



TESTING THE UTILITY OF GAUSSIAN MIXTURE MODELS IN PARATAXONOMIC CLASSIFICATION OF FOSSIL EGGS

Xavier Panadès I Blas¹, Roland Baddeley² & John Fennell²

¹ Museu de la Conca Dellà, C/ del Museu 4, 25650, Isona, Catalonia, Spain;
EC xpan100@gmail.com

²The University of Bristol, Department of Experimental Psychology, 12a, Priory Road, BS8
1TU, Bristol, UK, EC

Panadès I Blas, Xavier, Roland Baddeley & John Fennell. Testing the Utility of Gaussian Mixture Models in Parataxonomic Classification of Fossil Eggs. – Palarch's Journal of Vertebrate Palaeontology 17(3) (2021), 1-13. ISSN 1567-2158. 13 pages + 3 figures, 3 tables.

Keywords: Gaussian mixture models, Dinosaur eggshells, Parataxonomy, Biodiversity, Megaloolithidae, Dinosaur reproduction

ABSTRACT

Fossil eggs and eggshells are very seldom associated with skeletal or dermal remains and can rarely be assigned taxonomically. This has resulted in oologists creating an alternative classification system for fossil eggs and eggshells: the parataxonomy of fossil eggs. Under parataxonomy, fossil eggs are classified in oospecies, oogenera and oofamilies according to qualitative microstructural characters (*e.g.* shape of the shell units and ornamentation) and highly heritable quantitative characters (*e.g.* height of unit, HU, and width of unit, WU). *Megaloolithus* oospecies is an egg type previously partially attributed to both titanosaurian and hadrosaurid dinosaurs. These oospecies are defined by the shape of their units and the quantification of the morphometric variation and variability of their microstructures (HU and WU of eggshell units). Two competing interpretations of the observed morphological variation of the eggshell units have been proposed: 1) different megaloolithic morphologies are indicative of different dinosaur species; and 2) the same dinosaur species was responsible for all the variation seen in megaloolithic eggshell units. In this study, a Gaussian mixture model was applied to test both interpretations. This probabilistic model assumes all the data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. We compared the morphometric distributions of HU and WU in eggshells, eggs and clutches belonging to six *Megaloolithus*

oospecies from Catalonia to the defined morphometric variation seen in the eggshell microstructures of two extant turtle species and two crocodile species. The resulting Gaussian model was best defined for four distributions of HU and WU representing the extant turtle and crocodile oospecies, and three distributions consisting of one or more *Megaloolithus* oospecies. It was inferred that *Megaloolithus* oospecies from the Late Cretaceous of Catalonia depict the polymorphism of a monotypic taxon and were laid by a single dinosaur species.

INTRODUCTION

Parataxonomy creates a systematic taxonomic framework for oological material that cannot be directly or indirectly associated to osseous and embryonic remains *in ovo* and *in utero*, eggs and eggshells *in utero*, DNA in more recent material, or correlations of bones and palaeoöological material from independent deposits (Varricchio *et al.*, 2002; Sato *et al.*, 2005; Oskam *et al.*, 2010; Huynen *et al.*, 2010).

Parataxonomy defines oospecies via quantification of the polymorphism of egg characteristics (eggshell thickness, egg size, details of ornamentation, and external pore patterns) (Mikhailov 1991; 1997a; 1997b; Mikhailov *et al.*, 1996). For example, the dinosaurian megaloolithid oospecies have been defined by the intra- (between eggs and clutches) and inter-specific (oospecies) morphometric variability of the height and width of their units (Vianey-Liaud & Lopez-Martínez, 1997; Vianey-Liaud *et al.*, 2003; Panadès I Blas, 2005).

Megaloolithid oospecies have been used to estimate the diversity of dinosaurs during the Late Jurassic as well as the Late Cretaceous in areas where eggshell material is more abundant than diagnostic skeletal remains (*e.g.* Zhao & Ding, 1976; Vianey-Liaud & Lopez-Martínez, 1997; Magalhães Ribeiro, 2002; Vianey-Liaud *et al.*, 2003; Panadès I Blas, 2005; Grellet-Tinner *et al.*, 2004; Grigorescu, 2016; Khosla & Lucas, 2020). However, the role of megaloolithid palaeoöspecies as indicators of nesting palaeobiodiversity is debated, because palaeontologists disagree over the taxonomic validity of megaloolithid palaeoöspecies (Vianey-Liaud & Lopez-Martínez, 1997; Peitz, 2000; Vianey-Liaud *et al.*, 2003; Panadès I Blas, 2005). Indeed, traditionally megaloolithid eggs and eggshells are assigned to sauropods and hadrosaurids, based on direct association with embryonic remains *in ovo*, bones, and from correlations of bones and palaeoöological material from independent deposits (Chiappe *et al.*, 1998; Grigorescu, 2016). Two hypotheses have been proposed to explain the taxonomic validity of megaloolithid palaeoöspecies (Vianey-Liaud & Lopez-Martínez, 1997; Peitz, 2000; Vianey-Liaud *et al.*, 2003; Panadès I Blas, 2005):

1. Each megaloolithid palaeoöspecies is defined by dimensional variations of the same eggshell's fan-shape structural units, and represents dinosaur polytypic diversity. Therefore, the diversity of megaloolithid palaeoöspecies reflects the diversity of dinosaur palaeobiodiversity (Vianey-Liaud & Lopez-Martínez, 1997; Vianey-Liaud *et al.*, 2003; Panadès I Blas, 2005).

2. Or, megaloolithid palaeoötaxa are characterised by different variations of the same eggshell's fan-shape unit from different areas of the eggs, and describes the polymorphism of a monotypic taxon, and consequently are not useful indicators of palaeobiodiversity (Peitz, 2000).

Here, we test both hypotheses using a Gaussian Mixture Model. We argue that we can solve the taxonomic validity of *Megaloolithus* oospecies, by comparing the morphometric variability of their eggshells to the variability of oospecies of defined layers such as modern turtle and crocodile oospecies, using a Gaussian Mixture Model.

Gaussian mixture models predict the probability of distributions (set of points) to belong to a set distribution of classes (Reynolds, 2008). The Gaussian mixture model will determine whether the morphometric variability of *Megaloolithus* oospecies is equivalent to the variability of modern reptile oospecies, or not.

MATERIALS

Reptilian nests of extant species of turtle and crocodile were used as comparative material for the analysis, because the structure of their eggshells is the most similar to megaloolithid palaeoögenera (Mikhailov, 1991; 1997a; 1997b; Mikhailov

et al., 1996). Megaloolithid, crocodile, and turtle eggshells are composed of discrete units along the eggshell thickness that extend from the interior to the exterior of the thickness of the shell.

The eggs used were laid by the same female, and are of two different species of turtle and crocodile, to capture the regional intraspecific variation of HU and WU exhibited by reptile and megaloolithid oospecies (Mikhailov, 1991; 1997a; 1997b; Mikhailov *et al.*, 1996; Vianey-Liaud & López Martínez, 1997). The materials used were two clutches with three and two eggs of *Geochelone carbonaria* (Gc), two clutches with three eggs each of *Testudo hermanni* (Th), donated by Bristol Zoo, four clutches with three eggs each of *Crocodylus niloticus* (Cn) from La Ferme aux Crocodiles Zoo (France), and four clutches with three eggs of *C. porosus* (Cp) obtained from the Koorana Crocodile Farm (Queensland, Australia).

Upper Cretaceous megaloolithid eggshells and data on megaloolithid material from previous studies were used in this project. Three nests were sampled from Catalan localities: a nest of two eggs from Coll de Nargó (Na) no. IPS-33141, a nest of three eggs from Fumanya (Fu) no. IPS-27380, a nest with three eggs from Faidella (Fai) no. IPS-36341. The nests are stored at the Institut Català de Paleontologia Miquel Crusafont (Sabadell, Catalonia). Data on megaloolithid eggshells from Coll de Nargó (from Peitz, 2000) were computed in the analyses.

Six *Megaloolithus* oospecies are represented in Coll de Nargó, Faidella and Fumanya: *Megaloolithus baghesis* (Coll de Nargó), *M. aureliensis* and *C. cf. roussetensis* (Coll de Nargó), *M. mamillare* (Coll de Nargó, Basturs, and Fumanya), *M. siruguei* (Coll de Nargó, Faidella, and Fumanya), and *M. sp* (Fumanya), (Vianey-Liaud & López Martínez, 1997; Panadès I Blas, 2005; Vila *et al.*, 2010 & 2012). The random sampling of eggshell measurements by Peitz (2000) is ideal for this study because it will likely include most palaeoöspecies present in the localities. Random sampling gathers representative samples of larger populations (Smyth *et al.*, 2019).

The three nests were assigned parataxonometrically to *M. siruguei*. They exhibit the typical characters of *M. siruguei*: subspherical eggs of 20 × 18.5 cm; eggshell thickness ranging between 1.75 and 3.6 mm composed of elongated fan-shaped units five times higher than wide exhibiting intermittent fusion, with arched growth lines running from the base to the top of the units and crossing to adjacent units (Figure 1) (Vianey-Liaud *et al.*, 1994; Elez & López-Martínez, 2000; Vianey-Liaud & Zelenitsky, 2003; Sellés *et al.*, 2013; 2017).



Figure 1. Petrographic thin sections of eggshells from the pole 1 (A), pole 2 (B) and equator (C) of an egg from the nest from Faidella used here. Notice the different morphometrical and shape of units between the eggshells from the equator and the poles.

METHODS

Parataxonomy defines oospecies via the quantification of the diversity/variability of egg characters, height (HU) and width (WU) of units, egg size, details of ornamentation, and external pore patterns (Mikhailov, 1991; 1997a; 1997b; Mikhailov *et al.*, 1996). Only well-defined and consistent ranges of these characters should be applied for taxonomic inference (Mikhailov, 1991; 1997a; 1997b; Mikhailov *et al.*, 1996). Therefore, only HU and WU measurements will be used here because these are the most widely used in parataxonomic analysis (Mikhailov, 1991; 1997a; 1997b; Mikhailov *et al.*, 1996; Schleich *et al.*, 1994; López-Martínez *et al.*, 2000).

The intraspecific variation of HU and WU in reptile and megaloolithid oospecies appears to vary between equator and poles of eggs and may be highly variable in eggs and clutches (Schleich & Kastle, 1988; Mikhailov, 1991; 1997a; 1997b; Mikhailov *et al.*, 1996; Schleich *et al.*, 1994; López-Martínez *et al.*, 2000). For example, Schleich *et al.* (1994) and Panadès and Patnaik (2009) found that modern and fossil crocodile eggs are thicker in the equatorial region than at the poles, while López-Martínez *et al.* (2000) found that a single egg in a megaloolithid nest was thicker at the poles and thinner at the equator. However, none of the studies established whether morphological and morphometrical changes of the eggshell units are correlated.

Accordingly, samples were taken from the equator of *Megaloolithus* and reptile eggs, mounted as petrographic thin sections, and photographed under a Polarised Light Microscope (PLM) (Carpenter, 1999; García & Vianey-Liaud, 2001a, 2001b; Grellet-Tinner, 2000). Unlike the ovoidal eggs of megaloolithic layers and crocodiles, the poles and equator in the spherical eggs of turtles are not distinct. Two crossing circles were drawn around the middle of the eggs of turtles and the poles were marked on the opposite sites of one of the circles. The two eggshells were extracted, inspected under a light microscope for any pathology. Eggshells with double layering, diagenetic alterations and/or eggshells without completed units were discarded. The resultant thin sections have been deposited at the Museu de la Conca Dellà (Catalonia).

HU and WU from each eggshell unit were captured digitally using Image J 1.47a from photographs of calibrated petrographic thin sections (Rasband, 1997-2012). Under the PLM, HU was measured from the base to the highest point in a vertical line; and WU from the widest straight horizontal line of the unit. This method is more accurate than measuring eggshells from SEM (Scanning Electron Microscope) micrographs and/or using a calliper, since taking measurements by either of these methods can produce distorted data. With the naked eye, the structure of the eggshells, diagenetic alterations and pathologies cannot be discerned. Equally, obtaining eggshells measurements from SEM micrographs can be inaccurate since the specimens may not be at 90° and any protruding surfaces may distort the real measurements. The microstructures of eggshells are more distinguishable under polarised light (PLM), as it provides real two-dimensional visualisations of the morphometrical variations of the units at a perfect 90° angle (Mikhailov, 1991; 1997a; 1997b; Mikhailov *et al.*, 1996; López-Martínez, 1997; Carpenter, 1999).

Previous quantifications of megaloolithid oospecies have been limited to small samples that were taxonomically constrained, and they did not test independently the relations of the HU and WU distributions (*e.g.* García & Vianey-Liaud, 2001a and b; Vianey-Liaud & López-Martínez, 1997; Panadès I Blas, 2005). In order to investigate how the measurements might be related, independent of the site information (eggshells, eggs, and localities), a Gaussian mixture model was

used to establish the most probable number of distributions of megaloolithic HU and WU distributions using the taxonomically constrained samples of modern reptiles.

A Gaussian mixture model is a probabilistic model that assumes that data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters (Reynolds, 2008). Fitting the best mixture of Gaussians for a given dataset (as measured by the log likelihood) results in a probability distribution of classes that can be used to predict the probability (posterior) of new data points belonging to those classes (Reynolds, 2008). Fitting Gaussian mixture models is an example of an unsupervised learning method, however, the computing required for fitting a mixture of Gaussians is exponential for the number of latent Gaussian distributions, so approximate inference techniques are often used (Reynolds, 2008). While this does not guarantee *the* optimal solution, models do converge quickly to a local optimum (Reynolds, 2008).

To improve the quality of the results it is usual to fit many of these models and choose the model that best fits the data, often on the basis of log likelihood or similar approach (Reynolds, 2008). Here the Gaussian mixture model functions from the Netlab toolbox (Nabney & Bishop, 2004) are used: these functions initialise the model using a clustering process known as *k-means* and then use the expectation maximisation (EM) algorithm. EM is an iterative method for obtaining maximum likelihood estimates of parameters for models that depend on unobserved variables: in the present case a finite number of Gaussian distributions representing the relationship between measurements irrespective of collection site (Nabney & Bishop, 2004).

Mixture models were produced for two to ten centres, each repeated ten times. Each centre represented HU and WU distributions of modern reptiles and megaloolithic eggshells. From the ten repetitions for each of the models, the best log likelihood was obtained. From these, as shown in Equation 1, the minimum description length L , the best number of distributions accounting for the measurements, was calculated and used to choose the number of centres that best represented the data (Nannen, 2003:14). Here, D is the data set, M_k is the model, k is the number of Gaussians or centres and N is the number of points in the data set.

$$L(D) = \min[-\log P(D|M_k) + k \log \sqrt{N}] \quad (1)$$

A between groups Kolmogorov-Smirnov test was then carried out with the measurements of thickness and width as the dependent variables and centre as the factor in order to investigate the difference in thickness and width for each of the centres found by the mixture model. The classical one-dimensional Kolmogorov-Smirnov test is a non-parametric statistic for comparing two empirical distributions, which defines the largest absolute difference between the two cumulative distribution functions as a measure of disagreement. Adapting this test to more than one dimension is a challenge because there are $2d-1$ independent ways of defining a cumulative distribution function when d dimensions are involved.

We used a Matlab implementation (Lau, 2018) of a two-dimensional Kolmogorov-Smirnov test (Fasano & Franceschini, 1987; Peacock, 1983). Since we are doing multiple comparisons, we need to adjust the alpha level according to the number of comparisons being made. The simplest and most conservative correction is the Bonferroni correction, which is satisfactory for present purposes:

$$\text{Adjusted alpha level} = \alpha/n \text{ comparisons} = 0.05/7 = 0.0071$$

RESULTS

Comparison of the minimum description lengths for each of the Gaussian mixture models revealed that a model with seven centres provided the best description of the data, suggesting that the measurements could be considered to belong to seven distinct distributions (Figure 3). The regression line is a very good fit to both modern reptile and the dinosaur affirming that megaloolithid and modern reptile eggshells possess similar variability distributions. The regression line suggests that this is true for all of the measured shells, including the bigger scale megaloolithid eggs.

In fact, the variability of egg thickness among reptile and dinosaurs is very similar: both are highly variable (coefficient of variation of about 20%). The mean and variability (standard deviation) of dinosaur eggs are not different to reptiles, just bigger; the regression line is a fit to both the reptile and dinosaur data (Figure 2). The Kolmogorov-Smirnov test revealed a significant difference for both width and thickness between all the centres (Table 1-3).

DISCUSSION

The Gaussian Mixture Model has defined seven oospecies: four representing the extant turtle and crocodile oospecies, and three representing *Megaloolithus* oospecies at the end of the Cretaceous in Catalonia (Figure 3). The three *Megaloolithus*

Table 1. Descriptive statistics for each centre found by the Gaussian mixture model.

Centre	Thickness			Width			
	Mean	SD	SE	Mean	SD	SE	N
1	0.413	0.068	0.002	0.396	0.115	0.003	1602
2	0.656	0.086	0.007	0.426	0.202	0.017	136
3	0.369	0.081	0.004	0.181	0.040	0.002	452
4	2.524	0.314	0.017	0.607	0.120	0.007	335
5	3.412	0.707	0.056	0.819	0.177	0.014	157
6	2.945	0.160	0.009	1.076	0.267	0.015	299
7	1.207	0.179	0.014	0.481	0.143	0.011	157

Table 2. Statistical comparisons between the centres produced by the Gaussian Mixture Model under alpha/n comparisons.

KS statistics						
	Centre 2	Centre 3	Centre 4	Centre 5	Centre 6	Centre 7
Centre 1	0.998	0.895	0.998	0.998	1.000	0.735
Centre 2		0.982	1.000	0.981	1.000	0.998
Centre 3			1.000	0.969	1.000	0.954
Centre 4				1.000	1.000	1.000
Centre 5					0.974	0.998
Centre 6						1.000

Table 3. Statistical comparisons between the centres produced by the Gaussian Mixture Model under Adjusted alpha level.

P-value						
	Centre 2	Centre 3	Centre 4	Centre 5	Centre 6	Centre 7
Centre 1	7.07E-121	1.46E-92	1.23E-28	0	1.15E-72	1.10E-174
Centre 2		1.25E-62	6.33E-24	1.35E-101	1.84E-48	4.84E-103
Centre 3			9.99E-24	1.22E-95	1.14E-47	1.33E-90
Centre 4				9.49E-28	4.98E-21	1.41E-27
Centre 5					5.16E-63	2.95E-236
Centre 6						1.03E-65

oospecies are considered taxonomically valid. The oospecies are statistically different to each other and to the taxonomically constraint samples of turtle and crocodile oospecies (see Figure 2, Table 2, 3). It was expected that *Megaloolithus* eggs would exhibit similar intra- and inter- morphometric variability to reptiles because of the highly regional morphometric variation between the equator and poles of their eggshells (Schleich *et al.*, 1994; López-Martínez *et al.*, 2000; Panadès & Patnaik, 2009). Accordingly, we explored the two hypotheses to elucidate the taxonomic validity of *Megaloolithus* parataxa.

The Gaussian Mixture Model defined the six possible *Megaloolithus* parataxa used in the analyses as variations of the three *Megaloolithus* oospecies contained within them (Figure 3). The Gaussian Mixture Model also grouped the same sites under the same *Megaloolithus* oospecies from samples collected separately (Figure 3). Thus, hypothesis 2 is more plausible: the 6 *Megaloolithus* parataxa illustrate the polymorphism of a monotypic taxon and thus the three oospecies from the Late Cretaceous of Catalonia were laid by a single dinosaur species. In fact, most of Late Cretaceous sites exhibit one or two *Megaloolithus* parataxa (*e.g.* Zhao and Ding, 1976; Vianey-Liaud & Lopez-Martínez, 1997; Magalhães Ribeiro, 2002;

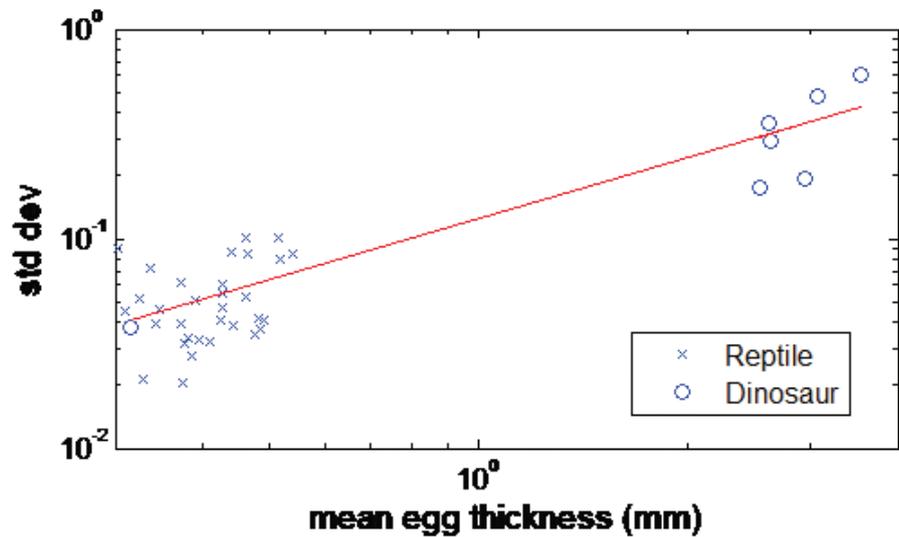


Figure 2. Linear regression graphs displaying the close relation between the variation of eggshell thickness between reptile and dinosaur eggs.

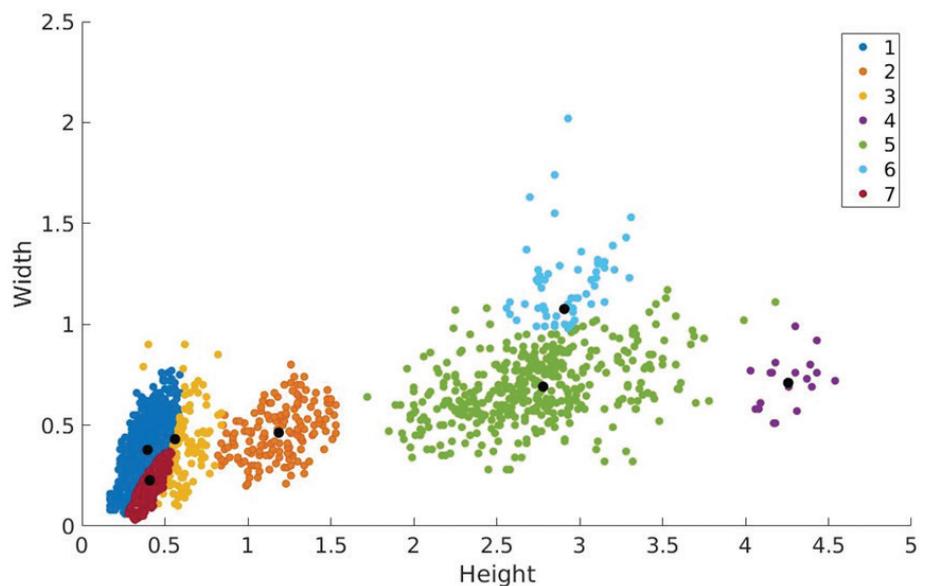


Figure 3. Representation of the eggshell data together with the centres identified by the Gaussian mixture model. The centres represented:

- 1) *Geochelone carbonata*
- 2) *Testudo hermanni*
- 3) *Crocodylus niloticus*
- 4) *Crocodylus porosus*
- 5) coll de Nargo
- 6) *Fumanya*
- 7) *Faidella*

Vianey-Liaud *et al.*, 2003; Chiappe *et al.*, 2005; Varricchio & Jackson, 2005; Panadès I Blas, 2005; Grellet-Tinner *et al.*, 2004; Grigorescu, 2016; Khosla & Lucas, 2020). If *Megaloolithus* parataxa were indicating polytypic dinosaur diversity as hypothesis 1 states, the Gaussian Mixture Model would have defined the six *Megaloolithus* parataxa as six distinct and statistically independent distributions from modern reptile oospecies. Thus, *Megaloolithus* oospecies described in Coll de Nargó, Fumanya, and Faidella are considered synonyms of the type species *M. mammillare* (Vianey-Liaud, 1994). This should not be extended to *Megaloolithus* parataxa in other localities in Catalonia and elsewhere. Their validity should be tested using statistical modelling with taxonomically constrained samples of modern and (if possible) fossil reptile oospecies. The validity and interpretations of parataxa established from limited data, statistical methods, and material, warrants caution.

However, there is a possibility that each *Megaloolithus* type was laid by two or more different dinosaur species. *Crocodylus acutus* and *C. moreletii* produced identical eggshells, while the species occur sympatrically and hybridise in coastal wetlands through the Caribbean region (Ray *et al.*, 2004; Cedeño-Vazquez *et al.*, 2006; Cedeño-Vazquez *et al.*, 2008). It may be plausible that different species of titanosaur and hadrosaur produced identical eggshells and cohabited along similar coastal regions. This cannot be assessed as only a very few solitary osseous remains have been found from the Late Cretaceous localities used here. Moreover, these have never been associated with any *Megaloolithus* parataxa (Vila *et al.*, 2012; Prieto-Márquez *et al.*, 2013).

Furthermore, we disagree with Peitz (2000) that *Megaloolithus* parataxa should be synonymised to the type species, to become valid biochronological markers. Changes in the structure through time in *Megaloolithus* can be mirroring environmental and faunal changes. In fact, the Gaussian Mixture Model detected that the thickness of *Megaloolithus* eggs significantly increase towards the end of the Cretaceous (Figure 3). This trend has been observed in localities that yield megaloolithic egg and eggshells (*e.g.* Vianey-Liaud *et al.*, 1994; 1997; Khosla & Lucas, 2020). Thickness of eggshell is determined by the resource quality or feeding behaviour of animals (Schaffner & Swart, 1991; Robert & Ball, 1998; Senkoylu *et al.*, 2005). However, Sellés *et al.* (2017) disputed that European titanosaurs altered their dietary behaviour, food type or quality of feeding resource at the end of the Cretaceous. Instead the authors suggest that reproductive stress as a result of ecological competition between dinosaurs led to thicker, probably pathological eggshells (Sellés *et al.*, 2017). Thicker eggshells in reptiles decreases gas exchange through the eggshell membranes and thus increases the chance of suffocation of the embryo (Wink *et al.*, 1994; Andrews, 2004). This may have produced high embryo mortality in *Megaloolithus* eggs and may explain the lack of embryonic remains. This pathology is common and may have contributed to the extinction of titanosaurid dinosaurs at the end of the Cretaceous. Entire nests of *Megaloolithus* eggs were unearthed empty in Portugal (Ribeiro *et al.*, 2014). The eggs were either not fertilized or the embryos died before or during egg burial (Ribeiro *et al.*, 2014). Further research, however, is required to ascertain the causes and possible consequences of an increase of eggshell thickness.

CONCLUSION

This study shows that Gaussian mixture models are a valid tool in parataxonomic classification of fossil eggs. The Gaussian mixture model has optimally quantified the polymorphism of the eggshell characters (HU and WU) of *Megaloolithus* parataxa from the Late Cretaceous of Catalonia (Iberia). The model has recognised the variability of taxonomically constrained samples of modern reptile oospecies

and determined that previously established *Megaloolithus* parataxa are variations of eggs laid by the same dinosaur species in three Catalan localities. The *Megaloolithus* opecies delimited by the model are considered taxonomically valid although they cannot be assigned to a dinosaur species. Further parataxonomic analyses using a Gaussian mixture model may determine the taxonomic validity of *Megaloolithus* parataxa from other geographical areas.

SOFTWARE

Gaussian mixture model functions from the Netlab toolbox is available from: <http://www.ncrg.aston.ac.uk/netlab/>.

ImageJ is available from: <http://imagej.nih.gov/ij/>.

ACKNOWLEDGMENTS

This study is dedicated to the memory of Prof. Nieves López Martínez (Universidad Complutense de Madrid). The authors would like to thank Dr Femke Holwerda (Royal Tyrrell Museum of Palaeontology) for the very useful reviews and great tips to improve the manuscript. Equally, the authors are very grateful to Dr Àngel Galobart from the ICP (Institut Català de Paleontologia Miquel Crusafont) for access to material, access to the Late Cretaceous sites, and constant support over the last few years. The following institutions are thanked for having provided oological material for the study: Bristol Zoo, La Ferme aux Crocodiles Zoo and the Koorana Crocodile Farm.

Xavier Panadès I Blas would like to thank Prof. Paul Buckland (BugsCEP) for his encouragement to continue with research; Dr Bernat Vila (ICP), Dr Albert Prieto-Marquez (ICP), and Dr Albert Sellés (ICP) for comments and discussions about dinosaur nesting palaeobiology; Dr Lluís Ardèvol i Oró (GeoPlay) for providing support for fieldwork around the Tremp basin and his help in understanding the Late Cretaceous geology of the Catalan Pyrenees; and Prof. Mike Benton (University of Bristol) for instilling in me that all contributions to biology are fundamental to understanding the history of life.

BIBLIOGRAPHY

- Andrews, R. 2004. Patterns of embryonic development. In: Deeming D.C. Ed. Reptilian incubation: environment, evolution and behaviour. – Nottingham, University Press: 75-102.
- Carpenter, K. 1999. Eggs, nests and baby dinosaurs. – Indiana, Indiana University Express.
- Cedeño-Vázquez, J.R., D. Rodriguez, S. Calmé, J.P. Ross, Densmore, IL. D. & J.B. Thorbjarnarson. 2008. Hybridization between *Crocodylus acutus* and *Crocodylus moreletii* in the Yucatán Peninsula: I. Evidence from mitochondrial DNA and morphology. – Journal of Experimental Zoology. Part A, Ecological Genetics and Physiology 309(A): 661-673.
- Cedeño-Vázquez, J.R., J.P. Ross, & S. Calmé. 2006. Population status and distribution of *Crocodylus acutus* and *Crocodylus moreletii* in South-eastern Quintana Roo, Mexico. – Herpetological Natural History 10; 53-66.
- Chiappe, L.M., R.A. Coria, L. Dingus, F.F. Jackson, A. Chinsamy & M. Fox. 1998. Sauropod dinosaur embryos from the Late Cretaceous of Patagonia. – Nature 396: 258-261.
- Chiappe, L.M., F. Jackson, R.A. Coria & L. Dingus. 2005. Nesting titanosaurs from Auca Mahuevo and adjacent sites: understanding sauropod reproductive behavior and embryonic development. In: Rogers K.C. & J. Wilson. Eds. The sauropods: evolution and paleobiology. – Berkeley and Los Angeles: University of California Press: 285-302.

- Elez, J. & N. López-Martínez. 2000. Interrelationships between growth of mineral phase and pore system in dinosaur eggshells. In: Bravo A.M. & T. Reyes. Eds. First international symposium on dinosaur eggs and babies, Sept 21-24. – Catalonia (Spain), Diputació de Lleida: 43-50.
- Fasano, G. & A. Franceschini. 1987. A multidimensional version of the Kolmogorov-Smirnov test. – Notices of the Royal Astronomical Society in astronomy 225: 155-170.
- García, G. & M. Vianey-Liaud. 2001a. Dinosaur eggshells as bio-chronological markers in Upper Cretaceous continental deposits. – Palaeogeography, Palaeoclimatology, Palaeoecology 169: 153-164.
- García, G. & M. Vianey-Liaud. 2001b. Nouvelles données sur les coquilles d'oeufs de dinosaures Megaloolithidae du Sud de la France: systématique et variabilité intraspecificue. – Comptes Rendus de l'Académie des Sciences Serie II Fascicule A-Sciences De La Terre et Des Planètes 332: 185-191.
- Grellet-Tinner, G., L. Chiappe & R. Coria. 2004. Eggs of titanosaurid sauropods from the Upper Cretaceous of Auca Mahuevo (Argentina). – Canadian Journal of Earth Sciences 41: 949-960.
- Grellet-Tinner, G. 2000. Phylogenetic interpretation of eggs and eggshells. In: Bravo A.M. & T. Reyes. First international symposium on dinosaur eggs and babies, Sept 21-24. – Catalonia (Spain), Diputació de Lleida: 61-75.
- Grigorescu, D. 2016. The 'Tuştea puzzle' revisited: Late Cretaceous (Maastrichtian) *Megaloolithus* eggs associated with *Telmatosaurus* hatchlings in the Haţeg Basin. – Historical Biology 29(5): 1-14
- Huynen, L., B.J. Gill, C.D. Millar & D.M. Lambert. 2010. Ancient DNA reveals extreme egg morphology and nesting behaviour in New Zealand's extinct moa. – Proceedings of the National Academy of Sciences 107(37): 16201-16206.
- Khosla, A. & S.G. Lucas. 2020. Introduction of Indian Late Cretaceous dinosaur eggs and eggshells of Peninsular India. In: Khosla A. & S.G. Lucas. Eds. Late Cretaceous dinosaur eggs and eggshells of Peninsular India. – Cham, Springer Nature Switzerland: 1-30.
- Lau, B. 2018. Matlab code for testing for differences between two multidimensional distributions (2-d K-S test, n-d energy test, Hotelling T^2 test). – Downloaded from <https://github.com/brian-lau/multdist> 20/01/2018
- López-Martínez, N., J.J. Moratalla & J.L. Sanz. 2000. Dinosaurs nesting on tidal flats. – Palaeogeography, Palaeoclimatology, Palaeoecology 160(1): 153-163.
- López-Martínez, N. 1997. Late Cretaceous dinosaur eggshells from the Tremp Basin, Southern Pyrenees, Lleida, Spain. – Journal of Paleontology 71 (6): 1157-1171.
- Magalhães Ribeiro, C.M. 2002. Ovo e fragmentos de cascas de ovos de dinossauros, provenientes de região de Peirópolis, Uberaba, Minas Gerais. – Arquivos do Museu Nacional, 60: 223-228.
- Mikhailov, K.E. 1991. Classification of fossil eggshells of amniotic vertebrates. – Acta Palaeontologica Polonica 36: 193-238.
- Mikhailov, K.E. 1997a. Avian eggshells: an atlas of scanning electron micrographs. – Tring, British Ornithologists' Club.
- Mikhailov, K.E. 1997b. Fossil and recent eggshell in amniotic vertebrates: fine structure, comparative morphology and classification. – Special Papers in Palaeontology 56: 1-80.
- Mikhailov, K.E., E.S. Bray & K.F. Hirsch. 1996. Parataxonomy of fossil egg remains (Veterovata): basic principles and applications. – Journal of Vertebrate Paleontology 16(4) : 763-769.
- Nabney, I. & C. Bishop. 2004. Netlab toolbox. – Downloaded from <http://www.ncrg.aston.ac.uk/netlab/>.
- Nannen, V. 2003. The Paradox of Overfitting. – Unpublished MA thesis, Rijksuniversiteit Groningen.

- Oskam, C.L., J. Haile, E. Mclay, P. Rigby, M.E. Allentoft, M.E. Olsen, C. Bengtsson, G.H. Miller, J.L. Schwenninger, C. Jacomb, R. Walter, B. A., J. Dortch, M. Parker-Pearson, M.T.P. Gilbert, R.N. Holdaway, E. Willerslev & M. Bunce. 2010. Fossil avian eggshell preserves ancient DNA. – *Proceedings of the Royal Society B-Biological Sciences* 277(1690): 1991-2000.
- Panadès I Blas, X. 2005. Diversity versus variability in Megaloolithid dinosaur eggshells. – *PalArch's Journal of Vertebrate Palaeontology* 2(1): 1-15.
- Panadès I Blas, X. & P. Rajeev. 2009. A complete crocodylian egg from the Upper Miocene (Chinji Beds) of Pakistan and its palaeobiographical implications. – *alArch's Journal of Vertebrate Palaeontology* 6(1):1-8.
- Peacock, J.A. 1983. Two-dimensional goodness-of-fit testing in astronomy. – *Monthly Notices of the Royal Astronomical Society* 202: 615-627.
- Peitz, C. 2000. Megaloolithid dinosaur eggs from the Maastrichtian of Catalunya (NE-Spain). Parataxonimic implications and stratigraphic utility. In: Bravo A.M. & T. Reyes. First international symposium on dinosaur eggs and babies, Sept 21-24. – Catalonia (Spain), Diputació de Lleida: 155-160.
- Prieto-Márquez, A., F.M. Dalla Vecchia, R. Gaete & À. Galobart. 2013. Diversity, relationships, and biogeography of the Lambeosaurine dinosaurs from the European Archipelago, with description of the new Aralosaurin *Canardia garonnensis*. – *PLoS ONE* 8(7) : e69835
- Rasband, W.S. 1997-2012. ImageJ, U. S. National Institutes of Health. – Downloaded from <http://imagej.nih.gov/ij/>.
- Ray, D.A., J.A. Dever, S.G. Plat, T.R. Rainwater, A.G. Finger, S.T. McMurry, M.A. Batzer, B. Barr, P.J. Stafford, J. Mcknight & L.D. Densmore. 2004. Low levels of nucleotide diversity in *Crocodylus moreletii* and evidence of hybridization with *C. acutus*. – *Conservation Genetics* 5: 449-462.
- Reynolds, D.A. 2008. Gaussian mixture models. *Encyclopaedia of biometric recognition*. – New York, Springer.
- Ribeiro, V., O. Mateus, F. Holwerda, R. Araújo & R. Castanhinha. 2013. Two new theropod egg sites from the Late Jurassic Lourinhã Formation, Portugal. – *Historical Biology* 26 (2): 206-217.
- Robert, R.J. & M.J. Ball. 1998. Egg shell quality problem: causes and solutions. – Armidale, University of New England.
- Sato, T., Y. Cheng, X. Wu, D.K. Zelenitsky & Y. Hsiao. 2005. A pair of shelled eggs inside a female dinosaur. – *Science* 308: 375.
- Schaffner, F.C. & P.K. Swart. 1991. Influence of diet and environmental water on the carbon and oxygen isotopic signatures of seabird eggshell carbonate. – *Bulletin of Marine Science* 48: 23-38.
- Schleich, H.H. & W. Kastle. 1988. Reptile egg-shells SEM Atlas. – Stuttgart/ New York, Fischer.
- Schleich, H.H., P. Gassner & T.M. Maskey. 1994. Ultrastrukturen und Variabilität de eishalen von *Gavialis Gangeticus* (Gmelin, 1789) (Crocodylia: Gavialidae). – *CFS Courier Forschungsinstitut Senckenberg* 173: 299-313.
- Sellés, A.G., A.M. Bravo, X. Delclòs, F. Colombo, X. Martí, J. Ortega-Blanco, C. Parellada & À. Galobart. 2013. Dinosaur eggs in the Upper Cretaceous of the Coll de Nargó area Lleida Province, south-central Pyrenees, Spain: oodiversity, biostratigraphy and their implications. – *Cretaceous Research* 40: 10-20.
- Sellés, A., B. Vila & A. Galobart. 2017. Evidence of reproductive stress in Titanosaurian sauropods triggered by an increase in ecological competition. – *Scientific Reports* 7: 13827.
- Senkoylu, N., H. Akyurek, H.E. Ersin, Samli & A. Agma. 2005. Assessment the impacts of dietary electrolyte balance levels on laying performance of commercial white layers. – *Pakistan Journal of Nutrition* 4: 423-427.

- Smyth, J.D., Olson, K. & M. Stange. 2009. Within-household selection methods: a critical review and experimental examination. In: Lavrakas, P., M. Trautgott, C. Kennedy, A. Holbrook, E. de Leeuw & W. Brady. Eds. *Experimental methods in survey research: techniques that combine random sampling with random assignment*. – New Jersey, John Wiley & Sons: 23-42.
- Varricchio, D.J., J.R. Horner & F.D. Jackson. 2002. Embryos and eggs for the Cretaceous theropod dinosaur *Troodon formosus*. – *The Journal of Vertebrate Paleontology* 22: 564-576.
- Varricchio, D.J. & F.D. Jackson. 2005. A phylogenetic assessment of prismatic dinosaur eggs from the Cretaceous Two Medicine Formation of Montana. – *Journal of Vertebrate Paleontology* 24(4) : 931-937.
- Vianey-Liaud, M., P. Mallan, O. Buscail & C. Montgelard. 1994. Review of French dinosaur eggshells: morphology, structure, mineral, and organic composition. In: Carpenter K. & Hirsch K. F. & Horner J.R. *Dinosaur eggs and babies*. – New York, Cambridge University Press: 151-183.
- Vianey-Liaud, M., K.F. Hirsch, A. Sahni & B. Sigé. 1997. Late Cretaceous Peruvian eggshells and their relationships with Laurasian and Eastern Gondwanian material. – *Geobios* 30: 75-90.
- Vianey-Liaud, M. & N. López Martínez. 1997. Late Cretaceous dinosaur eggshells from the Tremp Basin, southern Pyrenees, Lleida, Spain. – *Journal of Palaeontology* 71(6): 1157-1171.
- Vianey-Liaud, M. & D. Zelenitsky. 2003. Historical and new perspectives on the parataxonomy of fossil eggs. – *Palaeovertebrata* 32: 189-195.
- Vianey-Liaud, M., A. Khosla & G. García. 2003. Relationships between European and Indian dinosaur eggs and eggshells of the oofamily Megaloolithidae. – *The Journal of Vertebrate Paleontology* 23: 575-585.
- Vila, B., Jackson, F.D., J. Fortuny, A. Sellés & À. Galobart. 2010. D-modelling of Megaloolithid clutches: insights about nest construction and dinosaur behaviour. – *PLoS ONE* 5(5) : e10362.
- Vila, B., Galobart À., J.I. Canudo, J. Le Loeuff, J. Dinarès-Turell, V. Riera, O. Oms, T. Tortosa & R. Gaete. 2012. The diversity of sauropod dinosaurs and their first taxonomic succession from the latest Cretaceous of southwestern Europe: clues to demise and extinction. – *Palaeogeography, Palaeoclimatology, Palaeoecology* 350-352: 19-38.
- Wink, C.S., R.M. Elsey & M. Bouvier. 1990. Porosity of eggshells from wild and captive, pen-reared alligators (*Alligator mississippiensis*). – *Journal of Morphology* 203:35-39.
- Zhao, Z. & S. Ding. 1976. Discovery of the dinosaurian eggshells from Alxa, Ningxia and its stratigraphic significance. – *Vertebrata Palasiatica* 14: 42-44.

Submitted: 18 August 2020

Accepted for Publication (with revisions): 18 September 2020

Resubmitted: 1 February 2021

Published: 25 May 2021

© 2021 The authors. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.