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

DENDROCLIMATIC CALIBRATION AND VERIFICATION USING REGIONALLY AVERAGED AND SINGLE STATION PRECIPITATION DATA

T. J. BLASING
D. N. DUVICK

and

D. C. WEST

Environmental Sciences Division
Oak Ridge National Laboratory, Tennessee

ABSTRACT

The average ring-width index of two published chronologies from the eastern Tennessee climatic division was used as a single predictor variable in linear regression to reconstruct May-June precipitation. Regression equations obtained using regionally averaged precipitation data from stations within the climatic division were compared with regressions obtained using single-station (Knoxville, Tennessee) data for comparable periods. The (regionally averaged) division data always provided the better calibration statistics for the regression equations. When the regressions calibrated using division data were verified with independent data for the climatic division and for Knoxville, the better results were always obtained for the division data. When regressions calibrated using single-station data were verified, the independent division data once again provided better results than the independent Knoxville data. Regionally averaged precipitation data also provided more satisfactory results than single-station data in a similar experiment for central Iowa, and probably provides better results in general for this type of dendroclimatic experiment.

Der durchschnittliche Jahrringbreitenindex von zwei veröffentlichten Chronologien aus der östlichen Klimaregion von Tennessee wurden als einzige Prädiktorvariable zur Rekonstruktion der Mai/Juni-Temperatur bei einer linearen Regression benutzt. Die Regressions gleichungen auf der Grundlage regional gemittelter Niederschlagsdaten von Stationen innerhalb der Klimaregion wurden mit den Regressionen verglichen, die sich mit den Daten einer einzelnen Station (Knoxville, Tennessee) für vergleichbare Perioden ergeben haben. Die regional gemittelten Daten ergaben stets bessere Eichparameter für die Regressionsgleichungen. Wenn die mit Regionaldaten berechneten Regressionen mit unabhängigen Daten der Klimaregion und der Station Knoxville verifiziert werden, führen die Regionaldaten stets zu besseren Ergebnissen. Wenn die mit Einzelstationsdaten berechneten Regressionen verifiziert werden, führen wiederum die Regionaldaten zu besseren Ergebnissen als die unabhängigen Knoxville-Daten. Regional gemittelte Niederschlagswerte haben in einem ähnlichen Experiment für Central-Iowa ebenfalls befriedigendere Ergebnisse erbracht als Einzelstationsdaten. Sie sind vermutlich allgemein für derartige dendroklimatologische Untersuchungen besser geeignet.

Les indices dendrochronologiques moyens obtenus à partir de deux chronologies déjà publiées provenant de la zone climatique du Tennessee oriental ont été utilisées dans une régression linéaire comme prédicteur simple en vue de reconstituer les précipitations de mai — juin. Les équations de régression utilisant les données moyennes régionales obtenues à partir des stations situées dans la zone climatique sont comparées avec des régressions calculées à partir d'une seule station (Knoxville, Tennessee) au cours de périodes comparables. Les données régionales procurent toujours les meilleures calibrations statistiques pour les équations de régression. Lorsque les régressions calibrées d'après les valeurs régionales sont vérifiées en utilisant des données climatiques indépendantes de la région et de Knoxville, les meilleurs résultats sont toujours obtenus avec les données moyennes de la région. Lorsque les régressions calibrées à l'aide d'une seule station sont vérifiées avec des données indépendantes, les valeurs régionales fournissent une fois de plus de meilleurs résultats que les mesures de Knoxville. En Iowa central, les précipitations régionales moyennes fournissent également des résultats plus favorables que les données d'une seule station. Il est probable que ce fait soit général dans ce type d'expérience dendroclimatique.

Empirically derived equations of varying degrees of complexity have been used to obtain quantitative reconstructions of past climate from tree-ring data. In the simplest case, a linear regression equation is used to reconstruct past values of a single climatic variable from ring-width indices of a single tree-ring chronology, or from a mean of two or more chronologies which have been merged to form a single chronology. The slope and y-intercept of the regression lines are computed from a series of ring-width indices, x, and the corresponding values of the climatic variable to be estimated, y. This process is known as *calibration* of the regression equation, and the values of x and y used are called the calibration data. Usually, some of the climatic data, y, and the corresponding values of x, are not used in the calibration but are saved as an independent sample. Climatic data from this sample are then compared with their estimates to evaluate the accuracy of the regression equation when it is applied to data which are independent of the calibration. This process is called *verification.*

Ideally, the climatic data used in calibration and verification should represent climatic variations at the site where the trees were sampled, but usually the nearest weather station is at least several kilometers from the tree site and has very different characteristics of exposure, elevation, etc. Data at any single station may thus represent differences between the microclimate of the weather instruments and that of the tree site. The average climatic record of several stations over a large area is representative of regional climatic conditions, and does not reflect localized phenomena which may be characteristic of the climatic record at a single station. Tree-ring data may therefore have more variance in common with the regionally averaged climatic record than with the climatic record of the nearest weather recording station. For this reason, regionally averaged data from state climatic divisions are often used to calibrate climatic variables with one or more ring-width chronologies within a given climatic division.

We used published ring-width chronologies from Steiner's Woods (*Quercus alba*) and the Norris Watershed boundary (*Pinus echinata*) to estimate May-June precipitation for the eastern Tennessee climatic division. These chronologies were sampled and worked up by staff members of the Laboratory of Tree-Ring Research at the University of Arizona. The ring-width values are found in DeWitt and Ames (1978). The average ring-width index of these chronologies was used as a single predictor variable, x, in straight-line regression. The regionally averaged division precipitation data since 1931 were used to calibrate the regression equation, and verification was obtained from data at a single station, Knoxville, which is relatively near the tree sites and has a climatic record back to 1871. In this case, all available division data were used for calibration because convincing verification could be obtained from the longer Knoxville record. However, it is quite possible that even better verification statistics could have been achieved if division data had been available to use for verification.

To compare the merits of using Knoxville data with those of using division data for calibration and verification, we subjected the available data to several statistical analyses. We concluded from these that division data were superior for calibration and verification, at least for the case we investigated. Comparisons of division data with single-station data in central Iowa also led to the conclusion that division data were superior. We tentatively concluded that it is generally true that division data, or possibly other regionally averaged data, are superior to single-station data for dendroclimatic studies in isolated regions, and we recommend that exploratory dendroclimatic studies in such regions involve calibrations with regionally averaged data.

Statistics used in our comparisons are among those used by Fritts and others (1979). For verification, we used the correlation coefficient between the actual and estimated values of the climatic variable, y. However, for the independent data, the correlation coefficient only measures the pattern similarity of the two time series, and does not account for the magnitudes of the differences between observed values and their estimates. A statistic which accounts for these magnitudes is the reduction of error, RE, which is defined in equation (1).

$$RE = 1 - [\Sigma(y_i - \hat{y}_i)^2 / \Sigma(y_i - \bar{y})^2], \quad i = 1,n, \tag{1}$$

where y_i is one of n observed values, $\hat{y}_i$ is its estimate, and $\bar{y}$ is the mean of the observed values for those climatic data used to calibrate the regression equation (*not* the average value of the independent data). The value of RE may theoretically range from minus infinity to +1.0. As Fritts and others (1979) noted, this is a very rigorous statistic and "any positive value indicates there is some information in the reconstructions."

Calibration using Knoxville precipitation data from 1931-1972 resulted in an F-value of 8.45, with 17.5% of the variance of May-June precipitation accounted for and a correlation coefficient between actual and estimated values of +0.42. Verification on 60 years of independent data, beginning in 1871, produced an RE of +0.17 and a correlation coefficient of +0.44. From this we concluded that the calibration statistics hold up remarkably well when verified with independent data. Calibration with the regionally averaged division data resulted in an F-value of 13.44, with 25.2% of the precipitation variance accounted for by regression and a correlation coefficient of +0.50. This result was more statistically significant than that for the Knoxville calibration. However, these statistics can be somewhat misleading because the actual sum of squares due to regression was 8.06 for the division calibration and 8.57 for the Knoxville calibration. At Knoxville, extreme values of observed precipitation are more pronounced than in the division data, so the residual variance for the Knoxville calibration was 40.6 while that for the division calibration was 24.0. Because the sum of squares due to regression was about the same in each case, verification of both calibrations with the same set of independent data would be expected to provide similar results, and that is what happened. Verification of the division calibration with independent data from Knoxville, for 1871-1930, produced an RE of +0.17 and a correlation coefficient of +0.44, which are the same numbers obtained when the Knoxville calibration was verified on the same data. From this we hypothesize that if the division data before 1931 had been available for independent checking, then higher values of RE and correlation coefficient would have been obtained on independent data. This would have been primarily due to less residual variance rather than to any increase in variance accounted for by regression.

To test this hypothesis, we performed two experiments using data exclusively from the period 1931-1972, for which both division data and Knoxville data are available. The structure of these experiments was suggested in a paper by Gordon (1980). The first experiment involved using half the data for calibration and the other half for verification. In this case we used even-numbered years for calibration and verified on odd years, and then reversed the procedure. In both cases we calibrated with regionally averaged division climatic data, and compared verifications using division

Table 1. Dendroclimatic calibration and verification statistics for tree-ring data from within the eastern Tennessee climatic division. Two calibrations were obtained with May-June precipitation. The first used regionally averaged data for the climatic division; the second used single-station data from Knoxville. For each calibration, the verification results for both the division data and the Knoxville data are shown.

Calibration data	**division**		**Knoxville**	
Calibration period	even years	odd years	even years	odd years
Percentage variance due to regression	28.8	17.7	21.1	11.0
Correlation coefficient	+0.54	+0.42	+0.46	+0.33
Verification — division data	odd years	even years	odd years	even years
Reduction of error	+0.18	+0.25	+0.13	+0.23
Correlation coefficient	+0.40	+0.51	+0.40	+0.51
Verification — Knoxville data	odd years	even years	odd years	even years
Reduction of error	−0.06	+0.18	+0.05	+0.16
Correlation coefficient	+0.32	+0.44	+0.32	+0.44

data with verifications using Knoxville data. The results are summarized in Table 1. In both cases the division data produced better verification statistics than the Knoxville data, and in one case the use of Knoxville data even led to a negative reduction of error and generally insignificant verification. It was also noted that those years which produced better calibration statistics, the even years, also provided the best verification statistics.

We then ran separate calibrations using the even and odd years, respectively, for the Knoxville climatic data. These calibrations, and their verifications, are also summarized in Table 1. When the even years were calibrated, the percentage of variance accounted for by regression was 21.1 as compared with 28.8 for the division data. When odd-numbered years were calibrated, only 11.0% of the variance was accounted for by regression, and the corresponding F-ratio of 2.34 indicated an insignificant calibration. Thus, when the disadvantages of single-station data are combined with those of a small sample size (21 years) it might be erroneously concluded that no relationship exists between May-June precipitation and ring-width indices for the eastern Tennessee chronologies. When the regionally averaged division data were used for calibration, the variance accounted for increased to 17.7%, although the corresponding F-ratio of 4.10 would still not be deemed statistically significant for a sample size of 21 years. However, when the regression equations calibrated for the odd-numbered years were verified for the even years on both the Knoxville and division data, the correlation coefficients were always positive and significant (95% confidence level), indicating that the calibrated relationship was valid, although not strong enough to be significant for one particular small sample (21 odd years) of calibration data. Equations calibrated for the even years provided better calibration results, but did not verify as well over the odd-numbered years. The linear relationship between precipitation and ring-width index was simply more evident in the even years of the 42-year period considered. This difference is attributable to the small size (21 years) of the subsamples.

From Table 1, it is seen that the division data provided better verification statistics than did the Knoxville data even when the Knoxville data were used for calibration. This is true regardless of whether even- or odd-numbered years were used for calibration. From all the results discussed above and summarized in Table 1, we conclude that for exploratory dendroclimatic studies in isolated regions, investigators should consider the use of regionally averaged climatic data. The use of such data is more likely to reveal any dendroclimatic relationships that may be obscured by idiosyncracies inherent in single-station data.

We also used the "leave out one" procedure suggested by Gordon (1980) to generate 42 years of "independent" data. The procedure involved calibrating 42 regression equations, each calibration using 41 years of division data, and using each regression equation to estimate the datum that was left out of the calibration sample. For each calibration, a different year was left out, so a calibration-independent estimate of May-June precipitation was made for each year from 1931 to 1972. These 42 estimates were compared with division data and with Knoxville data. The Knoxville data produced an RE of +0.12 and a correlation coefficient of +0.35, but the division data produced an RE of +0.17 and a correlation coefficient of +0.41. This supports our earlier contention that when division data are calibrated, then division data will produce better verification than will single-station data.

Our general conclusions are supported by another analysis involving three ring-width chronologies from central Iowa. The white oak (*Quercus alba*) chronologies at

Table 2. Dendroclimatic calibration and verification statistics for tree-ring data from central Iowa. Two calibrations were obtained with annual (July-June) precipitation. The first used regionally averaged data for the central Iowa climatic division; the second used single-station data from Des Moines. For each calibration, the verification results for both the division data and the Des Moines data are shown.

Calibration data	**division**		**Des Moines**	
Calibration period	1932-1955	1956-1979	1932-1955	1956-1979
Percentage variance due to regression	69.4	68.3	48.5	52.7
Correlation coefficient	+0.83	+0.83	+0.70	+0.73
Verification — division data	1956-1979	1932-1955	1956-1979	1932-1955
Reduction of error	+0.64	+0.62	+0.62	+0.51
Correlation coefficient	+0.79	+0.80	+0.79	+0.80
Verification — Des Moines data	1956-1979	1932-1955	1956-1979	1932-1955
Reduction of error	+0.47	+0.40	+0.45	+0.29
Correlation coefficient	+0.70	+0.67	+0.70	+0.67

Pammel, Ledges, and Duvick Back Woods (Duvick 1979) were extended and updated for this study. The average ring width of these three chronologies is strongly correlated with annual precipitation, from July 1 of the year preceding ring formation through June 30 of the year the rings were formed. Regionally averaged precipitation data from the central Iowa climatic division and single-station data from Des Moines are both available for the 48-year period 1932-1979. (The year listed corresponds to the latter calendar year of the July-June period.) The 48 years of data were divided in half, this time using the early 24 years for calibration and the latter 24 years for verification. The procedure was then reversed. [There are many ways to divide the data. Some discussion of the possibilities is given by McCarthy (1976).] The results are summarized in Table 2. In all cases the division data led to better calibration and verification statistics than did the Des Moines data, even though Des Moines is centrally located with respect to the tree sites but is in the southwest corner of the climatic division. This supports our earlier conclusions which were based on analysis of data from eastern Tennessee.

REFERENCES

DeWitt, E. and M. Ames, editors
1978 Tree-ring chronologies of eastern North America. *Chronology Series 4,* Laboratory of Tree-Ring Research, The University of Arizona, Tucson.

Duvick, D. N.
1979 An attempt to verify dendroclimatic reconstructions using independent tree-ring chronologies. Masters thesis, Department of Geosciences, The University of Arizona.

Fritts, H. C., G. R. Lofgren, and G. A. Gordon
1979 Variations in climate since 1602 as reconstructed from tree rings. *Quaternary Research* 12: 18—46.

Gordon, G. A.
1980 Verification of climatic reconstructions. Paper prepared for the Proceedings of the Dendroclimatology Workshop, Norwich, England, July 7-11, 1980.

McCarthy, P. J.
1976 The use of balanced half-sample replication in cross-validation studies. *Journal of the American Statistical Association* 71: 596-604.