584-013-0840-2, 2013.

Guilyardi, E., Wittenberg, A., Fedorov, A., Collins, M., Wang, C., Capotondi, A., van Oldenborgh, G. J. and Stockdale, T.: Understanding El Niño in ocean-atmosphere general circulation models: Progress and challenges, *Bull. Am. Meteorol. Soc.*, 90(3), 325–340, doi:10.1175/2008BAMS2387.1, 2009.

Gyau-Boakye, P.: Environmental impacts of the Akosombo dam and effects of climate change on the lake levels, *Environ. Dev. Sustain.*, 3, 17–29, doi:10.1023/A:1011402116047, 2001.

Hannachi, A., Unkel, S., Trendafilov, N. T. and Jolliffe, I. T.: Independent component analysis of climate data: A new look at EOF rotation, *J. Clim.*, 22(11), 2797–2812, doi:10.1175/2008JCLI2571.1, 2009.

Hansen, J., Ruedy, R., Sato, M. and Lo, K.: Global surface temperature change, *Rev. Geophys.*, 48(4), RG4004, doi:10.1029/2010RG000345, 2010.

Hansen, J. W., Hodges, A. W. and Jones, J. W.: ENSO influences on agriculture in the southeastern United States, *J. Clim.*, 11(3), 404–411, doi:10.1175/1520-0442(1998)011<0404:EIOAIT>2.0.CO;2, 1998.

Hassanzadeh, E., Zarghami, M. and Hassanzadeh, Y.: Determining the Main Factors in Declining the Urmia Lake Level by Using System Dynamics Modeling, *Water Resour. Manag.*, 26(1), 129–145, doi:10.1007/s11269-011-9909-8, 2012.

Hegerl, G. C., Black, E., Allan, R. P., Ingram, W. J., Polson, D., Trenberth, K. E., Chadwick, R. S., Arkin, P. A., Sarojini, B. B., Becker, A., Dai, A., Durack, P. J., Easterling, D., Fowler, H. J., Kendon, E. J., Huffman, G. J., Liu, C., Marsh, R., New, M., Osborn, T. J., Skliris, N., Stott, P. A., Vidale, P. L., Wijffels, S. E., Wilcox, L. J., Willett, K. M. and Zhang, X.: Challenges in quantifying changes in the global water cycle, *Bull. Am. Meteorol. Soc.*, doi:10.1175/BAMS-D-13-00212.1, 2015.

Hidalgo, H. G. and Dracup, J. A.: ENSO and PDO Effects on Hydroclimatic Variations of the Upper Colorado River Basin, *J. Hydrometeorol.*, 4(1), 5–23, doi:10.1175/1525-7541, 2003.

Hijmans, R. J., Phillips, S., Leathwick, J. R. and Elith, J.: Package ‘dismo’, October, doi:10.1016/j.jhydrol.2011.07.022., 2017.

IPCC: IPCC Fifth Assessment Report, Fifth Assess. Rep., s. 10-12 [online] Available from: <https://www.ipcc.ch/report/ar5/index.shtml>, 2014.

Jöhnk, K. D., Straile, D. and Ostendorp, W.: Water level variability and trends in lake constance in the light of the 1999 centennial flood, *Limnologica*, 34(1–2), 15–21, doi:10.1016/S0075-9511(04)80017-3, 2004.

Kim, K. Y. and Wu, Q.: A comparison study of EOF techniques: Analysis of nonstationary data with periodic statistics, *J. Clim.*, 12(1), 185–199, doi:10.1175/1520-0442-12.1.185, 1999.

Lau, W. K. M. and Kim, K. M.: The 2010 Pakistan flood and Russian heat wave: Teleconnection of hydrometeorological extremes, *J. Hydrometeorol.*, 13(1), 392–403, doi:10.1175/JHM-D-11-016.1, 2012.

Lenters, J. D.: Long-term Trends in the Seasonal Cycle of Great Lakes Water Levels, *J. Great Lakes Res.*, 27(3), 342–353, doi:10.1016/S0380-1330(01)70650-8, 2001.

Lorenzo, D., E, Lorenzo, E. Di, Schneider, N., Cobb, K. M., Franks, P. J. S., Chhak, K., Miller, A. J., McWilliams, J. C., Bograd, S. J., Arango, H., Curchitser, E., Powell, T. M., Rivie, P., Di Lorenzo, E., Schneider, N., Cobb, K. M., Franks, P. J. S., Chhak, K., Miller, A. J., McWilliams, J. C., Bograd, S. J., Arango, H., Curchitser, E., Powell, T. M. and Rivière, P.: North Pacific Gyre Oscillation links ocean climate and ecosystem change, *Geophys. Res. Lett.*, 35(8), 1–6,

doi:10.1029/2007GL032838, 2008.

Marchant, R., Mumbi, C., Behera, S. and Yamagata, T.: The Indian Ocean dipole ? the unsung driver of climatic variability in East Africa, *Afr. J. Ecol.*, 45(1), 4–16, doi:10.1111/j.1365-2028.2006.00707.x, 2007.

Mekonnen, M. M. and Hoekstra, A. Y.: Four billion people facing severe water scarcity, *Sci. Adv.*, 2(2), e1500323, doi:10.1126/sciadv.1500323, 2016.

Mercier, F., Cazenave, A. and Maheu, C.: Interannual lake level fluctuations (1993–1999) in Africa from Topex/Poseidon: connections with ocean–atmosphere interactions over the Indian Ocean, *Glob. Planet. Change*, 32(2–3), 141–163, doi:10.1016/S0921-8181(01)00139-4, 2002.

Messenger, M. L., Lehner, B., Grill, G., Nedeva, I. and 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/ncomms13603, 2016.

Micklin, P. P.: Desiccation of the Aral Sea: A Water Management Disaster in the Soviet Union, *Science* (80-. ), 241(4870), 1170–1176, doi:10.1126/science.241.4870.1170, 1988.

Molinos, J. and Donohue, I.: Downscaling the non-stationary effect of climate forcing on local-scale dynamics: The importance of environmental filters, *Clim. Change*, 124(1–2), 333–346, doi:10.1007/s10584-014-1077-4, 2014.

Molinos, J. G., Viana, M., Brennan, M. and Donohue, I.: Importance of long-term cycles for predicting water level dynamics in natural lakes, edited by I. Álvarez, *PLoS One*, 10(3), e0119253, doi:10.1371/journal.pone.0119253, 2015.

Muller, M.: Cape Town’s drought: don’t blame climate change, *Nature*, 559(7713), 174–176, doi:10.1038/d41586-018-05649-1, 2018.

Nijssen, B., O’donnell, G. M., Hamlet, A. F. and Lettenmaier, D. P.: Hydrologic sensitivity of global rivers to climate change, *Clim. Change*, 50(1–2), 143–175, doi:10.1023/A:1010616428763, 2001.

O’Gorman, P. A., Allan, R. P., Byrne, M. P. and Previdi, M.: Energetic Constraints on Precipitation Under Climate Change, *Surv. Geophys.*, 33(3–4), 585–608, doi:10.1007/s10712-011-9159-6, 2012.

van Oldenborgh, G. J., Philip, S. Y. and Collins, M.: El Niño in a changing climate: a multi-model study, *Ocean Sci.*, 1(2), 81–95, doi:10.5194/os-1-81-2005, 2005.

Pasquini, A. I., Lecomte, K. L. and Depetris, P. J.: Climate change and recent water level variability in Patagonian proglacial lakes, Argentina, *Glob. Planet. Change*, 63(4), 290–298, doi:10.1016/j.gloplacha.2008.07.001, 2008.

Pekel, J.-F., Cottam, A., Gorelick, N. and Belward, A. S.: High-resolution mapping of global surface water and its long-term changes, *Nature*, 540(7633), 418–422, doi:10.1038/nature20584, 2016.

[Phinyomark, A., Hettinga, B. A., Osis, S. and Ferber, R.: Do intermediate- and higher-order principal components contain useful information to detect subtle changes in lower extremity biomechanics during running?, \*Hum. Mov. Sci.\*, doi:10.1016/j.humov.2015.08.018, 2015.](#)

R Core Team: R: A Language and Environment for Statistical Computing, R Found. Stat. Comput. [online] Available from: <http://www.r-project.org/>, 2017.

Ramanathan, V., Crutzen, P. J., Kiehl, J. T. and Rosenfeld, D.: Aerosols, climate, and the hydrological cycle., *Science* (80-. ), 294(5549), 2119–24, doi:10.1126/science.1064034, 2001.

Ricko, M., Carton, J. A. and Birkett, C.: Climatic Effects on Lake Basins. Part I: Modeling Tropical Lake Levels, *J. Clim.*, 24(12), 2983–2999, doi:10.1175/2010JCLI3602.1, 2011.

Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoin, H. K., Landerer, F. W. and

Lo, M.-H.: Emerging trends in global freshwater availability, *Nature*, doi:10.1038/s41586-018-0123-1, 2018.

Ropelewski, C. F. and Halpert, M. S.: Global and Regional Scale Precipitation Patterns Associated with the El Niño/Southern Oscillation, *Mon. Weather Rev.*, 115(8), 1606–1626, doi:10.1175/1520-0493(1987)115<1606:GARSPP>2.0.CO;2, 1987.

Stager, J. C., Ryves, D., Cumming, B. F., David Meeker, L. and Beer, J.: Solar variability and the levels of Lake Victoria, East Africa, during the last millenium, *J. Paleolimnol.*, 33(2), 243–251, doi:10.1007/s10933-004-4227-2, 2005.

Stager, J. C., Ruzmaikin, A., Conway, D., Verburg, P. and Mason, P. J.: Sunspots, El Niño, and the levels of Lake Victoria, East Africa, *J. Geophys. Res. Atmos.*, 112(15), doi:10.1029/2006JD008362, 2007.

Stenseth, N. C., Ottersen, G., Hurrell, J. W., Mysterud, A., Lima, M., Chan, K. S., Yoccoz, N. G. and Ådlandsvik, B.: Studying climate effects on ecology through the use of climate indices: The North Atlantic Oscillation, El Niño Southern Oscillation and beyond, *Proc. R. Soc. B Biol. Sci.*, 270(1529), 2087–2096, doi:10.1098/rspb.2003.2415, 2003.

Tao, S., Fang, J., Zhao, X., Zhao, S., Shen, H., Hu, H., Tang, Z., Wang, Z. and Guo, Q. and Turner, B. L.: Rapid loss of lakes on the Mongolian Plateau, *Proc. Natl. Acad. Sci. U. S. A.*, 112(7), 2281–2286, doi:10.1073/pnas.1411748112, 2015.

Tierney, J. E., Russell, J. M., Huang, Y., Damsté, J. S. S., Hopmans, E. C. and Cohen, A. S.: Northern hemisphere controls on tropical southeast African climate during the past 60,000 years., *Science (80-. )*, 322(5899), 252–255, 2008.

Tierney, J. E., Smerdon, J. E., Anchukaitis, K. J. and Seager, R.: Multidecadal variability in East African hydroclimate controlled by the Indian Ocean., *Nature*, 493(7432), 389–92, doi:10.1038/nature11785, 2013.

Timmermann, A., Oberhuber, J., Bacher, A., Esch, M., Latif, M. and Roeckner, E.: Increased El Niño frequency in a climate model forced by future greenhouse warming, *Nature*, 398(6729), 694–697, doi:10.1038/19505, 1999.

Tomasion, M. and Valle, F. D.: Natural climatic changes and solar cycles: An analysis of hydrological time series, *Hydrol. Sci. J.*, 45(3), 491, doi:10.1080/02626660009492344, 2000.

Vadeboncoeur, Y., McIntyre, P. B. and Vander Zanden, M. J.: Borders of biodiversity: life at the edge of the world's large lakes, *Bioscience*, 61(7), 526–537, doi:10.1525/bio.2011.61.7.7, 2011.

Vinukollu, R. K., Meynadier, R., Sheffield, J. and Wood, E. F.: Multi-model, multi-sensor estimates of global evapotranspiration: climatology, uncertainties and trends, *Hydrol. Process.*, 25(26), 3993–4010, doi:10.1002/hyp.8393, 2011.

Vörösmarty, C. J. and Sahagian, D.: Anthropogenic Disturbance of the Terrestrial Water Cycle, *Bioscience*, 50(9), 753, doi:10.1641/0006-3568(2000)050[0753:adottw]2.0.co;2, 2006.

Vörösmarty, C. J., Green, P., Salisbury, J. and Lammers, R. B.: Global water resources: Vulnerability from climate change and population growth, *Science (80-. )*, 289(5477), 284–288, doi:10.1126/science.289.5477.284, 2000.

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S. E., Sullivan, C. A., Liermann, C. R. and Davies, P. M.: Global threats to human water security and river biodiversity, *Nature*, 467(7315), 555–561, doi:10.1038/nature09440, 2010.

Wang, B., Liu, J., Kim, H. J., Webster, P. J. and Yim, S. Y.: Recent change of the global monsoon precipitation (1979-2008), *Clim. Dyn.*, 39(5), 1123–1135, doi:10.1007/s00382-011-1266-z, 2012.

[Wang, J., Yang, B., Ljungqvist, F. C., Luterbacher, J., Osborn, T. J., Briffa, K. R. and Zorita, E.: Internal and external forcing of multidecadal Atlantic climate variability over the past 1,200 years, Nat. Geosci., doi:10.1038/ngeo2962, 2017.](#)

[Wang, S. Y., Gillies, R. R., Jin, J. and Hipps, L. E.: Coherence between the great salt lake level and the Pacific quasi-decadal oscillation, J. Clim., 23\(8\), 2161–2177, doi:10.1175/2009JCLI2979.1, 2010.](#)

Wrzesiński, D., Ptak, M. and Plewa, K.: Effect of the north Atlantic oscillation on water level fluctuations in lakes of northern Poland, Geogr. Pol., doi:10.7163/gpol.0119, 2018.

Wurtsbaugh, W. A., Miller, C., Null, S. E., Justin De Rose, R., Wilcock, P., Hahnenberger, M., Howe, F. and Moore, J.: Decline of the world's saline lakes, Nat. Geosci., 10(11), 816–821, doi:10.1038/NCEO3052, 2017.

[Zhang, G., Xie, H., Kang, S., Yi, D. and Ackley, S. F.: Monitoring lake level changes on the Tibetan Plateau using ICESat altimetry data \(2003–2009\), Remote Sens. Environ., 115\(7\), 1733–1742, doi:10.1016/j.rse.2011.03.005, 2011.](#)