CHRONOLOGY STRIPPING AS A TOOL FOR ENHANCING THE STATISTICAL QUALITY OF TREE-RING CHRONOLOGIES

ANTHONY FOWLER and GRETEL BOSWIJK

School of Geography and Environmental Science
The University of Auckland
Private Bag 92019
Auckland, New Zealand

ABSTRACT

Replication is a key principle in tree-ring research. Dendrochronologists strive to maximise sample size to enhance the "signal" in tree-ring chronologies, often relying on crossdating to provide an effective quality control filter. However, is crossdating alone a sufficient quality test for incorporating a series into a site chronology? We address this question using an objective and automated "chronology stripping" method designed to maximise the chronology's "Expressed Population Signal" (EPS), by iteratively removing series which lower chronology EPS. A 15-site data set of *Agathis australis* (D. Don) Lindley is used to demonstrate the method. Results suggest that modest benefits may be gained by chronology stripping, but the quality control implicit in crossdating is indeed effective, at least for *Agathis australis*.

Keywords: tree-rings, Expressed Population Signal, kauri, *Agathis australis,* New Zealand.

INTRODUCTION

Replication enhances the climate signal in tree-ring chronologies (Fritts 1976). The premise is that if trees have a common response to climate that is obscured by noise, combining multiple series (usually by averaging) will preserve the climate signal and suppress the noise. The larger the sample set, the stronger the expected common signal preserved in the chronology—although this is tempered by knowledge that the marginal return of an additional series declines exponentially as sample size increases (Briffa and Jones 1990).

Where possible, dendrochronologists seek to enhance the quality of tree-ring chronologies by careful sample selection (*e.g.* healthy trees growing in similar environments). The subsequent crossdating process then applies implicit quality control. Material discarded because of crossdating problems is likely to include radii that, for various possible reasons, do not exhibit the prevalent response to climate forcing. Significantly, the data filtering that occurs here is somewhat arbitrary because it depends on the crossdating skill and experience of the researcher, the tools they have at

their disposal, and the "richness" of the data environment in which they are working.

Our own work with *Agathis australis* (D. Don) Lindley (kauri) indicates that experience, the application of computer crossdating tools, and an expanding data base of tree-ring chronologies has significantly improved our crossdating success rate. Many previously rejected samples can now be successfully crossdated. But this presents a conundrum. The importance of replication prompts us to include all crossdated material in deriving tree-ring chronologies, but because this "new" material is of lower quality, we are concerned that including it could actually lower the quality of the chronologies.

The central question of this paper is: does crossdating implicitly impose such a high quality standard that crossdating alone is sufficient justification for including a sample in the construction of a tree-ring chronology? We address this question by developing a "chronology stripping" procedure in which the statistical merits of each sample are assessed and samples that do not improve the statistical quality of the chronology are iteratively removed until we have an "optimum" data set. This
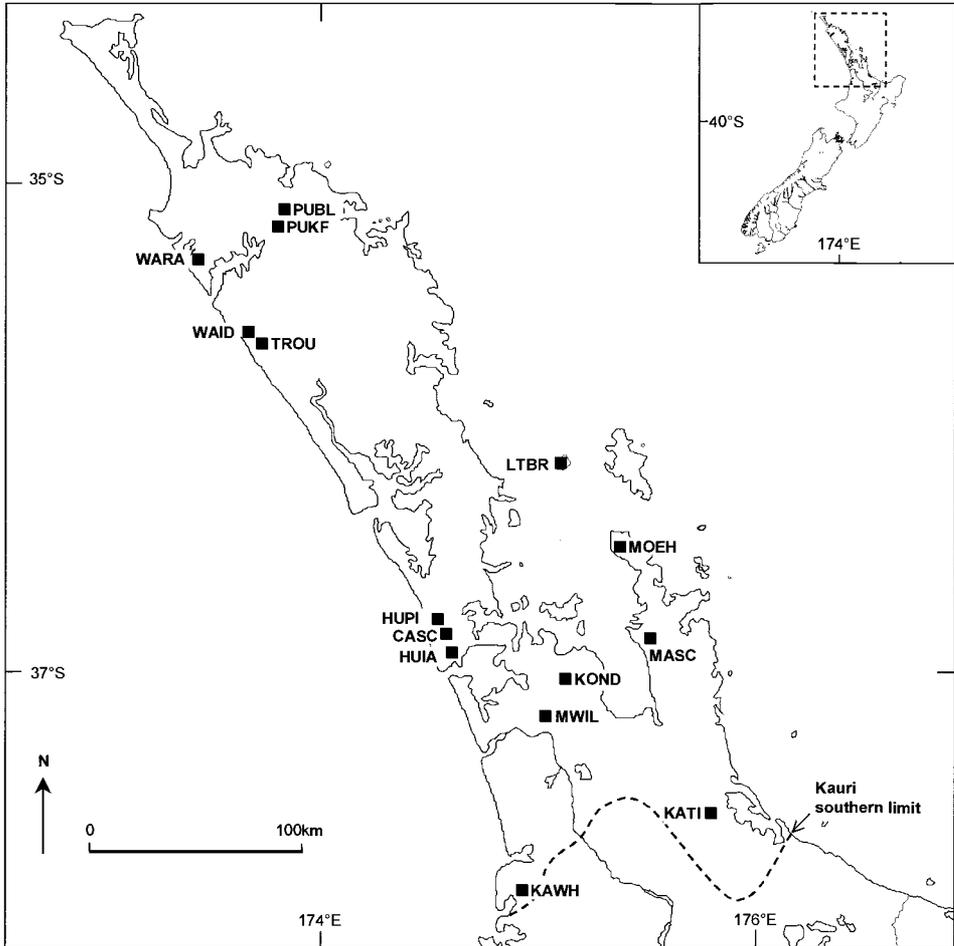
Figure 1. Location of 15 modern kauri chronology sites. Site codes relate to sites as follows: CASC (Cascades Kauri Park); HUIA (Huia); HUPI (Huapai Scientific Reserve); KATI (Katikati); KAWH (Kawhia); KONI (Konini Forks); LTBR (Little Barrier Island); MASC (Manaia Sanctuary); MOEH (Mt. Moehau); MWIL (Mt. William); PUBL (Puketi Bluff); PUKF (Puketi Forest); TROU (Trounson Kauri Park); WAID (Waipaoa Forest dead trees); WARA (Warawara Forest).

approach allows for an objective and automated approach to quality control, although an element of judgement is still required. Other quality control procedures are used by dendrochronologists, including minimum acceptable correlations and minimum series lengths. We acknowledge these here but do not consider them in this paper.

DATA

The utility of the chronology stripping methodology is explored using data from 15 modern kauri tree-ring sites (Figure 1). These data were developed by several workers over the last few decades (Dunwiddie 1979; La Marche *et al.* 1979; Ahmed and Ogden 1985; Palmer 1982; Fowler 1984; Buckley *et al.* 2000; Boswijk *et al.* 2000; Fowler and Boswijk 2000; Boswijk 2001; Fowler and Boswijk 2001; Boswijk and Ogden 2001; Boswijk and Palmer 2001; Fowler *et al.* in press).

Kauri grow in closed canopy forests and are subject to competitive pressures throughout their lives. They are prone to wind damage (often losing major branches) and frequently exhibit growth-form anomalies. These competitive/injury/growth characteristics have clear expression in complex

ring-width time series. Decline in ring widths with age is typical, but is neither ubiquitous nor monotonic. Century-scale periods of growth suppression or release can occur at any time, probably related to evolving competitive pressures. Decadal-scale variation is normally superimposed on these century-scale trends. Decadal-scale departures are usually radii-specific; indicative of competition, injury, or growth idiosyncrasies, rather than variable climate forcing. The climate signal is primarily expressed in high-frequency year-on-year variations. Wide rings are most commonly associated with cool and dry growing seasons (characteristic of El Niño events in kauri's growth region), whereas narrow rings are associated with warm and wet (La Niña) conditions (Buckley *et al.* 2000; Fowler *et al.* 2000).

In preparation for chronology stripping, the AR-STAN program developed by Cook (Cook 1985; Holmes *et al.* 1986) was used to standardise all series. Two sets of tree-ring indices were derived using splines with 50% frequency response at 20 years and 200 years. The "flexible" 20-year splines maximise intra-site correlations (compared with intra-site correlations for 10, 15, 20, 50, 100, and 200 year splines). The resulting tree-ring chronologies are suitable for high-frequency climate applications, such as sunspots and high-frequency elements of the El Niño–Southern Oscillation phenomenon. They are also probably ideal for dating purposes (crossdating and archaeological). Chronologies derived using the more conservative 200-year splines are more typical of those used for climate trend analysis, such as global warming applications. Two versions of the data set were constructed to examine the sensitivity of the chronology stripping results to the "conservativeness" of the standardisation applied.

CHRONOLOGY STRIPPING

Chronology quality was assessed using the Expressed Population Signal (EPS) statistic. This was derived from mathematical first principles for chronologies computed from one series per tree by Wigley *et al.* (1984) and extended to a more general form of variable cores per tree by Briffa and Jones (1990). The latter form was adopted here.

EPS is derived from the cross-correlation matrix of all series (see Appendix for details) and has a possible range from zero to one, where one represents a hypothetical perfect chronology.

EPS increases with sample size and with the strength of the mean correlation between series from different trees (Briffa and Jones 1990). Thus, for an arbitrary EPS threshold, fewer trees are required to meet the threshold as the mean inter-tree correlation increases. Briffa and Jones (1990) reported a range of 4–25 trees required to meet an EPS threshold of 0.85, the lower limit being for semi-arid conifers in the western USA and the upper limit for deciduous sites in the UK. Kauri is at the lower end of this range (Figure 2A). Using 20-year spline standardisation, we would expect to need about six trees to reach the 0.85 threshold, rising to about 11 trees for the more conservative 200-year spine standardisation.

The EPS-based analyses were undertaken in two parts. First, a chronology stripping exercise identified "poor" quality cores, defined as cores that lowered the chronology EPS when their standardised time series were included. Second, analysis of the quality of the chronology through time was undertaken to assess the impact of declining sample depth. The chronology stripping analysis was undertaken separately for each site as follows:

a) All series standardised using splines with 50% frequency response at 20 and 200 years.

b) EPS of the chronology calculated with all series present.

c) Iterated through all series, calculating EPS as each series is removed (and replaced) in turn, flagging the series yielding the highest EPS (when removed).

d) Discarded the series (if any) that, when removed, yielded the highest EPS.

e) Repeated steps c) and d) until no further stripping out of series increased EPS.

f) Repeated steps c), d), and e) but checking if replacing previously removed series increased EPS.

g) Iterated through each year of the chronology calculating the stripped and unstripped EPS for all series present at that year.
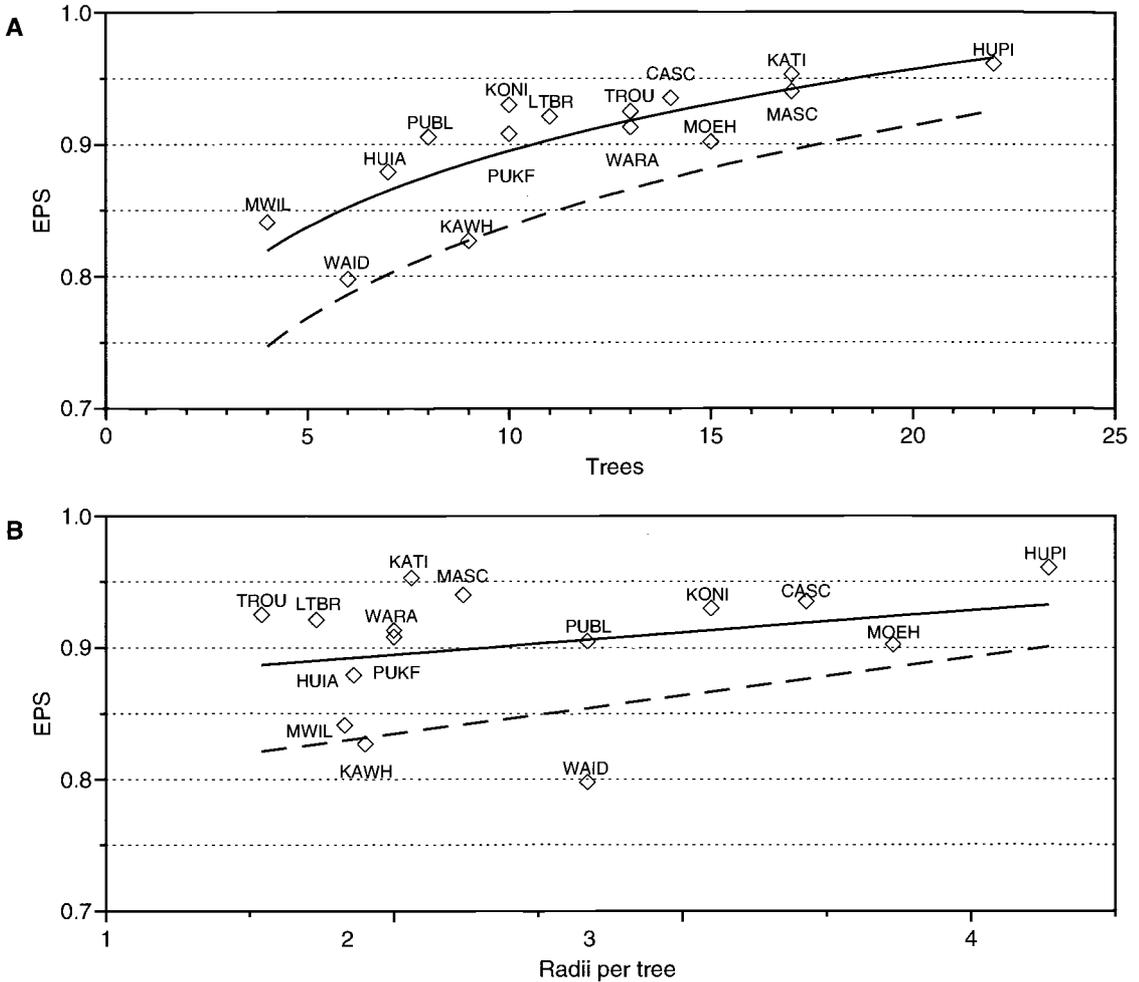
h) Subjectively assessed the merits of the stripping.

Figure 2. EPS statistics prior to chronology stripping in relation to the number of trees in the site chronology (A) and the mean number of radii per tree (B). The plotted points and the solid regression lines are for chronologies built from data standardised using 20-year splines. The dashed regression lines relate to the 200-year splines (points not plotted). Regression lines are of power form ($y = a\,x^z$) in Panel A and linear in Panel B.

The replacement check (f) ensured that no stripped series later proved useful when other series had been stripped. No such replacements were required for any site. The year-by-year iteration, comparing stripped and unstripped EPS for each year (g), was undertaken to assess evolving statistical quality of the series. It was also the basis for subjective analysis of the merits of stripping (h). The rationale for this last step was that although a sub-sample of the data may yield a higher overall EPS, it may do so by stripping series that are critical for sub-periods when sample

depth is so low that loss of material is considered unacceptable.

The EPS statistic is calculated from the sample size and the mean within- and between-tree correlations (see Appendix). The implementation here involved calculating the correlation matrix for all series at a site for the full overlap of each series pair, provided that there was a minimum of 30 year overlap. As each series was tested (by removal), all component parts of the EPS equation were recalculated. Similarly, for the analysis of evolving quality through time, EPS was recalculated each

year for only those series that overlapped with the year in question.

RESULTS AND DISCUSSION

A context for the chronology stripping results is provided by the pre-stripping EPS statistics for the 15 sites (Figure 2). EPS is dependent on the number of trees, the standardisation method used, and the number of radii per tree. It also indicates the scale of inter-site variability of EPS. Sample depth is clearly the most important consideration. Variation in the number of trees is responsible for EPS variation of the order of 0.15, whereas increasing the mean number of radii per tree by one can increase EPS by up to 0.03. Inter-site variability (*i.e.* for sites of similar sample size) is of the order of 0.05–0.10, and is most pronounced with lower sample size (<10). The effect of the standardisation method (vertical difference between the two curves) is about 0.05.

Chronology stripping results for the HUPI chronology (Figure 3) were derived from the data set using 20-year splines run on 94 series from the 22 largest crossdated trees (DBH ≥ 0.72 m) with 2–7 cores per tree. Pre-stripping EPS was 0.961, the highest recorded at any site. Chronology stripping removed 26 series (1–3 from 15 trees), but no tree was reduced to less than two series. Series sample depth was reduced by about one third throughout the data range, but the number of trees by a maximum of one (Figure 3B), because of stripping of all cores from a tree for some time periods. Stripping increased EPS to 0.963.

Change in sample depth (Figure 3A), particularly the changing number of trees, dominates the evolution of EPS through time (Figure 3B). The impact of stripping on EPS is practically indistinguishable after AD 1620, despite the 27% paring of the data set. Prior to AD 1620, stripping maintained EPS above 0.75 and substantially above the values obtained for the non-stripped data. Close inspection of the few series covering this early period suggests that the effectiveness of stripping is a consequence of poor agreement at this time between the three trees extending back to the mid-1500s. This is clearly demonstrated by the differ-

ences in the stripped and unstripped HUPI chronologies for A.D. 1530–1560 (Figure 3C).

Stripping results for HUPI in the case of series standardised using 200-year splines are broadly similar (Figure 4). A total of 27 series from 17 trees were stripped, raising EPS from 0.942 to 0.948. About half (13) of the stripped series are common to both analyses and the first six series removed in each analysis were also removed at some stage in the other. A significant difference in the 200-year spline analysis is the absence (Figure 4) of the early period EPS divergence evident in Figure 3B. In fact, stripping significantly reduced EPS from 1566–1616, even though overall chronology EPS was raised. In view of low sample depth during 1566–1616, we would subjectively "reject" the stripping results in favour of preserving sample depth at this time.

Stripping results for the 15 sites are summarised in Table 1. Of the 30 analyses undertaken, stripping increased chronology EPS in all but one case (TROU, 20-year splines). With three exceptions (HUPI, MOEH, CASC) few series were stripped and, in all cases, very few trees were completely removed. In six cases we would subjectively choose not to apply the stripping results because of undesirable effects on EPS for periods of low sample depth.

Although stripping raised EPS in most cases, the increase was almost universally very small. The mean EPS increase for the 200-year splines (0.017) is more than twice as large as that for the 20-year splines (0.007), reflecting the sensitivity of EPS to the standardisation used, but is low compared to variations in EPS related to sample depth, standardisation applied, and inter-annual variation (Figure 2). The mean EPS improvement is roughly equivalent to adding an additional 1–2 trees to the chronology (20- and 200-year splines, respectively). EPS increases for MOEH (0.045) and MWIL (0.053), both for 200-year splines, are noticeably larger.

CASC, HUPI, and MOEH are noteworthy in being stripped of relatively large numbers of series. In the case of MOEH, this probably reflects the complicated composition of the data (Boswijk and Palmer 2001) and the fact that it started off with a relatively large number of series, especially rel-
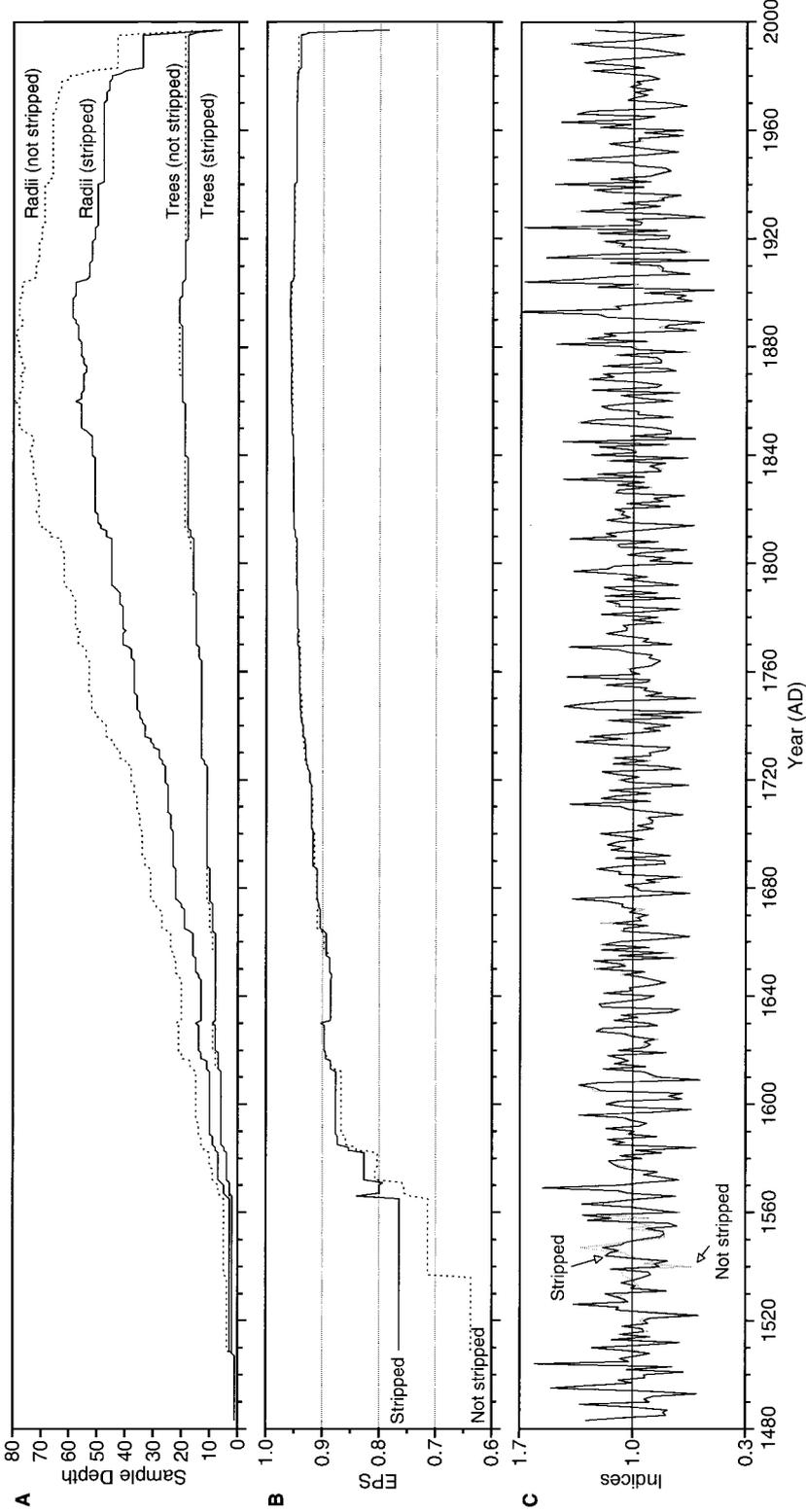
Figure 3. Chronology stripping results for HUPI. Data standardised using 20-year splines. Dashed lines are prior to stripping, solid lines after. Panel A shows sample depth (trees, radii) through time, Panel B the evolution of EPS through time, and Panel C the resulting chronology indices. Indices were computed by averaging radii indices into tree indices, then averaging tree indices.
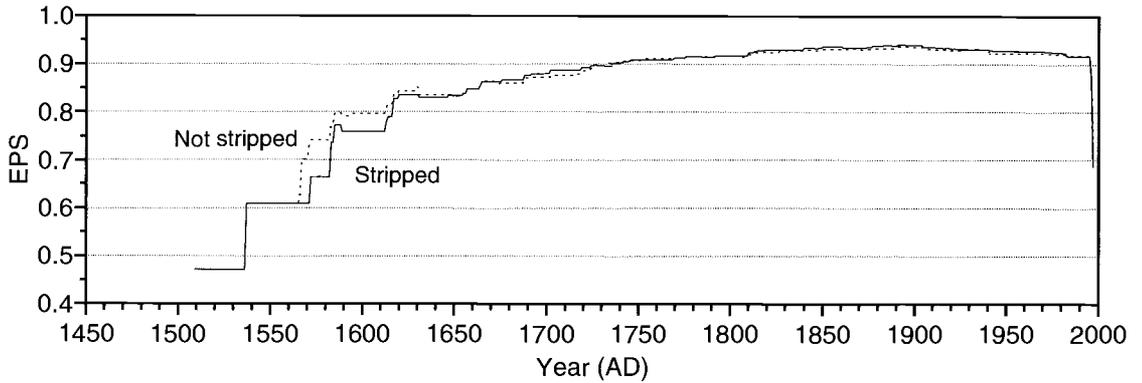
Figure 4. Chronology stripping results for HUPI. Data standardised using 200-year splines.

Table 1. Chronology stripping results. The upper entry for each site gives the results from stripping data standardised using 20-year splines. The lower entry is for the 200-year splines. Rbar$_{bt}$ and Rbar$_{wt}$ are pre-stripping mean between- and within-tree correlations. "Reject" means that stripping is subjectively judged to be inappropriate because of detrimental impacts on one or more periods of low sample depth (even though chronology-wide EPS increases).

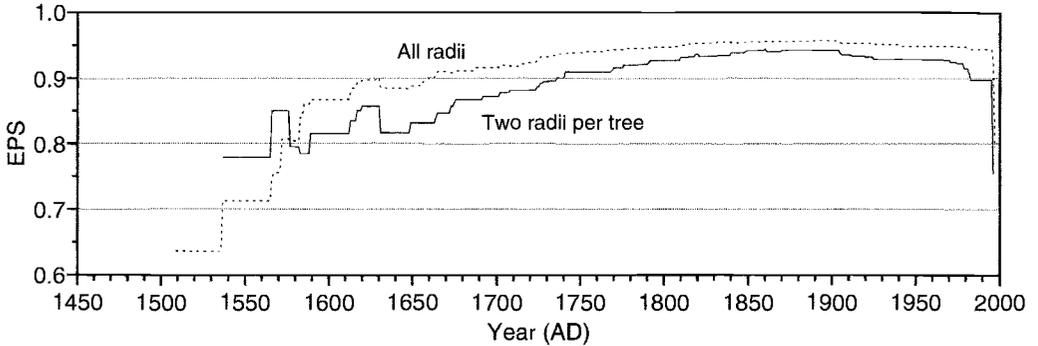| Site	Pre-stripping Rbar$_{bt}$/Rbar$_{wt}$	Pre-stripping series/trees	Stripped series/trees	EPS Original → Stripped	EPS Change	Reject
CASC	0.363/0.570	48/14	10/0	0.935 → 0.941	0.006	X
	0.244/0.383		13/0	0.908 → 0.918	0.010	X
HUIA	0.426/0.622	13/7	1/0	0.879 → 0.880	0.001	
	0.315/0.535		2/1	0.819 → 0.836	0.017	
HUPI	0.391/0.651	94/22	26/0	0.961 → 0.963	0.002	
	0.264/0.484		27/0	0.942 → 0.948	0.006	X
KATI	0.493/0.589	35/17	4/1	0.953 → 0.954	0.001	X
	0.239/0.368		5/1	0.897 → 0.900	0.003	
KAWH	0.250/0.458	19/10	5/1	0.827 → 0.845	0.018	
	0.147/0.249		2/1	0.731 → 0.750	0.019	X
KONI	0.443/0.667	31/10	6/0	0.930 → 0.937	0.007	
	0.286/0.549		6/0	0.875 → 0.893	0.018	
LTBR	0.440/0.603	19/11	1/0	0.921 → 0.922	0.001	
	0.337/0.521		1/0	0.883 → 0.884	0.001	
MASC	0.373/0.562	38/17	4/0	0.940 → 0.941	0.001	
	0.214/0.413		6/1	0.881 → 0.899	0.018	
MOEH	0.293/0.653	56/15	23/1	0.902 → 0.923	0.021	
	0.178/0.507		18/2	0.842 → 0.890	0.045	
MWIL	0.396/0.570	11/6	3/2	0.841 → 0.854	0.013	
	0.287/0.363		4/2	0.781 → 0.834	0.053	
PUBL	0.408/0.641	24/9	5/1	0.905 → 0.914	0.009	
	0.322/0.518		6/1	0.878 → 0.891	0.013	
PUKF	0.414/0.586	20/10	1/0	0.908 → 0.909	0.001	
	0.282/0.427		4/1	0.852 → 0.860	0.008	X
TROU	0.432/0.583	20/13	0/0	0.925 → 0.925	0	
	0.250/0.337		6/2	0.850 → 0.858	0.008	
WAID	0.293/0.550	16/6	3/0	0.798 → 0.813	0.015	
	0.172/0.293		3/0	0.712 → 0.738	0.026	
WARA	0.376/0.676	28/14	4/1	0.913 → 0.915	0.002	
	0.265/0.487		5/1	0.872 → 0.879	0.007	

Figure 5. Comparison of HUPI time-evolving EPS statistics for all radii and for a reduced data set of the first two radii per tree. Data standardised using 20-year splines. The two radii per tree line does not extend as far back in time because some of the oldest series were not present in the 44 series subset used.

ative to the number of trees. Similarly, HUPI had a relatively large data set that must have included considerable redundancy in its original state, given the negligible impact on the site chronology of removing 26 series. CASC has the third largest sample size.

The data redundancy evident in the HUPI, MOEH, and CASC data sets suggests that chronology stripping may provide useful information to guide future sampling. For example, the fact that HUPI was stripped of a large proportion of its series, but no tree was reduced to less than two radii may indicate redundancy in taking more than two cores per tree. Figure 2B suggests that this may not be the case, but to test it further the HUPI data set was reduced to its first two radii per tree (44 radii) and evolving EPS compared to that for all 94 radii (Figure 5). The resulting significant decrease in EPS indicates that the redundancy hypothesis is not valid and there is significant benefit for kauri in obtaining more than two radii. It follows that the 26/27 radii stripped from HUPI were stripped for quality reasons rather than data redundancy issues.

CONCLUSIONS

Chronology stripping can increase the statistical quality of tree-ring chronologies. The fact that stripping raised EPS in 29 of 30 analyses indicates that crossdating does not guarantee that a series will be a useful addition to a chronology. However, since in most cases, very few series were

stripped, it is also clear that crossdating does indeed provide quite effective implicit quality control. This conclusion is reinforced by the relatively modest improvements in EPS as a result of stripping.

One caveat we attach to this benign conclusion is that significant quality issues may be identified by the stripping process, even when chronology-wide EPS changes are minor. This is apparent in Figure 3B, where stripping increased EPS by 0.044–0.127 prior to about A.D. 1570, even though chronology-wide EPS improvement is very small (0.002).

The stripping method used here assesses the quality of a series across its entire length (because correlations are calculated across the full overlap of each series pair). As a consequence, a series may be stripped because of relatively weak overall correlation, even when it has a relatively high correlation for a part of that time. Stripping such a series may be undesirable if the period of higher correlation is for a time period of low sample depth. The converse situation is also plausible, where a series is not stripped, but is in fact detrimental to a period of low sample depth. An alternative methodology to that presented here would be required to resolve such issues of time-evolving quality.

Not withstanding the above reservations, we conclude that chronology stripping is a useful addition to quality control, particularly in identifying series that should be reviewed before proceeding with chronology building. Any benefits are likely

to be modest and we present no evidence that challenges the veracity of the quality control implicit in the crossdating process. However, because our analysis is specific to kauri, our conclusion may be too.

ACKNOWLEDGMENTS

REFERENCES CITED

Ahmed, M., and J. Ogden
   1985   Modern New Zealand tree-ring chronologies 3. *Agathis australis* (Salib.)—kauri. *Tree-Ring Bulletin* 45:11–24.
Boswijk, G.
   2001   *Tree-ring Analysis of Kauri (*Agathis australis*) from Huia, Waitakere Ranges, Auckland New Zealand Tree-Ring Site Report No. 5.* University of Auckland, Department of Geography Working Paper No. 14.
Boswijk, G., A. Fowler, and J. Ogden
   2000   *Tree-ring Analysis of Kauri (*Agathis australis*) from Manaia Sanctuary, Coromandel Peninsula. New Zealand Tree-Ring Site Report No. 1.* University of Auckland, Department of Geography Working Paper No. 9.
Boswijk, G., and J. Ogden
   2001   *Tree-ring Analysis of Kauri (*Agathis australis*) from Mt Hobson (Hirakimata), Great Barrier Island. New Zealand Tree-Ring Site Report No. 4.* University of Auckland, Department of Geography Working Paper No. 12.
Boswijk, G., and J. Palmer
   2001   *Tree-ring Analysis of Kauri (*Agathis australis*) from Mt Moehau, Coromandel Peninsula. New Zealand Tree-Ring Site Report No. 3.* University of Auckland, Department of Geography Working Paper No. 11.
Briffa, K., and P. Jones
   1990   Basic chronology statistics and assessment. In *Methods of Dendrochronology: Applications in the Environmental Sciences,* edited by E. Cook, and L. Kairiukstis, Kluwer Academic, Dordrecht; pp. 137–152.
Buckley, B., J. Ogden, J. Palmer, J. Salinger, and A. Fowler
   2000   Dendroclimatic interpretation of tree-rings in *Agathis australis* (kauri) 1. Climate correlation functions and master chronology. *Journal of the Royal Society of New Zealand* 30:263–275.

Cook, E. R.
   1985   *A Time Series Analysis Approach to Tree-ring Standardization.* Unpublished Ph.D. dissertation, University of Arizona, Tucson, USA.
Fowler, A.
   1984   A *Dendroclimatological Study of Kauri (*Agathis australis*).* Unpublished M.A. thesis, University of Auckland, Auckland, New Zealand.
Fowler, A., and G. Boswijk
   2000   *Tree-ring Analysis of Kauri* (Agathis australis) *from Cascades Kauri Park, Waitakere Ranges, Auckland. New Zealand Tree-Ring Site Report No. 2.* University of Auckland, Department of Geography Working Paper No. 10.
Fowler, A., and G. Boswijk
   2001   *Tree-ring Analysis of Kauri (*Agathis australis*) from Huapai Scientific Reserve, Huapai, Auckland. New Zealand Tree-Ring Site Report No. 6.* University of Auckland, Department of Geography Working Paper No. 13.
Fowler, A., G. Boswijk, and J. Ogden
   in   Tree-Ring studies on *Agathis australis* (kauri): A
   press   synthesis of development work on Late Holocene chronologies. *Tree-Ring Research.*
Fowler, A., J. Palmer, J. Salinger, and J. Ogden
   2000   Dendroclimatic interpretation of tree-rings in *Agathis australis* (kauri) 2. Evidence of a significant relationship with ENSO. *Journal of the Royal Society of New Zealand* 30:277–292.
Fritts, H.C.
   1976   *Tree Rings and Climate.* Academic Press, London.
Holmes, R. L., R. K. Adams, and H. C. Fritts
   1986   Users manual for Program Arstan. In *Tree-ring Chronologies of Western North America: California, Eastern Oregon and Northern Great Basin.* The University of Arizona; pp. 50–60.
La Marche, V. C., R. L. Holmes, P. W. Dunwiddie, and L. G. Drew
   1979   *Tree-ring Chronologies of the Southern Hemisphere: 3. New Zealand.* Laboratory of Tree-Ring Research, University of Arizona, Tucson.
Palmer, J.
   1982   A *Dendrochronological Study of Kauri (*Agathis australis*).* Unpublished M.S. thesis, University of Auckland, Auckland, New Zealand.
Wigley, T. M. L., K. R. Briffa, and P. D. Jones
   1984   On the average value of correlated time series, with applications in dendroclimatology and hydrometeorology. *Journal of Climate and Applied Meteorology* 23:201–213.

APPENDIX

Expressed Population Signal (EPS) Calculation

The Expressed Population Signal (EPS) statistic used in this paper is that described in detail in Briffa and Jones (1990). The

relevant equations are provided below. See Wigley *et al.* (1984) for the derivation from first mathematical principles of the equation for the case of one series per tree and Briffa and Jones (1990) for the expansion to a more general form of variable cores per tree.

For the case of one series per tree, Wigley *et al.* (1984) derived EPS as:

$$EPS = \frac{t\bar{r}_{bt}}{t\bar{r}_{bt} + (1 - \bar{r}_{bt})} \quad \text{(Briffa and Jones Equation 3.44)}$$

$$(A1)$$

where is the number of trees and $\bar{r}_{bt}$ is the mean between-tree correlation.

According to Briffa and Jones (1990, p. 146) "a more general expression for EPS, where a chronology has more than one core per tree, can be obtained by replacing $\bar{r}_{bt}$ with $\bar{r}_{eff}$ in Equation 3.44 and using the number of trees sampled as t". This gives:

$$EPS = \frac{t\bar{r}_{eff}}{t\bar{r}_{eff} + (1 - \bar{r}_{eff})} \quad (A2)$$

where $\bar{r}_{eff}$ (the effective chronology signal) is a chronology-signal estimate incorporating both within- and between-tree signals (Briffa and Jones 1990, p. 142). It is calculated from mean within- and between-tree correlations as follows:

$$\bar{r}_{eff} = \frac{\bar{r}_{bt}}{\bar{r}_{wt} + \dfrac{1 - \bar{r}_{wt}}{c_{eff}}} \quad \text{(Briffa and Jones Equation 3.43)} \quad (A3)$$

where $\bar{r}_{bt}$ is the mean correlation of series between trees, $\bar{r}_{wt}$ is the mean correlation of series within trees, and c_{eff} is the effective number of cores per tree. The latter is derived from:

$$\frac{1}{c_{eff}} = \frac{1}{t} \sum_{i=1}^{t} \frac{1}{c_i} \quad \text{(Briffa and Jones Equation 3.42)} \quad (A4)$$