Journal of Hydrology 596 (2021) 126054

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Nonstationary flood-frequency analysis to assess effects of harvest and cover type conversion on peak flows at the Marcell Experimental Forest, Minnesota, USA

Zachary P. McEachran [a,*], Diana L. Karwan [a], Stephen D. Sebestyen [b], Robert A. Slesak [a,c], Gene-Hua Crystal Ng [d]

[a] University of Minnesota, Department of Forest Resources, 1530 Cleveland Ave N., St. Paul, MN 55108, United States
[b] USDA Forest Service, Northern Research Station, 1831 Highway 169 E, Grand Rapids, MN 55744, United States
[c] USDA Forest Service, Pacific Northwest Research Station, 3625 93rd Avenue SW, Olympia, WA 98512, United States
[d] University of Minnesota, Department of Earth and Environmental Sciences, 116 Church Street SE, Minneapolis, MN 55455, United States

A R T I C L E I N F O

Keywords:
Paired catchments
Bayesian parameter estimation
Flood-frequency
Flooding
Nonstationary
Forest harvesting

A B S T R A C T

Forest harvest, climate change, and their interaction can alter catchment peak discharges, which have ramifications for channel geomorphology and water quality. Catchments in the boreal-temperate transition zone may be especially vulnerable to these factors. We developed a new approach to peak flow analysis using two long-term, paired catchment experiments in that landscape at the Marcell Experimental Forest (MEF) in north-central Minnesota. We investigated how forest cover change affects the magnitude, occurrence probability, event decoupling, and seasonal decoupling of annual maximum peak flows. Event decoupling is when the annual maximum flows on control and treatment catchments occur in response to different events within a given year. Seasonal decoupling is when the annual maximum flows on control and treatment catchments occur in different seasons within a given year. Commonly used statistical methods (e.g., ANCOVA) support inferences only for effects on peak flow magnitude, disregarding effects on event occurrence probability. Statistical stationarity (i.e., time-invariant parameters) is generally assumed within the ANCOVA analysis, despite temporal changes in climatic and land cover controls. Further, use of these methods requires event coupling, which is not always valid for annual maximum flows in catchments with mixed precipitation regimes. To address these limitations, we developed new nonstationary flood-frequency analysis methods with Bayesian parameter estimation to analyze two harvesting experiments, including conversion from deciduous to coniferous species. We compared results from these models to traditional least-squares based methods (e.g., ANCOVA). We assessed event and seasonal decoupling using logistic regression with Bayesian parameter estimation. There was no effect of harvesting on the annual maximum flow according to ANCOVA and generalized least squares models, but these methods were not reliable due to the presence of outliers, and both event and seasonal decoupling. Flood-frequency results from one clearcutting experiment showed increases in the annual maximum flow across nearly all return intervals, with 80–85% confidence for peak discharges between the 10 and 50-year peak flow. The 50-year return interval peak increased from an expected 17 cfs to 34 cfs after harvest. Harvesting induced substantial event decoupling in the catchment pairs, and seasonal decoupling in one catchment pair. As forest cover regenerated, decoupling probability decreased. Our coupling analysis results indicated that control and treatment catchments generated and processed flood peaks differently after harvest, and explained the poor fit of the linear models for which we assumed event coupling. Our study calls into question several key assumptions of traditional linear models in the analysis of paired-catchments given certain conditions. For example, where forest harvesting can affect the season of annual maximum peaks, and thus the phase of the generating precipitation event (snowmelt versus rainfall), it changes the relationship between the catchment pair in a way not accounted for in traditional linear models. We advocate the use of probabilistic methods that can incorporate nonstationarity and are robust to different mechanisms of catchment hydrologic change in response to forest harvest.

* Corresponding author at: 209C Green Hall, 1530 Cleveland Avenue N., St. Paul, MN 55108, United States.
E-mail address: mceac015@umn.edu (Z.P. McEachran).

https://doi.org/10.1016/j.jhydrol.2021.126054

Available online 10 February 2021

1. Introduction

Forest cover change and forest harvesting effects on catchment processes, especially on the high streamflow regime, have been an area of research for decades (Andréassian, 2004; Yu and Alila, 2019; Zon, 1927). Alterations in peak flows can have serious ramifications for channel geomorphology, stream ecology, as well as water quality and sediment dynamics (Auerswald and Geist, 2018; Poff et al., 2006). Furthermore, floods can be devastating for local communities (Bradshaw et al., 2007). Forests are generally thought to offer some level of flood protection, which can be influenced by disturbance such as forest harvesting (Alila and Green, 2014; Andréassian, 2004; Birkinshaw, 2014; Green and Alila, 2012). Moving forward, changing climate and a growing demand for wood products may increase forest disturbance, underlying the ecological and societal importance of understanding the relationship between forests and peak stream flows (Angel et al., 2018; Buongiorno et al., 2012; Buras and Menzel, 2018).

In North America, forest harvesting and forest cover change effects on peak streamflows are predominantly examined in high-relief regions with bedrock close to the soil surface (Alila et al., 2009; Jones and Grant, 1996; Kuraś et al., 2012; Thomas and Megahan, 1998). Small (<10 km^2) paired-catchment studies have provided the foundation for much of the current knowledge in forest hydrology, including peak flows and flooding (Andréassian, 2004). Widespread disagreement remains and often stems from interpretation of the statistical method used to analyze paired-catchment data (Buttle, 2011). Historical methods that have addressed the relationship between forest cover change and peak discharges in paired-catchment experiments have typically relied on linear statistical models (e.g., ANOVA, ANCOVA: Hewlett, 1982; Jones and Grant, 1996; Thomas and Megahan, 1998) that utilize traditional frequentist approaches: they are designed to test for differences in mean values and rely on chronological pairing, such that differences between catchments are compared for the same precipitation event (Alila et al., 2009). Such methods have given rise to the general paradigm that the effect of forest harvesting on peak flows exhibits a decreasing effect size with increasing peak flow, and thus a greater proportional impact associated with frequent, low-discharge peak flows (Andréassian, 2004; Bathurst et al., 2011; Thomas and Megahan, 1998). The historic, chronologically paired approach does not account for both event magnitude *and* frequency (e.g., as reviewed by Alilia et al 2009). Analyses accounting for frequency of peak flows, mostly conducted in mountainous, snowmelt-dominated catchments, have found sizable effects for large peak flows, and even an increasing effect size with increasing return interval (Alila et al., 2009; Green and Alila, 2012; Kuraś et al., 2012; Yu and Alila, 2019). It is unclear how findings in these landscapes apply to those found in low-relief glaciated regions such as those found in central North America, Russia, and Scandinavia. Further, catchments at these northern latitudes are expected to be especially sensitive to climate change due to their highly climate-dependent and seasonal hydrologic cycle (Tetzlaff et al., 2015; 2013).

When examining peak flows, changes between control and treatment catchments may manifest in several ways, such as changes in flood magnitude, probability of occurrence, or a change in the type of precipitation event associated with the largest flow events (e.g., rainfall vs snowmelt). A shift in the seasonality and phase of the driving precipitation event has important ramifications for understanding flood generation and catchment processes producing high flows, as these may be different for snowmelt versus rainfall events. For example, post-harvest changes in hydrology in snowmelt-dominated and snowmelt-influenced catchments have the potential to increase peak flows even for large, infrequent events due to catchment hydrologic processes (e.g., increased catchment saturation, alterations in soil thermodynamics, freezing, and infiltration at the time of snowmelt, increased snowpack snow-water equivalent, etc.; Murray and Buttle, 2003; Pomeroy et al., 1994).

Possible shifts in catchment processes in response to harvest highlight the need for statistical methods that offer flexibility to evaluate the temporal structure of the statistical model of peak discharges. The integration of nonstationary representation of flood generation and flood-frequency analysis methods offers a more detailed framework compared to previously used methods, allowing inference about how peak flow magnitudes *and* frequencies change in response to forest cover and climatic changes through time at the small catchment scale (Yu and Alila, 2019). Further, use of Bayesian methods to frame peak flow analysis in a probabilistic framework allows for uncertainty to be easily calculated in context of the models fitted to observations and our background (or "prior") knowledge about the system.

We quantified the effect of forest cover change on flood and peak flows in two sets of paired catchments at the Marcell Experimental Forest (MEF) in north-central Minnesota, USA in order to test the hypotheses that clearcutting will increase annual peak flows across all return intervals compared to pre-harvest mature forest conditions, and that annual peak flows will decrease with time since harvest as the forest regrows. Additionally, we examined consistency of results and implications for the effects of harvesting on floods by comparing our nonstationary flood frequency analysis with traditional paired-catchment ANCOVA methods.

2. Methods

2.1. Study sites

The MEF catchment studies were initiated by the U.S. Forest Service in 1961 to investigate the hydrologic and ecological role of peatlands in the boreal-temperate transition zone (Fig. 1a). Each of six small (<100 ha) upland-peatland experimental catchments, named S1 – S6, consists of an upland forest surrounding a central peatland, from which drains a first-order stream (Fig. 1b). Aspen (*Populus* spp.) and northern hardwoods dominate the uplands, while black spruce (*Picea mariana*) and tamarack (*Larix laricina*) dominate the central peatlands. The MEF is located on a moraine where multiple ice sub-lobes terminated over the course of the last glaciation (26–12 ka BP) (Verry and Janssens, 2011). Upland soils are Alfisols or Entisols and generally have a clay loam texture, with Histosols in the peatlands (Nyberg, 1987). The climate is continental, with moist, warm summers and dry, cold winters (Sebestyen et al., 2011a). The mean annual precipitation from 1961 to 2009 was 78 cm (±11 cm, standard deviation), and the annual mean temperature was 3.4 °C (±13 °C) (Sebestyen et al., 2011a). About one-third of annual precipitation falls as snow between November and March. No significant change was found in annual precipitation or maximum annual snow water equivalent (SWE) under black spruce canopy and in treeless (open) areas from 1961 to 2009 (Sebestyen et al., 2011a). Annual maximum SWE did decrease under aspen during this same time period (Sebestyen et al., 2011a).

Previous work has quantified some effects of clearcutting and forest regrowth on peak discharges at the MEF from two long-term paired-catchment experiments, however the full distribution of annual maximum discharge in response to forest harvesting and recovery has not been analyzed. Previous studies, largely summarized by Sebestyen et al. (2011b) and Verry et al. (1983), included: (1) ANCOVA analyses to separately analyze rainfall and snowmelt-derived peaks (Sebesteyen et al., 2011b; Verry et al., 1983), (2) a frequency comparison up to the 10-year return interval discharge for rainfall-generated peaks only (Verry et al., 1983), and (3) physically based modeling of catchment response to harvest using the Peatland Hydrologic Impact Model: PHIM – see Guertin et al. (1987) (Lu, 1994). The PHIM modeling study used flood-frequency analysis of simulated values of separated snowmelt and rainfall peaks for the S4/S5 MEF clearcutting experiment (S4/S5 experiment detailed in Section 2.1.2; Lu 1994), with results supporting the paradigm of decreasing effect size with increasing return interval. In contrast, Verry et al. (1983) found the 10-year recurrence interval of summer rainfall-generated peak discharge after clearcutting increased more than the 2-year summer rainfall-generated peak discharge. The response across the entire range of recurrence intervals was not included. In this study, we examined nonstationary probability distributions of annual peak streamflow for rainfall and snowmelt generated annual peaks and compared the results to traditional ANCOVA methods using long-term data from two paired-catchment experiments at the MEF.

2.1.1. S4/S5 experiment

The S4/S5 experiment began with catchment instrumentation in 1962. The control catchment, S5, is 52.6 ha in size, and includes uplands and five small satellite wetlands draining into the 6.1 ha central peatland. The treatment catchment, S4, is 34.0 ha in size, with an 8.1 ha central peatland. Clearcut harvest of merchantable timber in the S4 uplands took place over two sequential fall-winter seasons (1970–71, 1971–72), with final removal of non-merchantable timber in summer 1972. Aspen regenerated on the upland harvest areas and received fertilizer application, according to an intentional study design, in 1978 (Sebestyen et al., 2011a). The regenerated aspen forest was considered mature 40 years after harvesting, based on commercial rotations in the Great Lakes region (Bradford and Kastendick, 2010).

The S4 watershed straddles the Laurentian Divide. Approximately 70% of its annual water yield drains north to the Hudson Bay (outlet S4N) and 30% drains south to the Mississippi (outlet S4S) (Sebestyen et al., 2011b). Streamflow was measured using several gaging structures since 1962, including flumes and weirs. All stage data were recorded on stripcharts, which were then digitized and converted to streamflow using stage-discharge relationships (Sebestyen et al., 2020b; Supplemental Information, Section 1). Previous analyses of peak flows from the S4 catchment have only included the S4N outlet. For this analysis, we used the combined discharge values (S4S + S4N) to identify the total annual maximum peak flow from the S4 catchment (Sebestyen et al., 2020b). Annual maximum flow series were derived by taking the highest discharge value within each water year, which was defined as March 1 – February 28/29.

2.1.2. S2/S6 experiment

The pretreatment period in the S2/S6 experiment began in 1976 with a four-year calibration period between catchments. The control

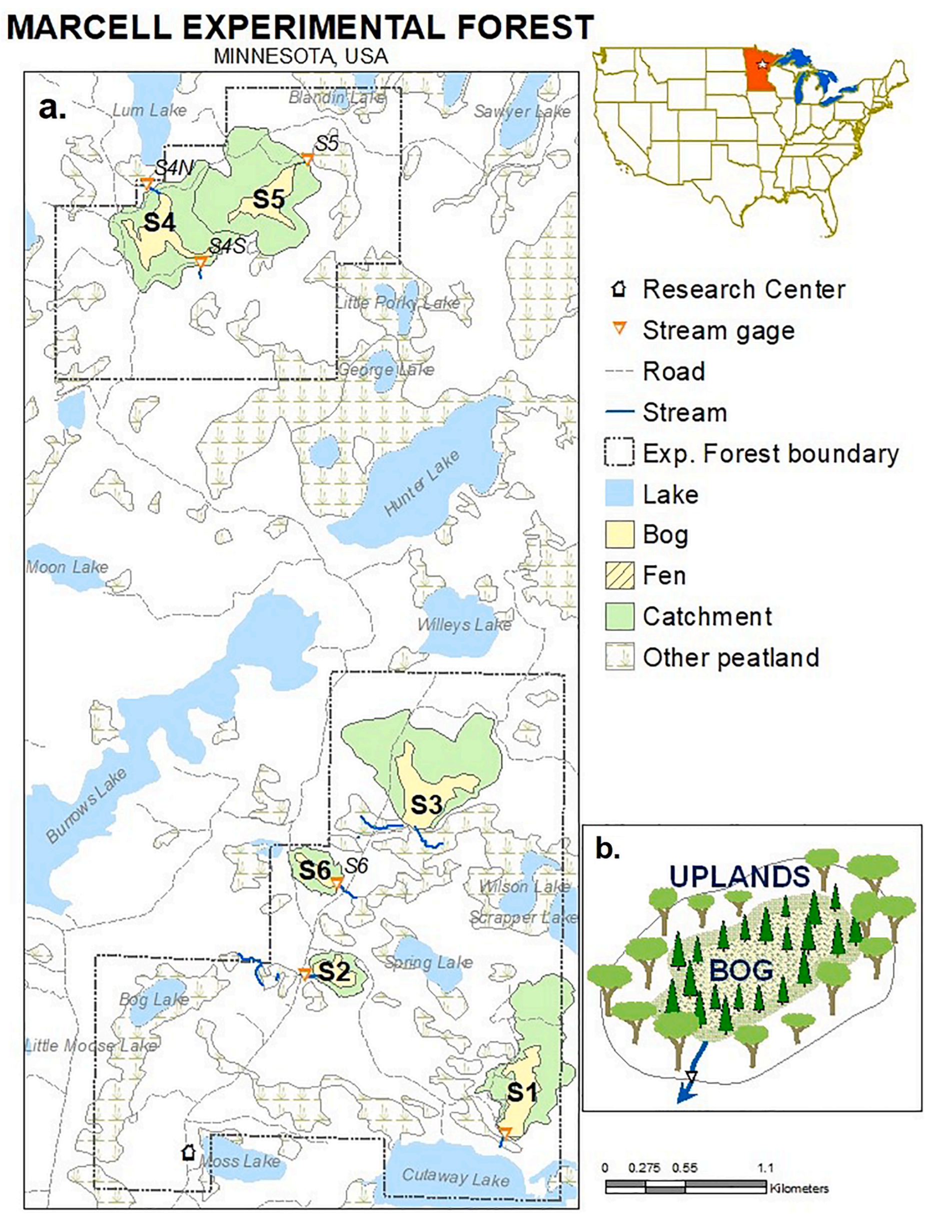

Fig. 1. a) Marcell Experimental Forest. b) a typical upland-peatland watershed such as those found at the MEF.

catchment, S2, is 9.7 ha in size, which includes a 3.2 ha central peatland. The treatment catchment, S6, is 8.9 ha in size, with a 2.0 ha central peatland. In March-June of 1980, the aspen uplands in the treatment catchment were clearcut-harvested, with 77% of the catchment area harvested. During the summers from June of 1980 to 1982, aspen regeneration was suppressed with grazing. In May of 1983, red pine (*Pinus resinosa*) and white spruce (*Picea glauca*) were planted in the clearcut area, with ~70% of the upland area in red pine and ~30% in white spruce. Annual maximum flow series were derived by taking the highest discharge value within each water year, which was defined as March 1 – February 28/29.

2.2. General approach

Three statistical methods were used to detect differences in annual maximum flows due to forest harvesting and regrowth, in each of the paired-catchment experiments. The first method, traditional linear regression, extended previous approaches at the MEF to examine differences in annual maximum flow magnitude (e.g., Sebestyen et al., 2011b). The second method, which we termed a "coupling analysis", utilized logistic regression analyses with parameters estimated using a Bayesian method. The coupling analysis was designed to determine the probability that the annual maximum flow in any given year was generated 1) by the same precipitation event on the control and treatment catchments, and 2) within the same season on the control and treatment catchments. Finally, we utilized nonstationary flood-frequency models of the annual maximum streamflow series fit to observed data using a Bayesian technique to explore differences in the entire probability distribution of annual maximum peak flow magnitudes across return intervals. The final analysis supported inferences about how flood magnitude and probability of occurrence change due to forest harvesting and regrowth. These three statistical approaches combine to offer complementary insights into the effects of forest cover change on the peak discharge magnitude, occurrence probability, and generating event/seasonality.

2.3. Analysis methods

2.3.1. Linear regression

The effects of different catchment treatments were assessed using least squares ANCOVA regression analysis, in which treatment and control catchment annual maximum discharge series were assumed to be related through a linear relationship. Slope and intercept effects were determined for pre- and post-treatment time periods (Table 1), following Sebestyen et al. (2011b). Models included a normally distributed error term with a constant residual variance per standard ANCOVA methods.

We detected outliers in both S4/S5 and S2/S6 experiment data through standardized residual and residual-versus-leverage plots and used a generalized least squares approach (GLS) to prevent the disproportionate influence of those outliers. The GLS approach is identical to least-squares ANCOVA, except it introduces a residual variance that linearly increases with the observed control catchment annual maximum flow. GLS allows for a looser relationship between the control and treatment catchments during higher flow events and thus reduces the influence of high-flow outliers. ANCOVA analyses were conducted using R Version 3.6.2 (R Core Team, 2019) using the base ordinary least squares regression function (lm) and the generalized least squares regression function (gls) found in the nlme package (Pinheiro et al., 2019).

Table 1
Stationary blocks used in the analysis of peak flows in Sebestyen et al., 2011b.

S4/S5 Experiment		S2/S6 Experiment	
1962–1970	Pretreatment	1976–1980	Pretreatment
1971–1979	Open/Young Forest	1981–1989	Open/Young Forest
1980–1989	Growing Forest	1990–1999	Growing Forest
1990–1999	Mid-Life Forest	2000–2016	Closed-Canopy
2000–2016	Mature Forest		

2.3.2. Coupling analysis

Control and treatment catchment annual maximum discharges may be generated due to different events and/or in different seasons, and thus become "decoupled" in any given year. Decoupling takes two forms: event decoupling and seasonal decoupling. Event decoupling occurs in a given year when the control and treatment catchments have their annual maximum flows due to different events in that year (i.e., different rainfall events, different snowmelt events, or one snowmelt and one rainfall). In any given year, the annual maximum flows were "event decoupled" if the date of annual maximum flow on the treatment and control catchments were greater than seven days apart based on an inspection of annual maximum flow hydrographs and MEF precipitation data (Supplemental Information, Section 2). The seven-day threshold was chosen from a visual inspection of discharge records which showed discharge generally decayed to approximately pre-event levels between sequential peaks greater than seven days apart (see Supplemental Information, Section 2). Further, the daily precipitation record at the MEF indicates the median number of consecutive days with measureable precipitation was 2 (±1.4 days, standard deviation) (Sebestyen et al., 2020c). These considerations indicate that most precipitation events at the MEF are relatively short such that stormflow peaks occurring greater than 7 days apart on control and treatment catchments were likely triggered by different events, rather than differences in runoff from the same precipitation event. A binary event decoupling series was created by denoting event decoupled years with 1 and coupled years with 0.

We define "seasonal decoupling" as the condition when annual maximum flows on the control and treatment catchments occurred in different seasons within a given water year. MEF field-collected snowpack data support the importance of seasonal snowpack melt in the months of March and April for streamflow (Sebestyen et al., 2020a). Snowpack depth and water equivalent are measured at snowcourses starting in February and continuing every two weeks until the course is free of snow (Sebestyen et al., 2020a). Based on previous work and MEF precipitation data, we defined a spring peak in March or April, versus a summer peak between June and November. The month of May was considered a transitional month because the dominant precipitation phase is rain, however catchments tend to be high in saturation from snowmelt. A year was considered seasonally decoupled if one catchment had a March/April spring peak and the other catchment had a June to November summer peak. The binary seasonal decoupling series was created by denoting seasonally decoupled years with 1 and coupled years with 0.

To determine the probability of event and seasonal decoupling, we analyzed the observed binary decoupling series within pretreatment and multiple post-treatment time periods (Table 1) for both catchment experiments using logistic regression with parameters estimated through a Bayesian technique, shown in Equation 1.

$$\begin{aligned} Decoupled_i &\sim Bernoulli(p_g) \\ logit(p_g) &= b_g \end{aligned} \quad (1)$$

The probability of decoupling in year i in the time series block g (for each block g in Table 1) is p_g, which corresponds to a log-odds (or "logit", $\log(p_g/(1-p_g))$ of b_g. We use Bayesian-based Markov chain Monte Carlo (McMC) sampling (Gelman and Rubin, 1992) to generate a full probabilistic estimate for b_g conditioned on the observed binary series. We implemented the JAGS ("Just Another Gibbs Sampler") algorithm (Plummer, 2003) with the "R2jags" package (Su and Yajima, 2015) in R for McMC sampling. To test whether the decoupling probability p_g increases due to treatment, we determined the proportion of McMC samples that predict the decoupling probability for each post treatment time period (p_g) to be less than in the pretreatment period ($p_{pretreatment}$).

This proportion then corresponds to a "Type-S error" probability, or, the probability of erroneously concluding that $p_{pretreatment} < p_g$, where g ≠ pretreatment (Gelman et al., 2012; Gelman and Tuerlinckx, 2000). We can claim with (1 – pvalue)*100% confidence that the probability of control and treatment annual maximum flows occurring due to independent precipitation events or within different seasons in any given year increases due to treatment.

2.3.3. Flood-Frequency analysis

2.3.3.1. Choice of distribution. We used a Bayesian approach to fit observed annual maximum flow to the generalized extreme value (GEV) distribution, which has the following cumulative distribution function for the annual maximum discharge Q:

$$F(Q;\mu,\sigma,\xi) = \begin{cases} exp\left[-\left(1+\xi\left(\frac{Q-\mu}{\sigma}\right)\right)^{-1/\xi}\right], \xi \neq 0 \\ exp\left[-exp\left(-\left(\frac{Q-\mu}{\sigma}\right)\right)\right], \xi = 0 \end{cases} \quad (2)$$

After estimating parameters for Equation 2 using McMC, we sampled across different probabilities of occurrence to estimate annual maximum flow magnitudes at given return intervals, as in flood-frequency analysis (Gumbel, 1941). To capture nonstationarities, we developed time-evolving estimates of the location parameter μ (possible range of -∞ to ∞) and scale parameter σ (possible range of 0 to ∞). The shape parameter ξ (possible range of -∞ to ∞; however, the mean of the GEV → ∞ for ξ > 1) was held stationary due to the narrow range of ξ for which the GEV has a finite (thus physically plausible) mean. Our narrow range of ξ is consistent with other nonstationary flood-frequency analyses of paired catchment data (Yu and Alila, 2019). The resulting nonstationary GEV models applied to treatment and control catchments were then compared to detect treatment effects. More details on the selection and fit of GEV models are provided in the Supplemental Information (Fig. S1; including comparison results in Table S1).

2.3.3.2. Nonstationary parameter evolution. Nonstationary models require a defined form for parameter trends through time (Koutsoyiannis, 2011). To capture the alterations in upland forest cover in each of the two catchment experiments, we developed the temporal trend structures shown in Fig. 2 for the location and scale parameters (μ and σ) of the GEV distribution (Eq. (2)). The shape parameter (ξ) for the GEV distribution was held constant in time. Bayesian-based McMC sampling with the JAGS algorithm (Plummer, 2003) in the "R2jags" package (Su and Yajima, 2015) was implemented to estimate the needed hyper-parameters associated with μ and σ (and stationary ξ) models for the treatment and control watersheds using annual maximum flow observations (Fig. 2). The advantage of the Bayesian-based McMC approach is that it provides full "posterior" distribution estimates of the hyper-parameters, which incorporate in a statistically consistent way both constraints from the maximum flow observations as well as "prior" information (represented by prior probability distributions) about the hyper-parameters based on previous knowledge of the site and/or physical limitations (Smith and Roberts, 1993).

For S4/S5, the general structure of the nonstationary model for the GEV location and scale parameters (μ and σ) is assumed to follow a piecewise linear trend with time *t* (water year) during its regrowth period, based on breakpoints at the year of complete upland clearcut (t = 1972; 71% of watershed area harvested) and the recovery year (t_r) (Equation 3; Fig. 2a):

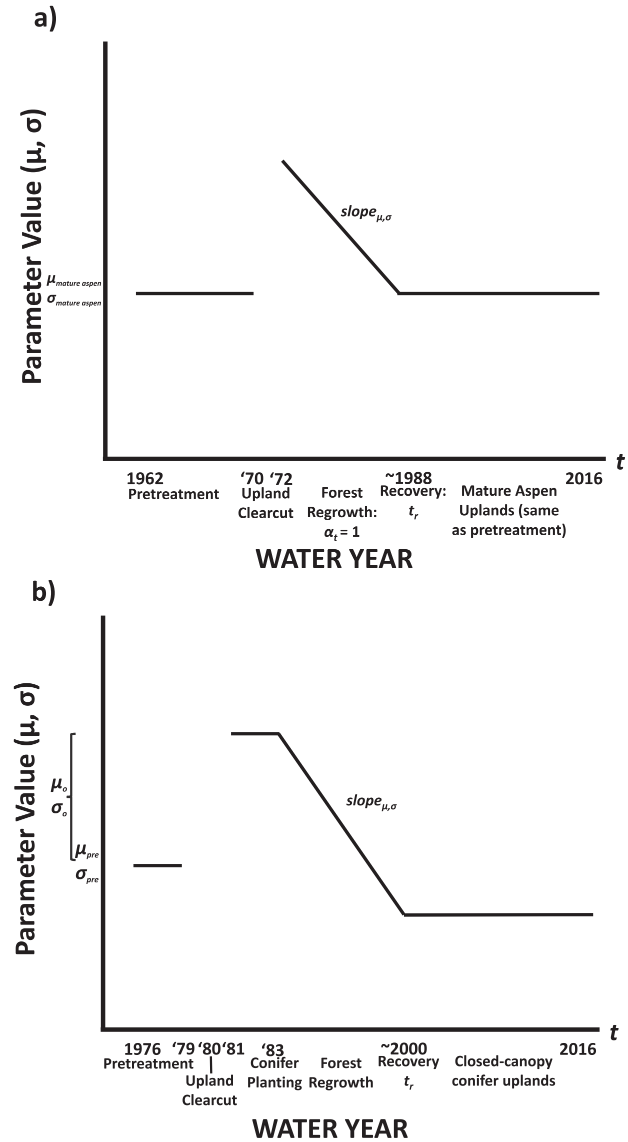

Fig. 2. Nonstationary parameter evolution for the location and scale parameters, as modeled for the a) S4/S5 and b) S2/S6 experiments.

$$\begin{bmatrix} \mu(t) \\ \sigma(t) \\ \xi \end{bmatrix} = \begin{bmatrix} \mu_{matureaspen} + slope_\mu * (t - t_r) * \alpha_t \\ \sigma_{matureaspen} + slope_\sigma * (t - t_r) * \alpha_t \\ \xi \end{bmatrix} \quad (3)$$

$$\text{with} \quad \alpha_t = \begin{cases} 1 & \text{if } t \geq 1972 \text{ and } t < t_r \\ 0 & otherwise \end{cases}$$

We omitted 1971 from the analysis to avoid complications due to only partial upland clearcut conditions in that year. Bayesian-based McMC sampling was used to estimate ξ, t_r, $\mu_{matureaspen}$, $slope_\mu$, $\sigma_{matureaspen}$, and $slope_\sigma$ (Equation 3).

In previous assessments, recovery of peak flow characteristics for post-harvest aspen regeneration were found ~ 16 years after harvest (Sebestyen et al., 2011b; Verry, 2004, 1986; Verry et al., 1983). The 16-year recovery was based on considerations of aspen growth rate and crown formation (Perala and Verry, 2011), and previous discharge regression analyses at MEF showing that snowmelt-associated peaks remained above pre-harvest conditions for 15 years after harvest (Verry, 2004). Thus, our McMC calculations included a moderately informative prior distribution for the t_r, consisting of a discrete approximation of a normal distribution centered on the 16th year after harvesting (1988 in

S4) with a standard deviation of two years. A discrete approximation of the normal distribution was used to represent the recovery year to an integer year post-harvest, because the time step of the input data was one year. The shape parameter was constrained between 0 and 1, and all other prior distributions were set as noninformative. For more information on choice of prior distributions, see the Supplemental information, Section 3.

For the S2/S6 experiment, we modeled the location and scale parameters (μ and σ) of the GEV distribution with a piecewise linear model with time t (calendar year) that includes stationary values in the pretreatment time period (1976–1979), different stationary values during upland grazing (1981–1983), decreasing values through time after conifer planting (1984 until the recovery year t_r), and a third set of stationary values associated with conifer canopy closure after the recovery year, ~ 2000. (Equation 4; Fig. 2b):

$$\begin{bmatrix} \mu(t) \\ \sigma(t) \\ \xi \end{bmatrix} = \begin{bmatrix} \mu_{pre} + \mu_o{}^*a_t + slope_\mu{}^*(t-1984)^*b_t + slope_\mu{}^*(t_r-1984)^*c_t \\ \sigma_{pre} + \sigma_o{}^*a_t + slope_\sigma{}^*(t-1984)^*b_t + slope_\sigma{}^*(t_r-1984)^*c_t \\ \xi \end{bmatrix}$$
and
$$a_t = 1 \; for \; t \geq 1981, \; else \; 0$$
$$b_t = 1 \; for \; t > 1984, \; else \; 0$$
$$c_t = 1 \; for \; t \geq t_r, \; else \; 0 \tag{4}$$

We omitted 1980 from the analysis to avoid complications due to only partial upland clearcut conditions. Analogous to the S4/S5 experiment, we applied Bayesian-based McMC sampling to estimate the stationary shape (ξ) parameter, μ_{pre} and σ_{pre}, the slope of the location and scale parameters as the conifer forest grows ($slope_\mu$ and $slope_\sigma$), the effect of harvesting on the location and scale parameters (μ_o and σ_o), and the recovery year t_r (Equation 4).

We included a moderately informative prior distribution in the McMC calculations that is a discrete approximation of a normal distribution centered on 17 years after harvesting (2000 in S6), when the canopy closed (Sebestyen et al. 2011b), and a two-year standard deviation. For information on the noninformative prior distributions used for all other hyperparameters, see the Supplemental Information, Section 3.

2.3.3.3. Treatment effects. A linear regression equation was fit to pretreatment observations for each catchment pair in order to parameterize the relationship between annual maximum series in the treatment and control catchments while isolating the non-climatic biophysical differences between them, similar to Alila et al. (2009). This regression was the same ANCOVA form explained in Section 2.3.1, fit only in the pretreatment time period. We then applied the regression equation to the post-treatment time (C_{postT}) control catchment annual maximum flow to predict the expected annual maximum flow series in the treatment catchment had the treatment not occurred ($\widetilde{T}_{postT}$):

$$\widetilde{T}_{postT,i} = b_0 + b_1{}^*C_{postT,i} \tag{5}$$

where i is the time index during the post-treatment period, and b_0 and b_1 are coefficients estimated using pre-treatment observations. The full time series of annual maximum flow in the treatment catchment had the treatment not occurred, from the pre-treatment time through the post-treatment time period, is then represented by the concatenation of the pre-treatment observations and $\widetilde{T}_{postT,i}$. The concatenated time series is represented as $\widetilde{T}_i$.

Parameters b_0 and b_1 in equation 5 were determined using Bayesian-based McMC sampling (with the JAGS package in R) to acknowledge uncertainty in the coefficient estimates and retain consistency with the parameter probabilistic treatment. Seventy McMC samples of b_0 and b_1, representing 70 possible parameter sets given the observed maximum flows, were used to generate 70 possible series of $\widetilde{T}_i$ as an uncertainty range of the expected annual maximum flow on the treatment catchment, had treatment not occurred. Each of the 70 expected annual maximum flow series was then used as data to fit nonstationary GEV distributions used for treatment catchments (Equations 3 and 4). Expected results for the no-treatment cases were compared to the observed with-treatment case in order to draw inferences about harvest and regrowth effects, hence utilizing the catchment pairing to control for climate covariates. We note that the number of series examined was limited to 70 due to the computational burden of fitting each nonstationary GEV model to the data using McMC. When Equation 5 predicted negative annual maximum discharges for S6, 0 was used as the predicted annual maximum discharge. This circumstance happened on average 2–3 years within the 40-year-long expected annual maximum flow series for S6.

Random samples from the posterior densities for location (μ), scale (σ), and shape (ξ) parameters in the time periods of interest were taken to construct 100,000 cumulative distribution functions for the expected and observed GEV model fits (Eqns. 3 and 4) in order to evaluate differences in annual maximum flow across return intervals. Effects of harvest were inferred for each treatment pair through comparison of the distribution functions during specific time intervals. Observed (with treatment) and expected (no treatment) models (Eqn. 3) were compared across S4/S5 for the year 1972, when the treatment effect of harvesting was largest. For the S2/S6 experiment, a similar comparison was constructed for the time interval after clearcutting and during grazing (1981–1983). Further, conifer conversion effect was inferred through observed and expected model comparison between 2000 and 2016, after conifer canopy closure on the treatment catchment. Expected (control) and observed (treatment) cumulative distribution functions for return intervals N = 1.5 to 50 years were compared through finding the proportion of the treatment cumulative distribution functions that are greater or less than the expected cumulative distribution functions as a "Type S" error probability for each return interval (Gelman and Tuerlinckx 2000), which represents the chance of incorrectly concluding that the N-year peak flow has increased or decreased when they have in fact decreased (increased) or stayed the same.

3. Results

3.1. ANCOVA, GLS

ANCOVA results for the S4/S5 experiment did not show any post-treatment slopes or intercepts different than the pretreatment time period (all $p > 0.10$), including in the decade immediately following upland harvesting ($p = 0.47$ for intercept, $p = 0.26$ for slope) (Fig. 3a). In general, no slopes or intercepts were different from each other according to a Type-II ANOVA (Langsrud, 2003), with slopes $p = 0.36$ and intercepts $p = 0.38$ (Fig. 3a). A GLS regression was used in addition to the ordinary least squares (OLS) regression due to detected outliers in the OLS. GLS results concurred with OLS, indicating no statistical differences of slopes or intercepts in the post-treatment time periods compared to pretreatment (all $p > 0.25$).

ANCOVA results for the S2/S6 experiment showed a significant decrease in regression slopes for the "Growing Forest" and "Closed Canopy" conifer conditions compared to the pretreatment time period ($p = 0.001$ and 0.004, respectively), but no significant difference between the open/young forest and pretreatment time periods (Fig. 3b). Our ANCOVA results on S2/S6 concur with previous work that found decreases in peak flows after conversion to conifer species (Sebestyen et al., 2011b). However, residuals versus leverage plots revealed outliers, and regression slopes corresponding to the "Growing Forest" and "Closed Canopy" time periods were no longer significantly different from pretreatment ($p = 0.15$ and $p = 0.46$, respectively) when using GLS. The GLS regression was a better fit according to diagnostic plots such as residuals versus fitted and residuals versus leverage plots.

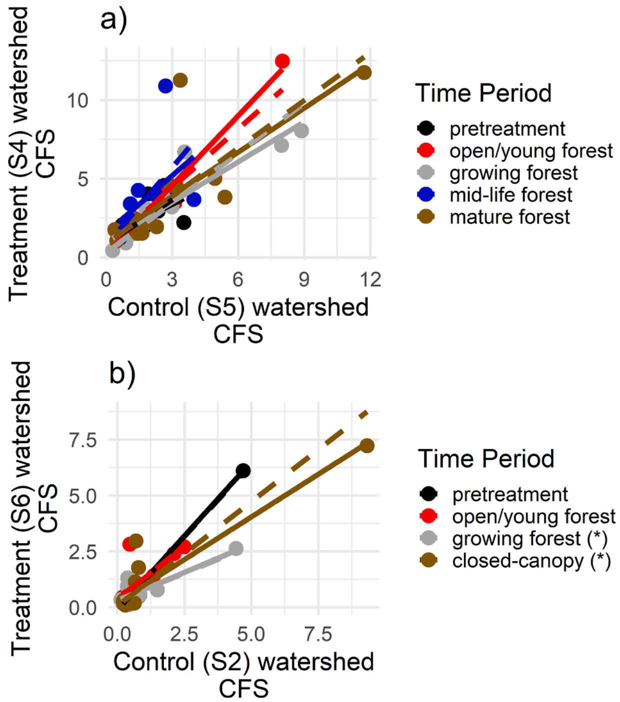

Fig. 3. a) S4 experiment ANCOVA (solid line) and GLS (dashed line) analyses of annual maximum flows. No significant differences in slopes or intercepts for the ANCOVA regression, nor for the GLS regression. b) S6 experiment ANCOVA (solid line) and GLS (dashed line) analyses of annual maximum flows. Significantly different regression slopes ($p < 0.10$) according to ANCOVA are indicated by (*). There were no significant treatment effects according to the GLS analysis.

3.2. Coupling analysis

Harvest induced a greater probability of the annual maximum flow being due to different independent events (>7 days apart) on control and treatment catchments. Event decoupling is captured in the logistic regression results for S4/S5 and S2/S6 (Table 2; Fig. 4a and 4c). The same general trend is seen in both experiments, but we are only > 90% confident for the S4/S5 experiment. Furthermore, harvest induced a greater probability of seasonal decoupling in the S4/S5 experiment (Fig. 4b, Table 2). Seasonal decoupling occurred in twelve years of the fifty-four year S4/S5 record, with 11 of those years occurring after harvest. In ten of these post-harvest seasonal decoupling years, S4 had a summer peak, and S5 had a spring peak. Seasonal decoupling was rarely observed for the S2/S6 experiment (only two times in the 41-year record), so seasonal decoupling analysis was not implemented for the S2/S6 experiment.

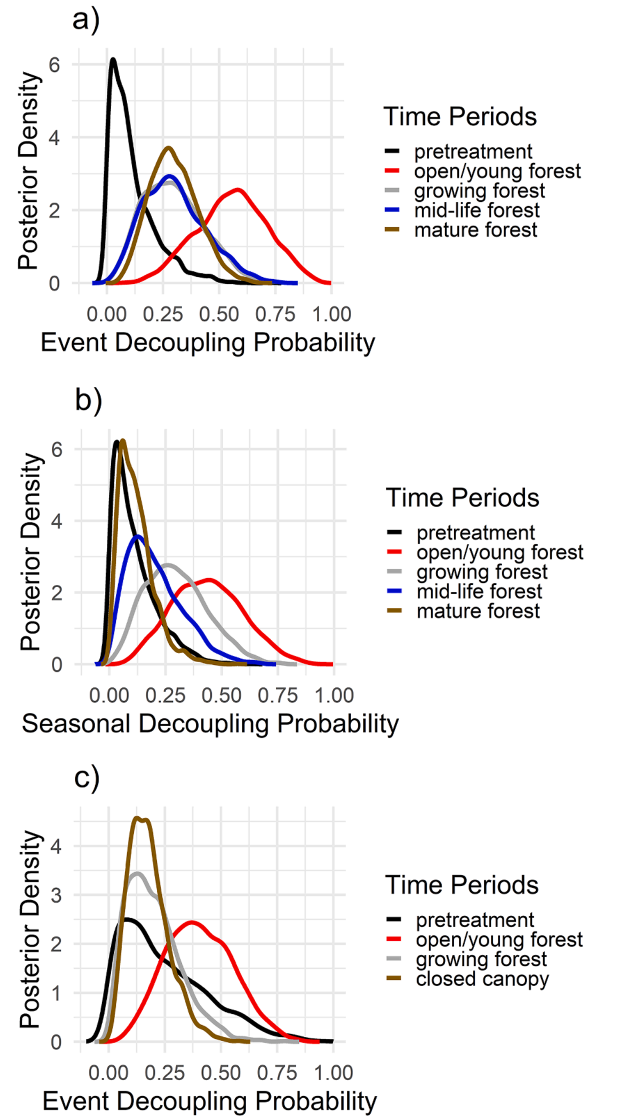

Fig. 4. a) Event decoupling probabilities for the S4/S5 experiment, from the logistic regression (Eq. (1)). b) Seasonal decoupling probabilities for the S4/S5 experiment. c) Event decoupling probabilities for the S2/S6 experiment. P-values for differences in all decoupling probabilities with respect to pretreatment are in Table 2.

Table 2
Probability of decoupling in the stationary time periods defined in Table 1 for both experiments. In parentheses are the probabilities of incorrectly stating that the decoupling probability in that time period is greater than the decoupling probability in the pretreatment time period. * = Type-S error rate less than 0.10.

Time Period	S4/S5: p_g ($P[p_g \leq p_{pretrt}]$)		S2/S6: p_g ($P[p_g \leq p_{pretrt}]$)
	p_g = probability that AM flow on S4 and S5 occur > 7 days apart	p_g = probability that AM flow on S4 and S5 occur in different seasons	p_g = probability that AM flow on S2 and S6 occur > 7 days apart
Pretreatment	0.111	0.110	0.249
Open/Young Forest	0.556 (0.013)*	0.436 (0.035)*	0.405 (0.245)
Growing Forest	0.296 (0.120)	0.297 (0.128)	0.199 (0.546)
Mid-Life Forest	0.300 (0.123)	0.202 (0.255)	NA
Mature Forest	0.297 (0.103)	0.119 (0.431)	NA
Closed Canopy	NA	NA	0.178 (0.563)

3.3. Flood frequency analysis

3.3.1. S4/S5 experiment flood frequency analysis

In the pretreatment calibration regression (Equation 5), expected annual maximum discharge on S4 given the annual maximum flow on S5 had an intercept estimate of 0.695 (95% credible interval: −0.108 to 1.475) and a slope estimate of 1.288 (95% credible interval: −0.311 to 1.752) (overall $r^2 = 0.32$). Expected values and their credible intervals, and the raw treatment annual maximum series, are shown in Fig. 5a.

The N-year discharge increased in the first year after cutting compared to the expected discharge if the harvest had not occurred (Fig. 6). We have between 80 and 85% confidence in discharge increases for return intervals >10 years (Type-S $p < 0.20$), whereas we have only 30% confidence for an increase in the 1.5-year discharge (Fig. 6b). Our

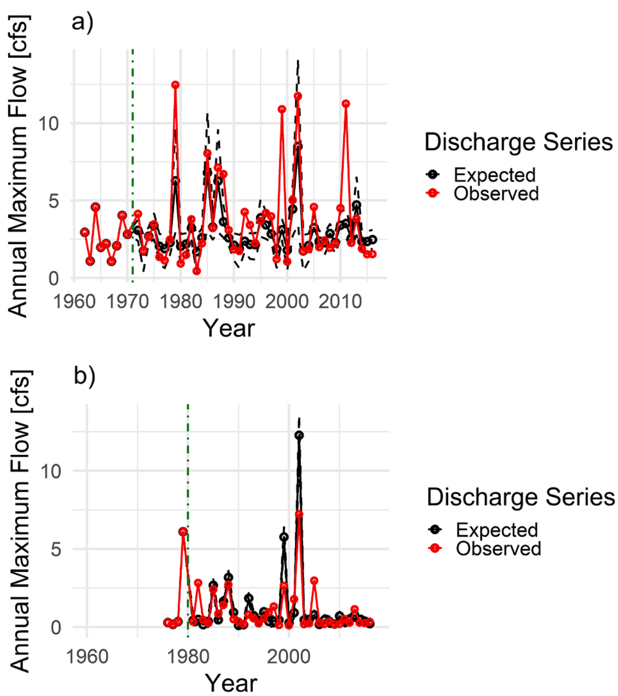

Fig. 5. Expected and observed series' for a) S4 and b) S5, plus uncertainty of the expected series shown as a 90% credible interval based on 70 expected series fit from the calibration regressions (dashed lines). The year of upland clearcut harvesting for each catchment is represented by a dashed green line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

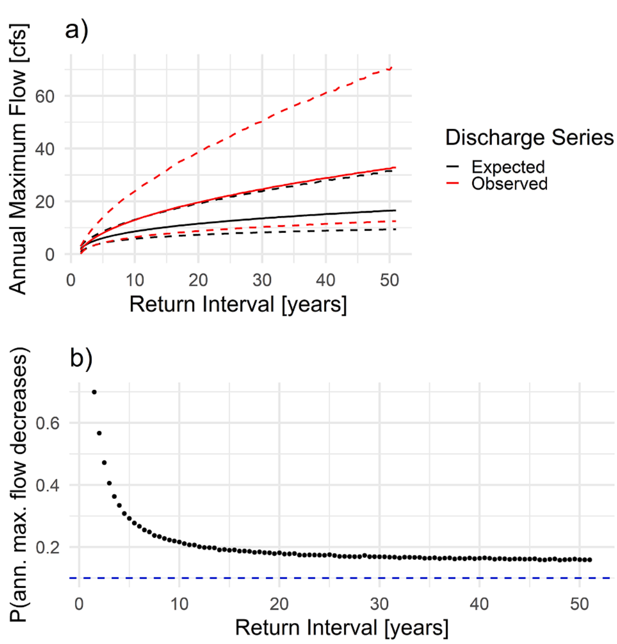

Fig. 6. a) S4 cumulative distribution for the first year after upland clearcut, compared to expected conditions derived from the calibration regression. Pictured with 90% credible intervals. b) Probability that we are incorrectly concluding that the N-year annual maximum flow is increasing, across return intervals up to N = 50. A dashed line is shown at p = 0.10 for reference.

results indicate a greater treatment effect for larger discharges. Averaged across all return intervals, we have 80% confidence that annual maximum flows increased, with increasing confidence for increases at larger return intervals ($p = 0.16$ for the 50-year annual maximum flow). Further, the increases in annual maximum flows for each return interval stem largely from an increase in the scale (σ; $p = 0.27$) and shape (ξ; $p = 0.19$) parameters due to harvesting (Table 3). The shape parameter controls the "heavy-tailedness" of the distribution of annual maximum flows. For the entire S4 treatment record, the shape parameter equaled 0.40, versus 0.20 on the entire record of the expected conditions had harvest not occurred, indicating that treatment induced greater probability of very high annual maximum flows. Even after recovery of the location and scale parameters to pretreatment values on the treatment catchment, the larger shape parameter for the treatment catchment causes larger annual maximum flows than what was expected in the absence of treatment, with ~ 88% confidence from the 10- to the 50-year return interval annual maximum flow. Our results indicate that post-harvest changes in annual maximum flows may be driven largely by changes to the variability of annual maximum flows, as the location parameter (related to mean flow) potentially even decreased ($p = 0.33$). We are more confident in increases to the scale parameter (related to variance; $p = 0.26$), and the shape parameter, related to the probability and variance of large annual maximum flows ($p = 0.19$). For expected and observed posterior densities for location, scale, and shape, see Fig. S2.

The recovery year for location and scale parameters is relatively unchanged from our moderately informative prior distribution. Based on model results, we are 90% confident that recovery of the location and scale parameters to pretreatment conditions occurs between 12 and 18 years after harvest. The shape parameter was held stationary within each model fit (Eqn. 3). However, we obtained different shape values when GEV models were fit for expected versus observed time series ($p = 0.19$; Eqn. 3). The difference in shape parameter indicates that the shape parameter may have increased with forest harvesting, and was therefore in truth nonstationary. Thus, we do not know when overall recovery happened because we only assumed location and scale were nonstationary within each individual model, and we have no means of testing when the shape parameter returned to what we would expect in the absence of treatment. Our results indicate high confidence (~88–89%) that annual maximum peak flows were higher on the treatment catchment than what would be expected in the absence of treatment even after recovery of the location and scale parameters. There is no way to investigate when recovery of the shape parameter occurred, if at all, within our modeling framework.

3.3.2. S2/S6 flood frequency analysis

Post-treatment annual maximum discharges on the control catchment were adjusted according to the calibration regression (Fig. 5b), which had an intercept of −0.120 (95% credible interval: −0.649 to 0.409) and a slope of 1.324 (95% credible interval: 1.106 to 1.550). For reference, overall $r^2 = 0.99$.

We did not find an increase in the annual maximum flow immediately following clearcut when the treatment effect should be highest based on Equation 4 (Fig. 7a, 7b). Across the cumulative distribution function, all p-values were between 0.35 and 0.45 up to the 50-year

Table 3
Type S error probabilities for alterations in the location, scale, and shape parameters in observed versus expected models of annual maximum flow for both harvesting experiments. Most noteworthy is the shape (ξ) parameter for the S4/S5 experiment, for which we have the most confidence of an increase (i.e., the smallest p-value).

Parameter	S4/S5	S2/S6	
	P[harvest year 1 parameter ≤ expected parameter]	P[harvest year 1 parameter ≤ expected parameter]	P[mature conifer parameter ≥ expected aspen parameter]
μ	0.677	0.357	0.346
σ	0.263	0.514	0.252
ξ	0.187	0.357	0.643

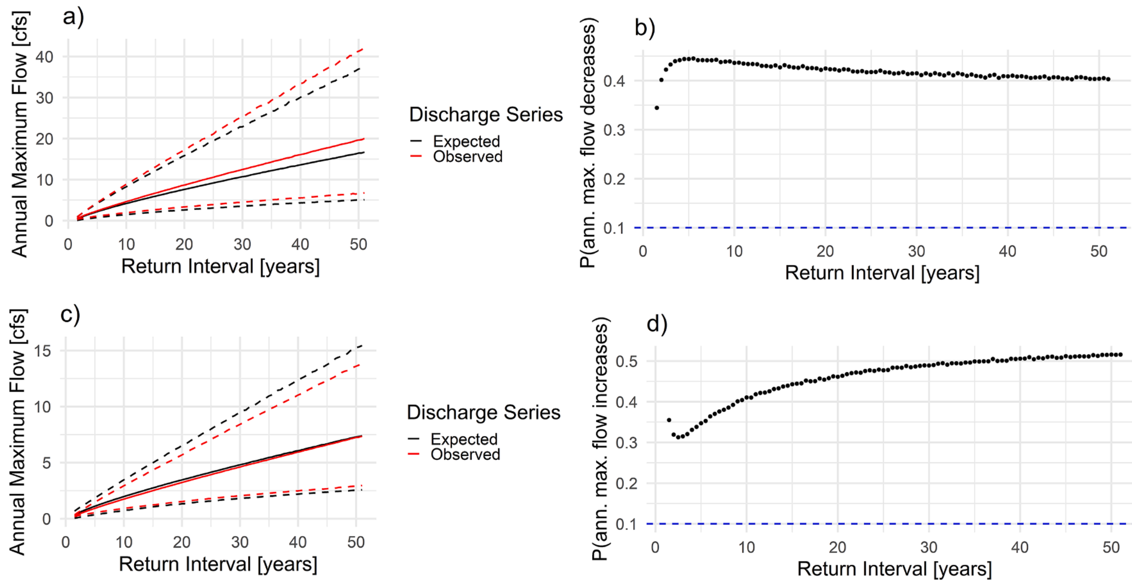

Fig. 7. a) S6 cumulative distribution for immediately after upland forest harvesting, compared to the expected values derived from the control watershed adjusted to the calibration regression. b) S6 probabilities that we are incorrect in stating that the N-year annual maximum flow decreased due to upland harvesting c) S6 cumulative distribution for when S6 had a closed-canopy conifer forest, compared to the expected values derived from the control watershed d) S6 probabilities that we are incorrect in stating that the N-year annual maximum flow increased due to conversion from deciduous to coniferous species.

return interval (Fig. 7b). Location, scale, and shape parameters stayed the same (Table 3; Fig. S3).

There was also no effect of conifer conversion on the annual maximum flow for the S6 catchment (Fig. 7c, 7d). Here, testing for a hypothesized decrease in annual maximum flows, we are not confident in any of the modeled decreases in the annual maximum flow across return intervals up to the 50-year event, with p being about 0.35 under the 10-year return interval discharge event, but going up to 0.5 for larger discharge events, indicating that we have no more confidence that flows decreased than we do that flows increased due to conifer conversion (Fig. 7d). Location, scale, and shape parameters stayed the same; unlike for S4/S5, we do not have evidence that the shape parameter was nonstationary (Table 3).

The recovery of the location and scale parameters to a new stationary state after conifer planting was heavily influenced by our semi-informative prior which centered recovery around canopy closure, ~17 years after conifer planting. However, because no effects of upland clearcutting or canopy conversion were found, recovery year has little meaning.

4. Discussion

4.1. ANCOVA, GLS

Based on the ANCOVA and GLS regression analyses, no significant increases were found for annual maximum flows in the time periods directly after upland clearcut harvest when the uplands were in open/young forest conditions, compared to the pretreatment time period (Fig. 3). This finding contrasts with previous ANCOVA results that have found significant effects of harvesting on snowmelt and rainfall-caused peaks for the S4 experiment (Verry et al., 1983). Several factors may explain this discrepancy between past studies and our results, including: (1) the longer time period of record in our analysis, (2) our use of both S4N and S4S gauges to compute total annual maximum discharge on S4, and (3) our examination of the annual maximum discharge including both rainfall and snowmelt events. Further, GLS was not previously utilized.

4.2. Coupling analysis

We observed significant event and seasonal decoupling on the S4/S5 pair in a given year. Seasonal decoupling was generally associated with streamflow generating processes that differed for control and treatment catchments. Summer annual maximum flows result directly from rainfall events. In past MEF research, March and April high flows were generally associated with snowmelt (Sebestyen et al., 2011b; Verry et al., 1983). With the two-week resolution of snowpack data, however, it remains unclear if individual March-April peaks are due to snowmelt, rain-on-snow, or rain on a snowmelt-saturated catchment. Two week data on snow depth and snow-water-equivalent do, however, support our assertions that every year does include measureable snowpack, and snowpack melt largely occurs in March and April. Hence, we can generalize that any annual maximum peak occurring in March or April is influenced by snowmelt, but possibly triggered by rain-on-snow or also influenced somehow by rain, whereas any annual maximum flow occurring after May 30 is due to summer rainfall. Thus, the seasonal decoupling observed in S4/S5 likely reflects a regime shift in the annual peak streamflow generating process after harvesting for S4, with peaks for S4 shifting from a snowmelt-influenced to a rainfall-dominated process.

Decoupling likely accounts for differences between our results and those of previous studies. Annual maximum discharge from the catchments is not the same as individual examinations of spring and summer peaks that have been analyzed previously (Verry et al., 1983; Sebestyen et al., 2011b). In our analysis, peak discharge on control and treatment watersheds was decoupled at times, hence not chronologically paired. Previous ANCOVA analysis of paired-catchment data, at the MEF and elsewhere, relied on chronological pairing of events, such that peaks generated during the same event were compared on control and treatment catchments. The significant event- and seasonal-decoupling in the S4/S5 pair calls into question the appropriateness of chronological pairing of annual maximum flows. As early as the pre-treatment period, catchment pairs had a non-zero probability of event-decoupling (Table 2). Event-decoupling was linked to open vs forest condition with a post-harvest increase in both the S4/S5 and S6/S2 pairs, followed by a decrease towards preharvest levels as forest cover recovered (Fig. 4a, 4c; Table 2). Event decoupling was not related to the cover type of the regrowing forest. The incidence of event decoupling for S6 with closed-canopy conifer uplands, compared to deciduous cover on S2 during the same time period, is no greater than in the pretreatment time period when both S2 and S6 had deciduous cover ($p = 0.563$). Thus, canopy conversion from deciduous to coniferous does not cause event decoupling.

The S4/S5 experiment showed evidence of seasonal-decoupling where S4 was more likely to have a summer annual maximum flow while S5 was more likely to be a spring annual maximum flow (Fig. S4). In the first decade following harvest, seven of ten of the S4 annual maximum flows occurred in the summer, while only three of ten occurred in the summer on S5. Seasonal decoupling was not observed significantly on S2/S6: only two of the forty years of record had seasonal decoupling, which was too few instances upon which to fit a logistic regression. The seasonal decoupling of S4/S5 is possibly due to internal subtle differences between the treatment (S4) and control (S5) catchments that were exacerbated by harvest. For example, S5 has 5 small satellite wetlands within the upland rim that act as distributed catchment storage in addition to the central peatland; S4 does not have these upland wetlands. This difference may cause S4 to be more responsive to summer rainfall events than S5. Harvest may have exacerbated these preexisting differences in catchment processing of flood peaks.

4.3. Flood-frequency analysis

The nonstationary flood-frequency analysis, based on the immediate post-harvest year, indicated increases in annual maximum flows up to the 50-year return interval within the S4 catchment in the first year after harvest (Fig. 6). The effect size of the harvest increased with increasing peak flow return interval (Fig. 6). Broadly, the increases in annual maximum flows are consistent with previous results that showed increases in snowmelt-caused peak discharges for 15 years and rainfall-caused peak discharges for 7 years after harvesting for the S4 experiment (Verry 2004; Sebesteyen et al. 2011b). Further, when Verry et al. (1983) compared stationary flood-frequency distributions for pre- versus post-treatment rainfall peaks for S4 versus S5, they found the 10-year rainfall-caused annual maximum discharge increased more than the 2-year discharge in the decade after harvesting. The seasonal-decoupling between S4 and S5, discussed in Section 4.3, further elucidates the findings of Verry et al. (1983). When only rainfall peaks were compared, the annual maximum from S5, which tended to be in the spring and associated with snowmelt, was not included.

We found an increasing effect size of forest harvesting for events at increasing event return intervals, which is similar to previous MEF work (Verry et al. 1983). Our results indicate that in the first year after harvest, the 2-year discharge was effectively the same on the harvested S4 catchment versus expected conditions: 2.98 cfs versus 3.08 cfs, while the 50-year discharge approximately doubles from 16.5 cfs to 32.5 cfs. Despite increasing uncertainty at larger return intervals, our confidence in the presence of an effect on peak flows remains relatively constant from the 10-year to the 50-year discharge, at about 80–85% (Fig. 6b). Thus, our results do not match the paradigm of decreasing effect size with increasing return interval that has been largely derived from paired-catchment studies in western North American catchments with very different landscapes and ecosystems (e.g., Thomas & Megahan, 1998; Buttle, 2011). One meta-analysis of paired catchments using both ANCOVA-based and flood-frequency-based methods supported a generally decreasing effect size of harvest with increasing discharge return interval, but one catchment pair in the analysis showed increasing effect size with increasing return interval (Bathurst et al., 2020). This finding may have been due to road construction in this catchment pair (Bathurst et al., 2020). Although S4 does have proportionally more forest roads than S5, forest roads and major skid trails occupied less than 2% of the catchment area after harvesting, likely not explaining the peak flow increases (Verry et al., 1983). For the first year after harvest (Fig. 6), we have between 80 and 86% confidence in the flow increases due to forest harvesting. However, our confidence in flow increases is ~88% for large discharge events after recovery of the location and scale parameters due to an increased shape parameter. Further, our results of increasing effect size with increasing return interval is consistent with recent flood-frequency analyses of paired catchment data in western North America (Alila et al., 2009; Green & Alila, 2012; Yu & Alila, 2019), and is consistent with the flood-frequency analysis of rainfall peaks at the MEF done by Verry et al. (1983).

The effect of harvest on large return-interval events was heavily influenced by the difference in the GEV model shape parameter on the treatment watershed. This shape parameter was held constant within each GEV model of annual maximum flow. However, when fit to the annual maximum flow series on treated S4, it was larger than when fit to the 70 series of expected conditions had treatment not occurred. The difference in shape parameter value for expected versus observed treatment models caused the increases in annual maximum peak flows to be present in the first year after harvesting. Further, increases in the annual maximum flow after the recovery of location and scale parameters on S4 are evident with high confidence (~88%). It should be noted that three of the four largest discharge events on S4 (1999, 2002, and 2011) occurred after the hypothesized 16-year recovery estimate for location and scale parameters, and were all in response to rainfall in June or July. Of these three peaks that occurred greater than 16 years after harvest, two were above the 90% credible interval for the expected value based on the pretreatment calibration (1999 and 2011: Fig. 5). Shape parameters, related to the tail of the annual maximum flow distribution, may thus be nonstationary and sensitive to forest harvesting, contrary to our initial assumption and that of others (e.g., Yu and Alila, 2019). However, fitting a nonstationary structure to all three parameters on 54 years of data would rapidly increase the ratio of parameters per data point, decreasing the usefulness of such models.

It is noteworthy that the flood-frequency analysis indicated high confidence for the effect of harvest for the S4/S5 experiment, but not for the S2/S6 experiment (where our confidence in annual maximum flow changes largely ranged between 55 and 65%). However, the investigation of the S2/S6 experiment is hindered by a short preharvest time period (4 years), which may have been inadequate for detecting a difference due to harvesting and conversion (Loftis et al., 2001; Sebestyen et al., 2011b). The annual maximum flow record for S2/S6 is shorter than for S4/S5, and more parameters were necessary to model S2/S6 compared to S4/S5 due to the grazing treatment on S6 prior to forest regrowth, and conversion to a different upland species (Fig. 2). This increased model complexity resulted in greater uncertainty for parameter estimates. We also consider the possibility that there truly was no effect of harvesting or conversion. Unlike for S4, we did not have strong evidence of seasonal decoupling for the S2/S6 experiment, which may indicate that S4 passed some hydrologic threshold due to disturbance while S6 did not. Nonlinear and threshold processes have been observed in northern headwater catchments, in which there are disproportional runoff responses to forcing inputs (Ali et al., 2015). Thresholds have been invoked for explaining increases in large peak flows and a sensitive upper tail for extreme value distributions (Alila et al., 2009). In one threshold-based investigation of a forest regenerating after disturbance, the threshold in the rainfall-stormflow relationship associated with rapid stormflow response did not change with forest regrowth, but the magnitude of stormflow response to precipitation decreased as the forest regrew (Wei et al., 2020). Watershed shape could also play a role in the divergent results between catchment pairs. S6 has an elongated shape, in contrast to S4, which is more circular and has two outlets (Fig. 1). More elongated watersheds tend to have relatively lower peak flows (Black, 1972). The peatland is also more centrally located within the catchment in S6 than S4 (Fig. 1), which could contribute to its ability to dampen peak flows out of the catchment. Finally, catchment differences within individual pairs may have contributed to the different flood-frequency findings in each experiment, as harvesting was found to exacerbate preexisting differences in these catchments via seasonal decoupling. For example, the control catchment S5 includes 5 small satellite wetlands within the uplands in addition to its central peatland. The preexisting differences in S4/S5 that may have contributed to our findings were not present with the S2/S6 pair.

No effect of canopy conversion from deciduous to coniferous upland species on the annual maximum flow was found in the flood-frequency analysis of the S2/S6 experiment. Uncertainty with low pretreatment sample size and number of parameters versus number of datapoints apply similarly to these results. Given this uncertainty, we note that the observed cumulative distribution function tended to decrease across all return intervals, which is the hypothesized direction of effect (i.e., conifer cover decreases flood peaks) (Fig. 7c, 7d). The ANCOVA analysis supported significant decreases in peak flows while the conifer forest grew and after canopy closure, but once the effects of high-flow outliers were dampened using GLS regression, the significance of these decreases disappeared. The annual maximum flow on both S2 and S6 was relatively low for the last decade of record (2006–2016), reflecting a prolonged period of relatively low flow (Fig. 5). This period of low flow coinciding roughly with conifer canopy closure did not offer much data on how larger or medium-sized peak flows respond to canopy conversion for a bulk of the closed-canopy conifer conditions. In our analysis, this caused larger flows that did occur in the closed-canopy period to exert stronger influence on the regression as outliers.

It is important to consider the limitations inherent in any investigation of flooding effects of land use. Limitations in our study include inherent factors in the paired-catchment design (e.g. past harvest treatment, pairings, etc.). Because of increasing complexity and different flood generation mechanisms at larger scales, it is unclear how these results apply for catchments at larger spatial scales (Blöschl, 2006; Rogger et al., 2017). Both experiments were partial clearcuts (~75% of catchment area) on only the upland portion of the catchments. The unharvested forested wetlands occupy significant catchment area and tend to attenuate flooding effects (Detenbeck et al., 2005). Significant portions of the harvests were also in the winter on frozen soils to reduce soil impacts and erosion. Furthermore, catchments were rapidly revegetated, leaving few years of "open" conditions on the uplands from which to derive inferences about the effects of open canopy conditions on peak flows. For example, the S4 clearcut was staggered between 1970 and 1972, and before the upland clearcut on S4 was complete in 1972, the upland portion previously cut had measureable regeneration, with 41,000 stems/ha that were 2 m tall (Verry et al., 1983). It may take several years for altered hydrology to equilibrate to a new stationary state after conversion to open conditions (Brown et al., 2005). Thus, there are potentially only several years of a transient state in between mature forest and recovery conditions for investigating these effects. The transient state of the effect of forest cover change is reflected in the large uncertainty bounds in nonstationary models.

Use of a calibration period provided another source of uncertainty. Because of the relatively short calibration periods in the paired catchment design (<10 years), we assumed that the relationship between the control and treatment catchments would have been stationary if the treatment had not occurred. Additionally, some "expected" values were extrapolated from the pretreatment calibration regression, despite uncertainty about how the calibration regressions may apply to these larger, extrapolated discharges. Changes in peak flow from the calibration period "expected" conditions are attributed solely to treatment – i.e., climate and biophysical differences between control and treatment catchments are encapsulated in a simple calibration equation. Thus, we attributed changes in flood-frequency to changes in forest cover as opposed to any other factor, such as rainfall frequency or changes in the control catchment as the forest aged. However, if the climate regime and/or weather patterns shift, the pretreatment calibration might not hold. The selection and appropriateness of a pretreatment calibration regression controlling for the biophysical differences in control and treatment catchments across the range of observed climatic conditions remains a large tenant in the paired catchment study design that has been debated for years, both criticized (Renne, 1967; Zégre et al., 2010) and defended (Hewlett et al., 1969; Neary, 2016). To incorporate our uncertainty about the calibration regressions, we sampled across 70 different slopes and intercepts fit with the calibration regressions. This incorporation of uncertainty in the calibration period is a considerable improvement to remedy issues of calibration uncertainty. However, some uncertainty about the calibration and its application for high flow regimes remains by virtue of the paired catchment study design.

5. Implications for analysis of paired-catchment studies

Several important ramifications for paired-catchment studies of the forest-peak flow relationship were revealed by this study. Ongoing disagreements about the effects of forest cover on flooding often stem from the statistical "lens" used to look at the question. We find it critical to articulate the benefits and limitations of various methods, clearly define the inference space for which the methods are valid, and relate statistical inference space to a physical inference space. First, methods that incorporate the probabilistic nature of peak flows are critical for finding differences due to landscape change, such as the nonstationary flood-frequency framework in our study. Increased use of probabilistic flood-frequency methods allows researchers to explicitly make inferences about effects of forest harvesting across multiple return intervals, which is not possible for traditional ANOVA/ANCOVA analyses. Next, ANCOVA is highly sensitive to high-flow outliers, which may not be adequately remedied through relaxing the constant residual variance assumption in the GLS analysis. Because our flood-frequency model of S4/S5 indicated that flow increases after forest harvesting can manifest as a change in occurrence probability for large flows (i.e., increase in shape parameter), it is critical to use a method that is robust to high-flow outliers and associated uncertainties, such as a probabilistic framework.

Our results underscore the importance of seasonal and event coupling in the assessment of forest harvest effects on peak flow. Peak coupling, also called chronological pairing, is assumed in ANCOVA, but not in flood-frequency analysis. The event and seasonal decoupling we found in the S4/S5 analysis highlights this important consideration in paired-catchment data analysis. The lack of chronological pairing for annual maximum flows likely explains the poor relationships between these discharges on control and treatment catchments in our ANCOVA/GLS analyses. The assumption of event coupling does not hold simply because the catchments are similar in size and physically adjacent. Methods robust to potential changes in pairing relationships should be further developed and utilized when analyzing paired-catchment data. Our coupling analysis illustrates the importance of analyzing the probabilistic distribution of peak flows as well as subtle changes in seasonality that can manifest as large changes at the tails of extreme value distributions.

Seasonal decoupling, particularly in catchments that receive both snow and rain, could further indicate differences in runoff generation and internal catchment processes between the paired catchments. In some catchments, for example mountain catchments that reliably have snowmelt-derived peaks every year (Alila et al., 2009), chronological pairing of annual maximum peaks may be possible. However, we found that harvesting may induce or enhance an alteration in the individual event and, in some cases, the event seasonality, and therefore flow generating process, that produces the annual maximum flow. We hypothesize the MEF is more likely than several western North American catchment pairs to experience seasonal decoupling due to its physiography and climate, and that other boreal-temperate transition catchments may be more sensitive to this decoupling response to harvest as the climate warms.

6. Conclusion

No significant effect of forest harvesting on annual maximum flows was found using ANCOVA/GLS or a nonstationary flood-frequency model of S2/S6, for which there were only 4 years of pretreatment data. However, increases in annual maximum flows across all return intervals were supported (confidence 80–85%) for the S4 clearcut and aspen regeneration experiment according to nonstationary flood-frequency analysis. ANCOVA-based methods were not adequate to fully investigate the effect of forest harvesting because annual maximum flows "decoupled" after harvest and tended to occur in different seasons, likely due to different streamflow generating processes (snowmelt-associated versus rainfall). Thus, even if flood magnitude and frequency may or may not be changed by harvesting, the precipitation event associated with the annual maximum flow may change. Our study demonstrates the value of using a probabilistic framework to investigate changes in peak flows. This method can support inferences about harvesting effects for specific return intervals, instead of being limited to differences in means (i.e., linear models). Increases in annual maximum flows after forest harvesting got larger with increasing return interval for the S4/S5 experiment, which indicates that if this effect truly occurred, it was not limited to small discharge events, as is the typical paradigm. More investigation is needed, particularly pairing computational, physically based models with Bayesian parameter estimation methods to minimize uncertainty about the form of nonstationarity introduced by harvesting, and to gain more physically based insight into this effect.

CRediT authorship contribution statement

Zachary P. McEachran: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Diana L. Karwan:** Conceptualization, Validation, Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition. **Stephen D. Sebestyen:** Validation, Data curation, Writing - review & editing. **Robert A. Slesak:** Conceptualization, Validation, Resources, Writing - review & editing, Project administration, Funding acquisition. **Gene-Hua Crystal Ng:** Methodology, Validation, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Minnesota Agricultural Experiment Station [grant number MN-42-080], the Minnesota Forest Resources Council, USDA Forest Service Northern Research Station, and the Peter F. Ffolliott Graduate Fellowship in Forest Resources. We acknowledge particularly the work of all MEF scientists and staff. We also acknowledge Genevieve Ali whose review greatly enhanced this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2021.126054.

References

Ali, G., Tetzlaff, D., Mcdonnell, J.J., Soulsby, C., Carey, S., Laudon, H., McGuire, K., Buttle, J., Seibert, J., Shanley, J., 2015. Comparison of threshold hydrologic response across northern catchments. Hydrol. Process. 29, 3575–3591. https://doi.org/10.1002/hyp.10527.

Alila, Y., Green, K.C., 2014. Reply to comment by Birkinshaw on "a paradigm shift in understanding and quantifying the effects of forest harvesting on floods in snow environments". Water Resour. Res. 50, 2769–2774. https://doi.org/10.1002/2013WR014198.

Alila, Y., Kuraś, P.K., Schnorbus, M., Hudson, R., 2009. Forests and floods: A new paradigm sheds light on age-old controversies. Water Resour. Res. 45 (W08416), 1–24. https://doi.org/10.1029/2008WR007207.

Andréassian, V., 2004. Waters and forests: From historical controversy to scientific debate. J. Hydrol. 291, 1–27. https://doi.org/10.1016/j.jhydrol.2003.12.015.

Angel, J., Swanston, C., Boustead, B.M., Conlon, K.C., Hall, K.R., Jorns, J.L., Kunkel, K.E., Lemos, M.C., Lofgren, B., Ontl, T.A., Posey, J., Stone, K., Takle, G., Todey, D., 2018. Midwest, in: Reidmiller, D.R., Avery, C.W., Easterling, D.R., Kunkel, K.E., Lewis, K.L. M., Maycock, T.K., Stewart, B.C. (Eds.), Impacts, Risks, and Adaptation in the United States: Fourth National Climate Assessment, Volume II2. U.S. Global Change Research Program, Washington, D.C., pp. 872–940. 10.7930/NCA4.2018.CH21.

Auerswald, K., Geist, J., 2018. Extent and causes of siltation in a headwater stream bed: catchment soil erosion is less important than internal stream processes. L. Degrad. Dev. 29, 737–748. https://doi.org/10.1002/ldr.2779.

Bathurst, J.C., Iroumé, A., Cisneros, F., Fallas, J., Iturraspe, R., Novillo, M.G., Urciuolo, A., de Bièvre, B., Borges, V.G., Coello, C., Cisneros, P., Gayoso, J., Miranda, M., Ramírez, M., 2011. Forest impact on floods due to extreme rainfall and snowmelt in four Latin American environments 1: Field data analysis. J. Hydrol. 400, 281–291. https://doi.org/10.1016/j.jhydrol.2010.11.044.

Bathurst, J.C., Fahey, B., Iroumé, A., Jones, J., 2020. Forests and floods: Using field evidence to reconcile analysis methods. Hydrol. Process. 34, 3295–3310. https://doi.org/10.1002/hyp.13802.

Birkinshaw, S.J., 2014. Comment on "'A paradigm shift in understanding and quantifying the effects of forest harvesting on floods in snow environments'" by Kim C. Green and Younes Alila. Water Resour. Res. 50, 2765–2768. https://doi.org/10.1002/2013WR013586.

Black, P.E., 1972. Hydrograph responses to geomorphic model watershed characteristics and precipitation variables. J. Hydrol. 17, 309–329. https://doi.org/10.1016/0022-1694(72)90090-X.

Blöschl, G., 2006. Hydrologic synthesis: Across processes, places, and scales. Water Resour. Res. 42, 3. https://doi.org/10.1029/2005WR004319.

Bradford, J.B., Kastendick, D.N., 2010. Age-related patterns of forest complexity and carbon storage in pine and aspen-birch ecosystems of northern Minnesota, USA. Can. J. For. Res. 40, 401–409. https://doi.org/10.1139/X10-002.

Bradshaw, C.J., Sodhi, N.S., Peh, K.-S.-H., Brook, B.W., 2007. Global evidence that deforestation amplifies flood risk and severity in the developing world. Glob. Chang. Biol. 13, 2379–2395. https://doi.org/10.1111/j.1365-2486.2007.01446.x.

Brown, A.E., Zhang, L., McMahon, T.A., Western, A.W., Vertessy, R.A., 2005. A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation. J. Hydrol. 310, 28–61. https://doi.org/10.1016/j.jhydrol.2004.12.010.

Buongiorno, J., Zhu, S., Raunikar, R., Prestemon, J.P., 2012. Outlook to 2060 for World Forests and Forest Industries: A Technical Document Supporting the Forest Service 2010 RPA Assessment. USFS Gen. Tech. Rep.

Buras, A., Menzel, A., 2018. Projecting Tree Species Composition Changes of European Forests for 2061–2090 Under RCP 4.5 and RCP 8.5 Scenarios. Front. Plant Sci. 9, 20. https://doi.org/10.3389/fpls.2018.01986.

Buttle, J.M., 2011. The Effects of Forest Harvesting on Forest Hydrology and Biogeochemistry. In: Levia, D.F., Carlyle-Moses, D., Tanaka, T. (Eds.), Forest Hydrology and Biogeochemistry: Synthesis of Past Research and Future Directions. Springer, Dordrecht, pp. 659–677. https://doi.org/10.1007/978-94-007-1363-5_33.

Gelman, A., Hill, J., Yajima, M., 2012. Why we (usually) don't have to worry about multiple comparisons. J. Res. Educ. Eff. 5, 189–211. https://doi.org/10.1080/19345747.2011.618213.

Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. 7, 457–511.

Gelman, A., Tuerlinckx, F., 2000. Type S error rates classical and Bayesian single and multiple comparison procedures. Comput. Stat. 15, 373–390. https://doi.org/10.1007/s001800000040.

Green, K.C., Alila, Y., 2012. A paradigm shift in understanding and quantifying the effects of forest harvesting on floods in snow environments. Water Resour. Res. 48 (W10503), 1–21. https://doi.org/10.1029/2012WR012449.

Guertin, D.P., Barten, P.K., Brooks, K.N., 1987. The Peatland Hydrologic Impact Model: Development and Testing. Hydrol. Res. 18, 79–100. https://doi.org/10.2166/nh.1987.0007.

Gumbel, E.J., 1941. The return period of flood flows. Ann. Math. Stat. 12, 163–190.

Hewlett, J.D., 1982. Forests and floods in the light of recent investigation, in: Proceedings of the Canadian Hydrology Symposium - Associate Committee on Hydrology.

Hewlett, J.D., Lull, H.W., Reinhart, K.G., 1969. In defense of experimental watersheds. Water Resour. Res. 5, 306–316. https://doi.org/10.1029/WR005i001p00306.

Jones, J.A., Grant, G.E., 1996. Peak flow responses to clear cutting and roads in small and large basins, western Cascades. Oregon. Water Resour. Res. 32, 959–974. https://doi.org/10.1029/95WR03493.

Koutsoyiannis, D., 2011. Hurst-Kolmogorov dynamics and uncertainty. J. Am. Water Resour. Assoc. 47, 481–495. https://doi.org/10.1111/j.1752-1688.2011.00543.x.

Kuraś, P.K., Alila, Y., Weiler, M., 2012. Forest harvesting effects on the magnitude and frequency of peak flows can increase with return period. Water Resour. Res. 48 (W01544), 1–19. https://doi.org/10.1029/2011WR010705.

Langsrud, Ø., 2003. ANOVA for unbalanced data: Use Type II instead of Type III sums of squares. Stat. Comput. 13, 163–167. https://doi.org/10.1023/A:1023260610025.

Loftis, J.C., MacDonald, L.H., Streett, S., Iyer, H.K., Bunte, K., 2001. Detecting cumulative watershed effects: The statistical power of pairing. J. Hydrol. 251, 49–64. https://doi.org/10.1016/S0022-1694(01)00431-0.

Lu, S.-Y., 1994. Forest Harvesting Effects on Streamflow and Flood Frequency in the Northern Lake States. University of Minnesota. Ph.D. Thesis.

Murray, C.D., Buttle, J.M., 2003. Impacts of clearcut harvesting on snow accumulation and melt in a northern hardwood forest. J. Hydrol. 271, 197–212. https://doi.org/10.1016/S0022-1694(02)000352-9.

Neary, D.G., 2016. Long-Term Forest Paired Catchment Studies: What Do They Tell Us That Landscape-Level Monitoring. Does Not? 7, 15. https://doi.org/10.3390/f7080164.

Nyberg, P.R., 1987. Soil Survey of Itasca County, Minnesota.

Perala, D., Verry, E., 2011. Forest Management Practices and Silviculture. Peatl. Biogeochem. Watershed Hydrol. Marcell Exp. For. 371–400 https://doi.org/10.1201/b10708-13.

Plummer, M., 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Proc. 3rd Int. Work. Distrib. Stat. Comput. (DSC 2003) 20–22. 10.1.1.13.3406.

Poff, N.L., Bledsoe, B.P., Cuhaciyan, C.O., 2006. Hydrologic variation with land use across the contiguous United States: geomorphic and ecological consequences for stream ecosystems. Geomorphology 79, 264–285. https://doi.org/10.1016/j.geomorph.2006.06.032.

Pomeroy, J.W., Granger, R.J., Pietroniro, A., Elliot, J.E., Toth, B., Hedstrom, N., 1994. Hydrological Pathways in the Prince Albert Model Forest, Final Report Submitted to the Prince Albert Model Forest Association. NHRI Contrib. Ser. No. CS-97004.

R Core Team, 2019. R: A language and environment for statistical computing.

Renne, R.R., 1967. Research Guidelines to Sound Watershed Development. J. Irrig. Drain. Div. ASCE 92, 53–58.

Rogger, M., Agnoletti, M., Alaoui, A., Bathurst, J.C., Bodner, G., Borga, M., Chaplot, V., Gallart, F., Glatzel, G., Hall, J., Holden, J., Holko, L., Horn, R., Kiss, A., Quinton, J. N., Leitinger, G., Lennartz, B., Parajka, J., Peth, S., Robinson, M., Salinas, J.L., Santoro, A., Szolgay, J., Tron, S., Viglione, A., 2017. Land use change impacts on floods at the catchment scale: Challenges and opportunities for future research. Water Resour. Res. 53, 5209–5219. https://doi.org/10.1002/2017WR020723.Received.

Sebestyen, S.D., Dorrance, C., Olson, D.M., Verry, E.S., Kolka, R.K., Elling, A.E., Kyllander, R., 2011a. Long-Term Monitoring Sites and Trends at the Marcell Experimental Forest, in: Kolka, R.K., Sebestyen, S.D., Verry, E.S., Brooks, K.N. (Eds.), Peatland Biogeochemistry and Watershed Hydrology at the Marcell Experimental Forest. CRC Press, Boca Raton, pp. 15–72.

Sebestyen, S.D., Verry, E.S., Brooks, K.N., 2011b. Hydrological Responses to Changes in Forest Cover on Uplands and Peatlands, in: Peatland Biogeochemistry and Watershed Hydrology at the Marcell Experimental Forest. CRC Press: Taylor & Francis Group, LLC, Boca Raton, FL, pp. 401–432.

Sebestyen, S. D., Verry, E. S., Elling, A. E., Kyllander, R. L., Lany, N. K., & Kolka, R. K., 2020a, Marcell Experimental Forest biweekly snow depth, frost depth, and snow water equivalent, 1962 - ongoing, 1 ed., USDA Forest Service, Environmental Data Initiative. 10.6073/pasta/2addb04291411d54c99f0740d5335e8f.

Sebestyen, S. D., Verry, E. S., Elling, A. E., Kyllander, R. L., Lany, N. K., & Kolka, R. K., 2020b. Marcell Experimental Forest breakpoint streamflow, 1962 - ongoing. USDA Forest Service. Environmental Data Initiative. Retrieved from: 10.6073/pasta/99051a7c663dc381a1f6a843bcdb78e9.

Sebestyen, S. D., Verry, E. S., Elling, A. E., Kyllander, R. L., Roman, D. T., Burdick, J. M., Kolka, R. K., 2020c. Marcell Experimental Forest daily precipitation, 1961 - ongoing ver 1. USDA Forest Service. Environmental Data Initiative. Retrieved from: 10.6073/pasta/75646a3bd41ba3219d0e578e8374eef7.

Smith, A.F.M., Roberts, G.O., 1993. Bayesian computation via the gibbs sampler and related Markov Chain Monte Carlo methods. J. R. Stat. Soc. Ser. B 55, 3–23. https://doi.org/10.1111/j.2517-6161.1993.tb01466.x.

Su, Y.-S., Yajima, M., 2015. R2jags: Using R to Run “JAGS.”.

Tetzlaff, D., Buttle, J., Carey, S.K., Mcguire, K., Laudon, H., Soulsby, C., 2015. Tracer-based assessment of flowpaths, storage and runoff generation in northern catchments: a review. Hydrol. Process. 29, 3475–3490. https://doi.org/10.1002/hyp.10412.

Tetzlaff, D., Soulsby, C., Buttle, J., Capell, R., Carey, S.K., Laudon, H., McDonnell, J., McGuire, K., Seibert, J., Shanley, J., 2013. Catchments on the cusp? Structural and functional change in northern ecohydrology. Hydrol. Process. 27, 766–774. https://doi.org/10.1002/hyp.9700.

Thomas, R.B., Megahan, W.F., 1998. Peak flow responses to clear-cutting and roads in small and large basins, western Cascades, Oregon: A second opinion. Water Resour. Res. 34, 3393–3403. https://doi.org/10.1029/98WR02500.

Verry, E.S., 2004. Land fragmentation and impacts to streams and fish in the Central and Upper Midwest. In: Ice, G.G., Stednick, J.D. (Eds.), A Century of Forest and Wildland Watershed Lessons. Society of American Foresters, Bethesda, MD, pp. 129–154.

Verry, E.S., 1986. Forest harvesting and water: The Lake States Experience. Water Resour. Bull. 22, 1039–1047.

Verry, E.S., Janssens, J., 2011. Geology, Vegetation, and Hydrology of the S2 Bog at the MEF: 12,000 Years in Northern Minnesota, in: Kolka, R.K., Sebestyen, S.D., Verry, E. S., Brooks, K.N. (Eds.), Peatland Biogeochemistry and Watershed Hydrology at the Marcell Experimental Forest. CRC Press: Taylor & Francis Group, LLC, Boca Raton, FL, pp. 93–134.

Verry, E.S., Lewis, J.R., Brooks, K.N., 1983. Aspen clearcutting increases snowmelt and storm flow peaks in North Central Minnesota. Water Resour. Bull. 19, 59–67.

Wei, L., Qiu, Z., Zhou, G., Kinouchi, T., 2020. Stormflow threshold behaviour in a subtropical mountainous headwater catchment during forest recovery period 1728–1740. 10.1002/hyp.13658.

Yu, X.J., Alila, Y., 2019. Nonstationary frequency pairing reveals a highly sensitive peak flow regime to harvesting across a wide range of return periods. For. Ecol. Manage. 444, 187–206. https://doi.org/10.1016/j.foreco.2019.04.008.

Zégre, N., Skaugset, A.E., Som, N.A., McDonnell, J.J., Ganio, L.M., 2010. In lieu of the paired catchment approach: Hydrologic model change detection at the catchment scale. Water Resour. Res. 46, 1–20. https://doi.org/10.1029/2009WR008601.

Zon, R., 1927. Forests and Water in the Light of Scientific Investigation. Forest Service, United States Department of Agriculture. US Govt. print. off.