

## Spring flood reconstruction from continuous and discrete tree ring series

Étienne Boucher,<sup>1</sup> Taha B. M. J. Ouarda,<sup>1</sup> Yves Bégin,<sup>1</sup> and Antoine Nicault<sup>1</sup>

Received 16 October 2010; revised 8 April 2011; accepted 19 April 2011; published 9 July 2011.

[1] This study proposes a method to reconstruct past spring flood discharge from continuous and discrete tree ring chronologies, since both have their respective strengths and weaknesses in northern environments. Ring width or density series provide uninterrupted records that are indirectly linked to regional discharge through a concomitant effect of climate on tree growth and streamflow. Conversely, discrete event chronologies constitute conspicuous records of past high water levels since they are constructed from trees that are directly damaged by the flood. However, the uncertainty of discrete series increases toward the past, and their relationships with spring discharge are often nonlinear. To take advantage of these two sources of information, we introduce a new transfer model technique on the basis of generalized additive model (GAM) theory. The incorporation of discrete predictors and the evaluation of the robustness of the nonlinear relationships are assessed using a jackknife procedure. We exemplify our approach in a reconstruction of May water supplies to the Caniapiscou hydroelectric reservoir in northern Quebec, Canada. We used earlywood density measurements as continuous variables and ice-scar dates around Lake Montausier in the James Bay area as a discrete variable. Strong calibration ( $0.57 < 0.61 < 0.75$ ) and validation ( $0.27 < 0.44 < 0.58$ )  $R^2$  statistics were obtained, thus highlighting the usefulness of the model. Our reconstruction suggests that, since  $\sim 1965$ , spring floods have become more intense and variable in comparison with the last 150 years. We argue that a similar procedure can be used in each case where discrete and continuous tree ring proxies are used together to reconstruct past spring floods.

**Citation:** Boucher, É., T. B. M. J. Ouarda, Y. Bégin, and A. Nicault (2011), Spring flood reconstruction from continuous and discrete tree ring series, *Water Resour. Res.*, 47, W07516, doi:10.1029/2010WR010131.

### 1. Introduction

[2] In northern regions, spring flood is a dominant component of the hydrological cycle. During that period, lake water levels and river discharge reach peaks that are rarely attained during the rest of the year [Church, 1988]. Spring flood waters may severely disturb riparian ecosystems [Hupp, 1988; Sigafos, 1964], alter channel and riverbank geomorphology [Ettema, 2002; Smith, 1979], and threaten the communities living alongside. However, in areas like northern Quebec (Canada) where major hydrogeneration systems have been constructed, spring floods are of economic importance: The annual infilling of the large reservoirs that feed the power plants is largely dependent on spring water supplies.

[3] An efficient water resources planning program should be based on an understanding of the decadal to centennial trends and patterns of hydrological variability [Intergovernmental Panel on Climate Change, 2007; Kundzewicz et al., 2007]. However, in many remote areas such as northern Quebec, long-term hydrological series that encompass a

large spectrum of the natural variability are nonexistent; the longest time series cover at most the last three to four decades. In such cases, natural proxies can be used to lengthen the existing records. Among the various proxies used to retrieve paleohydrological information, tree ring series offer an important advantage. They provide the analyst with annually resolved records that can be directly used as predictors for streamflow when climate has a concomitant (yet indirect) effect on both variables [Loaiciga et al., 1993]. On that basis, ring-width chronologies have been highly successful at reconstructing past hydrological conditions in areas of the world, such as semiarid parts of the United States [Hidalgo et al., 2000; Meko et al., 1995, 2001; Smith and Stockton, 1981; Stockton and Fritts, 1971; Stockton and Jacoby, 1976; Woodhouse and Lukas, 2006], where water deficit is the main factor limiting tree growth.

[4] The calibration of such an indirect statistical relationship between ring widths and streamflow is risky in boreal environments. First, water is not a limiting factor in these areas as it remains available for trees throughout the growing season. Second, and of utmost importance, for the reconstruction of spring flood events, ring widths integrate all of the environmental components influencing tree growth, i.e., hydroclimatic conditions, competition, disturbances, etc. [Cook and Kairiukstis, 1990]. Because spring

<sup>1</sup>Centre Eau-Terre-Environnement and Centre d'Études Nordiques, Institut National de la Recherche Scientifique, Québec, Canada.

floods usually occur on much shorter timescales, they often leave no visible traces in ring-width chronologies.

[5] To circumvent the latter situation, ice-scar chronologies are often used to document past spring floods in cold regions. These discontinuous records are constructed from exposed riparian trees that are frequently damaged by the direct abrasion of water and/or ice during extreme floods [Alestalo, 1973; Bégin, 2000a; Hupp, 1988; Schweingruber, 1996; Sigafos, 1964; Stoffel et al., 2010]. These marks have been proven to be useful paleoflood indicators in both lacustrine [Bégin, 1999, 2000a, 2000b; Bégin and Payette, 1988; Lemay and Bégin, 2008; Lepage and Bégin, 1996; Tardif and Bergeron, 1997] and fluvial environments [Boucher et al., 2009a; Gottesfeld and Johnson Gottesfeld, 1990; Hensch, 1973; McCord, 1996; Payette, 1980; Smith and Reynolds, 1983; Bégin, 2000]. However, a major difficulty with such series is that no precise statistical framework exists to quantitatively reconstruct past discharge using ice scars. An important caveat relates to the uncertainty of these records. For example, the magnitude of a paleoflood is often estimated using the proportion of living trees bearing a scar at year  $t$ . But how much confidence can one put in these estimated proportions? Moreover, how rapidly does that confidence degrade as the number of survivor trees diminishes back in time? This uncertainty must be quantified if discrete event series are to be used as proxies to reconstruct past spring flood events.

[6] The objective of the present paper is to propose an alternative approach for spring flood reconstructions in environments where trees are not clearly stressed by the lack of water. This procedure makes use of both continuous and discrete series simultaneously since each have their respective strengths and weaknesses in these environments. Continuous series are uninterrupted through time (each year, a ring is produced), and thus they can be easily coupled to annual streamflow observations. We use earlywood density measurements (wood formed during the late spring/early summer period) as a continuous paleoflood indicator. Nevertheless, the hydrological information obtained from these series remains indirect, as trees sampled to construct earlywood density chronologies are not in contact with a naturally fluctuating water body. On the other hand, discrete event series are constructed from trees that are directly affected by floods. However, these records are discontinuous and their uncertainty remains unquantified.

[7] The procedure described herein attempts to include the uncertainty of discrete tree ring series in the reconstruction process through the modeling of binomial confidence intervals. Generalized additive models (GAM) [Hastie and Tibshirani, 1990], which represent a generalization of the multiple linear regression (MLR) approach, are used as transfer functions. Despite the fact that GAMs have never been used in tree ring science, they present numerous advantages for the reconstruction of paleodischarge data because (1) they allow for the modeling of nonnormal response variables (discharge data usually follow a gamma distribution [Bobee and Ashkar, 1991]), and (2) they allow for optimal modeling of the nonlinear relationships that are frequently encountered in tree ring research [Graumich and Brubaker, 1986; Woodhouse, 1999]. We first detail the method and then we apply it in a case study in northern Quebec, with the objective of reconstructing spring water supplies to the

Caniapiscau reservoir, the head of one of the world's largest hydroelectric systems (the La Grande Complex).

## 2. Statistical Approach

### 2.1. Generalized Additive (Transfer) Model

[8] Over the last few years, most hydrological reconstructions from continuous tree ring series have been performed using multiple linear regressions (MLR) [Case and MacDonald, 2003; Gou et al., 2007; Hidalgo, et al., 2000; Meko et al., 1995, 2001; Meko and Graybill, 1995; Woodhouse, 2001; Woodhouse et al., 2006]. Let  $I_t$  be the proportion of trees bearing an ice scar at year  $t$  around some random lake (more details will be given in section 2.2.) and  $W_t$  a measure of ring width or ring density. Assuming, for instance, that both  $I_t$  and  $W_t$  are continuous, one could therefore formulate a MLR model

$$E(Q_t) = \beta_0 + \beta_1 I_t + \beta_2 W_t + \varepsilon, \quad (1)$$

where  $E(Q_t)$  denotes the expected discharge value,  $\beta$  are least square or maximum likelihood estimates of the parameters, and  $\varepsilon$  is a white-noise term. Such a linear model is appropriate in situations where (1)  $Q_t$  is Gaussian and (2) the relation between  $Q_t$  and tree ring series can be best approximated by a straight line. This might not be a suitable assumption if one is interested in reconstructing past hydrological conditions from ice-scoured trees found along water bodies. First, discharge data are usually best modeled as a gamma rather than a Gaussian distribution [Bobee and Ashkar, 1991] because the former remains strictly positive over its full domain. Second, ice-scouring indicators can be nonlinearly related to regional hydrological conditions. This nonlinearity is not only of statistical interest; it characterizes the natural interactions that occur in these dynamic systems. The most plausible causes for the disruption of the linear relationships are lagged hydrological responses between basins of various sizes, complex hydro-meteorological interactions that govern both the flooding and ice cover breakup processes, altitudinal variations in tree densities along shores, and irregular topographies of the shoreline. All of these situations discourage the systematic use of MLRs as a modeling tool for paleohydrological reconstructions when ice scars are used.

[9] Generalized additive models constitute a generalization of the MLR approach. They were first introduced by Hastie and Tibshirani [1990] and have since become quite popular in environmental science modeling. The general idea behind these models is to replace the traditional least-squares-estimated  $\beta$  parameter of MLRs by some smoothing function [Hastie and Tibshirani, 1990]. Thus, reformulating (1) into a GAM would give

$$g[E(Q_t)] = \beta_0 + f_1 I_t + f_2 W_t + \varepsilon, \quad (2)$$

where  $f_1$  and  $f_2$  are smooth terms and  $g$  is a link function.

[10] A major advantage of GAMs over MLRs is that the model structure can now handle, through  $g$ , responses of any exponential family distribution (e.g., binomial, Poisson, gamma, etc.). This is a great advantage in paleohydrology. The link function enables the modeling of gamma-

distributed discharge values and makes it possible to relax the hypothesis of normality that is inherent to MLR models. Within the GAM framework, it is also possible to model nonlinear relationships through  $f_1$  and  $f_2$ . Such scatterplot smoothers do not assume a rigid form for the dependency between the response and the covariates, and therefore they add a great deal of flexibility to the model, a crucial advantage in our context. Many different scatterplot smoothers can be used in equation (2) [Hastie and Tibshirani, 1990], e.g., loess and kernel smoothers, thin plate regression splines, cubic regression splines, penalized regression splines [Wood and Augustin, 2002], P-splines [Eilers and Marx, 1996], and adaptive and tensor product smoothers.

[11] When using GAMs, one seeks to find the best trade-off between the goodness of fit and the adjustment of a smoothing function (the so-called “wiggleness”) [Hastie and Tibshirani, 1990; Wood, 2000; Wood and Augustin, 2002]. For example, in order to find the best fit between  $I_t$ ,  $W_t$ , and  $Q_t$  (of size  $m$ ) using penalized regression splines, the GAM procedure attempts to minimize

$$\sum_{i=1}^m [Q_{t(i)} - \beta_0 - f_1(I_{t(i)}) - f_2(W_{t(i)})]^2 + \lambda_1 \int \left( \frac{\partial^2 f_1}{\partial I_{t(i)}^2} \right)^2 dI_t + \lambda_2 \int \left( \frac{\partial^2 f_2}{\partial W_{t(i)}^2} \right)^2 dW_t, \quad (3)$$

where  $\lambda_1$  and  $\lambda_2$  are smoothing parameters that control the degree of smoothness of the model (and hence its degree of freedom), and  $\lambda_1 \int (\partial^2 f_1 / \partial I_{t(i)}^2)^2 dI_t$  and  $\lambda_2 \int (\partial^2 f_2 / \partial W_{t(i)}^2)^2 dW_t$  correspond to the penalties that measure the degree of adjustment [Wood and Augustin, 2002]. A great deal of literature has focused on how to conveniently approximate  $\lambda$ , but an increasingly recognized strategy [Wood and Augustin, 2002] consists of combining a penalized iteratively reweighted least squares (P-IRLS) method with a procedure that attempts to find the value of  $\lambda$  that minimizes a general cross-validation (GCV) score.

## 2.2. Modeling Uncertainties Around $I_t$

[12] Up to this point, we assumed that  $I_t$  was certain and continuous to better describe the GAM. We will now take a look at how to include uncertain and discrete  $I_t$  values in the GAM.

[13] The proportions of ice-scarred trees ( $I_t$ ) around a lake, for each year  $t$ , are calculated using the Shroder [1980] method. Shroder [1980] proposed to count the number of trees bearing a scar at year  $t$  ( $R_t$ ) and to divide this quantity by the number of trees available ( $A_t$ ) for recording (i.e., the number of trees with a ring dating from year  $t$ ).  $I_t$  therefore provides the dendrochronologist with yearly estimates of past ice-flood magnitudes, since, in the simplest case, it can be expected that a higher proportion of trees will be scarred if water levels are higher at year  $t$ . However, evaluating the uncertainty of these estimates is important for two reasons. The first reason relates to the calibration of the transfer model. Large uncertainties would signify that, if more trees were collected within the same population, the value of these proportions could change dramatically. Therefore, coupling uncertain proportions to hydrological variables in a transfer model could lead to the perception of

trends that might not exist if the sampling effort had been more intense. The second reason relates to the reconstruction process. Although the data can be calibrated for the recent past (i.e., where most trees are alive and available for recording), performing a reconstruction inevitably forces the dendrochronologist to include precalibration  $I_t$  values that are less accurately estimated, as very few “old” trees (e.g., living trees that installed on the lakeshore 200 or 300 years ago) will have survived to the present.

[14] To model this source of uncertainty and include it in the reconstruction process, we propose to construct binomial confidence intervals around the proportion of ice-scarred trees. It must be noted that ice-scar chronologies often yield extreme proportions (e.g., proportions obtained from a very low  $R_t$  or  $A_t$ ), so the choice of method should account for that particularity. Many methods exist to calculate binomial confidence intervals; these methods have been compared for their coverage probabilities under a wide range of proportions and sample sizes [Agesti and Coull, 1998]. From these comparisons, it has been shown that the Wilson score method guarantees coverage probabilities that remain close to the nominal confidence levels for all combinations of sample sizes and proportions, even for extreme proportions (close to 1 or 0) or very small sample sizes (close to 1). The Wilson score method thus appears to be a logical choice over other methods (the Wald method, for example) that yield more conservative coverage probabilities [Agesti and Coull, 1998].

[15] In order to construct binomial confidence intervals around  $I_t$  values, we first consider the detection of a scar at year  $t$  on a given tree as a Bernoulli trial with the outcomes of 0 = “absence” and 1 = “presence” of a scar. Since there are  $A_t$  available trees, we consider  $A_t$  Bernoulli trials. Then, we let  $I_t = R_t/A_t$  (after Shroder [1980]) and hypothesize that the most probable values around  $I_t$  follow a binomial distribution of parameter  $A_t$  ( $A_t \in \mathbb{N}_0$ ) and  $I_t$  ( $I_t \in [0, 1]$ ) such that

$$\Pr(I_t = R_t/A_t) = \binom{A_t}{I_t} I_t^{R_t} (1 - I_t)^{A_t - R_t}, \quad (4)$$

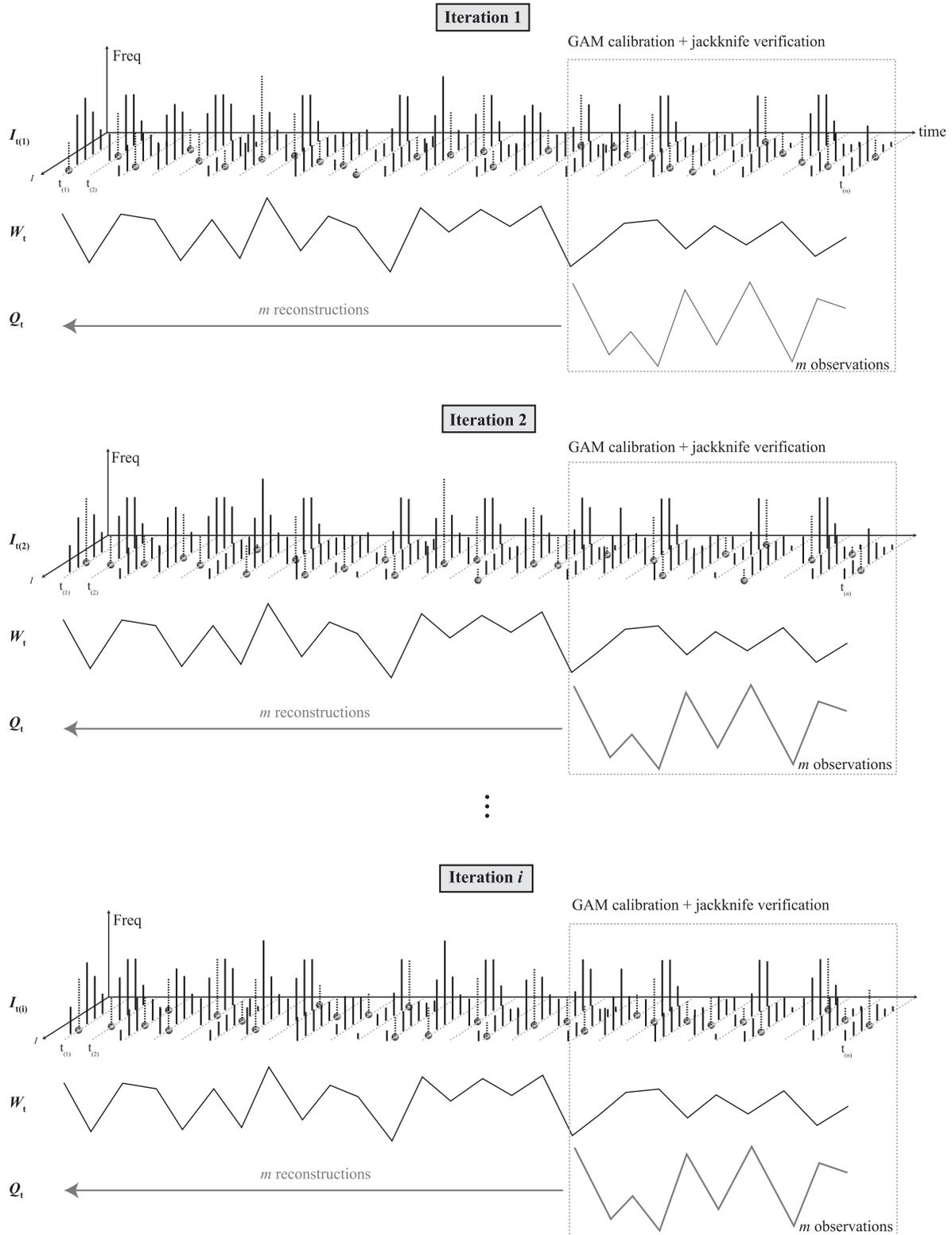
where  $R_t \in \mathbb{N}_{[0, A_t]}$ . We then define the confidence intervals after Wilson [1927]

$$\left\{ I_t + \frac{z_{\alpha/2}^2}{2A_t} \pm z_{\alpha/2} \sqrt{[I_t(1 - I_t) + z_{\alpha/2}^2/4A_t]/A_t} \right\} / \left( 1 + z_{\alpha/2}^2/A_t \right), \quad (5)$$

where  $z_{\alpha/2}^2$  corresponds to the quantile of the standard normal distribution.

[16] The integration of observed  $I_t$  values along with their confidence intervals into the GAM transfer function is achieved through a seven-step algorithm. The algorithm also includes a jackknife validation procedure [Miller, 1964; Shao and Tu, 1995]. The algorithm is detailed here and illustrated in Figure 1.

1. For each year  $t$ , randomly pick (with replacement) any value within the bounds of the Wilson confidence interval calculated using equation (5) and produce a synthetic series  $I_{t(1)}$  equal in length to the original  $I_t$  series.



**Figure 1.** Illustration of the statistical approach to calibrate and validate the transfer model. It is considered as an uncertain predictor. For each year  $t$ , there are many possible  $I_t$  values. Bars depict the range of possible  $I_t$  values, as they are expected from a binomial distribution of parameters ( $A_t$  and  $I_t$ ). Bar heights refer to the frequency of occurrence of each case when  $i$  values are drawn from the latter distribution. At each iteration  $i$ , a random  $I_t$  value is chosen (gray bars among the possible values (black bars)) and a synthetic  $I_t$  series is computed. The synthetic discrete series is coupled to  $W_t$  to predict  $Q_t$  and the model is validated using a jackknife procedure. The procedure is repeated  $i$  times and at each iteration, a different synthetic  $I_t$  series is recalculated.

2. Given  $I_{t(1)}$ ,  $W_t$ , and  $Q_t$ , build a training data set of size  $m - 1$  for calibration by leaving one observation out for model validation.

3. Fit a GAM using equation (2) with that training data set.

4. Make a prediction for  $Q_t$  by substituting  $W_t$  and  $I_{t(1)}$  values with the corresponding left-out cases (see step 5).

5. Reconstruct  $Q_t$  using the same GAM by substituting  $W_t$  and  $I_{t(1)}$  values with those of the preinstrumental period.

6. Repeat steps 2 through 5  $m$  times so that each of the  $m$  cases of the calibration data set are used for validation.

7. Repeat steps 1 through 6  $i$  times using a different synthetic  $I_{t(1)}$  series at each iteration.  $i$  should be as large as possible (i.e., at least 100).

[17] The latter procedure makes it possible to include discrete chronologies as predictor variables in a transfer model to reconstruct spring floods. The discrete variable is iteratively transformed into a continuous one with yearly values randomly picked within the bounds of the confidence intervals. The jackknife procedure enables the validation of the reconstruction model and the construction of confidence intervals around reconstructed values.

### 3. Reconstruction of Spring Water Supplies to the Caniapiscou Reservoir, Northern Quebec

[18] The statistical framework described in section 2 is now applied to a case study in northern Quebec (Canada). We use both continuous and discontinuous series to reconstruct spring water supplies to the Caniapiscou hydroelectric reservoir (Figure 2). Discontinuous series correspond to the proportion of trees bearing a scar at year  $t$  around Lake

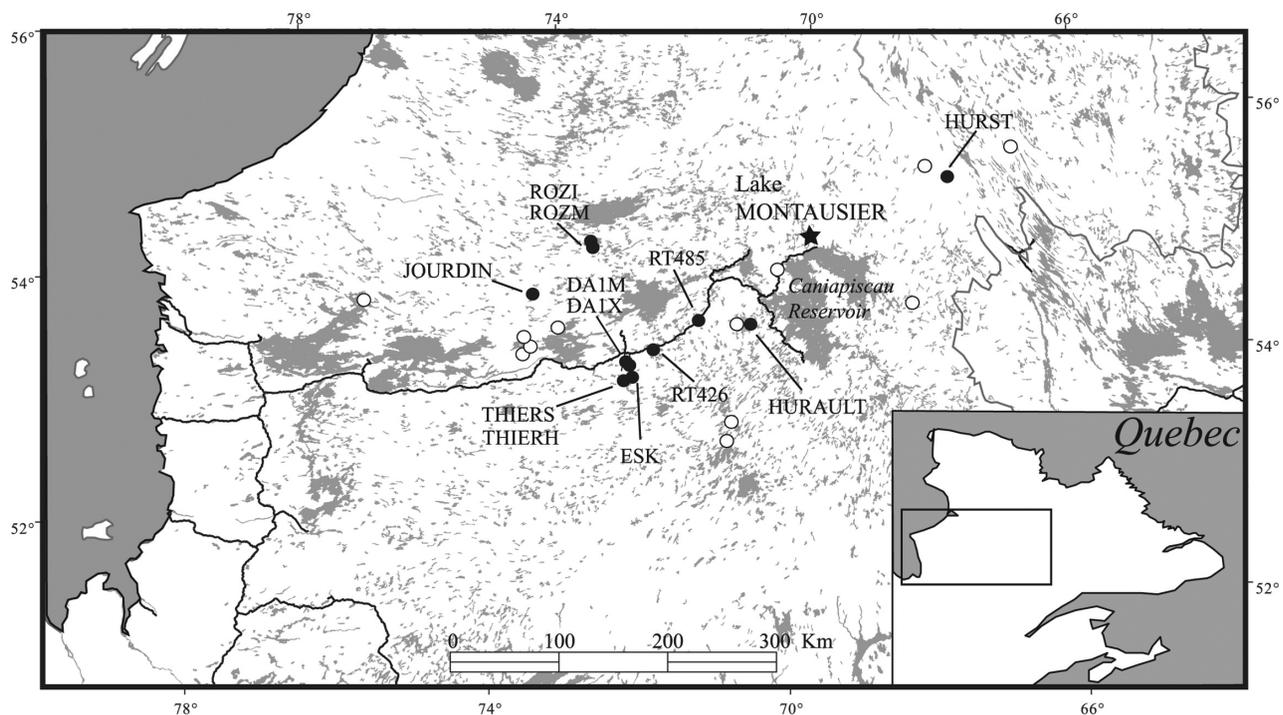
Montausier in northern Quebec. Continuous series are earlywood density series obtained from trees living in the nearby area.

#### 3.1. Water Supplies to the Caniapiscou Reservoir

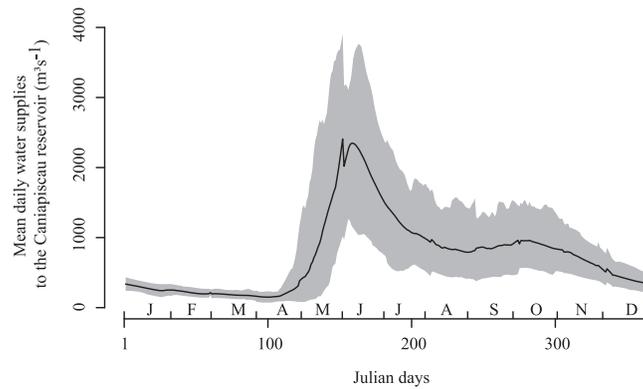
[19] Harmonized water supplies are used by Hydro-Quebec for hydrological predictions and water resources planning [Hydro-Quebec, 2006]. Averaging the daily values for the 1950–2007 period produces a hydrograph of the mean daily water supplies to the Caniapiscou reservoir over the last 58 years (Figure 3). Water supplies to the Caniapiscou reservoir are generally very low and stable during winter months ( $<250 \text{ m}^3 \text{ s}^{-1}$ ). The reservoir then starts to fill at the beginning of the spring flood, which generally occurs in late April or early May. May is clearly a month of increasing discharges (water supplies increase from  $\sim 300$  to  $\sim 2500 \text{ m}^3 \text{ s}^{-1}$ , on average), owing to the rapid snowmelt at that time of the year. The highest yearly peaks are normally attained in late May or early June. The spring flood then declines, and in August the discharge stabilizes around  $1000 \text{ m}^3 \text{ s}^{-1}$ .

[20] In this study, May ( $Q_{\text{May}}$ ) and June ( $Q_{\text{June}}$ ) water supplies (spring flood months) were computed from daily harmonized water supplies to the Caniapiscou reservoir [Hydro-Quebec, 2006]. We present a reconstruction of May water supplies since during that period, ice cover predominates on the lakes of the area.

[21] Observed water supply fluctuations are depicted in Figure 4. Inflows to the Caniapiscou reservoir have increased notably since approximately 1970. A central question remains, however, Is this increase unprecedented in recent



**Figure 2.** The LaGrande Hydroelectric Complex in northern Quebec. Black circles represent sites from which densitometric data was retrieved to reconstruct historical water supplies to the Caniapiscou reservoir. White circles are sites where densitometric data was not correlated to May water supplies and therefore not used in the present study. Lake Montausier's location is shown by a star.



**Figure 3.** Mean yearly hydrograph of water supplies to the Caniapiscou reservoir. The gray zone represent the fifth and ninety-fifth percentile.

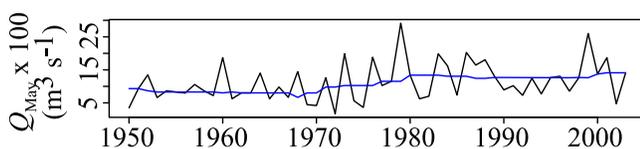
centuries? We exploited the hydrological information contained in tree rings to answer this question.

### 3.2. Tree Ring Data

#### 3.2.1. Discrete Tree Ring Series

[22] Lake Montausier ( $\sim 11.3 \text{ km}^2$ ;  $54^\circ 45' 03'' \text{N}$ ,  $70^\circ 09' 57'' \text{W}$ ) is an ungauged high-boreal water body located immediately to the northwest of the Caniapiscou reservoir, James Bay, northern Quebec (Figures 2 and 5). The Caniapiscou reservoir is at the head of one of the world's most important hydroelectric complexes: the La Grande Rivière (LG) Hydroelectric Complex. The construction of dams and the flooding of the reservoir in the mid-1980s greatly affected Lake Montausier's hydrological regime, lowering its water level by about half a meter. The original size of Lake Montausier's drainage area remains unknown at the time of writing because most of its basin was submerged after  $\sim 1985$ . Shoreline vegetation is characterized by the codominance of black spruce (*Picea mariana* (Mill. BSP)) and eastern larch (*Larix laricina* (Du Roi) K. Koch). Ice-scouring-tolerant shrubs such as green alder [*Alnus viridis* ssp. *crispa* (Ait.) Turill], diamondleaf willow (*Salix planifolia* Pursh), and American dwarf birch (*Betula glandulosa* Michx) are omnipresent along Lake Montausier's shores.

[23] Field sampling on Lake Montausier was conducted during the first week of June 2004. A total of 134 scar-bearing black spruce stems were selected. Trunks  $< 10 \text{ cm}$  diameter were not examined. Transversal cross-sections were taken at the center of each visible scar. Trees were also cut in multiple sections to ensure that no closed scars were left unsampled. In the laboratory, cross-sections were finely sanded and tree rings counted from the last year of growth (2004) to the center in order to approximate the date of tree establishment. A total of 304 scars were dendrochronologically dated, taking care not to replicate

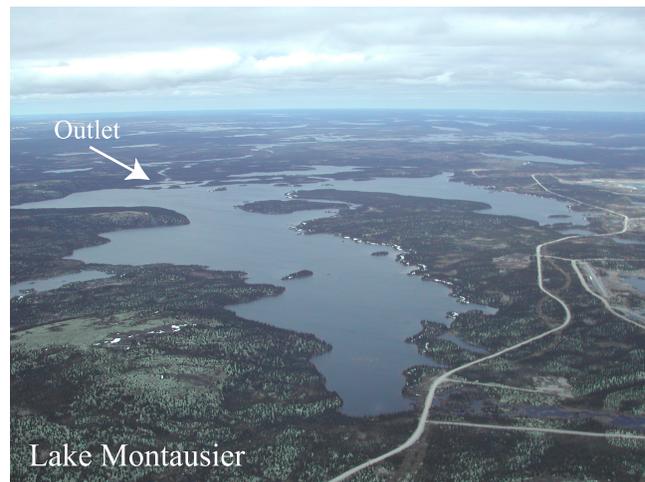


**Figure 4.** May waters supplies ( $Q_{\text{May}}$ ) to the Caniapiscou reservoir.

events found at multiple heights on the same tree. For each of the  $t$  years, the number of scars ( $R_t$ ) was counted along with the number of trees available (i.e., living) to record the events ( $A_t$ ).

#### 3.2.2. Continuous Tree Ring Series

[24] Earlywood measurements consisted of earlywood width (EW), minimal earlywood density (Dmin), and mean earlywood density (ED) measurements, all retrieved from a dendrochronological network in the James Bay area (Table 1, Figure 2). The network encompasses the La Grande Hydroelectric Complex. We retained 13 chronologies (all black spruce) that correlated ( $r > \pm 0.1$ ) to May water supplies (Figure 2). At each site, between 10 and 30 trees were sampled and a densitometric analysis was performed using the method described in Schweingruber [1978]. Density series were detrended using smoothing splines with a 30% variance cutoff. The sensitivity [Biondi and Qeadan, 2008] of these tree ring series is generally low (less than 0.2 to 0.2). However, the expressed population signal statistic (EPS) [Briffa and Jones, 1990] is fairly high ( $> 0.8$ ) in most cases, suggesting that the sampling effort is adequate.



**Figure 5.** Lake Montausier in early June 2004. Note that the ice cover has totally disappeared. The outlet is in the NW direction.

Table 1. Characteristics of Earlywood Density Measurements

Site <sup>a</sup>	DAIM		DAIX		ESKER		HURAULT		HURST		JOURDIN		ROZI		ROZM		ROZM		RT426		RT485		THIERH		THIERS		
	EW <sup>b</sup>	D <sub>min</sub> <sup>c</sup>	EW	D <sub>min</sub> <sup>c</sup>	EW	D <sub>min</sub> <sup>c</sup>	EW	D <sub>min</sub> <sup>c</sup>	ED <sup>d</sup>	D <sub>min</sub> <sup>c</sup>	EW	D <sub>min</sub> <sup>c</sup>	ED <sup>d</sup>	D <sub>min</sub> <sup>c</sup>													
Period	1766–2005	1743–2004	1877–2004	1746–2003	1738–2005	1768–2005	1831–2004	1769–2005	1769–2005	1769–2005	1769–2005	1764–2005	1757–2005	1779–2004	1804–2004												
Drainage	Mesic	Mesic	Mesic	Mesic	Mesic	Mesic	Mesic	Mesic	Mesic	Mesic	Mesic	Mesic	Mesic	Mesic	Mesic	Mesic	Mesic	Mesic	Mesic	Mesic	Mesic	Humid	Mesic	Mesic	Mesic	Mesic	
N trees	15	15	10	30	14	13	16	22	13	14	13	16	22	22	22	22	15	10	10	15	23	13	13	13	13		
Sensitivity	0.20	0.03	0.13	0.17	0.12	0.11	0.16	0.04	0.17	0.12	0.11	0.16	0.04	0.04	0.16	0.16	0.16	0.06	0.06	0.16	0.03	0.03	0.04	0.04	0.04		
EPS	0.86	0.78	0.68	0.87	0.77	0.81	0.85	0.84	0.77	0.77	0.81	0.85	0.84	0.87	0.87	0.87	0.77	0.06	0.06	0.06	0.87	0.87	0.82	0.82	0.82		
R	0.26	0.16	0.19	0.16	0.23	0.24	0.36	0.21	0.23	0.23	0.24	0.36	0.21	0.26	0.26	0.26	0.23	0.13	0.13	0.23	0.28	0.28	0.20	0.20	0.20		
R <i>W</i> <sub><i>t</i>(1)</sub>	0.35	-0.34	0.30	0.61	0.28	0.7	0.48	-0.44	0.28	0.28	0.7	0.48	-0.44	0.06	0.06	0.06	0.59	-0.41	-0.41	0.59	-0.08	-0.08	-0.08	-0.08	-0.08		
R <i>W</i> <sub><i>t</i>(2)</sub>	-0.16	-0.56	0.09	-0.03	0.04	0.04	0.01	-0.50	0.04	0.04	0.04	0.01	-0.50	-0.04	-0.04	-0.21	-0.39	-0.39	-0.21	-0.82	-0.82	-0.81	-0.81	-0.81	-0.81		
R <i>Q</i> <sub>May</sub>	0.19	-0.21	0.23	0.34	0.12	0.28	0.38	-0.33	0.12	0.12	0.28	0.38	-0.33	0.31	0.31	0.36	-0.38	-0.38	0.36	-0.41	-0.41	-0.38	-0.38	-0.38	-0.38		

<sup>a</sup>R corresponds to the mean Pearson correlation between trees in a site. *R W*<sub>*t*(1)</sub>, *R W*<sub>*t*(2)</sub>, and *R Q*<sub>May</sub> respectively represent the correlations with the two principal components (*W*<sub>*t*(1)</sub> and *W*<sub>*t*(2)</sub>) retained as predictor variables for spring flood and the correlation with *Q*<sub>May</sub>.

<sup>b</sup>EW, Earlywood width.

<sup>c</sup>D<sub>min</sub>, Minimal density of earlywood.

<sup>d</sup>ED, Earlywood mean density.

[25] The original set was reduced to two principal components (*W*<sub>*t*(1)</sub> and *W*<sub>*t*(2)</sub>), which explains 50% of the variance contained within the tree ring series.

### 3.3. Modeling Approach and Software

[26] Equation (2) was used to construct the GAM with the intent of reconstructing *Q*<sub>May</sub> over the last 150 years. Accordingly, since it is understood that May discharges cannot have a negative value, we modeled *Q*<sub>May</sub> as a gamma variable and represented *g* as an inverse link function. All computations were done in the *R* environment. Generalized additive models were performed using the *mgcv* 1.5.6 package (available at <http://cran.r-project.org/web/packages/mgcv/>). In all cases, penalized regression splines were used as smoothers for *I<sub>t</sub>* and *W<sub>t</sub>*.

[27] Validation statistics obtained via the jackknife procedure were the root-mean-square error (RMSE) and the reduction of error (RE)

$$RMSE = \sqrt{\frac{1}{m * i} \sum_{j=1}^{m * i} (Q_{May(j)} - \hat{Q}_{May(j)})^2}, \tag{6}$$

$$RE = 1 - \frac{\left[ \frac{1}{m * i} \sum_{j=1}^{m * i} (Q_{May(j)} - \hat{Q}_{May(j)})^2 \right]}{\left[ \frac{1}{m * i} \sum_{j=1}^{m * i} (Q_{May(j)} - \bar{Q}_{May})^2 \right]}, \tag{7}$$

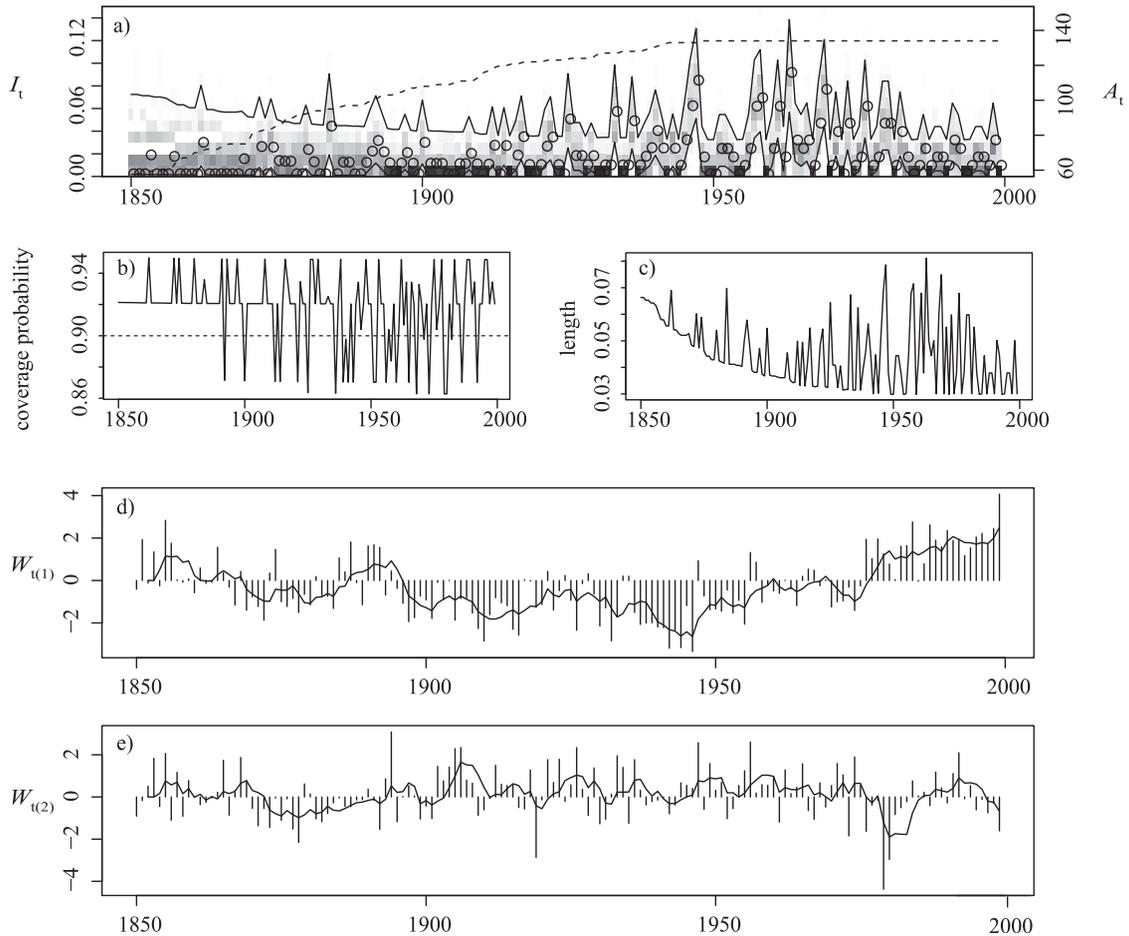
where *i* represents the number of synthetic series, *m* corresponds to the number of years in the calibration data set, *j* is the iteration number (*j*<sub>max</sub> = *m \* i*), *Q*<sub>May(*j*)</sub>, *Q*<sub>May(*j*)</sub>, and *Q*<sub>May</sub>, respectively, represent the observed, estimated, and mean water supplies. Median reconstructed *Q*<sub>May</sub> values are plotted with their 90% confidence interval.

## 4. Results and Discussion

### 4.1. Discrete and Continuous Tree Ring Record

[28] Between 1850 and 2000, the proportion of ice-scarred trees (*I<sub>t</sub>*) has considerably evolved (Figure 6a). Overall, observed *I<sub>t</sub>* values ranged from 0 (minimum) to 0.089 (maximum, A.D. 1963), with a mean value of 0.016 for the whole period. Since ~1930, however, *I<sub>t</sub>* values have increased notably. Before ~1930, the average proportion of scarred trees was 0.009. Since that time, that proportion has more than doubled, reaching 0.023 on average, with *I<sub>t</sub>* > 0.05 values recorded in 1933, 1936, 1946, 1947, 1957, 1958, 1963, 1969, and 1976.

[29] As anticipated, uncertainties around *I<sub>t</sub>* are larger toward the past because the estimation of proportions relies on an increasingly smaller number of available trees (*A<sub>t</sub>*) (Figure 6a). For example, *A<sub>t</sub>* decreases from 134 to 124 between 1999 and 1925 and drops to less than 108 before 1900 (Figure 6a). The length of 90% confidence intervals varies accordingly, changing from 0.04 to 0.07 for the same periods (Figure 6b). Coverage probabilities for the Wilson score confidence intervals since 1850 are depicted on Figure 6c, enabling us to verify that the intervals have an adequate size (not too large or too narrow). This analysis shows that the confidence intervals are close to the nominal



**Figure 6.** Discrete and continuous tree ring series used to predict  $Q_{\text{May}}$ . (a) The ice scar chronology of Lake Montausier. Observed proportions of scarred trees are represented by circles. The Wilson score confidence interval corresponds to the bold lines. The number of trees available ( $A_t$ ) is represented by the dashed line while the number of synthetic  $I_t$  series falling in each 0.01. The  $I_t$  class is represented on the raster image. A gradient from white (0 series) to black (100 series) illustrates the yearly dispersion of each possible  $I_t$  value. (b and c) The coverage probabilities and length of the confidence intervals. (d and e) Continuous series are the two first PCs ( $W_{t(1)}$  and  $W_{t(2)}$ ) of a network of earlywood density measurements (black circles in Figure 2).

confidence level (90%), even at the beginning of the chronology, where many  $I_t = 0$  are found. This suggests that the Wilson score method is well suited for modeling uncertainties around discrete event chronologies without being too permissive or too conservative, even for years where no events are recorded.

[30] The distribution of synthetic  $I_t$  series for each year within the discrete event chronology allows for a nuanced interpretation of past ice-scouring activity. The number of synthetic series falling into each equally incremented (0.01)  $I_t$  class provides an indication of the “plausibility” of each proportion value (Figure 6a). Darker colors indicate more certain estimations since they reflect a leptokurtic distribution of synthetic  $I_t$  values. In Figure 6a, one can easily observe that the proportion of scarred trees is less precisely estimated prior to  $\sim 1900$ . Further back in time, the raster image presents an unequivocal period of low  $I_t$  extending until about 1930. Finally, the post-1930 period corresponds to a clear rise in the ice scouring activity. After  $\sim 1985$ , however,  $I_t$  values drop substantially, probably as a result

of the modification of the lake’s hydrological regime due to the damming of its tributaries during the construction of the Caniapiscou hydroelectrical reservoir.

[31] The first principal component of the density series ( $W_{t(1)}$ ) presents an important low frequency oscillation component (Figure 6d).  $W_{t(1)}$  fluctuates around zero before  $\sim 1900$  and drops below that value between 1900 and  $\sim 1975$ . Although the values start rising after 1950,  $W_{t(1)}$  only becomes positive over the last 25 years. The second principal component ( $W_{t(2)}$ ) has a quite different pattern (Figure 6e). This series remains positive until  $\sim 1870$  but falls below zero during the last part of the nineteenth century.  $W_{t(2)}$  then oscillates around zero until  $\sim 2000$  and becomes slightly positive afterward. Additionally, the oscillation character of  $W_{t(2)}$  is of much higher frequency.

## 4.2. Nonlinear Relationships With $Q_{\text{May}}$

[32] In this section, the relationships between each tree ring indicator ( $I_t$ ,  $W_{t(1)}$ , and  $W_{t(2)}$ ) and  $Q_{\text{May}}$  are presented. The relationships are described for each predictor taken

**Table 2.** Comparison of  $R^2$ ,  $RMSE$ , and  $RE$  Statistics for the GAM and the MLR Model<sup>a</sup>

Model	GAM					MLR (eq)				
	$R^2_{Full}$	$R^2_c$	$R^2_v$	RMSE	RE	$R^2_{Full}$	$R^2_c$	$R^2_v$	RMSE	RE
gam.it	<b>0.43</b>	<b>0.22</b>	<b>0.15</b>	<b>503</b>	<b>0.15</b>	0.07	0.03	0	574	-
gam.wt1	<b>0.42</b>	<b>0.35</b>	<b>0.18</b>	501	<b>0.17</b>	0.26	0.22	0.15	501	0.11
gam.wt2	<b>0.49</b>	<b>0.36</b>	<b>0.22</b>	<b>448</b>	<b>0.26</b>	0.37	0.26	0.21	501	0.18
gam.it.wt	<b>0.75</b>	<b>0.61</b>	<b>0.44</b>	<b>398</b>	<b>0.36</b>	0.49	0.40	0.23	488	0.12

<sup>a</sup> $R^2_{Full}$  was calculated on the full period,  $R^2_c$  is the average value calculated on the training dataset, and  $R^2_v$  designates the validation coefficient (linear comparison between observed and simulated values). The strongest coefficients are in bold.

individually at first. Then, a true GAM is built to evaluate their additive effect. We compare the performances of our GAMs with a more traditional MLR equivalent built with the same input variables. Adjusted  $R^2$  values are presented, and all models were first constructed on the full calibration data set (i.e., of size  $m$ ). The results are presented in the second column of Table 2.

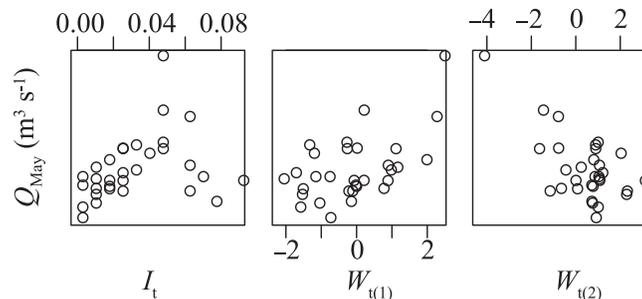
[33] The relationship between  $I_t$  and  $Q_{May}$  is highly nonlinear (Figures 7 and 8a). A GAM (gam.it) fitted on observed  $I_t$  values such that  $g[E(Q_{May})] = \beta_0 + f_1 I_t + \varepsilon$  depicts a humped relationship with  $Q_{May}$ . For example, while  $I_t \leq 0.05$ , the relationship between the proportion of scarred trees and  $Q_{May}$  is clearly positive-linear, meaning that as long as that critical value is not exceeded,  $Q_{May}$  increases with  $I_t$  in quite a predictable manner. However, for  $I_t > 0.05$ , this relationship reverses. In other words, years with an exceptionally high proportion of scarred trees do not occur during years of high  $Q_{May}$ . Instead, they seem to occur during years of low-to-moderate May discharges. Nevertheless, gam.it fitted the humped relationship fairly well (Table 2), with an  $R^2$  value of 0.43 and 50% of deviance explained by the variations in  $I_t$ , alone. By comparison, if the transfer model had been constructed from a linear regression approach, it would have been impossible to fit the parabolic relationship (adjusted  $R^2 = 0.07$ ,  $p > 0.05$ ). Therefore, the hump-shaped relationship between  $I_t$  and  $Q_{May}$  is an example of a complex interaction that is well captured within a GAM. This approach seems to be very efficient in the fitting of the nonlinear relationships that occur naturally, at spring, in cold-environment lakes. It is particularly useful in the context of ice flood modeling, where it is very difficult to have an a priori knowledge of the form of the “best fit” relationships that exist between dendrogeomorphological predictors (ice scars) and regional hydrological variables, as many factors imbricate.

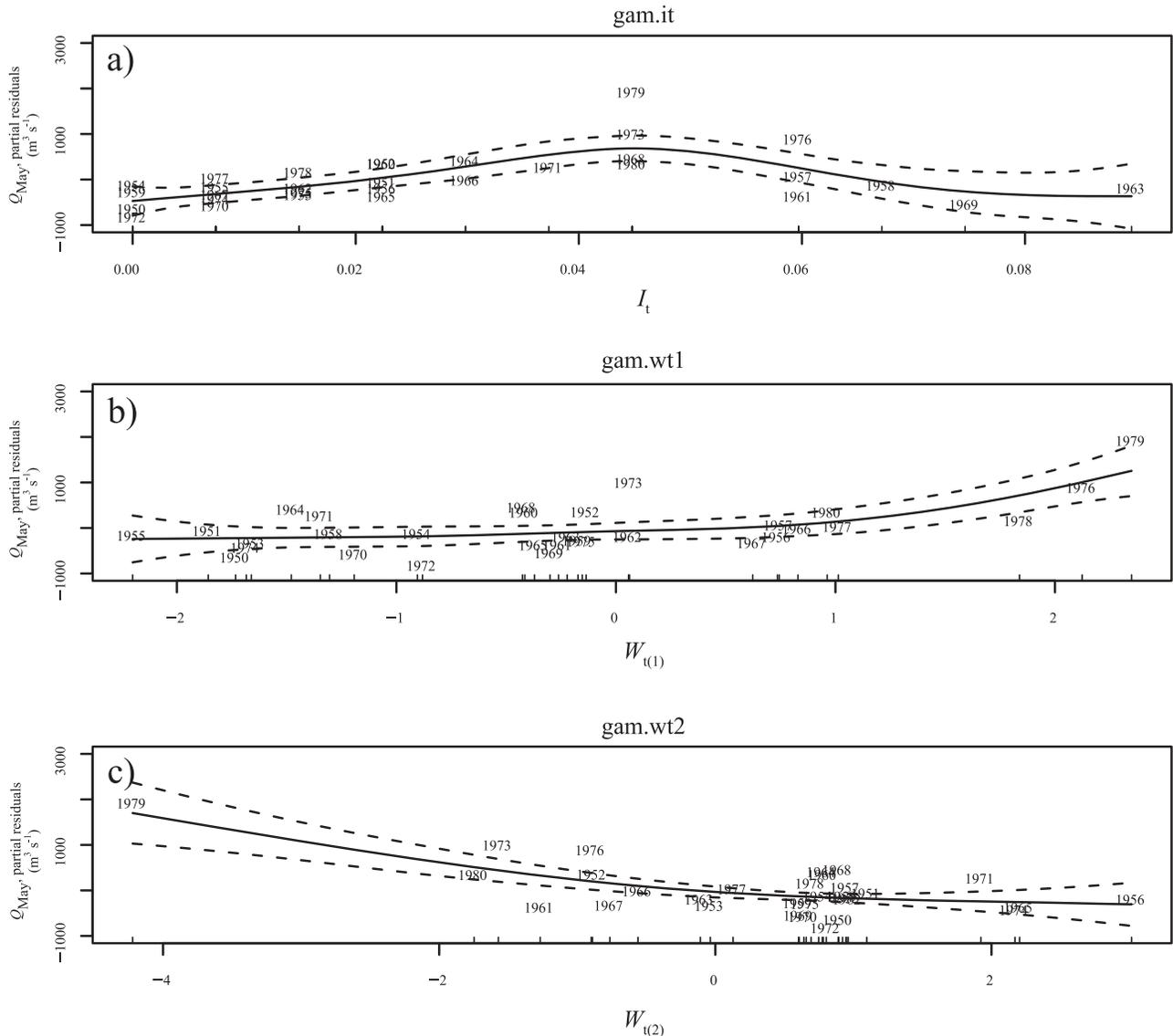
[34] Such a humped relationship, however, suggests that, when the Caniapiscaw water supplies are very low in May, both a small and a high proportion of scarred trees can be

expected (Figures 7 and 8a). This surprising situation may be caused by ice jams occurring at Lake Montausier’s outlet. Ice jams can temporarily raise and maintain the ice-covered lake level above its usual height, therefore provoking the massive scarification of the vegetation despite the fact that  $Q_{May}$  remains low. Ice jams are very common in these cold climates. In similar conditions, Boucher *et al.* [2009b] estimated that ice jams have occurred once every four years over the last 150 years. These events are typically generated in a context of rapid and sudden snow melt triggered by rain-on-snow events [Boucher *et al.*, 2009a]. For instance, the three years with the highest  $I_t$  values (i.e., 1958, 1963, and 1969) were all characterized by intense rain episodes in May. In each of these years, the maximum daily amount of rain was  $>25$  mm (Environment Canada, 2011, [http://climate.weatheroffice.gc.ca/Welcome\\_e.html?&](http://climate.weatheroffice.gc.ca/Welcome_e.html?&)), a situation that can generate serious ice jam problems.

[35] The relationships between  $W_{t(1)}$ ,  $W_{t(2)}$ , and  $Q_{May}$  are much more straightforward. For example, model gam.wt1 of the form  $g[E(Q_{May})] = B_0 + f_1(W_{t(1)}) + \varepsilon$  fits an exponential positive relationship (Figures 7 and 8b). Since the EW chronologies are all positively correlated to the first principal components (PC) (Table 1),  $W_{t(1)}$  probably represents the positive action of early spring precipitation on tree growth. Precipitation has a direct effect on cell elongation [Kozłowski *et al.*, 1991; Zhaner, 1968] and favor cell growth and turgescence [Catesson, 1990]. Hence, when water absorption by roots is greater than evaporation (which is usually the case in spring), the cellular volume increases, which leads to an earlywood that is larger and less dense. The adjusted  $R^2$  value for this relationship is 0.415, and the gam.wt1 model explains about 48% of the deviance. Again, the GAM approach performs slightly better than the linear regression, which yields an adjusted  $R^2$  value of 0.26.

[36] A second model, called gam.wt2 of the form  $g[E(Q_{May})] = B_0 + f_1(W_{t(2)}) + \varepsilon$ , enabled us to model the relationship between the second principal component and

**Figure 7.** Scatterplots of the relationships between  $Q_{May}$  and each tree ring predictor.



**Figure 8.** Nonlinear relationships between (a)  $I_t$ , (b)  $W_{t(1)}$ , and (c)  $W_{t(2)}$ , and May water supplies ( $Q_{\text{May}}$ ) to the Caniapiscau reservoir for the full calibration period (1950–1980,  $n = 31$ ). Partial residuals are presented here. They correspond to the Pearson residuals added to the smooth term for each covariate. The smoothed relationship (solid line) is presented with its 95% confidence interval (dashed line).

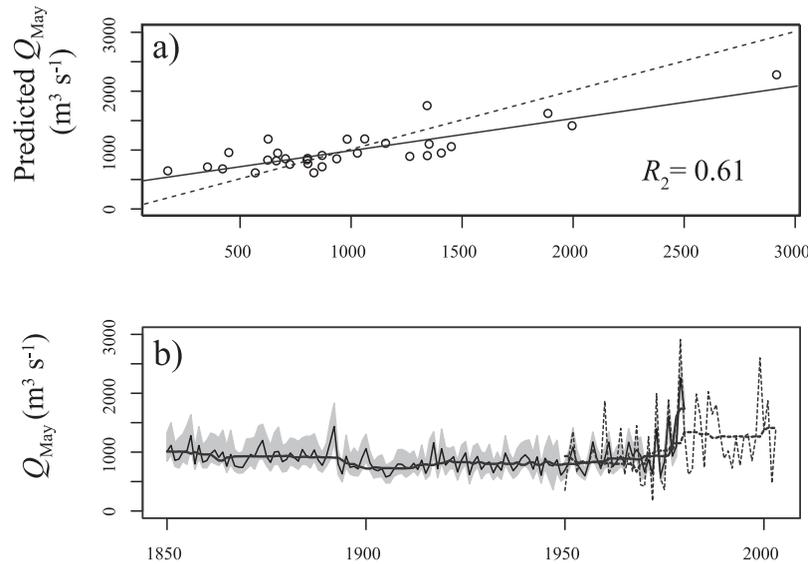
$Q_{\text{May}}$  (Figures 7 and 8c). The relationship is exponential negative. Since the  $D_{\text{min}}$  and ED chronologies all correlate negatively to  $W_{t(2)}$  (Table 2), it is likely that  $W_{t(2)}$  represents the negative impact of water on wood density. Wood cells produced during water-abundant conditions are known to be larger and thin-walled [Catesson, 1990], therefore lowering earlywood density. The model fits this negative relationship quite well, with an adjusted  $R^2$  value of 0.49 and about 52% of the deviance explained. A linear model could also have been used to fit this relationship, but the adjusted  $R^2$  values would have been smaller (0.37).

[37] Combining  $I_t$ ,  $W_{t(1)}$ , and  $W_{t(2)}$  into a single model (gam.it.wt) to predict  $Q_{\text{May}}$  yields even better adjustments (Table 1). With the three variables together, the model truly becomes additive and predicts  $Q_{\text{May}}$  quite well, with an adjusted  $R^2$  of about 0.75 and 81% of deviance explained.

The GAM approach provides a better fit than the MLR approach since, with the same predictors, the multiple linear regression yields significant, but unsatisfactory, results (adjusted  $R^2 = 0.49$ ,  $p < 0.01$ ).

### 4.3. Calibration and Validation of the Generalized Additive Model

[38]



**Figure 9.** (a) Observed versus simulated  $Q_{\text{May}}$  values. The 1:1 relationship corresponds to the dashed line. (b) Reconstruction of May water supplies to the Caniapiscou reservoir. The bold line and the gray shading correspond to the mean 10 year running average and the 90% confidence interval, respectively. Observations correspond to the dashed line.

calibration and validation adjusted  $R^2$  ( $R^2_c$ ,  $R^2_v$ ), RMSE, and RE values for each model.

[39] The GAM approach systematically yields better calibration and validation statistics than the traditional MLR approach (Table 2). This suggests that, in general, the relationships found by the GAM approach are robust and can be generalized to independent data. However, it is important to underline that, when comparing  $R^2_{\text{Full}}$  to both  $R^2_c$  and  $R^2_v$ , the values drop significantly when  $I_t$  is used either solely or jointly with  $W_t$  to predict  $Q_{\text{May}}$ . This is an inevitable consequence of our modeling approach. There are two explanations for that drop. The first is that the calibration period is quite short,  $m = 31$ . Thus, calibration of the model on a training  $m - 1$  size data set might be more sensitive to the jackknife procedure described in section 2. The second is that uncertainties around observed  $I_t$  values are included within the calibration process. At each iteration  $i$ , a new synthetic  $I_t$  series is included and a new GAM is recalibrated. Each iteration therefore contributes to an increasing departure from the  $R^2$  coefficients of the gam.it and gam.it.wt models originally computed from the fixed  $I_t$  series.

[40] For example, reusing gam.it.wt and incorporating the  $i = 100$  synthetic  $I_t$  series along with the corresponding  $W_{t(1)}$  and  $W_{t(2)}$  values in each of the  $m = 31$  training data sets yielded a median adjusted  $R^2$  value of 0.61 (Table 1). The fifth and ninety-fifth percentiles around the median adjusted  $R^2_c$  values are 0.57 and 0.75, respectively. Thus, the incorporation of uncertainties in the estimation of  $I_t$  and the iterative constitution of a calibration data set resulted in a more conservative fit. Then, at each iteration, a jackknife sample was removed and used for model validation. Following that procedure, we created  $i = 100$  jackknife validation data sets (of size  $m$ ) that were used to calculate validation statistics. As expected, the comparison between observed and simulated values, in each data set, yielded slightly smaller, although still significant,  $R^2$  values. The

median adjusted  $R^2_v$  was 0.44, and its 90% confidence interval was between 0.27 and 0.58. Graphical comparison between the observed and the median simulated  $Q_{\text{May}}$  values (Figure 9a) suggests that gam.it.wt may have slightly overestimated small discharges (e.g.,  $Q_{\text{May}} < 1000 \text{ m}^3 \text{ s}^{-1}$ ) and underestimated larger ones ( $Q_{\text{May}} > 1000 \text{ m}^3 \text{ s}^{-1}$ ). Nonetheless, gam.it.wt appears to be an acceptable model, as its RMSE coefficient (398.35) is smaller than the standard deviation of observed  $Q_{\text{May}}$  values (577.22) and the RE statistic is positive (0.36), indicating that our model has significant predictive skills.

#### 4.4. Reconstruction of $Q_{\text{May}}$ Since 1850

[41] Retrospective predictions were performed using gam.it.wt after replacing  $I_t$  and  $W_t$  values with those of the preinstrumental period. In total, gam.it.wt was recalibrated  $j = 3100$  times ( $j_{\text{max}} = mi = 31 \times 100$ ; see section 2.2), and at each time, a reconstruction was performed. The reconstruction presents the median  $Q_{\text{May}}$  value surrounded by a 90% confidence interval and extends back to 1850 (Figure 9b).

[42] Three distinct hydrological periods can be distinguished. The first period was between 1850 and 1900, when water supplies were of moderate amplitude. During that period, mean  $Q_{\text{May}}$  values oscillated around  $913 \pm 168 \text{ m}^3 \text{ s}^{-1}$  (mean  $\pm$  SD). This first phase was followed by a prolonged low flow period (1901–1965) when  $Q_{\text{May}}$  values oscillated around  $811 \pm 135 \text{ m}^3 \text{ s}^{-1}$ . Finally, water supplies rose after 1965. This recent period, which extends to today, is characterized by higher but also more variable  $Q_{\text{May}}$  values. The mean  $Q_{\text{May}}$  values of this period are around  $1190 \pm 668 \text{ m}^3 \text{ s}^{-1}$ .  $Q_{\text{May}}$  peaks during this period, i.e., 1973, 1976, 1979, 1983, 1986, and 1999, have been unequalled in magnitude for the last 150 years. Finally, we compared the mean 1950–2003  $Q_{\text{May}}$  to the mean 1850–2003  $Q_{\text{May}}$  value. Since the former is larger than the latter ( $1111 \text{ m}^3 \text{ s}^{-1}$  versus  $930 \text{ m}^3 \text{ s}^{-1}$ ), our data suggest that the mean value

recorded during the instrumental period may overestimate the average May water supply to the Caniapiscou reservoir.

[43] This is the first continuous spring flood reconstruction in the area. However, other discontinuous ice scar series exist in nearby environments such as Corvette Lake [Lemay and Bégin, 2008], Bienville Lake [Bégin, 1999, 2000a, 2000b; Lepage and Bégin, 1996], Clearwater Lake [Bégin, 2000a, 2000b; Bégin and Payette, 1988], and Duparquet Lake [Tardif and Bergeron, 1997]. Although spring floods have never been reconstructed from these chronologies, some share similarities with the present study, therefore highlighting the regional character of our reconstruction. For example, Tardif and Bergeron [1997] suggested that intermediate water levels occurred during the 1850–1900 period, according to an analysis of scar heights around Duparquet Lake in the southern boreal forest. An intermediate level of ice flood activity was also reconstructed for that period in Corvette [Lemay and Bégin, 2008] and Bienville [Bégin, 1999; Lepage and Bégin, 1996; Bégin, 2000] lakes in high-boreal and subarctic Quebec.

[44] The 1900–1965 low flow period was also observed in these lakes. However, Corvette, Bienville, Clearwater, and Duparquet lakes all seem to have experienced a period of higher ice-flooding activity during the 1930s–1940s, while in our reconstruction, that period is of much smaller magnitude. It is important to recall that ice flooding in these very large lakes occurs during years of very high spring discharge (e.g., high mean flood discharge or high peak discharge). In our study, we focused on reconstructing May water supplies mostly because Lake Montausier is a very small water body whose ice flooding activity is responsive to early spring hydroclimatic conditions. An important element to remember is that, at least for the Caniapiscou reservoir, the behavior of water supplies in May does not necessarily influence the magnitude of the rest of the flood for a given year. For example, we tested how  $Q_{\text{May}}$  relates to  $Q_{\text{June}}$ , flood peak, and flood volume using a correlation analysis and found no evidence of a significant relationship (all  $r < 0.1$ ,  $p < 0.05$ ) for the 1950–2003 period. Therefore, it is important to underline that we do not deny the existence of a high water level phase during the 1930s–1940s. However, our reconstruction suggests that  $Q_{\text{May}}$  may not have been significantly higher during that period, a result that does not, in itself, reduce the possibility that major spring floods could have occurred later in the season.

[45] Finally, our reconstruction revealed that the recent period (after  $\sim 1965$ ) has been characterized by a rapid increase in May water supplies, a situation that strongly suggests that spring floods are tending to occur increasingly earlier in the Caniapiscou area. To identify the general climatic processes driving these variations, we performed a standard correlation analysis with key climatic explanatory variables such as mean May temperature, sum of precipitation (total), and total accumulation of rain during the month of May (from the Schefferville station, 1957–2003). This analysis suggests that only mean May temperature is positively correlated ( $r = 0.54$ ) to  $Q_{\text{May}}$ . Thus, we conclude that  $Q_{\text{May}}$  is a temperature sensitive variable and that its evolution through the last 150 years reflect, at least partly, the evolution of May temperatures in the Caniapiscou reservoir area.

[46] Future research should aim at incorporating the various discrete and continuous chronologies available in the James Bay area in order to perform a first regional analysis of spring floods. Such an approach could, for example, study and model the oscillatory nature of tree ring series and their relationships with low-frequency climatic oscillation indices that are influential in the area (Arctic Oscillation, North Atlantic Oscillation). Empirical mode decomposition (EMD) methods [Lee and Ouarda, 2010] could be an interesting tool to carry out such an investigation.

## 5. Conclusions

[47] The acquisition of high-resolution historical information on past spring flood discharges in northern environments is of crucial importance, owing to the economic and hydrological significance of these events. Because of their high resolution, tree rings are generally well suited to retrieve this kind of information. However, in cold environments, some particularities must be underlined. For example, one is very unlikely to find sites where water is the only factor limiting tree growth. Thus, simply relying on continuous tree growth parameters such as ring widths or densities to reconstruct past spring discharges may be hazardous. We strongly believe that, in this context, the reliability of the reconstruction can be improved by including discontinuous tree ring series constructed from trees that are directly affected by water level fluctuations. Discrete indicators such as the proportion of ice-scarred trees around a lake are conspicuous evidence of past ice floods at the local scale. Such indicators have remained unexploited for the quantitative reconstruction of regional spring discharge beyond the instrumental period, despite the fact that a longer historical record would be of great interest for the planning of water resources. Moreover, the uncertainty associated with these indicators and the nonlinearities of their relationships with regional discharge have remained unaddressed until now.

[48] We developed a novel method that makes it possible to perform spring flood reconstructions in such environments. We take advantage of both continuous and discontinuous tree ring series, as they have their respective strengths and weaknesses. The procedure described herein makes use of computer-intensive techniques in order to (1) model the uncertainty associated with field-based estimation of ice-scouring indicators' yearly values (those of  $I_i$  in particular); (2) incorporate this changing uncertainty into the process of model calibration/reconstruction of past spring discharges; (3) use a GAM approach to model the nonlinear relationships that naturally characterize the interactions between tree ring indicators and spring discharges; (4) evaluate the calibration model's generalization capabilities using a jackknife procedure; and (5) perform hydrological reconstructions and compute their uncertainties.

[49] We tested our model in a case study in northern Quebec with the objective of reconstructing May water supplies to the Caniapiscou reservoir from continuous and discontinuous tree ring proxies. Satisfying calibration and verification results were obtained in this preliminary example. Although our model could be improved by elongating the calibration period, we strongly believe that a similar procedure can be used in other situations where spring

floods are reconstructed from discrete and continuous tree ring series. Instead of using ice scars, for example, one could easily use vessel anomalies [St. George and Nielsen, 2000, 2003; St. George et al., 2002; Yanosky, 1983] as a discrete tree ring proxy for spring floods. Additionally, our approach could serve as a basis for a large-scale regional analysis of spring floods in northern environments, which would incorporate all available discrete and continuous tree ring series.

[50] **Acknowledgments.** We gratefully acknowledge the following organizations for their financial support: NSERC (Natural Sciences and Engineering Research Council of Canada), CRC (Canada Research Chair program), FQRNT (Fonds québécois de la recherche sur la nature et les technologies), Hydro-Québec, Ouranos, ArcticNet, and the Northern Student Training Program of the Department of Indian and Northern Affairs of Canada.

## References

- Agresti, A., and B. A. Coull (1998), Approximate is better than “exact” for interval estimation of binomial proportions, *Am. Stat.*, *52*, 119–126.
- Alestalo, J. (1973), Dendrochronological interpretation of geomorphic processes, *Fennia*, *105*, 1–140.
- Bégin, Y. (1999), Tree-ring dating of extreme lake levels at the subarctic-boreal interface, *Quat. Res.*, *55*, 133–139.
- Bégin, Y. (2000), Augmenting the potential for reconstruction of lake ice-floods using tree-ring records, paper presented at International Workshop on River Environments Considering Hydraulic and Hydrologic Phenomena in Snowy and Cold Regions, INRS-EAU, Quebec, Canada.
- Bégin, Y. (2000a), Reconstruction of subarctic lake levels over past centuries using tree rings, *J. Cold Reg. Eng.*, *14*, 192–212.
- Bégin, Y. (2000b), Ice-push disturbances in high-boreal and subarctic lake-shore ecosystems since AD 1830, northern Québec, Canada, *Holocene*, *10*(2), 179–189.
- Bégin, Y., and S. Payette (1988), Dendroecological evidence of lake-level changes during the last three centuries in Subarctic Québec, *Quat. Res.*, *30*, 210–220.
- Biondi, F., and F. Qeadan (2008), Inequality in paleorecords, *Ecology*, *89*, 1056–1067.
- Bobée, B., and F. Ashkar (1991), *The Gamma Family and Derived Distributions Applied in Hydrology*, 203 pp., Water Resour. Publ., Littleton, Colo.
- Boucher, É., Y. Bégin, and D. Arseneault (2009a), Hydro-climatic analysis of mechanical breakups reconstructed from tree-rings, Necopastic watershed, northern Québec, Canada, *J. Hydrol.*, *375*, 373–382.
- Boucher, É., Y. Bégin, and D. Arseneault (2009b), Impacts of recurring ice jams on channel geometry and geomorphology in a small high-boreal watershed, *Geomorphology*, *108*, 273–281.
- Briffa, K. R., and P. D. Jones (1990), Basic chronology statistics and assessment, in *Methods of Dendrochronology: Applications in the Environmental Sciences*, edited by E. R. Cook and L. A. Kairiukstis, pp. 137–152, Kluwer Acad., Dordrecht, Netherlands.
- Case, R. A., and G. M. MacDonald (2003), Tree ring reconstructions of streamflow for three Canadian prairie rivers, *J. Am. Water Resour. Assoc.*, *39*(3), 703–716.
- Catesson, A.-M. (1990), Cambial cytology and biochemistry, in *The Vascular Cambium*, edited by M. Iqbal, pp. 63–112, Res. Stud. Press, Taunton, U. K.
- Church, M. (1988), Floods in cold climates, in *Flood Geomorphology*, edited by V. R. Baker, R. C. Kochel, and P. C. Patton, pp. 205–229, John Wiley, New York.
- Cook, E. R., and L. A. Kairiukstis (Eds.) (1990), *Methods of Dendrochronology: Applications in the Environmental Sciences*, 394 pp., Kluwer Acad., Dordrecht, Netherlands.
- Eilers, P. H. C., and B. D. Marx (1996), Flexible smoothing with B-splines and penalties, *Stat. Sci.*, *11*, 89–121.
- Ettema, R. (2002), Review of alluvial-channel responses to river ice, *J. Cold Reg. Eng.*, *16*, 191–217.
- Gottesfeld, A. S., and L. M. J. Johnson Gottesfeld (1990), Flood plain dynamics of a wandering river: Dendrochronology of the Maurice River, British Columbia, Canada, *Geomorphology*, *3*, 159–179.
- Gou, X., F. Chen, E. Cook, G. Jacoby, M. Yang, and J. Li (2007), Stream-flow variations of the Yellow River over the past 593 years in western China reconstructed from tree rings, *Water Resour. Res.*, *43*, W06434, doi:10.1029/2006WR005705.
- Graumlich, L. G., and L. B. Brubaker (1986), Reconstruction of annual temperature (1590–1979) for Longmire, Washington, derived from tree rings, *Quat. Res.*, *25*, 223–234.
- Hastie, T. J., and R. J. Tibshirani (Eds.) (1990), *Generalized Additive Models*, Chapman & Hall, New York.
- Henoch, W. E. S. (1973), Data (1971) on height, frequency of floods, ice jamming and tree-ring studies, *Tech. Rep. 5*, in *Hydrologic Aspects of Northern Pipeline Development, Rep. 73-3*, pp. 153–190, Task Force on North. Oil Dev., Dep. of the Environ., Ottawa, Ont.
- Hidalgo, H. G., T. C. Piechota, and J. A. Dracup (2000), Alternative principal components regression procedures for dendrohydrologic reconstructions, *Water Resour. Res.*, *36*(11), 3241–3249, doi:10.1029/2000WR900097.
- Hupp, C. R. (1988), Plant ecological aspects of flood geomorphology and paleoflood history, in *Flood geomorphology*, edited by V. R. Baker, R. C. Kochel, and P. C. Patton, pp. 335–356, John Wiley, New York.
- Hydro-Québec (2006), Plan stratégique 2006–2010, report, 62 pp., Montreal, Quebec, Canada.
- Intergovernmental Panel on Climate Change (2007), *Climate Change 2007: Impacts, Adaptation and Vulnerability*, 976 pp., Cambridge Univ. Press, Cambridge, U. K.
- Kozlowski, T. T., P. J. Kramer, and S. G. Pallardy (1991), *The Physiological Ecology of Woody Plants*, edited by J. Roy, 454 pp., Academic, San Diego, Calif.
- Kundzewicz, Z. W., L. J. Mata, N. W. Arnell, P. Döll, P. Kabat, B. Jiménez, K. A. Miller, T. Oki, Z. Sen, and I. A. Shiklomanov (2007), Freshwater resources and their management, in *Climate Change: Impacts, Adaptation and Vulnerability: Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, pp. 173–210, Cambridge Univ. Press, Cambridge, U. K.
- Lee, T., and T. B. M. J. Ouarda (2010), Long-term prediction of precipitation and hydrologic extremes with nonstationary oscillation processes, *J. Geophys. Res.*, *115*, D13107, doi:10.1029/2009JD012801.
- Lemay, M., and Y. Bégin (2008), Hydroclimatic analysis of an ice-scar tree-ring chronology of a high-boreal lake in Northern Québec, *Canada*, *39*, 451–464.
- Lepage, H., and Y. Bégin (1996), Tree-ring dating of extreme water level events at Lake Bienville, Subarctic Québec, Canada, *Arct. Antarct. Alp. Res.*, *28*, 77–84.
- Loaiciga, H. A., L. Haston, and J. Michaelsen (1993), Dendrohydrology and long-term hydrologic phenomena, *Rev. Geophys.*, *31*(2), 151–171, doi:10.1029/93RG00056.
- McCord, V. A. (1996), Fluvial process dendrogeomorphology: Reconstruction of flood events from the southwestern United States using flood-scarred trees, in *Tree-Rings, Environment, and Humanity: Proceedings of the International Conference*, edited by J. S. Dean, D. M. Meko, and T. W. Swetnam, pp. 689–702, Radiocarbon, Tucson, Ariz.
- Meko, D. M., and D. A. Graybill (1995), Tree-ring reconstruction of upper Gila River discharge, *J. Am. Water Resour. Assoc.*, *31*, 605–616.
- Meko, D. M., C. W. Stockton, and W. R. Boggess (1995), The tree-ring record of severe sustained drought, *J. Am. Water Resour. Assoc.*, *31*, 789–801.
- Meko, D. M., M. D. Therrell, C. H. Baisan, and M. K. Hughes (2001), Sacramento River flow reconstructed to AD 869 from tree rings, *J. Am. Water Resour. Assoc.*, *37*, 1029–1039.
- Miller, R. G. (1964), A trustworthy jackknife, *Ann. Math. Stat.*, *35*, 1594–1605.
- Payette, S. (1980), Les grandes crues glacielles de la rivière aux feuilles (Nouveau-Québec): Une analyse dendrochronologique, *Nat. Canadien.*, *107*, 215–225.
- Schweingruber, F. H. (1996), *Tree Rings and Environment Dendroecology*, 609 pp., Paul Haupt, Berne.
- Schweingruber, F. H., H. C. Fritts, O. U. Braecker, L. G. Drew, and E. Schär (1978), The X-ray technique as applied to dendroclimatology, *Tree-Ring Bull.*, *38*, 61–91.
- Shao, J., and D. Tu (1995), *The Jackknife and Bootstrap*, 516 pp., Springer, New York.
- Shroder, J. F., Jr. (1980), Dendrogeomorphology: Review and new techniques of tree-ring dating, *Prog. Phys. Geogr.*, *4*, 161–188.
- Sigafoos, R. S. (1964), Botanical evidence of floods and flood-plain deposition, *U.S. Geol. Surv. Prof. Pap. Rep. 485-A*, 35 pp.
- Smith, D. G. (1979), Effects of channel enlargement by river ice processes on bankfull discharge in Alberta, Canada, *Water Resour. Res.*, *15*(2), 469–475, doi:10.1029/WR015i002p00469.

- Smith, D. G., and D. M. Reynolds (1983), Tree scars to determine the frequency and stage of high magnitude river ice drives and jams, red deer, Alberta, *Can. Water Resour. J.*, *8*, 77–94.
- Smith, L. P., and C. W. Stockton (1981), Reconstructed stream flow for the Salt and Verde rivers from tree-ring data, *J. Am. Water Resour. Assoc.*, *17*, 939–947.
- St. George, S., and E. Nielsen (2000), Signatures of high-magnitude 19th century floods in *Quercus macrocarpa* tree rings along the Red River, Manitoba, Canada, *Geology*, *28*, 899–902.
- St. George, S., and E. Nielsen (2003), Palaeoflood records for the Red River, Manitoba, Canada, derived from anatomical tree-ring signatures, *Holocene*, *13*, 547–555.
- St. George, S., E. Nielsen, F. Conciatori, and J. Tardif (2002), Trends in *Quercus macrocarpa* vessel areas and their implications for tree-ring paleoflood studies, *Tree-Ring Res.*, *58*, 3–10.
- Stockton, C. W., and H. C. Fritts (1971), Long-term reconstruction of water level changes for Lake Athabasca by analysis of tree rings, *J. Am. Water Resour. Assoc.*, *9*, 1006–1027.
- Stockton, C. W., and G. C. Jacoby Jr. (1976), Long-term surface-water supply and streamflow trends in the Upper Colorado River Basin, *Lake Powell Research Project Bulletin*, 70 pp., Inst. of Geophys. and Planet. Phys., Univ. of Calif., Los Angeles.
- Stoffel, M., M. Bollschweiler, D. R. Butler, and B. H. Luckman (2010), *Tree Rings and Natural Hazards: A State-of-the-Art*, Springer, Dordrecht, The Netherlands.
- Tardif, J., and Y. Bergeron (1997), Ice-flood history reconstructed with tree-rings from the southern boreal forest limit, western Québec, *Holocene*, *7*, 291–300.
- Wilson, E. B. (1927), Probable inference, the law of succession, and statistical inference, *J. Am. Stat. Assoc.*, *22*, 209–212.
- Wood, S. N. (2000), Modelling and smoothing parameter estimation with multiple quadratic penalties, *J. R. Stat. Soc.*, *62*, 413–428.
- Wood, S. N., and N. H. Augustin (2002), GAMs with integrated model selection using penalized regression splines and applications to environmental modelling, *Ecol. Modell.*, *157*, 152–177.
- Woodhouse, C. A. (1999), Artificial neural networks and dendroclimatic reconstructions: An example from the Front Range, Colorado, USA, *Holocene*, *9*, 521–529.
- Woodhouse, C. A. (2001), A tree-ring reconstruction of streamflow for the Colorado Front Range, *J. Am. Water Resour. Assoc.*, *37*, 561–569.
- Woodhouse, C. A., and J. J. Lukas (2006), Multi-century tree-ring reconstructions of Colorado streamflow for water resource planning, *Clim. Change*, *78*, 293–315.
- Woodhouse, C. A., S. T. Gray, and D. M. Meko (2006), Updated streamflow reconstructions for the Upper Colorado River Basin, *Water Resour. Res.*, *42*(5), W05415, doi: 10.1029/2005WR004455.
- Yanosky, T. M. (1983), Evidence of floods on the Potomac River from anatomical abnormalities in the wood of flood-plain trees, *U.S. Geol. Surv. Prof. Pap.*, *1296*, 42 pp.
- Zhaner, R. (1968), Water deficits and growth of trees, in *Water Deficits and Plant Growth*, vol. 2, *Plant Water Consumption and Response*, edited by T. T. Kozolowski, pp. 151–254, Academic, New York.

---

Y. Bégin, É. Boucher, A. Nicault, and T. B. M. J. Ouarda, Centre Eau-Terre-Environnement, Institut National de la Recherche Scientifique, 490 de la Couronne, Québec, QC G1K 9A9, Canada. (boucher\_etienne@hotmail.com)