**Unique multi-century lake and river ice records provide insights into changing climate and variability**

**Sapna Sharma, John J. Magnuson, Ryan Batt, Luke Winslow, and Johanna Korhonen**

**Abstract**

We used direct human observations of ice phenology for Suwa Lake, Japan, and Torne River, Finland, from the past 5-7 centuries to characterize long-term changes in climate and variability. Increased rates of warming followed breakpoints of 1812 in Suwa (slopes) and 1867 in Torne (slopes), coinciding generally with the end of the Little Ice Age and the start of the Industrial Revolution. Variability in timing of ice formation and breakup using 30-year windows decreased or did not change over the extent of the records. Significant periodicities of oscillations did not persist over the entire time series, suggesting structural changes in teleconnections of climate drivers in response to a changing climate. Elevated atmospheric CO2 concentrations and air temperatures have the strongest relationships to later ice formation and earlier ice breakup, particularly after the breakpoints, suggesting the role of climate change in changing ice phenologies.

**Main text**

Direct annual observations by humans of climatic variables over 5 to 7 centuries are rare, as most long-term climate time series have been inferred using paleo-chronologies. We analyze the two longest known contemporary ice records from inland waters, the timing of ice breakup from 1443-2004 for Suwa Lake, Japan and the timing of ice breakup from 1692-2013 for Torne River, Finland. We identified changing trends, variability, and periodicity in climate using observational records that encompassed major global events, including the Little Ice Age, a ~ 40% increase in atmospheric CO2 concentrations since the start of the Industrial Revolution, and substantial land use change associated with urbanization, deforestation, and agriculture (IPCC 2013).

The Suwa Lake and Torne River ice records were collected primarily for religious or economic purposes, but provide unique opportunities to analyze and infer climatic change and variability over centuries. The long-term ice freeze record for Suwa was collected by Shinto priests. The Shinto legend is that the male God would cross the lake to visit the female God at her shrine across the lake. The footsteps on the ice of the male God left a sinusoidal ice ridge known as the *omiwatari* in Japanese. This important event was celebrated and recorded by at least 15 generations of Shinto priests since 1443 (Arakawa 1954; Mikami 2008; Takasaki, pers. comm). For Torne River, the tradition of recording the date of ice breakup was important to the river’s role in trade, transportation, food, and recreation (Loader et al. 2011).

Importantly, these two ice records, beginning several centuries before the start of the Industrial Revolution, allow us to identify whether there have been changes in trends, variability, and periodicities since the beginning of the Industrial Revolution. In addition, the direct human observations over 5-7 centuries, potentially allows detection of oscillations with longer periods than are possible with shorter time series with more certainty, identification of how relationships with drivers have changed over time, and to judge whether direct observations are consistent with findings using paleo-inferred climate records.

*Increasing trends of warming*

Both Suwa and Torne reveal warming trends with later ice freeze or earlier ice breakup with a significant shift to more rapid rates of change occurring in the 1800s. Breakpoints, defined as the year in which linear slopes changed significantly (see Supplementary Material for detailed methods), were 1812 for Suwa and 1867 for Torne (Figure 1a and b). For Suwa, the trend prior to the breakpoint was \_\_\_\_\_ days per decade and increased to \_\_\_\_\_ days to decade afterwards. The trends in ice breakup of Torne increased from \_\_\_\_\_\_\_ days per decade to \_\_\_\_\_\_\_ days per decade. The timing of the breakpoints is consistent with the near-end of the Little Ice Age and the start of the Industrial Revolution in both regions. For Suwa, previous studies based on diaries of temperature and precipitation (Mikami 2008), paleo-reconstructions from the sediment core of Lake Nakatsuna in central Japan (Adhikari and Kumon 2001), and cherry blossom flowering times in Tokyo (Primack et al. 2009), all indicate that climate began to warm in the 1810s following very cold temperatures in the region associated with the Little Ice Age. Since the 1950s, cherry blossoms are flowering earlier with the earliest blooming records occurring between 1971-2000 compared to any point in the past 1200 years, which has been attributed to urbanization and climate change (Aono and Kazui, 2008; Primack et al. 2009). Coincidentally, on Suwa, this was during the time period when the lake had the highest incidence of not freezing (Figure 1c). In the most recent 55-year period (1950-2004), Suwa did not freeze twelve years compared to only three years over a 255-year earlier period (1443-1700; Figure 1c). In contrast, the Torne River, at a higher latitude than Suwa Lake, froze every year since 1693. The 1867 breakpoint in the Torne occurs in the year with the coolest spring (April-May temperatures) in Happaranda and Stockholm since 1756 (Loader et al. 2011). Coincidentally, 1867 was also the year with the latest ice breakup in 3 Finnish lakes (Kuuisito 1987). Following the breakpoint and the end of the Little Ice Age within the region, spring air temperatures were reported to rise (Loader et al. 2011; Helama et al. 2013). Other factors within the region began changing in the 1880s including increased urbanization, deforestation, and human population growth from 1,000 to 10,000 people (Loader et al. 2011; Helama et al. 2013).

*Changing variability*

The long-term changes in variability in the timing of ice formation and breakup at an extent of 320-550 years and a grain size of 30 years do not indicate that variability is increasing (Figure 2a and b). Variability was quantified using two metrics: standard deviation and the mean first difference (Karl et al. 1995; see Supplementary Materials for detailed methods). Both metrics produced similar results. For Suwa, variability since 1800 has been lower than that observed in the late 1500s and early 1600s whereas variability in the late 1400s to early 1500s is similar to that of recent years (Figure 2a). In the Torne, variability in ice breakup date has generally decreased from 1700 to present (Figure 2b).

All other published observational studies of variability in ice dates and climate have been conducted at much shorter time scales, ranging from 20 to at most 150 years. At time scales between 20-50 years in extent, variability in air temperature and ice phenology has been shown to be increasing in recent years (Kratz et al. 2000; Schar et al. 2004; Weyhenmeyer et al. 2011). Karl (1995), however, noted that variance in climate decreases with a longer temporal extent, for example the variance in daily air temperature decreased over a 90-year record across the United States. Similarly, Benson et al. (2012) illustrated that inter-annual variability in ice phenology decreased or did not change when examining lake ice records at 100 and 150-year extents across the Northern Hemisphere. Although, variability appeared to increase in the most recent 50 years, it did not attain levels higher than the previous 100 years (Benson et al. 2012). Clearly, variability is dependent upon the grain and extent of the time series and long-term climatic variability appears not to be increasing when examining longer time series (Karl 1995; Benson et al. 2012).

*Changing periodicities*

The powers associated with periodicities among years have changed over the years for both Suwa and Torne based on a Morlet wavelet analysis (Figure 2). Significant periods of oscillation did not persist across the entire time series (Figure 2c and d). Overall for Suwa, significant periods of oscillations were apparent from 32 up to at least 64 years in the power spectrum (Figure 2c). More specifically, we approximate significant periods of oscillations ranging from 16 to 48 years between 1480 and 1570, 32 to 64 years between 1670 and 1830, 6 to 8 years between 1760 and 1780, and 16-32 years between 1780 and 1840. After 1850, there was only one significant period at 64 years in 1870-1930 (Figure 2c).

Overall for Torne, periods of oscillations with the highest power were apparent from 3-5 years, 8-10 years, and 24-32 years (Figure 2d). Between 1700 and 1750, there were two significant periods of oscillations between 3-5 years and 7-13 years. Between 1750 and 1860, significant periods ranging from 3-6 years, 20-41 years, and 58-68 years. Following the breakpoint, there were significant periods of oscillations from 8-11 years between 1890-1930 and 12-16 years between 1940 and 1960 (Figure 2d). Interestingly, both inland waters from geographically disparate regions exhibit significant periods of oscillation of 16 and 64 years in the mid-1700s to the mid-1800s. These longer periods of oscillation are not significant in the time series after 1850 in Suwa and Torne (Figures 2c and d).

We found no evidence that oscillations that were observed before 1860 persisted consistently afterwards in Suwa and Torne. This change in significant periodicities over time suggests a structural change in teleconnections among large-scale climate drivers, including NAO and ENSO, in a warming post-Industrial Revolution climate (Hurrell and van Loon 1997; Higuchi et al. 1999; Robertson et al. 2000; Yoo and D’Odorico 2002; Li et al. 2011). For example, it appears that a shift has occurred in the importance of inter-annual cycles of the NAO and that increasing CO2 concentrations may have stabilized the positive phase of the NAO (Yoo and D’Odorico 2002). In the latter half of the 20th century, the 6-to-10 year period explained more variance in the NAO index than earlier, whereas the contribution from inter-decadal periods was almost absent from 1940 to the 1970s (Hurrell and van Loon 1997; Higuchi et al. 1999). An ENSO reconstruction over the past 1100 years using the North American Drought Atlas has suggested a shift in ENSO cycles to shorter periods over time (Li et al. 2011). For example, between 900-1300 AD, 82-90 year ENSO cycles were evident, shifting to 50-60 year cycles in 1300-1500, 30-year cycles in 1500-1800, and 2-8 year cycles in the contemporary time period (Li et al. 2011). In Wisconsin lakes, the relationship between El Niňo events and later ice breakup appears to be weaker or even switch after 1940 (Livingstone 2000; Robertson et al. 2000). Prior to 1940, cooler air temperatures were evident in late winter of El Niňo years, whereas following 1940 warmer air temperatures were evident in late winter of El Niňo years (Robertson et al. 2000). Our findings using observed data generally support significant changes of climate drivers over the centuries.

*Drivers are changing over time*

Results from our breakpoint analysis suggest that both systems experienced shifts in the temporal trend of ice date. Despite the geographical distance and season separating these time series, breakpoints were identified at relatively similar points in time (1812 in Suwa Lake, 1867 in Torne River; Figure 1). Thus, shifts in ice date trends are unlikely to be driven solely by system-specific forcings. We explored linear relationships before and after the breakpoints between ice dates and climate drivers, i.e., air temperature, aerosol optic depth, atmospheric CO2, ENSO, NAO, and sunspots (See supplementary material for additional details).

Increases in atmospheric CO2 concentrations were significantly related to later dates of ice formation in Suwa and earlier ice breakup in Torne after the breakpoints (Figure 3). To our knowledge, this is the first study to document a direct statistical link between increasing atmospheric CO2 concentrations and later ice freeze and earlier ice breakup following the start of the Industrial Revolution. The timing of ice breakup may especially be sensitive to CO2 concentrations and warming may be particularly pronounced in winter and early spring at higher latitudes (Yoo and D’Odorico 2002). This suggests a role for climate change in the alterations of ice phenology following increased greenhouse gas emissions.

Warmer falls and winters were related to later freeze dates in Suwa after the breakpoint and warmer springs were related to earlier ice breakup date in Torne over the time series. Years with elevated air temperatures have correspondingly shorter duration of ice coverage across the Northern Hemisphere (Benson et al. 2012). The only other significant driver we observed was NAO after the breakpoint in Torne (Figure 3). Indices of large-scale climate drivers do not tend to explain as much variation in ice breakup as changing air temperatures (Sharma et al. 2013).

*Conclusions*

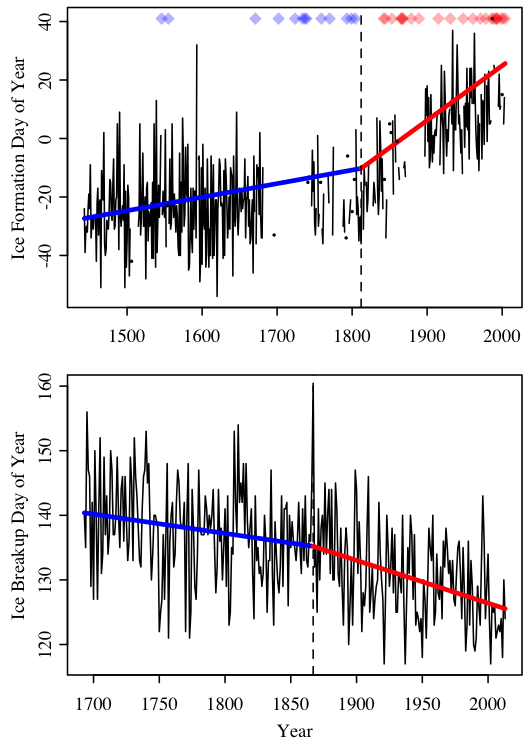
Direct, human observations of lake and river ice records from the past 5-7 centuries are unique. These multi-century ice records provide insights into changing climate and variability through a new lens. We identify that following the breakpoints coinciding with the start of the Industrial Revolution, warming rates increased, variability decreased or did not change significantly, significant short and long periods of oscillations disappeared, and later dates of ice formation and earlier breakup were related significantly to increases in atmospheric CO2 concentrations and air temperatures. Results from these direct human observations are generally consistent with findings using paleo-inferred climate records, reinforcing the importance of using a diverse suite of data types to assess long-term changes in climate and variability.

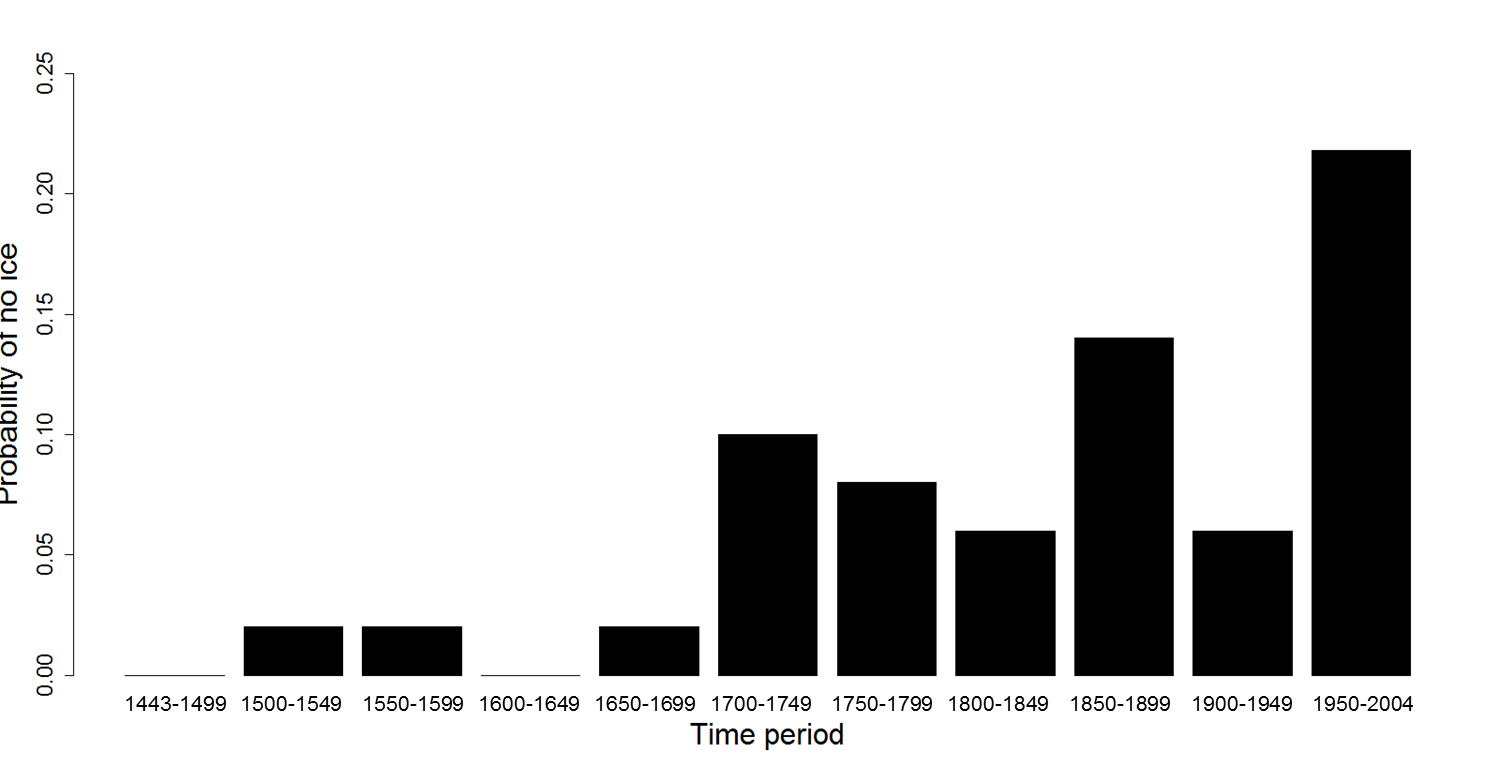
**Acknowledgments**

We thank Drs. Tadashi Arai and Takayuki Hanazato for providing a closer view of Lake Suwa and the generation of the remarkable dataset. We thank Drs. Barbara Benson and Jonathan Ruppert for their work in the early phases of the project and Mayu Takasaki for her translations of Japanese text regarding the legend of *omiwatari*. We acknowledge the Natural Sciences and Engineering Research Council, York University, and the North Temperate Lakes Long Term Ecological Research Network for funding.

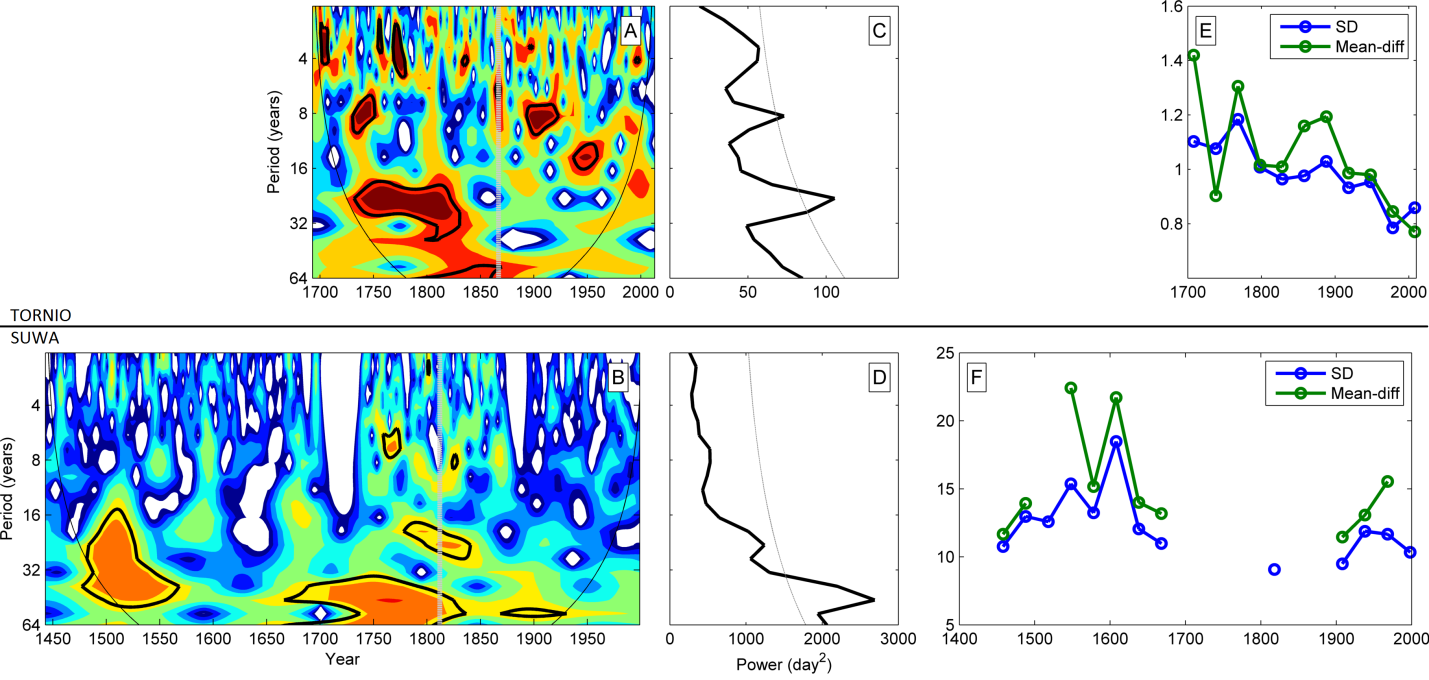
**References**

**Figures**

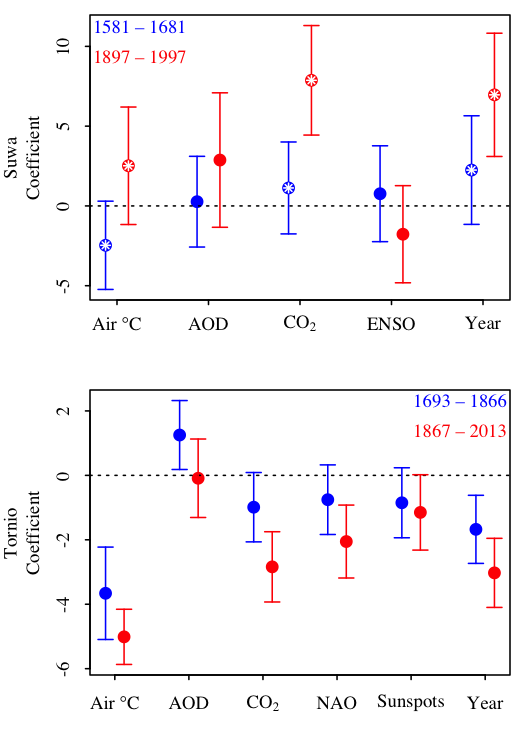
****



**Figure 1.** Long-term observational records of ice dates for Suwa Lake in Japan (**A**, ice formation) and for the Torne River in Finland (**B**, ice breakup) and **(C)** probability of no ice on Suwa Lake for each 50-year period from 1443-2004, except for the first and last period which have a few additional years. For **(A)** and **(B)**, solid black lines and black dots indicate ice dates (black dots are used when no observations were made during adjacent years). Blue and red diamonds indicate years when observations were made, but the lake did not freeze. Vertical dashed lines are placed at the year of the breakpoint in the timing of ice formation or breakup. Thick blue lines indicate the temporal trend before the breakpoint, and thick red lines indicate the temporal trend following the breakpoint.



**Figure 2.** Ice date variability (standard deviation (SD) and mean first difference (Mean-diff) using a 30-year window and the full ice record for Suwa Lake **(A)** and Torne River (**B**). Visualization of a continuous wavelet transform using a Morlet wavelet for Suwa (**C**) and Torne (**D**) across the full time series. The time-averaged global wavelet transform showing power distribution across frequency with a dotted line showing the 95% level of significance over the entire records for Suwa Lake (**E**) and Torne River (**F**).



**Figure 3.** Regression coefficients between drivers and ice dates for Suwa Lake (A) and Torne River (B). The drivers included: air temperature (AirT), aerosol optic depth (AOD), atmospheric CO2, El Niño Southern Oscillation index (ENSO, Suwa only), North Atlantic Oscillation index (NAO, Torne only), and Sunspots (Torne only). Points indicate estimated coefficients, with blue being the coefficient estimated from data before the breakpoint, and red the period after the breakpoint. Error bars are the 95% CI’s. Pairs of colored dots inlaid with white asterisks indicate significant (α = 0.05; family wise error rate maintained within each system) difference between coefficients prior to and following the breakpoint. For Suwa, positive coefficients suggest that the driver is related to later ice formation. Conversely for Torne, negative coefficients indicate that the driver is related to earlier ice breakup.

**Supplementary Material**

**Data Acquisition**

*Ice freeze and breakup dates*

We obtained ice freeze dates for a 550-year period from 1443-2004 for Suwa Lake and a 321-year record of river ice breakup dates for Torne River from 1692-2013 from the National Snow and Ice Data Center (NSIDC; Benson and Magnuson 2000). Suwa Lake is a relatively shallow lake (Zmax = 7.2 m) in Nagano Prefecture, Japan where ice-freeze dates have been recorded since 1443. Ice freeze dates occur before and after January 1st, therefore we converted dates to Day-Of-Year (DOY) where a zero value represents the calendar day January 1st. If the lake did not freeze, we used the latest observed ice freeze date for that lake (Benson et al. 2012). Data gaps (e.g., 1682-1736; 1872-1896) were excluded from the time series analyses. Ice breakup dates from River Torne have been recorded since 1693 and converted to Julian day including considerations for leap years.

*Large-scale climate drivers and weather*

We obtained data from paleo- and historical records that may be important to ice freeze date on Lake Suwa and ice breakup date on River Torne (Table S1). We acquired average annual sunspot number from 1700-2012 which represents a relative index of solar activity for the visible solar surface from the Solar Influences Data Analysis Center (SIDC-team 2013), volcanic aerosols measured as the annual global average aerosol optical depth inferred at 550 nm from 800-2000 AD derived from sulphate measured in ice cores from Antarctica and Greenland from National Oceanographic and Atmospheric Association (NOAA; Crowley and Unterman 2013), and atmospheric carbon dioxide concentrations from 1AD -2012 AD from contemporary atmospheric concentrations from Mauno Loa, Hawaii and splined CO2 records derived from Antarctic ice cores (Keeling et al. 2001). We also acquired an index of El Nino Southern Oscillation (ENSO) from 1301-2005 derived from the North American Drought Atlas (Li et al. 2011) and winter (December-March) North Atlantic Oscillation index from 1659-2013 (Luterbacher et al. 2002; Hurrell et al. 2013).

We acquired local and regional air temperatures for both locations. For Suwa, we obtained reconstructed growing season temperatures from Hokkaido, Japan derived from tree ring analysis from 1557-2007 (Davi *et al.* 2001) and local temperatures from Tokyo weather station from 1879-2013 (JMA 2013). For Torne, we acquired reconstructed records of monthly temperature adjusted to Haparanda, Sweden from 1802-2002 (nearby town to Torne; Klingbjer and Moberg 2003). These were updated with monthly records from the Haparanda weather station for 2003-2013 from the International Surface Temperature Initiative (Thorne *et al.* 2011). Reconstructed monthly and annual temperatures for central Europe were obtained for the time series between 1500 and 2007 (Dobrovolný *et al.* 2010). Finally, reconstructed growing season temperatures derived from tree ring analysis from central Sweden were acquired from 1107-2007 (Gunnarson *et al.* 2011; Supplementary Table S1).

**Data Analysis**

***A. Is there evidence of a breakpoint?***

*Continuous Segmented Regression*

We used segmented regression to test for abrupt changes in the trend of ice dates in Suwa and Torne. Specifically, we wanted to test when a shift in the temporal trend of ice date may have occurred. To estimate the timing and magnitude of a change in the slope of ice dates, we used continuous segmented regression (CSR) models. In CSR, trend lines on either side of the estimated breakpoint intersect (hence making them “continuous”), but are allowed to have different slopes. In general, a CSR takes the form

 [1]

where yi are observations of ice date,  is a latent variable representing potentially unobserved ice dates ( and yi only differ in Tobit model, described below), x­i are the years of the time series, β0 is the intercept of the regression (ice date on year 0), β1 is the temporal trend in ice date (change in ice date per change in year), the ak are the breakpoints (k was either 1 or 2), the βk+1 are the changes in the temporal trend at each of the k breakpoints, and the εi are the errors. Note that the βk+1 parameters indicate the effect of years elapsed since the breakpoint once the breakpoint has passed on ice date.

*Fitting CSR in Suwa (Tobit)*

The Suwa time series began in 1443 and ended in 2004 (x = 1, 2, … 562), and ice observations were made for 427 of the 562 years (See Fig X in Main Text). The day that Suwa froze ranged from day -54 to day 41 (negative values indicate freezing before January 1st of the designated “year”); however, there were 37 years when the lake did not freeze. Treating no-freeze years as missing data or as a constant date would result in biased results if we employed the regression techniques used for Torne. Thus, calculating trends and breakpoints for Suwa ice dates required a statistical approach distinct from that used in Torne. If the lake is considered as an instrument that measures a value, which we call ice date, that indicates the favorability of conditions for ice formation, and if we understand the lake instrument to censor these measurements at 41, then the no-freeze years can be encoded as ice dates of 41. We consider Suwa as an instrument with output of ice date that is censored at an upper limit, L = 41. As such, the observed yi are related to L and the latent variable in the following manner:

 [2]

To address this censoring of Suwa ice dates while fitting the parameters in Eq. 1, we used a Tobit regression model. For a Tobit regression model with an upper limit (right censoring) of the response variable, the log likelihood of observing data given the parameters β (as in Eq. 1) and σ2 (the variance of ε in Eq. 1), can be calculated as:

 [3]

where φ(.) and Φ(.) are the probability and cumulative density functions of the normal distribution, respectively. The first term is the standard normal likelihood, and applies to observations for which an ice date was observed. The second term reflects the probability of the observation being censored, and applies to no-freeze years. Given parameter values, Eq. 3 reflects the probabilities of observing the ice dates (yi) during freeze years, as well as the probabilities that ice date was censored (unobserved) during no-freeze years. Thus, the β in the Tobit regression model indicate the effect of unit change in X on the latent variable, . We used Tobit regression models as implemented by the vglm() function in the R package VGAM to fit parameters in Eq. 1 to Suwa data.

*Fitting CSR in Torne (OLS)*

The Torne time series began in 1693 and ended in 2013, thus x = 1, 2, … 321. Ice breakup dates for Torne ranged from day 117 to day 160, and the ice melted each year of the time series. For Torne, we fit CSR parameters using ordinary least squares using the lm() function in the statistical programming language R.

*Finding Breakpoint Locations*

When fitting models with one breakpoint, breakpoints were searched exhaustively, and the breakpoint location whose model had the lowest AIC was selected. The same procedure applied to fitting the two-breakpoint model for Torne, which was fit with OLS (Figure S1). Because the Tobit model requires substantially more computational power, and because the Suwa time series is longer (thus more possible breakpoint combinations), breakpoint combinations were not searched exhaustively for Suwa. Instead, we used the genetic optimization algorithm in the rgenoud package (Walter R. Mebane, Jr., Jasjeet S. Sekhon. 2011. Genetic optimization using derivatives: The rgenoud package for R. *Journal of Statistical Software* **42**:11, 1-26). This algorithm permits integer optimization, and is robust to rough likelihood surfaces and local minima.

*Model Selection*

We compared AIC values from CSR models containing one or two breakpoints to multiple regression models containing only year or only year and year2 as predictor variables. We fit models either by OLS (Torne) or by maximum likelihood of the corresponding Tobit regression (Suwa) (Table S1). For Torne and Lake Suwa, the biggest decrease in AIC between models of consecutive complexity was the comparison in AIC for the 2nd order polynomial and the one breakpoint model (ΔAIC for Torne and Suwa was -1.8 and -2.2, respectively; Figure S1 for Torne). Although modest, these changes in AIC suggest that a model with a single breakpoint was a reasonable descriptor of the trend in ice date in both systems.

***B. Has variability changed over the time series?***

For gross climate variability, we analyzed the ice dates using two metrics that are easily interpreted and compared to previously published work. One, the standard deviation of all values applied to the data within the 30-year window. And two, the average of the first difference of the ice dates, which is less prone to changes in the mean and the influence of high- and low-frequency variability (Karl et al. 1995). The ice date values were not normalized across the two lakes, so the variability is presented in units *days* and differs between the two lakes.

***C. Have the same oscillatory dynamics persisted over the time series?***

To examine the oscillatory dynamics in the ice data through time, we applied a continuous wavelet transform to the ice dates for both lakes. This allowed for the decomposition of the signal into its individual frequency components while still examining how they change through time (as opposed to Fourier Transform). The Morlet basis function was chosen for its frequency identifying characteristics. For the continuous wavelet transform (Figure 2C-D), significant periods were identified using a chi-squared test with 95% confidence intervals and assuming a red noise mean background spectrum. For the global wavelet transform (Figure 2 E-F), a 95% significance was also calculated using a red noise background spectrum, but time-averaged across all times outside of the cone of influence to give an overall significance level. For detail on the methods used in the continuous and global wavelet analysis, see Torrence and Compo (1998).

***D. How have drivers changed over time?***

We explored linear relationships between ice date and the following climate drivers (unless otherwise specified, drivers apply to both systems): air temperature (AirT), aerosol optic depth (AOD), atmospheric CO2, El Niño Southern Oscillation index (ENSO, Suwa only), North Atlantic Oscillation index (NAO, Torne only), and Sunspots (Torne only). In each system, the relationships between ice date and each of the climate drivers were explored using data before and after the breakpoint. In Suwa, the duration of the period on both sides of the breakpoint was 101 years (1581–1681 and 1897–1997), and the time periods were chosen to maximize duration on either side of the breakpoint while minimizing the inclusion of years for which ice data were missing. In Torne, we used the full time series on either side, giving 174 years in the first portion (1693 – 1866) and 146 years in the second portion (1867 – 2013). For Torne, air temperature data were not available prior to 1803, thus the early time period for analyses involving Torne AirT was only 64 years (1803 – 1866).

For each ice date – driver pair in each system, we performed separate linear regressions for the period before and the period after the breakpoint (periods as described above). This procedure resulted in twenty-two separate regressions, and each before-after pair of regressions is equivalent to fitting a single model of the form

y = β0 + β1\*x1 + β2\*x2 + β3\*x1\*x2 + ε [4]

where y is ice date, β0 is the intercept, x1 is the driver variable and β1 its effect, x2 is a dummy variable that is 0 before the breakpoint and 1 after, ergo β2 is the post-breakpoint intercept, and β3 is the change in the relationship between the driver and ice date after the breakpoint (i.e., β3 is the adjustment made to β1 after the breakpoint). As described for the breakpoint analysis, for Torne Eq. 4 was fit with ordinary least squares, and with the Tobit regression for Suwa. For each regression, we assumed there was a possibility that the residuals of the regression would be autocorrelated; to estimate coefficient standard errors in the presence of autocorrelation, we used a bootstrapping procedure where the randomized residuals retained the autocorrelation structure of the regression residuals.

To characterize the autocorrelation structure of the residuals in Eq. 4, we fit an autoregressive moving average (ARMA(p,q)) model with *p* AR parameters and *q* MA parameters. An ARMA(p,q) model has the general form

 [5]

The yt are the residuals ε from Eq. 4, and μ is the mean of the residuals, which is 0. The βi are the AR parameters, the αj are the MA parameters, and the εt-j are the residuals. We applied Eq. 5 to the ε of Eq. 4. We selected among ARMA models using AIC, and allowed model complexity to vary from ARMA(1,0) to ARMA(5,5) (including all orders in between). These models were fit and selected using the stepwise procedure implemented in the auto.arima function in the forecast R package (Rob J. Hyndman and Yeasmin Khandakar. 2008. Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software* **27**:3). We then simulated an ARMA process using the fitted ARMA parameters; the variance of the innovations in the simulated ARMA process was the maximum likelihood estimate acquired in fitting Eq. 5 to the residuals.

The ARMA-simulated residuals formed the basis of our bootstrapping procedure. These simulated residuals were then added to the fitted regression values, and the regression was re-fit. This procedure was repeated 1,000 times. The standard deviation of these 1,000 parameter estimates was then used as the standard error of the parameters in Eq. 4. In summary, our bootstrapping procedure was as follows:

1. Fit Equation 4
2. Divide residuals from #1 into before and after period
3. Fit time series model to each set of residuals (Eq. 5)
4. Simulate new sets of residuals from fitted time series model
5. Add simulated residuals from #4 to fitted values () from Eq. 4
6. Re-fit Eq. 4
7. Repeat steps 2-6 1,000 times
8. The standard error of parameters in #1 is the standard deviation of all estimates in #6

We calculated p-values corresponding to the probability that regression coefficients between drivers and ice dates differed before and after the breakpoints in the following manner:

Z = (β2 – β1)/(s.e.2 + s.e.1)-1/2 [6]

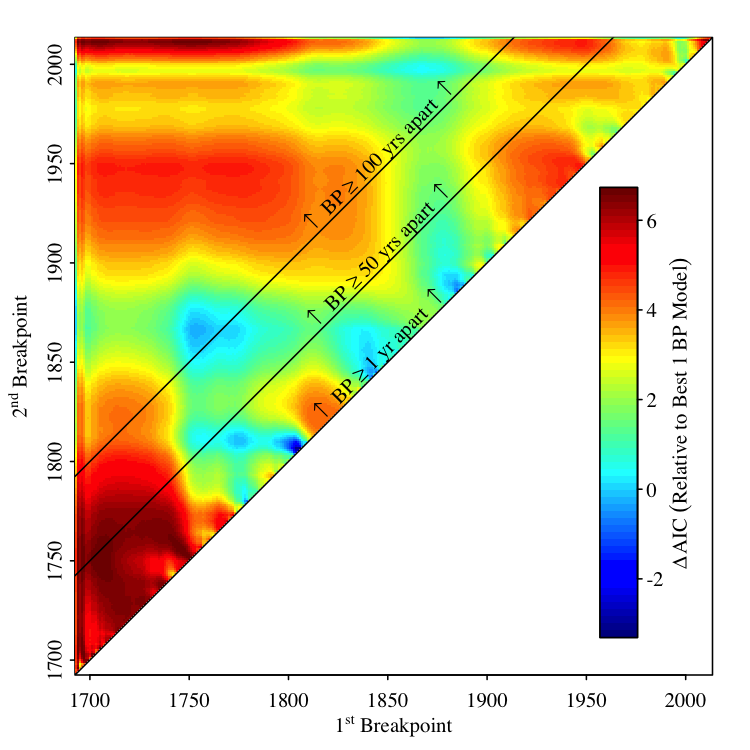
where Z is the z-score to be compared under the standard normal curve, the β are the regression coefficients, and s.e. are the standard errors of those coefficients. We also corrected these p-values to control for multiple tests and to maintain constant family wise error rates. We performed the Holm-Bonferroni correction, and no conclusions about significance among pairs changed.

**Table S1.** AIC values of fitted regression models relating ice date () to years elapsed (xi).

|  |  |  |
| --- | --- | --- |
| Model | Torne AIC | Suwa AIC |
|  | 2155.825 | 3536.38 |
|  | 2154.881 | 3515.407 |
|  | 2153.072 | 3513.241 |
|  | 2152.774\* | 3511.682\*\* |

\*Breakpoints restricted to being at least 10 years apart; See Figure S1.

\*\*Breakpoints restricted to being at least 50 years apart; if restricted to 25 years, AIC = 3510.898.



**Figure S1.** Relative probabilities of two versus one breakpoint in the Torne time series. Colors indicate change in AIC for the two breakpoint model relative to the one breakpoint model (Table S1) for all combinations of first and second breakpoint years in the two breakpoint model. Sloped lines2 indicate boundaries where the first and second breakpoints are separated by the indicated period of time. Note that when at least 25 years separates breakpoints, the one breakpoint model is always more parsimonious than the two breakpoint model (Table S1).