Math Geosci
DOI 10.1007/s11004-010-9275-8

CASE STUDY

A Discrete Classification Procedure Applied to Encrusted Ammonite Conches—A Contribution to Taphonometrics

Richard A. Reyment

Received: 27 October 2009 / Accepted: 4 April 2010

Abstract Oyster encrustations on shells of the Albian (Cretaceous) ammonite species *Knemiceras persicum* (Iran) and *K. syriacum* (Lebanon) are analysed by discrete discriminant analysis, a method based on distributional difference in qualitative characters appropriate for greatly unequal sample sizes. The two localities, which are approximately coeval and separated by a distance of 2000 km, are shown to have different taphonometrical backgrounds, as manifested in the mode of occurrence and preservation of the encrustations.

Keywords Epizoans · Ammonites · Discrete discrimination · Cretaceous · Albian · Middle East · Taphonometrics

1 Introduction

A problem occurs in interpreting the post-mortem history of cephalopod shells concerns, whether the fossils occur in situ or transported after death by water currents. A useful indicator of this problem is the presence of epizoans such as oysters and other encrusters on the shell. Kennedy et al. (2009) have given a detailed taxonomic and variational study of two species of *Knemiceras* from the Middle East, *Knemiceras persicum* Collignon (Iran) and *Knemiceras syriacum* (von Buch; Lebanon) to which the reader is referred for details bearing on the collections studied in this note and for stratigraphical analyses of the occurrences. The ammonites found in both places are remarkable in that a high percentage of conches bear epizoans, mainly oysters (Fig. 1).

The material studied in this paper comes from beds of Albian age (Kennedy et al. 2009). Encrusted *Knemiceras syriacum* (von Buch) from Jebel Liban were first reported by Hyatt (1903). The Iranian *Knemiceras* were originally thought to belong to

R.A. Reyment (✉)
Paleozoologiska Avdelningen, Naturhistoriska Riksmuseet, Box 50007, 10405 Stockholm, Sweden
e-mail: richard.reyment@nrm.se

Fig. 1 *Knemiceras persicum* Collignon. Example of oyster encrustations (left valves) on the left flank of a specimen, with half of the body chamber preserved, from the Albian (Lower Cretaceous) Kazdhumi Formation, Hamuran, Iran. (N. B. The left valve is always the encrusting one in oysters.) Oxford University Museum KY15 (specimen diameter = 150 mm). Collector: Peter Kent

K. syriacum, but Collignon (1983) proved this to be unlikely. Reyment and Kennedy (1991) could verify Collignon's opinion by multivariate morphometric analysis. Both occurrences are of a kind that is relatively rare among ammonites of the Cretaceous, and inasmuch as the taphonomic relationships of the two sequences seem, at first sight, to be similar, it becomes a matter of palaeoecological interest to put this impression to the test.

The categories of shells occurring in the collections are:

(1) One flank only encrusted by oysters (and very occasionally, barnacles and serpulids), with or without encrusts on the venter.
(2) Both flanks encrusted with or without encrusts on the venter.
(3) No encrusted shells.

The first case derives from oyster spat having settled on stranded shells. The second case represents encrusting organisms having settled on nekroplanktonically transported shells, on living shells fouled in life, or on reworked moulds (Reyment 2008). Two possible examples of the fouling of living individuals were observed in the Persian material, an indication of which is that the orientation of the length axis of the encrusting organism runs parallel to the tangent of the direction of coiling. The species of oysters forming the encrusts is *Exogyra* cf. *ardennuensis* d'Orbigny, according to Collignon (1983). The Russian palaeontologist J.A. Efremov coined the term Taphonomy in 1940 for the study of burial, sedimentation, and preservation of animals and plants. The statistical analysis of such data requires special methods encompassing factors not usually considered to form a united whole in neontological work. It is for this special aspect of palaeobiometry that I here coin the term 'taphonometrics'.

2 Statistical Analysis of the Oyster Encrustations

2.1 Hypothesis of Equal Frequencies

Suppose that we have a set of observed frequencies $F_1, F_2, F_3, \ldots$. A standard chi-squared test may be used to decide whether these frequencies differ significantly from

Table 1 Allocation rule results for the encrustation states for Lebanon and Iran

State	Π_1 Observed Lebanon (n_j)	Π_2 Observed Iran (m_j)	Rule-value (formula (2))	Allocation
1 0 0	6	7	1.000	Π_1
1 1 0	6	16	1.011	Π_1
1 1 1	6	42	1.024	Π_2
Sum	18	65		

Cut-off value for allocation rule = 1.018 (expression (3) in text)
Key:
1 0 0 denotes one flank of conch encrusted
1 1 0 denotes one flank of conch and venter encrusted
1 1 1 denotes both flanks of conch and venter encrusted
Π_i ($i = 1$ or 2) indicates the classification of the state according to the discrete discrimination rule of Kameo Matusita

an expected hypothesis. Reyment (1971, Chap. 7) illustrates examples of the test of common occurrence in palaeontology. In the present case, the hypothesis is that all of the encrusting categories recorded occur in equal proportions under the supposition that oyster spat have settled on conches under like conditions at both sites (Table 1). A test of association by chi-squared for the categories is given: one side fouled (1 0 0), one side and venter encrusted (1 1 0), both sides and venter fouled (1 1 1), and shells lacking epizoans (0 0 0). It indicates that the proportions for the Iranian specimens (ΣN = 81) give a very highly significant value of chi-squared of 32.10 for 3 d.f., with $P < 0.0001$, which indicates that the postulate that the frequencies for all states are equal, is not supported. Granted that about half of the Iranian conches are fouled on all sides (i.e. both flanks and venter), the result is not unexpected and it would seem likely that the majority of the shells were settled upon while floating post-mortem in shallow water. The result for the Lebanese material (ΣN = 27) yields a non-significant value of chi-squared of 1.000 for 3 d.f., with $P = 0.80$, which may indicate that the material has a mixed origin such as would be found where reworking is an important factor and where nekroplanktonic shells occur jumbled together with shells that were originally settled by epizoans while lying on the sea-floor, before and/or after reworking. These results suggest that the Iranian material is composed mainly of nekroplanktonically transported conches.

2.2 Discriminant Allocation

In many situations, discrete classification rules may be constructed through approximations to, and estimates of, multinomial densities. For the problem considered in this note, a different approach is used, in which likelihood ratios are not used as a foundation for classification, since the data are discrete and depart from a normal model. A problem also arises out of the small size of the Lebanese data set. The chi-squared method may be adversely affected by unequal sample sizes, particularly where one sample is much smaller than the other one; such is the case here. A useful, partly heuristic approach is a procedure based on distributional distance originally proposed by Matusita (1954, 1964) (also, Dillon and Goldstein 1978) for greatly unequal sample sizes. The notation of Goldstein and Dillon (1978) is adhered to in the following.

Consider data generated by discrete random variables $X_1, X_2, \ldots, X_p$ encompassing distinct values $s_1, s_2, \ldots, s_p$ in the sample space consisting of $s = \Pi_j s_j$ states generated by the vector $\mathbf{X}$ which has a multinomial distribution. Let $F_1 = \{p_j\}$ and $F_2 = \{q_j\}$, denote two discrete distributions defined on the same space. The states here are the patterns of occurrence of the encrustations as shown in Table 1 (three states: the state encompassing specimens lacking encrustations is not relevant for the classification exercise as opposed to the chi-squared model).

Matusita's concept of distributional distance relies on a measure of affinity defined as

$$\rho(F_1, F_2) = \sum_{j=1}^{s} (p_j q_j)^{1/2},$$

and a measure of distance between F_1 and F_2 given by

$$\|F_1 - F_2\|^2 = \sum_{j=1}^{s} \bigl(p_j^{1/2} - q_j^{1/2}\bigr)^2,$$

for expressing classification rules. The distance $\|\ \|$ is a metric with properties that are not shared by other distributional distances (Dillon and Goldstein 1978, p. 306). An allocation rule may be expressed as an ordered partition $D = (D_1, D_2)$ of the sample space. The rule is that D allocates an observation $\mathbf{x}$ to Π_i if $\mathbf{x} \in D_i$, for $i = 1$ or 2 (Goldstein and Dillon 1978, p. 11).

If $n < m$ so that $m(n+1)/n(m+1) > 1$ and $n_k = 0$ but $m_k \neq 0$, then the rule allocates the observation to F_1 of the inequality (1)

$$(m_k)^{1/2} < \sum_{j \neq k} (n_j m_j)^{1/2} \bigl[\bigl(m(n+1)\bigr)/\bigl(n(m+1)\bigr)\bigr]^{1/2} - 1, \tag{1}$$

which is satisfied, if the samples are very unequal in size, and m_k is small in relation to the other values of the smaller of the samples being compared. The values n and m constitute the sum of all observations of the states for each of the two samples ($n < m$). The results summarised in Table 1 place one of the samples with the second (larger) group (Iran) and two samples with the first group (Lebanon), in accordance with the allocation model of Matusita (1954, 1964). In other words, both sites contain shells some of which are thought to have a nekroplanktonic history, but the history of the Lebanese material bears witness to a more heterogeneous background such as would result from reworking and renewed burial of the fossils. The way in which (1) is computed is as follows.

Compute firstly the ratio (2)

$$\bigl\{m_j(n_j+1)^{1/2} + C\bigr\}/\bigl\{n_j(m_j+1)^{1/2} + C\bigr\}, \tag{2}$$

the operation of which requires finding the constant term C. This is done by computing the appropriate sum of square roots of the product $n_j m_j$, that is the sum of the square roots of all the states with the exclusion of the state to be allocated. Thirdly, compute a boundary-value (3) for the allocation rule:

$$\left[m(n+1)/n(m+1)\right]^{1/2}, \tag{3}$$

where n and m retain the same significance as above.

A problem that can arise in the statistical analysis of fossil material concerns how is one to handle sparse states and, in particular, the question as to whether all zeros are logically equal. When it comes to the application of statistical methodology to fossils, zero takes on a special significance in terms of the likelihood of turning up a non-zero observation under repeated sampling varying from sampling source to sampling source. A qualified zero therefore has a real significance. The case of encrusting organisms with n much smaller than m, is considered here. A zero registered for state j from F_1 could really represent a greater theoretically possible frequency than a very small value recorded in state j from F_2 (Dillon and Goldstein 1978, p. 312). This uncertainty leads automatically to the solution by the full multinomial rule, notably that the non-zero state is referred to F_2. The rule in Goldstein and Dillon (1978, p. 44), expressed in (1), relaxes the severity of an automatic assignation such as follows from the unqualified application of the full multinomial rule. This result of Goldstein and Dillon has been formally termed the Dillon Goldstein Classification Rule by Pires and Branco (1997).

Comparing the two methods used here, the standard chi-squared procedure and a discrete classification method, we see that results obtained by both agree, albeit on different statistical grounds. However, the discrete classification method goes a step further in that it outlines the source of the discrimination by means of an allocation rule. For states 1 0 0 and 1 1 0, the proportions are commensurable with Π_1. The decision for state 1 1 1 is for allocation to Π_2. The conclusion is that the taphonomic histories of the two occurrences differ with respect to the mode of occurrence of the encrustations.

3 Conclusions

The results for *Knemiceras* spp. attempt to interpret the differences in the presumed taphonomy of the Iranian and Lebanese ammonite occurrences. A hypothesis as to whether or not the Iranian *Knemiceras* were fouled in life and/or post-mortem must take into account the ecological constraint imposed by the life-history of oysters, particularly the fact that oysters are euryhaline organisms which inhabit a littoral environment. The results of the present study suggest that the environment inhabited by living ammonites was near-shore. Nekroplanktonic shells floating in a near-shore environment would have provided a firm settling ground for oyster spat.

As any person with professional sea-going maritime experience will confirm, once a ship starts to lose power seriously, it is time to put it into dry dock for a hull-scraping. Most of the fouling organisms are acquired when in port or riding at anchor. This raises the question as to how the fouling of living cephalopod shells takes place. The energy required to maintain a swimming speed for a heavily encrusted shell is so considerable that such a condition would not be sustainable for long by the hosting animal (Reyment 1988). For this reason, it seems likely that encrustation of conches on all sides takes place post-mortem, except in rare circumstances. It is common praxis to treat qualitative data as though they were continuous and to use the Linear Discriminant Function of R.A. Fisher. The LDF does not require normality, but its justification in a likelihood-ratio connection does. The method of distributional distances avoids this difficulty and introduces a taphonomical ingredient to the problem which facilitates interpretation of a complex palaeontological problem.

Once it has been found possible to distinguish between life-assemblages and reworked assemblages in a particular study, it may be interesting to test the application of a log-linear model to the question of determining the structure of the mixed assembly in more detail if the problem posed in quantifying the different taphonomic origins of the encrustations can be surmounted and the sample sizes are not too divergent (cf. Agresti 2007, p. 329). I thank a referee for suggesting such a possibility.

Acknowledgements I wish to thank Professor W. James Kennedy, Director of the Natural History Museum of Oxford University, Parks Road, Oxford for helping in many ways and for making the facilities of the museum available to me. A research grant from the Dunker Fund of Kungliga Fysiografiska Sällskapet, Lund is gratefully acknowledged.

References

Agresti A (2007) An introduction to categorical data analysis, 2nd edn. Wiley series in probability and statistics. Wiley, New York
Collignon M (1983) Faune Albo-Cénomanienne de la formation des marnes de Khazhdumi (1981) région du Fars-Khuzestan (Iran). Doc Lab Géol Lyon, Hors Ser 6:251–291
Dillon WR, Goldstein M (1978) On the performance of some multinomial classification rules. J Am Stat Assoc 73:305–313
Goldstein M, Dillon WR (1978) Discrete discriminant analysis. Wiley series in probability and mathematical statistics. Wiley, New York
Hyatt A (1903) Pseudoceratites of the Cretaceous. US Geol Surv Monogr 44:351
Kennedy WJ, Reyment RA, MacLeod N, Krieger J (2009) Species discrimination in the Lower Cretaceous (Albian) ammonite genus *Knemiceras* (Böhm, 1898). Palaeontogr, Abt A 290:1–63
Matusita J (1954) On estimation by the minimum distance method. Ann Inst Stat Math 7:67–77
Matusita J (1964) Distance and decision rules. Ann Inst Stat Math 16:305–315
Pires A, Branco J (1997) Comparison of multinomial classification rules. J Classif 14:137–145
Reyment RA (1971) Introduction to quantitative paleoecology. Elsevier, Amsterdam
Reyment RA (1988) A foraging model for shelled cephalopods. In: Wiedmann J, Kullmann J (eds) Cephalopoda—present and past. Schweizerbart'sche Verlagsbuchhandlung, pp 687–703
Reyment RA (2008) A review of the post-mortem dispersal of cephalopod shells. Palaeontol Electronica 11:1–13
Reyment RA, Kennedy WJ (1991) Phenotypic plasticity in a Cretaceous ammonite analyzed by multivariate statistical methods. Evol Biol 25:411–426