ESTIMATING MISSING MEASUREMENTS ON ANIMAL BONE

J.B. Tyldesley, E.A. Hardy and M.H. Williamson

Department of Biology,
University of York, York YO1 5DD.

Introduction

The objective of archaeological bone work is to give a description of the species, size, shape and number of animals found, from place to place and time to time. This objective is often frustrated by the fragmentary nature of the recovered bone.

The Coppergate excavation in York was a large open-area urban dig. The deposits ranged in age from Roman to modern, but the 10th century AD Viking layers were particularly thick and rich. The finds included quantities of animal bone and preservation of organic remains was excellent, since all the lower levels of the site (at the confluence of the rivers Ouse and Foss) were waterlogged. General accounts of the excavation and its place in the history of Viking age York have been given by Addyman (1980) and Hall (1980). Tyldesley (1982) has given an account of the descriptive statistics of the cattle bones from Coppergate. Some 3×10^5 items of animal bone were recovered, of which a small proportion is complete or nearly so. This gives the possibility of relating measurements on complete and on incomplete specimens, so assisting in the interpretation of smaller excavations from which complete bones are seldom recovered.

Material

For this preliminary study, cattle metacarpals were used, since they were common in the Coppergate deposits, as at many other sites. From well sealed and dated 10th century contexts, a group of 68 mature bones (epiphyses fully fused) was assembled, and of these 39 were complete, having the ten measurements shown in Fig. 1. These consisted of the measurements recommended by von den Driesch (1976), with some extra measurements chosen by Dr A. Turner.The Coppergate material was supplemented by 159 complete mature metacarpals found together in a pit by a local history group while searching for the foundations of an old rectory at Moor Monkton, a village on the NW side

	measurement number									
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1	1									
2	0.66	1								
3	0.53	0.74	1							
4	0.67	0.62	0.66	1						
5	0.67	0.87	0.83	0.71	1					
6	0.65	0.71	0.71	0.80	0.80	1				
7	0.70	0.77	0.65	0.67	0.75	0.72	1			
8	0.70	0.79	0.66	0.67	0.76	0.73	0.94	1		
9	0.57	0.83	0.74	0.55	0.87	0.70	0.71	0.72	1	
10	0.53	0.77	0.69	0.53	0.83	0.69	0.66	0.71	0.92	1

Table 1 : Correlation matrix for 159 fused cattle metacarpals.

of York. One of these bones has been dated by radiocarbon to 1590 130 ad, uncalibrated. (Harwell Low Level Measurements Laboratory, sample number HAR-4300).

Relation between the measurements on a metacarpal

Table 1 shows the correlation of each of the ten measurements (Fig. 1) with every other, for the 159 complete Moor Monkton bones. It is clear that each measurement is positively correlated with the others, so that by simple regression a missing measurement x_1 on a bone may be estimated from a measurement which is present x_2 on the same bone by

$$x_1 - \bar{x}_1 = \frac{\sigma_1}{\sigma_2} \; r_{12} \; (x_2 - \bar{x}) \qquad (1)$$

where $\bar{x}_1$ and $\bar{x}_2$ are the means of the respective measurements, σ_1 and σ_2 are the standard deviations of the measurements and r_{12} is the correlation coefficient between them. The best predictor measurement to choose is the one which has the highest correlation coefficient with the missing measurement. Table 2 gives some results for the mature Moor Monkton metacarpals. Regressions were made to estimate each variable, using both the best (most highly correlated) and worst (least highly correlated) single predictor variable. The root-mean-square residuals of the actual measurements from each regression line, give a measure of the quality of the estimates.

Using the conversational statistical package Statpack, multiple regressions were made for estimating each measurement from <u>all</u> the other measurements, e.g.

$$x_1 = a_0 + a_2x_2 + a_3x_3 + \ldots\ldots + a_{10}x_{10} \qquad (2)$$

The results of this are also shown in Table 2. The following points emerge:

a) Tolerable estimates can be produced using any measurement as predictor.

b) Estimates using the best predictor variable may be up to twice as good (measured by the reduction of r.m.s. residual) as those using the worst predictor variable.

c) Using all the other variables as predictors makes only a marginal improvement over using the best single variable.

d) Estimates of the overall length of the bone, measurement 1, are the least satisfactory. The smallest r.m.s. residual for any set of predictors is 5.4 mm, which is a substantial proportion of the standard deviation of the raw measurements, 8.3 mm.

These results are obtained using a homogenous group of bones of one origin and maturity class. They therefore indicate the best results which might be obtained by regression methods. Using regressions obtained from one set of bones to estimate missing values on another set will in general produce larger errors.

Estimating missing values for groups of bones

There is a large literature on missing values in general, which is well summarised by Kendall (1980, pp. 105-107). An obvious starting point is to separate the complete specimens from the group of bones, and use these to set up regressions and thence estimate the missing values. This has been found to work well when most of the specimens are complete, so that the regression coefficients are well estimated. In archaeological work, this is seldom the case, and indeed there may well be no complete specimens at all in a group of bones. One way round this is to estimate the standard deviations from those specimens in which each measurement appears (though others may be missing), and each correlation from those specimens in which the measurements appear pairwise. However, this method has been found by simulation studies sometimes to introduce serious bias, and is not recommended (Haitovsky, 1968).

Beale and Little (1973) in their review of the subject recommend a modification of the method of Buck (1960), the aim of which is to make best use of the information carried by both complete and incomplete specimens. The complete specimens are separated, and used to set up regressions for estimating missing measurements. Missing values for the incomplete specimens are supplied, using these regressions. The estimates are revised, and the procedure repeated until convergence is obtained. The disadvantage of this method is that for a bone with n measurements, 2^n-1 sets of regression coefficients have to be carried, most of which are either not required or make a negligible contribution to the accuracy of the result. For a bone such as the mandible with 17 measurements, the demands on computer time and capacity are unreasonable.

missing measurement			using all predictor measurements		using best single predictor measurements			using worst single predictor measurement		
no.	mean	s.d.	R	r.m.s. residual	no.	R	r.m.s. residual	no.	R	r.m.s. residual
1	198.9	8.3	0.77	5.4	7	0.70	5.9	10	0.53	7.0
2	57.5	3.4	0.90	1.5	5	0.87	1.7	4	0.62	2.7
3	33.4	2.4	0.84	1.3	5	0.83	1.3	1	0.53	2.0
4	22.6	1.2	0.84	0.7	6	0.80	0.8	10	0.53	1.1
5	55.6	3.3	0.95	1.1	9	0.87	1.6	1	0.67	2.5
6	25.2	1.7	0.87	0.9	5	0.80	1.0	1	0.65	1.3
7	31.9	1.7	0.94	0.6	8	0.94	0.6	3	0.65	1.3
8	31.0	1.7	0.95	0.5	7	0.94	0.6	3	0.66	1.3
9	28.1	2.0	0.95	0.6	10	0.92	0.7	4	0.55	1.6
10	27.4	2.0	0.93	0.5	9	0.92	0.8	4	0.53	1.7

Table 2 : Various regressions on 159 fused cattle metacarpals. All measurements in millimetres.

missing meas. no.	predictor measurements nos.	R	r.m.s. residual	R	using all predictor measurements r.m.s. residual
1	7,4,2	0.70	5.4	0.77	5.4
2	5,8,9	0.90	1.5	0.90	1.5
3	5,4	0.84	1.3	0.84	1.3
4	6,1,3,9	0.84	0.7	0.84	0.7
5	9,3,2,6,4	0.94	1.1	0.95	1.1
6	5,4,8	0.87	0.9	0.87	0.9
7	8,5,10	0.94	0.6	0.94	0.6
8	7,10,2,9	0.95	0.5	0.95	0.5
9	10,5,2	0.95	0.6	0.95	0.6
10	9	0.93	0.8	0.93	0.5

Table 3 : Compares the results of automatic regression on 159 fused cattle metacarpals with the use of all measurements as predictors. Measurements in millimetres.

Since Beale and Little's 1973 review, computer procedures have been developed which will choose the best set of predictor variables according to a specified criterion from a larger set of candidates. They work by either entering or removing variables one at a time, and noting the change in variance accounted for by the regression. On the whole such methods have had a bad press from professional statisticians. Kendall (1980, p. 98) mentions two difficulties. Forwards methods (variables entered one by one) and backwards methods (variables removed one by one) may not choose the same set of variables, and neither set may be the optimum. Another serious objection, in cases where many variables are entered into a regression in the hope that they may have some relevance to the problem in hand, is that by a statistical accident the program may replace variables that are physically relevant by others that happen to be highly correlated with them. In bone studies this objection does not apply, because such questions of cause and effect do not arise; any measurement which is effective as a predictor and is present is acceptable. We have therefore used the SPSS package "New regression" (Hull and Nie, 1981) on the Moor Monkton metacarpal data. The methods used are : forward entry, backward removal, and a third method which considers the possibility of removing an existing variable after each new variable has been entered. The criteria for entering and removing data are : probability of F (variance ratio) on entry less than 0.05 and probability of F to remove greater than 0.10. With these criteria the third (in and out) method does not remove any variables, so the results are the same as for the forward method. Table 3 shows results for the forward method. The choice of predictor variables by the backward method gives a list of predictors differing in a few details, but the multiple regression coefficients are the same to within 1% and the r.m.s. residuals to within 3%. The r.m.s. residuals are smaller than for a single predictor variable, but almost the same as are obtained using all variables as predictors. However the results cannot be directly applied to groups of bones with missing measurements, since on any particular specimen some or all of the chosen predictor variables may themselves be missing.

A Compromise Method

We have devised a method of estimating missing measurements which combines some of the advantages of

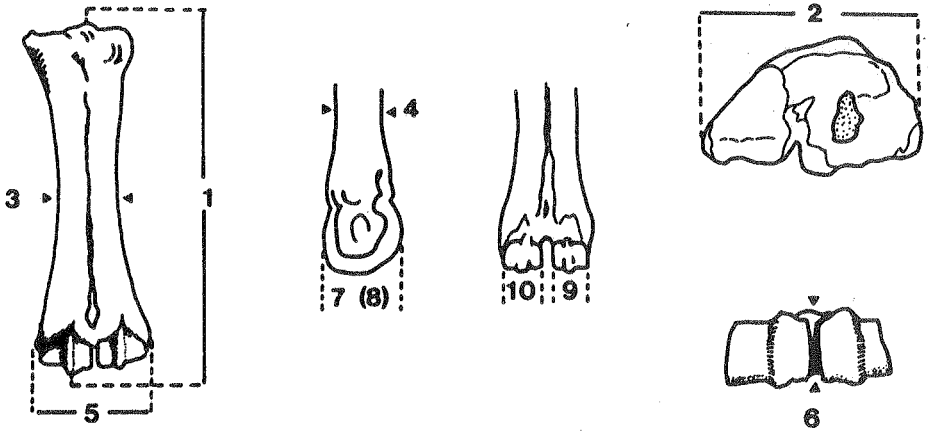

Fig. 1 The ten measurements made on a cattle metacarpal.

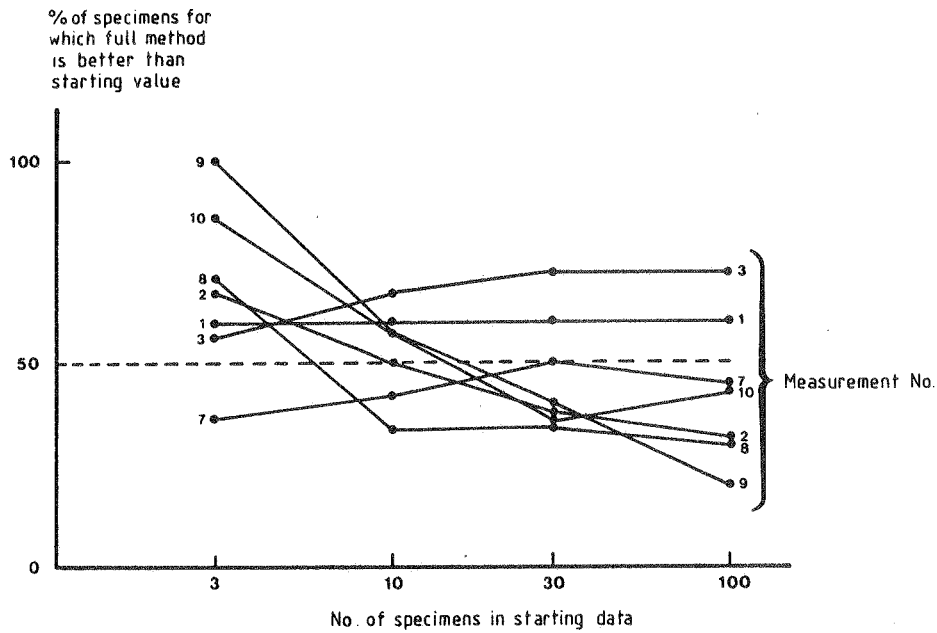

Fig. 2 The full method of estimating missing values compared with the use of the starting value, as affected by the number of specimens used to form the starting values.

using the single best estimator, and of multiple regression using a list of predictors automatically chosen. The steps are in outline:

a) First estimates of missing values are obtained using the best available single estimator, i.e. the measurement which has the highest correlation with the missing measurement using eq. 1. The coefficient may be obtained from the complete sets of measurements in the sample, or if these data are sparse or absent, from other comparable material.

b) The complete measurements are also used to find a list of best predictors for multiple regression, as described in the previous section.

c) Multiple regression (eq. 2) is used to obtain improved estimates. The list of regressors is that obtained in b. The data are the complete specimens and the incomplete specimens patched with first estimates, taken together.

d) Step c is repeated until convergence of the estimated measurements is obtained.

The advantages of this procedure are that the best regressors are used, and only as many sets of regression coefficients need be held as there are measurements on a bone. A disadvantage is that sometimes an estimated measurement will be used as a regressor, when an adequate alternative measured value is available. The procedure has been programmed using NAG (Numerical Algorithms Group) routines for the matrix operations involved. To test the method, we have used as data the measurements from 68 of the complete Moor Monkton bones, but with the pattern of measurement loss of the 68 Coppergate bones artificially applied. In this way it was possible to compare measured and estimated values, for a realistic pattern of measurement loss. The first estimates were formed using data from complete bones from different sources, including Coppergate and Moor Monkton. The following points emerged:

a) The values of final estimates are not affected by the source of the starting data, though the route by which the final values are attained is different.

b) Whether this new method is an improvement on the use of a single predictor variable depends on the particular measurement, and on the number of bones used to form the starting estimates.

Fig. 2 illustrates b above. It shows the proportion of specimens in our sample for which the new method produces a better estimate than a single regressor does. The result changes as the number of bones in the starting data

is increased, as this improves the single estimator results. For measurements 3 and 1 (the important overall length of the bone) the new method is always superior. For all the other measurements except 7, it is better for starting data drawn from five or less complete specimens.

Conclusion

The use of a single predictor variable for estimating missing values is straightforward and not to be despised, particularly when a large group of complete specimens is available for deriving the regression coefficients. In more difficult circumstances, when complete specimens are few or absent, and for estimating the overall length of long bones (and thence the stature of the animal) the new method shows considerable promise. Further testing is required, for more severe patterns of deletion than those found on Coppergate, and for other bones in the skeleton.

Acknowledgements

Dr A. Turner recorded the bones. Dr P. Newman made available the Moor Monkton material. Professor J.D. Currey gave much useful advice. The work was supported by the Science and Engineering Research Council under grant GR/A/9517.8 : Multivariate studies of changes in bone of selected species from archaeological deposits.

References

Addyman, P.V. 1980 Excavating Viking age York, ARCHAEOLOGY, 33, 14-22.

Beale, E.M.L. and Little R.J.A. 1975 Missing values in multivariate analysis. JOURNAL OF THE ROYAL STATISTICAL SOCIETY SERIES B, 37, 129-145.

Buck, S.F. 1960 A method of estimation of missing values in multivariate data suitable for use with an electronic computer. JOURNAL OF THE ROYAL STATISTICAL SOCIETY SERIES B, 22, 302-306.

Haitovsky, Y. 1968 Missing data in regression analysis. JOURNAL OF THE ROYAL STATISTICAL SOCIETY SERIES B, 30, 67-82.

Hall, R. 1980 JORVIK - VIKING AGE YORK, York Archaeological Trust, York.

Hull, C.H. and Nie, N.H. 1981 New regression. SPSS UPDATE 7-9, 3, 94-121, McGraw-Hill, New York.

Kendall, M. 1980 MULTIVARIATE ANALYSIS, 2nd edition, Charles Griffin, London.

Tyldesley, J.B. 1982 Exploration of cattle bone measurements by statistical techniques including principal component anaylsis. PROCEEDINGS OF THE 22nd SYMPOSIUM ON ARCHAEOMETRY, MICROCOMPUTER JAMBOREE, 5-10. University of Bradford, Bradford.

von den Driesch, A. A GUIDE TO THE MEASUREMENT OF ANIMAL BONES FROM ARCHAEOLOGICAL SITES. Peabody Museum Bulletin 1. University of Harvard, Harvard.