

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

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

Philip E. Higuera^{1*}, Daniel G. Gavin², Patrick J. Bartlein², Douglas J. Hallett^{3,4}

1. Department of Forest Ecology and Biogeoscience, University of Idaho, Moscow ID
83844, USA

2. Department of Geography, University of Oregon, Eugene OR 97493, USA

3. Biogeoscience Institute, University of Calgary, Calgary, Alberta, T2N 1N4, Canada

4. School of Environmental Studies, Queen's University, Kingston, Ontario, K7L 3N6,
Canada

* corresponding author

Running head: Fire history from sediment charcoal records

Additional key words: paleoecology, bias, sensitivity, charcoal analysis

1 **Summary (50 words):**

2 Charcoal peaks in lake sediments provide valuable records of fire history in stand-
3 replacing fire regimes. Despite increasing use of this proxy, data analysis methods vary
4 and have not been systematically compared. We demonstrate important biases between
5 methods and make recommendations based on analyses of simulated and empirical
6 datasets.

7
8 **Abstract**

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

1 when charcoal counts are low. We conclude by discussing how to evaluate when peak
2 detection methods are warranted with a given sediment record.

3

4 **Introduction**

5 High-resolution charcoal records are an increasingly common source of fire-
6 history information, particularly in ecosystems where tree-ring records are short relative
7 to average fire-return intervals (Gavin *et al.* 2007). Over the past several decades
8 numerous studies have used peaks in charcoal accumulation in sediment records to
9 estimate the timing of “fire episodes”, one or more fires within the sampling resolution of
10 the sediment record (Whitlock and Larsen 2001). Identifying fire episodes from charcoal
11 records is most promising when fires are (1) large, (2) burn with high severity, and (3)
12 recur with average intervals at least five times the sampling resolution of the sediment
13 record (Clark 1988b, Whitlock and Larsen 2001, Higuera *et al.* 2005, Higuera *et al.*
14 2007). Sediment-charcoal records are thus particularly valuable for studying stand-
15 replacing fire regimes in boreal and subalpine forests, where all three of these conditions
16 are typically met.

17 Interpreting fire episodes from sediment-charcoal records would be
18 straightforward if they were characterized by low levels of charcoal punctuated by
19 unambiguous peaks. In reality, however, charcoal records are complex and non-
20 stationary, i.e., their mean and/or variance change over time (Clark *et al.* 1996, Clark and
21 Patterson 1997, Long *et al.* 1998). Empirical and theoretical studies (e.g., Marlon *et al.*
22 2006, Higuera *et al.* 2007) suggest that non-stationarity in charcoal records can arise from
23 at least two sets of processes: (1) changes in the fire regime including the rate of burning,

1 the intensity of fires, the type of vegetation burned, and thus charcoal production per unit
2 time, and/or (2) changes in the efficiency of charcoal delivery to the lake center
3 (taphonomy) due to changing rates of slope wash and/or within-lake redeposition. The
4 latter process, known as sediment focusing, can greatly affect the sediment accumulation
5 rate as a lake infills (Davis *et al.* 1984, Giesecke and Fontana 2009) and may produce
6 long-term trends in charcoal records unrelated to changes in the fire regime. Recognizing
7 the importance of these processes, paleoecologists have applied a range of statistical
8 methods to charcoal data in order to isolate the signal related to “local” fire occurrence
9 (e.g., within 0.5-1.0 km; Gavin *et al.* 2003, Lynch *et al.* 2004a, Higuera *et al.* 2007) and
10 reconstruct fire history. Despite the proliferation of statistical methods for peak
11 identification, seemingly no study has discussed the assumptions underlying alternative
12 methods and their impacts on fire-history interpretations.

13 Here we address several key issues related to peak identification in high-
14 resolution, macroscopic charcoal records¹ by using simulated and empirical charcoal
15 records. We start by discussing some important statistical properties of macroscopic
16 charcoal records and then describe the motivation for statistical treatments. We briefly
17 review how different methods have been applied, and then introduce a typology of
18 methods, including their respective assumptions and justifications. Second, we illustrate
19 and quantify the biases that these techniques can introduce to fire-history interpretations
20 by applying them to simulated charcoal records. Third, we apply the same methods to
21 three previously-published charcoal records from conifer forests to demonstrate potential

¹ We refer to macroscopic charcoal records as those quantifying charcoal not passing through a sieve of 125 μm or larger.

1 biases in empirical records, and we introduce a technique to minimize some of these
2 biases. Finally, we conclude with recommendations of specific methodologies and a
3 discussion of how analysts can evaluate the suitability of records for peak identification
4 rather than other qualitative or quantitative analyses.

5 **Temporal variability in charcoal time series**

6 Charcoal time series can be generally characterized as “noisy”, and they contain
7 many forms of non-stationarity, including changing short-term variability superimposed
8 on a slowly varying mean (Long *et al.* 1998, Higuera *et al.* 2007). Changes in variability,
9 or *heteroscedasticity*, have implications for the particular goal of data analysis. When the
10 goal is to quantify changes in total charcoal input, as an index of biomass burning for
11 example, heteroscedasticity violates the assumptions of parametric statistics useful in this
12 context, e.g., analysis of variance and regression. In particular, in analysis of variance (or
13 in the t-test of the difference of means in the case of two periods) heteroscedasticity
14 increases the probability of Type I error, falsely inferring significant differences between
15 periods (Underwood 1997). Similarly, in regression analysis fitting a trend line to
16 charcoal data with changing variability over time can increase the variability of the slope
17 coefficient. Changes in variability (besides being interesting in their own right) can thus
18 lead to false conclusions about the significance of long-term trends or differences
19 between different parts of a record. In practice, heteroscedasticity is usually dealt with by
20 applying a “variance-stabilizing transformation” (Emerson 1983) that acts to homogenize
21 variance across a record. As will be illustrated below, when the goal of charcoal analysis
22 is peak identification, transformation can lead to the exaggeration of some peaks and

1 suppression of others. Consequently, the specific approach taken (whether to transform or
2 not), should depend upon the overall focus of an analysis. In this paper we focus on the
3 goal of detecting local fires through peak detection.

4

5 **Analytical methods for inferring local fire occurrence**

6 Following Clark's pioneering work (1988b, 1990) in which fire events
7 surrounding small lakes were identified from charcoal in thin-sections of laminated
8 sediments, similar approaches were developed for quantifying macroscopic charcoal
9 abundance and subsequently adopted by a large number of research groups (Table 1; See
10 also Whitlock and Larsen 2001). Most techniques quantify charcoal as either the total
11 number of pieces or surface area (mm^2) of charcoal in a particular size class, within
12 volumetric subsamples taken contiguously through sediment cores (typically at 0.5 to 1.0
13 cm resolution, corresponding to ca. 10-25 year resolution for most lakes). The resulting
14 concentration of charcoal (pieces cm^{-3} , or $\text{mm}^{-2} \text{cm}^{-3}$) in each level is multiplied by the
15 estimated sediment accumulation rate (cm yr^{-1}) to obtain the charcoal accumulation rate
16 (CHAR, $\text{pieces cm}^{-2} \text{yr}^{-1}$ or $\text{mm}^{-2} \text{cm}^{-2} \text{yr}^{-1}$). Sediment accumulation rates, and the age of
17 each sample, are estimated by an age-depth model based on radiometric dates, tephra
18 layers, and any additional sources of age information. The use of accumulation rates can
19 potentially correct for changing sediment accumulation rates that would dilute or
20 concentrate charcoal in a given volume of sediment, and as mentioned above, may also
21 be affected by sediment focusing processes. Usually, the CHAR series is interpolated to a
22 constant temporal resolution to account for unequal sampling intervals resulting from
23 variable sediment accumulation rates. This step is necessary to develop threshold

1 statistics that are not biased to a particular portion of a record, and to standardized within-
2 and between-site comparisons.² Hereafter we refer to the interpolated CHAR series as *C*.
3 The analytical choices and sources of error in the development of a charcoal record are
4 briefly summarized in Table 2 and discussed in detail by Whitlock and Larsen (2001).

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

17 In practice, determining the size of peaks that represents local fires involves a
18 three-step “decomposition” of the *C* series (Clark *et al.* 1996, Long *et al.* 1998; Fig. 1).
19 First, the slowly-varying mean, or “background” component, C_{back} , is modeled through a
20 curve-fitting algorithm, e.g., a locally-weighted regression that is robust to outliers (e.g.,
21 Cleveland 1979). The window size for this smoothing varies between studies but is

² When sampling intervals are not standardized within a record or between two records, then biases may be introduced when applying criteria uniformly. Interpolation helps minimize, but not remove, this bias, as noted in the last section of this paper.

1 typically between 100 and 1000 years. Background estimation may be preceded by
2 transforming C (e.g., logarithmically). Second, the background trend is removed from the
3 series by subtraction ($C - C_{back}$) or division (C / C_{back}), creating a series of residuals or
4 indices, respectively. This detrended series is frequently termed the “peak component,”
5 but in the case of indices, it is dimensionless rather than a portion of C , as implied by
6 “peak component.” Here we use the term “peak series” and notation C_{peak} to refer to the
7 detrended series. Third, a threshold is applied to C_{peak} to separate variability related to
8 local fire occurrence from variability unrelated to local fire occurrence (e.g., random
9 variability and sediment mixing). Peaks exceeding the threshold are the basis for fire
10 frequency and fire return interval calculations.

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

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

16 In comparison with the little attention given to alternative detrending models,
17 recent papers have more carefully addressed the task of determining threshold values for
18 peak identification. Comparison of peaks with known fire events (dated from historical
19 records or tree-rings) may help in selecting a threshold, but historical records often
20 represents only a fraction of a charcoal record, and a wide range of thresholds may still
21 be appropriate (e.g., Gavin *et al.* 2006). Clark *et al.* (1996) addressed this issue by using a
22 sensitivity analysis to test of the number of peaks as a function of changing threshold
23 values. They reasoned that if C_{peak} was comprised of a population of small values (e.g.,

1 background charcoal) and a smaller population of large values (local fires), then the
2 sensitivity test should detect the split between these populations. Gavin *et al.* (2006) built
3 on this sensitivity test by modeling C_{peak} as a mixture of two Gaussian distributions with
4 different means, variances, and proportional contribution to the total population. The
5 lower distribution is assumed to represent the majority of time during which C is small
6 and is affected mainly by distant fires, redeposition, mixing, and random variability; i.e.,
7 the “noise” unrelated to specific fires. The upper distribution, ideally distinct from the
8 lower distribution, describes the variability due to local fires and can be considered the
9 “signal” of interest. Gavin *et al.* (2006) suggested that the threshold be at the upper end of
10 the noise distribution, and Higuera *et al.* (2008) further specified that the threshold be at
11 the 95th, 99th, or 99.9th percentile of the noise distribution. If the noise and signal
12 distributions are distinct, then the variance of the signal distribution (σ^2_S) would be much
13 larger than that of the noise distribution (σ^2_N). A signal-to-noise index (SNI; Higuera *et*
14 *al.* 2009), calculated as $\sigma^2_S/(\sigma^2_S+\sigma^2_N)$ approaches one when the noise distribution is
15 tightly defined with a narrow standard deviation. SNI values less than ≈ 0.5 suggest poor
16 separation of large peaks from the noise-attributable variation. We note these details here
17 because the Gaussian mixture approach assumes that the distribution of C_{peak} values is
18 right-skewed, and therefore variance-stabilizing expressions of C_{peak} (all but the NR
19 model) work against defining a distinct noise distribution.

20 Unless variance of C_{peak} does not change through time (i.e., it is homoscedastic),
21 selecting a threshold based on the entire series could lead to systematic biases towards
22 detecting small or large peaks (depending on which size dominates the record). While
23 variance-stabilization approaches were developed to address this issue, Higuera *et al.*

1 (2008a, 2009) introduced a new approach intended to be more adaptable by applying the
2 Gaussian mixture model introduced by Gavin *et al.* (2006) to a 500-year moving window
3 of C_{peak} centered on each time step in the series. This technique is termed a “local
4 threshold,” and it accounts for potentially changing variance of C_{peak} by selecting a
5 threshold based on σ^2_N in a user-defined subsection of the record. Using smaller sample
6 sizes to compute the Gaussian mixture distribution increases the chance of erratic model
7 fits in a portion of the cases. Thus, it is important to smooth the local thresholds (typically
8 to the same frequency as that use to define C_{back}) such that it varies smoothly over time
9 and be cognizant of the total number of samples in each local population (a minimum of
10 ca. 30 is recommended; Higuera *et al.* 2009). This decomposition approach is similar to
11 peak-detection methods in other applications (e.g., Mudelsee 2006) in that it accounts for
12 changes in both the central tendency and variability in a series.

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

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

1
2

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

10 **Methods**

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

18

19 *Simulated Records*

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

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

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

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

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

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

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

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

8

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

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

10

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

13

14 *Empirical Records*

15 We selected three high-resolution charcoal records with differing variability in
 16 background charcoal and peak heights. Little Lake (Long *et al.* 1998) is located in
 17 Douglas-fir forest in the Oregon Coast Range. The 3.3-ha, 4.0-m deep lake is surrounded
 18 by a fen and has a small inflowing stream draining a 597-ha watershed (Marlon *et al.*
 19 2006; C. Long, personal communication, November 2009). The 11.3-meter core has
 20 overall C values similar to the simulated records (median = 14.4 pieces $\text{cm}^{-2} \text{yr}^{-1}$). Over

1 its 9000-yr record C_{back} varies between 0.94 and 44.04 pieces $\text{cm}^{-2} \text{yr}^{-1}$, and vegetation
2 was consistently dominated by Douglas-fir. Rockslide Lake (Gavin *et al.* 2006) is located
3 in subalpine forest in southeast British Columbia. The 3.2-ha, 14.1-m deep lake is fed by
4 an intermittent stream within an 86-ha watershed. The 2.1-m core has overall C values
5 lower than Little Lake (median = 0.49 pieces $\text{cm}^{-2} \text{yr}^{-1}$). Over its 5000-year record C_{back}
6 varies between 0.06 and 1.13 pieces $\text{cm}^{-2} \text{yr}^{-1}$, and vegetation was consistently dominated
7 by Engelmann spruce and subalpine fir. Finally, Ruppert Lake (Higuera *et al.* 2009) is
8 located in boreal forest of Alaska's south-central Brooks Range. The 3-ha, 7.0-m deep
9 lake has a ca. 200-ha watershed with subdued topography and a small inflowing stream.
10 The 4.8-m core has the lowest overall C values of all three records (median = 0.04 pieces
11 $\text{cm}^{-2} \text{yr}^{-1}$). Over the 14,000-year record C_{back} ranges from 0.00-0.22 pieces $\text{cm}^{-2} \text{yr}^{-1}$ with a
12 distinct increase around 5000 yr, coincident with the transition from a forest-tundra to
13 boreal forest vegetation. Overall, five different vegetation types dominated the landscape
14 around Ruppert Lake during the record.

15 For all records, we used the published age-depth relationship but reanalyzed each
16 series using the published resampling intervals of 10, 10, and 15 yr for Little, Rockslide,
17 and Ruppert lakes, respectively. We did not use the same analysis parameters as in the
18 published records, since our purpose was to test different parameters. We calculated
19 background charcoal using a locally weighted regression robust to outliers (lowess) in a
20 500-year window. The robust lowess model is less sensitive to non-stationarity and thus
21 may be applied to raw and transformed data (Cleveland 1979).

22

23 Data transformation, peak identification, and sensitivity analysis

1 We applied the four different detrending models to each simulated and empirical
2 record, and we used a modified Levene's test of equal variance (based on sample
3 medians; Brown and Forsythe 1974) to test the null hypothesis of equal variance between
4 two portions of each record. Sample sizes for p-value calculations were adjusted to
5 account for temporal autocorrelation in each record following Bretherton *et al.* (1999).
6 For simulated records, we compared the periods 10,000-6000 and 4000-0 yr BP. We
7 present only the median test result for 500 realizations of the simulated series.³ For
8 empirical records we subjectively selected periods where background charcoal had two
9 qualitatively different levels and then divided this period in half for comparison. At
10 Little, Rockslide, and Ruppert lakes, these periods corresponded to the last 8000, 5000,
11 and 10,000 years, respectively. The test statistic, W_{50} , is used as an index of
12 heteroscedasticity, and the associated p-value is used to assess the null hypothesis of
13 equal variance.

14 We identified peaks in simulated and empirical records using a Gaussian mixture
15 model that models the noise distribution within C_{peak} (described earlier). In this
16 application, the value of the mixture model is its ability to apply uniform treatments to all
17 records, making specific threshold-selection parameters of less importance. For all
18 analyses, we used the 99th percentile of the modeled noise distribution as the threshold

³ Results did not differ when analyzing 250, 500, or 1000 realizations (each 10,000 yr long), suggesting that the inherent variability captured.

1 value.⁴ Thresholds were defined both globally (a single mixture model fit to the entire
2 record) or locally (fitting the mixture models to 500-year windows centered on each
3 sample, and then smoothing the series of resulting threshold values).

4 For the simulated records, we quantified the sensitivity of peak identification to
5 the four detrending models with a sensitivity ratio, s . We defined s as the number of
6 peaks detected in the first half of each record divided by the total number of peaks
7 detected in the second half of each record. If an analytical method is insensitive to
8 variations in C_{back} , then s will equal one. Values of s significantly greater or less than one
9 indicate a systematic bias in the set of analytical methods. We used a Monte Carlo
10 approach to estimate the value of s for each of the 16 analysis combinations (2 simulation
11 scenarios x 4 detrending models x 2 threshold-determination techniques = 16). For each
12 combination, s was estimated by the average s from 500 simulations, and the 2.5th and
13 97.5th percentiles were used to estimate 95% confidence intervals around s . If the 95%
14 confidence intervals overlapped one, then the ratio was considered no different from one
15 and the method was deemed insensitive to the variation in background charcoal.

16 We performed two additional analyses on the empirical records. First, we
17 explored the effect of the four detrending models on the capacity of the Gaussian mixture
18 model to identify a distinct noise distribution. For simplicity, we chose to use only a
19 globally-fit model applied to the Rockslide Lake record, the least variable record; similar
20 examples could be based on subsections of other records. As a metric of how distinct
21 peaks were from C_{back} , we examined the SNI (defined earlier) of the fitted Gaussian

⁴ 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.

1 mixture model. Second, to illustrate the impact of the eight alternative decomposition
2 methods and the minimum-counts test, we applied each method to the empirical records.
3 We quantified the percent of peaks that fail to pass the minimum count test under each
4 decomposition method, and to illustrate how interpretations may differ, we summarized
5 peaks (after removing those failing to pass the minimum-count test) with 1000-yr
6 smoothed peak frequency curves (peaks 1000-yr^{-1} , smoothed to 1000 yr with a lowess
7 filter).

8

9 **Results**

10 *Simulated Records*

11 As designed, simulated charcoal records from Scenario 1 were homoscedastic
12 (500-sample median $W_{50} = 0.45$, median $p = 0.502$), while records from Scenario 2 were
13 heteroscedastic ($W_{50} = 21.87$, $p < 0.001$; Table 4, Fig. 2). For both scenarios, the choice
14 of decomposition method had a major effect on the variability in the resulting peak series,
15 C_{peak} (Table 4, Fig. 2). Under Scenario 1, only the NR model resulted in a stationary
16 series ($W_{50} = 0.44$, $p = 0.508$; Fig. 2; Table 4). The TR and NI models greatly inflated
17 variance when background charcoal was low ($W_{50} = 46.72, 25.62$; $p \leq 0.001$), and the TI
18 model further inflated variance ($W_{50} = 98.54$; $p < 0.001$). No model stabilized variance in
19 records from Scenario 2 (Table 4, Fig. 2). The NR model preserved heteroscedasticity in
20 the original record ($W_{50} = 21.81$), the TR and NI models reduced heteroscedasticity (W_{50}
21 $= 7.12, 10.81$), while the TI model increased heteroscedasticity ($W_{50} = 57.92$). The
22 skewness of C_{peak} also varied greatly among models. For both Scenario 1 and Scenario 2,
23 the NI model produced the most skewed peak series (3.47 and 3.03), followed closely by

1 the NR model (2.59, 2.97), and the TI (1.21, 1.55) and TR (0.95, 1.20) models (Table 4,
2 Fig. 2).

3 In simulated records, threshold type was more important than detrending model
4 when evaluating the sensitivity of peak identification to changes in variance (Fig. 3).
5 Locally-defined thresholds were insensitive to the presence of heteroscedasticity
6 (Scenario 2 vs. Scenario 1) and detrending model (s for all scenarios did not differ from
7 1). In contrast, using a globally-defined threshold produced unbiased results only under
8 three conditions. When C was characterized by constant variance (Scenario 1), a
9 globally-defined threshold was unbiased when C_{peak} was defined by residuals: median s
10 for NR and TR models was 1.00 (95% CI 0.76-1.33) and 1.17 (0.92-1.54), respectively.
11 Using an index to define C_{peak} inflated variance when C_{back} was low (Fig. 2), resulting in
12 1.86-2.55 times the number of detected peaks: s for NI and TI models was 1.86 (1.24-
13 3.00) and 2.55 (1.50-4.56), respectively. When variance in C increased with C_{back}
14 (Scenario 2), transforming C and using residuals produced unbiased results, as did
15 creating an index from the non-transformed series: s for TR and NI models was 0.85
16 (0.63-1.10) and 1.33 (0.98-1.79), respectively. Defining C_{peak} as the residuals of non-
17 transformed C (NR model) resulted in more peaks detected when C_{back} and variability
18 was high ($s = 0.63$ [0.43-0.87]), and transforming and using an index to define C_{peak} (TI
19 model) resulted in nearly twice as many peaks detected when C_{back} and variability were
20 low ($s = 1.70$ [1.15-2.63]).

21 Overall, analyses of simulated records illustrate the sensitivity of decomposition
22 methods to changes in the mean (Scenario 1) and changes in the mean and variance
23 (Scenario 2) of a series through time. Although simplified, the sensitivity of the simulated

1 records to the analytical method highlights biases that can arise from similar changes in
2 empirical records, even when of smaller magnitude and/or duration.

3

4 *Empirical Records*

5 The three empirical records differed greatly in their long-term variability in C (Fig
6 4). Little Lake had relatively low charcoal values until ca. 4000 yr BP when sediment
7 accumulation rate increased five-fold (0.07 to 0.35 cm yr⁻¹) in parallel with C . At
8 Rockslide Lake, sediment accumulation rates varied about two-fold (0.04 to 0.08 yr⁻¹)
9 and were largely independent of C . At Ruppert Lake, sediment accumulation rates varied
10 nearly five-fold (0.018 to 0.181 cm yr⁻¹) and while C followed the sediment accumulation
11 rate during the first few millennia, these variables were unrelated for the majority of the
12 record.

13 The raw records (C) at each site exhibited significant heteroscedasticity (Table 4),
14 with Little Lake exhibiting the most, followed by Ruppert and Rockslide lakes (W_{50} =
15 153.14, 84.62, 15.20, respectively; $p \leq 0.001$). The four detrending models had a large
16 effect on the variance in C_{peak} (Table 4; Fig. 4). The TR and NI models were most
17 effective at stabilizing variance, although results differed between sites. At Rockslide
18 Lake the TR model stabilized variance (500-2500 vs. 2500-0 yr BP, W_{50} = 0.49, p =
19 0.483); this model performed second best at Little Lake (800-4000 vs. 4000-0 yr BP, W_{50}
20 = 3.40, p = 0.066) and performed worst at Ruppert Lake (10,000-5000 vs. 5000-0 yr BP,
21 W_{50} = 66.52, $p < 0.001$). At Little Lake, the NI model stabilized variance (W_{50} = 0.01, p =
22 0.942), and at Ruppert Lake no model stabilized variance. Although the NR and TI
23 models reduced heteroscedasticity, they did not stabilize variance in any record.

1 Skewness in C_{peak} was largest when using the NR model (Little Lake, 17.73) or NI model
2 (Rockslide and Ruppert lakes, 5.28, 6.79, respectively). In contrast, log-transforming C
3 (TR and TI models) reduced skewness, and at Little Lake these models resulted in near-
4 symmetric distributions (0.28 and -0.56, respectively; Table 4).

5 The noise distribution fit by the Gaussian mixture model resulted in different
6 signal-to-noise (SNI) and skewness values, dependent on the decomposition model (Fig.
7 5). As applied to the Rockslide Lake record, the NR and NI models yielded the largest
8 signal-to-noise index (SNI; 0.97 and 0.98), whereas the TR and TI models had the
9 smallest SNI (0.81 and 0.94). Skewness, as a potential measure of the occurrence of high
10 values distinct from a noise distribution, was highest for the NI model (5.28). The TR
11 model, though having a moderate SNI, was the most symmetric (skewness = 1.68).

12 As with simulated records, peak detection in empirical records was more sensitive
13 to alternative decomposition models when using a global vs. local threshold, and this
14 sensitivity varied greatly between sites (Fig. 6). At Little Lake, where C_{back} varied the
15 most throughout the record, a global threshold detected 41 peaks with the NR model but
16 only 5 with the TI model, producing drastically different trends in 1000-yr mean fire
17 frequency. The TR and NI models produced an intermediate number of peaks (23 and 30)
18 with qualitatively similar trends over time. In contrast to Little Lake, peak detection with
19 any model varied by 6% at Rockslide Lake (33-35) and 13% at Ruppert Lake (64-72),
20 where variability in C_{back} was less. At all sites, locally-defined thresholds detected more
21 peaks and with less variability between models than with globally-defined thresholds,
22 even after minimum-count screening (described below). Again, differences among
23 models were greatest at Little Lake, where peak detection varied from 56-68 (21%). Peak

1 detection varied little at Rockslide Lake, 34-36 (6%), and slightly more at Ruppert Lake,
2 79-88 (11%). Differences at Little and Ruppert lakes largely reflect differences between
3 the two residual models (NR and TR) in comparison to the index models (NI and TI).

4 The minimum-count screening flagged between 0-14% of the total peaks detected
5 in any one record, with the least at Little Lake (median = 2.5%), followed by Rockslide
6 Lake (median = 9%) and Ruppert Lake (median = 11%). A greater proportion of the total
7 peaks detected was flagged when using a local vs. global threshold (median = 9% vs.
8 3%), and the variability between detrending models differed by threshold type. When
9 using a global threshold, a larger percentage of total peaks were flagged when using
10 index models (median for NI and TI = 11%) vs. residual models (median for NR and TR
11 = 3%). This difference was reduced using a local threshold (10% vs. 9%, respectively).

12

13 **Discussion**

14 Interpreting local fire history from sediment charcoal records involves a number
15 of analytical steps that decompose multiple signals into a series of peaks that bears
16 interpretation (Fig. 1). Accounting for non-stationarity in a record is a primary goal of
17 decomposition methods, and our analyses of simulated records illustrate the sensitivity of
18 alternative methods to two types of non-stationarity: a change in the mean (Scenario 1),
19 and a change in the mean and variance (Scenario 2) through time. In combination with
20 empirical records, our results highlight some critical methodological considerations that
21 have been broadly overlooked in the literature. Specifically, we emphasize the need for
22 careful consideration when proceeding through steps 4-6 of a decomposition method
23 (Fig. 1): (a) defining C_{peaks} , or detrending, (b) defining a threshold to detect peaks, and, in

1 the case of charcoal counts, (c) screening and removing peaks that could result from
2 insignificant changes in charcoal counts.

3

4 *Detrending to define a peak series*

5 An overriding conclusion from our study is that the impacts of different
6 detrending models are largely obviated by using a locally-defined threshold. In simulated
7 records, peak identification using a local threshold was robust to changes in background
8 charcoal, peak variance, and detrending model (Fig. 3). In empirical records, these
9 patterns largely held true, as reflected by less between-model variability when using local
10 vs. global thresholds. For example, the total number of peaks detected since 5000 yr BP
11 in Ruppert Lake varied by 7% vs. 30% when applying a local vs. global threshold to the
12 different detrending models (Fig. 6). Locally-defined thresholds outperformed global
13 thresholds because the mixture model used to determine thresholds constantly adapts to
14 variability in a record. Consequently, local thresholds are free from the assumption of
15 stable variance in peak heights, at least for time scales longer than the window width used
16 to define “local.” With no need to stabilize variance across a record, detrending before
17 applying a locally-defined threshold needs only to account for changes in the long-term
18 mean, and thus three of the four detrending models evaluated become obsolete. Even
19 stabilizing the mean, interestingly, may be unnecessary when using a locally-defined
20 threshold, but no studies have attempted this to date. Our results are consistent with
21 analyses done by Ali *et al.* (2009b), who applied the local threshold technique to three
22 different charcoal quantification metrics from individual cores (counts, area, and
23 estimated volume). Although the variability between the three metrics differed, the

1 locally-applied threshold produced similar results in each case. Overall, these findings
2 lend support to the recent adoption of local thresholds for peak-identification (Table 1) as
3 robust to changes in variance both within and between records.

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

12 When applying a global threshold, it is imperative to evaluate the presence or
13 absence of heteroscedasticity in a record before selecting a detrending model. If a record
14 has stable variance, then the NR model is the single appropriate model because it only
15 removes the mean trend of a series (Fig. 2). In simulated records, the only instance in
16 which the global threshold was unbiased was when the NR model was applied to
17 homoscedastic records (Scenario 1; Fig. 3). The closest analogy in the empirical records
18 is from Rockslide Lake, which had the least heteroscedasticity of the three records
19 evaluated (and was the shortest in length) and consequently was most robust to
20 alternative detrending methods. Applying variance-stabilizing models (TR, NI, TI) to
21 homoscedastic records is not only unwarranted, but it can result in severely biased peak
22 identification (Fig. 3) by simultaneously amplifying and suppressing peaks during
23 periods of low and high background charcoal, respectively (Fig. 2). This bias was

1 minimized when using the TR model, and it subsequently increased with the NI and TI
2 models. By amplifying peak sizes when C_{back} is low, index-based models applied to
3 homoscedastic records result in biased peak identification (Fig. 2, 6). This bias is most
4 extreme when using the TI model in combination with a global threshold, as this led to
5 more than twice as many peaks being detected during periods of low vs. high background
6 charcoal in our simulations (Fig. 3).

7 Unfortunately, most empirical charcoal records exhibit heteroscedasticity at some
8 time scale, particularly those spanning different biomes and/or many millennia. This
9 limits the utility of the NR model with a global threshold. All the empirical records in this
10 study, for example, had non-stable variance between the two periods of comparison
11 (Table 4; Fig. 4). In heteroscedastic records, both empirical and simulated records
12 support the TR or NI models as most appropriate for consideration when using a global
13 threshold. Although no model stabilized variance in the simulated records with
14 heteroscedasticity, the TR and NI performed the best, and in empirical records these
15 models stabilized variance across comparison periods in some records (Table 4). When
16 applied to heteroscedastic records, the NR and TI models are inappropriate for the
17 opposite reasons. Simply detrending by residuals (NR) fails to remove any
18 heteroscedasticity (Table 4), which biases peak identification towards periods of high
19 C_{back} (Fig. 3). Detrending with an index of transformed data (TI) *reverses* the pattern of
20 heteroscedasticity in what is essentially a “double whammy” of variance-stabilization
21 (Fig. 2, 4), biasing peak identification towards periods of low C_{back} (Fig 3). The
22 undesirable effects of the TI model are most apparent in the Little Lake record, where
23 four of the five peaks detect occurred during the period of low background charcoal (Fig.

1 6). The overall low number of peaks detect in this scenario also stands out as odd, and it
2 highlights the conceptual difficulty of interpreting fire history from a symmetric peak
3 series. If the variability above C_{back} does not differ from the variability below C_{back} , then
4 it is inconsistent to interpret the former as fire-related while interpreting the latter as
5 noise-related. When a globally-defined Gaussian mixture model is applied to a nearly
6 symmetric peak series (e.g., Little Lake under the TI model, skewness = -0.56; Table 4),
7 selecting a threshold at the 99th percentile cuts off 99% of the samples in the series (895
8 of 900 samples in the Little Lake record; Fig. 6).

9 Finally, we emphasize that the impacts of different detrending models will vary
10 between sites, depending on the mean and variability of charcoal accumulation rates in a
11 record. Rockslide Lake, for example, was largely robust to alternative decomposition
12 methods, whereas Little Lake displayed large variability between methods (Fig. 7).
13 Ruppert Lake was also clearly heteroscedastic, but overall lower C and C_{back} values as
14 compared to Little Lake resulted in less sensitivity to alternative detrending models (Fig.
15 6).

16

17 *Defining a threshold*

18 The Gaussian mixture model introduced by Gavin *et al.* (2006) is promising
19 because it provides a semi-objective, process-based means of selecting a threshold for
20 peak identification which in turn can be applied to multiple records. Using the mixture
21 model to identify a threshold depends upon three key assumptions: (1) variation in the
22 noise distribution, representing variability around the long-term trend (i.e. C_{back}), is
23 normally distributed; (2) the mean and variance of this noise distribution is stationary

1 within the period of analysis; and (3) there are enough samples within the period of
2 analysis to adequately characterize the noise distribution. The first assumption has
3 theoretical support from a charcoal simulation model (Fig. 3 in Higuera et. al. 2007), and
4 it is consistent with the distributions of peak charcoal observed in empirical records (e.g.
5 Higuera *et al.* 2009; Fig. 5). The mechanisms creating normally-distributed variability
6 around the long-term trend include sediment mixing, inter-annual variability in long-
7 distant charcoal input, sampling effects, and analytical error. Other mechanisms may
8 produce skewed variability, and to the extent that this is true, this is a limitation of the
9 Gaussian mixture model (discussed below).

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

1 power as the intervals decrease, and interpreting p-values become difficult with multiple
2 comparisons.

3 The application of the Gaussian mixture model to identify a threshold is also
4 aided by maximizing the separation between the noise distribution and fire-related peaks.
5 This is a key difference between the analytical approach taken for peak identification
6 compared to the analysis of long-term trends in total charcoal (e.g., Marlon *et al.* 2008,
7 Power *et al.* 2008, Marlon *et al.* 2009). Whereas homogenizing variance is desirable in
8 the context of the latter, this decreases separation between noise and fire-related samples,
9 i.e., it reduced the signal-to-noise index (SNI). For example, in the Rockslide Lake
10 record, the SNI was highest (0.97-0.988) using the NR and NI models, whereas it was
11 consistently lowered when applying variance-stabilizing transformations (0.81-0.94; Fig.
12 5). Skewness may also serve as a coarse index of how separated peak values are from
13 non-peak values. At Little Lake, a nearly symmetric peak distribution defined by the TI
14 model reflected little to no separation between peak and non-peak values. Thus, as a
15 general rule, a minimum level of skewness of around two would suggest a SNI sufficient
16 to aid in setting thresholds, but increased skewness beyond two does not necessarily
17 equate to an increased SNI. We also note that skewness alone is not justification for peak
18 interpretation, particularly if it is an artifact of the detrending processes.

19

20 *Interpreting small charcoal peaks*

21 Most decomposition methods resulted in the identification of small peaks that
22 failed to pass the minimum-count test at the 95% confidence level (0-14% of the total
23 peaks identified; Fig. 6). Some peaks fail to pass this test because they closely follow

1 other large peaks (i.e. a “double peak”), e.g., around 1200 yr BP at Ruppert Lake (Fig. 4
2 and 6). These peaks most likely represent non-significant variations in charcoal counts
3 due to natural and/or analytical variability. More challenging for sediment-based fire-
4 history reconstructions are periods of low charcoal abundance. In these cases, both
5 variance-stabilizing and local-threshold methods may result in detecting small peaks,
6 often associated with small charcoal counts in the raw record. The smaller the charcoal
7 peak, the more difficult it is to infer if the peak was caused by a local fire vs. a distant fire
8 and/or random variability in charcoal deposition and quantification. The minimum-count
9 test helps guard against falsely inferring a peak was caused by a local fire (Type I error).
10 When this probability is low, e.g. < 0.05 , it is highly unlikely that the two samples come
11 from the same population. Practically, Figure 7 illustrates the increase in counts (as a
12 proportion and absolute number) required to achieve a given level of confidence (95 or
13 99%) as a function of the number of charcoal pieces in the pre-peak sample. The lower
14 the pre-peak count, the greater the proportional increase in charcoal required before a
15 peak sample can be considered distinct with 95% confidence. For example, when pre-
16 peak counts are < 10 , peak counts must double before having a $< 5\%$ chance of coming
17 from the same populations. Much smaller proportional increases are required when
18 overall counts are large, e.g., only a 20% increase is required when pre-peak sample are
19 ca. 100. Thus, as a rough rule of thumb, it is highly desirable for researchers to use
20 sample volumes that will result in average non-peak samples of > 10 pieces, and peak
21 values of *at least* 20 pieces.

22 Even with large sample volume, in some cases the difference between a peak and
23 non-peak sample may be small. Interpreting variability when charcoal counts are low

1 highlights a limitation of the Gaussian mixture model briefly mentioned above. The
2 mixture model assumes normally-distributed noise, and thus it may fail when counts are
3 small, because the true noise distribution may be positively skewed (i.e., Poisson
4 distributions with a mean < 10 are positively skewed). If so, the Gaussian mixture model
5 would underestimate the threshold, resulting in an increased false-positive rate. Future
6 efforts modeling noise distributions within C_{peak} could address this limitation through the
7 use of non-Gaussian mixture models, which may be more appropriate for the heavy-tailed
8 distributions that characterize C and C_{peak} series (Coles 2001). For example, the signal
9 and noise distributions may be better fit by models in the generalized extreme value
10 family (e.g., Weibull, Fréchet, and Gumbel distributions), and the signal distribution may
11 be fit more appropriately by models in the Generalized Pareto family (e.g. Pareto, beta,
12 and exponential distributions; Katz *et al.* 2005). In the ideal case, the signal distribution
13 has little influence on the parameters of the noise distributions, because the noise
14 population typically dominates the mixed distribution. Nonetheless, improving the fit of
15 the signal distribution would be an improvement over current methods and deserves
16 exploration. In the mean time, we suggest that the minimum-count test serves well to
17 screen out small peaks, be they detected with a threshold from Gaussian mixture model
18 or otherwise.

19

20 *Recommendations and Conclusions*

21 Even when applying the most rigorous analytical techniques, there is no substitute
22 for careful inspection of a record to assess whether it can provide an unbiased fire history
23 in the first place. We highlight three key issues related to assessing the quality of

1 millennial-scale charcoal records when independent evidence supporting a particular
2 reconstruction is lacking. First, records should be interpreted in the context of a null
3 hypothesis of random variability. If a peak series lacks large values, is symmetric, or fails
4 to detect recent fires, then the record should be considered too noisy for peak
5 identification. The signal-to-noise index utilized here is intended to help evaluate this null
6 hypothesis, and current work is improving the application of this metric for this purpose
7 (Kelly *et al.*, in preparation). Records with low SNI value(s) or symmetric peak series
8 should either forgo peak identification methods, or be presented with a low, medium, and
9 high ranges of possible thresholds. Depending on the cause of a low SNI, these records
10 may still be appropriate and valuable for interpreting trends in biomass burning through
11 interpretations of C and/or C_{back} .

12 Second, if a record has large variability in sediment accumulation rates,
13 practitioners must consider the possibility that changing peak frequencies result from
14 changes in sample resolution. Resampling a record to the median or maximum deposition
15 time per sample coarsens or falsely increases the resolution during periods of high or low
16 sediment accumulation, and thus should create a more temporally unbiased time series.
17 However, this resampling may not be a simple solution if sedimentation varies widely,
18 because sediment mixing modifies the effect of changing sediment accumulation rate on
19 the effective resolution of a sediment record. For example, mixing the top 2-cm of
20 sediment during a period when the deposition time is 10 yr cm^{-1} would result in an
21 effective 1-cm resolution of 20 years. In contrast, the same 2-cm mixing depth during a
22 sediment accumulation rate of 20 yr cm^{-1} would result in an effective resolution of 40
23 years. An important cause of changing sediment accumulation rate is fluctuating within-

1 basin sediment focusing and sediment delivery by stream flow. Such processes can
2 change the effectiveness of sediment delivery, including charcoal, from lake margins to
3 the lake center. This results in the widely observed positive correlation between charcoal
4 accumulation and sediment accumulation rates (i.e., constant charcoal concentration
5 despite changing sediment accumulation rates) and a heteroscedastic charcoal record
6 (e.g., Fig. 4). In contrast, if charcoal were delivered entirely through airfall, increased
7 sediment accumulation rates would dilute charcoal concentrations, and the charcoal
8 record would not be heteroscedastic. Therefore, we strongly warn against interpreting fire
9 frequency changes in records with a several-fold change in sediment accumulation rate
10 (along with no evidence that charcoal concentrations are diluted by changing
11 sedimentation) and when inferred fire frequency closely tracks sediment accumulation
12 rates. While there may be a non-causal relationship between sediment accumulation and
13 fire frequency (e.g., via erosion or climate), this link must be explained with independent
14 evidence if fire history is to be interpreted. A viable alternative in these cases is to only
15 interpret high-resolution segments with constant sediment accumulation rates.

16 Finally, segments of records with low overall counts must be interpreted with
17 caution. The use of the minimum-count test presented here can help guide interpretation
18 in these cases, as can independent evidence of fire (e.g., pollen or macrofossils of fire-
19 dependent taxa). While charcoal records have successfully detected fires in non-forested
20 ecosystems (e.g. savannah and tundra: Duffin *et al.* 2008, Higuera *et al.* 2008b), at some
21 point along a fire-intensity spectrum, fires will not produce enough charcoal to create an
22 identifiable peak in a record (Higuera *et al.* 2005, Duffin *et al.* 2008). This may be the
23 case even with ample sample volume. Likewise, as local fire frequency decreases, so too

1 does the frequency of large charcoal peaks; this makes it more difficult to identify the
2 signal of local fires from the noise of long-distance transport and within-lake
3 redeposition.

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

20

21 **Acknowledgements**

22 We thank Doug Sprugel for raising these issues and Colin Long for sharing the Little
23 Lake data.

1

2 **References**

3 Ali, AA, H Asselin, AC Larouche, Y Bergeron, C Carcaillet, and PJH Richard (2008)

4 Changes in fire regime explain the Holocene rise and fall of *Abies balsamea* in the
5 coniferous forests of western Quebec, Canada. *Holocene* **18**, 693-703.

6 Ali, AA, C Carcaillet, and Y Bergeron (2009a) Long-term fire frequency variability in

7 the eastern Canadian boreal forest: the influences of climate vs. local factors. *Global*
8 *Change Biology* **15**, 1230-1241.

9 Ali, AA, PE Higuera, Y Bergeron, and C Carcaillet (2009b) Comparing fire-history

10 interpretations based on area, number and estimated volume of macroscopic charcoal
11 in lake sediments. *Quaternary Research* **72**, 462-468.

12 Allen, CD, RS Anderson, RB Jass, JL Toney, and CH Baisan (2008) Paired charcoal and

13 tree-ring records of high-frequency Holocene fire from two New Mexico bog sites.
14 *International Journal of Wildland Fire* **17**, 115-130.

15 Anderson, RS, CD Allen, JL Toney, RB Jass, and AN Bair (2008) Holocene vegetation

16 and fire regimes in subalpine and mixed conifer forests, southern Rocky Mountains,
17 USA. *International Journal of Wildland Fire* **17**, 96-114.

18 Anderson, RS, DJ Hallett, E Berg, RB Jass, JL Toney, CS de Fontaine, and A DeVolder

19 (2006) Holocene development of Boreal forests and fire regimes on the Kenai
20 Lowlands of Alaska. *Holocene* **16**, 791-803.

21 Beaty, RM, and AH Taylor (2009) A 14 000 year sedimentary charcoal record of fire

22 from the northern Sierra Nevada, Lake Tahoe Basin, California, USA. *Holocene* **19**,
23 347-358.

- 1 Bretherton, CS, M Widmann, VP Dymnikov, JM Wallace, and I Blade (1999) The
2 effective number of spatial degrees of freedom of a time-varying field. *Journal of*
3 *Climate* **12**, 1990-2009.
- 4 Briles, CE, C Whitlock, and PJ Bartlein (2005) Postglacial vegetation, fire, and climate
5 history of the Siskiyou Mountains, Oregon, USA. *Quaternary Research* **64**, 44-56.
- 6 Briles, CE, C Whitlock, PJ Bartlein, and PE Higuera (2008) Regional and local controls
7 on postglacial vegetation and fire in the Siskiyou Mountains, northern California,
8 USA. *Palaeogeography Palaeoclimatology Palaeoecology* **265**, 159-169.
- 9 Brown, MB, and AB Forsythe (1974) Robust Tests for Equality of Variances. *Journal of*
10 *the American Statistical Association* **69**, 364-367.
- 11 Brunelle, A, and RS Anderson (2003) Sedimentary charcoal as an indicator of late-
12 Holocene drought in the Sierra Nevada, California, and its relevance to the future.
13 *The Holocene* **13**, 21-28.
- 14 Brunelle, A, and C Whitlock (2003) Postglacial fire, vegetation, and climate history in
15 the Clearwater Range, northern Idaho, USA. *Quaternary Research* **60**, 307-318.
- 16 Brunelle, A, C Whitlock, P Bartlein, and K Kipfmüller (2005) Holocene fire and
17 vegetation along environmental gradients in the Northern Rocky Mountains.
18 *Quaternary Science Reviews* **24**, 2281-2300.
- 19 Carcaillet, C, Y Bergeron, PJH Richard, B Frechette, S Gauthier, and YT Prairie (2001)
20 Change of fire frequency in the eastern Canadian boreal forests during the Holocene:
21 does vegetation composition or climate trigger the fire regime? *Journal of Ecology*
22 **89**, 930-946.

- 1 Clark, JS (1988a) Effects of climate change on fire regimes in northwestern Minnesota.
2 *Nature* **334**, 233-235.
- 3 Clark, JS (1988b) Particle motion and the theory of charcoal analysis: source area,
4 transport, deposition, and sampling. *Quaternary Research* **30**, 67-80.
- 5 Clark, JS (1988c) Stratigraphic charcoal analysis on petrographic thin sections:
6 application to fire history in northwestern Minnesota. *Quaternary Research* **30**, 81-
7 91.
- 8 Clark, JS (1990) Fire and climate change during the last 750 yr in northwestern
9 Minnesota. *Ecological Monographs* **60**, 135-159.
- 10 Clark, JS, and WA Patterson. 1997. Background and local charcoal in sediments: scales
11 of fire evidence in the paleorecord. Pages 23-48 in J. S. Clark, H. Cachier, J. G.
12 Goldammer, and B. J. Stocks, editors. *Sediment Records of Biomass Burning and*
13 *Global Change*. Springer, New York.
- 14 Clark, JS, and PD Royall (1996) Local and regional sediment charcoal evidence for fire
15 regimes in presettlement north-eastern North America. *Journal of Ecology* **84**, 365-
16 382.
- 17 Clark, JS, PD Royall, and C Chumbley (1996) The role of fire during climate change in
18 an eastern deciduous forest at Devil's Bathtub, New York. *Ecology* **77**, 2148-2166.
- 19 Cleveland, WS (1979) Robust locally weighted regression and smoothing scatterplots.
20 *Journal of the American Statistical Association* **74**, 829-836.
- 21 Coles, S (2001) An introduction to statistical modeling of extreme values. Springer-
22 Verlag, London, UK.

- 1 Cook, ER, and K Peters (1997) Calculating unbiased tree-ring indices for the study of
2 climatic and environmental change. *The Holocene* **7**, 359-368.
- 3 Daniels, ML, RS Anderson, and C Whitlock (2005) Vegetation and fire history since the
4 Late Pleistocene from the Trinity Mountains, northwestern California, USA.
5 *Holocene* **15**, 1062-1071.
- 6 Davis, MB, RE Moeller, and J Ford. 1984. Sediment focusing and pollen influx. Pages
7 261-293 in E. Y. Haworth and J. W. G. Lund, editors. *Lake Sediments and*
8 *Environmental History*. University of Leicester Press, Leicester, U.K.
- 9 Detre, K, and C White (1970) Comparison of 2 Poisson-Distributed Observations.
10 *Biometrics* **26**, 851-&.
- 11 Duffin, KI, L Gillson, and KJ Willis (2008) Testing the sensitivity of charcoal as an
12 indicator of fire events in savanna environments: quantitative predictions of fire
13 proximity, area and intensity. *Holocene* **18**, 279-291.
- 14 Emerson, JD. 1983. Mathematical aspects of transformation
15 Pages 247-282 in D. C. Hoaglin, F. Mosteller, and J. W. Tukey, editors. *Understanding*
16 *Robust and Exploratory Data Analysis*. Wiley
- 17 Fowler, AM (2009) Variance Stabilization Revisited: A Case for Analysis Based on Data
18 Pooling. *Tree-Ring Research* **65**, 129-145.
- 19 Gavin, DG, LB Brubaker, and KP Lertzman (2003) An 1800-year record of the spatial
20 and temporal distribution of fire from the west coast of Vancouver Island, Canada.
21 *Canadian Journal of Forest Research* **33**, 573-586.
- 22 Gavin, DG, DJ Hallett, FS Hu, KP Lertzman, SJ Prichard, KJ Brown, JA Lynch, P
23 Bartlein, and DL Peterson (2007) Forest fire and climate change in western North

1 America: insights from sediment charcoal records. *Frontiers in Ecology and the*
2 *Environment* **5**, 499-506.

3 Gavin, DG, FS Hu, K Lertzman, and P Corbett (2006) Weak climatic control of stand-
4 scale fire history during the late Holocene. *Ecology* **87**, 1722-1732.

5 Giesecke, T, and SL Fontana (2009) Revisiting pollen accumulation rates from Swedish
6 lake sediments. *The Holocene* **18**, 293-305.

7 Hallett, DJ, and RS Anderson (2010) Paleo-fire reconstruction for high-elevation forests
8 in the Sierra Nevada, California with implications for wildfire synchrony and climate
9 variability in the late Holocene. *Quaternary Research* **73**, 180-190.

10 Hallett, DJ, DS Lepofsky, RW Mathewes, and KP Lertzman (2003a) 11,000 years of fire
11 history and climate in the mountain hemlock rain forests of southwestern British
12 Columbia based on sedimentary charcoal. *Canadian Journal of Forest Research* **33**,
13 292-312.

14 Hallett, DJ, RW Mathewes, and RC Walker (2003b) A 1000-year record of forest fire,
15 drought and lake-level change in southeastern British Columbia, Canada. *Holocene*
16 **13**, 751-761.

17 Hallett, DJ, and RC Walker (2000) Paleoecology and its application to fire and vegetation
18 management in Kootenay National Park, British Columbia. *Journal of*
19 *Paleolimnology* **24**, 401-414.

20 Higuera, PE, LB Brubaker, PM Anderson, TA Brown, AT Kennedy, and FS Hu (2008a)
21 Frequent Fires in Ancient Shrub Tundra: Implications of Paleorecords for Arctic
22 Environmental Change. *PLoS ONE* **3**, e0001744.

- 1 Higuera, PE, LB Brubaker, PM Anderson, FS Hu, and TA Brown (2009) Vegetation
2 mediated the impacts of postglacial climate change on fire regimes in the south-
3 central Brooks Range, Alaska. *Ecological Monographs* **79**, 201-219.
- 4 Higuera, PE, M Chipman, JA Allen, S Rupp, and FS Hu (2008b) Tundra fire regimes in
5 the Noatak National Preserve, northwestern Alaska, since 6000 yr BP. Pages 144 *in*
6 93th Annual Meeting of the Ecological Society of America, Milwaukee, WI.
- 7 Higuera, PE, ME Peters, LB Brubaker, and DG Gavin (2007) Understanding the origin
8 and analysis of sediment-charcoal records with a simulation model. *Quaternary*
9 *Science Reviews* **26**, 1790-1809.
- 10 Higuera, PE, DG Sprugel, and LB Brubaker (2005) Reconstructing fire regimes with
11 charcoal from small-hollow sediments: a calibration with tree-ring records of fire. *The*
12 *Holocene* **15**, 238-251.
- 13 Huber, UM, V Markgraf, and F Schäbitz (2004) Geographical and temporal trends in
14 Late Quaternary fire histories of Fuego-Patagonia, South America. *Quaternary*
15 *Science Reviews* **23**, 1079-1097.
- 16 Huerta, MA, C Whitlock, and J Yale (2009) Holocene vegetation-fire-climate linkages in
17 northern Yellowstone National Park, USA. *Palaeogeography, Palaeoclimatology,*
18 *Palaeoecology* **271**, 170-181.
- 19 Katz, RW, GS Brush, and MB Parlange (2005) Statistics of extremes: modeling
20 ecological disturbances. *Ecology* **86**, 1124-1134.
- 21 Long, CJ, and C Whitlock (2002) Fire and vegetation history from the coastal rain forest
22 of the western Oregon Coast Range. *Quaternary Research* **58**, 215-225.

- 1 Long, CJ, C Whitlock, and PJ Bartlein (2007) Holocene vegetation and fire history of the
2 Coast Range, western Oregon, USA. *Holocene* **17**, 917-926.
- 3 Long, CJ, C Whitlock, PJ Bartlein, and SH Millsbaugh (1998) A 9000-year fire history
4 from the Oregon Coast Range, based on a high-resolution charcoal study. *Canadian*
5 *Journal of Forest Research* **28**, 774-787.
- 6 Lynch, JA, JS Clark, NH Bigelow, ME Edwards, and BP Finney (2002) Geographic and
7 temporal variations in fire history in boreal ecosystems of Alaska. *Journal of*
8 *Geophysical Research* **108**, FFR8-1-FFR8-17.
- 9 Lynch, JA, JS Clark, and BJ Stocks (2004a) Charcoal production, dispersal and
10 deposition from the Fort Providence experimental fire: Interpreting fire regimes from
11 charcoal records in boreal forests. *Canadian Journal of Forest Research* **34**, 1642-
12 1656.
- 13 Lynch, JA, JL Hollis, and FS Hu (2004b) Climatic and landscape controls of the boreal
14 forest fire regime: Holocene records from Alaska. *Journal of Ecology* **92**, 447-489.
- 15 Marlon, J, PJ Bartlein, and C Whitlock (2006) Fire-fuel-climate linkages in the
16 northwestern USA during the Holocene. *The Holocene* **16**, 1059-1071.
- 17 Marlon, JR, PJ Bartlein, C Carcaillet, DG Gavin, SP Harrison, PE Higuera, F Joos, MJ
18 Power, and IC Prentice (2008) Climate and human influences on global biomass
19 burning over the past two millennia. *Nature Geoscience* **1**, 697-702.
- 20 Marlon, JR, PJ Bartlein, MK Walsh, SP Harrison, KJ Brown, ME Edwards, PE Higuera,
21 MJ Power, C Whitlock, RS Anderson, C Briles, A Brunelle, C Carcaillet, M Daniels,
22 FS Hu, M Lavoie, C Long, T Minckley, PJH Richard, SL Shafer, W Tinner, and C
23 Umbanhowar (2009) Wildfire responses to abrupt climate change in North America.

1 *Proceedings of the National Academy of Sciences of the United States of America*
2 **106**, 2519-2524.

3 Millspaugh, SH, and C Whitlock (1995) A 750-year fire history based on lake sediment
4 records in central Yellowstone National Park, USA. *The Holocene* **5**, 283-292.

5 Millspaugh, SH, C Whitlock, and P Bartlein (2000) Variations in fire frequency and
6 climate over the past 17000 yr in central Yellowstone National Park. *Geology* **28**,
7 211-214.

8 Minckley, TA, C Whitlock, and PJ Bartlein (2007) Vegetation, fire, and climate history
9 of the northwestern Great Basin during the last 14,000 years. *Quaternary Science*
10 *Reviews* **26**, 2167-2184.

11 Mohr, JA, C Whitlock, and CN Skinner (2000) Postglacial vegetation and fire history,
12 eastern Klamath Mountains, California, USA. *The Holocene* **10**, 587-602.

13 Mudelsee, M (2006) CLIM-X-DETECT: A Fortran 90 program for robust detection of
14 extremes against a time-dependent background in climate records. *Computers &*
15 *Geosciences* **32**, 141-144.

16 Power, MJ, J Marlon, N Ortiz, et al. (2008) Changes in fire regimes since the Last Glacial
17 Maximum: an assessment based on a global synthesis and analysis of charcoal data.
18 *Climate Dynamics* **30**, 887-907.

19 Power, MJ, C Whitlock, P Bartlein, and LR Stevens (2006) Fire and vegetation history
20 during the last 3800 years in northwestern Montana. *Geomorphology* **75**, 420-436.

21 Prichard, SJ, Z Gedalof, WW Oswald, and DL Peterson (2009) Holocene fire and
22 vegetation dynamics in a montane forest, North Cascade Range, Washington, USA.
23 *Quaternary Research* **72**, 57-67.

1 Shiue, W, K., and LJ Bain (1982) Experiment size and power comparisons for two-
2 sample Poisson tests. *Journal of Applied Statistics* **31**, 130-134.

3 Toney, JL, and RS Anderson (2006) A postglacial palaeoecological record from the San
4 Juan Mountains of Colorado USA: fire, climate and vegetation history. *Holocene* **16**,
5 505-517.

6 Tweiten, MA, SC Hotchkiss, RK Booth, RR Calcote, and EA Lynch (2009) The response
7 of a jack pine forest to late-Holocene climate variability in northwestern Wisconsin.
8 *Holocene* **19**, 1049-1061.

9 Underwood, AJ (1997) *Experiments in Ecology*. Cambridge University Press,
10 Cambridge.

11 Walsh, MK, C Whitlock, and PJ Bartlein (2008) A 14,300-year-long record of fire-
12 vegetation-climate linkages at Battle Ground Lake, southwestern Washington.
13 *Quaternary Research* **70**, 251-264.

14 Whitlock, C, W Dean, J Rosenbaum, L Stevens, S Fritz, B Bracht, and M Power (2008)
15 A 2650-year-long record of environmental change from northern Yellowstone
16 National Park based on a comparison of multiple proxy data. *Quaternary*
17 *International* **188**, 126-138.

18 Whitlock, C, and C Larsen. 2001. Charcoal as a fire proxy. in J. P. Smol, H. J. B. Birks,
19 and W. M. Last, editors. *Tracking Environmental Change Using Lake Sediments*.
20 Kluwer Academic Publisher, Dordrecht.

21
22
23

1 **Tables**

2

3 **Table 1.** Published fire-history studies in North America based on macroscopic charcoal
 4 (sieved or in thin sections) where the goal of analysis was to detect peaks associated with
 5 local fires. Studies are grouped by threshold type and the detrending models used for
 6 analysis. “N” indicates the total number of studies in each category.

Threshold type	Detrending model	Citation	Location (State or Province)	Particle Count or Area	Sediment volume per sample (cm ³)	Size class tallied	Background estimate	Threshold determination ¹	Threshold value ²
Global N = 37	Non-transform, Residuals (NR) N = 13	Clark (1990)	Minnesota, USA	A	Thin sections	> 60 μm long	15-yr moving average	Three lakes: TR (2-8)	> 42 to 68
		Millsbaugh and Whitlock (1995)	Wyoming, USA	C	5	125 μm	3-point, center 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-yr 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-yr window	Seven lakes: H, TR	> 40
		Carcaillet <i>et al.</i> (2001)	Québec, Canada	A	1	>150 μm	Inverse Fourier transform: 100 yr window	Three lakes: TR: not possible	> 1sd of the average of background
		Gavin <i>et al.</i> (2003)	British Columbia, Canada	C	ca. 12	150-500 μm	26-yr locally weighted minimum value	SC (12), TR (3); S1	0.22
		Lynch <i>et al.</i> (2002)	Alaska, USA	A	1	>180 μm	100-yr 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-yr locally weighted mean	Fifteen 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 to 3	>180 μm	100-yr locally weighted mean	Four lakes: S1	0.018, 0.085
	Gavin <i>et al.</i> (2006)	British Columbia, Canada	C	2 to 5	>125 μm	500-yr robust Lowess	Two lakes: H, TR (2); S1	GMM; 0.08, 5.00
	Prichard <i>et al.</i> (2009)	Washington, USA	C	ca. 10	>150-500 μm	750-yr locally weighted mean	TR (2): S1	GMM at 99 th percentile: not reported
	Ali <i>et al.</i> (2008)	Québec, Canada	A	1	> 160 μm	500-yr tricube locally-weighted regression	Lake, peat and soil charcoal compared; S1	GMM at 95 th percentile: 0.2
	Ali <i>et al.</i> (2009a)	Québec, Canada	A	1	> 160 μm	1000-yr tricube locally-weighted regression	Four lakes: S1	GSM at 95 th percentile: 0.015, 0.025, 0.040, 0.007
Transform, Residuals (TR)	Hallett and Anderson (2010)	California, USA	C	2.5	> 125 μm	500-yr robust Lowess	Two lakes: TR (1-2): S1	GSM at 95 th percentile: 0.03
Non-transform, Index (NI)	Higuera <i>et al.</i> (2005)	Washington, USA	C	3	150-500, > 500 μm	Series median (300-yr)	12 Small hollows: TR (1-3 per site)	1.63-1.75
	Tweiten <i>et al.</i> (2009)	Wisconsin, USA	C	1	> 150 μm	300-yr Lowess	not possible	1.3
Transform, Index (TI)	Long <i>et al.</i> (1998)	Oregon, USA	C	2.5	>125 μm	600-yr locally-weighted mean	TR, H (4)	1.12
	Hallett and Walker (2000)	British Columbia, Canada	C, A	1	>150 μm	500-yr locally weighted mean	TR (2)	1.0
	Millspaugh <i>et al.</i>	Wyoming, USA	C	5	>125 μm	750-yr locally-	S2	1.0

(2000)						weighted mean		
Mohr <i>et al.</i> (2000)	California, USA	C	5	>125 μm	120-yr locally-weighted mean	Two lakes: TR (3)	1.0	
Long and Whitlock (2002)	Oregon, USA	C	2.5	>125 μm	600-yr locally-weighted mean	H (2)	1.25	
Brunelle and Anderson (2003)	California, USA	C	5	>125 μm	500-yr locally-weighted mean	H(1)	1.1	
Brunelle <i>et al.</i> (2003, 2005)	Idaho and Montana, USA	C	5	>125 μm	750 or 600-yr 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-yr locally-weighted mean	Two lakes: SC (10-18)	1.0	
Hallett <i>et al.</i> (2003b)	British Columbia, Canada	C	10	>125 μm	50-yr locally-weighted mean	TR (4)	1.1	
Daniels <i>et al.</i> (2005)	California, USA	C	ca. 8	>125 μm	370-yr locally-weighted mean	TR (3)	1.0	
Briles <i>et al.</i> (2005)	Oregon, USA	C	2 to 5	>125 μm	500-yr locally-weighted mean	TR (3)	1.1	
Toney and Anderson (2006)	Colorado, USA	C	5	>125 μm	300-yr locally-weighted mean	H (1)	1.0	
Anderson <i>et al.</i> (2006)	Alaska, USA	C	5	>125 μm	300-yr locally-weighted mean	H (3-5); SC (4)	1.2	
Power <i>et al.</i> (2006)	Montana, USA	C	0.3 to 1?	>125 μm	150-yr locally-	H (1)	1.05	

						weighted mean			
		Marlon <i>et al.</i> (2006)	California, Oregon, Montana, Wyoming, Idaho, USA	C	1 to 5	>125 μm	500-yr 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 to 5	>125 μm	600-yr locally weighted mean	H (2-3); S1	1.2
		Anderson <i>et al.</i> (2008)	Colorado, New Mexico, USA	C	1 to 5	>125 μm	1000-yr locally-weighted mean	Three lakes and three bogs: H (1-3)	1.01
		Allen <i>et al.</i> (2008)	New Mexico, USA	C	1 to 5	>125 μm	300-yr 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 to 5	>125 μm	900-yr locally-weighted mean	Three lakes: H (lowest threshold with no fires in last 100 yr)	1.00, 1.05, 1.15
		Whitlock <i>et al.</i> (2008)	Wyoming, USA	C	1	>125 μm	150-yr locally-weighted mean	TR (2)	1.1
		Beaty and Taylor (2009)	California, USA	C	3	>125 μm	240-yr locally-weighted mean	TR (3)	1.0
Local	Non-transform, Residuals (NR)	Higuera <i>et al.</i> (2008a)	Alaska, USA	C	3 to 5	>150 μm	smoothed 500-yr median or mode	Two lakes (pre-modern): S2	GMM at 99 th percentile
N = 6	N = 6	Walsh <i>et al.</i> (2008)	Washington, USA	C	1	>125 μm	500-yr robust Lowess 700-yr locally-weighted regression	TR (3); S2	GMM at 95 th percentile
		Briles <i>et al.</i> (2008)	California, USA	C	2	>125 μm	locally-weighted regression	Two lakes; S2	GMM at 95 th percentile
		Marlon <i>et al.</i> (2009)	North America	C	Varies with several records used	>125 μm	500, 600, 800-yr robust Lowess	Thirty-five lakes: S2	GMM at 95 th percentile
		Higuera <i>et al.</i> (2009)	Alaska, USA	C	3 to 5	>150 μm	smoothed	Four lakes: H(1); S2	

						500-yr median or mode		GMM at 99 th percentile
	Huerta <i>et al.</i> (2009)	Wyoming, USA	C	5 and 50	>125 μm	500-yr robust Lowess	H(1); S2	GMM at 95 th percentile

-
- 1 ¹ Methods of threshold determination as follows:
2 TR: detected peaks compared to fires reconstructed from tree-rings (fire scars and/or stand ages)
3 SC: detect peaks compared to radiocarbon dates from local soil charcoal
4 H: detected peaks compared to historical fire record
5 S1: sensitivity analysis based on coefficient quantifying separation of peaks from background
6 S2: sensitivity analysis based on qualitative assessment of results using alternative threshold criteria
7 For TR, SC, and H methods, the number of independent fire records is shown in parentheses.
8
9 ² Threshold values: Units for NR and TR models are pieces $\text{cm}^{-2}\text{yr}^{-1}$ for counts or $\text{mm}^2\text{cm}^{-2}\text{yr}^{-1}$ for area.
10 Thresholds for NI and TI models are unitless index values.
11 GMM: Gaussian mixture model. For local thresholds, the percentile of the GMM defines a different
12 threshold value for each sample, and thus threshold values are not reported.
13 GSM: Gaussian single model, with mean of zero.
14
15

1 **Table 2.** Decisions typically required to develop a high-resolution lake-sediment
 2 macroscopic charcoal record, summarized from Whitlock and Larsen (2001). The aim of
 3 the current paper is to discuss data manipulations after completing these steps of
 4 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	Volume overestimate (core shrinkage).
	Sampling interval	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. Estimate charcoal accumulation rate (CHAR)	Age-depth model fitting to calculate sediment accumulation rates	Poor chronological control
6. Interpolate CHAR to a constant interval	Interval size (typical value is the median sample deposition time)	Loss of resolution in portions of the record

1 **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 $\text{cm}^{-2} \text{yr}^{-1}$ or $\text{cm}^2 \text{cm}^{-2} \text{yr}^{-1}$
$\log(C+1)$	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: $C_{peak} = C / C_{back}$
TR	Transform, Residual: $C_{peak} = \log(C+1) - C_{back}$, where $C_{back} = f(\log[C+1])$
TI	Transform, Index: $C_{peak} = \log(C+1) / C_{back}$, where $C_{back} = f(\log[C+1])$

2
3

1 **Table 4.** Stationarity of variance and skewness of C and C_{peak} series for different decomposition
2 models. The modified Levene's test statistic, W_{50} , and the probability of the null hypothesis of
3 equal variances, p , are based on comparisons between values from 10,000-6000 to 4000-0 yr BP
4 in simulated records, and equally split halves since 8000, 5000, and 10,000 yr BP for Little,
5 Rockslide, and Ruppert lakes, respectively. Bold (italic) values identify stationary series, those
6 that fail to reject the null hypothesis at $\alpha = 0.10$ (0.05), where a higher α is more conservative.
7 The skewness coefficient is a measure of the asymmetry of the entire peak series of each
8 respective model, where positive values indicate greater spread above the mean value and a zero
9 value indicates a symmetric distribution. The time series for each model is shown in Fig. 2 and 4
10 for the simulated and empirical records, respectively. Values for simulated records represent the
11 median value from 500 records constructed under each scenario.

12

Scenario or Site	W_{50} test statistic for equality of variances (p-value)					Skewness coefficient (2.5 th -97.5 th percentile) [within-row rank]			
	C	NR	TR	NI	TI	NR	TR	NI	TI
Scenario 1 (variance constant)	<i>0.45</i> (<i>0.502</i>)	<i>0.44</i> (<i>0.508</i>)	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)	<i>3.40</i> (<i>0.066</i>)	<i>0.01</i> (<i>0.942</i>)	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)	<i>0.49</i> (<i>0.483</i>)	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]

13

1 **Figure Legends**

2 **Fig. 1.** The set of decisions required for analyzing a charcoal time series with the goal of peak
3 detection for interpretation of fire episodes. These steps are implemented in the *CharAnalysis*
4 software (<http://Charanalysis.googlepages.com>; last accessed Oct 20, 2009).

5
6 **Fig. 2.** Simulated charcoal records reflecting alternative assumptions regarding the stability of
7 the variance through time. (a) Representative records from each scenario. Scenario 1 has
8 constant variance peak heights superimposed on a changing mean. Scenario 2 is a
9 heteroscedastic series in which the peak variance changes in proportion to C_{back} . The thick black
10 line in all figures is a 500-year loess smooth used to define the “background” levels (C_{back}). (b)
11 Series expressed on a log scale. (c-f) Detrended series based on four alternative methods.

12
13 **Fig. 3.** Sensitivity of peak identification to decomposition models and threshold type. The
14 sensitivity index, s , is the ratio of detected peaks from period 1 to period 2 in the two simulated
15 charcoal scenarios in Fig 2. The error bars indicate the 95% confidence interval from 500
16 realizations of the simulated records.

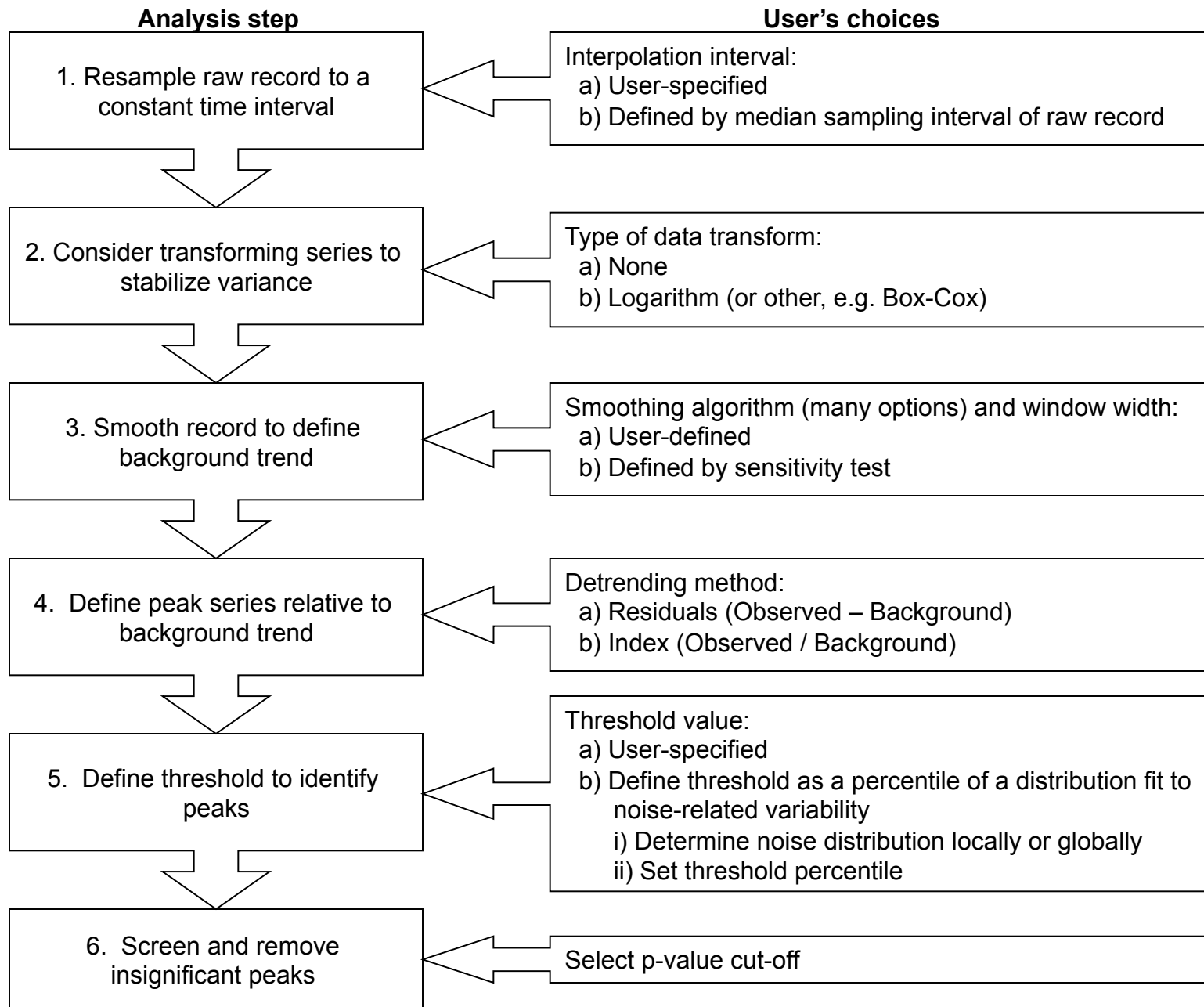
17
18 **Fig 4.** Empirical charcoal records detrended using the four decomposition models. Records are
19 from western Oregon (Little Lake; Long *et al.* 1998), southeast British Columbia (Rockslide
20 Lake, Gavin *et al.* 2006), and the Alaskan Brooks Range (Ruppert Lake; Higuera *et al.* 2009).

21
22 **Fig 5.** The Gaussian mixture model applied to the C_{peak} series from Rockslide Lake. Each panel
23 corresponds to a single detrending model. The Gaussian model representing the noise

1 distribution is shown by a thick gray line. The vertical line represents a typical threshold level for
2 peak identification, located at the 99th percentile of the lower distribution.

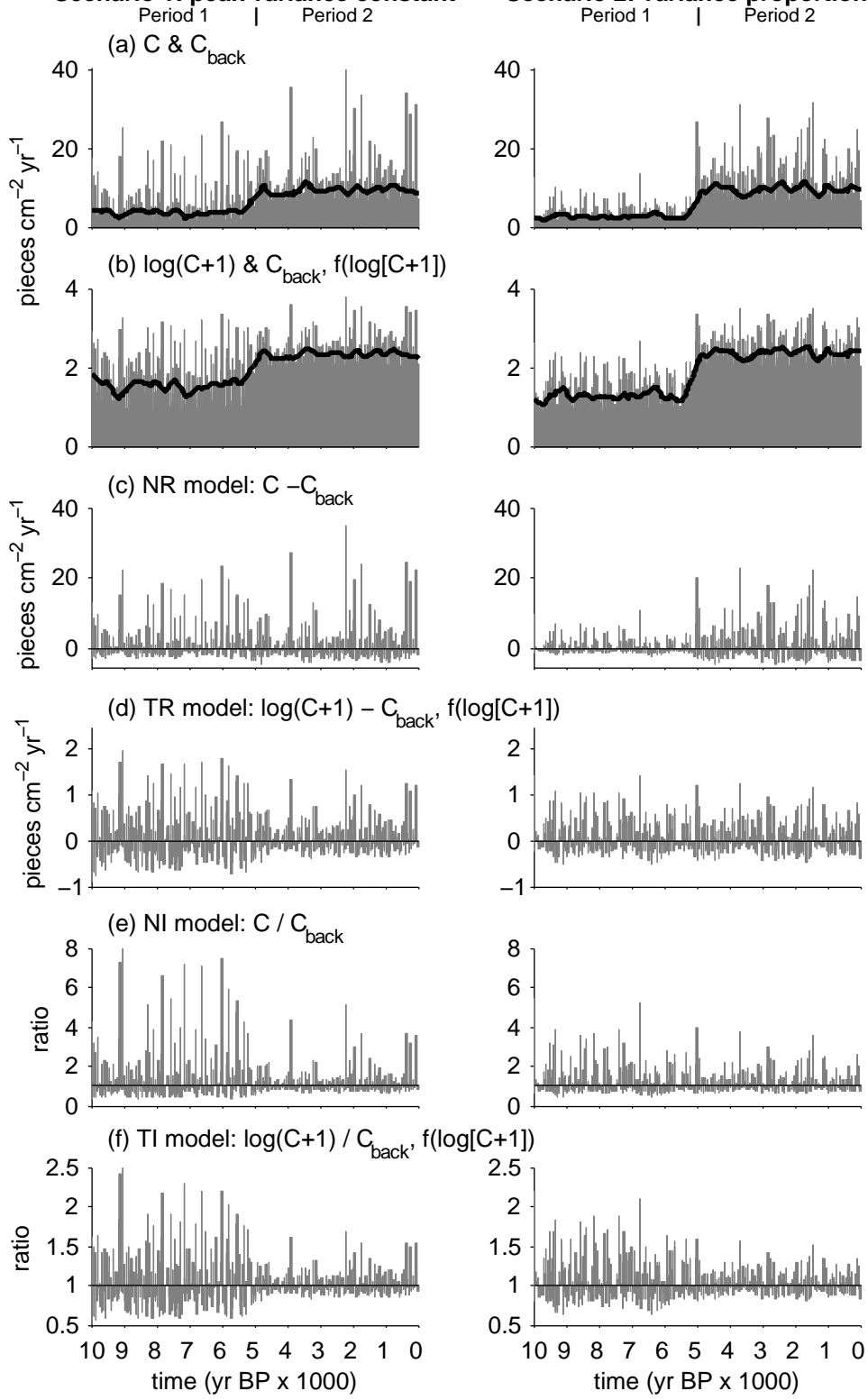
3
4 **Fig 6.** Inferred fire history from Little, Rockslide, and Ruppert lakes using alternative
5 decomposition methods. Each row corresponds to a different detrending model, as in Fig. 4, and
6 each panel includes peaks detected based on a global and local threshold. The location of peaks
7 exceeding the threshold value(s) are identified with “+” and “.” symbols, where the latter
8 identifies peaks that failed to pass the minimum-count test. The proportion of total peaks failing
9 to pass the minimum-count test is displayed on the right hand side of each panel. Smoothed lines
10 represent the 1000-yr average fire frequency for a given decomposition method, and all panels
11 are scaled from 0 to 15 (fires per 1000 yr) on the y-axis.

12
13 **Fig 7.** Minimum increase in charcoal counts required to confidently separate pre-peak from peak
14 samples. The required increase is displayed as a total number and a proportion, and it depends
15 upon (a) the confidence desired ($\alpha = 0.01$, 99% or $\alpha = 0.05$, 95% confidence), and (b) the number
16 of pieces in the smaller, pre-peak sample. The curves are developed from the test for assessing
17 whether two samples are from the same Poisson distribution (Detre and White 1970). Lines for
18 two significance levels are shown and presented both as absolute counts and percentage increase
19 of the lower count.

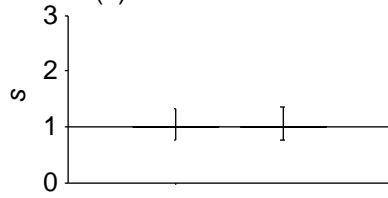


Scenario 1: peak variance constant

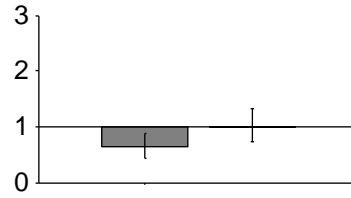
Scenario 2: variance proportional



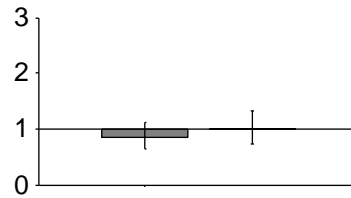
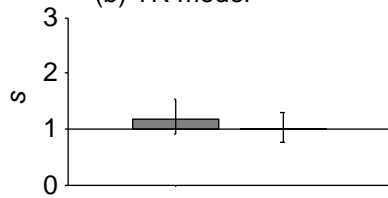
Scenario 1:
peak variance constant
(a) NR model



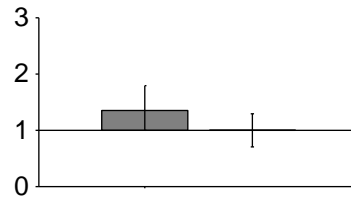
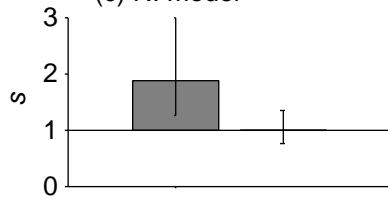
Scenario 2:
peak variance proportional



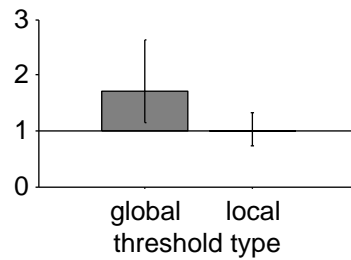
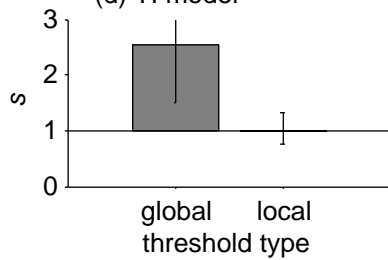
(b) TR model

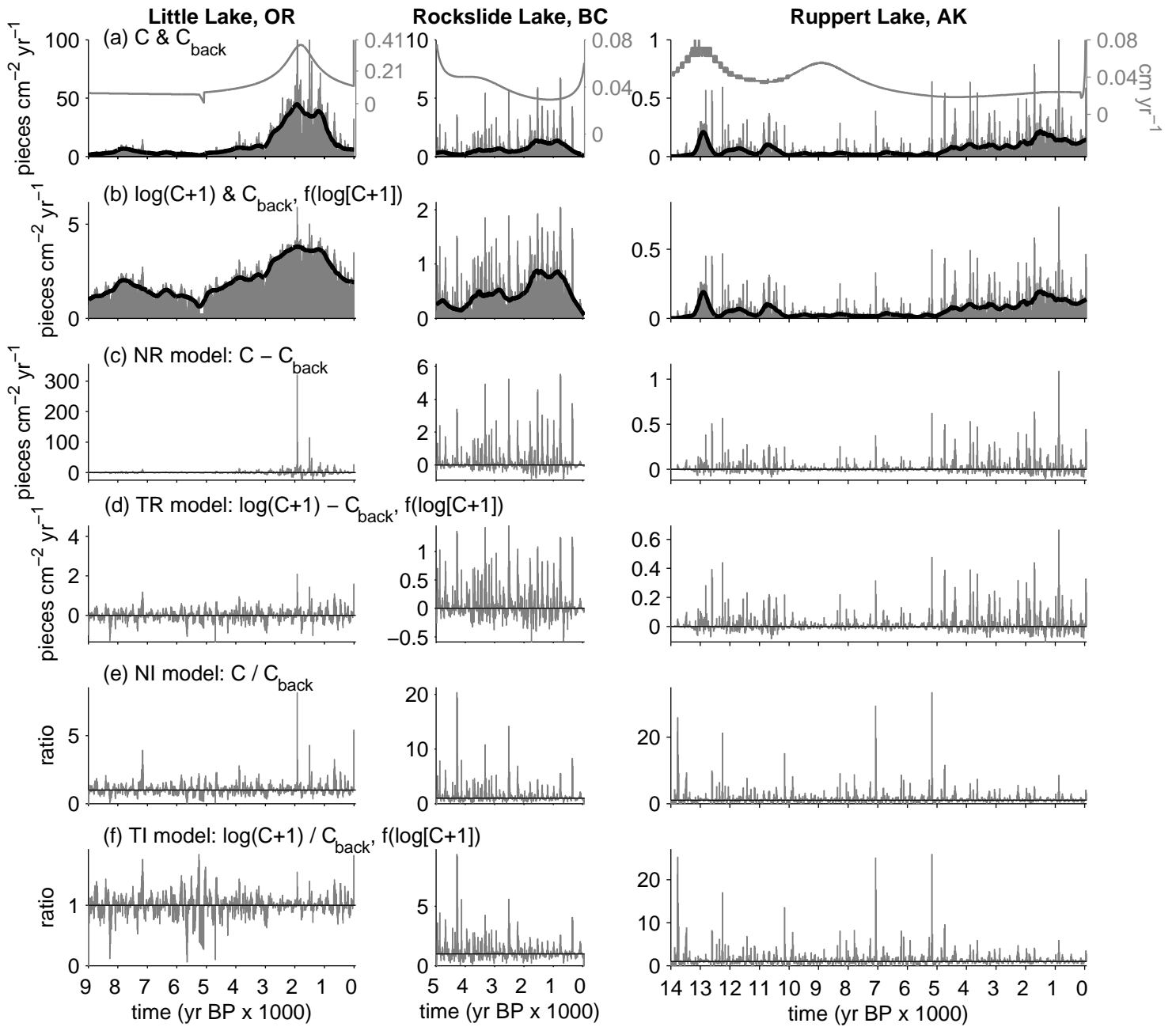


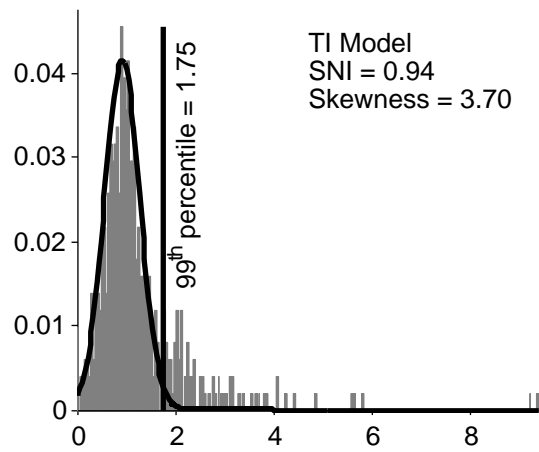
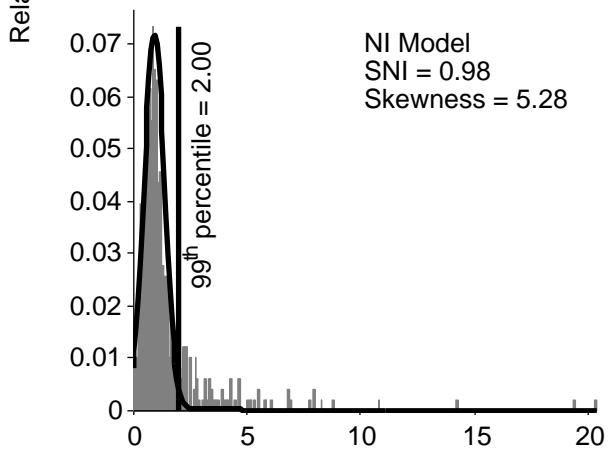
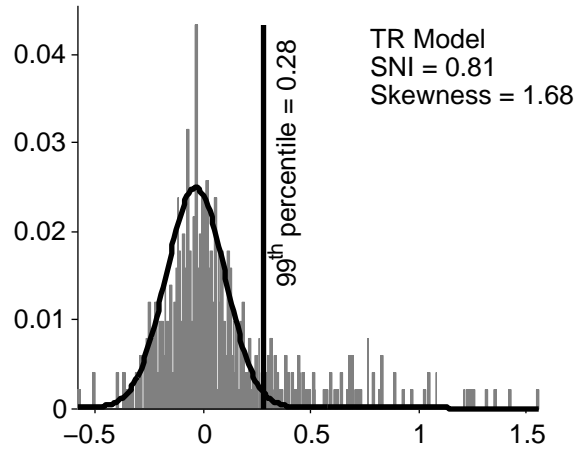
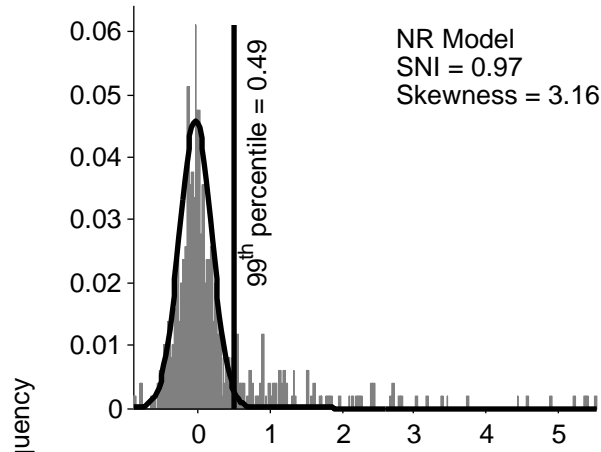
(c) NI model



(d) TI model







Relative frequency

Peak series

