TREE-RING BULLETIN, *Vol. 41, 1981*

TEST OF A NEW METHOD FOR REMOVING THE GROWTH TREND FROM DENDROCHRONOLOGICAL DATA

W. G. WARREN

and

S. L. MACWILLIAM

Forintek Canada Corp.
Vancouver, B.C.

ABSTRACT

Tests of the compound increment function, introduced by Warren (1980) as a means for removing the growth trend from dendrochronological data, are herein reported. In particular, the inter- and intra-site correlations of the residuals generated by the new method are compared with those generated by standard exponential fits. It is also shown that, in the presence of non-climatically induced responses, such as might arise from thinning, exponential fits can lead to spuriously high intra-site correlations. Accordingly, and because the new method's virtual elimination of negative and very low positive correlations, it appears to be the more satisfactory for portraying the growth trend.

Es wird über Tests der von Warren (1980) eingeführten zusammengesetzten Zuwachsfunktion als Mittel zur Eliminierung des Wachstumstrendes aus dendrochronologischen Daten berichtet. Insbesondere werden die zwischen-und innerstandörtlichen Korrelationen der Restvarianzen, wie sie mit dem neuen Verfahren und mit dem herkömmlichen Exponentialausgleich entstehen, miteinander verglichen. Ferner wird gezeigt, daß der Exponentialausgleich bei nicht-klimatisch verursachten Reaktionen der Bäume, wie z. B. als Folge von Durchforstungen, zu hohen unechten innerstandörtlichen Korrelationen führen kann. Aus diesem Grunde sowie wegen seiner Fähigkeit zur Beseitigung negativer und sehr schwach positiver Korrelationen erscheint das neue Verfahren geeigneter für die Wiedergabe des Wachstumstrends.

Des tests de la fonction dite "compound increment function" introduite par WARREN (1980) comme un moyen pour éliminer la tendance de croissance présente dans les données dendrochronologiques sont décrits ci-après. En particulier, les corrélations inter- et intra- site des résidus générés par la nouvelle méthode sont comparées avec celles fournies par les lissages exponentiels classiques. Il est également montré qu'en présence de réactions non induites par le climat, comme cela peut être le cas par éclaircie, les lissages exponentiels peuvent conduire à de hautes corrélations intra-site qui sont erronées. En conséquence, et parce que cette nouvelle méthode élimine les corrélations négatives ainsi que les positions très faibles, elle apparaît comme la plus satisfaisante pour représenter la tendance de croissance.

INTRODUCTION

A new approach to removing the growth trend from dendrochronological ring-width data has recently been described by the senior author (Warren 1980). In brief, it assumes that the ring width for the jth year of a record can be represented by:

$$y_j = \sum_{i=0}^{n} \delta_{ij} a_i (x_j - t_i)^{b_i} \exp[-c_i(x_j - t_i)], \quad \delta_{ij} = \begin{cases} 0, & x_j \leq t_i \\ 1, & x_j > t_i \end{cases} \tag{1}$$

and can be described as a compounding of supposed increment functions. The t_i denote the times, relative to the start of the record, that a "pulse" is added to the trend; $t_0 \leq 0$ thus refers to the age of the tree relative to the start of the record.

Warren (1980) also described a method for estimating the number of parameters in the model and their values. A time base, N, is selected and the basic increment function transformed to:

$$\ell n(y) = \tilde{a} + b\ell n\,(x - t) - c\,(x - t)$$

is fitted to the first N years of the record. The fit is extrapolated to the next N years. If the data points do not fall significantly above the extrapolation, these N points are added to the previous set and the function fitted over the enlarged range. If the data points are judged to be significantly above the extrapolation, a growth curve, of the same form, is fitted to the residuals and added to the original, as indicated in (1). The possibly compound curve is then extrapolated to the next N years and the process repeated.

Certain difficulties have to be overcome; full details are given in the aforementioned paper. It also is stated therein that, of the fitting-system parameters, as opposed to the model parameters, N is believed to be the most critical. Application of the method to a total of 10 increment cores from trees in three sites belonging to a single watershed led to the selection of N = 30 as the most suitable compromise, at least for this specific material. We now report the results of applying the method, with N = 30, to all 73 cores obtained from these sites.

It was previously remarked (Warren 1980) that, heuristically, the stronger the cross correlation of the residuals, or indices, between cores, the more likely that the residuals are associated with common environmental changes, in particular climatic fluctuations. We define the residual as the observed minus the calculated ring width, i.e. $y_{obs} - y_{cal}$. The index value is taken as y_{obs}/y_{cal}. The latter seems preferable since its variance will tend to be the more stable. In what follows, index values are used unless otherwise specified. Thus, the stronger the correlation, the more likely one has been successful in removing the growth trend. In the course of our investigation, we found cause to question this conjecture and have determined that there are circumstances where it is not necessarily true, and where spuriously high correlations can arise. One section of this report is, therefore, devoted to this aspect.

DATA

The three sites will be designated PR, CC and VB. Of the 73 available cores, 25 are from each of PR and CC and the remaining 23 from VB.

The growth trend was estimated by the new method, with N = 30, and by the exponential fit option of the computer program developed by Fritts et al. 1969. The index values stemming from each fitting method were calculated and the cross correlation matrices computed for each site. The cores terminated in 1976 and ranged in length from 107 to 257 years for PR (with all but one $\leq$ 187 years), from 107 to 227 years for CC and from 67 to 167 years for VB. Accordingly the correlation matrices are based on 107 years for PR and CC and 67 years for VB. For CC, one core was found to have effectively zero correlation with all other cores under either fitting method. It is consequently believed that this core was incorrectly dated and it was, therefore, excluded from all further calculations. There remain 72 cores, of which 24 are from CC. There are then 24x25/2 = 300 correlation coefficients for PR, 276 for CC and 253 for VB.

RESULTS

It must be remembered that, for any site, the computed correlations are not mutually independent. Certain inferences can, however, be drawn. The averages and standard deviations of the computed correlations are:

	PR	CC	VB
Exp. fit	0.519 (0.170)	0.481 (0.171)	0.410 (0.195)
New fit	0.480 (0.133)	0.444 (0.112)	0.450 (0.153)

(standard deviations in parentheses)

The averages for the two fitting methods are quite close, differing by less than 10% for each site. The standard deviations derived from the new fitting method are, however, noticeably and consistently less, by about 25%, than those based on the exponential fits.

The histograms of these statistics are typified by Figure 1 for CC. Those derived from the exponential fits have a somewhat longer tail to the left than those that

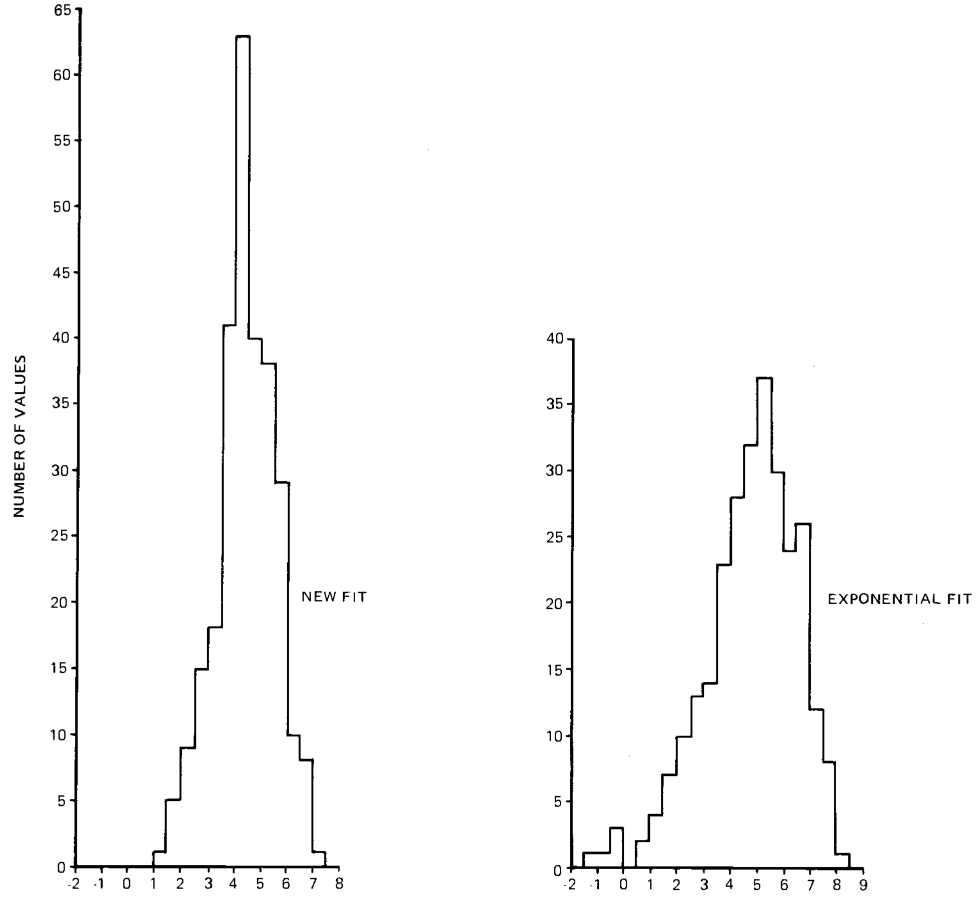

Figure 1. CC correlation coefficients for the exponential and new fits.

correspond to the new fits; in other words, a somewhat greater proportion of negative and very low positive correlations have been obtained via the exponential fits. Indeed there is no instance of a calculated negative correlation being associated with the new fits.

For formal significance at the 5% level a sample correlation based on 67 and 107 pairs of observations must exceed 0.204 and 0.156, respectively. We use a one-sided test since the meaningful alternative to the null hypothesis is that of positive correlation. The numbers of formally non-significant correlations are then:

	PR	CC	VB
Exp. fit	7/300 (11)	11/276 (16)	37/253 (46)
New fit	3/300 (6)	1/276 (4)	18/253 (29)

($2^1/_2$%-level frequencies are given in parentheses)

It can be seen that the number of non-significant correlations has been approximately halved under the new fitting method relative to the exponential fits.

On the other hand, the exponential fits also show a somewhat greater proportion

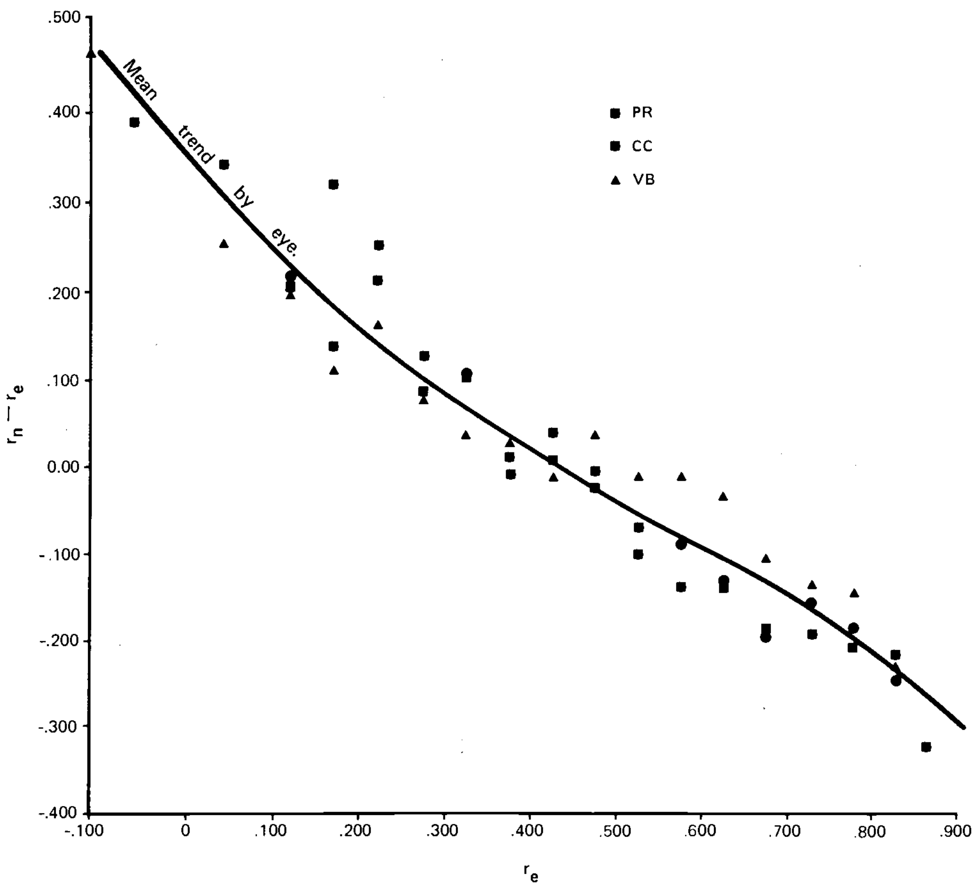

Figure 2. Average differences between the exponential- and new-fit correlations for each site, by values of the exponential-fit correlations.

of the higher correlation coefficients. This disappointing feature merits closer examination.

Instead of looking at the overall picture, let us consider the individual correlation comparisons, i.e. the difference between the exponential- and new-fit correlations for the ith and jth cores for all possible pairs (i, j) within a given site. Let r_e and r_n denote the exponential- and new-fit correlations, respectively. Figure 2 shows the average values of $r_n - r_e$ grouped by values of r_e. Note that $r_n - r_e > 0$ means that the correlation obtained via the new fit is greater than that obtained via the exponential fit.

The pattern is essentially the same for the three sites. On the average, when r_e is 0, r_n is about 0.35; when r_e is c. 0.45, so is r_n and when r_e is c. 0.8, r_n is about 0.6. More specifically:

Av. r_e:	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
Approx. Av. r_n:	0.35	0.35	0.35	0.375	0.425	0.475	0.50	0.55	0.60

Thus the lower correlations under the exponential fit tend to be replaced by somewhat stronger correlations under the new fit; intermediate correlations (0.4-0.5) tend to be unchanged, while the higher correlations are somewhat reduced.

At this point, an examination was made of the actual data plots to determine what, if anything, characterized cores of high correlation under the exponential fit. For PR and CC these seemed to be associated with those cores that exhibited an acceleration in growth starting about 30 and 42 years, respectively, prior to the end of the record. For VB the situation was less clear, but the high correlation cores appeared to be associated with accelerated growth starting about 60 years from the end of the record. It should be noted that evidence of such accelerated growth rates showed up in some, but not all, cores from a given site.

This gives rise to the conjecture that, because of their monotonicity, the exponential fits to cores with an accelerated growth rate towards the end of the record will result in spuriously high inter-core correlations. A Monte Carlo study, described in the next section, was undertaken to explore the validity of this conjecture.

ON THE POSSIBILITY OF SPURIOUSLY HIGH CORRELATIONS

The starting point for a Monte Carlo study was the development of growth trends with a single "release" according to equation (1) with n = 1, namely:

$$y_j = \sum_{i=0}^{1} \delta_{ij} a_i (x_j - t_i)^{b_i} \exp[-c_i(x_j - t_i)]$$

The simple increment function:

$$y = a(x - t)^b \exp[-c(x - t)]$$

is fully determined by, inter alia, the value of t, the maximum value of y and the value of x at which the maximum occurs, and the value of y at some other given value of x, which can be selected to control the rate at which the growth falls off from its maximum. In our study, t_0 was set to -35, and the maximum value of y to 2.0 at x = 20. The year in which accelerated growth commenced was set at one of 125, 150, 175 with the value of y at that point reduced to 0.6. Various values were used for the magnitude and spread of the "release pulse". The complete record was taken as 200 years. A typical example is given in Figure 3.

In any one realization two such curves, not necessarily distinct, were so prescribed. Individual data points were then generated by adding normally distributed random variables to these defined trends. Standard normal random variables (mean zero, standard deviation unity) were generated pairwise with prescribed correlation (0.5). After being transformed so the trend and random variable would have constant coefficient of variation (in practice, multiplied by y/D, with D = 4) one member of the pair was added to the first curve and one to the second. This was done independently for each of the 200 years (Figure 3). The possibility of serial correlation has thus been ignored; however, its presence would be expected to have little, if any, effect on the aspect being studied.

An exponential fit was then made to each series of 200 points, the method being that of the program developed by Fritts et al. 1969 (Figure 3). Finally, cross correlations were calculated between the residuals of the model curves (i.e. the transformed normal random variables) and between the residuals of the fitted exponentials, for the final 100 years of the record.

Since the growth curve of VB followed a slightly different pattern from PR and CC, our study also included records of 150 years, where t_0 was set to -15 and the maximum value of y to 2.0 at y = 20. The year in which accelerated growth

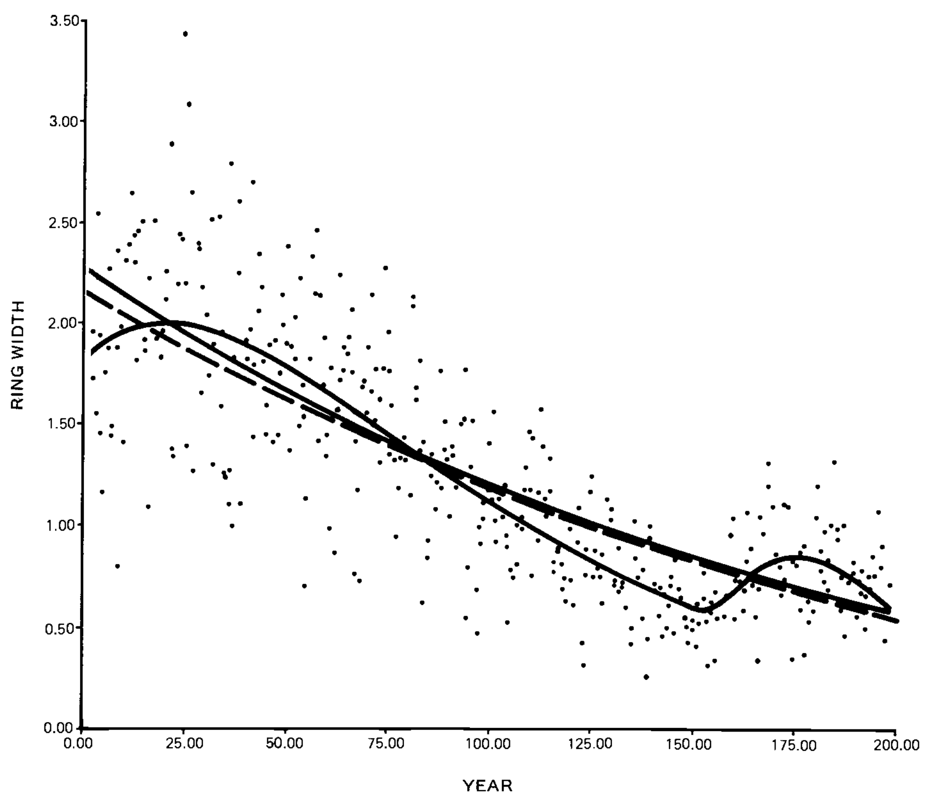

Figure 3. RP and CC simulated growth trends with generated data points and exponential fits (underlying trends identical for both cores).

commenced was set at 90 with the value of y at that point reduced to 1.0. The values for the magnitude and spread of the "release pulse" were varied. As above, two curves were so prescribed, and the data points generated for each of the 150 years. An exponential fit was then made to each series. Cross correlations were calculated for the final 75 years, between the residuals of the model curves and the residuals of the fitted exponentials. A typical example of these curves is given in Figure 4.

It should be noted that the correlation derived from the model curves will differ from 0.5 because of 1) the usual errors induced by sampling and 2) the fact that the variables were transformed and are no longer identically distributed, having constant coefficient of variation rather than constant standard deviation. Since the magnitude of the assumed growth trends does not change greatly over the last 75 to 100 years of the simulated records, the latter should have little effect and, indeed, our checks have shown that the resulting correlation coefficients are consistent with what would be obtained by random sampling of a correlation coefficient of 0.5 with sample size 100.

The results of the Monte Carlo study are summarized in Table 1. Correlation coefficients from different realizations of the same conditions exhibit fairly substantial variability. Indeed, in random samples of size 100 from a bivariate normal with correlation 0.5, it can be shown that the interval required to cover, symetrically, 95% of the outcomes is approximately 0.337 to 0.634 (or 0.309 to 0.655 with a sample size of 75). Thus, for several cases, more than one realization was undertaken and the resulting correlations averaged.

The values presented in Table 1 are not inconsistent with the trends shown in Figure 2. Indeed the exponential fit correlations from the Monte Carlo study are, if anything, greater than necessary to explain the trend observed for $r_e > 0.6$.

It must be pointed out, however, that in those cases reported in Table 1, where the core 1 trend is not identical to the core 2 trend, the "release peak" has been scaled to ensure approximately equal y at the end point (x = 200 or x = 150). We have other

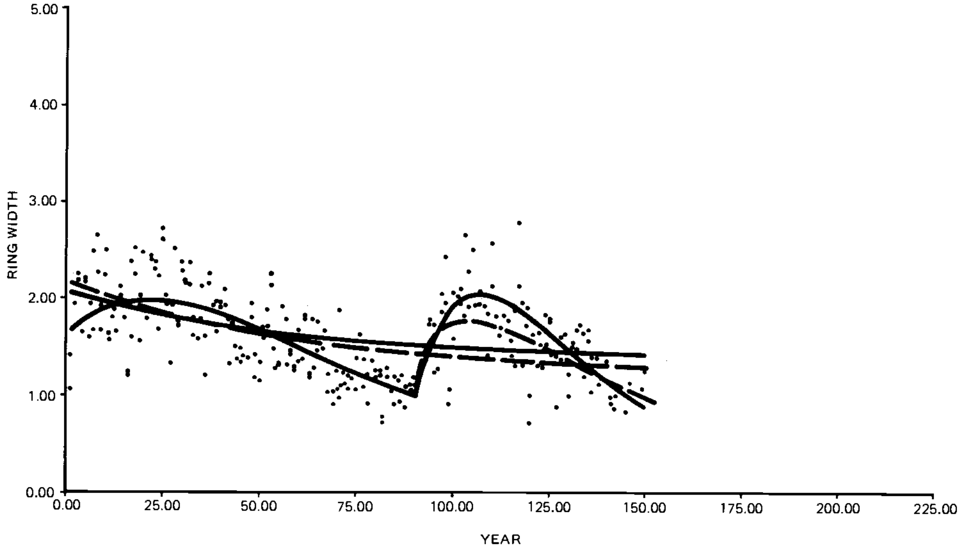

Figure 4. VB simulated growth trends with generated data points and exponential fits.

Table 1. Mean cross correlations of simulated growth trends versus their exponential fit.

	Core 1		Core 2		Correlation +		
Case	Year of release	Year of release peak	Year of release	Year of release peak	Model	Exp. fit	No. realizations
1	125	155	125	155	.4830	.7348	5
2	125	155	125	160	.4904	.8166	4
3	125	155	125	180	.4808	.6348	4
4	* 150	180	150	180	.4762	.6669	1
5	* 150	180	150	180	.4494	.7666	1
6	150	180	150	175	.5454	.7523	4
7	150	180	150	160	.5347	.5182	4
8	175	205	175	205	.6264	.7144	1
** 9	* 90	110	90	110	.3994	.5623	4
** 10	* 90	110	90	110	.3119	.7970	4
** 11	* 90	110	90	110	.4314	.7877	4

* Cases differ in magnitude of "release peak"

\+ Correlations averaged over the realizations

** VB Simulation

cases which show that for releases initiated at different times, and thus, with the trends not taking approximately the same value at the end of the record, r_e is less than r_n, perhaps substantially so.

Nevertheless, the Monte Carlo study has served to show that if trees are released at approximately the same time, then spuriously high values of the cross correlations can easily result from fitting exponential curves. This does not prove that the same thing happened with the actual data under study, although the data are consistent with such a hypothesis.

MATCHED CORES

For completeness it should be mentioned that the cores from each site did not all come from different trees. In most instances, an individual tree is represented by two cores, although there are a few cases of a single core, the other having been lost through breakage or some other unavoidable cause. One would expect to see, in general, a much stronger correlation between cores from the same tree than between cores from different trees, and, indeed, this is observed under both the new and exponential fits. Comparisons are as follows:

	PR	CC	VB
Exp. fit	0.706	0.623	0.685
New fit	0.590	0.565	0.677

If we discount those cases where, on the basis of our Monte Carlo study, we might judge the exponential fit correlation to be spuriously high, we obtain:

	PR	CC	VB
Exp. fit	0.670	0.615	0.667
New fit	0.646	0.605	0.707

We cannot, therefore, differentiate between the two fitting methods on the basis of this information.

LENGTHENING THE SERIES

Our study has not been restricted to the correlation coefficients for the period common to all cores within a single site. Correlation matrices have been computed for a sequence of increasing period lengths, although the number of cores is reduced at each stage. Mean correlations, together with the number of cores, are given in Table 2.

It will be seen that, as the period is extended, the correlations are remarkably stable; in particular, there is no marked tendency for them to decrease, while the critical value for testing formal significance at the 5% level, i.e. for providing evidence of a positive association drops from 0.204 at $n = 67$, and 0.157 at $n = 107$ to 0.133 at $n = 157$, to 0.122 at $n = 187$ and to 0.110 at $n = 227$.

For the reasons outlined in the previous section, the calculation of cross correlation associated with the exponential fits have not been extended back in time.

Table 2. Mean cross correlations for each site with increasing period lengths

	Site		
Length of Period (yrs.)	PR	CC	VB
67			0.451 (23)
87			0.457 (21)
97			0.444 (16)
107	0.480 (25)	0.444 (24)	0.439 (14)
117	0.477 (24)	-	0.433 (12)
127	0.496 (23)	0.426 (23)	0.432 (9)
137	0.492 (20)	-	0.471 (8)
147	0.510 (17)	0.420 (22)	0.505 (5)
157	0.506 (11)	0.430 (19)	0.522 (3)
167	0.517 (10)	-	0.637 (2)
177	0.545 (5)	0.421 (17)	-
187	0.497 (2)	0.412 (16)	-
197	-	0.414 (14)	-
207	-	0.412 (11)	-
217	-	0.403 (9)	-
227	-	0.415 (4)	-

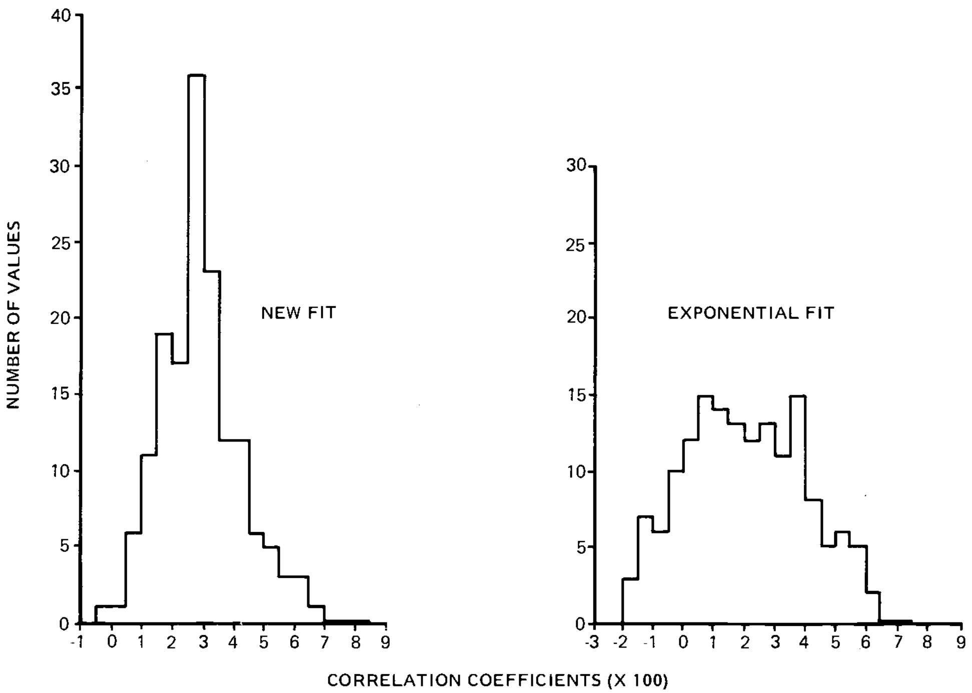

Figure 5. CC vs. PR: intersite correlations for the exponential and new fits.

Table 3. Exponential- and new-fit intersite correlations for selected cores from each site.

	PR X CC		PR X VB		CC X VB	
	Exp. fit	New fit	Exp. fit	New fit	Exp. fit	New fit
Mean	0.200	0.295	0.328	0.369	0.379	0.398
No. values < 5% critical level	69	20	46	20	31	10
No. negative values	27	1	16	0	5	0
No. of years	107		67		67	
Total no. of values	156		156		144	

INTERSITE CORRELATIONS

Since the three sites are in the same watershed, it can be assumed that they would be subject to the same general climatic fluctuations. Hence it is of interest to evaluate the cross correlations between the sites. This has been done for selected cores from PR and CC over the final 107-year period, and for cores from VB over the final 67-year period. Thirteen cores from PR and 12 cores from each of CC and VB were used in this part of the study. These were chosen as being reasonably strongly correlated within site on the grounds that such would be more likely to reflect changing climatic factors and hence have a better chance of exhibiting significant between-site correlations.

The histograms of the correlations derived from the new and exponential fits are typified by that for CC and PR given in Figure 5. A 5% critical value of 0.156 was used for the correlations between PR and CC. The value 0.200 was used for the correlations over the shorter time span (67 vs. 107 years) of PR vs. VB and CC vs. VB. The number of values less than 5% critical value, the number of negative values, and the mean correlation for each pair of sites are given in Table 3. It is seen, in all cases, that the mean correlation for the new fit is greater than that of the exponential fit. We also find that the number of values less than the 5% critical value for the new fit is at least half that of the exponential fit, in each case. Of all these values less than 5% critical value, there is only one negative correlation among the new fits versus 48 negative correlations among the exponential fits. Between the upper tails, i.e. for correlations greater than 0.4, there is only a negligible difference.

The new fits thus seem to be distinctly better in their ability to relate the between-site indices than the exponential fits.

It is, perhaps, worth remarking that, although correlations of 0.3 may seem remarkably weak, they do, with a sample of size 100, provide definite evidence of a positive association.

CONCLUSION AND DISCUSSION

In our tests, the compound increment function has shown up favourably in comparison with the exponential as a means for removing the growth trend from dendrochronological data. In particular, it reduced the incidence of negative and very low positive correlation coefficients of the index values both within and, perhaps more importantly, between sites. Further, exponential fits have been shown to yield spuriously high correlations for pairs of cores that exhibit accelerated growth rate consequent on the trees being released, perhaps as the result of thinning, at roughly the same time. If these cases can be discounted, the new fits are no worse, in terms of the derived correlations, than the exponential and seem physically the more meaningful. Another important point is that the magnitude of the correlation was not reduced by extending the time period back to cover longer periods.

These tests, as promising as they are, still do not guarantee the universally good performance of the suggested method. Its application to other sites, with records extending back for up to 400 years, where exponential fitting has led to very high cross correlations for one site and very low correlations for another, is contemplated and will be reported in due course.

REFERENCES

Fritts, H. C.
1976 *Tree rings and climate.* Academic Press, New York.

Fritts, H. C., J. E. Mosimann, and C. P. Bottorff
1969 A revised computer program for standardizing tree-ring series. *Tree-Ring Bulletin* 29: 15-20.

Warren, W. G.
1980 On removing the growth trend from dendrochronological data. *Tree-Ring Bulletin* 40: 35-44.