TREE-RING BULLETIN, *Vol. 54, 1997*

RESEARCH REPORT

COMPUTER ASSISTED, INDEPENDENT OBSERVER VERIFICATION OF TREE-RING MEASUREMENTS

HENRI D. GRISSINO-MAYER

Laboratory of Tree-Ring Research
The University of Arizona
Tucson, Arizona 85721, U.S.A.

ABSTRACT

The importance of tree-ring analyses to forestry and other disciplines (e.g., climate and ecology) requires that tree-ring measurements be as accurate as possible. Accuracy of tree-ring measurement, although often emphasized, may not be stressed as much as other areas of the research. Nonetheless, poor quality measurement data compromise the reliability of interpretations. Possible errors include individual ring mismeasurement errors, consecutive ring errors, multiple ring misdating errors, and multiple ring systematic errors. Verification of measurements can be accomplished by an independent observer who remeasures certain segments from randomly selected cores or cross sections and then uses the computer program VERIFY5 to quantitatively and qualitatively compare both sets of measurements. This program features analyses — such as comparative statistics, least squares analysis, and outlier detection — that can isolate specific measurement errors. Such errors can be minimized by: (1) proper training in the hardware and software used; (2) training in certain rudimentary dendrochronological skills; (3) developing a deliberate measuring pace; (4) consistently using VERIFY5 by an independent observer; (5) using a quality control program (such as COFECHA) to further check measurements and crossdating.

Die Wichtigkeit von Jahrring Messungs-Analysen in der Forstwirtschaft und andere Diszipline (z.B. Klima und Ökologie) fordert dass Jahrring Messungen so genau wie möglich seien sollten. Genauigkeit von Jahrring Messungen, obwohl oft erwähnt, wird aber nicht genug unterstreicht im Verhältnis zu andere Gebiete der Untersuchungs Arbeit. Nichtsdestoweniger, Daten von Messungen schlechter Qualität, beeinflussen die Zuverlässigkeit der Auslegungen. Mögliche Fehler einschliessen eine einzelene individuelle Fehlmessung eines Jahrringes, mehrere auf einander folgende Fehlmessungen von Jahrringe, mehrere datierungs Fehler der Jahrringe, und mehrere systematische Jahrrings fehler. Eine Verifikation der Messungen wird dabei erreicht, dass eine Unabhängige Person zufällige Segmente von Bohrkerne oder Querschnitte nachmisst. Danach wird das Computerprogramm VERIFY 5 benutzt um beide Messungen quantitativ und qualitativ zu Vergleichen. Dieses Programm bietet Analysen an — so wie Komparative Statistik, die das kleinste Quadrat Methode und detektion von ekstrem Wehrte — die spezifische Fehlmessungen isolieren können. Solche Fehler können minimiert werden bei: (1) ordentliches trainieren in dem verwendetem "Hardware" und "Software"; (2) das trainieren von bestimmte grundlegende dendrochronologische Fähigkeiten; (3) die Entwicklung von einen bewussten Rhythmus bei den Messungen; (4) die konsequente Verwendung von VERIFY 5 bei einer unabhängigen Person; (5) das Benutzten von ein Kontroll Programm (so wie COFECHA) um weiter die Messungen und Synchronisierungen nachzuprüfen.

L'importance de la dendrochronologie en foresterie et dans d'autres disciplines (climat, écologie) exige que les mesures des cernes soient aussi précises que possible. La précision des mesures dendrochronologiques, quoique souvent mise en évidence, ne doit pas prendre le pas sur les autres aspects de la recherche. Néanmoins, des mesures de mauvaise qualité compromettent la liabilité des interprétations. Parmi les inexactitudes possibles, on trouve l'erreur portant sur un seul cerne, sur plusieurs cernes consécutifs, la présence ou l'absence d'un cerne. La vérification des mesures peut être réalisée par un observateur qui remesure certains segments d'échantillons prélevés au hasard et utilise le programme VERIFY5 pour comparer quantitativement et qualitativement les deux séries de données. Ces analyses comprennent des statistiques comparatives, l'analyse des moindres carrés et d'autres témoins qui peuvent isoler les erreurs de mesures spécifiques. Ces erreurs peuvent être minimisées par (1) un entraînement approprié du matériel et des programmes utilises; (2) un entraînement basique; (3) maintenir un rythme de mesures régulier; (4) utiliser constamment VERIFY5 par un observateur independant; (5) utiliser un programme de contrôle de qualité (tel que COFECHA) pour de futurs contrôles, mesures et synchronisations.

INTRODUCTION

Dendrochronology and its subdisciplines (e.g., dendroecology and dendroclimatology) play prominent roles in many facets of applied research. Tree-ring data are routinely used to study past climates and climate change (Fritts 1976; Grissino-Mayer 1995; Stahle and Cleaveland 1992), stand productivity (LeBlanc 1990; LeBlanc et al. 1987; Telewski and Lynch 1991), past and present effects of insect defoliators (Brubaker 1978; Swetnam and Lynch 1993; Weber and Schweingruber 1995), and the effects of anthropogenic (e.g., logging and prescribed burning) and natural (e.g., wildfires and windthrow) disturbances on forest stand processes (Grissino-Mayer et al. 1995; Peterson et al. 1994; Swetnam and Dieterich 1985). The implicit assumption in research that makes heavy use of tree-ring data is that all tree-ring widths have been measured as accurately as possible. As more and more laboratories invest heavily in expensive measurement equipment and begin collecting and utilizing large amounts of tree-ring data, this assumption will play an increasingly significant role in the quality of the data generated and the reliability of the interpretations based on the final analyses.

Given the large number of measurements used in standard tree-ring studies, it is inevitable that measurement errors will occur. The effects of such errors, however, are often minimized by replication (Fritts 1976). Increasing the number of radii collected per tree and the number of trees collected per site not only enhances the desired ecological or climatic signal while reducing unwanted noise (e.g., a disturbance that affects only one tree), but also minimizes effects of user-introduced error. Occasionally, however, the effects of measurement errors can be profound, especially when few trees are available for study or when portions of the tree-ring chronology have low sample depths. Therefore, human related measurement errors can effectively reduce the quality of the data and the interpretations based on them, and their occurrence should be minimized.

Measurement errors usually occur due to unfamiliarity with the measurement systems, lack of basic dendrochronological knowledge, and fatigue and/or carelessness. To minimize measurement errors, a verification process should be adopted that involves remeasurement by an independent observer of certain ring segments from randomly selected samples (Fritts 1976). The independent observer then qualitatively and quantitatively compares both series of measurements using specially designed software. The independent observer is usually a trained dendrochronologist familiar with the measurement system and the software used to verify the original measurements. During the verification process, the independent observer *carefully* remeasures selected tree-ring segments to ensure that the verification measurements are error-free. Other methods for measurement verification have no doubt been adopted by various laboratories, but all methods involve some sort of remeasurement and comparison. In this paper, I describe the various types of measurement errors, the likely sources of such errors, and methods and software developed to help identify measurement error. While I describe these error sources on systems that measure from the inside to the bark, they can just as easily occur on systems that measure from the bark to the inside.

TYPES AND SOURCES OF MEASUREMENT ERROR

Type 1: Individual Ring Measurement Error

This type of error is simply a mismeasurement of one individual ring, usually due to unfamiliarity with the hardware used to capture ring measurement data. It is perhaps the most common, the easiest to prevent, and the easiest to identify. Certain measurement systems must be reset to zero millimeters before beginning measurement of a radius. Occasionally, the measurer may forget to reset the system prior to measuring the first ring of a series, resulting in a measurement that is too large (and often unrealistic). This type of error may occur when measuring erratic ring sequences caused by proximity to an injury, branch, or root or by species with nonconcentric ring-growth patterns. All laboratories should have specific guidelines for measuring erratic ring sequences so these errors can be minimized. Occasionally, a "missing" ring of width zero or a "false" ring may be accidentally inserted due to involuntary reflex actions by the measurer. This type of error is usually identified by the computer program COFECHA, because these measurements often fall outside the standard deviation thresholds used by the program (Holmes 1983, 1996). Once such errors are identified, the measurer should always remeasure the errant rings.

Type 2: Consecutive Ring Measurement Error

This type of measurement error occurs across two consecutive tree rings, usually caused by inadvertently measuring past the latewood boundary of the first ring. Errors of this type are common and easily prevented, but may be difficult to identify depending on the severity of the error. Overmeasurement of the first ring causes the next ring to be undermeasured. If the measurer realizes that a ring has been overmeasured, they should immediately correct the problem. Because these measurements usually fall within the standard deviation thresholds used by COFECHA, such errors will seldom be identified by COFECHA.

Type 3: Multiple Ring Misdating Error

This type of error occurs across a longer sequence of rings due to an inaccurate measurement in the first ring of the sequence. Depending on the length over which the misdating occurs, this type of error may or may not be identifiable. For example, the measurer may stop to reposition the sample and inadvertently place the cross hair of the microscope on an incorrect ring. Because the rings are marked every 10 years (Stokes and Smiley 1968; Swetnam et al. 1985), this error should be identified when decade rings are encountered on the measuring stage but are not indicated as being decade rings by the software either as output on the computer screen or as an audible beep. Occasionally, this error may not be recognized until after many decades have been measured. Sometimes, the entire sample is measured and the error never caught. Even worse, the measurer may eventually realize that an error has occurred and simply place the cross hair of the microscope on the next decade ring, not realizing that the error occurred many decades previously. This error may also occur due to carelessness. Occasionally, a nine-year "decade" occurs if the person who dated the sample neglects to mark a missing ring. More commonly, an eleven-year "decade" occurs if a false ring is measured as a true ring. If this error extends over a considerable length (e.g., greater than 20 years), COFECHA is usually successful at identifying such errors.

Type 4: Multiple Ring Systematic Error

This type of error occurs across all rings in a sequence, and is usually caused by individuals measuring faster than the system can capture data. Practically all systems in existence today, especially older systems attached to slower personal computers, have limitations on the speed with which they capture data. These systems are complex, linking a movable stage micrometer to either a linear or rotary encoder, then to a digital display, and eventually to the personal computer system via either a game port or serial port. These measurements are then captured by the software on the computer and eventually recorded on either a floppy diskette or hard disk. Because of this complex linkage, there is often a delay in the system's capturing each measurement. If measuring proceeds too rapidly, the measurements may be systematically truncated because the hardware and software cannot keep pace. This type of error can be minimized by emphasizing the importance of developing a slow and deliberate measuring pace, coming to a complete stop at each ring boundary before entering the measurement, and delaying slightly before measuring the succeeding ring to allow the system to reset (if necessary).

COMPUTER ASSISTED VERIFICATION

Verification of ring measurements is especially crucial when training a person to measure. It is recommended that the first ten series the person measures be remeasured by an independent observer for comparison. If the samples are short (e.g., fewer than 100 years), the independent observer may wish to increase the number of series initially reexamined. Errors should be pointed out and thoroughly explained. Once the measurer is trained, the independent observer occasionally should randomly select a certain number or percentage of samples (e.g., one of every ten), then remeasure a segment of each selected sample. Depending on the number of samples to be measured, as well as the length of the series being measured, the observer can adjust the number of samples selected and the lengths of the segments to be remeasured. For example, rather than remeasure a continuous 50-year segment from 10 samples that average 250 years in length, the observer may randomly select 20-year segments from each century for each sample. The independent observer should also be aware that accuracy may drop towards the end of the measurement process, especially for longer sequences. This is due to the "expectancy factor" — as the measurer nears the end of a particularly long series, measurement speed is likely to increase, causing accuracy to decrease. These portions should be targeted for verification. In addition, the independent observer may wish to select samples measured over a longer duration rather than selecting samples that were all measured in one day. This is necessary because measurement accuracy of individuals can vary from day to day, depending on physical and emotional factors.

The independent observer can then compare the original measurements for each series against the verification measurement data using the computer program VERIFY5. Developed for IBM-compatible computers, the program reads in both sets of measurements for the series being verified (the original set and the verification set made by the independent observer) then conducts four sets of analyses over the period common to both data sets. Each analysis reveals information on specific types of errors to help investigators assess the level of measurement accuracy.

Statistical Comparison

The first analysis compares descriptive statistics between both sets of measurements (Table 1). The assumption in this analysis is that the descriptive statistics should be equivalent, or nearly so, because the same ring-width measurements are contained (or supposed to be contained) in each data set. Because the data sets are not independent, no formal statistical tests (e.g., Student's t tests) can be applied. This analysis is therefore strictly comparative, but informative. Because one or two errors in individual ring measurements rarely cause great differences in statistical descriptors, this analysis is more useful for isolating more serious errors. For example, if a multiple ring systematic error has occurred, measures of central tendency (i.e., mean and median) will be lower in the original measurement file.

Table 1. Descriptive statistics for two actual measurement files indicating excellent agreement between both sets of measurements.

Attribute	San Juan Site Measurement File	San Juan Site Remeasurement File
Average width	1.211	1.215
Median	1.005	0.995
Variance	0.572	0.576
Standard deviation	0.756	0.759
Coeff of variation	0.625	0.625
Relative skewness	1.823	1.819
Pearson's Gamma	0.816	0.868
Kurtosis	7.613	7.470
Mean sensitivity	0.357	0.351
Autocorrelation	0.743	0.747

Least Squares Regression

The second analysis is an ordinary least squares regression in which the original set of measurements is treated as the predictand, and the verification set of measurements is treated as the predictor (Table 2). The assumption in this analysis is that the regression should indicate nearly perfect agreement because both sets of measurements should be equivalent. The slope of the regression should be near one, the y-intercept near zero, the correlation coefficient (r) and the coefficient of determination (r^2) near one, and the root mean square error near zero. If a y-intercept deviates from zero, this value can be interpreted as the average amount of systematic error introduced by the measurer. For example, a y-intercept of 0.12 indicates that the original measurements were truncated by an average of 0.12 units of measurement, indicating the person was measuring too rapidly for the system used. A correlation coefficient (and its related r^2) much less than one could indicate a possible multiple ring misdating error.

Table 2. Results of a least squares regression, with the original measurements as the dependent variable (y) and the verification measurements as the independent variable (x), indicating major errors in the original file. The y-intercept is non-zero, the root mean square error is large, and the correlation coefficient and coefficient of determination deviate from the desired value of ≈ 1.0.

Attribute	Value
Slope of regression	1.020
Intercept of regression	0.090
Root mean square error	0.857
Coefficient of variation	64.461
Correlation coefficient (r)	0.672
Coeff of determination (r^2)	0.452

Verification

The third analysis (Table 3) is a more formal determination of verification based on methods described by Fritts (1976:250-252). In this analysis, the absolute difference between yearly measurements from both data sets is calculated, then squared. These values are summed over all observations, yielding the sum of differences and the sum of squared differences. The average difference is obtained by dividing the sum of differences by the number of observations. This value should be very near zero if the original set of measurements is accurate. A larger value indicates the average amount of introduced error per measurement and roughly approximates the intercept of the regression line in the previous analysis. The average squared difference is obtained by dividing the sum of squared differences by the number of observations. In data sets with minimum measurement error, the average squared difference normally approximates zero (i.e., is less than .01). This value is the threshold that determines verification accuracy. Two levels of acceptance are reported, the .05 and .01 levels. The .05 level should be used for species with wavy or sinuous ring boundaries (e.g., ring porous species such as *Quercus*), while the .01 level should be used for species with nearly linear ring boundaries (e.g., conifers and diffuse porous species such as *Populus*). These levels are not based on formally defined distributions and therefore should not be confused with confidence levels.

These acceptance levels were determined experimentally (Fritts 1976), and may be too liberal for more accurate measurement systems (e.g., the new systems now being used for image analysis of tree rings). They are merely suggestions to provide some standards for acceptable levels of verification. Acceptance at the .01 level would occur if all ring measurements were inaccurate on average by as much as .1 units of measurement. A more appropriate level of acceptance, especially for conifer tree rings, may be the .005 level (tolerance of inaccuracy per measurement equal to .07 units of measurement) or the .0025 level (tolerance of inaccuracy per measurement equal to .05 units of measurement).

Table 3. Results of a simulated verification. The large average squared difference indicates significant errors in the original data set, and the analysis rejects the hypothesis that measurements in both data sets are equivalent.

Attribute	Value
Sum of differences	18.970
Sum of squared differences	76.272
Average difference	0.182
Average squared difference	0.733
Acceptance level: 0.05 REJECT	Acceptance level: 0.01 REJECT

Detection of Outliers

This analysis is useful for detecting all four types of measurement errors. Hence, this portion of the output should be inspected even if the measurements are acceptable according to the verification results. In this analysis, the results of the previous linear regression are tabulated (Table 4). Predicted values are generated during the regression as a function of the verification values (made by the independent observer), then residuals are calculated by subtracting the predicted value from the original, initial value. Residuals are then standardized by dividing each by the root mean square error (Table 2) obtained from the linear regression. These standardized residuals approximate Student's t-values where values that fall in the range $|t| \geq 2.0$ are rare. In VERIFY5, rings are flagged (and should therefore be reinspected) when the difference between measurements for any particular year is greater than 2.5 times the average difference. These rings roughly correspond to years when $|t| \geq 2.0$. This analysis always identifies outliers, even when the measurements are accurate, because all distributions have values that occupy the "tail" regions (e.g., outside the .025 level). The independent observer should inspect all flagged measurements to ensure that the errors fall within tolerance levels and, if deemed necessary, remeasure the flagged rings.

IDENTIFYING MEASUREMENT ERRORS

Individual and Consecutive Ring Measurement Error

Individual ring errors during any particular year are indicated by large differences and residuals between both measurements and standardized residuals outside the range $|t| \geq 2.0$. In Table 4, the years 1890 and 1918 are flagged by VERIFY5 (asterisks) because the original measurements for these years are statistical outliers. Consecutive ring measurement errors are indicated when a large positive difference for a year is followed by a negative difference of similar magnitude. In Table 4, the measurer overmeasured the ring for 1905 by 1.03 millimeters, and subsequently undermeasured the ring for 1906 by 1.05 millimeters. The measurement for 1911 was similarly overmeasured by 0.73, while the measurement for year 1912 was undermeasured by 0.69. These large measurement errors may mask the detection of smaller measurement errors, such as those indicated for the years 1898 and 1915 (Table 4). When substantial errors occur, the independent observer should reanalyze the same series once the corrections have been made to determine whether these smaller measurements errors are within acceptable tolerance levels.

Multiple Ring Misdating Error

Because this error results from mismeasurement of just one or (rarely) two rings, descriptive statistics usually do not reveal that a possible error has occurred. The first indication that a multiple ring misdating error has occurred is evident in the least squares regression analysis. The regression will show that the y-intercept deviates from zero (Table 2) because a partial subset of the observations is mismatched due to the measurement error. Correlation coefficients will be below 1.0. Results of the verification tests may reject the assumption that the

Table 4. Identification (*) of outliers by VERIFY5. In this example, the measurer did not reset the system to zero before measuring the years 1890 and 1918, resulting in unrealistic values. Consecutive ring errors are indicated for the years 1905-06 and 1911-12.

Year	Original Value	Verify Value	Diff	Predict Value	Resid	Stand Resid
*1890	8.14	1.68	6.46	1.80	6.34	7.39
1891	2.33	2.29	0.04	2.43	-0.10	-0.11
1892	1.68	1.73	-0.05	1.85	-0.17	-0.20
1893	0.69	0.66	0.03	0.76	-0.07	-0.09
1894	0.99	0.99	0.00	1.10	-0.11	-0.13
1895	1.97	1.97	0.00	2.10	-0.13	-0.15
1896	0.97	1.00	-0.03	1.11	-0.14	-0.16
1897	1.64	1.66	-0.02	1.78	-0.14	-0.17
1898	1.84	1.77	0.07	1.90	-0.06	-0.06
1899	0.70	0.67	0.03	0.77	-0.07	-0.09
1900	0.95	0.99	-0.04	1.10	-0.15	-0.18
1901	1.60	1.57	0.03	1.69	-0.09	-0.11
1902	0.83	0.87	-0.04	0.98	-0.15	-0.17
1903	2.15	2.14	0.01	2.27	-0.12	-0.14
1904	1.90	1.86	0.04	1.99	-0.09	-0.10
*1905	3.84	2.81	1.03	2.96	0.88	1.03
*1906	1.85	2.90	-1.05	3.05	-1.20	-1.40
1907	4.74	4.66	0.08	4.84	-0.10	-0.12
1908	0.43	0.43	0.00	0.53	-0.10	-0.12
1909	1.02	1.03	-0.01	1.14	-0.12	-0.14
1910	1.12	1.13	-0.01	1.24	-0.12	-0.14
*1911	2.03	1.30	0.73	1.42	0.61	0.72
*1912	0.71	1.40	-0.69	1.52	-0.81	-0.94
1913	0.97	0.96	0.01	1.07	-0.10	-0.12
1914	1.27	1.25	0.02	1.37	-0.10	-0.11
1915	0.68	0.80	-0.12	0.91	-0.23	-0.26
1916	0.67	0.71	-0.04	0.81	-0.14	-0.17
1917	1.06	1.04	0.02	1.15	-0.09	-0.11
*1918	6.49	0.91	5.58	1.02	5.47	6.39
1919	0.62	0.59	0.03	0.69	-0.07	-0.08

two data sets are similar (Table 3). The detection of outliers in the fourth analysis visually confirms the measurement error. In Table 5, all measurements fall within acceptable tolerance levels until the year 1955, after which large differences, residuals, and standardized residuals appear. This indicates a misdating error during measurement. Reevaluation of the 1955 ring reveals that the measurer inadvertently measured two rings, 1955 (0.45 mm) and 1956 (0.52 mm), as one ring (0.96 mm). All subsequent measurements are "shifted" down one year in the third column; for example, the 1959 measurement in Column 2 appears as the 1960 measurement in Column 3.

Table 5. VERIFY5 comparison of original and verification measurements indicating a multiple ring misdating error. The small differences between measurements prior to 1955 and the large values for all statistics beginning with the 1955 ring indicate that the 1955 and 1956 rings were inadvertently measured as one ring.

Year	Original Value	Verify Value	Diff	Predict Value	Resid	Stand Resid
1940	1.16	1.15	0.01	1.16	0.00	0.02
1941	2.19	2.24	-0.05	2.22	-0.03	-0.16
1942	1.54	1.50	0.04	1.50	0.04	0.24
1943	1.55	1.57	-0.02	1.57	-0.02	-0.09
1944	1.27	1.21	0.06	1.22	0.05	0.31
1945	1.21	1.26	-0.05	1.26	-0.05	-0.32
1946	1.10	1.09	0.01	1.10	0.00	0.01
1947	1.20	1.25	-0.05	1.25	-0.05	-0.32
1948	1.35	1.23	0.12	1.24	0.11	0.67
1949	1.40	1.39	0.01	1.39	0.01	0.05
1950	0.81	0.78	0.03	0.80	0.01	0.07
1951	0.82	0.83	-0.01	0.85	-0.03	-0.15
1952	1.19	1.19	0.00	1.20	-0.01	-0.04
1953	0.50	0.47	0.03	0.50	0.00	0.02
1954	0.78	0.78	0.00	0.80	-0.02	-0.10
*1955	**0.96**	**0.45**	**0.51**	**0.48**	**0.48**	**2.80**
1956	0.56	0.52	0.04	0.54	0.02	0.09
1957	0.41	0.55	-0.14	0.57	-0.16	-0.95
1958	0.44	0.44	0.00	0.47	-0.03	-0.16
1959	0.66	0.44	0.22	0.47	0.19	1.12
1960	0.47	0.65	-0.18	0.67	-0.20	-1.17
1961	0.61	0.51	0.10	0.53	0.08	0.44
1962	0.44	0.61	-0.17	0.63	-0.19	-1.11
*1963	0.91	0.43	0.48	0.46	0.45	2.62
1964	1.08	1.03	0.05	1.04	0.04	0.23
1965	1.37	1.13	0.24	1.14	0.23	1.34
1966	1.40	1.30	0.10	1.30	0.10	0.56
*1967	0.97	1.40	-0.43	1.40	-0.43	-2.49
*1968	1.25	0.96	0.29	0.97	0.28	1.61
*1969	0.74	1.25	-0.51	1.25	-0.51	-2.98
1970	0.78	0.80	-0.02	0.82	-0.04	-0.21

Multiple Ring Systematic Error

The first indication that this type of error has occurred is usually found in the differences between the two means. For example, the mean of the original measurements may be 1.112 mm, while the mean for the verification measurements could be 1.215 mm. This type of error usually produces the worst results in the linear regression analysis (Table 2) and usually features the lowest correlations of all types of errors. Without fail, these measurements will not verify as shown in the third analysis (Table 3). The outlier analysis visually confirms that this type of error has occurred (Table 6). Because this error is systematic across most or all rings, few rings will be flagged by VERIFY5 as being errant. However, closer inspection reveals that the differences between both sets of measurements are usually negative (Table 6, Column 4), indicating that the rings were systematically undermeasured.

Table 6. VERIFY5 comparison of original and verification measurements indicating a multiple ring systematic error. The preponderance of negative differences indicates truncated measurements resulting from measuring too fast for the measuring system.

Year	Original Value	Verify Value	Diff	Predict Value	Resid	Stand Resid
1905	2.78	2.81	**-0.03**	2.69	0.09	0.82
1906	2.93	2.90	**0.03**	2.78	0.15	1.38
1907	4.53	4.66	**-0.13**	4.51	0.02	0.17
1908	3.98	4.06	**-0.08**	3.92	0.06	0.54
1909	2.76	2.79	**-0.03**	2.67	0.09	0.81
*1910	1.79	2.10	**-0.31**	1.99	-0.20	-1.87
1911	2.81	2.88	**-0.07**	2.76	0.05	0.46
1912	1.44	1.64	**-0.20**	1.54	-0.10	-0.93
1913	0.67	0.83	**-0.16**	0.74	-0.07	-0.68
*1914	1.29	1.66	**-0.37**	1.56	-0.27	-2.49
1915	1.47	1.48	**-0.01**	1.38	0.09	0.80
1916	2.42	2.57	**-0.15**	2.46	-0.04	-0.33
1917	2.48	2.55	**-0.07**	2.44	0.04	0.41
*1918	1.17	1.59	**-0.42**	1.49	-0.32	-2.96
1919	1.99	1.86	**0.13**	1.76	0.23	2.14
1920	1.78	1.96	**-0.18**	1.86	-0.08	-0.69
1921	1.83	2.10	**-0.27**	1.99	-0.16	-1.50
1922	0.98	1.11	**-0.13**	1.02	-0.04	-0.36
1923	0.74	0.77	**-0.03**	0.68	0.06	0.51
1924	0.76	0.78	**-0.02**	0.69	0.07	0.60
1925	0.73	0.85	**-0.12**	0.76	-0.03	-0.31
1926	0.77	0.94	**-0.17**	0.85	-0.08	-0.75
1927	1.40	1.58	**-0.18**	1.48	-0.08	-0.75
1928	0.70	0.74	**-0.04**	0.66	0.04	0.41
1929	0.36	0.34	**0.02**	0.26	0.10	0.90
1930	0.29	0.36	**-0.07**	0.28	0.01	0.08
1931	0.63	0.83	**-0.20**	0.74	-0.11	-1.05
1932	0.35	0.44	**-0.09**	0.36	-0.01	-0.09
1933	0.54	0.65	**-0.11**	0.57	-0.03	-0.25
1934	0.03	0.00	**0.03**	-0.07	0.10	0.94

PREVENTING MEASUREMENT ERROR

Measurement errors will always occur. However, certain guidelines should be emphasized to help minimize the possibility of measurement error.

1. *Stress the importance of established guidelines.* All laboratories should have a manual for measuring tree rings, and this manual should be easily accessible in the measuring area. The manual should cover how physical breaks in tree rings should be treated, how rings with erratic boundaries should be measured, and the meaning of the marks on the rings.
2. *Teach the basics of measuring.* Take the time to efficiently train individuals in using the measurement hardware and software. Measurement systems (and tree-ring analysis in general) may seem particularly strange to new personnel.
3. *Teach the basics of tree-ring analysis.* Measurers should know some basics of tree-ring analysis (e.g., how tree rings are crossdated) and the goals of the research project. This will give them a greater appreciation for measurement accuracy and let them know that they are an integral part of the research.
4. *Emphasize a deliberate pace.* Teach the individual to develop and maintain a deliberate pace during measurement. The pace should not be too slow to "stall" the research or too fast to lessen its quality. The entire measurement process should "flow" as the individual becomes more comfortable with the hardware and software.
5. *Emphasize fatigue avoidance.* Fatigue is perhaps responsible for most measurement errors. The measurement environment should be ergonomically suitable for repetitious work. The chair should be comfortable and of the correct height to see into the microscope without strain. Instruct the individual to take short breaks, perhaps after each series or every 100 rings is measured.
6. *Teach the "language" of tree-ring dating.* The dater should provide, either on the side of the core mount or attached on a note, instructions for measuring particularly troublesome segments of samples (e.g., severely suppressed growth areas). It should not be assumed that the measurer understands the complex notation that accompanies such difficult ring segments. Standard abbreviations should be used. For example, the notation "1580a, 1581m, 1585m, 1586m, 1587a, 1592a, 1593m" could inform the measurer which rings are absent (a) and which are microrings (m).
7. *Emphasize the consequences of inaccuracy.* Teach the measurer the types of errors that may occur if care is not exercised during measurement. Such errors include not resetting the system, inadvertently moving samples on the measuring stage, and measuring too rapidly. Show them the results of the verification tests and point out possible trouble spots *before* they become serious.
8. *Emphasize communication.* The measurer should flag obvious dating errors (e.g., a nine ring "decade" in which a missing ring was not marked) for checking by the dater or the supervisor.

Such checks by an independent observer should be routine when measurers are first being trained, but cannot continue indefinitely. Eventually, independent verification must be terminated assuming the measurer has been sufficiently trained. All tree-ring measurements then should be processed through a quality control program such as COFECHA. COFECHA not only helps ensure the quality of the crossdating, it also analyzes all measurements to detect possible measurement errors. COFECHA flags any rings with measurements 3.0 standard deviations larger or 4.5 standard deviations smaller than the mean of all rings for that particular year (Holmes 1996). Such rings should *always* be remeasured to ensure their accuracy.

ACKNOWLEDGMENTS

This research was partially funded by a grant to the International Tree-Ring Data Bank from the Paleoclimatology Program of the Office of Global Programs, National Oceanic and Atmospheric Administration. I especially thank Rex Adams for his advice and suggestions, which have ensured that measurements made by Laboratory of Tree-Ring Research personnel are as accurate as possible, and for critiquing early drafts of this manuscript. This paper also benefited from discussions with Paul Sheppard. Suggestions made by the three anonymous reviewers greatly improved this manuscript. The program VERIFY5, as well as all programs in the International Tree-Ring Data Bank Program Library Version 2.0, can be obtained free of charge from the author.

REFERENCES CITED

Brubaker, L.B.
1978 Effects of defoliation by Douglas-fir tussock moth on ring sequences of Douglas-fir and grand fir. *Tree-Ring Bulletin* 38:49-60.

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

Grissino-Mayer, H.D.
1995 *Tree-Ring Reconstructions of Climate and Fire History at El Malpais National Monument, New Mexico*. Ph.D. dissertation, The University of Arizona, Tucson. University Microfilms International, Ann Arbor.

Grissino-Mayer, H.D., C.H. Baisan, and T.W. Swetnam
1995 Fire history in the Pinaleño Mountains of southeastern Arizona: effects of human-related disturbances. In "Biodiversity and Management of the Madrean Archipelago: The Sky Islands of Southwestern United States and Northwestern Mexico," edited by L.F. Debano, G.J. Gottfried, R.H. Hamre, C.B. Edminster, P.F. Ffolliott, and A. Ortega-Rubio, pp. 399-407. *General Technical Report* RM-GTR-264. USDA Forest Service, Fort Collins.

Holmes, R.L.
1983 Computer-assisted quality control in tree-ring dating and measurement. *Tree-Ring Bulletin* 43:69-78.
1996 The Dendrochronology Program Library. In *The International Tree-Ring Data Bank Version 2.0 User's Manual*, edited by H.D. Grissino-Mayer, R.L. Holmes, and H.C. Fritts, pp. 40-74. Laboratory of Tree-Ring Research, The University of Arizona, Tucson.

LeBlanc, D.C.
1990 Red spruce decline on Whiteface Mountain, New York. I. Relationships with elevation, tree age, and competition. *Canadian Journal of Forest Research* 20:1408-1414.

LeBlanc, D.C., D.J. Raynal, and E.H. White
1987 Acidic deposition and tree growth: I. The use of stem analysis to study historical growth patterns. *Journal of Environmental Quality* 16:325-333.

Peterson, D.L., S.S. Sackett, L.J. Robinson, and S.M. Haase
1994 The effects of repeated prescribed burning on *Pinus ponderosa* growth. *International Journal of Wildland Fire* 4:239-247.

Stahle, D.W., and M.K. Cleaveland
1992 Reconstruction and analysis of spring rainfall over the southeastern U.S. for the past 1000 years. *Bulletin of the American Meteorological Society* 73:1947-1961.

Stokes, M.A., and T.L. Smiley
1968 *An Introduction to Tree Ring Dating*. University of Chicago Press, Chicago. Reprinted 1996, The University of Arizona Press, Tucson.

Swetnam, T.W., and J. Dieterich
1985 Fire history of ponderosa pine forests in the Gila Wilderness, New Mexico. In "Proceedings — Symposium and Workshop on Wilderness Fire, Missoula, Montana, November 15-18, 1983," edited by J.E. Lotan, B.M. Kilgore, W.C. Fischer, and R.W. Mutch, pp. 390-397. *General Technical Report* INT-182. USDA Forest Service, Ogden.

Swetnam, T.W., and A.M. Lynch
1993 Multicentury, regional-scale patterns of western spruce budworm outbreaks. *Ecological Monographs* 63:399-424.

Swetnam, T.W., M.A. Thompson, and E. K. Sutherland
1985 Using dendrochronology to measure radial growth of defoliated trees. *Agricultural Handbook* 639, USDA Forest Service.

Telewski, F.W., and A.M. Lynch
1991 Measuring growth and development of stems. In *Techniques and Approaches in Forest Tree Ecophysiology*, edited by J.P. Lassoie and T.M. Hinckley, pp. 503-555. CRC Press, Boca Raton.

Weber, U., and F.H. Schweingruber
1995 A dendroecological reconstruction of western spruce budworm outbreaks (*Choristoneura occidentalis*) in the Front Range, Colorado, from 1720 to 1986. *Trees* 9:204-213.