Dendrochronologia, 15, 1997: 89-99

TIMOTHY J. OSBORN - KEITH R. BRIFFA - PHILIP D. JONES
Climatic Research Unit, University of East Anglia, Norwich, United Kingdom

ADJUSTING VARIANCE FOR SAMPLE-SIZE IN TREE-RING CHRONOLOGIES AND OTHER REGIONAL MEAN TIMESERIES

Key-Words: Sample size, mean chronology, regional mean, variance bias, variance correction, temperature timeseries.
Parole chiave: Dimensione del campione cronologie medie, media regionale, parametri di variazione, correzione della varianza, serie temporali della temperatura.

Abstract

This note explores the frequently encountered problem that a timeseries formed as an average of a sample of individual timeseries has a variance that depends upon the size of the sample. Methods for adjusting the timeseries to reduce this variance bias are demonstrated – first a simple one and then extensions to it to allow for time-dependent and time-scale-dependent effects. The discussion and techniques are applicable to the construction of tree-ring chronologies from an average of individual cores (or trees) and to the calculation of regional tree-growth timeseries by averaging individual chronologies.

1. Introduction

Consider a sample of timeseries $[x_i'(t), i = 1, \ldots, N]$ where N is the maximum sample size. Each series may cover only part of the time period of interest, or may at times have missing data. At a particular time, t, the available sample size, $n(t)$, will be less than or equal to N. To simplify the following discussion, each timeseries is expressed as anomalies from its long-term mean, $\overline{x}_i$, so that $x_i(t) = x_i'(t) - \overline{x}_i$.

Now consider a *mean* timeseries that is computed from the available sample size at each time:

$$X(t) = \sum_{i=1}^{n(t)} x_i(t)/n(t) \qquad (1)$$

The variance, $S_x^2(t)$, of this mean timeseries is dependent upon the available sample size via a relationship demonstrated by KAGAN (1966; see YEVJEVICH 1972 for an English version) and later applied to dendroclimatology by WIGLEY ET ALII (1984):

$$S_x^2(t) = \overline{s^2}(t)\left[\frac{1 + (n(t) - 1)\overline{r}(t)}{n(t)}\right] \qquad (2)$$

where $\overline{s^2}(t)$ is the mean of the variances of the individual timeseries and $\overline{r}(t)$ is the mean correlation between all pairs of timeseries. The latter parameter is also used in the computation of some basic chronology statistics described by BRIFFA, JONES (1990), who also extensively discuss the estimation of $\overline{r}$ [it may be estimated from the common overlap period of all the samples, but if this is short it is preferable that $\overline{r}$ be based on the average of individual correlations that have each been calculated over their maximum possible overlap (BRIFFA, JONES 1990)] Both $\overline{s^2}$ and $\overline{r}$ depend upon the particular sample of available timeseries; since the sample is time-dependent, so are they. To begin with, however, assume that they are constant. Under this assumption, $\overline{r}$ is computed from the full sample size, regardless of which series are available at which times (section 3a considers the case where this assumption is dropped for $\overline{r}$). The assumption is valid for $\overline{s^2}$ if the individual timeseries have equal or near-equal variances (e.g., if they are normalised timeseries).

The variance of X increases if the variance of the individual timeseries increases or if the common signal between them (i.e., $\overline{r}$) increases; these are real signals that may be of interest. An artificial signal will be introduced into the variance of X, however, if the sample size varies through time. In the particular case where the individual timeseries are all independent (i.e., $\overline{r} = 0$), equation (2) reduces to

$$S_X{}^2(t) = \overline{s^2}/n(t) \qquad (3)$$

Comparing this result with equation (2) leads to the definition of the effective independent sample size,

$$n'(t) = \frac{n(t)}{1+(n(t)-1)\overline{r}} \qquad (4)$$

and this can be applied to the more general case of $\overline{r} \geq 0$ to yield

$$S_X{}^2(t) = \overline{s^2}/n'(t) \qquad (5)$$

Figure 1 shows $1/n'$ as a function of n for a variety of $\overline{r}$ values. Given that the variance of our aggregate is proportional to $1/n'$, it is important to note that it will be relatively independent of sample size if $\overline{r}$ is high. If $\overline{r}$ is low, however, variance will increase strongly as n falls below about 10. To further illustrate this, consider two examples: one with $\overline{r}$ = 0.05 and the other with $\overline{r}$ = 0.5. If the sample size is reduced from 15 to 10, the variance would be expected to increase by 28% in the former case, but by just 3% in the latter. Further reduction (to n = 5) yields further expected increases of 66% and 9%, respectively. It is only when the sample size is reduced down to n = 2 that the variance is significantly affected when $\overline{r}$ = 0.5 (in this case, variance increases by an additional 25%).

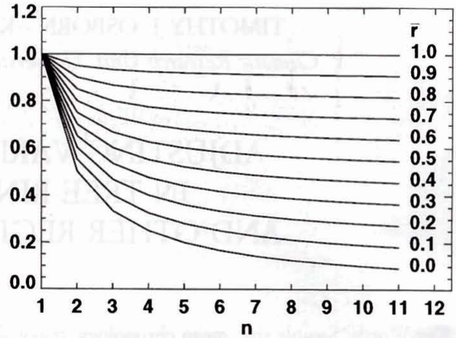

Figure 1 - Dependence of the inverse of the effective independent sample size ($1/n'$) on the sample size (n), for a range of $\overline{r}$ values. The latter are marked next to the corresponding curves.

Within- and between-site tree-ring chronology $\overline{r}$ values vary considerably. They are influenced by location (ecology; climate signal coherence); the type of tree-ring variable (ring-width or density); tree species; and the standardisation approach used. A typical within-site $\overline{r}$ value for a western

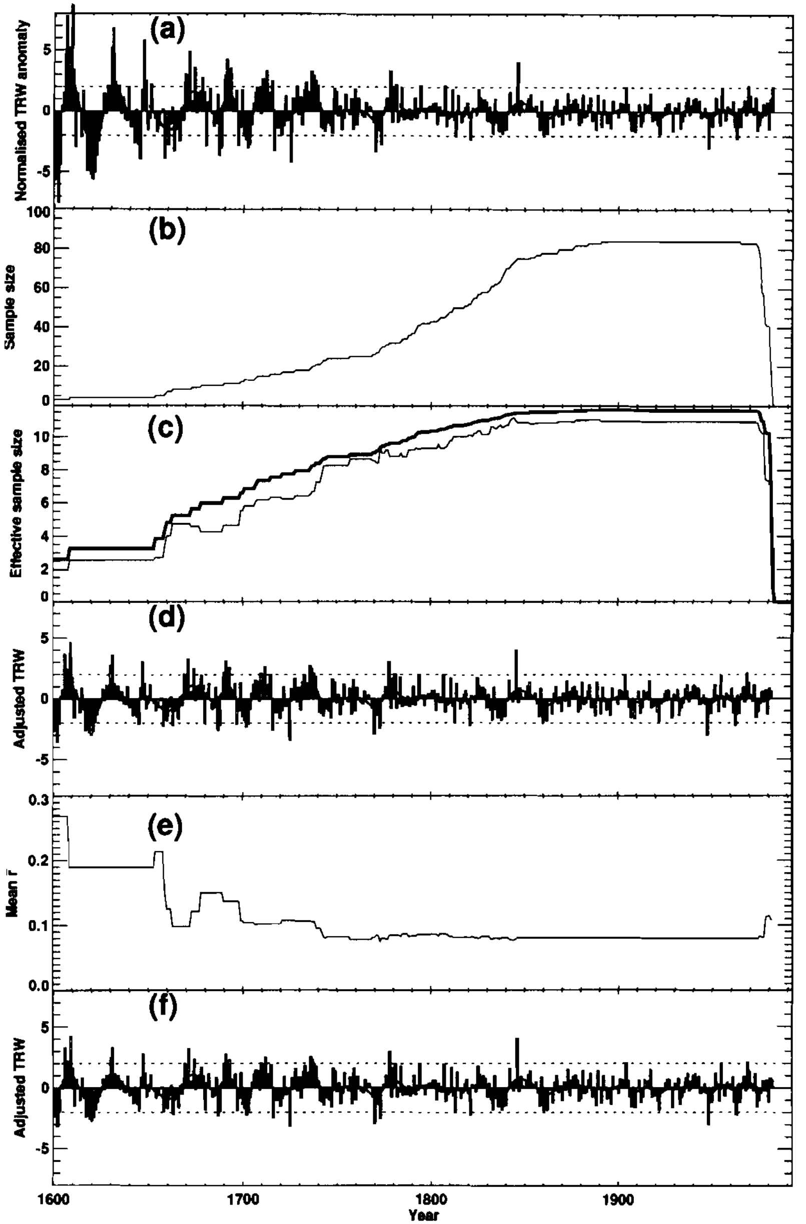

Figure 2 - Timeseries relating to a group of southern European tree-ring-width chronologies: (a) normalised mean of all chronologies (20-yr smoothed curve also shown); (b) sample size (i.e., number of chronologies with data); (c) effective independent sample size (thick line using constant $\bar{r}$, thin line using time-dependent $\bar{r}$); (d) mean timeseries after scaling using constant $\bar{r}$ adjustment; (e) time-dependence of $\bar{r}$; (f) mean timeseries after scaling using time-dependent $\bar{r}$ adjustment.

U.S. conifer ring-width chronology is around 0.6 whereas for an eastern U.S. hardwood chronology it is only 0.3 (DE WITT, AMES 1978). Deciduous sites in Europe have $\bar{r}$ values that can be as low as 0.2 (BRIFFA, JONES 1990). In a sample of ring-width and maximum-latewood-density site chronologies at 49 locations in northern Siberia (BRIFFA ET ALII 1996), within-site $\bar{r}$ for the ring-width data ranged from 0.28 to 0.71, and from 0.29 to 0.74 for density.

In a group of 17 regional chronologies (almost all spruce) in northern North America, each with between 2 and 8 individual site series (SCHWEINGRUBER ET ALII 1993; BRIFFA ET ALII 1994), the between-site $\bar{r}$ ranged from 0.04 to 0.47 for ring-width and from 0.26 to 0.76 for density (notably higher).

In general, $\bar{r}$ falls further as the region of interest gets larger. Figure 2a shows an example with low $\bar{r}$ – the average of N = 87 tree-ring width chronologies from high-altitude regions of Southern Europe (see BRIFFA ET ALII 1998a for details of dataset and region). The time-dependent available sample size (Figure 2b) indicates that the increased variance pre-1750 is likely to be due to the reduced sample size in the early period. For the sample as a whole, the mean intersample correlation is $\bar{r}$ = 0.07. The effective sample size, computed from equation (4), shows large changes over the 1600-1750 period (Figure 2c, thick line).

2. Simple adjustment

Previous methods (e.g., SHIYATOV ET ALII 1990; COOK, HOLMES no date; PARKER ET ALII 1992) have attempted to remove such artificial time-dependent variance changes. All follow an empirical approach, however, that could also remove any variance changes that might be real (i.e., a signal that needs to be retained). The method presented here is theoretically based and does not have this disadvantage.

If the correlation matrix is computed using the overlapping sections of each pair of timeseries, a single $\bar{r}$ value associated with the mean timeseries can be found by averaging the sub-diagonal correlations. Equation (4) then provides the time-dependent effective sample size (if supplied with the time-dependent available sample size). We would then expect the variance of the mean timeseries to vary according to equation (5). If we adjust the mean timeseries by

$$Y(t) = X(t)\sqrt{n'(t)} \qquad (6)$$

then we would expect that the variance $S_Y{}^2$ of $Y(t)$ would be independent of sample size (but would still have any real variance signals that are present in the data).

Equation (6) results in an adjusted timeseries that has a variance inflated to that of a timeseries computed from a sample size of one throughout (i.e., $S_Y{}^2 = \overline{s^2}$). If that is what is required, or if the mean timeseries is subsequently to be normalised over some reference period anyway, then this is adequate. Such is the case in our example (Figure 2) where the original timeseries was normalized using the mean and standard deviation of the 1901-70 period. Figure 2d shows the adjusted timeseries re-normalised over the same period. Much of the early high variance has been reduced; some remains however.

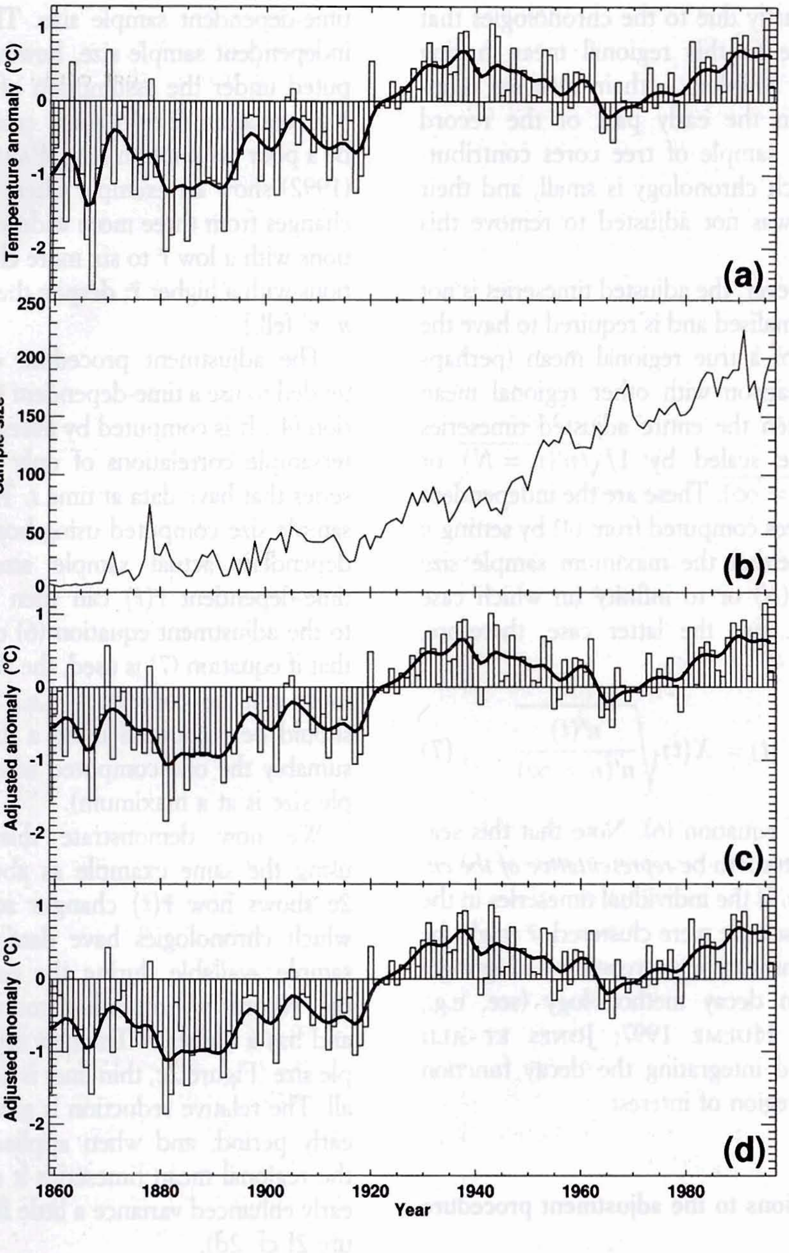

Figure 3 - Timeseries relating to instrumental temperature anomalies north of 65°N: (a) mean timeseries from all grid boxes with data; (b) sample size; (c) mean timeseries after scaling all timescales using constant $\overline{r}$ adjustment; (d) mean timeseries after scaling using a timescale-dependent $\overline{r}$ adjustment. 10-yr smoothed curves also shown in (a), (c) and (d).

This is partly due to the chronologies that contribute to this regional mean having the same problem – their variance is increased in the early part of the record when the sample of tree cores contributing to each chronology is small, and their variance was not adjusted to remove this bias.

If, however, the adjusted timeseries is not to be normalised and is required to have the variance of a true regional mean (perhaps for comparison with other regional mean series), then the entire adjusted timeseries should be scaled by $1/\sqrt{n'(n=N)}$ or $1/\sqrt{n'(n=\infty)}$. These are the independent sample sizes computed from (4) by setting n equal to either the maximum sample size available (N) or to infinity (in which case $n' = 1/\bar{r}$). For the latter case, therefore, we have

$$Y(t) = X(t)\sqrt{\frac{n'(t)}{n'(n=\infty)}} \qquad (7)$$

instead of equation (6). Note that this scaling requires $\bar{r}$ to be *representative of the entire region*. If the individual timeseries in the available sample were clustered, $\bar{r}$ might be biased. This can be overcome by adopting a correlation decay methodology (see, e.g., OSBORN, HULME 1997; JONES ET ALII 1997), and integrating the decay function over the region of interest.

3. Extensions to the adjustment procedure

(a) *Time-dependent parameters*

The adjustments suggested in section 2 [i.e., equation (6) or (7)] make a time-dependent adjustment that is based on a time-dependent sample size. The effective independent sample size, however, is computed under the assumption of a constant $\bar{r}$ in equation (2). In some cases this may be a poor assumption [e.g., PARKER ET ALII (1992) show an example where the sample changes from three more widely spread stations with a low $\bar{r}$ to six more clustered stations with a higher $\bar{r}$; despite the increase in n, n' fell.].

The adjustment procedure can be extended to use a time-dependent $\bar{r}$(t) in equation (4). It is computed by averaging the intersample correlations of only those timeseries that have data at time t. The effective sample size computed using both the time-dependent actual sample size and the time-dependent $\bar{r}(t)$ can then be applied to the adjustment equation (6) or (7). Note that if equation (7) is used, the denominator must still be constant – i.e., $n'(n = \infty)$ should be computed using a fixed $\bar{r}$ (presumably the one computed when the sample size is at a maximum).

We now demonstrate this extension using the same example as above. Figure 2e shows how $\bar{r}(t)$ changes according to which chronologies have data. The small sample available during the early part of the record is somewhat more clustered and has a higher $\bar{r}$. The new effective sample size (Figure 2c, thin line) is weaker overall. The relative reduction is greater for the early period, and when applied to adjust the regional mean timeseries it reduces the early enhanced variance a little further (Figure 2f cf. 2d).

(b) *Timescale-dependent parameters*

A somewhat different example is given in

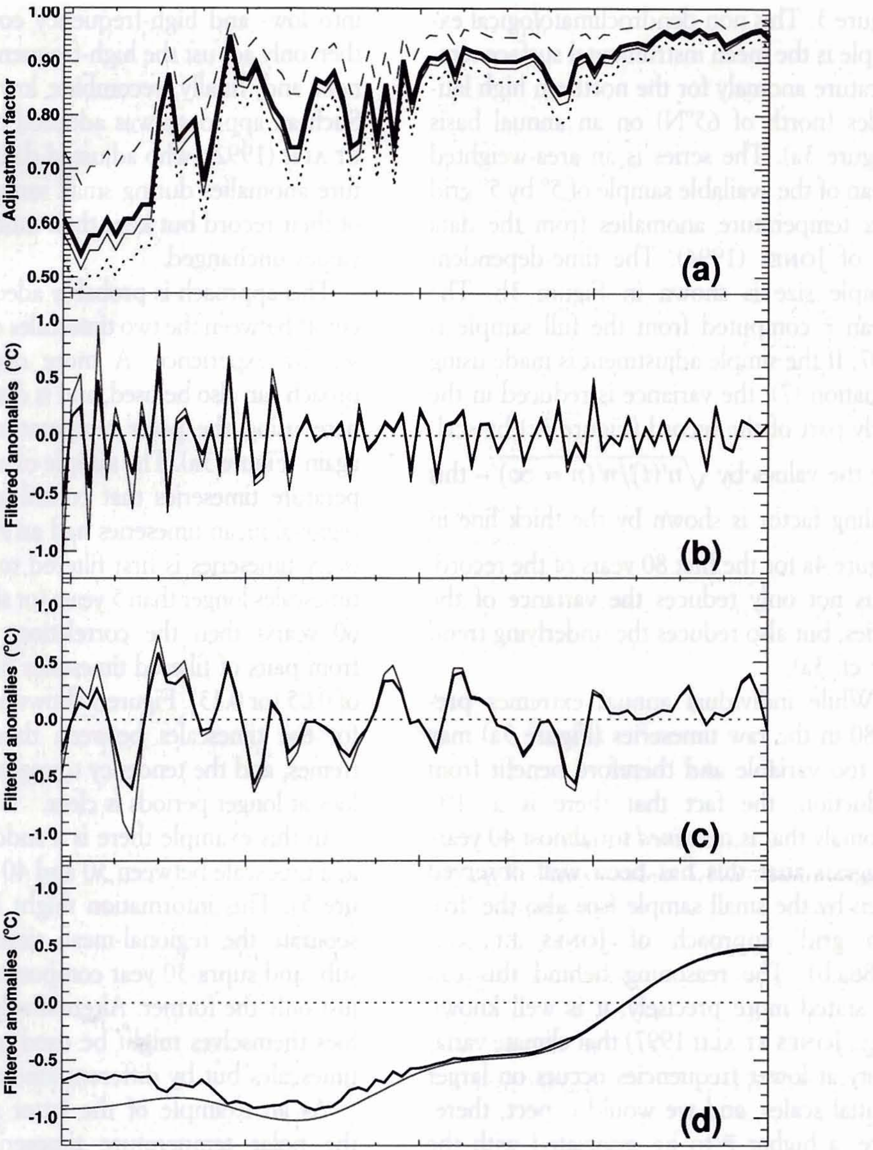

Figure 4 - Timeseries relating to instrumental temperature anomalies north of 65°N: (a) scaling factors used to correct variance bias (thick line using timescale-independent $\bar{r}$, dotted, thin and dashed lines using $\bar{r}$ values associated with timescales 0-5, 5-40 and 40+ years, respectively); (b) mean temperature timeseries filtered to emphasise timescales 0-5 years (thin line unadjusted, thick line after adjustment); (c) as (b) but for 5-40 years; (d) as (b) but for 40+ years.

Figure 3. This non-dendroclimatological example is the mean instrumental surface temperature anomaly for the northern high latitudes (north of 65°N) on an annual basis (Figure 3a). The series is an area-weighted mean of the available sample of 5° by 5° grid box temperature anomalies from the data set of JONES (1994). The time-dependent sample size is shown in Figure 3b. The mean $\bar{r}$ computed from the full sample is 0.07. If the simple adjustment is made using equation (7), the variance is reduced in the early part of the record (Figure 3c) by scaling the values by $\sqrt{n'(t)/n'(n=\infty)}$ – this scaling factor is shown by the thick line in Figure 4a for the first 80 years of the record. This not only reduces the variance of the series, but also reduces the underlying trend (3c cf. 3a).

While individual annual extremes pre-1880 in the raw timeseries (Figure 3a) may be too variable and therefore benefit from reduction, the fact that there is a -1°C anomaly that is *sustained* for almost 40 years suggests that this has been well observed even by the small sample (see also the 'frozen grid' approach of JONES ET ALII 1986a,b). The reasoning behind this can be stated more precisely: it is well known (e.g., JONES ET ALII 1997) that climate variability at lower frequencies occurs on larger spatial scales, and we would expect, therefore, a higher $\bar{r}$ to be associated with the low-frequency components of the mean timeseries. Since the adjustment to remove variance biases is smaller for higher $\bar{r}$, the low-frequency components should be adjusted to a lesser extent than has been done in Figure 3c.

One option to allow for such timescale dependence is to separate the timeseries into low- and high-frequency components, then only adj ust the high-frequency component and, finally, recombine low and high. Such an approach was adopted by PARKER ET ALII (1992), who adjusted daily temperature anomalies during small sample periods of their record but kept their monthly mean values unchanged.

This approach is probably adequate if the cutoff between the two timescales can be chosen by experience. A more objective approach can also be used, and is demonstrated here using the polar temperature example again (Figure 3a). The sample of annual temperature timeseries that contributed to this regional mean timeseries had an $\bar{r}$ of 0.07. If every timeseries is first filtered to damp out timescales longer than 5 years (or shorter than 60 years) then the correlations computed from pairs of filtered timeseries hare a mean of 0.05 (or 0.13). Figure 5 shows how $\bar{r}$ varies for the timescales between these two extremes, and the tendency towards higher values at longer periods is clear.

In this example there is a sudden change at a timescale between 30 and 40 years (Figure 5). This information might be used to separate the regional-mean timeseries into sub- and supra-30 year components and adjust only the former. Alternatively, the $\bar{r}$ values themselves might be used to adjust all timescales but by different amounts.

As an example of the latter suggestion, the polar temperature timeseries (Figure 3a) has been filtered to separate it into three period bands: variations on timescales of less than 5 years, 5 to 40 years, and greater than 40 years. The early parts of these filtered timeseries are shown by the thin lines in Figures 4b, 4c and 4d respectively. The individual grid-box timeseries that contribute to this regional mean are also filtered

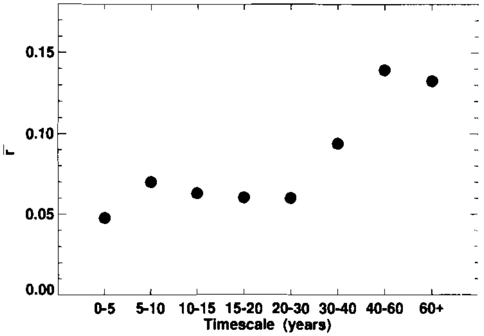

Figure 5 - Dependence of $\bar{r}$ on timescale of temperature variability in anomalies north of 65°N.

in the same way and then cross-correlated to obtain a value of $\bar{r}$ for each timescale (0.048, 0.062 and 0.134).

Each timescale is then adjusted by the scaling factors shown in Figure 4a (dotted = 0 to 5 years, thin = 5 to 40 years, dashed = 40+ years), computed from (7) with the appropriate value of $\bar{r}$ for each timescale. During the first decade of the record, low-frequency variations are reduced by 30%, while high-frequency variations are reduced by 50%. The timescale-independent approach reduced all timescales by a little more than 40% during the same decade (thick line, Figure 4a). The thick lines in Figures 4b, c, d show the filtered timeseries after scaling has been applied.

Combining (by addition) the three filtered and adjusted timeseries yields the series shown in Figure 3d. Compared to the result of the timescale-independent adjustment (Figure 3c), the low-frequency/trend is a little stronger (although still reduced compared to the raw timeseries – Figure 3a), and the early high-frequency changes are suppressed slightly more. (Ideally, the filters used should minimise leakage between frequency bands.)

4. Conclusions

This paper suggests a simple but theoretically-based approach to correcting bias in the variance of tree-ring chronologies, either at a site or regional-average level, that arises as a result of changes in sample replication (sample depth). A correction or adjustment is not necessary in all cases; the magnitude of the bias is discussed in terms of intersample correlation ($\bar{r}$) and sample size.

On the basis of limited experimentation, the simple scaling of the average timeseries by

$$\sqrt{\frac{n(t)}{1+\big(n(t)-1\big)\bar{r}}}$$

produces a noticeable improvement in the apparent stability of variance through time. Some alternatives are also suggested that allow one to account for time-dependent and timescale-dependent effects. The adjustment procedures investigated here have been applied under the assumption that all individual timeseries have similar variances. If that is not the case, then the mean variance of the sample of available timeseries may have a secondary artificial time dependence. An additional adjustment can be made to remove this artefact.

One conclusion that comes from the timescale-dependent adjustment procedure is that high-frequency variations should normally be adjusted more strongly. This demonstrates that variance-adjustment is of primary importance when the high-frequency aspects of the timeseries are to be analysed (e.g., in the case of extreme value ranking – Briffa et alii 1998b).

ACKNOWLEDGEMENTS

TJO acknowledges the support of the UK Natural Environment Research Council (GR9/02522), KRB is currently funded by the European Environment and Climate Programme (DGXII grant ENV4-CT95-0127: ADVANCE-10K) and PDJ by the US Dept. of Energy (Environmental Sciences Division grant DE-FG02-86ER60397).

LITERATURE

Briffa K.R., Jones P.D., 1990. Basic chronology statistics and assessment. In: Cook E.R., Kairiukstis L.A., Methods of Dendrochronology, Kluwer, Dordrecht, 137-152.

Briffa K.R., Jones P.D., Schweingruber F.H., 1994. Summer temperatures across northern North America: regional reconstructions from 1760 using tree-ring densities. *J. Geophys. Res.*, 99D: 25835-25844.

Briffa K.R., Jones P.D., Schweingruber F.H., Shiyatov S.G., Vaganov E.A., 1996. Development of a northern Eurasian chronology network: rationale and preliminary results of comparative ring-width and densitometric analyses in northern Russia. In: Dean J.S., Meko D.M., Swetnam T.W., Tree rings, environment and humanity, Radiocarbon, University of Arizona, Tucson, 25-41.

Briffa K.R., Schweingruber F.H., Jones P.D., Osborn T.J., Shiyatov S.G., Vaganov E.A., 1998a. Reduced sensitivity of recent tree-growth to temperature at high northern latitudes. *Nature*, 391: 678-682.

Briffa K.R., Jones P.D., Schweingruber F.H., Osborn T.J., 1998b. Influence of volcanic eruptions on northern hemisphere summer temperatures over the past 600 years. *Nature*, 393: 450-455.

Cook E.R., Holmes R.L., no date. Users Manual for Program ARSTND. Laboratory of Tree-Ring Research, University of Arizona, Tucson, AZ, USA.

De Witt E., Ames M., 1978. Tree Ring Chronologies of Eastern North America. Chronology Series IV 1, Laboratory of Tree-Ring Research, University of Arizona, Tucson, 42 pp.

Jones P.D., 1994. Hemispheric surface air temperature variations: a reanalysis and an update to 1993. *J. Climate*, 7: 1794-1802.

Jones P.D., Raper S.C.B., Bradley R.S, Diaz H.F., Kelly P.M., Wigley T.M.L., 1986a. Northern Hemisphere surface air temperature variations, 1851-1984. *J. Clim. Appl. Met.*, 25: 161-179.

Jones P.D., Raper S.C.B., Wigley T.M.L., 1986b. Southern Hemisphere surface air temperature variations, 1851-1984. *J. Clim. Appl. Met.*, 25: 1213-1230.

Jones P.D., Osborn T.J., Briffa K.R., 1997. Estimating sampling errors of large-scale temperature averages. *J. Climate*, 10: 2548-2568.

Kagan R.L., 1966. An evaluation of the representativeness of precipitation data. *Gidrometeoizdat*, 213 pp. (In Russian).

Osborn T.J., Hulme M., 1997. Development of a relationship between station and grid-box rainday frequency for climate model evaluation. *J. Climate*, 10: 1885-1908.

Parker D.E., Legg T.P., Folland C.K., 1992. A new daily central England temperature series, 1772-1991. *Int. J. Climatol.*, 12: 317-342.

Schweingruber F.H., Briffa K.R., Nogler P., 1993. A tree-ring densitometric transect from Alaska to Labrador: comparison of ring-width and maximum-latewood-density chronologies in the conifer belt of northern North America. *Int. J. Biometeorology*, 37: 151-169.

Shiyatov S., Mazepa V., Cook E.R., 1990. Correcting for trend in variance due to changing sample size. In Cook E.R., Kairiukstis L.A., Methods of Dendrochronology, Kluwer, Dordrecht, 133-137.

Wigley T.M.L., Briffa K.R., Jones P.D., 1984. On the average value of correlated time series with applications in dendroclimatology and hydrometeorology. *J. Clim. Appl. Met.*, 23: 201-213.

Yevjevich V., 1972. Probability and Statistics in Hydrology. Water Resources Publications, Fort Collins, CO, 302 pp.

SUMMARY

Adjusting variance for sample-size in tree-ring chronologies and other regional mean timeseries

It is demonstrated, using tree-ring and temperature examples, that a mean timeseries formed from an average of a sample of individual timeseries has a variance that increases when the sample is small. This variance bias is artificial and is not a real signal. Methods for adjusting the timeseries to reduce this variance bias are developed and applied – first a simple one and then extensions to allow for time-dependent and timescale-dependent parameter changes.

ZUSAMMENFASSUNG

Varianzanpassung von Jahrringchronologien und anderen regionalen Zeitreihen in Abhängigkeit vom Stichprobenumfang.

Es wird am Beispiel von Jahrring- und Temperaturreihen gezeigt, daß die Varianz einer aus einzelnen Zeitreihen gemittelten Zeitreihe bei kleinen Stichproben zunimmt. Eine derartige Varianz ist künstlich und kein reales Signal. Es werden Methoden der Varianzkorrektur zur Verminderung dieses einseitigen Einflusses entwickelt und angewandt, zunächst eine einfache und danach erweiterte Versionen für zeitabhängige und zeitskalenabhängige Parameterveränderungen.

RIASSUNTO

Correzione della varianza in funzione della dimensione del campione in cronologie anulari e in altre serie temporali regionali

Si dimostra, usando esempi di cronologie e temperature, che una serie media temporale formata da un campione di serie singole ha una varianza che cresce quando il campione è piccolo.

Questa varianza è artificiale e non costituisce un segnale reale. In questo lavoro sono sviluppati e applicati i metodi per correggere le serie temporali al fine di ridurre la distorsione di stima della varianza. In sostanza si applica prima un metodo semplice e poi una sua estensione al fine di fissare variazioni nei parametri, che dipendono dal tempo e dalla scala temporale.

Indirizzo degli Autori:

TIMOTHY J. OSBORN - KEITH R. BRIFFA - PHILIP D. JONES
Climatic Research Unit
University of East Anglia
Norwich NR4 7TJ
U.K.
tel: +44 1603 592089
fax: +44 1603 507784
e-mail: t.osborn@uea.ac.uk