An empirical resampling method for determining optimal high-pass filters used in correlation-based tree-ring crossdating.

Anthony M. Fowler[1], Martin C. Bridge[2,3], Gretel Boswijk[1]

1. School of Environment, The University of Auckland, Auckland, New Zealand.
2. University College London, Institute of Archaeology, London, UK.
3. Oxford Dendrochronology Laboratory, Oxford, UK.

Abstract

Visual crossdating of tree-ring series focusses on high-frequency variations. Automated correlation-based crossdating tools mimic this by transforming raw ring widths into indices that emphasise the high frequency signal, prior to calculating the goodness-of-fit between series. Here we present a resampling methodology to determine the relative merits of alternative simple high-pass filters and demonstrate it using two tree-ring data sets (British Isles oak, New Zealand kauri). Results indicate that: a) high-pass filtering is a critical step; b) the efficacy of alternative filters is variable, and; c) efficacy appears to be species specific. These results have implications for crossdating in the two contexts investigated, and also for future software developments, especially the desirability of flexible implementations of high-pass filtering.

Key words: dendrochronology, crossdating, oak, kauri.

1. Introduction

Visual crossdating may seem routine for a skilled and experienced dendrochronologist working on a familiar species. They are likely to have developed an intimate understanding of what a correct-date match looks like, and the ability to readily distinguish it from the multitude of mostly poor matches at misaligned positions. Some of the latter may occasionally be strong enough to warrant close examination and caution may lead to rejection of some date-aligned samples because the “goodness-of-fit” is too weak for confidence. Determining this goodness-of-fit may involve direct visual inspection of sample pairs under the microscope, or comparison of time series plots of ring widths, perhaps log-transformed or converted into derived indices. Alternatively, crossdating may use abstracted (i.e. reduced) information, such as A.E. Douglas’s skeleton plot technique (Speer, 2010). The specific approach used will be influenced by the dendrochronologist’s training and will probably evolve with experience and experimentation. Moreover, because visual pattern recognition is subjective, different dendrochronologists looking at the same data will inevitably assess goodness-of-fit somewhat differently.

Although there are diverse ways to visually compare temporal patterns, high-pass filtering is ubiquitous. This is explicit in the abstraction methods, where the derived series are essentially reduced to ring-width variations relative to a few adjacent rings (e.g. skeleton plotting), or perhaps first-order differences. It is implicit in other visual approaches where the researcher’s “view” of the sample is limited to a relatively short sequence at any particular point in time. In this case, although short-term trends may be taken into consideration, our experience is that it is always the high-frequency variation about that trend which is most important, especially notably wide and narrow rings, and sometimes sub-decadal signature patterns.

Notwithstanding the diversity of visual approaches to assessing goodness-of-fit, and the inherent subjectivity of associated pattern matching, a generic conceptualisation of the crossdating process is possible (Fig. 1). Consider the case of two time series being compared to each other at many overlapping positions. We know that all but possibly one of these overlap positions is misaligned and we expect to see mostly no or weak agreement between the series at these positions (the frequency curve in Fig. 1). Even if our assessment of goodness-of-fit is purely subjective we implicitly can assess some matches as "-ve", denoting situations of disagreement (i.e. wide rings on one series mostly corresponding to narrow rings on the other), some as showing no meaningful association, and others as "+ve". Because experience and a basic understanding of probability inform us that strong +ve matches at misaligned positions are rare, but possible, the experienced dendrochronologist is likely to impose some sort of goodness-of-fit threshold for what they will accept as possibly indicating a correct match. This threshold will necessarily be somewhat vague for subjective pattern matching and it is likely to evolve with experience in what correct matches look like for the material being investigated. Arrows "A–C" in Fig. 1 indicate three hypothetical date-aligned goodness-of-fit positions: "A" is a match likely to be rejected because it is weaker than the threshold (i.e. there is too much risk that it is a spurious chance match); "B" is a stronger-than-threshold match, but within the bounds of plausible misaligned relationships, and; "C" is the ideal case of a very strong match outside of any seen for misaligned cases.

Although subjective visual crossdating remains a fundamental component of crossdating, various attempts have been made to supplement it, and deal with the undesirable subjectivity, by deriving suitable objective statistics to measure goodness-of-fit. The first of these was B. Huber's '*Gleichläufigkeit'*, developed in the 1930s, which quantifies the percentage of years that two series conjointly increase or decrease (Dean, 1997) – essentially the sign of the first-order difference. The statistic also provided the first means for estimating the statistical significance of a particular match and was automated in the late 1960s (Eckstein, 1972). Shortly after, Baillie and Pilcher (1973, BP73 hereafter) presented the "Belfast Method" of statistical crossdating, which uses Pearson's product-moment correlation coefficient as the goodness-of-fit statistic and Student's *t* as a measure of statistical significance. In this case, goodness-of-fit is calculated on transformed indices, representing relative changes about the local level[1].

Both *Gleichläufigkeit*, and especially the Belfast Method, can be viewed as attempts to objectively automate the concepts underpinning visual pattern matching. They are simplifications, that reduce a sophisticated approach, in which multiple threads of evidence can be synthesised (albeit subjectively), to a single objective statistic. From this perspective, they can be viewed as supplementary crossdating tools, suitable for identifying candidate matching positions that can then be explored in depth using the visual approach (Baillie, 1982). In this context an objective statistic is ideal, because it permits near-instant automated application that crudely mimics visual matching at many thousands of positions. Moreover, conjoint computer-based and visual crossdating is arguably a more sophisticated methodology, because it permits alternative high-pass filtering methods to be explored, some of which are beyond the scope of visualisation of raw ring widths.

[1] Local level is a generic term for evolving time series trend, typically calculated using some form of moving window. The running average is one such. Others include the running median, splines, and digital filters. BP73 used a five-year running mean and calculated annual indices as natural logarithms of percentage change from the mean (Table 1).

Although the conceptual merits of the Belfast Method are widely accepted, including the desirability of high-pass filtering of raw ring-width data prior to calculating goodness-of-fit, specific implementation details related to filtering have been challenged. For example, Munro (1984) showed that crossdating efficacy is filter-dependent, and Wigley et al (1987) noted that running means are problematic, because they introduce phase distortions that increase the frequency of relatively high correlations at mismatched positions. Moreover, it seems likely that filter efficacy will be influenced by species-specific, and perhaps location-specific, characteristics of the ring-width series – such as the frequency of missing rings, autocorrelation, and heteroscedasticity. If so, then a high-pass filtering method that works well in one situation may be sub-optimal when applied elsewhere, resulting in weaker crossdating. More frequent false positives (mismatched positions flagged as statistically significant) and lower statistical significance for date-aligned matches, compared to results obtained using a superior filter, may be a consequence.

This research is based on the assumption that high-pass filters have variable efficacy. This variability may be inherent (some filters are simply better than others), and may also be species and/or place specific. In this context, our aim is to develop and demonstrate an empirical resampling method to objectively quantify the efficacy of alternative filters, based on the conceptualisation of crossdating presented in Fig. 1. We do this in the context of two tree-ring data sets: a) the British Isles oak archaeological sites database compiled by Fowler and Bridge (2015), and; b) the living trees subset of the New Zealand kauri data set (Boswijk et al, 2014). Both data sets were built using, at least in part, the BP73 crossdating methodology, but they are sufficiently different in terms of their respective ring-width data and how BP73 has been applied to provide a useful contrast. We limit our investigation to five simple high-pass filters that are commonly used or could be easily implemented in crossdating software. Our results will have direct relevance to crossdating methodology in the two specific cases investigated, may indirectly facilitate high-pass filter selection in other cases, and will usefully inform future software developments related to computer-assisted crossdating.

2. Data

Oak

The British Isles oak database remains as it was used in Fowler and Bridge (2017), itself updated from that used in Fowler and Bridge (2015). It contains 2024 sites covering the 1000–2010 CE time period. Although not important in the current paper, sites from inner-London have been excluded as they are likely to contain timbers imported into the area from a wide hinterland. The site chronologies used are quite variable in the number of constituent timbers, ranging from as few as three timbers to over 50 in some cases. Most are either from living trees, or from standing buildings. The database will continue to grow and be refined and may thus change in subsequent publications.

Kauri

Kauri (*Agathis australis* (D.Don) Lindl) is a member of the Araucariaceae and the only *Agathis* species endemic to New Zealand. Kauri occurs naturally in the upper North Island with the southern limit at about 38S. It was abundant in lowland forest from sea level to 300 m in Northland, Auckland and Waikato, and up to 700 m in parts of the Coromandel Range (Ecroyd, 1982) and tolerated a range of conditions from lowland bogs to ridge crests. Landscape change since human arrival in the 13th century has resulted in fragmented

forest patches, particularly from fire and logging during the 19th and early 20th century. Most large areas of kauri forest are now preserved as part of the conservation estate.

The trees are large, up to 30 m tall with a straight, thick trunk up to 3 m diameter (Ecroyd, 1982), and can achieve ages >1000 years. Longevity and preservation of kauri wood in bogs made kauri a focus for dendrochronological study for palaeoclimate reconstruction and forest ecology. Since the 1980s, a kauri tree-ring database has been developed at the University of Auckland from three sources of wood: modern (living) trees; archaeological (= historical) wood mostly from 19th century structures; and Holocene age subfossil kauri recovered from lowland bogs (swamp kauri). Most of this data are included in a master tree ring chronology spanning 2488 BCE – 2002 CE, with a low number of older floating subfossil kauri sequences and chronologies (Boswijk et al, 2014). Of relevance here, the modern chronology (AGAUm13) spans 734 years (1269 – 2002 CE). It is derived from ring width data from 228 trees (565 radii) located at 18 sites distributed throughout the natural range of kauri. Series length is typically between 100 and 300 years, with a low number exceeding 300 years. Sample depth is at a maximum in about 1900 CE (>400 radii), steadily decreasing to <50 radii by 1500 CE.

Much of the chronology development carried out since 1999 was undertaken using the Belfast method, integrating statistical (e.g. BP73) and visual matching. Kauri has particular growth characteristics that can present challenges to dating. These include lobate growth, suppression episodes, and wedging resulting in locally absent rings. False rings can occur but wholly missing rings are rare. Because issues such as locally absent rings affect outcomes from automated crossmatching, visual matching using line plots has been important to reconciling ring series as well as checking suggested matches.

3. Methods

High-pass filters

Fig. 2a shows 101 years of mean oak ring-widths for Hengrave Hall and Abbey Farm, two East-Anglia archaeological sites, about 15 km apart. Both have a multi-decadal declining trend, similar decadal-scale patterns (except for about a 20-year period centred on 1460 CE), and inter-annual variability in close agreement. Figures 2b–d show the same data transformed using three high-pass filters: BP73 indices; first-order differences; and logged first-order ratios (Table 1: BP73, FOD, LBFOR). Each filter removes almost all of the decadal to multi-decadal trend, leaving residual high-frequency series that are visually and statistically in closer agreement than is the case for raw ring widths. The scatter diagrams (Panels e–h) show that high-pass filtering has much reduced the scatter, with variance explained (R^2) by linear regression increasing from 32% to 50–53%.

Although the three filters used in Fig. 2 have similar effects for the specific case shown, they are conceptually quite different. The BP73 method converts each ring width into a log-transform of its percentage size relative to the five years centred on that ring. Logged first-order ratios are similar, but measure relative change between adjacent rings. First-order differencing measures absolute width changes from one ring to the next. Under some circumstances the filters result in marked differences in the derived series. For example, ratio-based methods may be superior, in the sense of stronger agreement between correctly dated

series, if local variance depends on local level (heteroscedasticity), but they may be compromised if ring-widths become very narrow, because relative change is bounded low (zero) but not high.

In addition to the three high-pass filters shown in Fig. 2 we investigated two others (Table 1: D, LBR). Both involve fitting a binomial smoothing curve to the data, analogous to the BP73 running mean, then calculating indices as differences (D) or logged ratios (LBR) relative to the fitted curve. The alternative curve fitting was in response to the previously noted criticism of the running mean (Munro, 1984; Wigley et al 1987) and to allow experimentation with the flexibility of the fitted curve. Similarly, the implementation of the BP73 running mean filter was made flexible so that alternative window widths could be explored.

Each of the three ratio-based filters (BP73, LBR, LBFOR) includes a $\log_e$ transform. This is partly for consistency with BP73, but also has the advantage of reducing the scope for extreme outliers where the fitted curve (LBR) or immediately preceding ring (LBFOR) is near zero. However, missing rings (kauri only) are a special case needing workarounds to avoid potential divide-by-zero and $\log_e(0)$ errors. For consistency with how BP73 has been applied in the kauri case, the BP73 algorithm was not changed – instead missing rings were given a value of one. For the other two ratio-based filters, minimum and maximum ratio bounds were set prior to the $\log_e$ transform, with 0.01 and 100 used as initial minimum and maximum values. In the rare case of multiple missing rings, zero divided by zero cases were set to $\mathrm{Log}_e(1)$. Because the fitted binomial smoothing curve reaching zero is very rare (requiring a block of missing rings as long as the number of binomial filter weights), divide by zero issues are largely confined to the LBFOR filter.

Resampling methodology for assessing the merits of high-pass filters

As noted above, the merits of different high-pass filters depend on the specific characteristics of the samples in question. It is possible that an optimal filter for one species may not be so for others, nor perhaps even for the same species under different growing conditions. However, for any specific case (e.g. British Isles oak or New Zealand kauri), we contend that an objective method of determining an optimal filter can be derived by random resampling at misaligned and date-aligned positions, provided that a large crossdated tree-ring database is available. In the context of Fig. 1, resampling of inter-series correlations at misaligned positions empirically defines the frequency distribution and permits formal definition of a suitable “threshold” (e.g. the 0.999 quantile). Resampling of date-aligned correlations provides an analogous frequency distribution and the overlap of the two distributions indicates the efficacy of the high-pass filter. In short, a good filter will minimise the spread of the misaligned distribution (especially the right tail) and maximise the separation of the two distributions. By increasing the percentage of date-aligned correlations to the right of the threshold quantile date-aligned matches are more likely to be flagged. The methodology is illustrated below, building from the Hengrave Hall case (Fig. 2) and expanding to the full British Isles oak database. Note that, because resampling is for fixed window widths, significance can be determined from the frequency distribution of correlation coefficients.

The British Isles oak database has 906 sites with extant tree-ring data at 1450 CE, 315 of which have complete data for 1450–1500 (Fig. 3). Correlating BP73 indices for Hengrave Hall with each of these other sites gives a mean correlation of 0.31. Abbey Farm is the highest of these and there is a spatial pattern of declining correlations to the north and west. Because the Hengrave Hall series is 146 years long we can extend the analysis to 45 different 101 year windows, centred on 1417 through 1461. This brings in

additional candidate comparison sites to those shown in Fig. 3, but the latter drop out at about the same rate, giving approximately 14,000 potential (non-independent) inter-site date-aligned correlations. Fig. 4a shows bootstrapped correlation frequency distributions for these and for misaligned correlations (right and left curves respectively), the latter drawn from millions of potential combinations of sites and dates. Fig. 4b is an equivalent analysis for raw ring widths. Comparing the two sets of plots demonstrates three significant impacts of BP73 high-pass filtering:

a) The frequency distribution for misaligned correlations is now centred on zero and has a shape approaching the normal distribution.

b) The spread of the frequency distribution for misaligned correlations is reduced, therefore reducing the threshold for statistical significance.

c) The separation between the date-aligned and misaligned correlation frequency distributions is increased. This "right-shift" of the frequency distribution for date-aligned correlations (relative to misaligned) increases the proportion of correct dates distinguishable as statistically-significant matches[2].

Extending the Hengrave Hall analysis to all sites across the full oak database permits random sampling of several million date-aligned pairs and many times that number of misaligned pairs. Fig. 5a shows bootstrap correlation frequency distributions for unfiltered data with fixed series lengths of 101 years at misaligned and date-aligned positions. The former is derived from one million random pairs and the vertical dashed line in Fig. 5a is the associated 0.999 quantile (R = 0.793). The distribution for date-aligned pairs is derived from 10,000 pairs and the shaded area beneath the right tail of the curve denotes date-aligned correlations greater than the 0.999 quantile for misaligned positions. Fig. 5b shows comparable results for fixed series lengths of 51 years. Here we adopt the percentage of date-aligned correlations greater than the 0.999 quantile for misaligned positions as our statistic for high-pass filter efficacy (Filter Efficacy Score, FES), with the results for unfiltered data providing baseline FES against which the performance of high-pass filters can be compared. For the inter-site oak case discussed here this is less than one percent for both series lengths tested. Fig. 5c,d shows the impact on FES of BP73 high-pass filtering. Note that right-shift is slightly higher for the shorter series (median 0.324 vs. 0.306), but the impact of this on the overlap of the respective frequency distributions is overwhelmed by conjoint changes in spread. This represents the familiar increasing difficulty of crossdating shorter sequences, because higher correlations are required for a correct match to emerge from the "background" of potential spurious (i.e. misaligned) correlations. All FES analyses undertaken here were repeated for series lengths of 51 and 101 years (FES_{51} and FES_{101} respectively) for each of the high-pass filters listed in Table 1, and for both the oak and kauri data sets.

[2] The fact that date-aligned correlations for raw ring widths tend to be higher than correlations on BP73 indices in this case is not relevant to the right-shift argument, nor is it a characteristic feature when the analysis is extended to multiple sites. This Hengrave Hall result probably reflects the effect of the negative growth trend for the site (Fig. 2a). Because this is a characteristic tree-ring width series growth trend, raw ring-width correlations will have a positive bias at both date-aligned and misaligned positions.

4. Results & Discussion

Oak

Oak inter-site baseline FES_{51} and FES_{101} (i.e. for raw ring widths) are <1% (Fig. 5a,b). This means that very few correlations of date-aligned raw ring widths will be statistically significant against the background of random correlations at misaligned positions. BP73 high-pass filtering markedly improves this situation. Most importantly, the spread of the frequency distribution for correlations at misaligned positions is much reduced, with an associated reduction in the 0.999 quantile correlations, from about 0.8 (Fig. 5a,b) to less than 0.5 (Fig. 5c,d). Similar reduction in spread of the date-aligned correlations results in a separation of each pair of frequency distributions, with an associated increase in the percentage of statistically significant date-aligned correlations. The percentage increases with series length.

As previously noted, one of the criticisms of the BP73 approach is that the five-year running mean is a suboptimal high-pass filter (Munro, 1984; Wigley et al., 1987). To investigate this, we compared running means to binomial filters with a similar approach of logging relative changes in ring width about the local level (LBR in Table 1). Multiple running window widths were investigated for each filter. Solid lines in Fig. 6 show filter efficacy scores for running mean windows of width 3–15 years, and for series lengths of 51 and 101 years. Dashed lines show equivalent results for the LBR filter, where window width refers to the number of filter weights. Two results are noteworthy. First, five years is confirmed as the optimal window width for BP73. Second, the LBR filter is clearly superior, with efficacy highest at seven filter weights. Comparing detailed results for the latter (Fig. 7) with those for BP73 (Fig. 5c,d) indicates that the improved performance is almost entirely due to the reduced spread of the frequency curve for misaligned correlations, with no meaningful change in the right shift. In short, the poorer BP73 performance is caused by a higher frequency of spurious high correlations at misaligned positions (e.g. 44% more correlations greater than 0.3 for series lengths of 101 years).

Efficacy scores for the other high-pass filters are plotted against the right axis of Fig. 6. It is noteworthy that the difference-based filters (FOD, D) are generally poorer performers than the relative change ones (BP73, LBR, LBFOR), especially for the 101-year series length, provided that the optimal window width is used for the first two of these. Of the three relative-change filters, LBFOR achieves the highest efficacy scores for both series lengths investigated. This suggests that, for British Isles oak, relative change from one ring to the next, rather than variation about the local level, is marginally optimal for inter-site crossdating.

The inter-site analyses above were repeated, but this time correlating site chronologies against the all-site British Isles master. In this case baseline FES values are higher (FES_{51} = 21%, FES_{101} = 14%), due to less decadal-scale trend in the master chronology compared to individual sites. Results for each of the high-pass filters are shown in Fig. 8 and detailed results for one of the strongest (9-weight LBR) are given in Fig. 9. The most obvious difference between these and the inter-site results (Fig. 6, Fig. 7) is the more than doubling of the FESs, caused by the enhanced right-shift of the date-aligned correlations. The frequency distributions for correlations at misaligned positions are essentially identical to those for the inter-site analyses and the right-shift enhancement is simply a consequence of the stronger common signal in the master chronology due to massive replication. Inter-comparison of the BP73 and LBR filters (Fig. 8) gives very similar results to the inter-site analysis (Fig. 6). The LBR filter is again superior to BP73, although the difference is smaller in this

case. The optimal number of weights for the binomial smoothing curve is greater than indicated by the inter-site results, but is also less clear-cut. Nine to eleven weights appear to be the best choice, but sensitivity is low. The LBFOR filter again performs well, but is no longer clearly differentiated from other options. In fact, although all of the alternative filters perform better than BP73, there is little to differentiate between them, and improvements over BP73 are relatively small (0.3–4.3% increases in FES).

The results presented here confirm and quantify the importance of high-pass filtering for correlation-based oak crossdating. Without it, a statistically significant inter-site dating match will be extraordinarily rare (Figure 5a,b). BP73 high-pass filtering improves the situation significantly (Fig. 5c,d), but alternative simple filters, especially those based on relative changes in ring width, generally perform better (Fig. 6, Fig. 8). Although incremental improvements over BP73 are fairly small, they are not trivial. For example, comparing BP73 to LBFOR for typical series lengths of about 100 years, the latter would correctly flag about 14% more date-aligned inter-site correlations as statistically significant[3]. The increase is only 2% in the site vs. master case, in part because BP73 FES_{101} is already high (91.6%).

Kauri

The results reported above are specific to British Isles oak site chronologies, and applicability to other species, to other regions, and to non-site chronologies cannot reasonably be assumed. This is because different species in different environments may have characteristics (e.g. the frequency of missing rings) that influence which high-pass filter is optimal. To investigate this we investigated how the various high-pass filters perform in the case of New Zealand kauri for the more challenging context of crossdating individual radii against a kauri master chronology. The additional challenge relates to kauri radii often having missing rings and also frequently exhibiting notable heteroscedasticity (Fowler, 2009). For the purpose of this experiment we analysed radii from living trees and crossdated them against the Boswijk et al (2014) mean ring width master chronology for living trees (AGAUm13, Section 2).

Fig. 10a shows the AGAUm13 mean ring-width chronology, with 19-weight binomial smoothing curve, for the period 1700–1900. Panel b shows ring widths for radius CAS018A for the same period, also with fitted curve. CAS018A has a fairly common ring-width pattern for kauri of declining trend with age, with an associated decrease in evolving variance (i.e. heteroscedasticity). There are several quite narrow rings in the last few decades, including a single missing ring at 1839. Other time series plots in Fig. 10 are CAS018A indices, calculated using the five high-pass filters in Table 1. BP73 uses the original 5-year running mean with the 1939 ring changed to one, the D and LBR indices were calculated using the 19-weight binomial smoothing curve (discussed below), and the LBFOR and LBR indices were bounded by minimum and maximum ratios of 0.01 and 100. Scatter plots immediately to the right of each time series are indices for AGAUm13 (x-axis) plotted against CAS018A (y-axis), in each case using the same index calculation method used by the paired time series plot (e.g. Panel h is BP73 indices). The two additional scatter plots (Panels n, o) are additional LBFOR and LBR analyses with the tighter ratio bounds of 0.2 and 5 shown by the dashed lines in Panels e and g. Note that scatters are plotted over the full data range in each case, making visual comparison across plots potentially misleading. For example, the only difference between Panels k and n is the scaling of the two circled points.

[3] BP73 FES = 38.0. LBFOR FES = 43.3. Percent increase = 100 x (43.3 - 38.0) / 38.0 = 13.9%

297 FES dependence on the running mean window width (BP73) and on the number of binomial smoothing
298 curve weights (LBR) is shown in Fig. 11. As with oak (Figs. 6, 8), five years is the optimal window width for
299 the BP73 method, although FES decline as the window width increases is much reduced. The optimal
300 number of binomial weights is in the range 17–21, which is notably higher than for oak. Both results indicate
301 that kauri sub-decadal trend provides some useful information for crossdating purposes. Comparing the 7-
302 weight and 19-weight LBR results indicates that the higher FES of the latter is solely caused by reduced
303 spread of correlations at misaligned positions. In other words, widening the high-pass filter window does not
304 improve agreement at date-aligned positions, but it does reduce the frequency of statistically-significant
305 correlations at misaligned positions. On this basis, a 19-weight binomial smoothing curve was adopted for
306 the D and LBR indices (Fig. 10f,g).

307 Thin dotted lines in Fig. 11 show kauri LBR FES results with the same settings used for oak applied. In sharp
308 contrast to the oak results (Figs. 6, 8), the LBR filter is not superior to BP73. Moreover, results for the
309 LBFOR filter (FES_{51} = 54%, FES_{101} = 81%), one of the best for oak, are clearly inferior. Part of the
310 explanation for the relatively poor performance of the alternative ratio-based filters (LBFOR, LBR) lies with
311 how very narrow rings, especially missing rings, are handled. As noted in Section 3, logged ratio-based
312 methods encounter issues with missing rings, which led to the bounded ratios approach adopted here. This
313 was not important in the oak case because we were dealing with mean ring-width site chronologies with few
314 near-zero values, but it is problematic for kauri because missing rings on radii are not uncommon. The
315 nature of the problem is clear in the case of the 1839 missing ring in Fig 10. For all three ratio-based indices
316 (BP73, LBFOR, LBR) the missing ring is a clear outlier, as is 1840 in the case of LBFOR. The net effect is to
317 inflate correlations where missing or very narrow rings are aligned. For example, the linear regression R^2 of
318 0.20 for LBFOR indices (Panel i) reduces to 0.13 if the two circled outliers are removed. In response to this,
319 different ratio bounds were explored, with a view to minimising the outlier effect. We explored maximum
320 (minimum) ratio bounds from 1000 (0.001) to 5 (0.2), the latter assumed to be reasonable limits to the
321 bounding band, because anything more restrictive would begin to clip non-zero rings (see Fig. 10e,g). In
322 terms of the FES, performance monotonically improved as the ratio bounding band was reduced, down to
323 the 5 (0.2) limits. Fig. 10n,o show the impact of imposing this limit on the CAS018A LBFOR and LBR indices
324 (others are unaffected) and the heavy dashed lines in Fig. 11 show the improved LBR performance,
325 compared to the wider 100 (0.01) ratio bounds. The higher revised LBR FES_{101} (91%) than BP73 (88%)
326 appears to be almost entirely due to ratio bounding, which BP73 does not have.

327 The performance of the D, FOD, and LBFOR (with revised ratio bounds) high-pass filters are plotted against
328 the right axis of Fig. 11. All three have lower FESs than LBR and both filters that deal only with changes from
329 one year to the next have lower FESs than BP73. The latter is consistent with the argument that kauri
330 crossdating benefits from a broader perspective than how growth in one year relates to growth in the last.
331 Also interesting is the relatively good performance of the D filter. The fact that this is only marginally lower
332 than LBR suggests that absolute changes about the local level may be as useful for crossdating as relative
333 changes. This is a particularly interesting result because it seems likely that performance of the D high-pass
334 filter is degraded by typical heteroscedasticity, i.e. variance of local width about the local level scales with the
335 local level, in kauri ring-width time series (Fowler, 2009). The CAS018A FOD and D indices (Fig. 10d,f)
336 illustrate this very well, with higher variance over the period 1700–1745 associated with higher local level
337 (Fig. 10b). Rescaling the indices over this period to give a similar range to those after 1745 reduces the

scatter in Fig. 10l and increases the linear regression R^2 from 0.16 to 0.20. Note that ratio-based indices usually reduce heteroscedasticity (Fritts, 1976), but they often introduce “reversed” heteroscedasticity in the case of kauri, although this is usually less pronounced (Fowler, 2009; Fig. 10c,e,g).

In summary, the kauri data exposed a critical weakness of the ratio-based high-pass filtering methods in the presence of missing rings, highlighting the potential dangers of transferring methods to situations different to those in which they were developed. Ironically, the BP73 method, with the simple expedient of changing zeros to ones, proved relatively robust and one of the best performing indices until the ratio bounds were significantly tightened. In agreement with oak, high-pass filtering is again critical, but both the width of the filtering window and the optimal filter type differ. Kauri crossdating is best served by a slightly wider filter window and a simple difference-based filter shows promise.

5. Summary and Conclusions

The research presented here is framed in terms of computer-assisted crossdating, where automated correlation-based methods may usefully flag potential crossdating positions, but subsequent visual inspection of plots (and wood where possible) remains a fundamental requirement. In view of the time-consuming nature of the visual pattern-matching second stage, automation should ideally flag relatively few potential dating positions, with maximum likelihood that any correct date will be flagged. In terms of the conceptualisation of crossdating presented here, that means minimising the spread of the frequency distribution of correlations at misaligned positions (although actually it is only the right tail of the distribution that matters) and maximising separation of the distribution for date-aligned positions (what we have referred to as “right-shift”). The merit of high-pass filtering to achieve these ends is not in doubt, essentially mimicking the visual approach to crossdating, but our results usefully quantify how critical it is. Certainly, at the individual site (oak) and radii (kauri) levels, explored here, attempts to crossdate raw ring widths would clearly be a futile undertaking, with less than a 5% chance of date-aligned series being flagged as statistically significant.

Although high-pass filtering is clearly important, exactly what form of filter might be optimal is not obvious. A key part of our research was therefore to develop a resampling methodology that statistically approximates the conceptualisation of crossdating presented in Fig. 1, and provide a mechanism to objectively quantify the efficacy of alternative high-pass filters. We adopted the percentage of correlations at date-aligned positions above the 0.999 quantile of correlations at misaligned positions as our metric – the Filter Efficacy Score (FES). The relative merits of five simple high-pass filters were then determined by comparing FESs for identical resampling experiments.

The high-pass filters investigated here are all relatively simple. This is because our main focus was on developing the methodology for assessing the relative merit of filters (i.e. FES), but also because we thought it best to start with filters that are already implemented in existing crossdating software, or could be added relatively easily because they follow a similar approach. BP73 is commonly used, D and LBR follow the same approach of fitting a flexible smoothing curve, and FOD and LBFOR are very easy to implement. The mix also allowed comparison of filters based on relative changes (BP73, LBR, LBFOR) and absolute changes (D, FOD). Logging ratios was included in LBR and LBFOR for consistency with BP73 and bounding

of ratios was added to deal with issues related to missing rings. The use of a binomial smoothing curve in D and LBR instead of a running mean (BP73) was in response to statistical criticisms related to the latter. All five filters worked reasonably well, as initially implemented, when applied to the British Isles oak database, but significant issues related to missing rings emerged in the case of kauri. It was apparent that setting ratio bounding too high (e.g. 0.01, 100) resulted in outlier indices for missing rings, with highly undesirable impacts on the correlation frequency distributions, which in turn lowered FESs for the two ratio-bounded filters (LBR, LBFOR). This problem was solved by reducing the bounding range (to 0.2–5, Fig. 11).

For oak, our results indicate that LBFOR is the best performing filter overall, although only marginally so in some circumstances. For kauri, LBR appears to be best, but D is not far behind and perhaps may match LBR if heteroscedasticity is minimised using an approach similar to Cook and Peters (1997). Moreover, comparison of the oak and kauri results reveals some notable differences in the performance of the various high-pass filters. First, optimal window widths for the LBR (Figs. 6, 8, 11) and D filters are different for oak and kauri. The results for BP73 show comparable changes, although in this case the 5-year moving average window remains optimal. Second, the two filters that measure only changes from one year to the next (FOD, LBFOR) perform well for oak, but are poor candidates for kauri. Third, there are some notable differences in the performance of the relative change (BP73, LBR, LBFOR) and absolute change filters (D, FOD). The former are clearly superior for oak inter-site crossdating (Fig. 6), but not for kauri (although they still perform reasonably well).

The inter-species differences, noted above, demonstrate potential issues when transferring filters that have a proven track record for one species to another. It follows that a “one size fits all” approach to high-pass filtering is inappropriate and that detailed analysis is required to determine what will work best in different circumstances (e.g. Fig. 10). However, we think it important to stress that, as long as correlation-based crossdating is used as a precursor step to visual pattern matching, the main impact is likely to be to speed up the crossdating process. Although high-pass filtering is certainly critical, it appears that filter choice is a matter of refinement. At worst, using a sub-optimal filter will simply make the crossdating task somewhat more difficult, by increasing the frequency of false positives and reducing statistical significance associated with correct matches. Moreover, in practical terms, what can be achieved depends on how correlation-based crossdating is implemented in readily-available software. In that context, we encourage flexible coding that allow experimentation and fine tuning.

406 **References**

407 Baillie, M.G.L. 1982. *Tree Rings and Archaeology*. University of Chicago Press, Chicago.

408 Baillie, M.G.L., Pilcher, J.R. 1973. A simple crossdating program for tree-ring research. *Tree-Ring Bulletin,*
409 **33**, 7–14.

410 Boswijk, G., Fowler, A.M., Palmer, J.G., Fenwick, P., Hogg, A., Lorrey, A., Wunder, J. 2014. The late
411 Holocene kauri chronology: assessing the potential of a 4500-year record for palaeoclimate
412 reconstruction. *Quaternary Science Reviews*, **90**, 128–142.

413 Cook, E. R., Peters, K. 1997. Calculating unbiased tree-ring indices for the study of climatic and
414 environmental change. *The Holocene*, **7**, 359–368.

415 Dean J.S. 1997. "Dendrochronology". Chapter 2 in Taylor R.E. and Aitken M.J. (eds.) *Chronometric Dating in*
416 *Archaeology.* Springer, New York.

417 Eckstein, D. 1972. Tree-ring research in Europe. *Tree-Ring Bulletin*, **32**, 1–18.

418 Ecroyd, C.E. 1982. Biological flora of New Zealand 8. *Agathis australis* (D. Don) Lindl. (Araucariaceae) Kauri.
419 *New Zealand Journal of Botany*, **20**, 17–36.

420 Fowler A. M. 2009. Variance stabilization revisited: a case for analysis based on data pooling. *Tree-Ring*
421 *Research*, **65**, 129–145.

422 Fowler, A.M., Bridge, M.C. 2015. Mining the British Isles oak tree-ring data set. Part A: rationale, data,
423 software, and proof of concept. *Dendrochronologia*, **35**, 24–33.

424 Fowler, A.M., Bridge, M.C. 2017. Empirically-determined statistical significance of the Baillie and Pilcher
425 (1973) *t* statistic for British Isles oak. *Dendrochronologia*, **42**, 51–55.

426 Fritts, H. C., 1976. *Tree Rings and Climate*. Academic Press, London.

427 Munro, M.A.R., 1984. An improved algorithm for crossdating tree-ring series. *Dendrochronologia*, **44**, 17–27.

428 Speer J.H. 2010. *Fundamentals of Tree-Ring Research*. University of Arizona Press, Tucson.

429 Wigley, T.M.L., Jones, P.D., Briffa, K.R. 1987. Cross-dating methods in dendrochronology. *Journal of*
430 *Archaeological Science*, **14**, 51–64.

431

Figure captions

Fig. 1. Conceptualisation of visual crossdating. Goodness-of-fit is the dendrochronologist's subjective assessment of the agreement between two series being compared at multiple misaligned positions. Arrows "A", "B", and "C" denote possible date-aligned positions and "Threshold" is the, again subjective, standard required for a date to be considered plausible. Modified after Fowler and Bridge (2017).

Fig. 2. Hengrave Hall and Abbey Farm time series and cross-correlations (1400–1500). **a)** Simple ring-width plots. **b)** BP73 high-pass-filtered data: natural logs of ring widths expressed as percentages of the five years each ring is centred on (BP73). **c)** First-order differences (FOD). **d)** Natural logs of bounded first-order ratios (LBFOR). **e–h)** Scatterplots of the data in a–d, with linear regression lines and associated correlations. Details of high-pass filters are in Table 1.

Fig. 3. Correlation of Hengrave Hall (grey circle with black frame) BP73 indices with 314 other sites with full data available for the 101 years centred on 1450. Colour intensity shows correlation strength (see histogram) and the circle area (scaled to Students *t*) is a guide to statistical significance. The large circle northeast of Hengrave Hall is Abbey Farm (Fig. 2). Screen capture from *Oak Mapper* (Fowler & Bridge, 2015).

Fig. 4. Bootstrap (N=100,000) correlation frequency distributions for Hengrave Hall against other sites in the British Isles oak database at date-aligned and misaligned positions. **a)** BP73 indices. **b)** Raw ring widths. Series length (span) is fixed at 101 years.

Fig. 5. Oak inter-site bootstrap frequency distributions of correlations at misaligned (N=1,000,000) and date-aligned (N=10,000) positions. **a)** Unfiltered data and fixed series length (span) of 101 years. **b)** Unfiltered, span 51. **c)** BP73 filtered, span 101. **d)** BP73 filtered, span 51. Vertical dashed lines are empirically-derived 0.999 quantiles (P = 0.001) of correlations for the misaligned series. Shading denotes date-aligned correlations higher than the 0.999 quantile for correlations at misaligned positions – the Filter Efficiency Score (FES).

Fig. 6. Comparison of inter-site filter efficiency scores (percentage of all-site date-aligned correlations above the 0.999 quantile for misaligned correlations). Plotted values correspond to the shaded areas in Figure 5 and are derived from analysis of identical all-site bootstrap (N=1,000,000) correlation frequency distributions for the British Isles oak database, for series lengths of 51 and 101 years. Solid lines are results derived using the BP73 high-pass filter method (BP73_051, BP73_101), but with variable running mean window widths. Dashed lines are equivalent results for logged bounded ratios, derived for a fitted binomial filter, where window width is the number of filter weights (LBR_051, LBR_101). Coloured bars on the right are results for three other high-pass filters: first-order differences (FOD_051, FOD_101); logged bounded first-order ratios (LBFOR_051, LBFOR_101); and differences relative to a fitted 7-weight binomial filter (D_051, D_101).

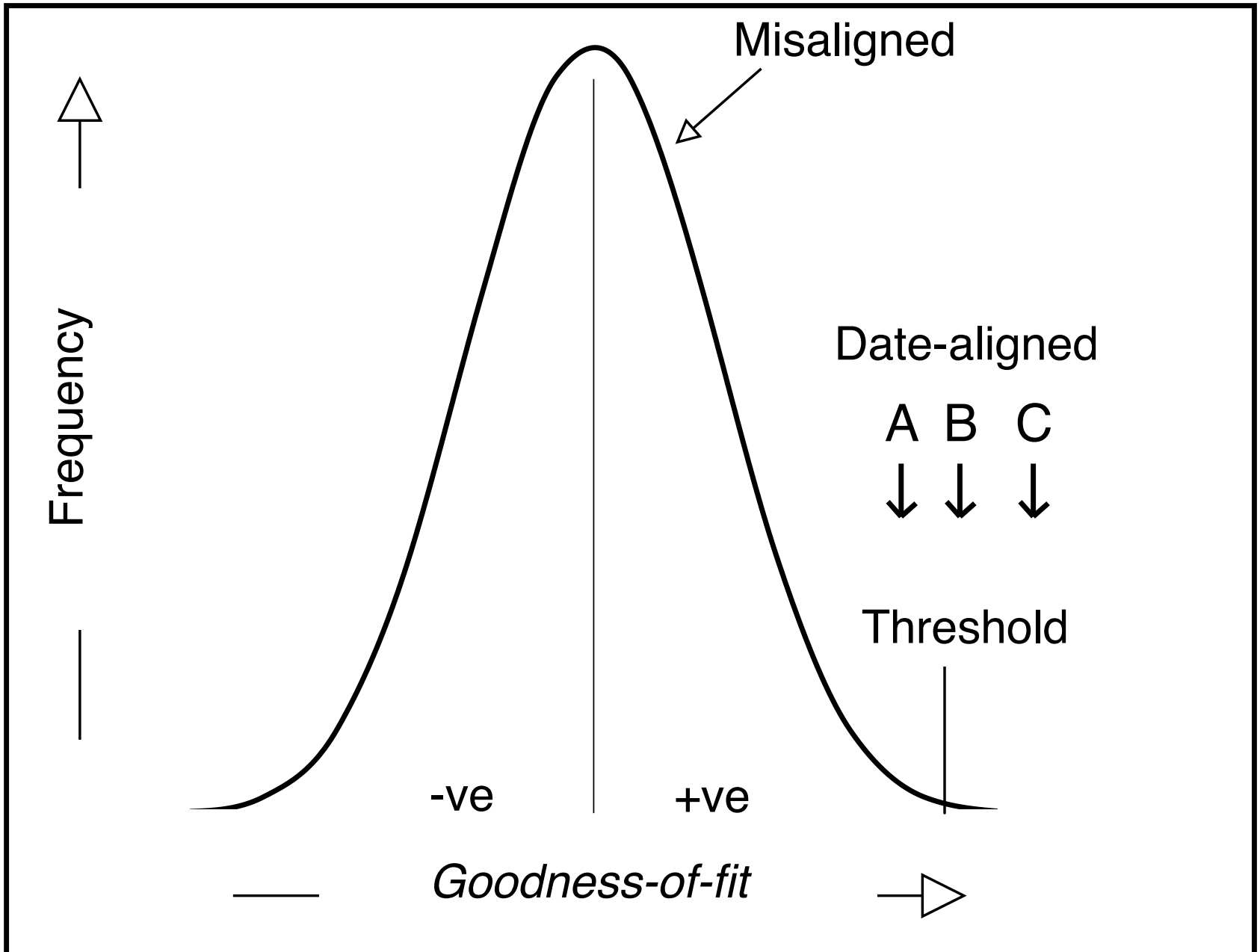
Fig. 7. Same as Fig. 5, for a 7-weight LBR high-pass filter (Table 1).

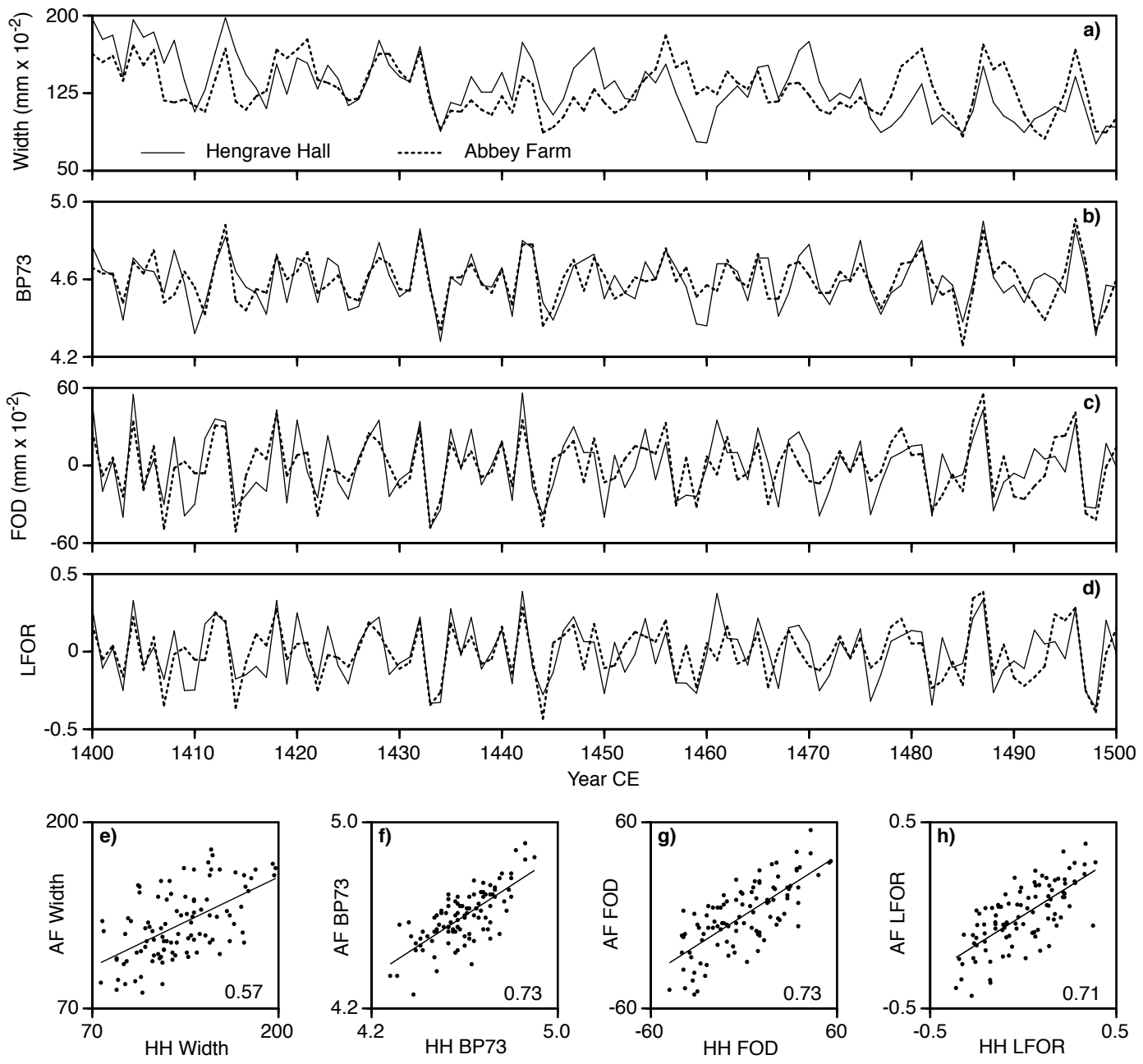
Fig. 8. Same as Fig. 6, for correlations between sites and an all-site British Isles master chronology.

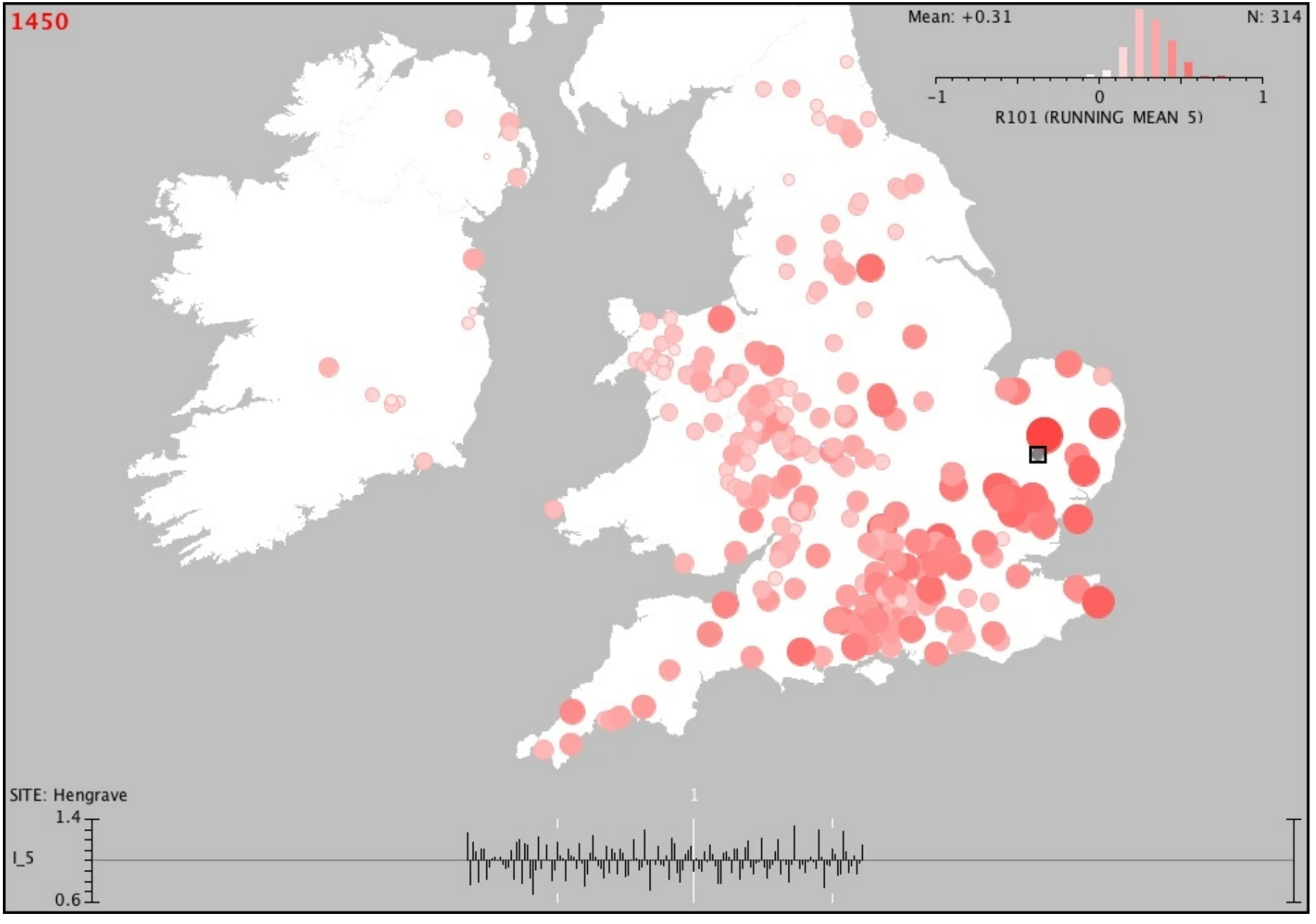
Fig. 9. Same as Fig. 5c and 5d, for a 9-weight LBR high-pass filter and for correlations between sites and an all-site British Isles master chronology.

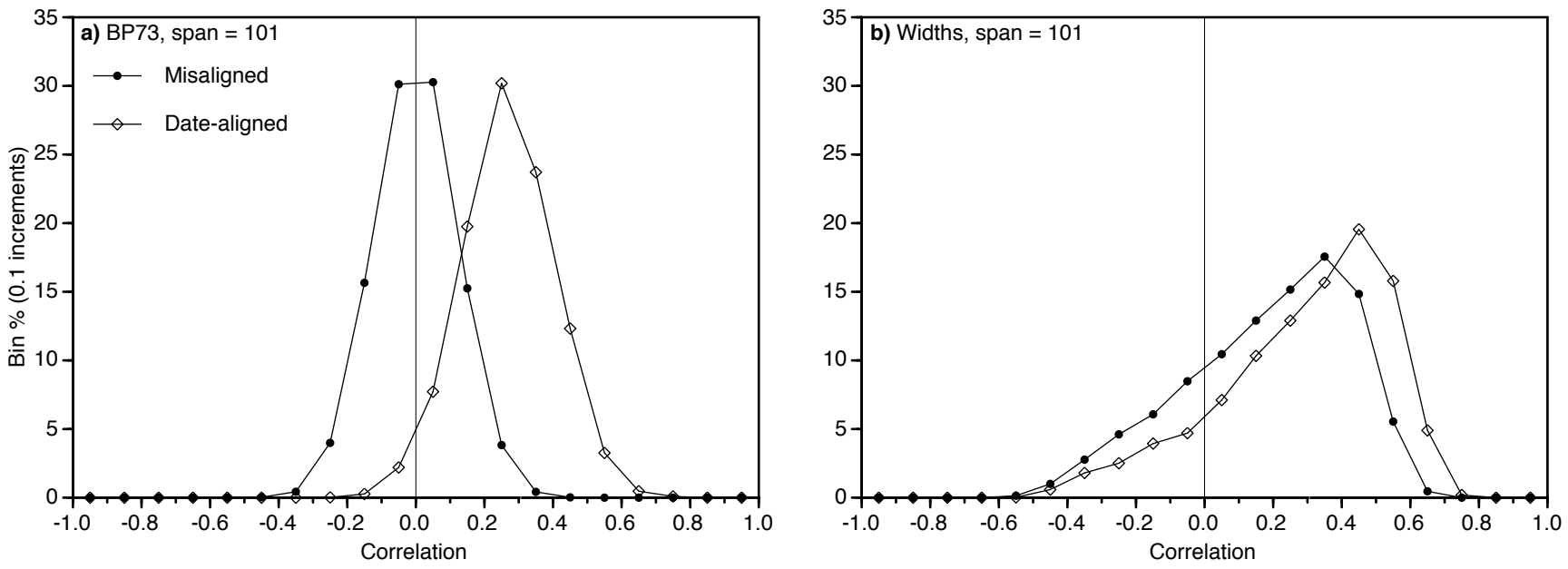
Fig. 10. Example of high-pass filter performance for kauri. **a,b)** Ring-width plots for AGAUm13 and radius CAS018A. **c–g)** CAS018A indices for the five high-pass filters (Table 1). Ratio bounds of 0.01 and 100 applied to LBFOR and LBR. **h–m)** Scatter plots for CAS018A indices (y-axis) against AGAUm13 indices calculated in the same way. Linear regression lines are shown and the regression R^2 is given in the bottom right corner of each plot. Circled data points are 1839 (missing ring) and 1840. **n,o)** Revised scatter plots for ratio bounds of 0.2 and 5.

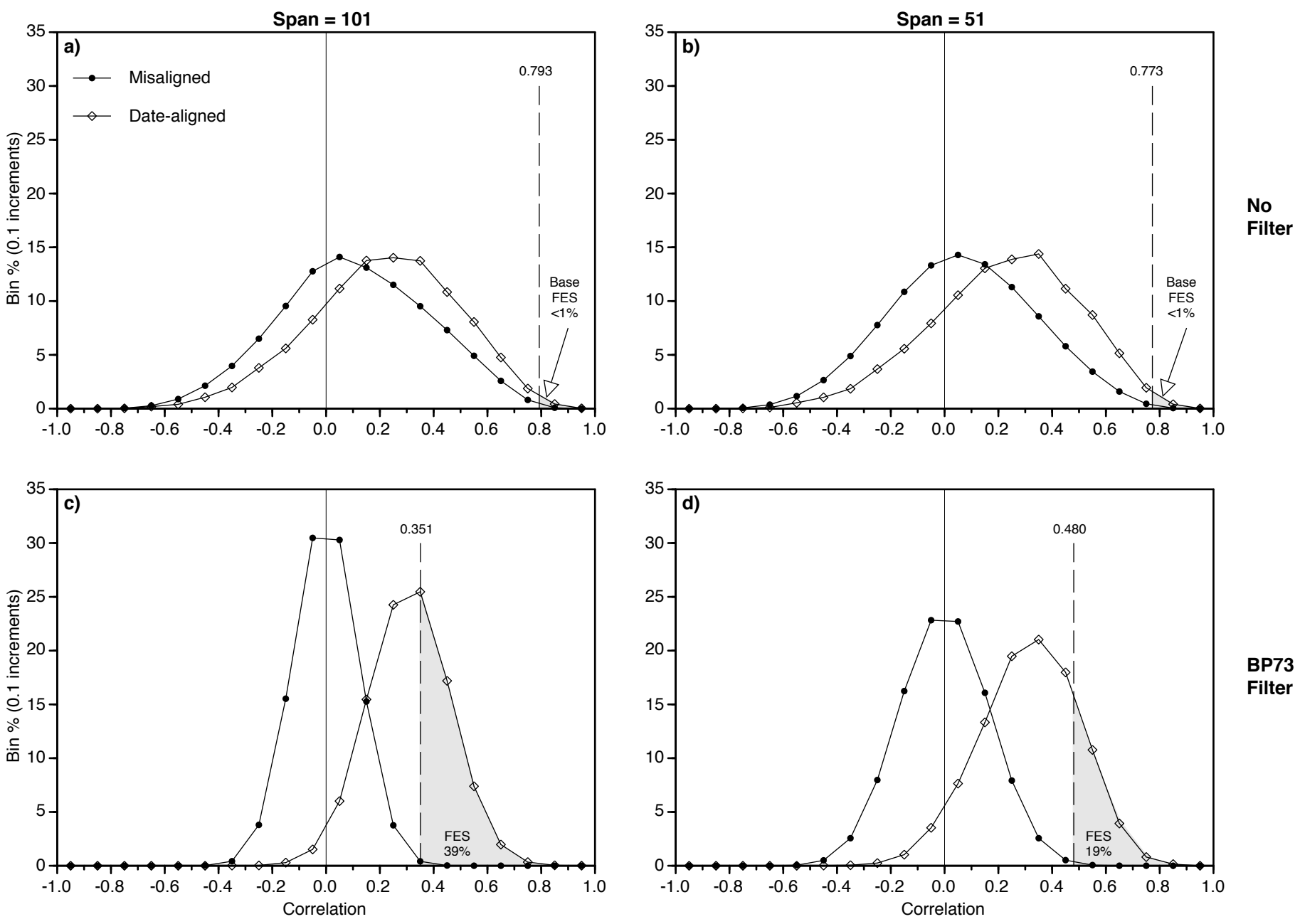
Fig. 11. Same as Fig. 6 for Kauri radii-AGAUm13 cross-correlations. Thin dashed lines are LBR results for ratio bounds of 0.001 and 100. Heavy dashed lines are for revised ratio bounds (0.2, 5), as are LBFOR results.

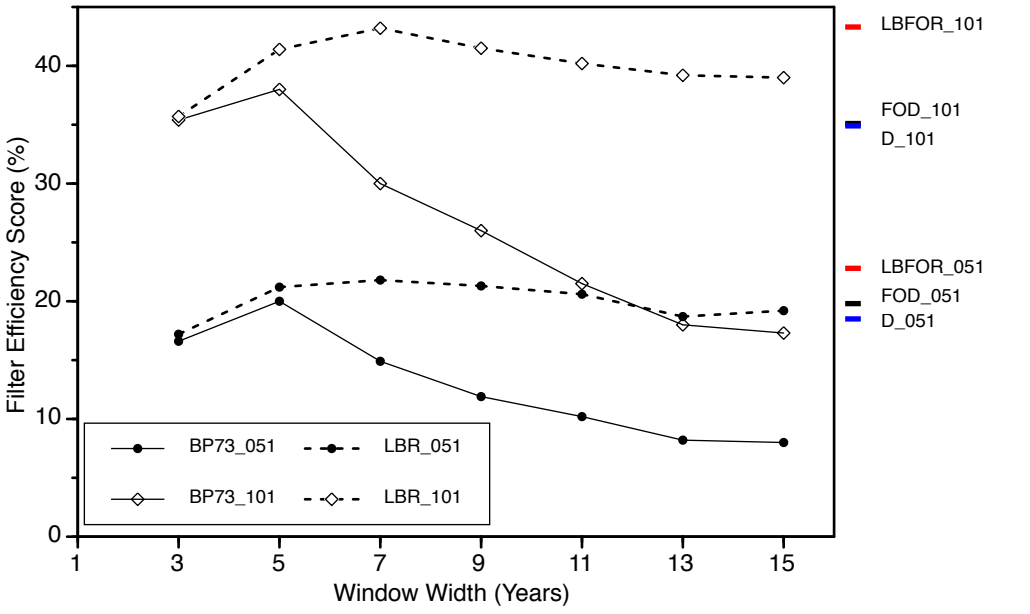

Misaligned
Frequency
Date-aligned
A B C
Threshold
-ve
+ve
Goodness-of-fit

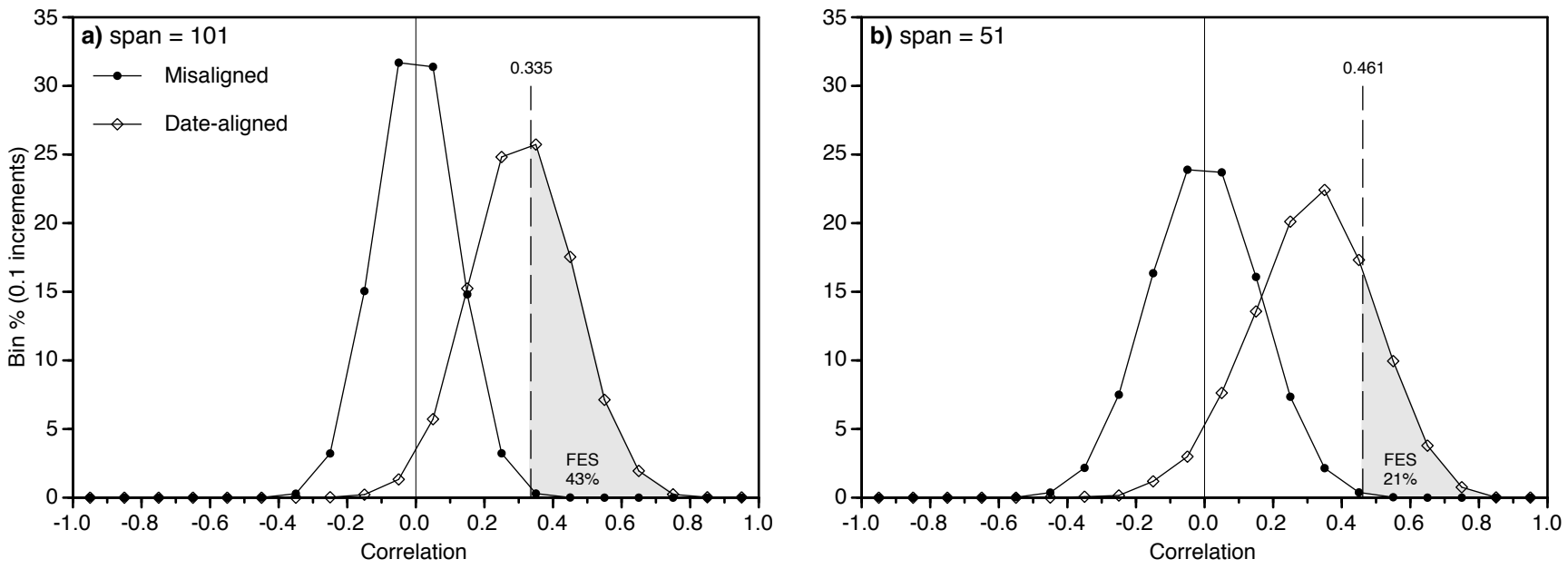

a)
Width (mm x 10-2)
200
125
50
Hengrave Hall
Abbey Farm
b)
BP73
5.0
4.6
4.2
c)
FOD (mm x 10-2)
60
0
-60
d)
LFOR
0.5
0
-0.5
1400
1410
1420
1430
1440
1450
1460
1470
1480
1490
1500
Year CE
e)
AF Width
HH Width
0.57
f)
AF BP73
HH BP73
0.73
g)
AF FOD
HH FOD
0.73
h)
AF LFOR
HH LFOR
0.71

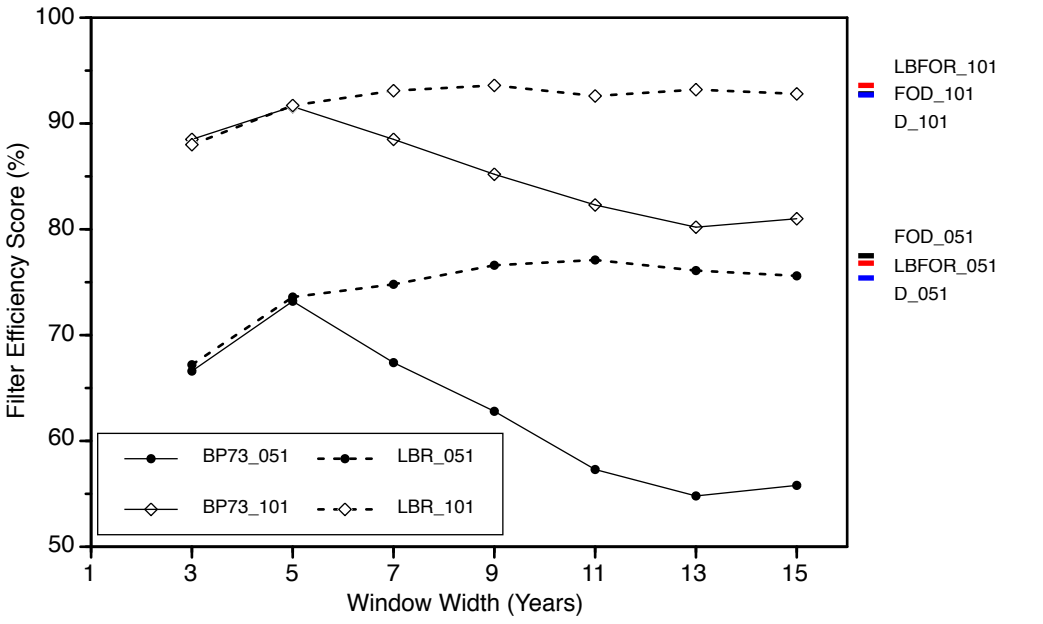

1450
Mean: +0.31
N: 314
-1
0
1
R101 (RUNNING MEAN 5)
SITE: Hengrave
1.4
I_5
0.6
1

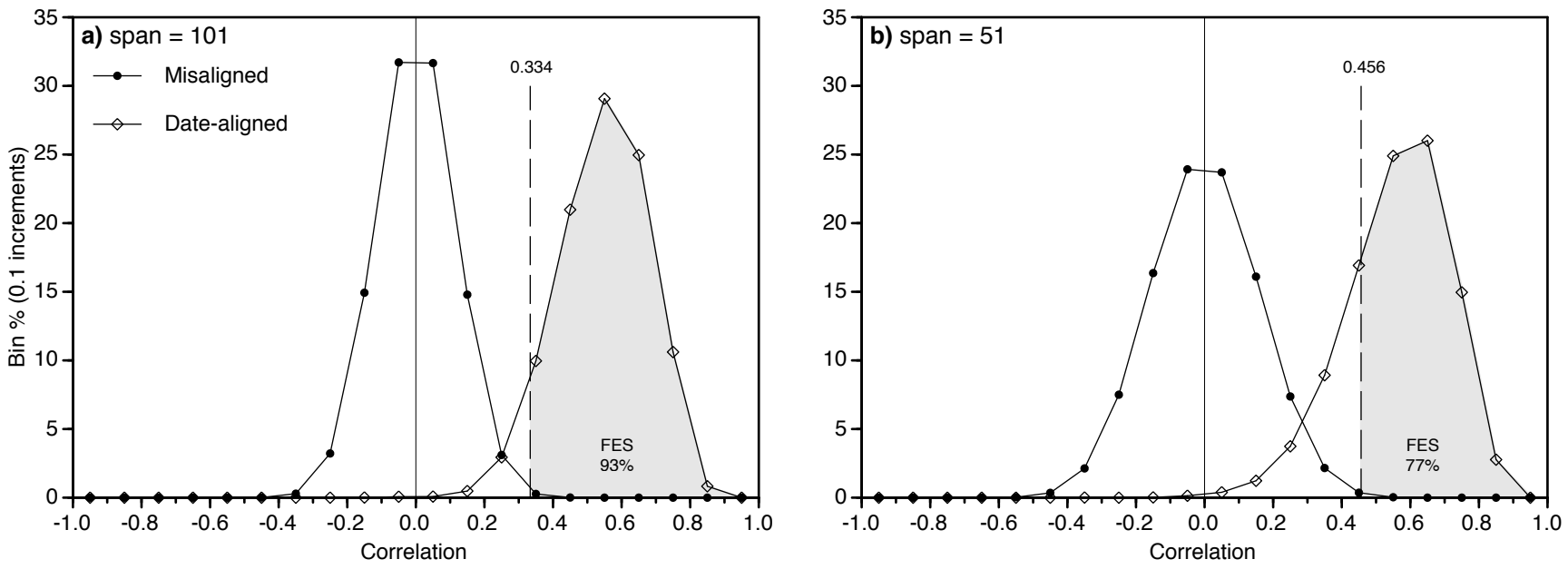

a) BP73, span = 101
Misaligned
Date-aligned
Bin % (0.1 increments)
Correlation
b) Widths, span = 101
Correlation

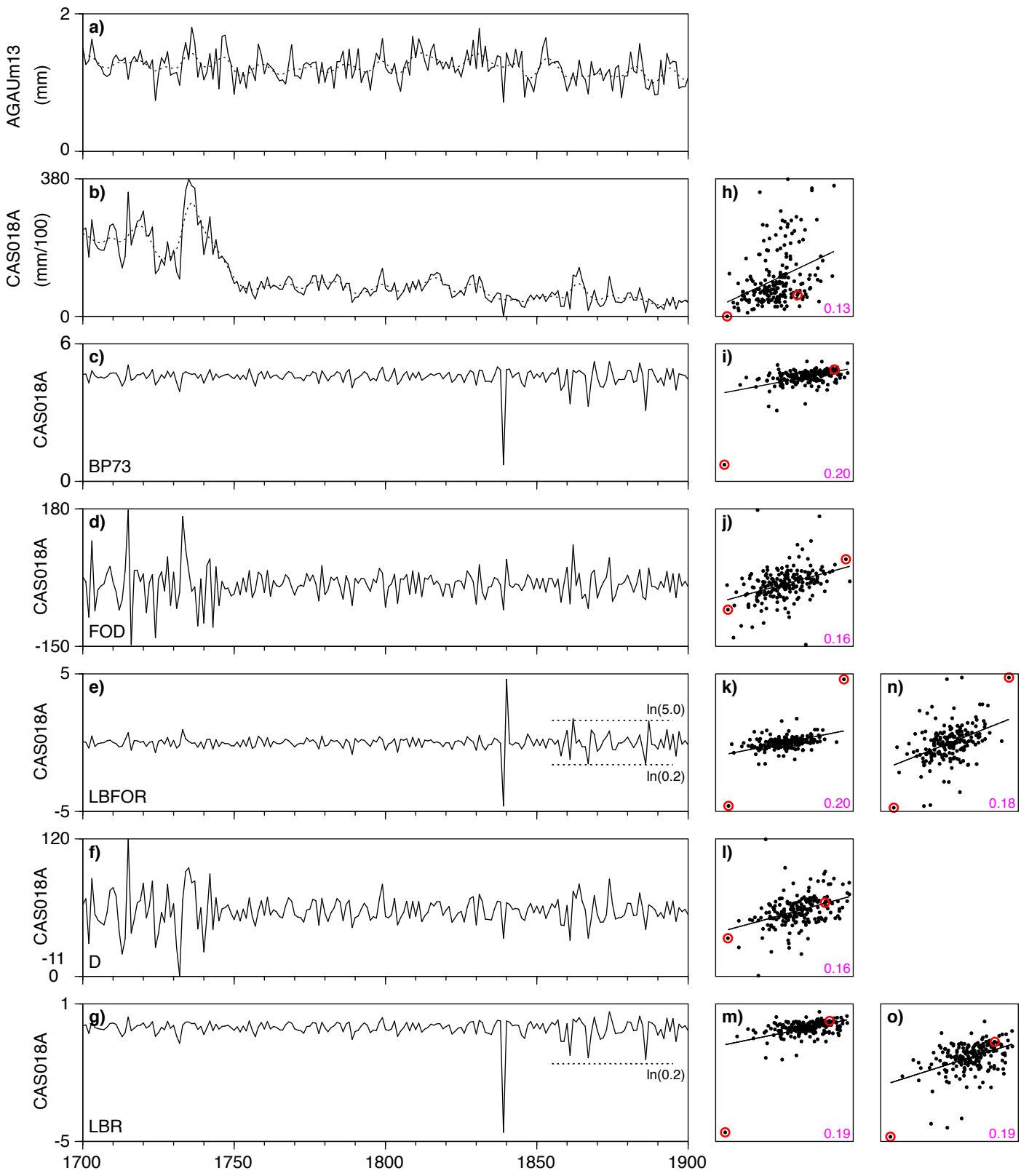

Span = 101
Span = 51
a)
b)
c)
d)
Misaligned
Date-aligned
0.793
0.773
0.351
0.480
Base
FES
<1%
FES
39%
FES
19%
Bin % (0.1 increments)
Correlation
No
Filter
BP73
Filter


Filter Efficiency Score (%)
Window Width (Years)
0
10
20
30
40
1
3
5
7
9
11
13
15
BP73_051
LBR_051
BP73_101
LBR_101
LBFOR_101
FOD_101
D_101
LBFOR_051
FOD_051
D_051


a) span = 101
Misaligned
Date-aligned
0.335
FES
43%
Bin % (0.1 increments)
Correlation
b) span = 51
0.461
FES
21%
Correlation


100
90
80
70
60
50
Filter Efficiency Score (%)
1
3
5
7
9
11
13
15
Window Width (Years)
BP73_051
LBR_051
BP73_101
LBR_101
LBFOR_101
FOD_101
D_101
FOD_051
LBFOR_051
D_051


a) span = 101
Misaligned
Date-aligned
0.334
FES
93%
Bin % (0.1 increments)
Correlation
b) span = 51
0.456
FES
77%
Correlation


a)
AGAUm13
(mm)
2
0
b)
CAS018A
(mm/100)
380
0
c)
CAS018A
BP73
6
0
d)
CAS018A
FOD
180
-150
e)
CAS018A
LBFOR
ln(5.0)
ln(0.2)
5
-5
f)
CAS018A
D
120
-11
0
g)
CAS018A
LBR
ln(0.2)
1
-5
h)
0.13
i)
0.20
j)
0.16
k)
0.20
n)
0.18
l)
0.16
m)
0.19
o)
0.19
1700
1750
1800
1850
1900

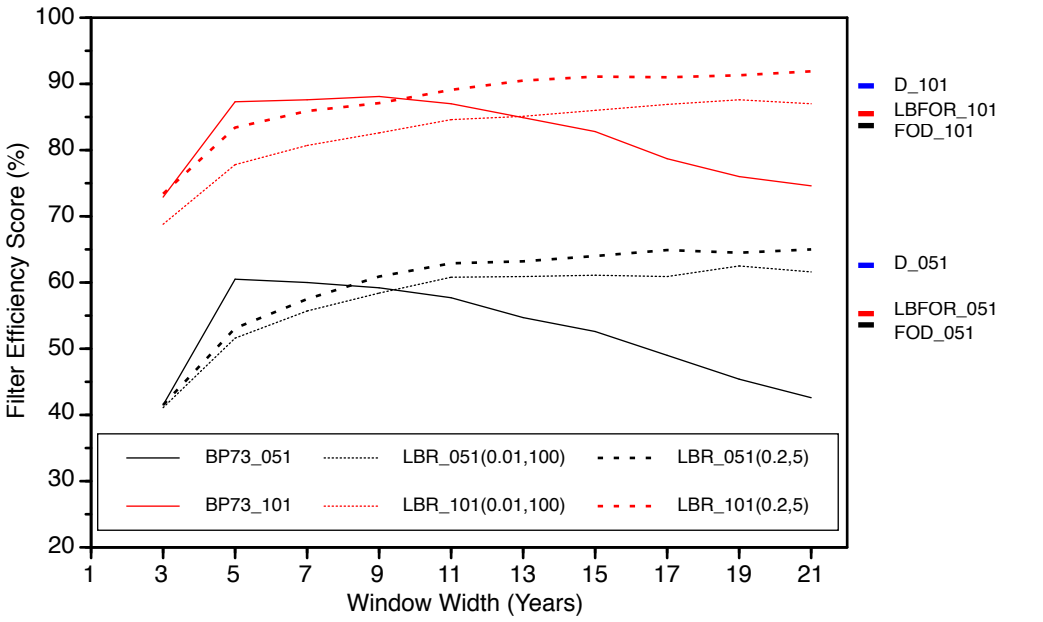

Filter Efficiency Score (%)
100
90
80
70
60
50
40
30
20
1
3
5
7
9
11
13
15
17
19
21
Window Width (Years)
D_101
LBFOR_101
FOD_101
D_051
LBFOR_051
FOD_051
BP73_051
LBR_051(0.01,100)
LBR_051(0.2,5)
BP73_101
LBR_101(0.01,100)
LBR_101(0.2,5)

Table 1. High-pass filters.

Filter	Equations	Details
BP73	$I_i = Log_e\left(\frac{100\,W_i}{RMk_i}\right)$ $\boldsymbol{if}\ (W_i == 0)\ W_i = 1$	Baillie and Pilcher (1973) method. Log_e of ring width (W_i) expressed as a percentage of the mean of the k years each ring is centred on (RMk_i). The running mean window width (k) is fixed at five years in the original BP73, but is adjustable in this implementation. Where RMk_i cannot be calculated, because the running-mean window extends beyond the end of the series, the first or last available running mean value is used. Zero rings are replaced by ones.
FOD	$I_i = W_i - W_{i-1}$	First-order differences. $I_0 = 0$ (no change from unknown prior ring). Arrays are zero-indexed, making I_0 the first element.
LBFOR	$I_i = Log_e\left(\frac{W_i}{W_{i-1}}\right)$ $R_{min} \leq \frac{W_i}{W_{i-1}} \leq R_{max}$	Log_e of bounded first-order ratios. $I_0 = 0$ (no change from unknown prior ring). R_{min} and R_{max} are W_i/W_{i-1} bounds (e.g. 0.01, 10), used to mitigate outliers and to avoid divide-by-zero and $Log_e(0)$ math errors. If the denominator (W_{i-1}) is zero: $I_i = 0$ if the numerator (W_i) is also zero, else $I_i = Log_e(R_{max})$.
LBR	$I_i = Log_e\left(\frac{W_i}{BSCk_i}\right)$ $R_{min} \leq \frac{W_i}{BSC_i} \leq R_{max}$	Log_e of bounded ratios. Ring widths dividing by the corresponding value from a k-weight binomial smoothing curve ($BSCk_i$). Where $BSCk_i$ cannot be calculated, because the window extends beyond the end of the series, the first or last available value for the fitted curve is used. The number of smoothing curve weights (k) is flexible. Ratio bounding is as for LBFOR (above).
D	$I_i = W_i - BSCk_i$	Differences. Ring widths minus the corresponding value from a k-weight binomial smoothing curve ($BSCk_i$). End padding as for LBR (above).