

# Peak detection in sediment–charcoal records: impacts of alternative data analysis methods on fire-history interpretations

Philip E. Higuera<sup>A,E</sup>, Daniel G. Gavin<sup>B</sup>, Patrick J. Bartlein<sup>B</sup>  
and Douglas J. Hallett<sup>C,D</sup>

<sup>A</sup>Department of Forest Ecology and Biogeosciences, University of Idaho,  
Box 83844-1133, Moscow, ID 83844, USA.

<sup>B</sup>Department of Geography, University of Oregon, Eugene, OR 97493, USA.

<sup>C</sup>Biogeoscience Institute, University of Calgary, 2500 University Drive NW,  
Calgary, AB, T2N 1N4, Canada.

<sup>D</sup>School of Environmental Studies, Queen's University, BioSciences Complex,  
3134, Kingston, ON, K7L 3N6, Canada.

<sup>E</sup>Corresponding author. Email: phiguera@uidaho.edu

**Abstract.** Over the past several decades, high-resolution sediment–charcoal records have been increasingly used to reconstruct local fire history. Data analysis methods usually involve a decomposition that detrends a charcoal series and then applies a threshold value to isolate individual peaks, which are interpreted as fire episodes. Despite the proliferation of these studies, methods have evolved largely in the absence of a thorough statistical framework. We describe eight alternative decomposition models (four detrending methods used with two threshold-determination methods) and evaluate their sensitivity to a set of known parameters integrated into simulated charcoal records. Results indicate that the combination of a globally defined threshold with specific detrending methods can produce strongly biased results, depending on whether or not variance in a charcoal record is stationary through time. These biases are largely eliminated by using a locally defined threshold, which adapts to changes in variability throughout a charcoal record. Applying the alternative decomposition methods on three previously published charcoal records largely supports our conclusions from simulated records. We also present a minimum-count test for empirical records, which reduces the likelihood of false positives when charcoal counts are low. We conclude by discussing how to evaluate when peak detection methods are warranted with a given sediment–charcoal record.

**Additional keywords:** bias, paleoecology, sensitivity.

## Introduction

High-resolution charcoal records are an increasingly common source of fire-history information, particularly in ecosystems where tree-ring records are short relative to average fire-return intervals (Gavin *et al.* 2007). Over the past several decades, numerous studies have used peaks in charcoal accumulation in sediment records to estimate the timing of ‘fire episodes’, one or more fires within the sampling resolution of the sediment record (Whitlock and Larsen 2001). Identifying fire episodes from charcoal records is most promising when fires: (1) are large; (2) burn with high severity; and (3) recur with average intervals at least five times the sampling resolution of the sediment record (Clark 1988b; Whitlock and Larsen 2001; Higuera *et al.* 2005, 2007). Sediment–charcoal records are thus particularly valuable for studying stand-replacing fire regimes in boreal and subalpine forests, where all three of these conditions are typically met.

Interpreting fire episodes from sediment–charcoal records would be straightforward if they were characterised by low levels of charcoal punctuated by unambiguous peaks. In reality, however, charcoal records are complex and non-stationary,

i.e. their mean and variance change over time (Clark *et al.* 1996; Clark and Patterson 1997; Long *et al.* 1998). Empirical and theoretical studies (e.g. Marlon *et al.* 2006; Higuera *et al.* 2007) suggest that non-stationarity in charcoal records can arise from at least two sets of processes: (1) changes in the fire regime, including the rate of burning, the intensity of fires, the type of vegetation burned, and thus charcoal production per unit time; or (2) changes in the efficiency of charcoal delivery to the lake centre (taphonomy) due to changing rates of slope wash or within-lake redeposition. The latter process, known as sediment focussing, can greatly affect the sediment accumulation rate as a lake fills in over time (Davis *et al.* 1984; Giesecke and Fontana 2008) and may produce long-term trends in charcoal records unrelated to changes in the fire regime. Recognising the importance of these processes, paleoecologists have applied a range of statistical methods to charcoal data in order to isolate the signal related to ‘local’ fire occurrence (e.g. within 0.5–1.0 km; Gavin *et al.* 2003; Lynch *et al.* 2004a; Higuera *et al.* 2007) and reconstruct fire history. Despite the proliferation of statistical methods for peak identification, seemingly no study has

discussed the assumptions underlying alternative methods and their impacts on fire-history interpretations.

Here, we address several key issues related to peak identification in high-resolution, macroscopic charcoal records<sup>A</sup> by using simulated and empirical charcoal records. We start by discussing some important statistical properties of macroscopic charcoal records and then describe the motivation for statistical treatments. We briefly review how different methods have been applied, and then introduce a typology of methods, including their respective assumptions and justifications. Second, we illustrate and quantify the biases that these techniques can introduce to fire-history interpretations by applying them to simulated charcoal records. Third, we apply the same methods to three previously published charcoal records to demonstrate potential biases in empirical records, and we introduce a technique to minimise some of these biases. Finally, we conclude with recommendations of specific methodologies and a discussion of how analysts can evaluate the suitability of records for peak identification rather than other qualitative or quantitative analyses.

### Temporal variability in charcoal time series

Charcoal time series can be generally characterised as ‘noisy’, and they contain many forms of non-stationarity, including changing short-term variability superimposed on a slowly varying mean (Long *et al.* 1998; Higuera *et al.* 2007). Changes in variability (i.e. heteroscedasticity) have implications for the particular goal of data analysis. When the goal is to quantify changes in total charcoal input, as an index of biomass burning for example, heteroscedasticity violates the assumptions of parametric statistics useful in this context, e.g. analysis of variance and regression. In particular, in analysis of variance (or in the *t*-test of the difference of means in the case of two periods), heteroscedasticity increases the probability of Type I error, falsely inferring significant differences between periods (Underwood 1997). Similarly, in regression analysis, fitting a trend line to charcoal data with changing variability over time can increase the variability of the slope coefficient. Changes in variability (besides being interesting in their own right) can thus lead to false conclusions about the significance of long-term trends or differences between different parts of a record. In practice, heteroscedasticity is usually dealt with by applying a ‘variance-stabilising transformation’ (Emerson 1983) that acts to homogenise variance across a record. As will be illustrated below, when the goal of charcoal analysis is peak identification, transformation can lead to the exaggeration of some peaks and suppression of others. Consequently, the specific approach taken (whether to transform or not) should depend on the overall focus of an analysis. In this paper, we focus on the goal of detecting local fires through peak detection.

### Analytical methods for inferring local fire occurrence

Following the pioneering work of Clark (1988b, 1990) in which fire events surrounding small lakes were identified from charcoal in thin-sections of laminated sediments, similar

approaches were developed for quantifying macroscopic charcoal abundance and subsequently adopted by a large number of research groups (Table 1; see also Whitlock and Larsen 2001). Most techniques quantify charcoal as either the total number of pieces or surface area (mm<sup>2</sup>) of charcoal in a particular size class, within volumetric subsamples taken contiguously through sediment cores (typically at 0.5- to 1.0-cm resolution, corresponding to ~10–25-year resolution for most lakes). The resulting concentration of charcoal (pieces cm<sup>-3</sup>, or mm<sup>-2</sup> cm<sup>-3</sup>) in each level is multiplied by the estimated sediment accumulation rate (cm year<sup>-1</sup>) to obtain the charcoal accumulation rate (CHAR, pieces cm<sup>-2</sup> year<sup>-1</sup> or mm<sup>-2</sup> cm<sup>-2</sup> year<sup>-1</sup>). Sediment accumulation rates, and the age of each sample, are estimated by an age–depth model based on radiometric dates, tephra layers, and any additional sources of age information. The use of accumulation rates can potentially correct for changing sediment accumulation rates that would dilute or concentrate charcoal in a given volume of sediment, and as mentioned above, may also be affected by sediment focussing processes. Usually, the CHAR series is interpolated to a constant temporal resolution to account for unequal sampling intervals resulting from variable sediment accumulation rates. This step is necessary to develop threshold statistics that are not biased to a particular portion of a record, and to standardise within- and between-site comparisons.<sup>B</sup> Hereafter, we refer to the interpolated CHAR series as *C*. The analytical choices and sources of error in the development of a charcoal record are briefly summarised in Table 2 and discussed in detail by Whitlock and Larsen (2001).

At this point, most *C* series can be characterised as irregular time series with discrete peaks superimposed on a slowly varying mean. Although the size of any individual peak reflects the size, location, and charcoal production of individual fires, the average size of peaks may change through time, contributing to a slowly changing variance. This non-stationarity may arise, as discussed above, owing to variations in charcoal production per unit time or variable taphonomic and sedimentation processes. Without knowledge of whether non-stationarity is due to changes in taphonomy and sedimentation or to real changes in fire history, it is reasonable to stabilise the variance of peak heights so as to not ‘pass over’ periods of low charcoal. This motivates the manipulation of *C* to produce a stationary series in which all local fires would theoretically result in a similar range of peak sizes. Doing so would allow for the application of a single global threshold value to the final series to separate fire-related from non-fire-related peaks.

In practice, determining the size of peaks that represents local fires involves a three-step ‘decomposition’ of the *C* series (Clark *et al.* 1996; Long *et al.* 1998; Fig. 1). First, the slowly varying mean, or ‘background’ component,  $C_{back}$ , is modelled through a curve-fitting algorithm, e.g. a locally weighted regression that is robust to outliers (e.g. Cleveland 1979). The window size for this smoothing varies between studies but is typically between 100 and 1000 years. Background estimation may be preceded by transforming *C* (e.g. logarithmically). Second, the background trend is removed from the series by subtraction ( $C - C_{back}$ ) or

<sup>A</sup>We refer to macroscopic charcoal records as those quantifying charcoal not passing through a sieve of 125 µm or larger.

<sup>B</sup>When sampling intervals are not standardised within a record or between two records, then biases may be introduced when applying criteria uniformly. Interpolation helps minimise, but not remove, this bias, as noted in the last section of this paper.

**Table 1. Published fire-history studies in North America based on macroscopic charcoal (sieved or in thin sections) where the goal of analysis was to detect peaks associated with local fires**  
 Studies are grouped by threshold type and the detrending models used for analysis. *n* indicates the total number of studies in each category. Methods of threshold determination as follows: TR, detected peaks compared with fires reconstructed from tree-rings (fire scars, stand ages or both); SC, detect peaks compared with radiocarbon dates from local soil charcoal; H, detected peaks compared with historical fire record; S1, sensitivity analysis based on coefficient quantifying separation of peaks from background; S2, sensitivity analysis based on qualitative assessment of results using alternative threshold criteria; for TR, SC and H methods, the number of independent fire records is shown in parentheses. Threshold values: units for NR and TR models are pieces per square centimetre per year for counts or square millimetres per square centimetre per year for area. Thresholds for NI and TI models are unitless index values. For local thresholds, the percentile of the Gaussian mixture model (GMM) defines a different threshold value for each sample, and thus threshold values are not reported. GSM, Gaussian single model, with mean of 0

Threshold type	Detrending model	Citation	Location (state or province)	Particle count (C) or area (A)	Sediment volume per sample (cm <sup>3</sup> )	Size class tallied	Background estimate	Threshold determination	Threshold value
Global, <i>n</i> = 37	Non-transform-residuals (NR), <i>n</i> = 13	Clark (1990)	Minnesota, USA	A	Thin sections	>60 µm long	15-year moving average	Three lakes: TR (2–8)	> 42 to 68
		Millsbaugh and Whitlock (1995)	Wyoming, USA	C	5	125 µm	Three-point, centre weighted average	Five lakes: TR (2–8)	≥3.4, >4.6, ≥5
		Clark <i>et al.</i> (1996)	New York, USA	A	Thin sections	>60 µm long	Inverse Fourier transform: 30-year window	TR (11)	>60
		Clark and Royall (1996)	New York, Wisconsin, Pennsylvania, Maine, USA; Ontario, Canada	A	Thin sections	>60 µm long	Inverse Fourier transform: 10-year window	Seven lakes: H, TR	>40
		Carcaillet <i>et al.</i> (2001)	Québec, Canada	A	1	>150 µm	Inverse Fourier transform: 100-year window	Three lakes: TR: not possible	> 1 s.d. of the average of background
		Gavin <i>et al.</i> (2003)	British Columbia, Canada	C	~12	150–500 µm	26-year locally weighted minimum value	SC (12), TR (3); S1	0.22
		Lynch <i>et al.</i> (2002)	Alaska, USA	A	1	>180 µm	100-year locally weighted mean	Three lakes: H (2); S1	0.07; upper 12% tail of residuals
		Lynch <i>et al.</i> (2004a)	Alaska, USA; Manitoba, Northwest Territories, Ontario, Canada	A	1	>180 µm	100-year locally weighted mean	15 lakes: H (1–2) for each lake; S1	0.03–0.33; upper 8–13% of residuals
		Lynch <i>et al.</i> (2004b)	Alaska, USA	A	1–3	>180 µm	100-year locally weighted mean	Four lakes: S1	0.018, 0.085
		Gavin <i>et al.</i> (2006)	British Columbia, Canada	C	2–5	>125 µm	500-year robust lowess	Two lakes: H, TR (2); S1	GMM; 0.08, 5.00

Prichard <i>et al.</i> (2009)	Washington, USA	C	~10	>150–500 µm	750-year locally weighted mean	TR (2): S1	GMM at 99th percentile: not reported; GMM at 95th percentile: 0.2
Ali <i>et al.</i> (2008)	Québec, Canada	A	1	>160 µm	500-year tricube locally weighted regression	Lake, peat and soil charcoal compared; S1	GSM at 95th percentile: 0.015, 0.025, 0.040, 0.007
Ali <i>et al.</i> (2009a)	Québec, Canada	A	1	>160 µm	1000-year tricube locally weighted regression	Four lakes: S1	GSM at 95th percentile: 0.015, 0.025, 0.040, 0.007
Hallett and Anderson (2010)	California, USA	C	2.5	>125 µm	500-year robust lowess	Two lakes: TR (1–2): S1	GSM at 95th percentile: 0.03
Higuera <i>et al.</i> (2005)	Washington, USA	C	3	150–500, >500 µm	Series median (300-year)	12 small hollows: TR (1–3 per site)	1.63–1.75
Tweitten <i>et al.</i> (2009)	Wisconsin, USA	C	1	>150 µm	300-year lowess	Not possible	1.3
Long <i>et al.</i> (1998)	Oregon, USA	C	2.5	>125 µm	600-year locally weighted mean	TR, H (4)	1.12
Hallett and Walker (2000)	British Columbia, Canada	C, A	1	>150 µm	500-year locally weighted mean	TR (2)	1.0
Millsbaugh <i>et al.</i> (2000)	Wyoming, USA	C	5	>125 µm	750-year locally weighted mean	S2	1.0
Mohr <i>et al.</i> (2000)	California, USA	C	5	>125 µm	120-year locally weighted mean	Two lakes: TR (3)	1.0
Long and Whitlock (2002)	Oregon, USA	C	2.5	>125 µm	600-year locally weighted mean	H (2)	1.25
Brunelle and Anderson (2003)	California, USA	C	5	>125 µm	500-year locally weighted mean	H (1)	1.1
Brunelle and Whitlock (2003); Brunelle <i>et al.</i> (2005)	Idaho and Montana, USA	C	5	>125 µm	750 or 600-year locally weighted mean	Four lakes: TR (3–5)	1.15–1.30
Hallett <i>et al.</i> (2003a)	British Columbia, Canada	C	10	>125 µm	400-year locally weighted mean	Two lakes: SC (10–18)	1.0
Hallett <i>et al.</i> (2003b)	British Columbia, Canada	C	10	>125 µm	50-year locally weighted mean	TR (4)	1.1
Daniels <i>et al.</i> (2005)	California, USA	C	~8	>125 µm	370-year locally weighted mean	TR (3)	1.0
Briles <i>et al.</i> (2005)	Oregon, USA	C	2–5	>125 µm	500-year locally weighted mean	TR (3)	1.1

(Continued)

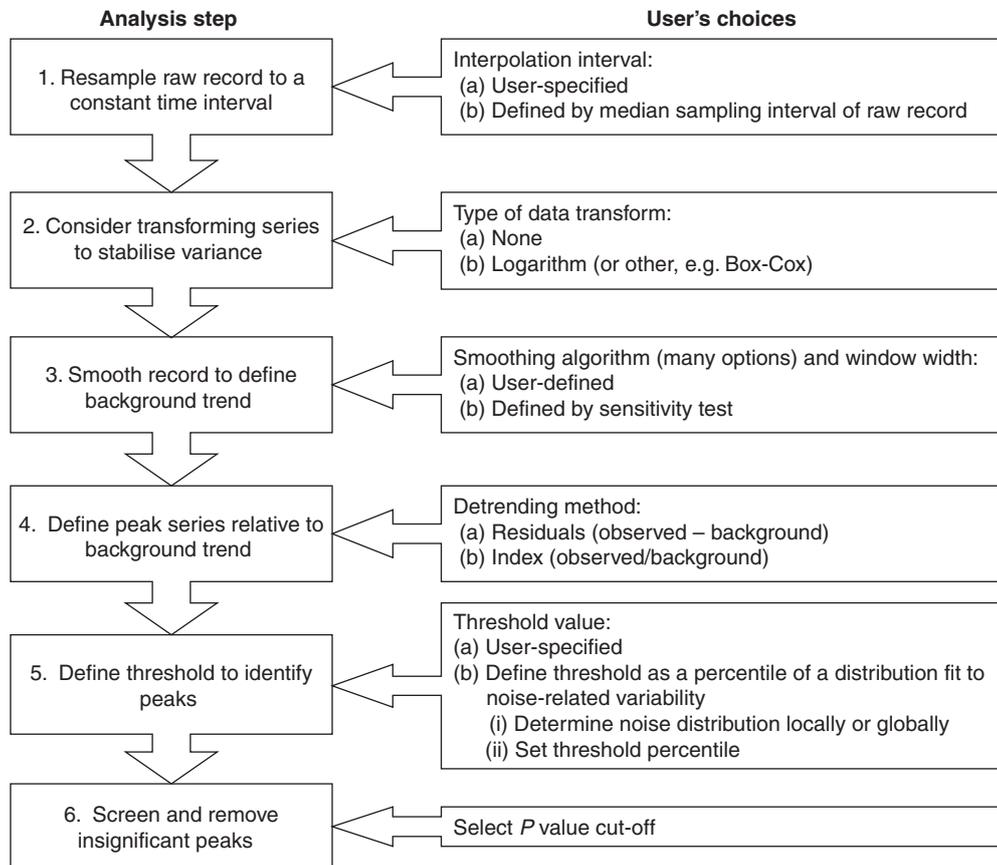
Table 1. (Continued)

Threshold type	Detrending model	Citation	Location (state or province)	Particle count (C) or area (A)	Sediment volume per sample (cm <sup>3</sup> )	Size class tallied	Background estimate	Threshold determination	Threshold value
		Toney and Anderson (2006)	Colorado, USA	C	5	>125 µm	300-year locally weighted mean	H (1)	1.0
		Anderson <i>et al.</i> (2006)	Alaska, USA	C	5	>125 µm	300-year locally weighted mean	H (3–5); SC (4)	1.2
		Power <i>et al.</i> (2006)	Montana, USA	C	0.3–1?	>125 µm	150-year locally weighted mean	H (1)	1.05
		Manlon <i>et al.</i> (2006)	California, Oregon, Montana, Wyoming, Idaho, USA	C	1–5	>125 µm	500-year locally weighted mean	15 lakes: H (1–5); TR (1–5); varies by site	1.0–1.3
		Long <i>et al.</i> (2007)	Oregon, USA	C	3–5	>125 µm	600-year locally weighted mean	H (2–3); S1	1.2
		Anderson <i>et al.</i> (2008)	Colorado, New Mexico, USA	C	1–5	>125 µm	1000-year locally weighted mean	Three lakes and three bogs: H (1–3)	1.01
		Allen <i>et al.</i> (2008)	New Mexico, USA	C	1–5	>125 µm	300-year locally weighted mean	Two bogs: TR (1–4), H (fire-free interval)	1.01
		Minckley <i>et al.</i> (2007)	Oregon and California, USA	C	4–5	>125 µm	900-year locally weighted mean	Three lakes: H (lowest threshold with no fires in last 100 years)	1.00, 1.05, 1.15
		Whitlock <i>et al.</i> (2008)	Wyoming, USA	C	1	>125 µm	150-year locally weighted mean	TR (2)	1.1
		Beatty and Taylor (2009)	California, USA	C	3	>125 µm	240-year locally weighted mean	TR (3)	1.0
Local, <i>n</i> = 6	Non-transform-residuals (NR), <i>n</i> = 6	Higuera <i>et al.</i> (2008)	Alaska, USA	C	3–5	>150 µm	Smoothed 500-year median or mode	Two lakes (pre-modern): S2	GMM at 99th percentile
		Walsh <i>et al.</i> (2008)	Washington, USA	C	1	>125 µm	500-year robust lowess	TR (3); S2	GMM at 95th percentile
		Briles <i>et al.</i> (2008)	California, USA	C	2	>125 µm	700-year locally weighted regression	Two lakes; S2	GMM at 95th percentile
		Manlon <i>et al.</i> (2009)	North America	C	Varies with several records used	>125 µm	500, 600, 800-year robust lowess	35 lakes: S2	GMM at 95th percentile
		Higuera <i>et al.</i> (2009)	Alaska, USA	C	3–5	>150 µm	Smoothed 500-year median or mode	Four lakes: H (1); S2	GMM at 99th percentile
		Huerta <i>et al.</i> (2009)	Wyoming, USA	C	5 and 50	>125 µm	500-year robust lowess	H (1); S2	GMM at 95th percentile

**Table 2. Decisions typically required to develop a high-resolution lake-sediment macroscopic charcoal record, summarised from Whitlock and Larsen (2001)**

The aim of the current paper is to discuss data manipulations after completing these steps of developing a charcoal record

Step	Decisions	Potential issues and sources of error
(1) Sediment collection	Coring location	Gaps in record
(2) Sediment subsampling	Sediment volume per sample Sampling interval	Volume overestimate (core shrinkage) Sample volume too small, resulting in low charcoal counts Interval too long to distinguish consecutive fire events
(3) Sediment sieving	Sieve sizes	Incomplete sieving Sample spillage
(4) Charcoal quantification	Count or area	Misidentification Breakage of charcoal results in inflated counts
(5) Estimation of charcoal accumulation rate (CHAR)	Age–depth model fitting to calculate sediment accumulation rates	Poor chronological control
(6) Interpolation of CHAR to a constant interval	Interval size (typical value is the median sample deposition time)	Loss of resolution in portions of the record



**Fig. 1.** The set of decisions required for analysing a charcoal time series with the goal of peak detection for interpretation of fire episodes. These steps are implemented in the *CharAnalysis* software (<http://code.google.com/p/charanalysis/>, accessed 30 November 2010).

division ( $C \div C_{back}$ ), creating a series of residuals or indices respectively. This detrended series is frequently termed the ‘peak component’, but in the case of indices, it is dimensionless rather than a portion of  $C$ , as implied by ‘peak component’. Here, we use the term ‘peak series’ and notation  $C_{peak}$  to refer to the

detrended series. Third, a threshold is applied to  $C_{peak}$  to separate variability related to local fire occurrence from variability unrelated to local fire occurrence (e.g. random variability and sediment mixing). Peaks exceeding the threshold are the basis for fire-event frequency and fire-event return interval calculations.

**Table 3.** Selected abbreviations used in the text and corresponding definitions

Abbreviation	Definition
Components of a charcoal record	
$C$	Resampled charcoal in a charcoal series, expressed as pieces $\text{cm}^{-2} \text{year}^{-1}$ or $\text{cm}^2 \text{cm}^{-2} \text{year}^{-1}$
$\log(C + 1)$	Natural logarithm of resampled charcoal, after one is added to guard against negative values
$C_{back}$	Background charcoal, defined as a function of resampled charcoal
$C_{back}$ , where $C_{back} = f(\log[C + 1])$	Background charcoal, defined as a function of log-transformed, resampled charcoal
$C_{peak}$	Detrended, or 'peak' series of a charcoal record, after trends in background charcoal are removed
Detrending models	
NR	No-transform-residual: $C_{peak} = C - C_{back}$
NI	No-transform-index: $C_{peak} = C \div C_{back}$
TR	Transform-residual: $C_{peak} = \log(C + 1) - C_{back}$ , where $C_{back} = f(\log[C + 1])$
TI	Transform-index: $C_{peak} = \log(C + 1) \div C_{back}$ , where $C_{back} = f(\log[C + 1])$

Here, we present a typology of four possible decomposition approaches based on whether the raw or transformed  $C$  series is used and whether  $C_{peak}$  is calculated as residuals or index values relative to  $C_{back}$  (see Table 3 for abbreviations). The *no-transform-residual model* (NR model hereafter) is a simple subtraction:  $C - C_{back}$ . The *no-transform-index model* (NI model) is a ratio:  $C/C_{back}$ . Because background charcoal is in the denominator, the NI model cannot be applied when background charcoal equals zero, which occurs in non-forested or treeline ecosystems (e.g. Huber *et al.* 2004; Higuera *et al.* 2009; Hallett and Anderson 2010). The *transform-residual model* (TR model) first log-transforms  $C$  (after adding 1 to guard against negative values) before calculating the background:  $\log(C + 1) - C_{back}$ , where  $C_{back} = f(\log[C + 1])$ . Finally, the *transform-index model* (TI model) is the ratio of the log-transformed series:  $\log(C + 1) \div C_{back}$ , where  $C_{back} = f(\log[C + 1])$ .

Nearly all studies have used the NR or the TI model in charcoal peak analyses (Table 1), but there has been no discussion of the assumptions underlying each model. The NR model implicitly assumes that charcoal peaks from local fires are created through additive processes. That is, charcoal introduced from a fire is added to the total amount of background charcoal (i.e. charcoal delivery to the core site during periods without local fires). Background charcoal may change as redeposition processes change (e.g. wind-mixing of littoral sediment, higher fire frequencies), but the total amount of charcoal produced per fire remains unchanged. Variance stabilisation is the goal of the NI, TR and TI models, which implicitly assume that charcoal peaks from local fires are created through multiplicative processes; i.e. the total amount of charcoal introduced from a local fire is some multiple of background charcoal. Similar variance-stabilisation goals used in dendrochronology are typically based on the NI or TR models, rather than the methods more recently adopted for charcoal records (NR and TI models; Table 1). As in dendrochronology (Cook and Peters 1997; Fowler 2009), the choice of detrending model has an important impact on the resulting detrended series.

In comparison with the little attention given to alternative detrending models, recent papers have more carefully addressed the task of determining threshold values for peak identification. Comparison of peaks with known fire events (dated from historical records or tree-rings) may help in selecting a threshold, but historical records often represent only a fraction of a

charcoal record, and a wide range of thresholds may still be appropriate (e.g. Gavin *et al.* 2006). Clark *et al.* (1996) addressed this issue by using a sensitivity analysis to test the number of peaks as a function of changing threshold values. They reasoned that if  $C_{peak}$  comprised a population of small values (e.g. background charcoal) and a smaller population of large values (local fires), then the sensitivity test should detect the split between these populations. Gavin *et al.* (2006) built on this sensitivity test by modelling  $C_{peak}$  as a mixture of two Gaussian distributions with different means, variances, and proportional contributions to the total population. The lower distribution is assumed to represent the majority of time during which  $C$  is small and is affected mainly by distant fires, redeposition, mixing and random variability (i.e. the 'noise' unrelated to specific fires). The upper distribution, ideally distinct from the lower distribution, describes the variability due to local fires and can be considered the 'signal' of interest. Gavin *et al.* (2006) suggested that the threshold be at the upper end of the noise distribution, and Higuera *et al.* (2008) further specified that the threshold be at the 95th, 99th or 99.9th percentile of the noise distribution. If the noise and signal distributions are distinct, then the variance of the signal distribution ( $\sigma_S^2$ ) would be much larger than that of the noise distribution ( $\sigma_N^2$ ). A signal-to-noise index (SNI; Higuera *et al.* 2009), calculated as  $\sigma_S^2/(\sigma_S^2 + \sigma_N^2)$ , approaches one when the noise distribution is tightly defined with a narrow standard deviation. SNI values less than  $\sim 0.5$  suggest poor separation of large peaks from the noise-attributable variation. We note these details here because the Gaussian mixture approach assumes that the distribution of  $C_{peak}$  values is right-skewed, and therefore variance-stabilising expressions of  $C_{peak}$  (all but the NR model) work against defining a distinct noise distribution.

Unless variance of  $C_{peak}$  does not change through time (i.e. it is homoscedastic), selecting a threshold based on the entire series could lead to systematic biases towards detecting small or large peaks (depending on which size dominates the record). Although variance-stabilisation approaches were developed to address this issue, Higuera *et al.* (2008, 2009) introduced a new approach intended to be more adaptable by applying the Gaussian mixture model introduced by Gavin *et al.* (2006) to a 500-year moving window of  $C_{peak}$  centred on each time step in the series. This technique is termed a 'local threshold', and it

accounts for potentially changing variance of  $C_{peak}$  by selecting a threshold based on  $\sigma_N^2$  in a user-defined subsection of the record. Using smaller sample sizes to compute the Gaussian mixture distribution increases the chance of erratic model fits in a portion of the cases. Thus, it is important to smooth the local thresholds (typically to the same frequency as that used to define  $C_{back}$ ) such that they vary smoothly over time and be cognisant of the total number of samples in each local population (a minimum of  $\sim 30$  is recommended; Higuera *et al.* 2009). This decomposition approach is similar to peak-detection methods in other applications (e.g. Mudelsee 2006) in that it accounts for changes in both the central tendency and variability in a series.

Finally, Gavin *et al.* (2006) introduced a test to screen peaks detected by a threshold that may nevertheless result from statistically insignificant changes in charcoal abundance. This ‘minimum-count test’ applies specifically to studies quantifying charcoal through numbers, as opposed to area, and it examines the possibility that the differences in counts between two samples may result simply from sampling effects. If charcoal count and volume data are available, then it is possible to assess the minimum increase in charcoal count required to be statistically greater than a previous sample, assuming measured counts are Poisson-distributed around the ‘true’ (unknown) count for a given sample volume. The probability that two sample counts,  $X_1$  and  $X_2$ , from sediment volumes  $V_1$  and  $V_2$ , may originate from the same Poisson distribution is estimated from the  $d$  statistic:

$$d = \frac{\left| X_1 - (X_1 + X_2) \left( \frac{V_1}{V_1 + V_2} \right) \right| - 0.5}{\sqrt{(X_1 + X_2) \left( \frac{V_1}{V_1 + V_2} \right) \left( \frac{V_2}{V_1 + V_2} \right)}} \quad (1)$$

where the significance of  $d$  is assessed from the cumulative normal distribution (Detre and White 1970; Shiu and Bain 1982). This test does not incorporate additional errors in counts from laboratory error (Table 2), and so significance thresholds higher than 0.05 may be warranted. We incorporate the minimum-count test here because the possibility of sampling-related errors increases with the variance-stabilisation models (NI, TR, TI) owing to the inflation of very small changes in  $C$  at times when  $C_{back}$  is small (Cook and Peters 1997).

## Methods

To illustrate how analytical choices affect peak identification, we applied the methods introduced above to simulated and empirical charcoal records. With simulated records, where the underlying processes creating a charcoal record are known, we evaluated the sensitivity of each of the four decomposition and the two threshold-determination methods (global and local thresholds) to two hypothetical scenarios (described below). We analysed the empirical records in the same manner but also applied the minimum-count test to illustrate the impacts of this technique.

### Simulated records

Simulated charcoal records were generated from statistical distributions to reflect two scenarios for the relationship between  $C$  and  $C_{back}$ . In both scenarios, the rate of peak

occurrence (implicitly representing local fires) was constant, but  $C_{back}$  increased half-way through the 10 000-year record. In Scenario 1, charcoal peak heights had a constant variance that was independent of  $C_{back}$ , representing the assumption that charcoal from local fires is added to a charcoal record through additive processes; thus variability is stationary throughout the record. In Scenario 2, peak heights varied in direct proportion to  $C_{back}$ , representing a multiplicative relationship between charcoal from local fires and  $C_{back}$ ; thus the charcoal series is heteroscedastic.

Simulated records with 20-year time steps,  $x(i)$ ,  $i = 0, 20, 40, \dots, 10\,000$ , were constructed in three steps, and we use the notation  $C_b$  and  $C_p$  to refer to the known populations of background and peak charcoal respectively, whereas the estimated populations are referred to with  $C_{back}$  and  $C_{peak}$ , as introduced above. First, background charcoal,  $C_b$ , was prescribed as constant values that increased from a minimum of 50 to a maximum of 100 pieces per  $5\text{ cm}^3$  between 5500 and 4500 simulated years before present (BP). Specifically, the concentration of background charcoal in any 20-year sample,  $x(i)$ , was defined as:

$$C_b(i) = \frac{\min(C_b) + \max(C_b)}{1 + \exp[-lrx(i)]} \quad (2)$$

where  $l = 45$  and  $r = 0.009$  and determine the location (in time) and rate of change in  $C_b$  respectively. Second, a charcoal series characterised by right-skewed high-frequency variation,  $C_p$  (pieces per  $5\text{ cm}^3$ ), was calculated from random numbers using a power function, as follows:

$$C_p(i) = b[-\log \varepsilon(i)]^c \quad (3)$$

where  $b = 35$  and determines the location of the distribution,  $c = 1.25$  and creates a distribution slightly more skewed than a log-normal distribution (as found in many empirical records; Marlon *et al.* 2009), and  $\varepsilon(i) \approx N(0, 1)$ , a random number from a normal distribution with mean 0 and standard deviation 1. Third, the background and peak series (pieces  $\text{cm}^{-3}$ ) were added, and then multiplied by the sediment accumulation rate,  $s_{acc}$  ( $\text{cm year}^{-1}$ ), to obtain the final series of charcoal accumulation rates (CHAR, pieces  $\text{cm}^{-2}\text{ year}^{-1}$ ),  $C$ . For Scenario 1, no further treatment was performed, and:

$$C(i) = s_{acc}[C_b(i) + C_p(i)] \quad (4)$$

For Scenario 2,  $C$  was scaled to background charcoal,  $C_b$ , as follows:

$$C(i) = s_{acc} \left[ \frac{C_b(i)}{\max(C_b)} C_b(i) + C_p(i) \right] \quad (5)$$

As a result, peak heights in Scenario 2 increased proportionally to  $C_b$ , and the structure of the variance changed through the time series.

### Empirical records

We selected three high-resolution charcoal records with differing variability in background charcoal and peak heights. Little Lake

(Long *et al.* 1998) is located in Douglas-fir forest in the Oregon Coast Range. The 3.3-ha, 4.0-m-deep lake is surrounded by a fen and has a small inflowing stream draining a 597-ha watershed (Marlon *et al.* 2006; C. Long, pers. comm., November 2009). The 11.3-m core has overall  $C$  values similar to the simulated records (median = 14.4 pieces  $\text{cm}^{-2} \text{year}^{-1}$ ). Over its 9000-year record,  $C_{\text{back}}$  varies between 0.94 and 44.04 pieces  $\text{cm}^{-2} \text{year}^{-1}$ , and vegetation was consistently dominated by Douglas-fir. Rockslide Lake (Gavin *et al.* 2006) is located in subalpine forest in south-east British Columbia. The 3.2-ha, 14.1-m-deep lake is fed by an intermittent stream within an 86-ha watershed. The 2.1-m core has overall  $C$  values lower than those of Little Lake (median = 0.49 pieces  $\text{cm}^{-2} \text{year}^{-1}$ ). Over its 5000-year record,  $C_{\text{back}}$  varies between 0.06 and 1.13 pieces  $\text{cm}^{-2} \text{year}^{-1}$ , and vegetation was consistently dominated by Engelmann spruce and subalpine fir. Finally, Ruppert Lake (Higuera *et al.* 2009) is located in boreal forest of Alaska's south-central Brooks Range. The 3-ha, 7.0-m-deep lake has an ~200-ha watershed with subdued topography and a small inflowing stream. The 4.8-m core has the lowest overall  $C$  values of all three records (median = 0.04 pieces  $\text{cm}^{-2} \text{year}^{-1}$ ). Over the 14 000-year record,  $C_{\text{back}}$  ranges from 0.00 to 0.22 pieces  $\text{cm}^{-2} \text{year}^{-1}$  with a distinct increase ~5000 years BP, coincident with the transition from a forest-tundra to boreal forest vegetation. Overall, five different vegetation types dominated the landscape around Ruppert Lake during the record.

For all records, we used the published age–depth relationship but reanalysed each series using the published resampling intervals of 10, 10 and 15 years for Little, Rockslide and Ruppert lakes respectively. We did not use the same analysis parameters as in the published records, because our purpose was to test different parameters. We calculated background charcoal using a locally weighted regression robust to outliers (lowess) in a 500-year window. The robust lowess model is less sensitive to non-stationarity and thus may be applied to raw and transformed data (Cleveland 1979).

#### *Data transformation, peak identification, and sensitivity analysis*

We applied the four different detrending models to each simulated and empirical record, and we used a modified Levene's test of equal variance (based on sample medians; Brown and Forsythe 1974) to test the null hypothesis of equal variance between two portions of each record. Sample sizes for  $P$  value calculations were adjusted to account for temporal autocorrelation in each record following Bretherton *et al.* (1999). For simulated records, we compared the periods 10 000–6000 and 4000–0 years BP. We present only the median test result for 500 realisations of the simulated series.<sup>C</sup> For empirical records, we subjectively selected periods during which background charcoal had two qualitatively different levels and then divided this period in half for comparison. At Little, Rockslide and Ruppert lakes, these periods corresponded to the last 8000, 5000 and 10 000 years respectively. The test statistic,  $W_{50}$ , is used as an index of heteroscedasticity, and the associated  $P$ -value is used to assess the null hypothesis of equal variance.

We identified peaks in simulated and empirical records using a Gaussian mixture model that models the noise distribution within  $C_{\text{peak}}$  (described earlier). In this application, the value of the mixture model is its ability to apply uniform treatments to all records, making specific threshold-selection parameters of less importance. For all analyses, we used the 99th percentile of the modelled noise distribution as the threshold value.<sup>D</sup> Thresholds were defined both globally (a single mixture model fit to the entire record) or locally (fitting the mixture models to 500-year windows centred on each sample, and then smoothing the series of resulting threshold values).

For the simulated records, we quantified the sensitivity of peak identification to the four detrending models with a sensitivity ratio,  $s$ . We defined  $s$  as the number of peaks detected in the first half of each record divided by the total number of peaks detected in the second half of each record. If an analytical method is insensitive to variations in  $C_{\text{back}}$ , then  $s$  will equal one. Values of  $s$  significantly greater or less than one indicate a systematic bias in the set of analytical methods. We used a Monte Carlo approach to estimate the value of  $s$  for each of the 16 analysis combinations (2 simulation scenarios  $\times$  4 detrending models  $\times$  2 threshold-determination techniques = 16). For each combination,  $s$  was estimated by the average  $s$  from 500 simulations, and the 2.5th and 97.5th percentiles were used to estimate 95% confidence intervals around  $s$ . If the 95% confidence intervals overlapped one, then the ratio was considered no different from one and the method was deemed insensitive to the variation in background charcoal.

We performed two additional analyses on the empirical records. First, we explored the effect of the four detrending models on the capacity of the Gaussian mixture model to identify a distinct noise distribution. For simplicity, we chose to use only a globally fitted model applied to the Rockslide Lake record, the least variable record; similar examples could be based on subsections of other records. As a metric of how distinct peaks were from  $C_{\text{back}}$ , we examined the SNI (defined earlier) of the fitted Gaussian mixture model. Second, to illustrate the impact of the eight alternative decomposition methods and the minimum-counts test, we applied each method to the empirical records. We quantified the percentage of peaks that fail to pass the minimum count test under each decomposition method. To illustrate how interpretations may differ, we summarised peaks (after removing those failing to pass the minimum-count test) with 1000-year smoothed peak frequency curves (peaks per 1000 years, smoothed to 1000 years with a lowess filter).

## Results

### *Simulated records*

As designed, simulated charcoal records from Scenario 1 were homoscedastic (500-sample median  $W_{50} = 0.45$ , median  $P = 0.502$ ), whereas records from Scenario 2 were heteroscedastic ( $W_{50} = 21.87$ ,  $P < 0.001$ ; Table 4, Fig. 2). For both scenarios, the choice of decomposition method had a major effect on the variability in the resulting peak series,  $C_{\text{peak}}$  (Table 4, Fig. 2). Under Scenario 1, only the NR model resulted in a

<sup>C</sup>Results did not differ when analysing 250, 500 or 1000 realisations (each 10 000 years long), suggesting that the inherent variability was captured.

<sup>D</sup>Note that the exact threshold criterion used here has no consequence on our interpretations, because interpretations are based on relative changes across a record. For example, analysis using the 95th percentile produced identical patterns.

**Table 4. Stationarity of variance and skewness of  $C$  and  $C_{peak}$  series for different decomposition models**

The modified Levene's test statistic,  $W_{50}$ , and the probability of the null hypothesis of equal variances,  $P$ , are based on comparisons between values from 10 000 to 6000 and 4000 to 0 years BP in simulated records, and equally split halves since 8000, 5000 and 10 000 years BP for Little, Rockslide, and Ruppert lakes respectively.  $C$  refers to the interpolated charcoal series; see Table 1 for abbreviations of the specific decomposition models. Bold and italic values respectively identify stationary series, those that fail to reject the null hypothesis at  $\alpha = 0.10$  and  $0.05$ , where a higher  $\alpha$  is more conservative. The skewness coefficient is a measure of the asymmetry of the entire peak series of each respective model, where positive values indicate greater spread above the mean value and a 0 value indicates a symmetric distribution. The time series for each model is shown in Figs 2 and 4 for the simulated and empirical records respectively. Values for simulated records represent the median value from 500 records constructed under each scenario

Scenario or site	$W_{50}$ test statistic for equality of variances ( $P$ value)						Skewness coefficient (2.5–97.5th percentile) [within-row rank]						
	$C$	NR	TR	NI	TI	NR	TR	NI	TI	NR	TR	NI	TI
Scenario 1 (variance constant)	<b>0.45 (0.502)</b>	<b>0.44 (0.508)</b>	46.72 (<0.001)	25.62 (<0.001)	98.54 (<0.001)	2.59 (1.98–4.17) [1]	0.96 (0.68–1.29) [2]	3.47 (2.40–6.14) [1]	1.22 (0.79–1.74) [2]				
Scenario 2 (variance proportional)	21.87 (<0.001)	20.90 (<0.001)	7.12 (0.008)	10.81 (0.001)	57.92 (<0.001)	2.97 (2.12–5.05) [1]	1.20 (0.89–1.52) [2]	3.03 (2.17–4.98) [1]	1.55 (1.01–2.10) [2]				
Little Lake	153.14 (<0.001)	37.72 (<0.001)	3.40 (0.066)	<b>0.01 (0.942)</b>	68.79 (<0.001)	17.73 [1]	0.28 [3]	4.17 [2]	–0.56 [4]				
Rockslide Lake	15.20 (<0.001)	5.71 (0.018)	<b>0.49 (0.483)</b>	6.20 (0.013)	14.70 (<0.001)	3.16 [3]	1.68 [4]	5.28 [1]	3.70 [2]				
Ruppert Lake	84.62 (<0.001)	59.51 (<0.001)	66.52 (<0.001)	5.62 (0.018)	9.11 (0.003)	4.10 [3]	3.14 [4]	6.79 [1]	6.37 [2]				

stationary series ( $W_{50} = 0.44, P = 0.508$ ; Fig. 2; Table 4). The TR and NI models greatly inflated variance when background charcoal was low ( $W_{50} = 46.72, 25.62; P < 0.001$ ), and the TI model further inflated variance ( $W_{50} = 98.54; P < 0.001$ ). No model stabilised variance in records from Scenario 2 (Table 4, Fig. 2). The NR model preserved heteroscedasticity in the original record ( $W_{50} = 21.81$ ), the TR and NI models reduced heteroscedasticity ( $W_{50} = 7.12, 10.81$ ), whereas the TI model increased heteroscedasticity ( $W_{50} = 57.92$ ). The skewness of  $C_{peak}$  also varied greatly among models. For both Scenario 1 and Scenario 2, the NI model produced the most skewed peak series (3.47 and 3.03), followed closely by the NR model (2.59, 2.97), and then the TI (1.21, 1.55) and TR (0.95, 1.20) models (Table 4, Fig. 2).

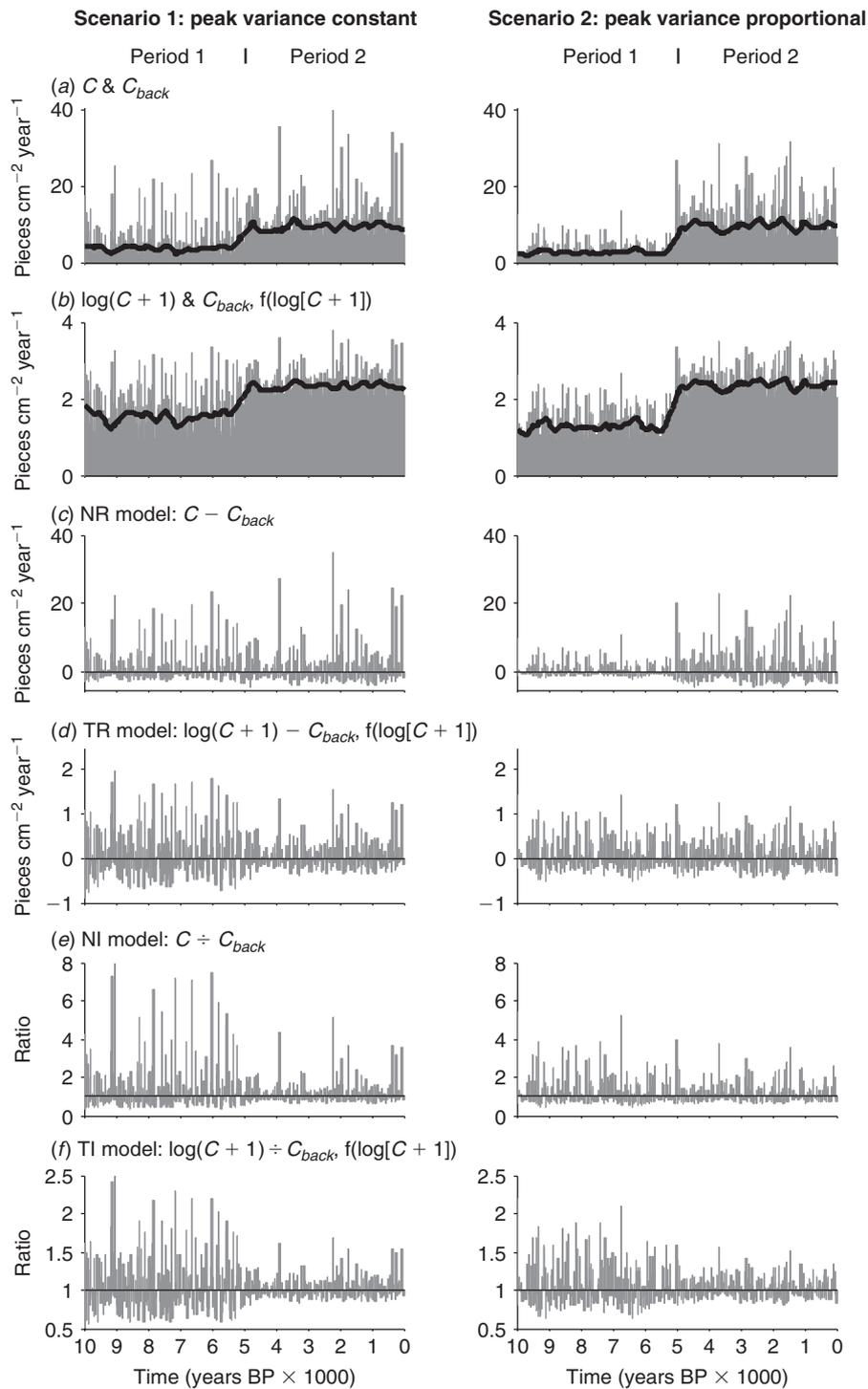
In simulated records, threshold type was more important than detrending model when evaluating the sensitivity of peak identification to changes in variance (Fig. 3). Locally defined thresholds were insensitive to the presence of heteroscedasticity (Scenario 2 v. Scenario 1) and detrending model ( $s$  for all scenarios did not differ from 1). In contrast, using a globally defined threshold produced unbiased results only under three conditions. When  $C$  was characterised by constant variance (Scenario 1), a globally defined threshold was unbiased when  $C_{peak}$  was defined by residuals: median  $s$  for NR and TR models was 1.00 (95% CI 0.76–1.33) and 1.17 (0.92–1.54) respectively. Using an index to define  $C_{peak}$  inflated variance when  $C_{back}$  was low (Fig. 2), resulting in 1.86–2.55 times the number of detected peaks:  $s$  for NI and TI models was 1.86 (1.24–3.00) and 2.55 (1.50–4.56) respectively. When variance in  $C$  increased with  $C_{back}$  (Scenario 2), transforming  $C$  and using residuals produced unbiased results, as did creating an index from the non-transformed series:  $s$  for TR and NI models was 0.85 (0.63–1.10) and 1.33 (0.98–1.79) respectively. Defining  $C_{peak}$  as the residuals of non-transformed  $C$  (NR model) resulted in more peaks detected when  $C_{back}$  and variability was high ( $s = 0.63$  (0.43–0.87)), and transforming and using an index to define  $C_{peak}$  (TI model) resulted in nearly twice as many peaks detected when  $C_{back}$  and variability were low ( $s = 1.70$  (1.15–2.63)).

Overall, analyses of simulated records illustrate the sensitivity of decomposition methods to changes in the mean (Scenario 1) and changes in the mean and variance (Scenario 2) of a series through time. Although simplified, the sensitivity of the simulated records to the analytical method highlights biases that can arise from similar changes in empirical records, even when of smaller magnitude, duration or both.

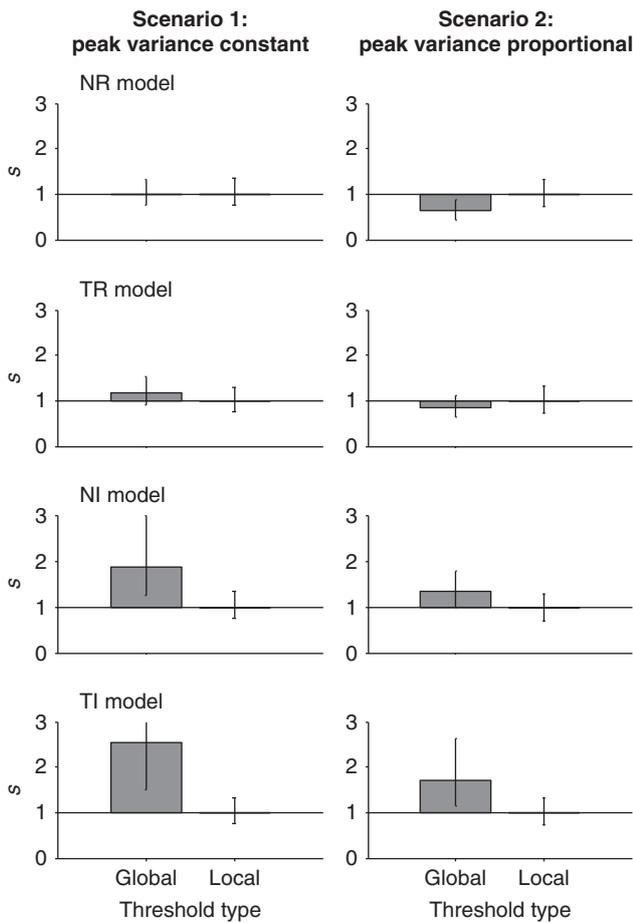
### Empirical records

The three empirical records differed greatly in their long-term variability in  $C$  (Fig. 4). Little Lake had relatively low charcoal values until ~4000 years BP, when sediment accumulation rate increased five-fold (0.07 to 0.35 cm year<sup>-1</sup>) in parallel with  $C$ . At Rockslide Lake, sediment accumulation rates varied approximately two-fold (0.04 to 0.08 cm year<sup>-1</sup>) and were largely independent of  $C$ . At Ruppert Lake, sediment accumulation rates varied nearly ten-fold (0.018 to 0.181 cm year<sup>-1</sup>) and although  $C$  followed the sediment accumulation rate during the first few millennia, these variables were unrelated for the majority of the record.

The raw records ( $C$ ) at each site exhibited significant heteroscedasticity (Table 4), with Little Lake exhibiting the



**Fig. 2.** Simulated charcoal records reflecting alternative assumptions regarding the stability of the variance through time. (a) Representative records from each scenario. Scenario 1 has constant variance in peak heights superimposed on a changing mean. Scenario 2 is a heteroscedastic series in which the peak variance changes in proportion to  $C_{back}$ . The thick black line in all figures is a 500-year lowess smooth used to define the 'background' levels ( $C_{back}$ ). (b) Series expressed on a log scale. (c–f) Detrended series based on four alternative methods. Abbreviations:  $C$ , interpolated charcoal series;  $C_{back}$ , background charcoal series; NR, non-transform, residuals; TR, transform, residuals; NI, non-transform, index; TI, transform, index.



**Fig. 3.** Sensitivity of peak identification to decomposition models and threshold type. The sensitivity index,  $s$ , is the ratio of detected peaks from period 1 to period 2 in the two simulated charcoal scenarios in Fig. 2. The error bars indicate the 95% confidence interval from 500 realisations of the simulated records.

most, followed by Ruppert and Rockslide lakes ( $W_{50} = 153.14, 84.62, 15.20$  respectively;  $P \leq 0.001$ ). The four detrending models had a large effect on the variance in  $C_{peak}$  (Table 4; Fig. 4). The TR and NI models were most effective at stabilising variance, although results differed between sites. At Rockslide Lake, the TR model stabilised variance (500–2500 v. 2500–0 years BP,  $W_{50} = 0.49, P = 0.483$ ); this model performed second best at Little Lake (800–4000 v. 4000–0 years BP,  $W_{50} = 3.40, P = 0.066$ ) and performed worst at Ruppert Lake (10 000–5000 v. 5000–0 years BP,  $W_{50} = 66.52, P < 0.001$ ). At Little Lake, the NI model stabilised variance ( $W_{50} = 0.01, P = 0.942$ ), and at Ruppert Lake, no model stabilised variance. Although the NR and TI models reduced heteroscedasticity, they did not stabilise variance in any record. Skewness in  $C_{peak}$  was largest when using the NR model (Little Lake, 17.73) or NI model (Rockslide and Ruppert lakes, 5.28, 6.79 respectively). In contrast, log-transforming  $C$  (TR and TI models) reduced skewness, and at Little Lake, these models resulted in near-symmetric distributions (0.28 and  $-0.56$  respectively; Table 4).

The noise distribution fitted by the Gaussian mixture model resulted in different SNI and skewness values, dependent on the

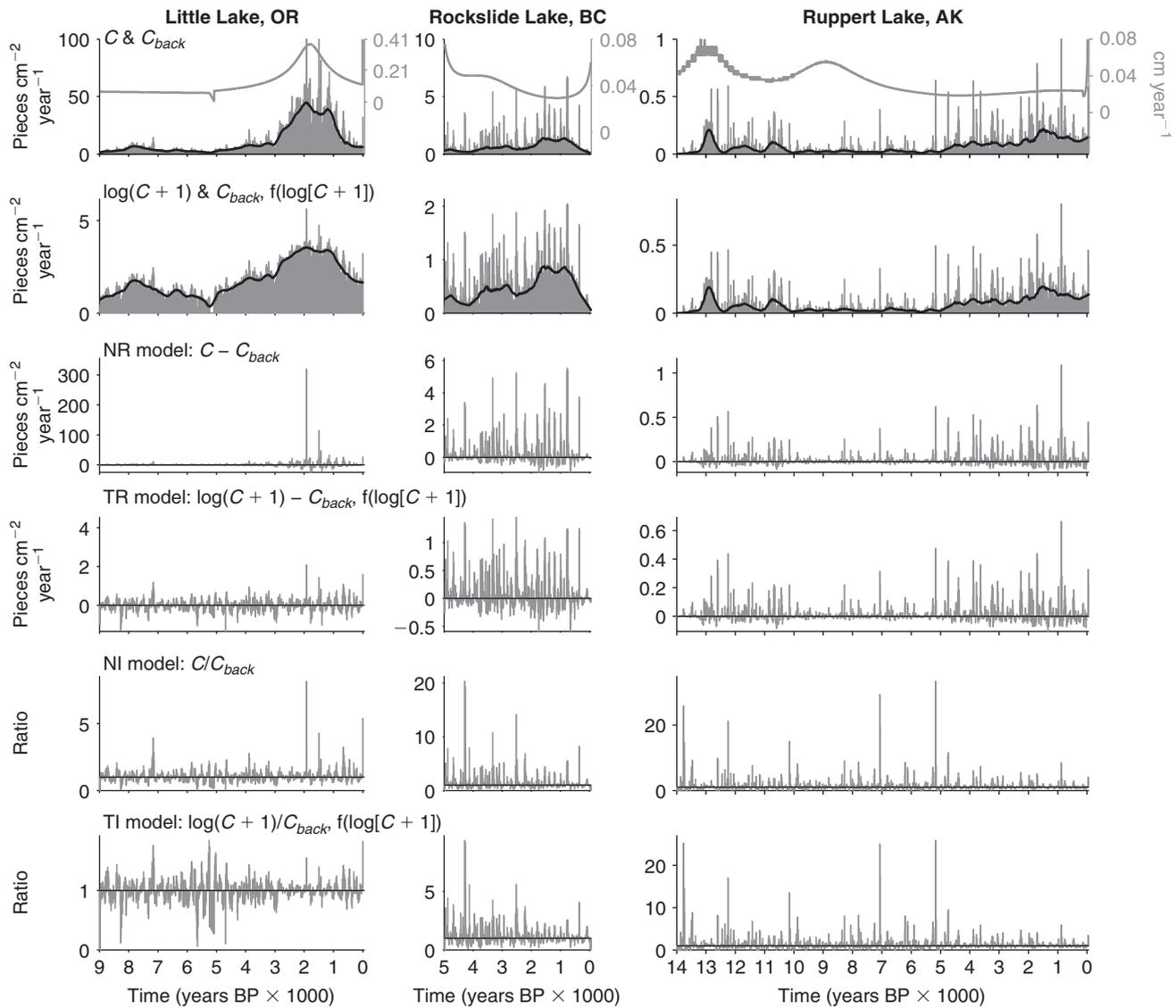
decomposition model (Fig. 5). As applied to the Rockslide Lake record, the NR and NI models yielded the largest SNI (0.97 and 0.98), whereas the TR and TI models had the smallest SNI (0.81 and 0.94). Skewness, as a potential measure of the occurrence of high values distinct from a noise distribution, was highest for the NI model (5.28). The TR model, though having a moderate SNI, was the most symmetric (skewness = 1.68).

As with simulated records, peak detection in empirical records was more sensitive to alternative decomposition models when using a global v. local threshold, and this sensitivity varied greatly between sites (Fig. 6). At Little Lake, where  $C_{back}$  varied the most throughout the record, a global threshold detected 41 peaks with the NR model but only five with the TI model, producing drastically different trends in 1000-year mean fire-event frequency. The TR and NI models produced an intermediate number of peaks (23 and 30) with qualitatively similar trends over time. In contrast to the variance at Little Lake, peak detection with any model varied by 6% at Rockslide Lake (33–35) and 13% at Ruppert Lake (64–72), where variability in  $C_{back}$  was less. At all sites, locally defined thresholds detected more peaks and with less variability between models than did globally defined thresholds, even after minimum-count screening (described below). Again, differences among models were greatest at Little Lake, where peak detection varied from 56 to 68 (21%). Peak detection varied little at Rockslide Lake, 34–36 (6%), and slightly more at Ruppert Lake, 79–88 (11%). Differences at Little and Ruppert lakes largely reflect differences between the two residual models (NR and TR) in comparison with the index models (NI and TI).

The minimum-count screening flagged between 0 and 14% of the total peaks detected in any one record, with the least at Little Lake (median = 2.5%), followed by Rockslide Lake (median = 9%) and Ruppert Lake (median = 11%). A greater proportion of the total peaks detected was flagged when using a local v. global threshold (median = 9 v. 3%), and the variability between detrending models differed by threshold type. When using a global threshold, a larger percentage of total peaks was flagged when using index models (median for NI and TI = 11%) v. residual models (median for NR and TR = 3%). This difference was reduced using a local threshold (10 v. 9% respectively).

### Discussion

Interpreting local fire history from sediment–charcoal records involves several analytical steps that decompose multiple signals into a series of peaks that bears interpretation (Fig. 1). Accounting for non-stationarity in a record is a primary goal of decomposition methods, and our analyses of simulated records illustrate the sensitivity of alternative methods to two types of non-stationarity: a change in the mean (Scenario 1), and a change in the mean and variance (Scenario 2) through time. In combination with empirical records, our results highlight some critical methodological considerations that have been broadly overlooked in the literature. Specifically, we emphasise the need for careful consideration when proceeding through steps 4–6 of a decomposition method (Fig. 1): (i) defining  $C_{peak}$ , or detrending; (ii) defining a threshold to detect peaks; and, in the case of charcoal counts; (iii) screening and removing peaks that could result from insignificant changes in charcoal counts.

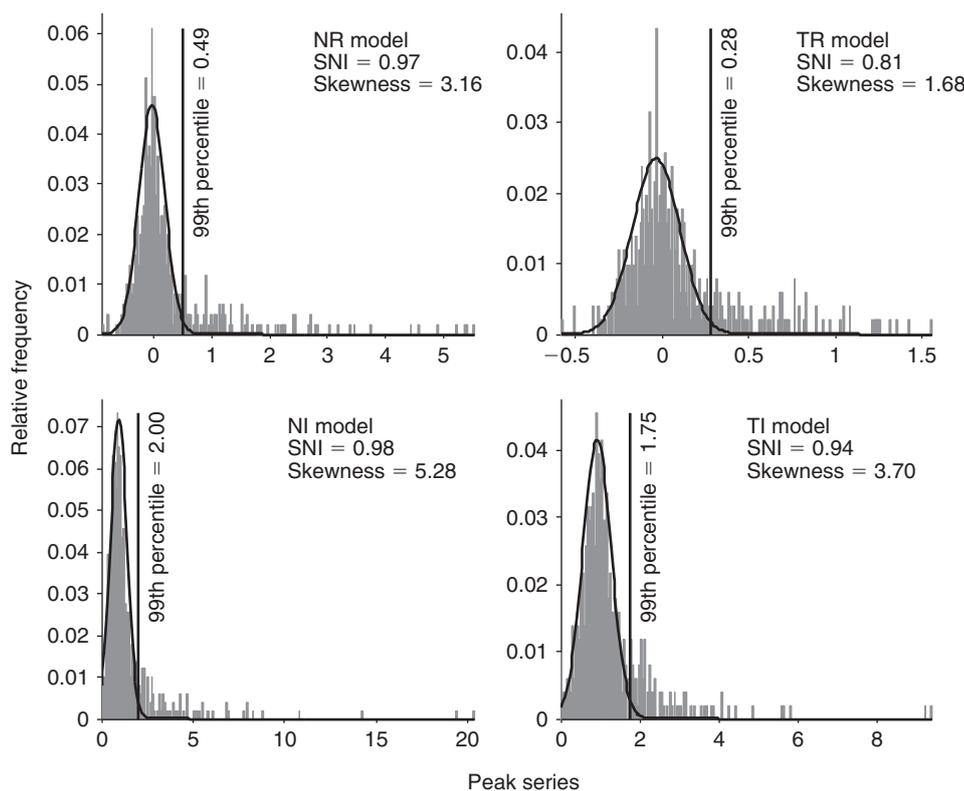


**Fig. 4.** Empirical charcoal records detrended using the four decomposition models. Records are from western Oregon (Little Lake; Long *et al.* 1998), south-east British Columbia (Rockslide Lake; Gavin *et al.* 2006), and the Alaskan Brooks Range (Ruppert Lake; Higuera *et al.* 2009). See Fig. 2 for abbreviations.

#### *Detrending to define a peak series*

An overriding conclusion from our study is that the impacts of different detrending models are largely obviated by using a locally defined threshold. In simulated records, peak identification using a local threshold was robust to changes in background charcoal, peak variance and detrending model (Fig. 3). In empirical records, these patterns largely held true, as reflected by less between-model variability when using local *v.* global thresholds. For example, the total number of peaks detected since 5000 years BP in Ruppert Lake varied by 7 *v.* 30% when applying a local *v.* global threshold to the different detrending models (Fig. 6). Locally defined thresholds outperformed global thresholds because the mixture model used to determine thresholds constantly adapts to variability in a record. Consequently, local thresholds are free from the assumption of stable

variance in peak heights, at least for time scales longer than the window width used to define 'local'. With no need to stabilise variance across a record, detrending before applying a locally defined threshold needs only to account for changes in the long-term mean, and thus three of the four detrending models evaluated become obsolete. Even stabilising the mean, interestingly, may be unnecessary when using a locally defined threshold, but no studies have attempted this to date. Our results are consistent with analyses done by Ali *et al.* (2009b), who applied the local threshold technique to three different charcoal quantification metrics from individual cores (counts, area and estimated volume). Although the variability between the three metrics differed, the locally applied threshold produced similar results in each case. Overall, these findings lend support to the recent adoption of local thresholds for peak identification



**Fig. 5.** The Gaussian mixture model applied to the  $C_{peak}$  series from Rockslide Lake. Each panel corresponds to a single detrending model. The Gaussian model representing the noise distribution is shown by a thick black line. The vertical line represents a typical threshold level for peak identification, located at the 99th percentile of the noise distribution. (SNI, signal-to-noise index.)

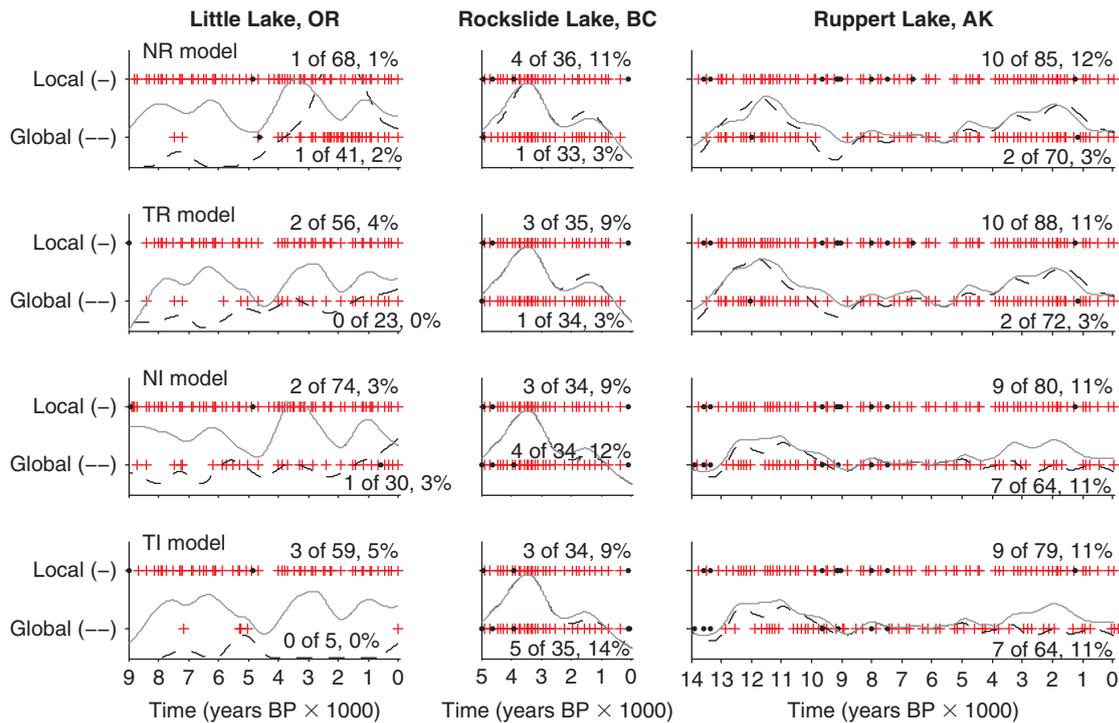
(Table 1) as robust to changes in variance both within and between records.

Although local thresholds effectively eliminate the need to stabilise variance, understanding the impacts of detrending models when combined with global thresholds remains important, mainly owing to the prior use of these approaches (Table 1). Our results suggest that reanalysis of some previously published records is justified, as has been initiated in some larger-scale synthesis studies (Marlon *et al.* 2009). In particular, analyses using a global threshold and the NR model with clearly heteroscedastic records or a global threshold with the TI model should be reconsidered, given the potential for systematically biased peak detection during periods of high or low  $C_{back}$ .

When applying a global threshold, it is imperative to evaluate the presence or absence of heteroscedasticity in a record before selecting a detrending model. If a record has stable variance, then the NR model is the single appropriate model because it removes only the mean trend of a series (Fig. 2). In simulated records, the only instance in which the global threshold was unbiased was when the NR model was applied to homoscedastic records (Scenario 1; Fig. 3). The closest analogy in the empirical records is from Rockslide Lake, which had the least heteroscedasticity of the three records evaluated (and was the shortest in length) and consequently was most robust to alternative detrending methods. Applying variance-stabilising models (TR, NI, TI) to homoscedastic records is not only unwarranted, but it can result in

severely biased peak identification (Fig. 3) by simultaneously amplifying and suppressing peaks during periods of low and high background charcoal accumulation rates respectively (Fig. 2). This bias was minimised when using the TR model, and it subsequently increased with the NI and TI models. By amplifying peak sizes when  $C_{back}$  is low, index-based models applied to homoscedastic records result in biased peak identification (Figs 2, 6). This bias is most extreme when using the TI model in combination with a global threshold, as this led to more than twice as many peaks being detected during periods of low *v.* high background charcoal in our simulations (Fig. 3).

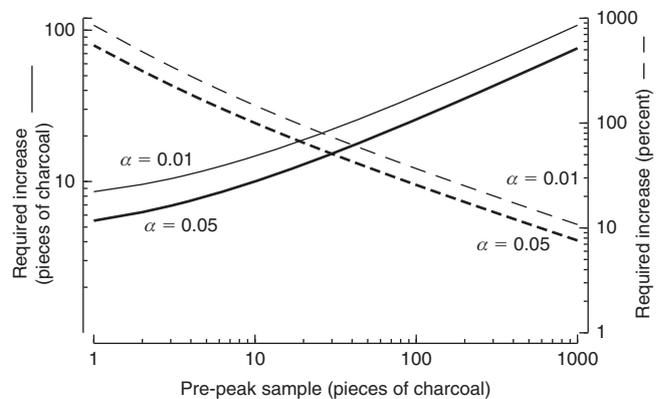
Unfortunately, most empirical charcoal records exhibit heteroscedasticity at some time scale, particularly those spanning different biomes, many millennia or both. This limits the utility of the NR model with a global threshold. All the empirical records in this study, for example, had non-stable variance between the two periods of comparison (Table 4; Fig. 4). In heteroscedastic records, both empirical and simulated records support the TR or NI models as most appropriate for consideration when using a global threshold. Although no model stabilised variance in the simulated records with heteroscedasticity, the TR and NI performed the best, and in empirical records, these models stabilised variance across comparison periods in some records (Table 4). When applied to heteroscedastic records, the NR and TI models are inappropriate for the opposite reasons. Simply detrending by residuals (NR) fails to remove



**Fig. 6.** Inferred fire history from Little, Rockside and Ruppert lakes using alternative decomposition methods. Each row corresponds to a different detrending model, as in Fig. 4, and each panel includes peaks detected based on a global and local threshold. The location of peaks exceeding the threshold value(s) are identified with ‘+’ and ‘.’ symbols, where the latter identifies peaks that failed to pass the minimum-count test. The proportion of total peaks failing to pass the minimum-count test is displayed on the right-hand side of each panel. Smoothed lines represent the 1000-year mean fire-event frequency for a given decomposition method, and all panels are scaled from 0 to 15 (fire-events per 1000 years) on the y-axis.

any heteroscedasticity (Table 4), which biases peak identification towards periods of high  $C_{back}$  (Fig. 3). Detrending with an index of transformed data (TI) reverses the pattern of heteroscedasticity in what is essentially a ‘double whammy’ of variance stabilisation (Figs 2, 4), biasing peak identification towards periods of low  $C_{back}$  (Fig. 3). The undesirable effects of the TI model are most apparent in the Little Lake record, where four of the five peaks detected occurred during the period of low background charcoal (Fig. 6). The overall low number of peaks detected in this scenario also stands out as odd, and it highlights the conceptual difficulty of interpreting fire history from a symmetric peak series. If the variability above  $C_{back}$  does not differ from the variability below  $C_{back}$ , then it is inconsistent to interpret the former as fire-related while interpreting the latter as noise-related. When a globally defined Gaussian mixture model is applied to a nearly symmetric peak series (e.g. Little Lake under the TI model, skewness =  $-0.56$ ; Table 4), selecting a threshold at the 99th percentile cuts off 99% of the samples in the series (895 of 900 samples in the Little Lake record; Fig. 6).

Finally, we emphasise that the impacts of different detrending models will vary between sites, depending on the mean and variability of charcoal accumulation rates in a record. Rockside Lake, for example, was largely robust to alternative decomposition methods, whereas Little Lake displayed large variability between methods (Fig. 7). Ruppert Lake was also clearly heteroscedastic, but overall lower  $C$  and  $C_{back}$  values as compared with Little Lake resulted in less sensitivity to alternative detrending models (Fig. 6).



**Fig. 7.** Minimum increase in charcoal counts required to confidently separate pre-peak from peak samples. The required increase is displayed as a total number and a proportion, and it depends on (1) the confidence desired ( $\alpha = 0.01$ , 99%, or  $\alpha = 0.05$ , 95% confidence), and (2) the number of pieces in the smaller, pre-peak sample. The curves are developed from the test for assessing whether two samples are from the same Poisson distribution (Detre and White 1970). Lines for two significance levels are shown and presented both as absolute counts and percentage increase of the lower count.

### Defining a threshold

The Gaussian mixture model introduced by Gavin *et al.* (2006) is promising because it provides a semi-objective, process-based means of selecting a threshold for peak identification, which in

turn can be applied to multiple records. Using the mixture model to identify a threshold depends on three key assumptions: (1) variation in the noise distribution, representing variability around the long-term trend (i.e.  $C_{back}$ ), is normally distributed; (2) the mean and variance of this noise distribution is stationary within the period of analysis; and (3) there are enough samples within the period of analysis to adequately characterise the noise distribution. The first assumption has theoretical support from a charcoal simulation model (fig. 3 in Higuera *et al.* 2007), and it is consistent with distributions of peak charcoal observed in empirical records (e.g. Higuera *et al.* 2009; Fig. 5). The mechanisms creating normally distributed variability around the long-term trend include sediment mixing, interannual variability in long-distance charcoal input, sampling effects, and analytical error. Other mechanisms may produce skewed variability, and to the extent that this is true, this is a limitation of the Gaussian mixture model (discussed below).

The second assumption, that the properties of the noise distribution are stable, becomes increasingly difficult to satisfy as more samples are included in the population. The two ways to satisfy this assumption are to define the threshold over a period of stable mean and variance (i.e. use a locally defined threshold), or define  $C_{peak}$  with one of the two recommended variance-stabilising methods (TR, NI). The shorter the period over which a threshold is defined, the more difficult it becomes to satisfy the third assumption, that the Gaussian distribution adequately describes the empirical data. Thus, the analyst has to make a trade-off between satisfying assumptions two and three. In practice, one can test the third assumption with a goodness-of-fit statistic, which quantifies the probability that the empirical data came from the modelled Gaussian distribution (e.g. Higuera *et al.* 2009). The modified Levene's test used in the present study can be used to test for equal variance between different periods in a peak series. Future application of this test could be done on shorter, overlapping intervals, although one faces reduced statistical power as the intervals decrease, and interpreting  $P$  values becomes difficult with multiple comparisons.

The application of the Gaussian mixture model to identify a threshold is also aided by maximising the separation between the noise distribution and fire-related peaks. This is a key difference between the analytical approach taken for peak identification compared with the analysis of long-term trends in total charcoal (e.g. Marlon *et al.* 2008, 2009; Power *et al.* 2008). Whereas homogenising variance is desirable in the context of the latter, this decreases separation between noise and fire-related samples, i.e. it reduced the SNI. For example, in the Rockslide Lake record, the SNI was highest (0.97–0.99) using the NR and NI models, whereas it was consistently lowered when applying variance-stabilising transformations (0.81–0.94; Fig. 5). Skewness may also serve as a coarse index of how separated peak values are from non-peak values. At Little Lake, a nearly symmetric peak distribution defined by the TI model reflected little to no separation between peak and non-peak values. Thus, as a general rule, a minimum level of skewness of approximately two would suggest a SNI sufficient to aid in setting thresholds, but increased skewness beyond two does not necessarily equate to an increased SNI. We also note that skewness alone is not justification for peak interpretation, particularly if it is an artefact of the detrending processes.

### Interpreting small charcoal peaks

Most decomposition methods resulted in the identification of small peaks that failed to pass the minimum-count test at the 95% confidence level (0–14% of the total peaks identified; Fig. 6). Some peaks fail to pass this test because they closely follow other large peaks (i.e. a 'double peak'), e.g. ~1200 years BP at Ruppert Lake (Figs 4, 6). These peaks most likely represent non-significant variations in charcoal counts due to natural or analytical variability. More challenging for sediment-based fire-history reconstructions are periods of low charcoal abundance. In these cases, both variance-stabilising and local-threshold methods may result in detecting small peaks, often associated with small charcoal counts in the raw record. The smaller the charcoal peak, the more difficult it is to infer if the peak was caused by a local fire *v.* a distant fire or random variability in charcoal deposition and quantification. The minimum-count test helps guard against falsely inferring that a peak was caused by a local fire (Type I error). When this probability is low, e.g.  $<0.05$ , it is highly unlikely that the two samples come from the same population. Practically, Fig. 7 illustrates the increase in counts (as a proportion and absolute number) required to achieve a given level of confidence (95 or 99%) as a function of the number of charcoal pieces in the pre-peak sample. The lower the pre-peak count, the greater the proportional increase in charcoal required before a peak sample can be considered distinct with 95% confidence. For example, when pre-peak counts are  $<10$ , peak counts must double before having a  $<5\%$  chance of coming from the same population. Much smaller proportional increases are required when overall counts are large, e.g. only a 20% increase is required when pre-peak samples are  $\sim 100$ . Thus, as a rough rule of thumb, it is highly desirable for researchers to use sample volumes that will result in average non-peak samples of  $>10$  pieces, and peak values of *at least* 20 pieces.

Even with large sample volume, in some cases the difference between a peak and non-peak sample may be small. Interpreting variability when charcoal counts are low highlights a limitation of the Gaussian mixture model briefly mentioned above. The mixture model assumes normally distributed noise, and thus it may fail when counts are small, because the true noise distribution may be positively skewed (i.e. Poisson distributions with a mean  $<10$  are positively skewed). If so, the Gaussian mixture model would underestimate the threshold, resulting in an increased false-positive rate. Future efforts modelling noise distributions within  $C_{peak}$  could address this limitation through the use of non-Gaussian mixture models, which may be more appropriate for the heavy-tailed distributions that characterise  $C$  and  $C_{peak}$  series (Coles 2001). For example, the signal and noise distributions may be better represented by models in the generalised extreme value family (e.g. Weibull, Fréchet and Gumbel distributions), and the signal distribution may be represented more appropriately by models in the generalised Pareto family (e.g. Pareto,  $\beta$  and exponential distributions; Katz *et al.* 2005). In the ideal case, the signal distribution has little influence on the parameters of the noise distributions, because the noise population typically dominates the mixed distribution. Nonetheless, improving the fit of the signal distribution would be an improvement over current methods and deserves exploration. In the meantime, we suggest that the minimum-count test serves well to screen out small peaks, be they detected with a threshold from a Gaussian mixture model or otherwise.

### Recommendations and conclusions

Even when applying the most rigorous analytical techniques, there is no substitute for careful inspection of a record to assess whether it can provide an unbiased fire history in the first place. We highlight three key issues related to assessing the quality of millennial-scale charcoal records when independent evidence supporting a particular reconstruction is lacking. First, records should be interpreted in the context of a null hypothesis of random variability. If a peak series lacks large values, is symmetric, or fails to detect recent fires, then the record should be considered too noisy for peak identification. The SNI utilised here is intended to help evaluate this null hypothesis, and we refer readers to recent work that has improved the application of this metric for this purpose (Kelly *et al.* 2010). Records with low SNI value(s) or symmetric peak series should either forgo peak identification methods, or be presented with a low, medium and high range of possible thresholds. Depending on the cause of a low SNI, these records may still be appropriate and valuable for interpreting trends in biomass burning through interpretations of  $C$ ,  $C_{back}$  or both.

Second, if a record has large variability in sediment accumulation rates, practitioners must consider the possibility that changing peak frequencies result from changes in sample resolution. Resampling a record to the median or maximum deposition time per sample coarsens or falsely increases the resolution during periods of high or low sediment accumulation, and thus should create a more temporally unbiased time series. However, this resampling may not be a simple solution if sedimentation varies widely, because sediment mixing modifies the effect of changing sediment accumulation rate on the effective resolution of a sediment record. For example, mixing the top 2 cm of sediment during a period when the deposition rate is  $10 \text{ year cm}^{-1}$  would result in an effective 1-cm resolution of 20 years. In contrast, the same 2-cm mixing depth under a sediment deposition rate of  $20 \text{ year cm}^{-1}$  would result in an effective resolution of 40 years. An important cause of changing sediment accumulation rate is fluctuating within-basin sediment focussing and sediment delivery by stream flow. Such processes can change the effectiveness of sediment delivery, including charcoal, from lake margins to the lake centre. This results in the widely observed positive correlation between charcoal accumulation and sediment accumulation rates (i.e. constant charcoal concentration despite changing sediment accumulation rates) and a heteroscedastic charcoal record (e.g. Fig. 4). In contrast, if charcoal were delivered entirely through airfall, increased sediment accumulation rates would dilute charcoal concentrations, and the charcoal record would not be heteroscedastic. Therefore, we strongly warn against interpreting fire frequency changes in records with a several-fold change in sediment accumulation rate (along with no evidence that charcoal concentrations are diluted by changing sedimentation) and when inferred fire frequency closely tracks sediment accumulation rates. Although there may be a non-causal relationship between sediment accumulation and fire frequency (e.g. via erosion or climate), this link must be explained with independent evidence if fire history is to be interpreted. A viable alternative in these cases is to interpret only high-resolution segments with constant sediment accumulation rates.

Finally, segments of records with low overall counts must be interpreted with caution. The use of the minimum-count test

presented here can help guide interpretation in these cases, as can independent evidence of fire (e.g. pollen or macrofossils of fire-dependent taxa). Although charcoal records have successfully detected fires in non-forested ecosystems (e.g. savannah and tundra: Duffin *et al.* 2008; Higuera *et al.*, in press), at some point along a fire-intensity spectrum, fires will not produce enough charcoal to create an identifiable peak in a record (Higuera *et al.* 2005; Duffin *et al.* 2008). This may be the case even with ample sample volume. Likewise, as local fire frequency decreases, so too does the frequency of large charcoal peaks; this makes it more difficult to separate the signal of local fires from the noise of long-distance transport and within-lake redeposition.

Following the work of Clark (1988a, 1988c) in detecting fires from charcoal in laminated lake sediments, high-resolution charcoal records have proliferated in the absence of a thorough statistical framework for interpretation. This study is a first attempt to provide such a framework. We conclude from discussion above that applying a local threshold, in conjunction with the minimum-count test, is likely to provide the best interpretation of fire history from high-resolution macroscopic charcoal records. In most cases, the simplest detrending model (NR) is appropriate in this context, but there may be scenarios where the TR or NI variance-stabilisation methods are justified. We emphasise the need for careful consideration when selecting, applying, and interpreting variance-stabilising methods, and we encourage practitioners to evaluate the sensitivity of these choices on fire-history interpretations. Despite the challenges of inferring fire history from sediment charcoal records, significant progress has been made to improve the rigor of analysis and interpretations. In combination with the growing database of high-resolution charcoal records worldwide, charcoal records should continue to contribute uniquely to our understanding of fire regimes, the controls and ecological impacts of fire, and the role of fire in the Earth system.

### Acknowledgements

We thank Doug Sprugel for raising these issues, Colin Long for sharing the Little Lake data, and two anonymous reviewers and the guest editor for helpful comments on earlier versions of this manuscript. Empirical data are available through the Global Paleofire Database (<http://www.ncdc.noaa.gov/paleo/impd/gcd.html>, accessed 29 November 2010) or from the authors upon request. Funding was provided by a National Park Ecological Research Fellowship and the University of Idaho (P. E. Higuera) and grants DEB-02-12917 (D. G. Gavin) and ATM-0714146 (P. J. Bartlein) from the US National Science Foundation.

### References

- Ali AA, Asselin H, Larouche AC, Bergeron Y, Carcaillet C, Richard PJH (2008) Changes in fire regime explain the Holocene rise and fall of *Abies balsamea* in the coniferous forests of western Quebec, Canada. *The Holocene* **18**, 693–703. doi:10.1177/0959683608091780
- Ali AA, Carcaillet C, Bergeron Y (2009a) Long-term fire frequency variability in the eastern Canadian boreal forest: the influences of climate v. local factors. *Global Change Biology* **15**, 1230–1241. doi:10.1111/J.1365-2486.2009.01842.X
- Ali AA, Higuera PE, Bergeron Y, Carcaillet C (2009b) Comparing fire-history interpretations based on area, number and estimated volume of macroscopic charcoal in lake sediments. *Quaternary Research* **72**, 462–468. doi:10.1016/J.YQRES.2009.07.002
- Allen CD, Anderson RS, Jass RB, Toney JL, Baisan CH (2008) Paired charcoal and tree-ring records of high-frequency Holocene fire from two

- New Mexico bog sites. *International Journal of Wildland Fire* **17**, 115–130. doi:10.1071/WF07165
- Anderson RS, Hallett DJ, Berg E, Jass RB, Toney JL, de Fontaine CS, DeVolder A (2006) Holocene development of boreal forests and fire regimes on the Kenai Lowlands of Alaska. *The Holocene* **16**, 791–803. doi:10.1191/0959683606HOL966RP
- Anderson RS, Allen CD, Toney JL, Jass RB, Bair AN (2008) Holocene vegetation and fire regimes in subalpine and mixed conifer forests, southern Rocky Mountains, USA. *International Journal of Wildland Fire* **17**, 96–114. doi:10.1071/WF07028
- Beatty RM, Taylor AH (2009) A 14 000-year sedimentary charcoal record of fire from the northern Sierra Nevada, Lake Tahoe Basin, California, USA. *The Holocene* **19**, 347–358. doi:10.1177/0959683608101386
- Bretherton CS, Widmann M, Dymnikov VP, Wallace JM, Blade I (1999) The effective number of spatial degrees of freedom of a time-varying field. *Journal of Climate* **12**, 1990–2009. doi:10.1175/1520-0442(1999)012<1990:TENOSD>2.0.CO;2
- Briles CE, Whitlock C, Bartlein PJ (2005) Post-glacial vegetation, fire, and climate history of the Siskiyou Mountains, Oregon, USA. *Quaternary Research* **64**, 44–56. doi:10.1016/J.YQRES.2005.03.001
- Briles CE, Whitlock C, Bartlein PJ, Higuera PE (2008) Regional and local controls on post-glacial vegetation and fire in the Siskiyou Mountains, northern California, USA. *Palaeogeography, Palaeoclimatology, Palaeoecology* **265**, 159–169. doi:10.1016/J.PALAEO.2008.05.007
- Brown MB, Forsythe AB (1974) Robust tests for equality of variances. *Journal of the American Statistical Association* **69**, 364–367. doi:10.2307/2285659
- Brunelle A, Anderson RS (2003) Sedimentary charcoal as an indicator of late-Holocene drought in the Sierra Nevada, California, and its relevance to the future. *The Holocene* **13**, 21–28. doi:10.1191/0959683603HL591RP
- Brunelle A, Whitlock C (2003) Post-glacial fire, vegetation, and climate history in the Clearwater Range, northern Idaho, USA. *Quaternary Research* **60**, 307–318. doi:10.1016/J.YQRES.2003.07.009
- Brunelle A, Whitlock C, Bartlein P, Kipfmüller K (2005) Holocene fire and vegetation along environmental gradients in the Northern Rocky Mountains. *Quaternary Science Reviews* **24**, 2281–2300. doi:10.1016/J.QUASCIREV.2004.11.010
- Carcaillet C, Bergeron Y, Richard PJH, Frechette B, Gauthier S, Prairie YT (2001) Change of fire frequency in the eastern Canadian boreal forests during the Holocene: does vegetation composition or climate trigger the fire regime? *Journal of Ecology* **89**, 930–946. doi:10.1111/J.1365-2745.2001.00614.X
- Clark JS (1988a) Effects of climate change on fire regimes in north-western Minnesota. *Nature* **334**, 233–235. doi:10.1038/334233A0
- Clark JS (1988b) Particle motion and the theory of charcoal analysis: source area, transport, deposition, and sampling. *Quaternary Research* **30**, 67–80. doi:10.1016/0033-5894(88)90088-9
- Clark JS (1988c) Stratigraphic charcoal analysis on petrographic thin sections: application to fire history in north-western Minnesota. *Quaternary Research* **30**, 81–91. doi:10.1016/0033-5894(88)90089-0
- Clark JS (1990) Fire and climate change during the last 750 years in north-western Minnesota. *Ecological Monographs* **60**, 135–159. doi:10.2307/1943042
- Clark JS, Patterson WA (1997) Background and local charcoal in sediments: scales of fire evidence in the paleorecord. In 'Sediment Records of Biomass Burning and Global Change'. (Eds JS Clark, H Cachier, JG Goldammer, BJ Stocks) pp. 23–48. (Springer: New York)
- Clark JS, Royall PD (1996) Local and regional sediment charcoal evidence for fire regimes in presettlement north-eastern North America. *Journal of Ecology* **84**, 365–382. doi:10.2307/2261199
- Clark JS, Royall PD, Chumbley C (1996) The role of fire during climate change in an eastern deciduous forest at Devil's Bathtub, New York. *Ecology* **77**, 2148–2166. doi:10.2307/2265709
- Cleveland WS (1979) Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* **74**, 829–836. doi:10.2307/2286407
- Coles S (2001) 'An Introduction to Statistical Modeling of Extreme Values.' (Springer-Verlag: London, UK)
- Cook ER, Peters K (1997) Calculating unbiased tree-ring indices for the study of climatic and environmental change. *The Holocene* **7**, 361–370. doi:10.1177/095968369700700314
- Daniels ML, Anderson RS, Whitlock C (2005) Vegetation and fire history since the Late Pleistocene from the Trinity Mountains, north-western California, USA. *The Holocene* **15**, 1062–1071. doi:10.1191/0959683605HL878RA
- Davis MB, Moeller RE, Ford J (1984) Sediment focusing and pollen influx. In 'Lake Sediments and Environmental History'. (Eds EY Haworth, JWG Lund) pp. 261–293. (University of Minnesota Press: Minneapolis, MN)
- Detre K, White C (1970) Comparison of two Poisson-distributed observations. *Biometrics* **26**, 851–854. doi:10.2307/2528732
- Duffin KI, Gillson L, Willis KJ (2008) Testing the sensitivity of charcoal as an indicator of fire events in savanna environments: quantitative predictions of fire proximity, area and intensity. *The Holocene* **18**, 279–291. doi:10.1177/0959683607086766
- Emerson JD (1983) Mathematical aspects of transformation. In 'Understanding Robust and Exploratory Data Analysis'. (Eds DC Hoaglin, F Mosteller, JW Tukey) pp. 247–282. (Wiley: New York)
- Fowler AM (2009) Variance stabilization revisited: a case for analysis based on data pooling. *Tree-Ring Research* **65**, 129–145. doi:10.3959/2008-7.1
- Gavin DG, Brubaker LB, Lertzman KP (2003) An 1800-year record of the spatial and temporal distribution of fire from the west coast of Vancouver Island, Canada. *Canadian Journal of Forest Research* **33**, 573–586. doi:10.1139/X02-196
- Gavin DG, Hu FS, Lertzman K, Corbett P (2006) Weak climatic control of stand-scale fire history during the late Holocene. *Ecology* **87**, 1722–1732. doi:10.1890/0012-9658(2006)87[1722:WCCOSF]2.0.CO;2
- Gavin DG, Hallett DJ, Hu FS, Lertzman KP, Prichard SJ, Brown KJ, Lynch JA, Bartlein P, Peterson DL (2007) Forest fire and climate change in western North America: insights from sediment charcoal records. *Frontiers in Ecology and the Environment* **5**, 499–506. doi:10.1890/060161
- Giesecke T, Fontana SL (2008) Revisiting pollen accumulation rates from Swedish lake sediments. *The Holocene* **18**, 293–305. doi:10.1177/0959683607086767
- Hallett DJ, Anderson RS (2010) Paleo-fire reconstruction for high-elevation forests in the Sierra Nevada, California, with implications for wildfire synchrony and climate variability in the late Holocene. *Quaternary Research* **73**, 180–190. doi:10.1016/J.YQRES.2009.11.008
- Hallett DJ, Walker RC (2000) Paleoecology and its application to fire and vegetation management in Kootenay National Park, British Columbia. *Journal of Paleolimnology* **24**, 401–414. doi:10.1023/A:1008110804909
- Hallett DJ, Lepofsky DS, Mathewes RW, Lertzman KP (2003a) 11 000 years of fire history and climate in the mountain hemlock rain forests of south-western British Columbia based on sedimentary charcoal. *Canadian Journal of Forest Research* **33**, 292–312. doi:10.1139/X02-177
- Hallett DJ, Mathewes RW, Walker RC (2003b) A 1000-year record of forest fire, drought and lake-level change in south-eastern British Columbia, Canada. *The Holocene* **13**, 751–761. doi:10.1191/0959683603HL660RP
- Higuera PE, Sprugel DG, Brubaker LB (2005) Reconstructing fire regimes with charcoal from small-hollow sediments: a calibration with tree-ring records of fire. *The Holocene* **15**, 238–251. doi:10.1191/0959683605HL789RP
- Higuera PE, Peters ME, Brubaker LB, Gavin DG (2007) Understanding the origin and analysis of sediment-charcoal records with a simulation model. *Quaternary Science Reviews* **26**, 1790–1809. doi:10.1016/J.QUASCIREV.2007.03.010

- Higuera PE, Brubaker LB, Anderson PM, Brown TA, Kennedy AT, Hu FS (2008) Frequent fires in ancient shrub tundra: implications of paleo-records for arctic environmental change. *PLoS ONE* **3**, e0001744. doi:10.1371/JOURNAL.PONE.0001744
- Higuera PE, Barnes JL, Chipman ML, Urban M, Hu FS (in press) Tundra fire history over the past 6000 years in the Noatak National Preserve, north-western Alaska. *Alaska Park Science*.
- Higuera PE, Brubaker LB, Anderson PM, Hu FS, Brown TA (2009) Vegetation mediated the impacts of post-glacial climate change on fire regimes in the south-central Brooks Range, Alaska. *Ecological Monographs* **79**, 201–219. doi:10.1890/07-2019.1
- Huber UM, Markgraf V, Schäbitz F (2004) Geographical and temporal trends in Late Quaternary fire histories of Fuego-Patagonia, South America. *Quaternary Science Reviews* **23**, 1079–1097. doi:10.1016/J.QUASCIREV.2003.11.002
- Huerta MA, Whitlock C, Yale J (2009) Holocene vegetation–fire–climate linkages in northern Yellowstone National Park, USA. *Palaeogeography, Palaeoclimatology, Palaeoecology* **271**, 170–181. doi:10.1016/J.PALAEO.2008.10.015
- Katz RW, Brush GS, Parlange MB (2005) Statistics of extremes: modeling ecological disturbances. *Ecology* **86**, 1124–1134. doi:10.1890/04-0606
- Kelly RF, Higuera PE, Barrett CM, Hu FS (2010) A signal-to-noise index to quantify the potential for peak detection in sediment–charcoal records. *Quaternary Research*. [Published online ahead of print 2010] doi:10.1016/J.YQRES.2010.07.011
- Long CJ, Whitlock C (2002) Fire and vegetation history from the coastal rain forest of the western Oregon Coast Range. *Quaternary Research* **58**, 215–225. doi:10.1006/QRES.2002.2378
- Long CJ, Whitlock C, Bartlein PJ, Millsaugh SH (1998) A 9000-year fire history from the Oregon Coast Range, based on a high-resolution charcoal study. *Canadian Journal of Forest Research* **28**, 774–787. doi:10.1139/CJFR-28-5-774
- Long CJ, Whitlock C, Bartlein PJ (2007) Holocene vegetation and fire history of the Coast Range, western Oregon, USA. *The Holocene* **17**, 917–926. doi:10.1177/0959683607082408
- Lynch JA, Clark JS, Bigelow NH, Edwards ME, Finney BP (2002) Geographic and temporal variations in fire history in boreal ecosystems of Alaska. *Journal of Geophysical Research* **107**, 8152. doi:10.1029/2001JD000332
- Lynch JA, Clark JS, Stocks BJ (2004a) Charcoal production, dispersal and deposition from the Fort Providence experimental fire: interpreting fire regimes from charcoal records in boreal forests. *Canadian Journal of Forest Research* **34**, 1642–1656. doi:10.1139/X04-071
- Lynch JA, Hollis JL, Hu FS (2004b) Climatic and landscape controls of the boreal forest fire regime: Holocene records from Alaska. *Journal of Ecology* **92**, 477–489. doi:10.1111/J.0022-0477.2004.00879.X
- Marlon J, Bartlein PJ, Whitlock C (2006) Fire–fuel–climate linkages in the north-western USA during the Holocene. *The Holocene* **16**, 1059–1071. doi:10.1177/0959683606069396
- Marlon JR, Bartlein PJ, Carcaillet C, Gavin DG, Harrison SP, Higuera PE, Joos F, Power MJ, Prentice IC (2008) Climate and human influences on global biomass burning over the past two millennia. *Nature Geoscience* **1**, 697–702. doi:10.1038/NGEO313
- Marlon JR, Bartlein PJ, Walsh MK, Harrison SP, Brown KJ, Edwards ME, Higuera PE, Power MJ, Anderson RS, Briles C, Brunelle A, Carcaillet C, Daniels M, Hu FS, Lavoie M, Long C, Minckley T, Richard PJH, Scott AC, Shafer DS, Tinner W, Umbanhowar CE, Whitlock C (2009) Wild-fire responses to abrupt climate change in North America. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 2519–2524. doi:10.1073/PNAS.0808212106
- Millsaugh SH, Whitlock C (1995) A 750-year fire history based on lake sediment records in central Yellowstone National Park, USA. *The Holocene* **5**, 283–292. doi:10.1177/095968369500500303
- Millsaugh SH, Whitlock C, Bartlein P (2000) Variations in fire frequency and climate over the past 17 000 years in central Yellowstone National Park. *Geology* **28**, 211–214. doi:10.1130/0091-7613(2000)28<211:VIFFAC>2.0.CO;2
- Minckley TA, Whitlock C, Bartlein PJ (2007) Vegetation, fire, and climate history of the north-western Great Basin during the last 14 000 years. *Quaternary Science Reviews* **26**, 2167–2184. doi:10.1016/J.QUASCIREV.2007.04.009
- Mohr JA, Whitlock C, Skinner CN (2000) Post-glacial vegetation and fire history, eastern Klamath Mountains, California, USA. *The Holocene* **10**, 587–601. doi:10.1191/095968300675837671
- Mudelsee M (2006) CLIM-X-DETECT: a Fortran 90 program for robust detection of extremes against a time-dependent background in climate records. *Computers & Geosciences* **32**, 141–144. doi:10.1016/J.CAGEO.2005.05.010
- Power MJ, Whitlock C, Bartlein P, Stevens LR (2006) Fire and vegetation history during the last 3800 years in north-western Montana. *Geomorphology* **75**, 420–436. doi:10.1016/J.GEOMORPH.2005.07.025
- Power MJ, Marlon J, Ortiz N, Bartlein PJ, Harrison SP, Mayle FE, Ballouche A, Bradshaw RHW, Carcaillet C, Cordova C, Mooney S, Moreno PI, Prentice IC, Thonicke K, Tinner W, Whitlock C, Zhang Y, Zhao Y, Ali AA, Anderson RS, Beer R, Behling H, Briles C, Brown KJ, Brunelle A, Bush M, Camill P, Chu GQ, Clark J, Colombaroli D, Connor S, Daniau AL, Daniels M, Dodson J, Doughty E, Edwards ME, Finsinger W, Foster D, Frechette J, Gaillard MJ, Gavin DG, Gobet E, Haberle S, Hallett DJ, Higuera P, Hope G, Horn S, Inoue J, Kaltenrieder P, Kennedy L, Kong ZC, Larsen C, Long CJ, Lynch J, Lynch EA, McGlone M, Meeks S, Mensing S, Meyer G, Minckley T, Mohr J, Nelson DM, New J, Newnham R, Noti R, Oswald W, Pierce J, Richard PJH, Rowe C, Goni MFS, Shuman BN, Takahara H, Toney J, Turney C, Urrego-Sanchez DH, Umbanhowar C, Vandergoes M, Vanniere B, Vescovi E, Walsh M, Wang X, Williams N, Wilmshurst J, Zhang JH (2008) Changes in fire regimes since the Last Glacial Maximum: an assessment based on a global synthesis and analysis of charcoal data. *Climate Dynamics* **30**, 887–907. doi:10.1007/S00382-007-0334-X
- Prichard SJ, Gedalof Z, Oswald WW, Peterson DL (2009) Holocene fire and vegetation dynamics in a montane forest, North Cascade Range, Washington, USA. *Quaternary Research* **72**, 57–67. doi:10.1016/J.YQRES.2009.03.008
- Shiue W-K, Bain LJ (1982) Experiment size and power comparisons for two-sample Poisson tests. *Journal of Applied Statistics* **31**, 130–134. doi:10.2307/2347975
- Toney JL, Anderson RS (2006) A post-glacial palaeoecological record from the San Juan Mountains of Colorado, USA: fire, climate and vegetation history. *The Holocene* **16**, 505–517. doi:10.1191/0959683606HL946RP
- Tweiten MA, Hotchkiss SC, Booth RK, Calcote RR, Lynch EA (2009) The response of a jack pine forest to late-Holocene climate variability in north-western Wisconsin. *The Holocene* **19**, 1049–1061. doi:10.1177/0959683609340993
- Underwood AJ (1997) ‘Experiments in Ecology.’ (Cambridge University Press: Cambridge, UK)
- Walsh MK, Whitlock C, Bartlein PJ (2008) A 14 300-year-long record of fire–vegetation–climate linkages at Battle Ground Lake, south-western Washington. *Quaternary Research* **70**, 251–264. doi:10.1016/J.YQRES.2008.05.002
- Whitlock C, Larsen C (2001) Charcoal as a fire proxy. In ‘Tracking Environmental Change Using Lake Sediments’. (Eds JP Smol, HJB Birks, WM Last) (Kluwer Academic Publisher: Dordrecht)
- Whitlock C, Dean W, Rosenbaum J, Stevens L, Fritz S, Bracht B, Power M (2008) A 2650-year-long record of environmental change from northern Yellowstone National Park based on a comparison of multiple proxy data. *Quaternary International* **188**, 126–138. doi:10.1016/J.QUAINT.2007.06.005

Manuscript received 18 November 2009, accepted 19 March 2010