Panades I Blas, X., Baddeley, R. J., & Fennell, J. G. (2021). Testing the Utility of Gaussian Mixture Models in Parataxonomic Classification of Fossil Eggs. *PalArch's Journal of Vertebrate Palaeontology*, *17*(3), 1-13. https://archives.palarch.nl/index.php/jv

Peer reviewed version

Testing the utility of gaussian mixture models in parataxonomic classification of fossil eggs

1

2

3

6 **Summary**

7 Fossil eggs and eggshells are very seldom associated with skeletal or dermal remains and can

8 rarely be assigned taxonomically. This has resulted in oologists creating an alternative

9 classification system for fossil eggs and eggshells: the parataxonomy of fossil eggs. Under

10 parataxonomy, fossil eggs are classified in oospecies, oogenera and oofamilies according to

11 qualitative characters (e.g. shape of the shell units and ornamentation) and highly heritable

12 quantitative characters (e.g. height of unit, HU, and width of unit, WU). *Megaloolithus*

13 oospecies, which have previously been partially attributed to both titanosaurian

14 and hadrosaurid dinosaurs. These oospecies are defined by are defined by the shape of their

15 units, quantification of the morphometric variation and variability of their microstructures

16 (HU and WU of eggshell units). Two competing interpretations of the observed

17 morphological variation of the eggshell units have been proposed: 1) different megaloolithic

18 morphologies are indicative of different dinosaur species; and 2) the same dinosaur species

19 was responsible for all the variation seen in megaloolithic eggshell units. In this study, a

20 Gaussian mixture model was applied to test both interpretations. This probabilistic model

21 assumes all the data points are generated from a mixture of a finite number of Gaussian

22 distributions with unknown parameters. We compared the morphometric distributions of HU

23 and WU in eggshells, eggs and clutches belonging to 7 *Megaloolithus* oospecies from

24 Catalonia to the defined morphometric variation seen in the eggshell microstructures of two

25 extant turtle species and two crocodile species. The resulting Gaussian model was best

1

26   defined for four distributions of HU and WU representing the extant turtle and crocodile

27   oospecies, and three distributions consisting of one or more *Megaloolithus* oospecies. Thus,

28   *Megaloolithus* oospecies from the Late Cretaceous of Catalonia depict the polymorphism of a

29   monotypic taxon and were laid by a single dinosaur species.

1. Introduction

31   Parataxonomy creates a systematic taxonomic framework for oological material that cannot

32   be directly or indirectly associated to osseous and embryonic remains *in ovo* and in *utero*,

33   eggs and eggshells *in utero*, DNA, and correlations of bones and palaeooölogical material

34   from independent deposits (Varricchio *et al.* 2002; Sato *et al.* 2005; Oskam *et al.* 2010;

35   Huynen *et al.* 2010).

36   Parataxonomy defines oospecies via quantification of the polymorphism of egg

37   characters (eggshell thickness, egg size, details of ornamentation, and external pore patterns)

38   (Mikhailov 1991, 1997a, 1997b; Mikhailov *et al.,* 1996). For example, the dinosaurian

39   megaloolithid oopecies have been defined by the intra- (between eggs and clutches) and inter-

40   specific (oospecies) morphometric variability of the height and width of their units (Vianey-

41   Liaud and Lopez-Martinez, 1997; Vianey-Liaud *et al.,* 2003; Panadès I Blas, 2005).

42   Megaloolithid oospecies have been used to estimate the diversity of dinosaurs during

43   the Late Jurassic as well as the Late Cretaceous in areas where eggshell material is more

44   abundant than diagnostic skeletal remains (e.g. Zhao and Ding, 1976; Vianey-Liaud and

45   Lopez-Martinez, 1997; Vianey-Liaud *et al.,* 2003; Panadès I Blas, 2005; …..). However, the

46   role of megaloolithid palaeooöospecies as indicators of nesting palaeobiodiversity is debated,

47   because palaeoologists disagree with the taxonomic validity of megaloolithid

48   palaeooöospecies (Vianey-Liaud and Lopez-Martinez, 1997; Peitz, 2000; Vianey-Liaud *et al,.*

49   2003; Panadès I Blas, 2005). Two hypotheses have been proposed to explain the taxonomic

2

50   validity of megaloolithid palaeooöpecies (Vianey-Liaud and Lopez-Martinez, 1997; Peitz,

51   2000; Vianey-Liaud *et al.,* 2003; Panadès I Blas, 2005):

52   1. Each megaloolithid palaeooöspecies is defined by dimensional variations of the

53   same eggshell's fan-shape unit, and represents dinosaur polytypic diversity. Therefore, the

54   diversity of megaloolithid palaeooöspecies reflects the diversity of dinosaur

55   palaeobiodiversity (Vianey-Liaud and Lopez-Martinez, 1997; Vianey-Liaud *et al.,* 2003;

56   Panadès I Blas, 2005).

57   2. Or, megaloolithid palaeooötaxa are characterised by different variations of the same

58   eggshell's fan-shape unit from different areas of the eggs, and describes the polymorphism of

59   a monotypic taxon, and consequently are not useful indicators of palaeobiodiversity (Peitz,

60   2000).

61   Here, we test both hypotheses using a Gaussian Mixture Model. We argued that we

62   can solve the taxonomic validity of *Megaloolithus* oospecies, by comparing the morphometric

63   variability of their eggshells to the variability of oospecies of defined layers such as modern

64   turtle and crocodile oospecies, using a Gaussian Mixture Model.

65   Gaussian mixture models predict the probability of distributions (set of points) to

66   belong to a set distribution of a classes (Reynolds, 2015). The Gaussian mixture model will

67   determine whether the morphometric variability of *Megaloolithus* oospecies is equivalent to

68   the variability of modern reptile oospecies, or no.

69

70   **2. Materials**

71   Megaloolithid and modern reptile eggs and eggshells, and data of megaloolithid material from

72   previous studies were used in this project. Three nests were sampled from Catalan localities: a

73   nest of two eggs from Coll de Nargó (Na) number IPS-33141, a nest of three eggs from

74   Fumanya (Fu) number IPS-27380, a nest with three eggs from Faidella (Fai) number IPS-

3

75 36341; and the fossil turtle nest from El Canyet number IPS-36381. The three nests are stored

76 at the Institut Català de Paleontologia Miquel Crusafont (Sabadell, Catalonia). Data of

77 megaloolithid eggshells from the Catalan localities Suterranya are from Panadès I Blas, 2005;

78 Basturs and Coll de Nargó from Peitz, 2000; and Biscarri clutch from López-Martínez *et al.,*

79 2000.

80   Reptilian nests of extant species of turtle and crocodile were used as comparative

81 material for the analysis, because the structure of their eggshells is the most similar to

82 megaloolithid palaeoogenera (Mikhailov, 1991, 1997a, 1997b; Mikhailov *et al.,* 1996).

83 Megaloolithid, crocodile, and turtle eggshells are composed of discreet units along the

84 eggshell thickness that extend from the interior to the exterior of the thickness of the shell.

85   The eggs used were laid by the same female, and are of two different species of turtle

86 and crocodile, to capture the regional intraspecific variation of HU and WU exhibited by

87 reptile and megaloolithid oospecies (Mikhailov, 1991, 1997a, 1997b; Mikhailov *et al.,* 1996;

88 Vianey-Liaud and López Martínez, 1997). The materials used were two clutches with three

89 and two eggs of *Geochelone carbonaria* (Gc), two clutches with three eggs each of *Testudo*

90 *hermanni* (Th), donated by Bristol Zoo; four clutches with three eggs each of *Crocodylus*

91 *nilocitus* (Cn) from La Ferme aux Crocodiles Zoo (France), four clutches with three eggs of

92 *C. porosus* (Cp) obtained from the Koorana Crocodile Farm (Queensland, Australia).

93   Megaloolithid and modern reptile eggs and eggshells, and data of megaloolithid

94 material from previous studies were used in this project. Three nests were sampled from

95 Catalan localities: a nest of two eggs from Coll de Nargó (Na) number IPS-33141, a nest of

96 three eggs from Fumanya (Fu) number IPS-27380, a nest with three eggs from Faidella (Fai)

97 number IPS-36341. The nests are stored at the Institut Català de Paleontologia Miquel

98 Crusafont (Sabadell, Catalonia). Data of megaloolithid eggshells from the Catalan localities

4

99   Suterranya (from Panadès I Blas, 2001); Basturs and Coll de Nargó (from Peitz, 2000);

100   Biscarri clutch (from López-Martínez *et al.,* 2000) were computed in the analyses.

101   Seven *Megaloolithus* oospecies were represented in the material: *Megaloolithus*

102   *baghesis* (Coll de Nargó), *M. mamillare* (Coll de Nargó, Basturs), *M. siruguei* (Biscarri, Coll

103   de Nargó, Faidella, and Fumanya), and *M. pseudomamillare, M. cf. mamillare, M. trempii,*

104   and unknown oospecies (Suterranya) (Vianey-Liaud and López Martínez, 1997; Panadès I

105   Blas, 2005; Vila *et al.,* 2010, 2011). The three nests have never been assigned

106   parataxonomically to a megaloolithid oospecies, but they exhibit the typical characters of *M.*

107   *siruguei*: subspherical eggs of 20×18.5 cm; eggshell thickness ranging between 1.75 and 3.6

108   mm; elongated fan-shaped units five times higher than wide exhibiting intermittent fusion;

109   arched growth lines running from the base to the top of the units and crossing to adjacent

110   units (Figure 1), (Vianey-Liaud *et al.,* 1994; Elez and López-Martínez, 2000; Vianey-Liaud

111   and Zelenitsky, 2003; Sellés *et al.,* 2013).

112   Reptilian nests of extant species of turtles and crocodile were used as comparative

113   material for the analysis. The eggshells units of turtles and crocodile and megaloolithid

114   eggshells exhibit a similar biomineral structure: a series of discreet units along the eggshell

115   thickness that extend from the interior to the exterior of the thickness of the shell (Mikhailov,

116   1991, 1997a, 1997b; Mikhailov *et al.,* 1996).

117   The eggs used were laid by the same female, of two different species of turtles and

118   crocodile eggs, to capture the regional intraspecific variation of HU and WU exhibited by

119   reptile and megaloolithid oospecies (Mikhailov, 1991, 1997a, 1997b; Mikhailov *et al.,* 1996;

120   López-Martínez, 1997). The materials used were two clutches with three and two eggs of *G.*

121   *carbonaria* (Gc)*,* two clutches with three eggs each of *T. hermanni* (Th)*,* donated by Bristol

122   Zoo; four clutches with three eggs each of *C. nilocitus* (Cn) from La Ferme aux Crocodiles

5

123   Zoo (France)*,* four clutches with three eggs of *C. porosus* (Cp) obtained from the Koorana

124   Crocodile Farm (Queensland, Australia).

125   **3. Methods**

126   Parataxonomy defines oospecies via the quantification of the diversity/variability of egg

127   characters: height (HU) and width (WU) of units, egg size, details of ornamentation, and

128   external pore patterns (Mikhailov, 1991, 1997a, 1997b; Mikhailov *et al.,* 1996). Only well-

129   defined and consistent ranges of these characters should be applied for taxonomic inference

130   (Mikhailov, 1991, 1997a, 1997b; Mikhailov *et al,.* 1996). Measurements of HU and width

131   WU of units will exclusively be used here because these are the most widely used in

132   parataxonomic analysis (Mikhailov, 1991, 1997a, 1997b; Mikhailov *et al,.* 1996; Schleich *et*

133   *al.,* 1994; López-Martínez *et al.,* 2000).

134   The intraspecific variation of HU and WU in reptile and megaloolithid oospecies is

135   regional (equator and poles) in eggs and highly variable in eggs and clutches (Mikhailov,

136   1991, 1997a, 1997b; Mikhailov *et al.,* 1996; Schleich *et al.,* 1994; López-Martínez *et al.,*

137   2000). For example, Schleich *et al.* (1994) and Panadès and Patnaik (2009) found that modern

138   and fossil crocodile eggs are thicker in the equatorial region than in the poles, while López-

139   Martínez *et al.* (2000) found that a single egg in a megaloolithid nest was thicker in the poles

140   and thinner in the equator. However, none of the studies established whether morphological

141   and morphometrical changes of the eggshells units are correlated.

142   Accordingly, two *Megaloolithus* and reptile eggshells from the equator and the poles

143   from each egg were sampled, prepared in petrographic thin sections, and photographed under

144   a Polarised Light Microscope (PLM) (Carpenter, 1999; García and Vianey-Liaud, 2001a,

145   2001b; Grellet-Tinner *et al.,* 2000). Unlike the ovoidal eggs of megaloolithic layers and

146   crocodiles, the poles and equator in the spherical eggs of turtles are not distinct. Two crossing

147   circles were drawn around the middle of the eggs of turtles and the poles were marked on the

6

148   opposite sites of one of the circles. The two eggshells were extracted, inspected under a light

149   microscope for any pathologies. Eggshells with double layering, diagenetic alterations and/or

150   eggshells without completed units were discarded (Carpenter, 1999; García and Vianey-

151   Liaud, 2001a, 2001b; Grellet-Tinner *et al.,* 2010). Thin sections were deposited at the Museu

152   de la Conca Dellà (Catalonia).

153   HU and WU from the each eggshell units were captured digitally using Image J 1.47a

154   from photographs of calibrated petrographic thin sections (Rasband, 1997-2012). HU was

155   measured from the base to the highest point in a straight vertical line; and WU, from the

156   widest straight horizontal line of the unit under the PLM. This method is more accurate than

157   measuring the eggshells using a calliper and under S.E.M. (Scanning Electron Microscope)

158   and may produce distorted data (Mikhailov, 1991, 1997a, 1997b; Mikhailov *et al.,* 1996;

159   López-Martínez, 1997; Carpenter, 1999). The microstructures of eggshells are more

160   distinguishable under the Polarised Light Microscope (PLM), as it provides realistic

161   bidimensional visualisations of the morphometrical variations of the units under a perfect 90°

162   angle (Mikhailov, 1991, 1997a, 1997b; Mikhailov *et al.,* 1996; López-Martínez, 1997;

163   Carpenter, 1999).

164   Taking measurements of eggshells on the field using a calliper and from S.E.M

165   micrographs is imprecise. While the structure of the eggshells cannot be visualised, and

166   specimens diagenetically altered, over- and/or underlay by sediment, and pathologies could be

167   undetected; obtaining eggshells measurements is also inaccurate: the specimens are not in a

168   straight 90° angle and the protruding surfaces distort the real measurements (Mikhailov, 1991,

169   1997a, 1997b; Mikhailov *et al.,* 1996; López-Martínez, 1997; Carpenter, 1999).

170   Previous quantifications of megaloolithid oospecies were limited to small samples that

171   were taxonomically constraint, and they did not test independently the relations of the HU and

172   WU distributions (e. g. García and Vianey-Liaud, 2001a and b; Vianey-Liaud and López-

7

173 Martínez, 1997; Panadès I Blas, 2005). In order to investigate how the measurements might

174 be related, without the site information (eggshells, eggs, and localities), a Gaussian mixture

175 model was used to establish the most probable number of distributions of megaloolithic HU

176 and WU distributions using the taxonomically constrained samples of modern reptiles.

177 A Gaussian mixture model is a probabilistic model that assumes that data points are

178 generated from a mixture of a finite number of Gaussian distributions with unknown

179 parameters (Reynolds, 2008). Fitting the best mixture of Gaussians for a given dataset (as

180 measured by the log likelihood) results in a probability distribution of classes that can be used

181 to predict the probability (posterior) of new data points belonging to those classes (Reynolds,

182 2008). Fitting Gaussian mixture models is an example of an unsupervised learning method,

183 however, the computing required for fitting a mixture of Gaussians is exponential for the

184 number of latent Gaussian distributions, so approximate inference techniques are often used

185 (Reynolds, 2008). While this does not guarantee *the* optimal solution, models do converge

186 quickly to a local optimum (Reynolds, 2008).

187 To improve the quality it is usual to fit many of these models and choose the model

188 that best fits the data, often on the basis of log likelihood or similar approach (Reynolds,

189 2008). Here the Gaussian mixture model functions from the Netlab toolbox (Nabney and

190 Bishop, 2004) are used: these functions initialise the model using a clustering process known

191 as *k-means* and then use the expectation maximisation (EM) algorithm. EM is an iterative

192 method for obtaining maximum likelihood estimates of parameters for models that depend on

193 unobserved variables: in the present case a finite number of Gaussian distributions

194 representing the relationship between measurements irrespective of collection site (Nabney

195 and Bishop, 2004).

196 Mixture models were produced for 2 to 10 centres, each repeated 10 times. Each

197 centre represented HU and WU distributions of modern reptiles and megaloolithic eggshells.

198   From the 10 repetitions for each of the models, the best log likelihood was obtained. From

199   these, as shown in Equation 1 (Nannen, 2003, p. 14), the minimum description length (or the

200   best number of distributions accounting for the measurements), L calculated, was used to

201   choose the number of centres that best represented the data. Here, D is the data set, M_k is the

202   model, k is the number of Gaussians or centres and N is the number of points in the data set.

203

204   (1)

205

206   A between groups Kolmogorov-Smirnov test was then carried out with the

207   measurements of thickness and width as the dependent variables and centre as the factor in

208   order to investigate the difference in thickness and width for each of the centres found by the

209   mixture model. The classical one-dimensional Kolmogorov-Smirnov test is a non-parametric

210   statistic for comparing two empirical distributions, which defines the largest absolute

211   difference between the two cumulative distribution functions as a measure of disagreement.

212   Adapting this test to more than one dimension is a challenge because there are 2d-1

213   independent ways of defining a cumulative distribution function when d dimensions are

214   involved.

215   We used a Matlab implementation (Lau, 2018) of a two dimensional Kolmogorov-

216   Smirnov test (Fasano and Franceschini, 1987; Peacock, 1983). Since we are doing multiple

217   comparisons, we need to adjust the alpha level according to the number of comparisons being

218   made. The simplest and most conservative correction is the Bonferroni correction, which is

219   satisfactory for present purposes:

220   Adjusted alpha level = alpha/n comparisons = 0.05/7 = 0.0071

221

222   **4. Results**

9

223   Comparison of the minimum description lengths for each of the Gaussian mixture models

224   revealed that a model with seven centres provided the best description of the data, suggesting

225   that the measurements could be considered to belong to seven distinct distributions (Figure 3).

226   The regression line is a very good fit to both modern reptile and the dinosaur affirming that

227   megaloolithid and modern reptile eggshells possess similar variability distributions, but it

228   scales with the bigger scale of megaloolithid eggs.

229   In fact, the variability of egg thickness among reptile and dinosaurs is very similar:

230   both are highly variable (coefficient of variation of about 20%). The mean and variability

231   (standard deviation) of dinosaur eggs are not different to reptiles, just bigger; the regression

232   line is a fit to both the reptile and dinosaur data (Figure 2).

233   The Kolmogorov-Smirnov test revealed a significant difference for both width and

234   thickness between all the centres (Tables 2 and 3).

235

236

237

238   **5. Discussion**

239

240   The Gaussian Mixture Model has defined 7 oospecies: 4 representing the extant turtle and

241   crocodile oospecies, and 3 representing *Megaloolithus* oospecies at the end of the Cretaceous

242   in Catalonia (Figure 3). The 3 *Megaloolithus* oospecies are considered taxonomically valid

243   because they are statistically different and between them and to the taxonomically constraint

244   samples of turtle and crocodile oospecies (see Figure 2, Tables 2 and 3). It was expected that

245   *Megaloolithus* eggs exhibit similar intra- and inter- morphometric variability to reptiles

246   because of the highly regional morphometric variation (equator and poles) of their eggshells

247   (Schleich *et al.*, 1994; López-Martínez *et al.*, 2000; Panadès and Patnaik, 2009). Accordingly,

10

248   we explored the two hypotheses to elucidate the taxonomic validity of *Megaloolithus*

249   parataxa.

250   The Gaussian Mixture Model defined the 7 *Megaloolithus* parataxa used in the

251   analyses as variations of the 3 *Megaloolithus* oospecies contained within them (Figure 3). The

252   Gaussian Mixture Model also grouped the same sites under the same *Megaloolithus* oospecies

253   from samples collected separately from (Figure 3). Thus, hypothesis 2 is more plausible: the 7

254   *Megaloolithus* parataxa illustrate the polymorphism of a monotypic taxon and thus the three

255   oospecies from the Late Cretaceous of Catalonia were laid by a single dinosaur species.

256   Otherwise, if *Megaloolithus* parataxa were indicating polytypic dinosaur diversity as

257   hypothesis 1 states, the Gaussian Mixture Model would have defined the 7 *Megaloolithus*

258   parataxa as seven distinct and statistically independent distributions to modern reptile

259   oospecies.

260   However, there may be a possibility that each *Megaloolithus* type was laid by two or

261   more different dinosaur species. *Crocodylus acutus* and *C. moreletii* produced identical

262   eggshells, while the species occur sympatrically and hybridise in coastal wetlands through the

263   Caribbean region (Ray *et al.*, 2004; Cedeño-Vazquez *et al.*, 2006; Cedeño-Vazquez *et al.*,

264   2008). It may be plausible that different species of titanosaur and hadrosaur produced

265   identical eggshells and interacted sympatrically along similar coastal regions. This cannot be

266   assessed as only a very few solitary osseous remains have been found from the Late

267   Cretaceous localities used here. Moreover, these have never been associated with any

268   *Megaloolithus* parataxa (Vila *et al.*, 2012; Prieto-Márquez *et al.*, 2013).

269   Furthermore, The Gaussian Mixture Model detected that the thickness of *Megaloolithus*

270   eggs significantly increase towards the end of the Cretaceous (Figure 3). Increased eggshell

271   thickness in reptiles decreases gas exchange through the eggshell membranes and thus

272   increases the chance of suffocation of the embryo (Wink *et al.*, 1994; Andrews, 2004). This

11

273 may have produced a high embryo mortality in *Megaloolithus* eggs, and may explain the lack

274 of embryonic remains. This can also have partially contributed to the extinction of

275 titanosaurid dinosaurs at the end of the Cretaceous in Catalonia.

276   Thickness of eggshell is determined by the resource quality or feeding behaviour of

277 animals (Schaffner and Swart, 1991; Robert and Ball, 1998; Senkoylu *et al.*, 2005). Recent

278 studies discarded that titanosaurian sauropods did not alter their dietary behaviours, food type

279 or quality of feeding resources of at the end of the Cretaceous in Catalonia (Sellés *et al.*,

280 2017). Thus, other environmental factors produced physiological alterations that affected their

281 reproductive cycle and dramatically increased the thickness of their eggshells. Further

282 research is required to ascertain the causes and possible consequences of an increase of

283 eggshell thickness in titanosaur dinosaurs at the end of the Cretaceous.

284 **6. Conclusion**

285 This study shows that Gaussian mixture models are a valid tool in parataxonomic

286 classification of fossil eggs. The Gaussian mixture model has optimally quantified the

287 polymorphism of the eggshell characters (HU and WU) of *Megaloolithus* parataxa from the

288 Late Cretaceous of Catalonia (Iberia). The model has recognised the variability of

289 taxonomically constraint samples of modern reptile oospecies and determined that previously

290 established *Megaloolithus* parataxa are variations of eggs laid by the same dinosaur species.

291 The *Megaloolithus* oopecies delimited by the model are considered taxonomically valid

292 although they cannot be assigned to a dinosaur species. Further parataxonomical analyses

293 using a Gaussian mixture model will determine the taxonomic validity of *Megaloolithus*

294 parataxa from other geographical areas.

295 **7. Software**

296 Gaussian mixture model functions from the Netlab toolbox is available from:

297 http://www.ncrg.aston.ac.uk/netlab/.

298

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

300

301 **Acknowledgments**

302 **Add acknowledgements if you have them**

303

304 **References**

305 CARPENTER K. (1999). *Eggs, nests and baby dinosaurs*. Indiana, Indiana University

306 Express.

307

308 CEDEÑO-VÁZQUEZ J. R., RODRIGUEZ D., CALMÉ S., ROSS J. P., DENSMORE III L.

309 D. AND THORBJARNARSON J. B. (2008). Hybridization between *Crocodylus acutus* and

310 *Crocodylus moreletii* in the Yucatán Peninsula: I. Evidence from mitochondrial DNA and

311 morphology. *Journal of experimental zoology. Part A, Ecological genetics and physiology*

312 **309(A)**, 661–673.

313

314 CEDEÑO-VÁZQUEZ J. R., ROSS J. P. AND CALMÉ S. (2006). Population status and

315 distribution of *Crocodylus acutus* and *Crocodylus moreletii* in South-eastern Quintana Roo,

316 Mexico. *Herpetological Natural History* **10**, 53–66.

317

318 DEEMING D. C. (2004). *Reptilian Incubation: Environment, Evolution and Behaviour.*

319 Nottingham, Nottingham University Press.

320

321 ELEZ J. AND LÓPEZ-MARTÍNEZ N. (2000). Interrelationships between growth of mineral

322 phase and pore system in dinosaur eggshells. In: *First International Symposium on Dinosaur*

323 *Eggs and Babies; Sept 21– 24 Isona. Catalonia (Spain).* pp. 43-50.

324

325 FASANO, G. AND FRANCESCHINI, A. (1987). A multidimensional version of the

326 Kolmorogov-Smirnov test. *Notices of the Royal Astronomical Society in astronomy* **225**,155-

327 170.

328

329 GARCÍA G. AND VIANEY–LIAUD M. (2001a). Dinosaur eggshells as bio-chronological

330 markers in Upper Cretaceous continental deposits. *Palaeogeography, Palaeoclimatology,*

331 *Palaeoecology***169**, 153–164.

332

333 GARCÍA G. AND VIANEY–LIAUD M. (2001b). Nouvelles donné sur les coquilles díoeufs

334 de dinosaures Megaloolithidae du Sud de la France: systématique et variabilité

335 intraspecifique. *Comptes Rendus de l'Academie des Sciences Serie II Fascicule A-Sciences*

336 *De La Terre et Des Planetes* **332**,185–191.

337

338 GRELLET-TINNER G. (2000). Phylogenetic interpretation of eggs and eggshells. In: *First*

339 *International Symposium on Dinosaur Eggs and Babies; Sept 21– 24 Isona. Catalonia*

340 *(Spain).* pp. 61-75.

341

342 HUYNEN L., GILL B. J., MILLAR C. D., AND LAMBERT D. M. (2010). Ancient DNA

343 reveals extreme egg morphology and nesting behaviour in New Zealand's extinct moa.

344 *Proceedings of the National Academy of Sciences* **107(37)**, 16201–16206.

345

14

346   LAU, B. (2018). *Matlab code for testing for differences between two multidimensional*

347   *distributions (2-d K-S test, n-d energy test, Hotelling T^2 test).* Downloaded from

348   https://github.com/brian-lau/multdist 20/01/2018

349

350   LÓPEZ-MARTÍNEZ N., MORATALLA J. J., SANZ J. L. (2000). Dinosaurs nesting on tidal

351   flats. *Palaeogeography, Palaeoclimatology, Palaeoecology* **169160**, 153–163.

352

353   MIKHAILOV K. E. (1991). Classification of fossil eggshells of amniotic vertebrates. Act

354   *Acta Palaeontologica Polonica* **36**, 193–238.

355

356   MIKHAILOV K. E. (1997a). *Avian Eggshells: an Atlas of Scanning Electron Micrographs*.

357   Tring, British Ornithologists' Club.

358

359   MIKHAILOV K. E. (1997b). Fossil and recent eggshell in amniotic vertebrates: Fine

360   structure, comparative morphology and classification. *Special Papers in Palaeontology* **56**, 1–

361   80.

362

363   MIKHAILOV K. E., BRAY E. S. AND HIRSCH K. F. (1996). Parataxonomy of fossil egg

364   remains (Veterovata): basic principles and applications. *Journal of Vertebrate Paleontology*

365   **16(4)**, 763–769.

366

367   NABNEY I. AND BISHOP C. (2004). *Netlab toolbox*. Downloaded from

368   http://www.ncrg.aston.ac.uk/netlab/.

369

370   NANNEN V. (2003). "*The Paradox of Overfitting" [master's thesis].* Groningen,

371   Rijksuniversiteit Groningen.

372

373   OSKAM C. L., HAILE J., MCLAY E., RIGBY P., ALLENTOFT M. E., OLSEN M. E.,

374   BENGTSSON C., MILLER G. H., SCHWENNINGER J.L., JACOMB C., WALTER R.,

375   BAYNES A., DORTCH J., PARKER-PEARSON M., GILBERT M. T. P., HOLDAWAY R.

376   N., WILLERSLEV E., AND BUNCE M. (2010). Fossil avian eggshell preserves ancient

377   DNA. *Proceedings of the Royal Society B-Biological Sciences* **277(1690)**, 1991-2000.

378

379   PANADÈS I BLAS X. (2005). Diversity versus Variability in Megaloolithid Dinosaur

380   Eggshells. *www.PalArch.nl Vertebrate Palaeontology* **2(1)**, 1–15

381

382   PANADÈS I BLAS X. AND RAJEEK P. (2009). A complete crocodylian egg from the

383   Upper Miocene (Chinji Beds) of Pakistan and its palaeobiographical implications.

384   *www.PalArch.nl Vertebrate Paleontology* **6(1)**, 1-8.

385

386   PEACOCK, J. A. (1983). Two-dimensional goodness-of-fit testing in astronomy.

387   *Monthly Notices of the Royal Astronomical Society* **202**, 615–627.

388

389   PEITZ C. (2000). Megaloolithid dinosaur eggs from the Maastrichtian of Catalunya (NE–

390   Spain) Parataxonimic implications and stratigraphic utility. In: *First International Symposium*

391   *on Dinosaur Eggs and Babies; Sept 21– 24 Isona. Catalonia (Spain).* pp. 155–160.

392

393   PRIETO-MÁRQUEZ A., DALLA VECCHIA F. M., GAETE R. AND GALOBART À.,

394   (2013). Diversity, relationships, and biogeography of the Lambeosaurine dinosaurs from the

395   European Archipelago, with description of the new Aralosaurin *Canardia garonnensis*.

396   *PLoS ONE* **8(7)**, e69835

397

398   RASBAND, W.S. (1997-2012). *ImageJ, U. S. National Institutes of Health*. Downloaded

399   from http://imagej.nih.gov/ij/. Maryland: USA.

400

401   RAY D. A., DEVER J. A., PLATT S. G., RAINWATER T. R., FINGER A. G., MCMURRY

402   ST, BATZER M. A., BARR B, STAFFORD P. J., MCKNIGHT J., AND DENSMORE L. D.

403   (2004). Low levels of nucleotide diversity in *Crocodylus moreletii* and evidence of

404   hybridization with *C. acutus*. *Conservation Genetics* **5**.449–462.

405

406   REYNOLDS D. A. (2008). *Gaussian Mixture Models*. *Encyclopaedia of Biometric*

407   *Recognition*. New York City, Springer.

408

409   ROBERT R. J. AND BALL M. J. (1998). *Egg Shell Quality Problem: Causes and Solutions*.

410   Armidale, University of New England.

411

412   SATO T., CHENG Y., WU X., ZELENITSKY D. K. AND HSIAO Y. (2005). A pair of

413   shelled eggs inside a female dinosaur. *Science* **308**, 375.

414

415   SCHAFFNER F. C. AND SWART P. K. (1991). Influence of diet and environmental water

416   on the carbon and oxygen isotopic signatures of seabird eggshell carbonate. *Bulletin of*

417   *Marine Science* **48**, 23–38.

17

418

419   SCHLEICH H. H. AND KASTLE W. (1988). *Reptile egg–shells SEM Atlas*. Stuttgart/New

420   York, Fischer.

421

422   SELLÉS A. G., BRAVO A. M., DELCLÒS X., COLOMBO F., MARTÍ X., ORTEGA-

423   BLANCO J., PARELLADA C., AND GALOBART À. (2013). Dinosaur eggs in the Upper

424   Cretaceous of the Coll de Nargó area Lleida Province, south-central Pyrenees, Spain:

425   oodiversity, biostratigraphy and their implications. *Cretaceous Research* **40**, 10–20.

426

427   SELLÉS A., VILA B., AND GALOBART A. (2017). Evidence of Reproductive Stress in

428   Titanosaurian Sauropods Triggered by an Increase in Ecological Competition. *Scientific*

429   *Reports* **7**, 1.

430

431   SENKOYLU N., AKYUREK H., ERSIN SAMLI H. E., AND AGMA, A., (2005).

432   Assessment the impacts of dietary electrolyte balance levels on laying performance of

433   commercial white layers. *Pakistan Journal of Nutrition* **4**,423–427.

434

435   VARRICCHIO D. J., HORNER J. R., AND JACKSON F. D. (2002). Embryos and eggs for

436   the Cretaceous theropod dinosaur *Troodon formosus*. *The Journal of Vertebrate Paleontology*

437   **22**, 564–576.

438

439   VIANEY-LIAUD M. AND LÓPEZ MARTÍNEZ N. (1997). Late Cretaceous dinosaur

440   eggshells from the Tremp Basin, southern Pyrenees, Lleida, Spain. *Journal of Palaeontology*.

441   **71(6)**, 1157–1171.

442

443   VIANEY-LIAUD M. AND ZELENITSKY D. (2003). Historical and new perspectives on the

444   parataxonomy of fossil eggs. *Palaeovertebrata* **32**, 189–195.

445

446   VIANEY-LIAUD M., KHOSLA A., AND GARCÍA G. (2003). Relationships between

447   European and Indian dinosaur eggs and eggshells of the oofamily Megaloolithidae. *The*

448   *Journal of Vertebrate Paleontology* **23**, 575–585.

449

450   Vila B, Jackson FD, Fortuny J, Sellés AG, Galobart, À. (2010).D Modelling of Megaloolithid

451   Clutches: Insights about Nest Construction and Dinosaur Behaviour. *PLoS ONE* 5**(5)**, e10362.

452

453   GALOBART À. (2011) The chronology of dinosaur oospecies in south-western Europe:

454   refinements from the Maastrichtian succession of the eastern Pyrenees. *Cretaceous Research*

455   **32**, 378–86.

456

457   VILA B., GALOBART À., CANUDO J. I., LE LOEUFF J., DINARÈS-TURELL J., RIERA

458   V., OMS O., TORTOSA T., AND GAETE. R. (2012). The diversity of sauropod dinosaurs

459   and their first taxonomic succession from the latest Cretaceous of southwestern Europe: Clues

460   to demise and extinction. *Palaeogeography, Palaeoclimatology, Palaeoecology* **350-352**, 19-

461   38.

462

463   WINK C. S., ELSEY R. M., AND BOUVIER M. (1990). Porosity of eggshells from wild and

464   captive, pen-reared alligators (*Alligator mississippiensis*). *Journal of Morphology* **203**, 35-39.

465

466   ZHAO Z, DING S. (1976). Discovery of the dinosaurian eggshells from Alxa,

467   Ninguia and its stratigraphic significance. *Vertebrata PalAsiatica* **14**, 42–44.

19

468

469

470

471

472

473

474

475

476

477

478

479

480

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

	Thickness			Width			
Centre	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

482

483

Classification: Protected A

484

485

486

487

488

489

490

491

492

493

494

495

496 Table 2. Statistical comparisons between the centres produced by the Gaussian Mixture

497 Model under alpha/n comparisons.

498

499 KS statistic

	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

500

501

502

503

504

505

506

507

508

509

510

511

512   Table 3. Statistical comparisons between the centres produced by the Gaussian Mixture

513   Model under Adjusted alpha level.

514

515   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

516
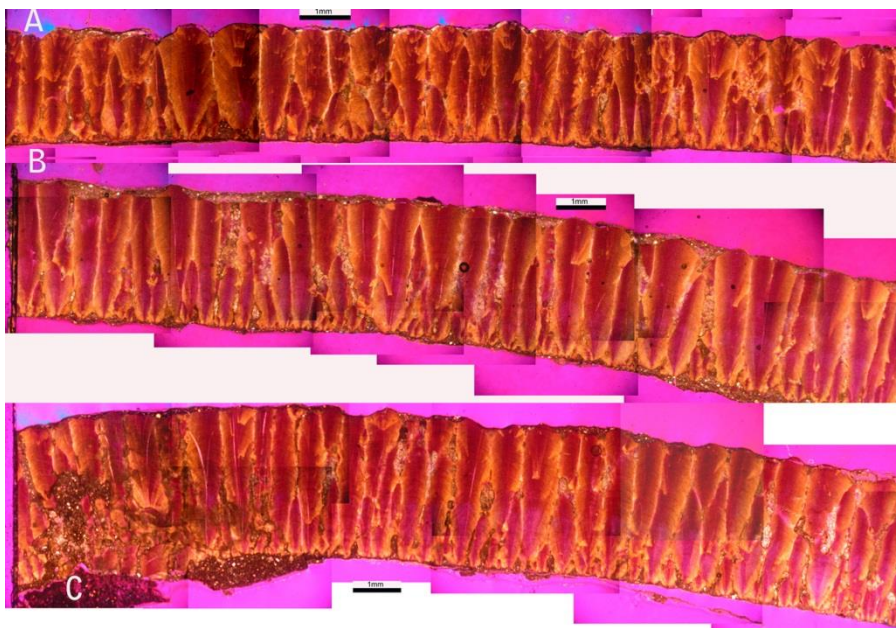
517

22

518

519

520

521

522

523

524

525

526

527

528

529 Figure 1. Petrographic thin sections of eggshells from the pole 1 (A), pole 2 (B) and equator

530 (C) of an egg from the nest from Faidella used here. Notice the different morphometrical and

531 shape of units between the eggshells from the equator and the poles.

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547   Figure 2. Linear regression graphs displaying the close relation between the variation of

548   eggshell thickness between reptile and dinosaur eggs.
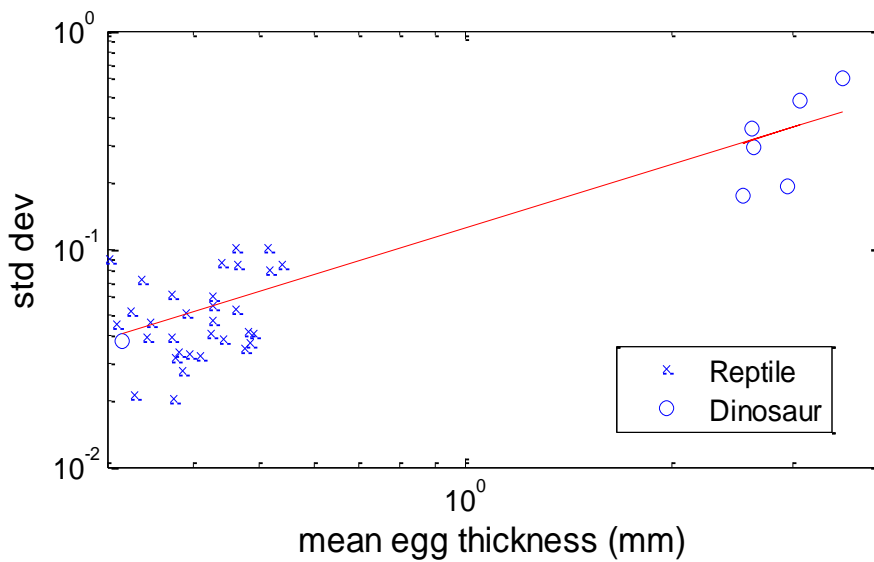
549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571   Figure 3. Representation of the eggshell data together with the centres identified by the

572   Gaussian mixture model. The centres represented:

573   1 Cp = *C. porosus*

574   1 Th = *T. hermanni*

575   1 Gc = *G. carbonata*

576   1 Cn = *C. nilocitus*
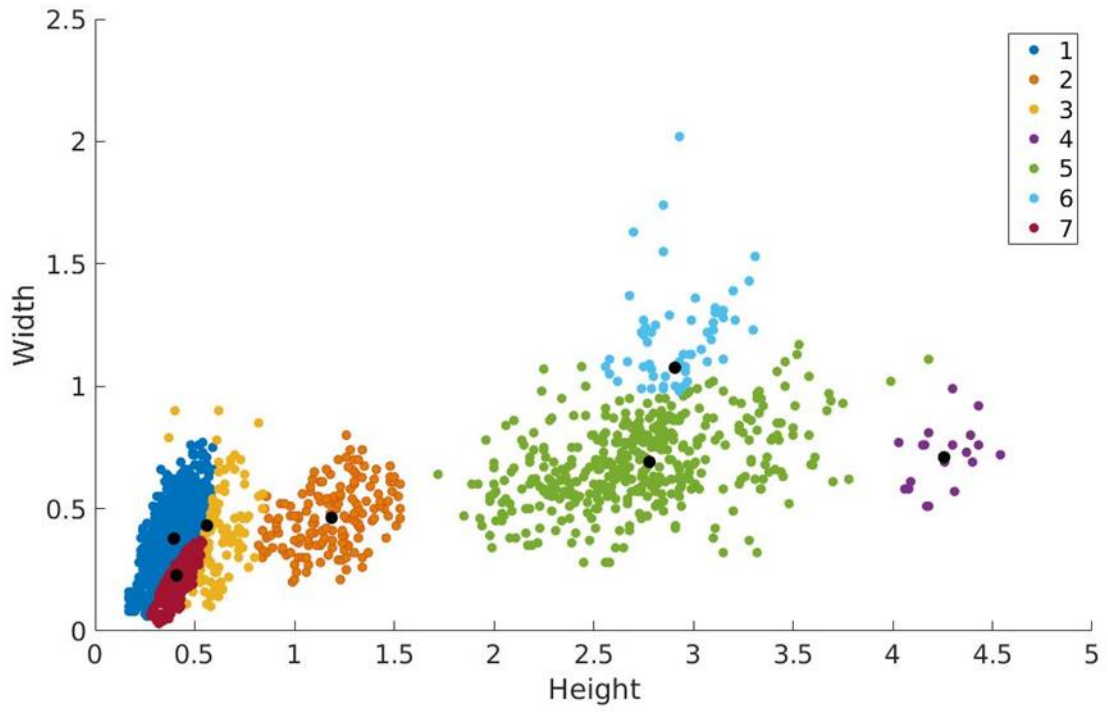
577   2 St = Suterranya

578   3 Fd = Faidella

579   4 Bs = Basturs

580   5 Fd = Faidella

581   6 CdN (nest) = Coll de Nargó

582   7 CdN (locality) = Coll de Nargó.

583

25

584

585

586

Classification: Protected A