GEOCHRONOMETRIA Vol. 23, pp 39-43, 2004 – *Journal on Methods and Applications of Absolute Chronology*

NEW METHODS OF DATING IN ARCHAEOLOGY

VALENTIN DERGACHEV and SERGEY VASILIEV

Ioffe Physico-Technical Institute, Russian Academy of Sciences
Politekhnicheskaya 26, St.Petersburg 194021, Russia
(e-mail: v.dergachev@pop.ioffe.rssi.ru)

Key words: ARCHAEOLOGICAL WOOD SAMPLES, CROSS-DATING, SCYTHIAN EPOCH, REGRESSION ANALYSIS, MULTIPLE CORRELATION

Abstract: A modification of statistical method to determine the relative age of wood samples based on the analysis of dendroseries has been developed. The main difference of the method proposed here from the one used earlier lies in the fact that there is the possibility to match several samples simultaneously. Using the methods of statistical analysis, we have the possibility to ascertain what is reliability of determination of the relative age. Approaches used in our method are typical for regression analysis. The correlation coefficient between the analyzed sample and the group of matching samples (the comparison groups) serve as a measure concordance. It is so-called multiple correlation coefficient. The Fisher F-statistics is used to determine the significance of matching. The method is applied for cross-dating of wood samples of archaeological monuments in the Southern Siberia.

1. INTRODUCTION

Some archeological monuments, containing wooden fragments, are investigated in Siberia (e.g., Alekseev *et al*., 2002). As a rule, these fragments are logs, whose conditions are determined by the monument age, local climate and soil nature. There are samples in a good state (e.g., Pasyryk) and quite eroded ones. The main and the most reliable dating method of archeological monuments is radiocarbon dating. However there might be monuments containing constructions of a close age, e.g., groups of the burials of historically short period. In this case it is reasonable to perform radiocarbon dating of a construction containing the best preserve samples or a construction containing the most wealth archaeological material. The age of other constructions in respect to the main construction is possible to be inferred through comparison of dendroseries which would contain the information regarding ring widths, such as, e.g., growth indices. The main problem in determining the relative age of wooden samples in archeology is their poor state.

We have performed a series of investigations, where the methods of determination of the relative age of samples, making use of the dendroseries comparison, have been proposed (Dergachev and Vasiliev, 2000; Dergachev *et al*., 2001; Vasiliev *et al*., 2001). The method, making possible to match the relative age of several samples simultaneously, is presented in this work.

The methods of statistical analysis allow to estimate what is the reliability of the relative age finding. The methods used in our approach are related to regression analysis. The coefficient correlation between the considered sample and the comparison group serves as a measure of concordance. It is a so-called multiple correlation coefficient. To determine the relative position of the analyzed sample on the time scale we find the value of the relative age shift, which result in maximal value of multiple correlation coefficient.

The Fisher F-statistics is used to determine the significance of correlation. The method proposed allows us to improve the reliability in determination of the relative age. The samples from the barrow Bystrovka (Altay) are used in the present paper to demonstrate our technique.

2. PRELIMINARY ANALYSIS OF THE SAMPLES

Dendroseries, obtained as a result of measurements of ring widths of wooden samples, usually contain long-term trends of a complicated shape. The nature for this trend lies in both the processes of tree development and climate changes. Traditionally methods, used to compare the samples, implicitly suggest the absence of any long-term trends. Because of this reason the growth indices, obtained as a running mean, or after some filtration, are usually used. Our technique is similar. Firstly, we remove a long-term trend from the data of measurements of ring widths of wooden samples. At the same time we plot the trend line through points spaced at 10 years intervals. Vertical coordinates of such points are determined by the method of the least squares. Thereafter we remove this trend and obtain the standardized remainder of the series used and we work with this data.

As a case in point, the data for the sample B-2001 along with the trend line (smoothed curve) are shown in **Fig. 1.** The remainder of the series of data (B-2001) are presented in **Fig. 2.** Step choice in finding of the trend defines the frequency band that would be excluded from the considered standardized remainder. If 10-year step is selected then cycles with periods >20 years are removed. Data, obtained after a long-term trend has been removed, contain all the features necessary for further analysis.

It is especially important not to introduce any distortions complicating a further analysis when the filtration procedure was performed. The main characteristics determining the quality of filtration are amplitude-frequency and phase-frequency characteristics of the filter (see, e.g., **Figs 3** and **4**). From examination of **Fig. 3** follows that the amplitude of harmonics with periods > 20 years (to the left from the vertical line, the trend area) is essentially decreased. **Fig. 4** allows us to conclude that for periods < 20 years (to the right from the vertical line, the region of filtered data) any phase distortions are actually absent.

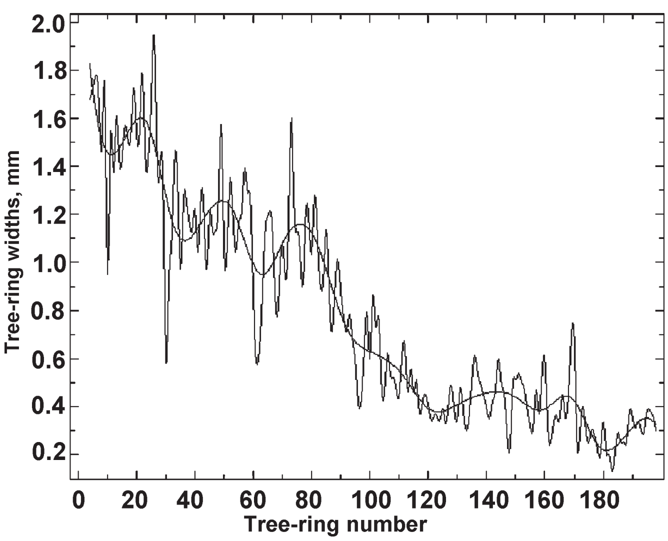

Fig. 1. *Annual ring widths for the sample B-2001 (Bystrovka barrow, Altay). Long-term trend is shown by smoothed curve.*

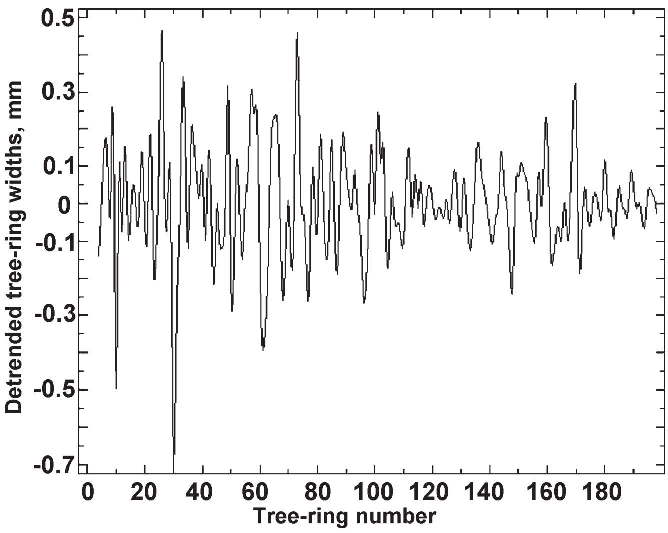

Fig. 2. *Standardized B-2001 remainder after the trend has been removed.*

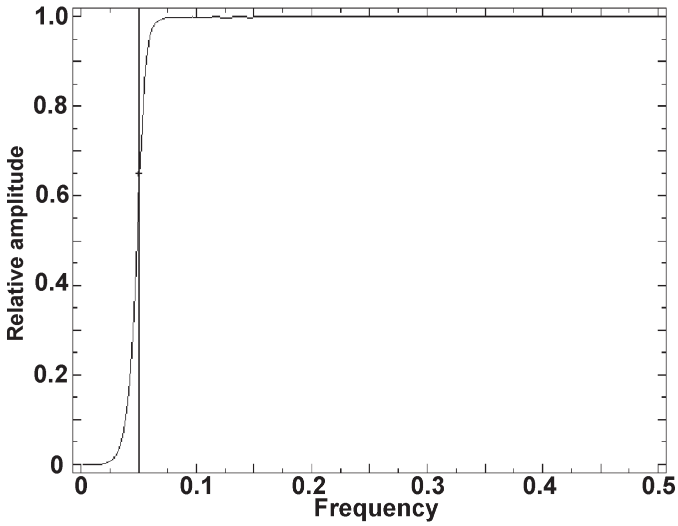

Fig. 3. *Amplitude-frequency characteristic of filter used. Harmonics are sufficiently suppressed to the left from the vertical line (periods > 20 years, trend region).*

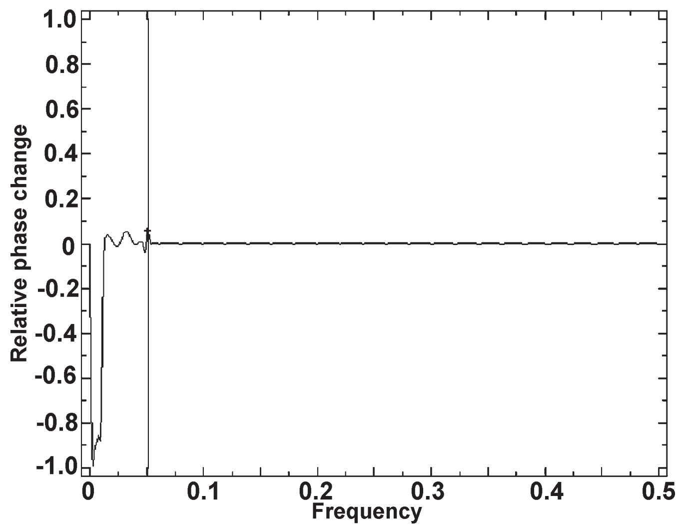

Fig. 4. *Phase-frequency characteristics of filter used. Phase distortions are absent to the right from the vertical line (periods < 20 years, the region of filtered data).*

3. CORRELATION OF SERIES

To determine a relative age of samples the correlation coefficient is used (**eq. 6.1**)). To compare two dendroseries a correlation coefficient is calculated at their different relative position (different shift). The relative location of series when their correlation coefficient is maximal, determines the relative age of samples (**Fig. 5**). The locations of outer rings are compared. A shift of relative location of the outer ring to the right on the diagram corresponds to a right shift on the time scale. Thus, in accordance with **Fig. 5**, the sample B-2003 is 5 years younger than B-2001. Note that such little difference in age is hardly to be obtained in the framework of radiocarbon method.

4. EXAMINATION OF THE CORRELATION COEFFICIENT SIGNIFICANCE

F-statistics is used to examine a significance of the correlation coefficient (**eq. 6.2**). At given the length **n** of compared series and the obtained value of the correlation coefficient $\mathbf{r_{xy}}$ this statistics allows one to conclude, whether the obtained value of $\mathbf{r_{xy}}$ is not equal to zero. It is more convenient to use the criterion (**6.2**) in our investigation, considering it from other point of view. Namely, how is the minimal value of the correlation coefficient $\mathbf{r_{xy}}$ at the known length of series **n**? The matter is that series are shifted relative to each other, when they are compared. As a result of it the number of overlapped rings **n** changes. Considering $\mathbf{r_{xy}}$ as a function of parameter **n** from (**6.2**) one can obtain an expression for values of the correlation coefficient corresponding to the 0.95 confidence coefficient (**eq. 6.3**). A usage of this equation is demonstrated in **Fig. 6**. When B-2005 series is shifted by 45 years relative B-2001 series, the values of correlation coefficient are significant.

5. MULTIPLE CORRELATION

Correlation is not always significant when two series are compared. We can demonstrate this with a simple example of comparison of two series. The results of comparison of two series (B-2004 and B-2010) are presented in **Fig. 7**. It is seen that their pair correlation is not significant. In subsequent study of the relative age of the B-2004 sample we have added the new series B-2001 into the comparison group and have calculated the multiple correlation coefficient. Main definitions are given in section 6 (**eqs 6.4 - 6.8**). The results of our analysis of these two series are presented in **Fig. 8**. It is shown that the multiple correlation between the B-2004 sample and the comparison group (B-2010, B-2001) is significant. Therefore to improve the reliability of estimates of the relative age of samples one should use a multiple comparison, when multiple correlation coefficient is calculated, not pair correlation coefficient. The results of our analysis of the Bystrovka barrow are given in general in **Table 1** and in **Fig. 9**.

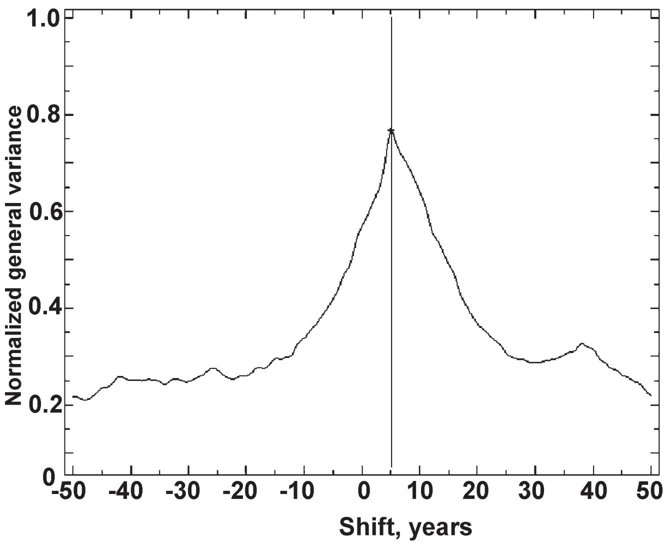

Fig. 5. *Correlation between two series as a function of their relative position. The maximum on graph corresponds to the largest coefficient correlation.*

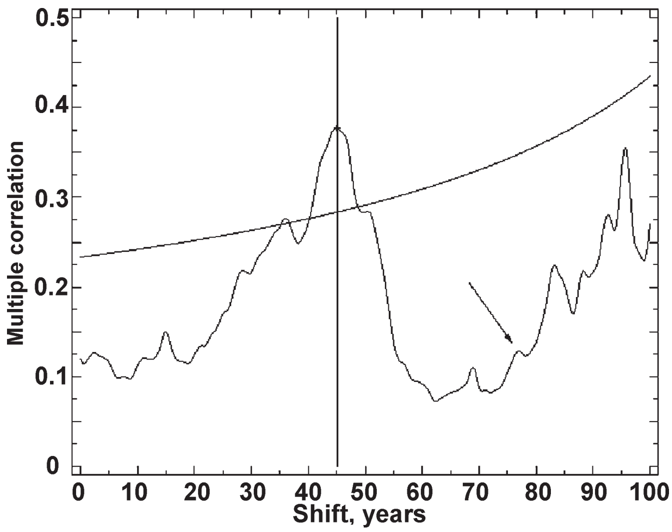

Fig. 6. *Significance of correlation (see section 6, Eq. (6.3)) in the study of relative age of the samples B-2005 and B-2001. Monotonic line corresponds to 0.95 confidence coefficient.*

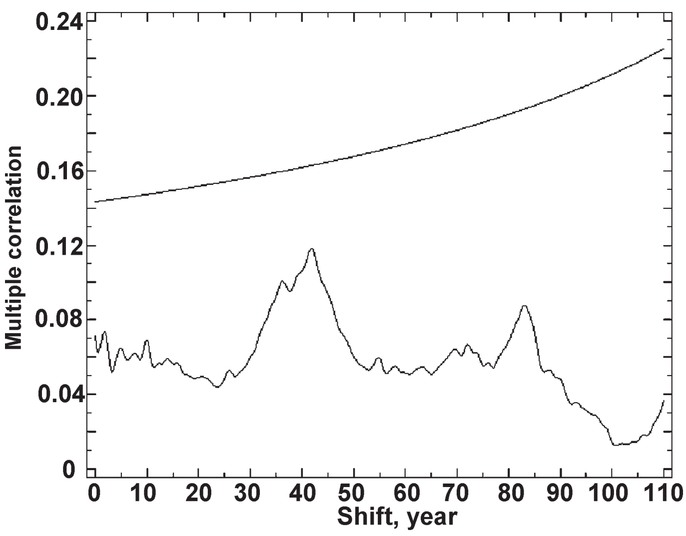

Fig. 7. *Pair correlation between the series B-2004 and B-2010. The correlation is not significant.*

Table 1. *Comparison groups used in analysis of the Bystrovka barrow*

	COMPARISON GROUPS						
Sample	**2001**	**2003**	**2004**	**2005**	**2008**	**2009**	**2010**
2003	+						
2004	+						+
2005	+	+					+
2008		+					+
2009	+						+
2010	+	+	+	+			

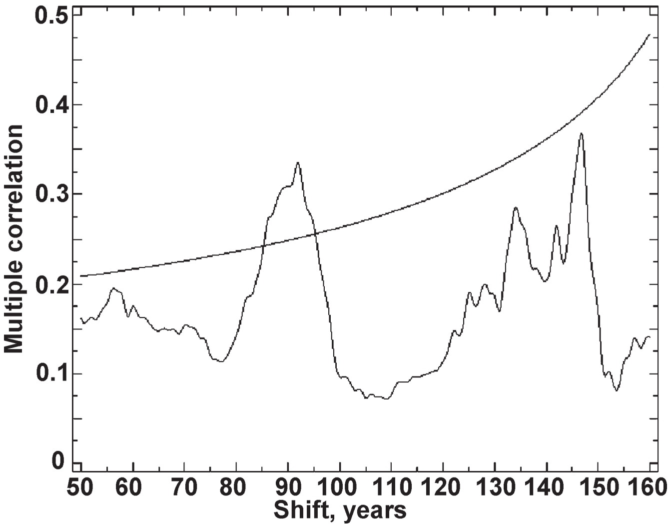

Fig. 8. *Multiple correlation. Series B-2004 is compared with series B-2010, B-2001. Correlation is significant if the series B-2004 and B-2001 are shifted about 90 years.*

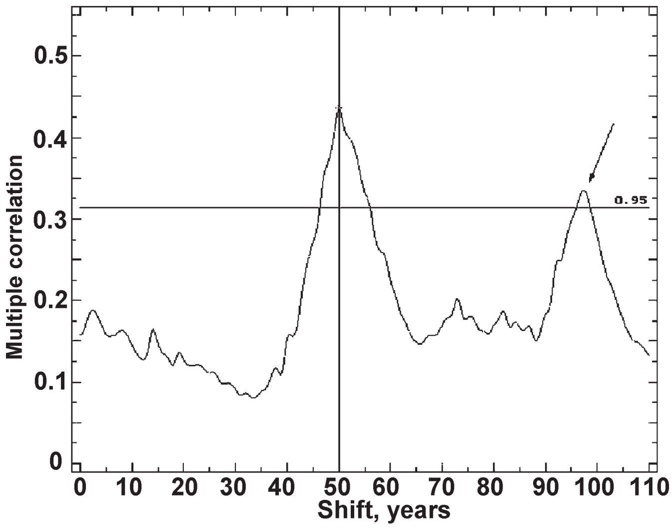

Fig. 10. *Results of deriving of the relative age of B-2010. Two maxima are significant.*

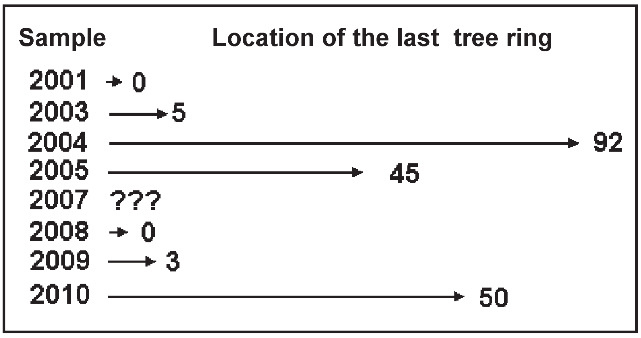

Fig. 9. *The results of Bystrovka barrow dating. Right-band numbers correspond to relative age of samples.*

Comparison groups for six considered samples are presented in Table 1. Series included into comparison group are marked by '+' sign. When the location of the sample 2010 was determined, the comparison group included four series: 2001, 2003, 2004 and 2005. Naturally, the comparison group size was increased gradually and in such a way that the whole data set would be self-consistent. The relative age of all the samples in this case is uniquely determined. It is interest to consider the case of a deriving of the relative age of the sample B-2010 (**Fig. 10**). The dependence of correlation coefficient on the relative position of series has two maxima. However, the second maximum is much lower suggesting that the relative age in this case is also uniquely determined.

6. STATISTICAL ANALYSIS OF SERIES

The correlation coefficient $\mathbf{r_{xy}}$ between **x** and **y** variables is defined as following, e.g., (Bolch and Huang, 1972):

$$r_{xy} = \frac{\sum_{k=1}^{n}(x_k - \overline{x})(y_k - \overline{y})}{\left(\sum_{k=1}^{n}(x_k - \overline{x})\right)^{1/2}\left(\sum_{k=1}^{n}(y_k - \overline{y})\right)^{1/2}}, \quad (6.1)$$

where **n** is the series length, $\bar{x}$ and $\bar{y}$ are the averaged values of these variables.

The examination of a significance of pair correlation coefficient is determined making use of the **F**-statistics. If the length of two compared series equal **n**, the correlation coefficient $\mathbf{r_{xy}}$ is significant if the next conditions are fulfilled:

$$F_{0.95}(1, n-2) < \frac{r_{xy}^2}{(1 - r_{xy}^2)/(n-2)}, \quad (6.2)$$

where $\mathbf{F_{0.95}}$ **(l,m)** is the value of **F**-statistics with **(l,m)** degrees of freedom and 0.95 confidence coefficient. The correlation coefficient has to be sufficiently large to satisfy the condition (**eq. 6.2**). The values of multiple correlation coefficient, corresponding to 0.95 confidence coefficient, as it follows from (**eq. 6.2**), is determined by the number of overlapped rings **n**:

$$(r^2_{xy})_{0.95} = \frac{F_{0.95}(1, n-2)}{F_{0.95}(1, n-2) + n - 2}. \quad (6.3)$$

The meaning of multiple correlation coefficient one can explain as follows. Let's consider one compared series $\mathbf{y_k}$, where $\mathbf{k=1,\dots, n}$, and several series in comparison group: $\mathbf{x_k^i}$, $\mathbf{i = 1,\dots, p}$, where $\mathbf{p}$ is the size of this comparison group. If the series $\mathbf{y_k}$ and $\mathbf{x_k^i}$ are correlated then, on the average, the following equation should be valid:

$$y_k = \sum_{i=1}^{p} c_i\, x_k^i + \xi_k, \quad (6.4)$$

where ξ_k is a noise component. The correlation coefficient between left and right parts of (**eq. 6.4**) depends on the values of coefficients $\mathbf{c_i}$. It is shown in mathematical statistics that a set of coefficients, when correlation is maximal, does exist. This maximal value of correlation coefficient between left and right sides of **(eq. 6.4)** is called as a multiple correlation coefficient.

To calculate the multiple correlation coefficient a special correlation matrix **R** is used:

$$R = \begin{bmatrix} 1 & r_{12} & \dots & r_{1p} \\ r_{21} & 1 & \dots & r_{2p} \\ \dots & \dots & \dots & \dots \\ r_{p1} & r_{p2} & \dots & 1 \end{bmatrix} \quad (6.5)$$

Here **p** is the number of variables, $\mathbf{r_{kl}}$ is the correlation coefficient between k-th and l-th variables. Square of the multiple correlation coefficient between variable **k** and other variables is expressed through the elements of the inverse to **R** matrix:

$$r^2_{k\,|\,1,\dots,k-1,k+1,\dots,p} = 1 - \frac{1}{R^{-1}_{kk}} \quad (6.6)$$

A significance of the multiple correlation coefficient is examined making use of the **F**-statistics. For series of length **n** the correlation coefficient is significant if

$$F_{0.95}(p, n-p-1) < \frac{r^2_{mult}/p}{(1 - r^2_{mult})/(n-p-1)}, \quad (6.7)$$

where $\mathbf{F_{0.95}(l,m)}$ is the value of **F**-statistics with **(l,m)** freedom degrees and 0.95 confidence probability. Really the correlation coefficient has to be sufficiently large to satisfy (**eq. 6.7**). The values of multiple correlation coefficient, corresponding to the 0.95 confidence probability (as follows from **eq. 6.7**) are determined by the number of overlapped rings **n** and the number of series **p** in the comparison group:

$$(r^2_{mult})_{0.95} = \frac{F_{0.95}(p, n-p-1)}{F_{0.95}(p, n-p-1) + n - p - 1}. \quad (6.8)$$

7. CONCLUSIONS

The method of deriving the relative age of wooden samples developed in the present work can be applied along with the radiocarbon dating to date archeological monuments, including the wood constructions, the difference of ages of which permits to construct dendrochronological rows of dates. A sensitivity of the method is increased when one used multiple correlation of several samples. The method allows to control the reliability of matching of samples using the estimations of significance of the multiple correlation coefficient. This method can be applied to analyze deformed samples or destroyed samples.

ACKNOWLEDGEMENTS

The work is supported by RFBR, project No. 03-04-48769 and INTAS, project No. 03-51-4445.

REFERENCES

Alekseev A.Yu., Bokovenko N.A., Boltrik Yu., Chugunov K.A., Cook G., Dergachev V.A., Kovaliukh N., Possnert G., van der Plicht J., Scott E.M., Sementsov A., Skripkin V., Vasiliev S. and Zaitseva G. 2002: Some problems in the study of the chronology of the ancient nomadic cultures in Eurasia (9th – 3rd centuries BC. *Geochronometria* 21: 143-150.

Bolch B.W. and Huang C.J. 1972: *Multivariate statistical methods for business and economics*. New Jersey.

Dergachev V.A. and Vasiliev S.S., 2000: Cross-dating methods in dendrochronology and construction dates of Altai barrows. In: Paraskevopoulos K.M., ed., *Physics in Culture I. The Solid State Physics in the Study of Cultural Heritage*: An Aristotel University Publication. Thessaloniki, Greece: 217-223.

Dergachev V.A., Vasiliev S.S., Sementsov A.A., Zaitseva G.A., Chugunov K.A. and Sljusarenko I.Ju. 2001: Dendrochronology and radiocarbon dating methods in archaeological studies of Scythian sites. *Radiocarbon*. 43(2A): 417-424.

Vasiliev S.S., Bokovenko N.A., K.A. Chugunov K.A., Dergachev V.A., Sementsov A.A., Sljusarenko I.Ju. and Zaitseva G.I. 2001: Tree-rings "wiggle mathching" and statistics in the chronological studies of Scythian age site in Asia. *Geochronometria* 20: 61-68.