Research Paper

Impact of high flows of an Arctic river on ring widths of floodplain trees

David M Meko,[1] iD Irina P Panyushkina,[1] Leonid I Agafonov[2]
and Julie A Edwards[1]

Abstract
The tree-ring signal for flooding along the Ob River, a large Arctic River in western Siberia, is investigated using a combination of floodplain tree-ring sites from riparian and non-riparian settings. A conceptual model is presented contrasting tree-growth responses of riparian and non-riparian trees to unusually severe flooding. A set of five riparian (*Salix* and *Populus*) tree-ring chronologies is developed and used in combination with existing floodplain non-riparian *Larix* and *Pinus* chronologies in a binary classification tree (CT) model to classify high-flood years, defined as a Salekhard water-level gage reading in the seasonal window from May 1 to August 31 of above 470 cm for 82 or more consecutive days. Correlation and regression identifies a nonlinear relationship of riparian ring widths to discharge and flooding: higher annual discharge generally leads to higher growth, but the relationship reverses in extreme-flood years. Micrographs highlight the suppression of width and occasional distortion of cell anatomy in selected trees. CT modeling guided by cross-validation yields a CT model with a primary split on the riparian ring width and secondary split on the non-riparian ring width. The model successfully identifies four of the eight most severe high-flow years, 1934–2014. The model further identifies two years (1885 and 1914) before the start of the gaged record in 1934 as high-flow years. No appreciable difference is found in frequency of high-flow years before and after 1956, when the first major reservoirs began filling upstream of the Lower Ob. The CT modeling approach is proposed as a novel approach to dealing with nonlinearity in reconstructing flood history of Arctic rivers from tree rings.

Introduction

Riparian tree rings can provide information on river discharge and flood history before the start of gaged discharge records. Although dendrohydrologists have generally favored non-riparian trees for streamflow reconstruction (Meko and Woodhouse, 2011), studies have identified a variety of hydrologic signals in riparian tree-ring width and cell anatomy (Boucher et al., 2011; Copini et al., 2016; Kames et al., 2016; St George and Nielsen, 2002; Yanosky, 1983), and have exploited these signals for reconstruction of time series of flood-related variables (e.g. Kames et al., 2016; Therrell and Bialecki, 2015). Methods of extracting flood records from tree rings are reviewed by Ballesteros-Cánovas et al. (2015). The timing of flooding relative to tree phenology is important, as flooding may leave no trace on ring width if the flood recedes before the start of spring growth (e.g. Denneler et al., 2010).

The Ob River, in western Siberia, is one of the three large Eurasian Arctic rivers that contribute a combined 45% of the freshwater inflow to Arctic Ocean (Shiklomanov et al., 2000). The Ob is attractive as a test site for riparian dendrohydrology because of the presence of undisturbed (by man) stands of riparian trees, previous work in tree-ring chronology development (Agafonov, 1998; Agafonov et al., 2016), and the importance of discharge and flooding to ecology and climatology. The Ob's inflow of heat and freshwater to the Arctic Ocean is critical to the deep circulation of the ocean and global climate system (Aagaard and Carmack, 1989).

In a previous study, the authors investigated the discharge signal in non-riparian tree species *Larix sibirica* and *Pinus sibirica* growing on slightly elevated locations in the floodplain within a few hundred meters of the Ob channel, and developed a reconstruction of December–July average discharge at Salekhard, 1705–2012, from a network of 11 tree-ring chronologies (Agafonov et al., 2016). The conceptual model for that reconstruction was that tree growth in those setting is reduced by cool air temperature at the start of the cambial growth season, and that high discharge and associated broad flooding keep air temperature low in the vicinity of the tree sites through channeling of solar energy into evaporation of water from the broad floodplain rather than into heating of the land surface and air (Agafonov, 1998; Agafonov and Mazepa, 2001).

In this paper, we explore how ring widths of riparian trees along the Lower Ob can be used in combination with the non-riparian tree-ring signal just mentioned to identify extreme Ob River flooding. A few studies have looked specifically at the response of riparian trees along the Ob River to flooding

[1]Laboratory of Tree-Ring Research, University of Arizona, USA
[2]Institute of Plant and Animal Ecology, Ural Branch of the Russian Academy of Sciences, Russia

Corresponding author:
David M Meko, Laboratory of Tree-Ring Research, University of Arizona, 1215 E. Lowell Street, Tucson, AZ 85721-0045, USA.
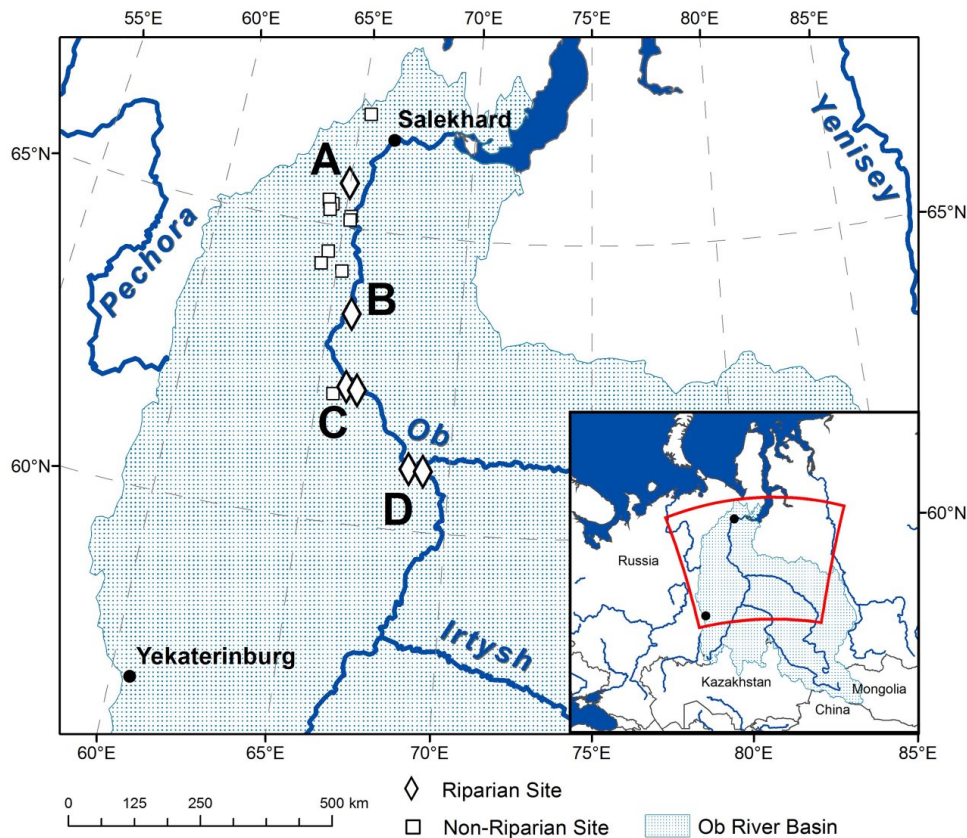Email: dmeko@LTRR.arizona.edu

Figure 1. Ob River basin and sampling points. Symbols mark six locations of riparian tree-ring sites (diamonds) and eleven non-riparian sites (squares) from Agafonov et al. (2016). Riparian trees were sampled near map points A, B, C, and D (see list in Table 1 for details). Stream gage location is at Salekhard.

(e.g. Agafonov, 1995, 1996; Bokk, 1985), but as yet have not used the signal for flood reconstruction. We apply new collections of tree-ring sites to develop five standard chronologies of riparian sites, investigate the statistical signal for discharge and flood duration in the chronologies, and explore the use of a binary classification tree (CT) model to identify high-flood years. The model is then used to infer occurrence of high-flood years prior to the 1934 start of the gaged streamflow record at Salekhard.

Study area

The Ob River Basin, with a drainage area of 2,972,497 km², stretches across the West Siberia Plains of Russia and includes parts of Kazakhstan and China (Figure 1). The Ob flows into the Arctic Ocean through the Gulf of the Ob and Kara Sea. Headwaters are in the Altai-Sayan Mountains of Inner Asia. These mountains contribute some runoff from glacier melt and much more from melting of deep mountain snowpack (Revenga et al., 1998). Additional major components of runoff are snowmelt from the vast area of the West Siberian Plains, and rainfall in the warm season. The basin has a continental climate with a warm-season rainfall peak. Flow of the Ob is largely snowmelt driven, as shown by the mean monthly hydrograph for the gage at Salekhard, near the mouth of the Ob (Agafonov et al., 2016). Monthly discharge typically peaks in June, with seasonal flooding dependent on the various runoff sources and modulated by timing of the spring ice break (Agafonov et al., 2016). Permafrost is not common in the Ob River Basin; less than 2% of the basin is underlain by continuous (90% or more of exposed land) permafrost (Zhang et al., 2005).

All tree-ring chronologies used in this study are located along the Lower Ob, the reach below the confluence of the Irtysh River

with main stem of the Ob (Figure 1). The collections, made in boat expeditions led by L. Agafonov between 1997 and 2017, are described in the next section. Any flood and discharge signal in these tree-ring data will include effects of upstream reservoirs in modifying the flow. Total reservoir storage capacity of the Ob is about 15% of the mean discharge (Yang et al., 2004). Three large (>1 km³ storage) reservoirs are located upstream of the confluence of the Irtysh and the main stem of the Ob (Adam et al., 2007). The exact impact of these reservoirs, two of which began filling in 1956, on floods along the Lower Ob is not known, but studies using hydrologic models (Adam et al., 2007) as well as gaged streamflow (McClelland et al., 2004) suggest appreciable modification of seasonal flows.

Materials and methods

Monthly discharge data for the Ob River at Salekhard (Figure 1) for 1936–2009 were downloaded from the ArcticRIMS database (http://rims.unh.edu/) for the purpose of summarizing correlations of riparian and non-riparian tree rings with seasonally averaged discharge. Time series of discharge averaged over 12 months and over the four warm-season months from May to August underscore the higher flows associated with spring flooding, as well as the high interannual variability of flows (Figure 2).

Daily water levels at Salekhard for 1934–2014 were obtained from the State Hydrological Institute in St Petersburg, Russia, for the purpose of deriving time series related to duration of inundation of riparian trees. Ob River discharge at Salekhard closely tracks Ob water levels throughout the Lower Ob River Basin (Agafonov and Mazepa, 2001). We defined 'high flow' time series as the maximum number of consecutive days each year with water level at Salekhard above some specified gaged level, and developed a graphic to display the high-flow history.
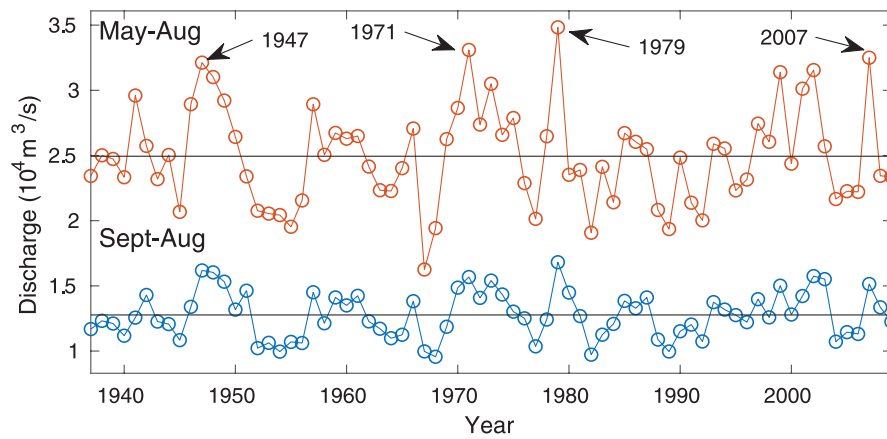
Figure 2. Ob River discharge at Salekhard for two hydrological seasons, 1937–2009. Long-term means (horizontal black lines) are 12,770 m³/s for September–August, and 24,944 104 m³/s for May–August. Series are significantly correlated ($r = 0.88$, $N = 73$, $p < 0.01$). Seasonally averaged discharge is slightly positively autocorrelated (Supplemental Table S1, available online).

Table 1. Statistics of riparian tree-ring width chronologies.

Map point[a]	Site ID[b]	Sample size N_C/N_T[c]	Span[d]	R_{bar}[e]	N_T EPS > .85[f]	EPS year[g]
A	SPS	16/8	1936–2017 (1936)	0.41	9	–
B	POP	–	–	–	–	–
C	OLA	30/24	1901–2014 (1836)	0.57	5	1919
C	SBA	57/35	1880–2014 (1880)	0.51	6	1884
D	OBE	14/11	1834–1997 (1834)	0.35	11	1952
D	SBG	19/10	1939–1997 (1939)	0.65	4	1942

[a]Map points in Figure 1 are A = Tokhotgort (65.747796 N, 64.982858E), B = Kazym (63.557119 N; 65.826463E), C = Oktyabrskoye (OLA–62.316644 N, 66.019693E and SBA–62.280589 N, 66.422280E), and D = Belogorye (61.057127 N, 68.574762E). No chronology was developed at site POP; rings from those trees were used only for micrographs to illustrate ring-width suppression in 1979.
[b]Site ID: SPS is *Salix dasyclados*; SBA and SBG are *Salix alba*; OLA, OBE, and POP are *Populus tremula*.
[c]Number of cores (N_C) and number of trees (N_T) in the developed site chronology.
[d]First and last year of standard chronology, with approximate first year of oldest tree collected at site in parentheses (chronology truncated to exclude data with questionable dating).
[e]Average between-series correlation of detrended ring-width series over uniform common period 1948–1997.
[f]Required number of trees for critical expressed population signal (EPS; Wigley et al., 1984) of 0.85 or higher.
[g]Earliest year sample size large enough for EPS of 0.85; sample for SPS never reaches that threshold.

Principal components (PCs) of 11 standard chronologies of *Larix sibirica* and *Pinus Sibirica* that summarize non-riparian tree-ring variations along the Lower Ob were downloaded from Supplemental Material of Agafonov et al. (2016), who had used PC1 and PC4 to reconstruct the average December–July discharge at Salekhard, 1705–2012.

For this study, we developed five new site chronologies to characterize the riparian tree growth along the Lower Ob. Core samples were collected from riparian trees of willow (*Salix*) and poplar (*Populus*) species at four distinct locations along the Lower Ob River between 1997 and 2014 (A–D in Figure 1). Samples were dated by visually matching patterns of wide and narrow rings, and rings were then measured to an accuracy of 0.01 mm using a sliding stage and microscope (Stokes and Smiley, 1996). Dating and measurements were revised as needed with guidance from the quality-control program COFECHA (Holmes, 1983).

Site tree-ring chronologies were developed at sample points A, C, and D (Figure 1). Cores from a few *Salix* at an additional point (B; Figure 1) were used only to develop micrographs to illustrate ring-width suppression and cell anomalies in extreme-flood years. Micrographs were also produced from selected cores of *Salix* and *Populus* at the other sites. Cores from only one or two trees at each sites were selected for this purpose from wood samples stored at the Institute of Plant and Animal Ecology in Ekaterinburg. Ring-width chronologies were standardized to dimensionless indices with program ARSTAN (Cook et al., 2007). To remove trend

possibly associated with tree size or age, each measured ring width series was fit with a cubic smoothing spline (Cook and Peters, 1981) whose frequency response is 0.5 at a wavelength equal to the length of the series. 'Core indices' were then computed as the ratio of ring width to the value of the fitted curve. The standard site chronology in any year was computed as the robust mean of the available core indices at the site. During chronology development with ARSTAN, variance-stabilization was implemented in two stages to reduce the likelihood of spurious trends in variance of the site chronologies: the first was adjustment of the site chronology to remove trend because of temporal changes in the number of cores averaged to produce the site chronology (Osborn et al., 1997) and the second was removal of any residual trend in the chronology variance by detrending of the absolute departures of the site chronology from its long-term mean with a cubic smoothing spline of the same flexibility as used to detrend the ring-width series (Meko et al., 1993). The adequacy of sample replication was summarized by the expressed population signal (EPS; Wigley et al., 1984). Computed EPS values indicate that only the *Salix* chronology SBA and the *Populus* chronology OLA have sufficient replication for reliable characterization of the population tree-ring signal in years before the start of the daily water-level measurements at Salekhard (Table 1). Replication is strongest for SBA, which begins in 1880 and reaches its critical EPS of 0.85 in 1884.

For subsequent analysis, riparian tree-ring growth was summarized by the average of the two longest standard chronologies,
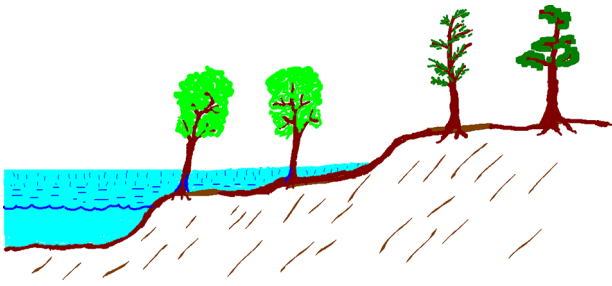
Figure 3. Sketch contrasting exposure of riparian and non-riparian trees to seasonal flooding. The riparian sites (*Salix* and *Populus*) are typically inundated each year, sometimes to a meter or more up the trunk. The non-riparian sites (*Larix* and *Pinus*) are above inundation even in high-flood years.

SBA and OLA, 1884–2014, and by the first principal component (PC1) of all five riparian chronologies, 1939–1997. The principal components analysis (PCA) for that series was run on the correlation matrix of the standard chronologies (Mardia et al., 1979).

Selected samples of rings of *Salix* and *Populus* were examined with thin sections to illustrate cell anatomy in flood years. We cut the samples to a thickness of ~20 μm, stained them with safranin, and then took digital images at 40× magnification.

We evaluated the relationship of riparian ring widths to river discharge and water level in three ways. First was with scatterplots and the Pearson correlation coefficient (Wilks, 1995). Significance of correlation was tested using a Monte Carlo method: for a tree-ring series x and hydrologic series y, the sample correlation, r, of x with y was compared with the distribution of sample correlations of y with 10,000 simulations of x generated by exact simulation (Percival and Constantine, 2006). Exact simulations retain the spectral properties, and so the autocorrelation, of the observed series. A sample r larger than 97.5% of the correlations for the simulations was taken to indicate significance of correlation at $\alpha = 0.05$. The Monte Carlo approach circumvents assumptions (normality and lack of autocorrelation) necessary for theoretical confidence intervals for the Pearson correlation.

Second, stepwise regression (Weisberg, 1985) was used to check whether riparian ring widths explain significant additional variance of average discharge (seasonal) beyond that already explained by the PCs of the 11 non-riparian chronologies used by Agafonov et al. (2016) to reconstruct December–July Ob River discharge. The predictor pool contained the same PCs of non-riparian chronologies used by Agafonov et al. (2016), augmented by the average of the two well-replicated riparian chronologies. Predictors were entered or removed in stepwise regression using a p-to-enter of 0.05 and p-to-remove of 0.10.

Finally, nonlinear aspects of the relationship of the floodplain ring-width indices to water level were tested using a binary CT model (Clark and Pregibon, 1992; Hastie et al., 2013). This type of model has previously been applied in dendroclimatology to classify years of a very wet summer monsoon in the American Southwest (Meko and Baisan, 2001). In our application, the predictand is a binary variable indicating a year of persistent high daily water level, and the predictors are tree-ring variables of non-riparian and riparian ring width. The different exposure of the riparian and non-riparian sites to inundation (Figure 3) is critical to the reasoning for the CT model. The riparian trees are inundated to different levels and durations each year; growth is hypothesized to be favored by flooding, up to some level above which the flooding is detrimental to growth. The non-riparian trees, on the contrary, following Agafonov et al. (2016), are hypothesized to generally grow slower with increased discharge and broader coverage of the

floodplain by water. Those non-riparian tree-ring sites are not inundated by flooding (Agafonov, 1998).

CT modeling was done following Hastie et al. (2013) and was programmed in MATLAB for a binary predictand (high flood or not) and two numeric predictors, representing width of riparian and non-riparian trees. The algorithm for such a setup searches each of the predictors to identify the predictor and the threshold giving the best possible classification of the predictand. 'Best' was judged by minimum deviance (Hastie et al., 2013) over the two resulting nodes. Those nodes are further split to classify their subsets of observation. The process is repeated until no further splits are possible: all nodes are pure or another split would result in too few observations in a terminal node. We did not specify a minimum node size, so that a terminal node could conceivably contain just one observation. The end result is a full 'classification tree', with many splits and terminal nodes. Accuracy of classification by the tree was judged by the weighted classification error averaged over the terminal nodes.

The process just described will typically yield an overly complex CT that overfits the data. To avoid this problem, we applied K-fold cross-validation (Hastie et al., 2013) in an initial run of tree-fitting to identify a reasonable cutoff number of splits, m, beyond which further splitting is likely to result in a decrease of skill when the model is applied to data outside the calibration set. We then re-fit the m-split model using all available observations to arrive at our final CT. This CT was then applied to classify flood years before the start of the gaged daily water-level measurements, back to the start of the available riparian tree-ring predictor.

Results and discussion

High-flood years

The flood history of the Ob River as measured at Salekhard is characterized by occasional years with weeks to months of persistent high water level (Figure 4). Each year the Ob River floods to a different maximum elevation along the floodplain and the water remains high for differing durations. May 1 to August 31 was selected as the window for the color mapping because water-level data are sparse before May, and water-level variations after August are probably unimportant to the growth of the current year's ring in the riparian trees. The highest annual daily gage reading has a range of about 2 m, from 454 cm on May 24, 1954, to 654 cm on May 30, 1987.

The color map in Figure 4 emphasizes the variety of flood elevation and duration in high-flow years. Year 1979 is exceptional for flooding of long duration to a high gage reading; the high yellow bar for 1979 indicates ~100 consecutive days above 550 cm. Year 1941 is characterized by unusually long inundation above 550 cm, but has much shorter inundation than 1979 at lower gage readings. Year 2007 is characterized by long inundation above 500 cm, but no inundation above 550 cm. These varied signatures of floods are likely driven by differences in spatial distribution and timing of precipitation and melt over the basin, modulated by timing of ice break and, likely to some extent, after 1956, by reservoir operation. An extremely long period of inundation is most probable with sequencing of pulsing of melt from the high mountains (Altai Mountains) and the West Siberian Plains, followed by runoff from precipitation in an unusually wet spring and early summer.

For any gage reading, or y-axis level in the color map, a single time series can be extracted of maximum number of consecutive days of inundation. Such a time series for gage reading of 580 cm emphasizes the flood severity in 1979 (Figure 4a). Using a threshold of $N > 20$ consecutive days of water level above 580 cm, the four years 1941, 1971, 1979, and 1999 are defined as 'high-flood' years at 580 cm. The daily progress of water level change usually
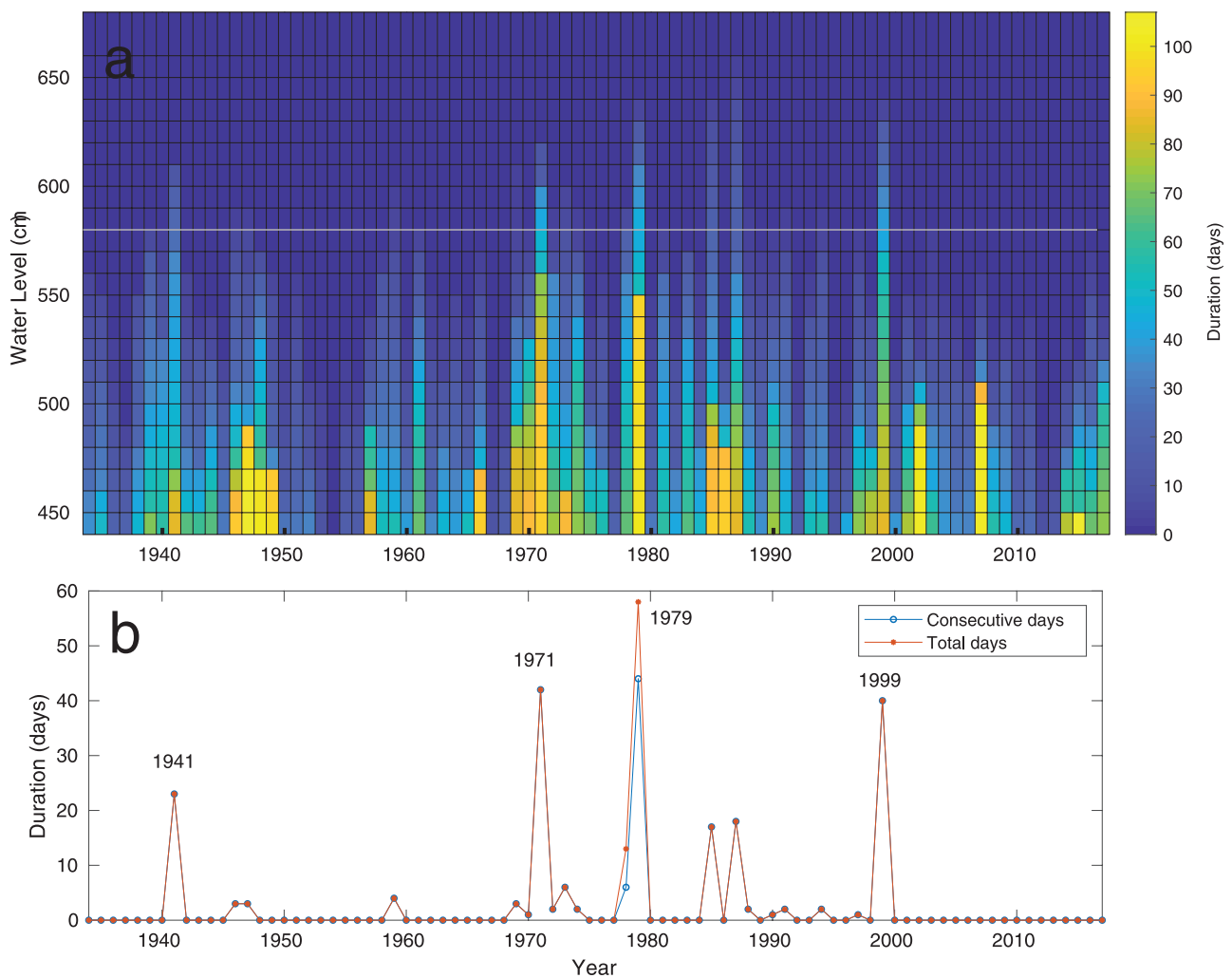
Figure 4. Flood history of Ob River as summarized by duration of high water level: (a) Color map of maximum number of consecutive days each year with water level at Salekhard above a specified gage reading during interval from May 1 to August 31. For example, water level in 1979 persisted above gage reading of 580 cm (white horizontal line) for more than 40 consecutive days, and above 550 cm for 100 consecutive days. (b) Time series representing one slice of color map, specifically for gage reading of 580 cm. Red line gives corresponding count of total days, which usually is the same as number of consecutive days.

leads to one peak, so that the number of consecutive days above x equals the total number of days above x. An exception is 1979, when a secondary pulse of runoff pushed the water back above 580 cm after it had receded from an earlier persistent interval above 580 cm.

Covariation of riparian and non-riparian tree rings

The five standard riparian chronologies listed in Table 1 are weakly intercorrelated over their 1939–1997 common period. All elements of the correlation matrix except two are above $r = 0.18$. Correlation ranges from $r = -0.10$ between OBE and SBG and to $r = 0.54$ between SBA and SBG. That lowest correlation is for a pair of chronologies of different species, including one chronology (OBE) with weak replication. The highest correlation is between two Salix chronologies, each with a strong common signal as measured by the mean between-tree correlation (Table 1). PC1 of the PCA on the five chronologies has positive weights on all sites and explains 44% of the tree-ring variance (Supplemental Table S2, available online). PC2 is a contrast of growth in *Salix* and *Populus*.

The time series variations in riparian PC1, 1939–1997, are tracked closely by the two-site mean of chronologies SBA and OLA, here used to represent riparian tree-ring variation over the much longer period 1884–2014, and referred to from now on as

Rip2 (Figure 5a). The four previously identified high-flow years correspond to lows in Rip2, consistent with long inundation being detrimental to growth.

Non-riparian tree growth is insignificantly correlated ($r = 0.03$, $N = 129$ years) with riparian tree growth over the 1884–2012 common period of Rip2 with PC1 of the non-riparian chronologies. Two of the high-flood years (1971 and 1979), however, correspond to low growth in both series (Figure 5).

Tree-ring signal for river discharge and floods

The riparian and non-riparian trees have distinctly different statistical relationships to annual discharge over a 1937–2009 common period (Figure 6). For the non-riparian trees, both the cloud of points and the correlation highlight a significant negative relationship, consistent with the conceptual model proposed by Agafonov et al. (2016) of tree-ring response to cool air temperature driven by broad flooding. In contrast, the correlation for the riparian tree-ring series is near zero and insignificant. If the four high-flood years 1941, 1971, 1979, and 1999 are omitted, this correlation rises from $r = 0.03$ to $r = 0.21$, which still is not significant at $\alpha = 0.05$ for the given sample size. The main part of the cloud of points in the riparian scatterplot suggests, however, a positive relationship of growth with annual discharge (Figure 6a). This scatterplot therefore suggests a nonlinear relationship, with a
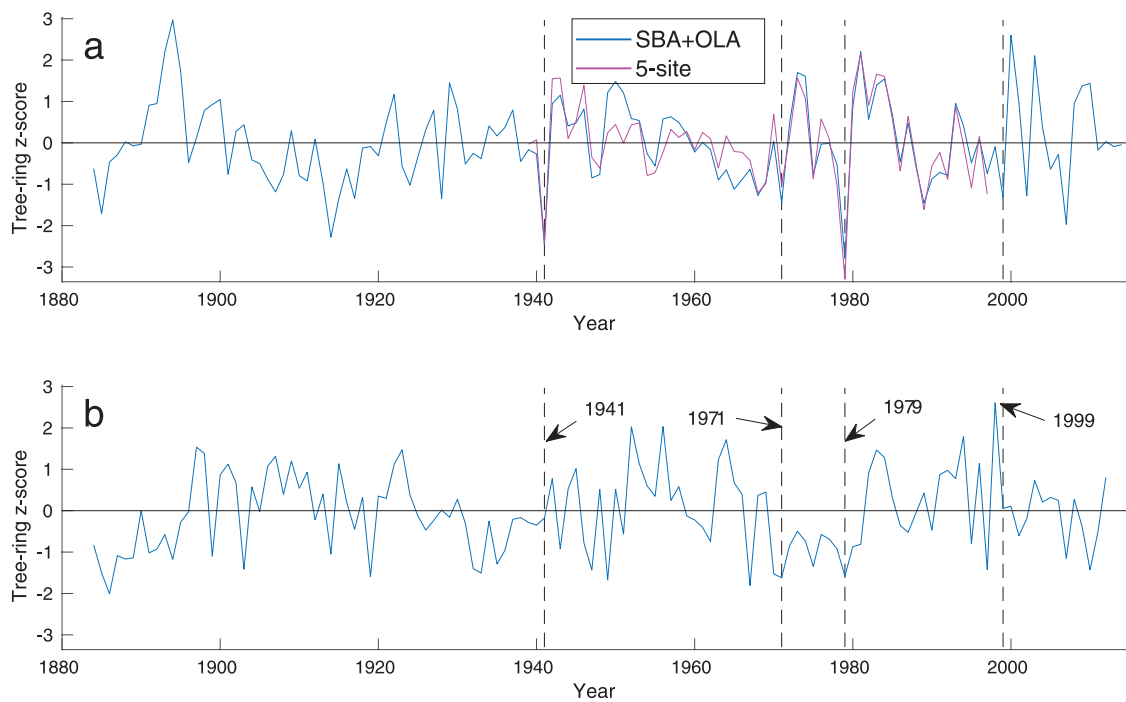
Figure 5. Time series of tree-ring width variation in riparian and non-riparian trees. All series plotted as *z* scores computed using common-period (1939–1997) means and standard deviations. (a) PC1 of the five riparian chronologies listed in Table 1 (pink), and the two-site average of *Salix* chronology SBA and *Populus* chronology OLA (blue). (b) PC1 of 11 *Larix* and *Pinus* chronologies used by Agafonov et al. (2016) for reconstructing annual flows of Ob River at Salekhard. Vertical dashed lines mark the four years with the longest duration (consecutive days) of water level at Salekhard above gage reading of 580 cm (Figure 4).
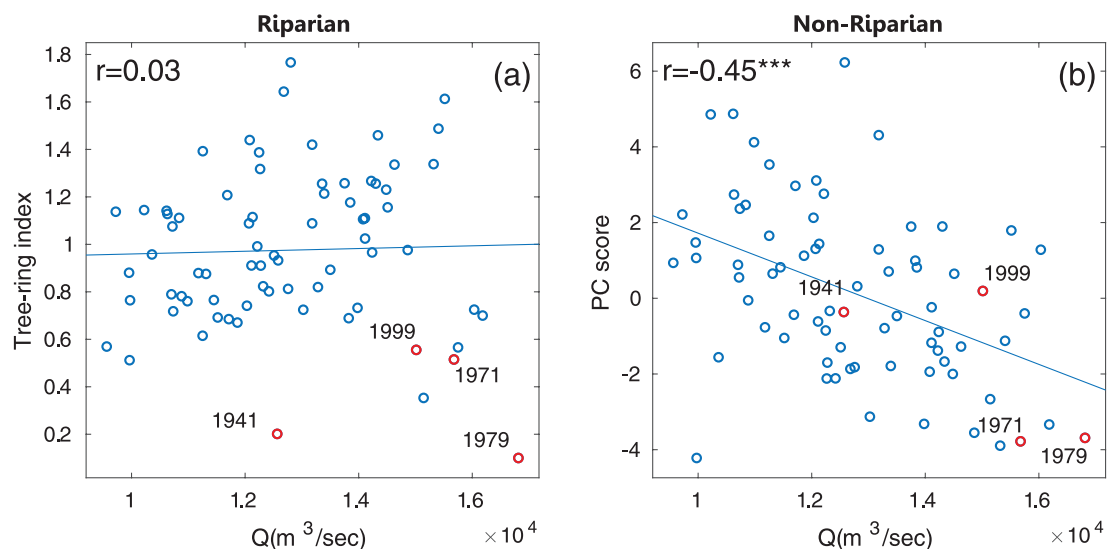

Figure 6. Scatterplots showing differing responses of riparian and non-riparian trees to Ob River annual discharge: (a) Average of riparian standard chronologies SBA and OLA plotted against annual (September–August) discharge of Ob River at Salekhard. (b) PC1 of 11 standard chronologies from *Larix* and *Pinus* plotted against the same discharge series. Scatterplots cover years 1937–2009. Red points, with years annotated, are the four years with the longest duration of water level above gage reading of 580 cm at Salekhard (see Figure 5). Significance of correlation is flagged (***$p < 0.001$).

generally positive growth response to discharge breaking down in high-flow years, defined here as an unusually long period of inundation to high elevations on the floodplain. The background positive relationship can be explained by the strong dependence of riparian trees on abundant moisture. Moderate drying out of the root zone, combined with increased evaporative demand, in hot dry years could induce water stress in the trees in the critical period of cambial growth. Positive growth response to moisture has been used in reconstruction of discharge in northern latitudes of the USA Great Plains (e.g. Meko et al., 2015).

Micrographs illustrate the severe suppression in ring width for the two different species at four widely separate locations along the Ob in 1979 (Figure 7) and for the different species at one particular location in 1999 (Figure 8). Three of the four samples show a micro-ring in 1979. This is likely an effect of long inundation on reducing the access of roots to sufficient oxygen and delaying leaf development (Kozlowski, 2002; Rood et al., 2010). Sample D-OBE9 is different in showing no suppression of growth, but a disruption in the normal shape and density of vessels near the transition of earlywood to latewood. Such impacts of
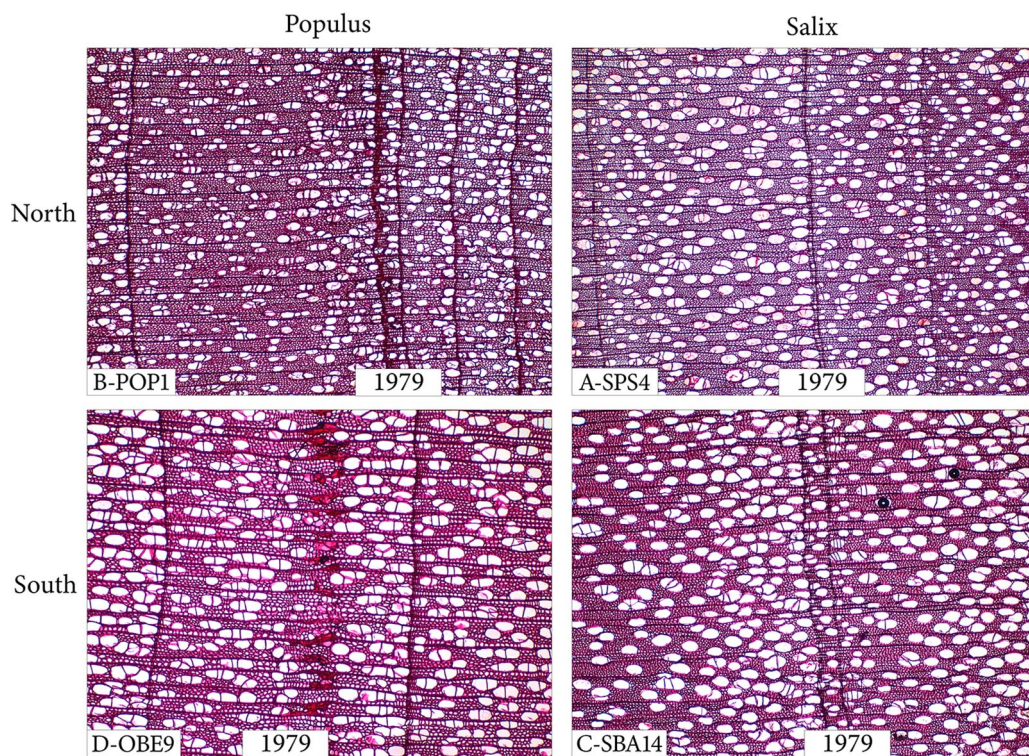
Figure 7. Micrographs illustrating tree-ring anomalies in 1979 in riparian *Salix* and *Populus* at different locations along the Ob River. Label at lower left indicates sampling point (Figure 1) as well as site (Table 1) and tree number. For example, upper right photo is *Salix* from SPS site, tree 4, map point A in the northern part of our sampling transect.
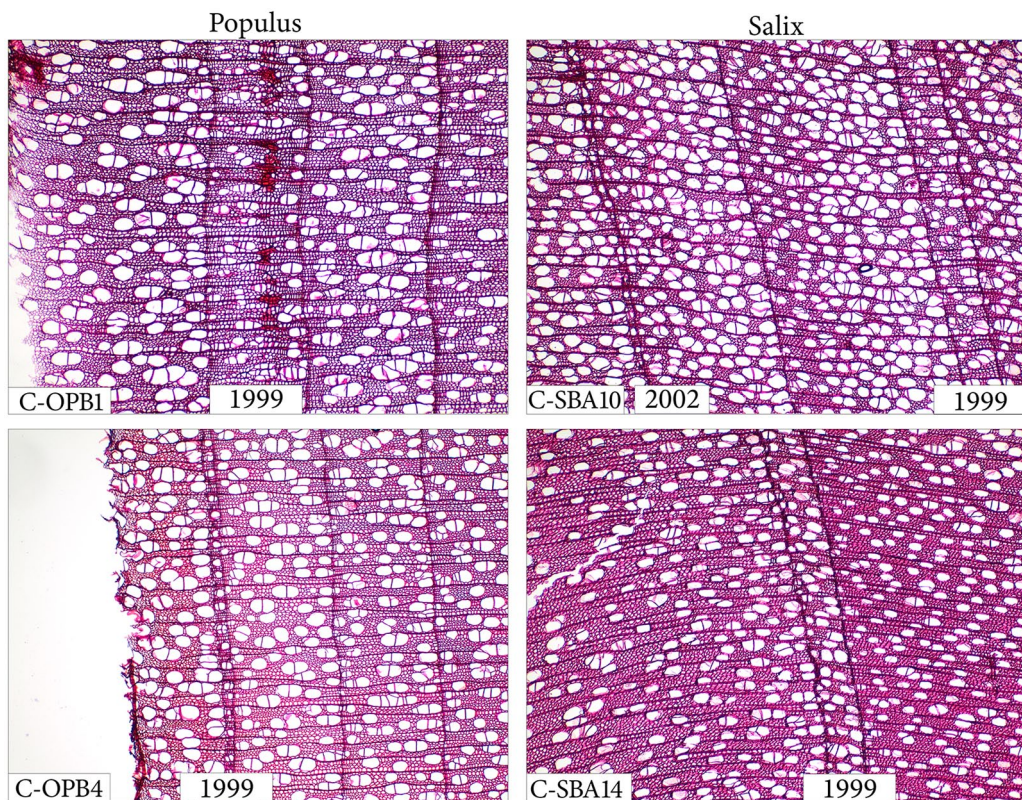

Figure 8. Micrographs illustrating tree-ring anomalies of 1999 in riparian *Salix* and *Populus* at same sampling point (C) along the Ob River.

floods on vessels have been reported elsewhere (e.g. Copini et al., 2016; Kames et al., 2016; Schweingruber, 1996). Studies have also documented mechanisms of some species to cope with inundation and mitigate its effects on growth (Frye and Grosse, 1992; Glenz et al., 2006). The micrographs for 1999 similarly highlight

ring-width suppression and, in one sample, anomalous cell anatomy. In sample C-OPB1, the typical vessel diameter in 1999 appears much smaller than in the adjacent rings. Tree-ring studies elsewhere have noted that inundation tends to produce vessels with reduced diameter (e.g. Kames et al., 2016). The development

of annual time series of average vessel size for correlation with discharge or water level was beyond the scope of our study, but should be a goal in future work.

Tree-to-tree differences in cell anatomy and width suppression in specific flood years, as illustrated in the micrographs, might be expected from sampling variability, as a multitude of climatic and non-climatic factors influence growth. Attention to sample replication can help deal with that problem in developing riparian chronologies for flood reconstruction. Our data set clearly includes some sites requiring greater sample size, as suggested by the EPS statistic (Table 1). A more systematic factor in tree-to-tree differences is likely to be the differing elevations of the sampling locations of individual trees. If possible, future collections should include accurate data on tree elevation, as a meter or so difference in elevation could make an important difference in length and height of inundation. Such resolution is not possible with the handheld GPS (global positioning system) units typically used in tree-ring fieldwork.

Some studies have reported multiyear tree-ring suppression following flood episodes (Yanosky, 1983). Micrographs as well as measured tree-ring widths showed no evidence for a multiyear flood signal in our samples. The riparian tree-ring index plotted in Figure 6a was above normal (1.0) in the years immediately following the four annotated marked flood years. A stepwise linear regression of non-riparian tree-ring PC1 (see Figure 6b) on the annual September–August discharge lagged 0 and −1 year relative to PC1 allowed only lag-0 discharge to enter as a significant predictor.

Could riparian chronologies improve on the accuracy of reconstruction of seasonal or annual average Ob River discharge obtainable in linear regression models based on the much longer non-riparian chronologies? Agafonov et al. (2016) were able to reconstruct 31% of the variance of December–July average discharge at Salekhard from a model calibrated on 1937–2009 and a pool of potential predictors consisting of PCs 1–4 of the 11 non-riparian chronologies; PCs 1 and 4 entered into that stepwise model. We used the identical pool of potential predictors augmented by riparian series Rip2, the same calibration period, and the same stepwise regression settings as Agafonov et al. (2016), and found that Rip2 failed to enter the equation. This result is not surprising considering the breakdown of the weak positive bivariate linear relationship between riparian chronologies and discharge in extreme-flood years (Figure 6a).

Classifying high-flow years from tree rings

The objective of the CT modeling was not reconstruction of annual or seasonal discharge, but classification of extreme events represented as high-flood years – years with likely extended periods of inundation to unusually high elevations in the floodplain. Using the color map in Figure 4 as a guide, we selected the gage reading of 470 cm as the threshold gage height for a high flood and defined the 10% of years, 1934–2014, with the highest number of consecutive days of water level above 470 cm between May 1 and August 31 as 'high-flood' years. This criterion marks eight high-flood years, which were then used as a binary predictand (1 for high-flood, 0 for not high-flood) in the CT modeling. These eight years are 1947, 1969, 1971, 1979, 1985, 1986, 2002, and 2007. Only 1947 precedes the building of the first major dam upstream.

The two predictands for the CT model are the riparian series Rip2 defined previously and the non-riparian series PC1 from the PCA by Agafonov et al. (2016) on the 11 non-riparian chronologies. Accordingly, the CT model attempts to classify high-flow years from summary time series of riparian and non-riparian tree growth. We note that because the non-riparian chronologies end in 2012, we needed to estimate their PC1 time series score for 2013 and 2014 to be able to calibrate the CT model on the full overlap, 1934–2014, of

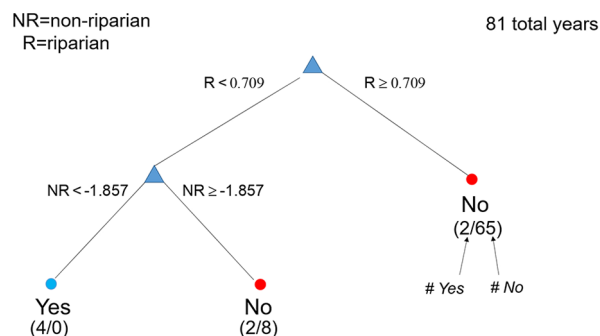Long inundation (Yes) or not (No) at 470 cm gage reading

Figure 9. Model classifying high-flood years during 1934–2014 (81 years) from ring widths of riparian and non-riparian trees. 'High-flood' events defined as the 10% of years with the most consecutive days above gage reading of 470 cm at Salekhard in the time window May 1–August 31. The 10% threshold is 82 or more consecutive days. Riparian predictor is the mean standard chronology of sites SBA and OLA. Non-riparian predictor is PC1 of 11 tree-ring chronologies used by Agafonov et al. (2016) to reconstruct annual Ob River flows. Triangles mark decision nodes and circles mark terminal nodes. Labels 'Yes' and 'No' refer to classification as a high-flood year. Terminal node numbers (*m/n*) in parentheses are number of high-flood years and non-high-flood years. For example, the lower left node, corresponding to very low growth in the riparian trees and low growth in the non-riparian trees, has four observations, and all of those were high-flood years.

Rip2 and the water-level data. For this exercise, we padded PC1 by substituting its long-term mean as values for 2013 and 2014.

We chose three-fold cross-validation to evaluate model performance in selecting an appropriate number of splits in the CTs. With three-fold cross-validation, in this case, the 81 observations are first randomly partitioned into three separate subsets of 27 observations, each subset is successively dropped out, a model is fit to the remaining 54 observations, and performance is evaluated by accuracy of prediction for the 27 omitted observations. Preliminary tree-growing with three-fold cross-validation indicated that beyond two or three splits, the model would be overfit. We therefore ran the final model out to only two splits. The first split in the CT model is on the riparian series, Rip2, and the second split is on the non-riparian series, PC1 (Figure 9). The three terminal nodes classify a year as high-flow or not depending on values of Rip2 and PC1 in a particular year. For example, the combination of a very narrow riparian ring (Rip2 < 0.709) and a very narrow non-riparian ring (PC1 < −1.857) classifies a year as high-flow (left-most node). Note that PC1 is not a tree-ring index, but a PC score; the threshold PC1 = −1.857 is about the 20th percentile (very narrow) of non-riparian growth. Four observations fall into the leftmost node, and all are correctly classified (misclassification error 0/4 = 0). The other two nodes classify a year as not-high-flood, with different misclassification errors. The rightmost terminal node is chosen if the riparian ring is not very narrow (Rip ⩾ 0.709), includes 67 of the total 81 calibration years, and, because two of the observations at the node were actually high-flood years, has a misclassification probability of 2/67, or 0.0299. The middle terminal node also classifies a year as non-high-flood, but with a greater misclassification probability (2/10 = 0.20).

CT modeling allowed to proceed beyond two splits can result in models with high classification accuracy on the full calibration data. Our conceptual model leaves open the possibility that a moderately severe high-flood year could occur with a combination of high riparian growth and low non-riparian growth – further splits down the right side of the two-split model in Figure 9. An example of the model run to four splits is shown in

Table 2. High-flow classification of years in precalibration-period interval 1884–1933 from CT model for terminal nodes plotted in Figure 9.

CT node	High-flood?[a]	p_{mc}[b]	N[c]	Years[d]
Left	Yes	0	2	1885, 1914
Middle	No	0.20	8	1906, 1907, 1911, 1913, 1915, 1917, 1924, 1928
Right	No	0.03	40	All other years, 1884–1933

CT: classification tree.
[a]Whether or not a year at this node is classified as high-flow.
[b]Misclassification probability (calibration).
[c]Number of years at node.
[d]Listing of years at node.

the Supplemental Figure S1, available online. We repeated the cross-validation many times to ensure that rejection of a more complicated model was not just a quirk of the randomization procedure used to assign members of the calibration and validation subsets.

The CT model in Figure 9 was applied to classify high-flow years before the start of the gaged water levels in 1934. Such extended classification is 'reconstruction' of the probability of high-flow events, and was possible back to 1884, the start of the two-site-mean riparian series Rip2. Classification consists of dropping observations down the tree, such that the values of Rip2 and PC1 will determine the path taken and the terminal node. Results show that for the 50 years (1884–1933) before the start of gaged flows, only two years (1885 and 1914) are classified as high-flow (Table 2). The other 48 years fall into the middle or rightmost terminal nodes and are classified as not high-flow; eight of those years have a 20% chance of being high-flow, and the remaining years have only a 3% of being high-flow.

The choice of 470 cm as a gauge-elevation threshold defining 'high flood' for the CT analysis is arbitrary, and other thresholds will generally identify somewhat different sets of years as those with longest duration of high flooding. For example, 1999 is among the four years with longest duration above a threshold of 580 cm (see Figure 4), but is not among the eight years with longest duration above 470 cm used in the CT modeling. Such differences can occur depending on the variable rates of flood recession in different years. Detecting the importance of this factor on tree growth would require accurate elevations for individual trees.

Seven of the eight observed high-flow years in the 1934–2014 gaged record occurred after construction of the first large dam upstream in 1956. The frequency of years classified as high-flow by tree rings is about the same before 1956 as afterward: three years (1885, 1914, 1947), or 4.2% of the 72 years 1884–1955; and three years (1971, 1979, 2007), or 5% of the 59 years 1956–2014. We expected that reservoir operations would tend to increase winter flows but decrease flows in the warm season, including June, the usual month of peak flooding. Yang et al. (2004) noted opposing trends, 1936–1990, in annual discharge in the Low Ob (positive trend) and upstream reaches (negative trend), and concluded that impact of reservoirs in reducing June discharge upstream may have been offset by increasing trend in runoff from snowmelt and rainfall in the Lower Ob Basin.

Conclusion

Ring-width records from riparian and non-riparian trees in the floodplain of the Lower Ob River have complementary signals for discharge and flooding. Previous studies have shown that low growth in non-riparian species *Larix sibirica* and *Pinus sibirica* is related to high discharge and large flooded area through cooling of the air in the critical period of tree-ring development. Here, we show that growth of riparian species *Salix dasyclados, Salix alba*, and *Populus tremula* responds positively or not at all to increased discharge, but is severely reduced in high-flood years, defined by long duration of high water level over May–August. The latter effect is probably because of persistent inundation of roots and lower trunks and consequent oxygen deprivation. The combined information of tree-ring width data from riparian and non-riparian trees has been exploited here with modeling by binary CTs to identify high-flood years. A model developed with three-fold cross-validation effectively identified four of the eight high-flood years over a 1934–2014 calibration period. Applied to earlier tree-ring data, two years, 1885 and 1914, in the 1884–1933 period of earlier riparian tree-ring record are also classified as high-flood years.

The modeling approach used here needs to be tested on other watersheds. Our analysis emphasizes the importance of high replication for capturing the common tree-growth signal at riparian sites, and the importance of cross-validation in development of the CT statistical model. Improvement in flood-year classification accuracy is also needed. The CT model used here successfully classified half of the high-flood years as defined by our particular metric. While some information is better than none, the practical value of a 50% successful classification would depend on the intended use of the results. Future application along the Ob should focus on finding more old living trees, and searching for wood sources (e.g. preserved stumps buried in sediment) to allow extension beyond the relatively short life span of the riparian species. Sampling of living trees should also include accurate elevation estimation at individual trees, as this is critical to the height and duration of inundation. Micrographs of some rings formed during high-flood years also suggest that development of cell-anatomy time series (e.g. number of vessels, average vessel area, size of earlywood vessels) might yield additional flood information from riparian trees in the study basin.

ORCID iD

David M Meko [iD] https://orcid.org/0000-0002-5171-2724

Data Accessibility Statement

Tree-ring site chronologies and hydrologic data used in this paper can be obtained by contacting the corresponding author.

Supplemental material

Supplemental material for this article is available online.

References

Aagaard K and Carmack EC (1989) The role of sea ice and other fresh water in the Arctic circulation. *Journal of Geophysical Research* 94(C10): 14485–14498.

Adam JC, Haddeland I, Su F et al. (2007) Simulation of reservoir influences on annual and seasonal streamflow changes for the Lena, Yenisei, and Ob Rivers. *Geophysical Research Letters* 112: D24114.

Agafonov LI (1995) Effects of hydrologic conditions and temperature on radial increment of deciduous tree species in the Lower Ob floodplain. *Ekologia* 26(6): 436–443.

Agafonov LI (1996) *Effects of hydrologic and climatic factors on the growth of woody plants in the Lower Ob floodplain*. PhD Thesis, Cand. Science (Biology) Dissertation, Yekaterinburg (in Russian).

Agafonov LI (1998) Indication of changes in hydrologic regime of the Lower Ob by means of tree-ring analysis. *Russian Journal of Ecology* 29(5): 311–317 (translated from Russian).

Agafonov LI and Mazepa VS (2001) Runoff of Ob River and summer air temperature in the North of Western Siberia Eurasia. *Proceedings of the Academy of Sciences: Geographic Series (Izvestiya AN. Seriya Geographicheskaya)* 1: 82–90 (in Russian).

Agafonov LI, Meko DM and Panyushkina IP (2016) Reconstruction of Ob River, Russia, discharge from ring widths of floodplain trees. *Journal of Hydrology* 543: 198–207.

Ballesteros-Cánovas JA, Stoffel M, St George S et al. (2015) A review of flood records from tree rings. *Progress in Physical Geography* 39(6): 794–816.

Bokk E (1985) Effect of spring floods on the dynamics of radial increment of white willow from the Ob floodplain. *Lesovedenie* 6: 30–36 (in Russian).

Boucher E, Ouarda TBMJ, Bégin Y et al. (2011) Spring flood reconstruction from continuous and discrete tree-ring series. *Water Resources Research* 47(7): W07516.

Clark LA and Pregibon D (1992) *Tree-based Models, in Statistical Models in S* (ed Chambers JM and Hastie TJ). Pacific Grove, CA: Wadsworth & Brooks/Cole, pp. 377–417.

Cook ER and Peters K (1981) The smoothing spline: A new approach to standardizing forest interior tree-ring width series for dendroclimatic studies. *Tree-Ring Bulletin* 41: 45–53.

Cook ER, Krusic PJ, Holmes RH et al. (2007) *Program ARSTAN*, Version 41d. Available at: www.ldeo.columbia.edu/tree-ring-laboratory

Copini P, den Ouden J, Robert EMR et al. (2016) Flood-ring formation and root development in response to experimental flooding of young *Quercus robur* trees. *Frontiers in Plant Sciences* 7: 775.

Denneler B, Bergeron Y and Bégin Y (2010) Flooding effect on tree-ring formation of riparian eastern white-cedar (*Thuja occidentalis L.*). *Tree Ring Research* 66(1): 3–17.

Frye J and Grosse W (1992) Growth responses to flooding and recovery of deciduous trees. *Zeitschrift fur Naturforschung. C: A Journal of Biosciences* 47(9–10): 683–689.

Glenz C, Schlaepfer R, Iorgulescu I et al. (2006) Review: Flooding tolerance of Central European tree and shrub species. *Forest Ecology and Management* 235: 1–13.

Hastie T, Tigshirani R and Friedman J (2013) *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. 2nd Edition. New York: Springer.

Holmes RL (1983) Computer-assisted quality control in tree-ring dating and measurement. *Tree-Ring Bulletin* 43: 69–78.

Kames S, Tardif JC and Bergeron Y (2016) Continuous earlywood vessels chronologies in floodplain ring-porous species can improve dendrohydrological reconstructions of spring high flows and flood levels. *Journal of Hydrology* 534: 377–389.

Kozlowski T (2002) Physiological-ecological impacts of flooding on riparian forest ecosystems. *Wetlands* 27(3): 550–561.

McClelland JW, Holmes RM, Peterson BJ et al. (2004) Increasing river discharge in the Eurasian Arctic: Consideration of dams, permafrost thaw, and fires as potential agents of change. *Journal of Geophysical Research* 109: D18.

Mardia K, Kent J and Bibby J (1979) *Multivariate Analysis*. London: Academic Press, 518 pp.

Meko DM and Baisan CH (2001) Pilot study of latewood-width of conifers as an indicator of variability of summer rainfall in the North American Monsoon region. *International Journal of Climatology* 21: 697–708.

Meko DM, Friedman JM, Touchan R et al. (2015) Alternative standardization approaches to improving streamflow reconstructions with ring-width indices of riparian trees. *The Holocene* 25(7): 1093–1101.

Meko DM and Woodhouse CA (2011) Application of streamflow reconstruction to water resources management, in Dendroclimatology. In: Hughes MK, Swetnam TW and Diaz HF (eds) *Progress and Prospects, Developments in Paleoenvironmental Research*, vol. 11. Dordrecht: Springer, pp. 231–261.

Meko DM, Cook ER, Stahle DW et al. (1993) Spatial patterns of tree-growth anomalies in the United States and southeastern Canada. *Journal of Climate* 6: 1773–1786.

Osborn TJ, Briffa KR and Jones PD (1997) Adjusting variance for sample-size in tree-ring chronologies and other regional mean time series. *Dendrochronologia* 15: 89–99.

Percival DB and Constantine WLB (2006) Exact simulation of Gaussian time series from nonparametric spectral estimates with application to bootstrapping. *Statistics and Computing* 16: 25–35.

Revenga AC, Murray S and Hammond A (1998) *Watersheds of the World: Ecological Value and Vulnerability*. Washington, DC: World Resources Institute and Worldwatch Institute, 172 pp.

Rood SB, Nielsen JL, Shenton L et al. (2010) Effects of flooding on leaf development, transpiration, and photosynthesis in narrow leaf cottonwood, a willow-like poplar. *Photosynthesis Research* 104(1): 31–39.

Schweingruber FH (1996) *Tree Rings and Environment*. Bern: Paul Haupt Publishers, 669 pp.

Shiklomanov IA, Shiklomanov AI, Lammers RB et al. (2000) The dynamics of river water inflow to the Arctic Ocean. In: Lewis EL, Jones EP, Lemke P et al. (eds) *The Freshwater Budget of the Arctic Ocean*. Dordrecht: Kluwer Academic Publishers, pp. 281–296.

St George S and Nielsen E (2002) Hydroclimatic change in southern Manitoba since A.D. 1409 inferred from tree rings. *Quaternary Research* 58: 103–111.

Stokes MA and Smiley TL (1996) *An Introduction to Tree-Ring Dating* (originally published 1968, University of Chicago Press). Tucson: University of Arizona Press, 73 pp.

Therrell MD and Bialecki MB (2015) A multi-century tree-ring record of spring flooding on the Mississippi River. *Journal of Hydrology* 529: 490–498.

Weisberg S (1985) *Applied Linear Regression*. 2nd Edition. New York: John Wiley, 324 pp.

Wigley TML, Briffa KR and Jones PD (1984) On the average value of correlated time series, with applications in dendroclimatology and hydrometeorology. *Journal of Applied Meteorology and Climatology* 23: 201–213.

Wilks DS (1995) *Statistical Methods in the Atmospheric Sciences* (International Geophysics Series, vol. 59). San Diego, CA: Academic Press, 467 pp.

Yang D, Ye B and Shiklomanov A (2004) Discharge characteristics and changes over the Ob River Watershed in Siberia. *Journal of Hydrometeorology* 5: 595–610.

Yanosky TM (1983) *Evidence of floods on the potomac river from anatomical abnormalities in the wood of flood-plain trees*. Technical report, U.S. Geological Survey Professional Paper 1296. Washington, DC: U.S. Geological Survey, 42 pp.

Zhang T, Frauenfeld OW, Serreze MC et al. (2005) Spatial and temporal variability in active layer thickness over the Russian Arctic drainage basin. *Journal of Geophysical Research* 110(4): D16.