Evolutionary Ecology Research, 2010, **12**: 363–376

The living and the fossilized: how well do unevenly distributed points capture the faunal information in a grid?

Juha Saarinen[1], Emilia Oikarinen[2], Mikael Fortelius[1]
and Heikki Mannila[3]

[1]*Department of Geosciences and Geography, University of Helsinki, Helsinki,*
[2]*HIIT, Department of Computer Science, University of Helsinki, Helsinki*
and [3]*Department of Computer and Information Science and HIIT,*
Aalto University, Aalto, Finland

ABSTRACT

Background: Fossil data sets are typically point-like, i.e. they provide information about a fossil fauna only for scattered localities. Modern distribution data are typically based on grid cells, and provide a (nearly) full description of the fauna.

Question: How good are estimates of the characteristics of the whole fauna that one obtains by looking only at point-like data similar to fossil data sets?

Data: Locality-based occurrence data for Neogene terrestrial mammals taken from the NOW database, and grid-based modern mammal distribution data for extant terrestrial mammals of Europe taken from the EMMA database.

Analytical methods: We used the EMMA dataset to generate subsets of the data, which contain only information about those grid cells that correspond to the location of at least one fossil locality in the NOW data. For each such dataset, we compared the different characteristics of the fauna (such as diversity, trophic structure, hypsodonty, occupancy, and representation of spatial patterns) to the whole EMMA dataset.

Results: Grid cells corresponding to fossil localities detect about 60% of the known species. Community structure can be estimated very accurately from the data on these localities.

Conclusions: Grid cells containing fossil localities give good estimates about the total fauna known from all grid cells, especially for relative measures such as community structure. Limitation to grid cells has a stronger effect on spatial patterns, but the main trends are still observed. Adding noise to the data has a negligible effect in most cases. The terrestrial mammals present in the NOW database can thus be regarded as a reasonably good representation of the real fauna from which they are drawn.

Keywords: grid data, incompleteness of the fossil record, NOW database, point data.

Correspondence: E. Oikarinen, Department of Geosciences and Geography, University of Helsinki, PO Box 64, FIN-00014 Helsinki, Finland. e-mail: emilia.oikarinen@helsinki.fi

INTRODUCTION

How incomplete the fossil record is has been a major issue in paleobiology at least since Charles Darwin drew attention to it. In recent decades, the advent of databases and the rise to prominence of analyses of past biodiversity have served to further emphasize its significance (e.g. Valentine, 1989; Benton, 1994; Benton and Simms, 1995; Foote, 1997; Foote and Sepkoski, 1999; Alroy *et al.*, 2001, 2008). The study of sampling biases in the fossil record has been primarily directed at correcting for them and thereby generating believable representations of changes in past taxon richness and the ecological structure of fossil communities. A simple but potentially important aspect that has received little attention to date is that the 'baseline' comparison of fossil to living biota typically involves comparing data with very different structure.

Fossil data are nearly always point-like data, derived from individual, spatially limited localities (sites, quarries) containing fossiliferous sedimentary strata. This is also true of some datasets for living organisms, such as data based on game parks and nature reserves. Such data differ from most distribution data for living animals, which tend to be collected and reported according to a cartographic grid covering large geographical areas more or less completely. Our recent efforts to map ecological and spatial patterns in the record of Neogene land mammals of Eurasia (e.g. Fortelius *et al.*, 1996, 2002, 2006; Jernvall and Fortelius, 2002, 2004; Eronen, 2006, Eronen *et al.*, 2009; Liu *et al.*, 2009) and in the record of living land mammals of Europe (Heikinheimo *et al.*, 2007) have drawn our attention to this issue. Here we test how much of the original grid data for European living mammal distributions can be captured by scattered data points, using both real and simulated locality distributions. We investigate the number of species, trophic structure, ecometrics, and spatial distribution patterns. We also investigate the effect of adding noise to the data.

MATERIALS AND METHODS

The grid data in this study are from the European mammal dataset (EMMA), which has been collected by the Societas Europaea Mammalogica (http://www.european-mammals.org) to prepare the *Atlas of European Mammals* (Mitchell-Jones *et al.*, 1999). The data in EMMA consist of presence/absence records of 194 mammal species for a set of 2670 grid cells covering Europe within the boundaries 32°W, 35°E, 81°N, and 30°N. The cell resolution of the grid is approximately 50 × 50 km, and the grid system is based on the Universal Transverse Mercator (UTM) projection and the Military Grid Reference System (MGRS).

We used the preprocessed EMMA dataset from Heikinheimo *et al.* (2007), except for Northern Europe (> 55°N) where there are no mammal fossil localities. Our preprocessed data provide information on 111 species in 1528 grid cells covering a window with boundaries of 11°W, 32°E, 55°N, and 35°N.

The fossil mammal localities chosen for this study are from the NOW (Neogene of the Old World) database [http://www.helsinki.fi/science/now (Fortelius, 2010)]. NOW localities with longitude > 30°E are omitted, since they do not coincide with EMMA grid cells. Localities lacking geographic coordinates were omitted for the same reason.

To assess the quality of estimates generated by point-like data, we compared the total diversity and other characteristics of the fauna in complete modern grid data to those of point-like data obtained by a random selection of 10–120 grid cells (repeated 1000 times) as well as based on spatially clustered fossil localities from the NOW database (Fig. 1). To

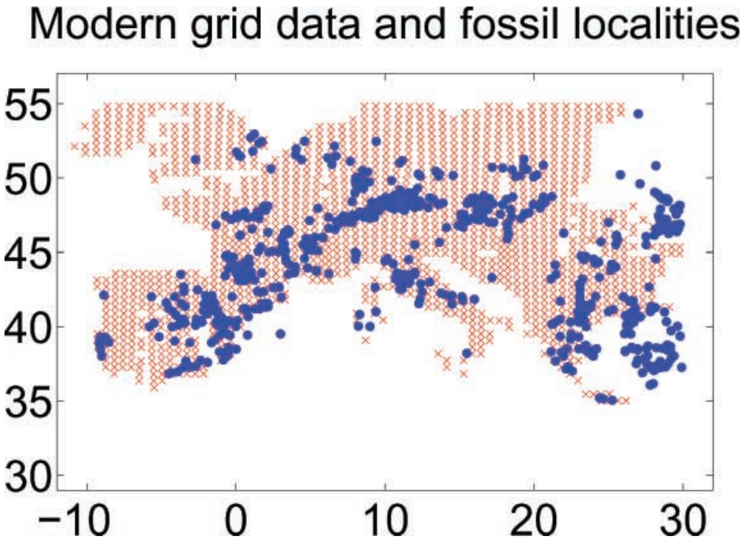

Fig. 1. The modern grid data (EMMA) represented by crosses and the fossil localities (from NOW data) represented by solid circles. The EMMA grid cells containing fossil localities were chosen to evaluate how much of the diversity indicated by grid data are also indicated by point-like data.

evaluate the effect of NOW locality distribution on apparent temporal patterns, NOW localities were chosen from each 'MN' unit of the Neogene Mammal biochronology (Mein, 1975, 1989; Steininger, 1999). The numbers of localities for each MN zone are presented in Fig. 2. The characteristics were also computed for perturbed NOW localities obtained by randomly shifting both longitude and latitude of each site within an interval of [−1, 1] degrees (repeated 100 times). In addition, we studied the effect of noise on the presence/absence information in the mammal database by allowing a 50% probability of not observing a species in a grid cell.

Due to the spatial distribution of fossil localities, there can be several localities in one grid cell. The data from such a cell are repeated as many times as there are fossil sites in the cell. We chose to allow repetitions so as to reflect the clumped nature of the fossil data and see how it affects results based on point-like data obtained from fossil sites. To discover the effects of the repetitions on the results, we also sampled the grid cells based on fossil locations but without repetitions.

The characteristics analysed were 'number of species' (diversity), 'occupancy' (fraction of species occurring in at least 20% of the cells), 'mean hypsodonty' (calculated from the species for which a hypsodonty score is present in the dataset), 'trophic structure' (fraction of herbivores), 'correlation between hypsodonty and longitude', and 'correlation between number of species and longitude'. The use of mean hypsodonty of fossil herbivores as a proxy for humidity has so far been based on 'large' mammals only, excluding 'small' mammals such as rodents (e.g. Fortelius *et al.*, 2002; Eronen *et al.*, 2010a, 2010b). Because the number of large mammal herbivores in the EMMA data is small, we have included rodents in the present study. Thus, although our results do represent the sampling effects on the distribution of a typical ecometric, they are not directly comparable with published hypsodonty results based on NOW data.

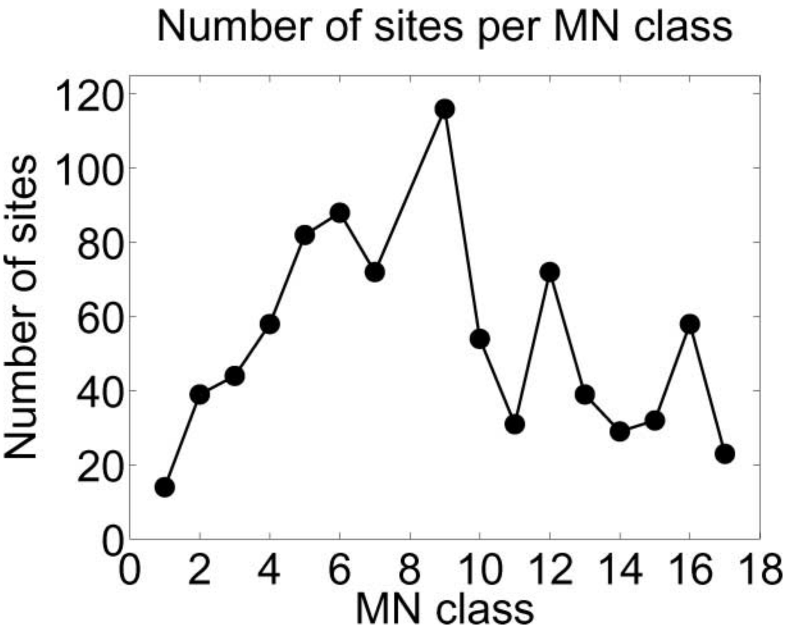

Fig. 2. Number of fossil localities per MN zone (simulated temporal sequence).

RESULTS AND DISCUSSION

More than half of the actual richness (111 species) is preserved in the randomly sampled cells even with a sample size of 10 cells (Fig. 3A). When the number of cells increases, they detect an even greater proportion of the known species. The proportion of known species asymptotes at 80%. Somewhat fewer of the species are detected by random selection based on NOW locations than by randomly selected locations. One explanation is that some areas of Europe are not covered by NOW sites and fewer species are likely to occur in grid cells containing NOW sites. However, if grid cells are sampled based on NOW locations without repetitions, the sample detects about the same proportion of species as random grid cells. That fact implies that the difference is most likely explained by the clumped nature of NOW sites, which makes grid cells containing several fossil locations dominate in the analysis.

On average, about 68% of the actual species diversity is preserved in the fossil data clustered by MN zones (76 species) (Table 1 and Fig. 3B).

A correlation between observed number of species and number of localities sampled has been reported in previous studies based on various versions and subsets of the NOW data (e.g. Fortelius *et al.*, 1996; Peláez-Campomanes and van der Meulen, 2009). The accumulation curve obtained for the EMMA data for NOW locations arranged by time (MN units) (Fig. 3B) is also somewhat similar to the curves obtained in previous studies, especially in the low values of the oldest MN units and the maxima around the MN units with the highest locality counts, MN 6 and MN 9. The details are quite different, however, suggesting that although a correlation clearly does exist, the NOW data contain significant additional information not related to sampling.

The results remain very similar when noise is added to the presence data, but the amount of information captured decreases. The decrease is stronger for the randomly sampled

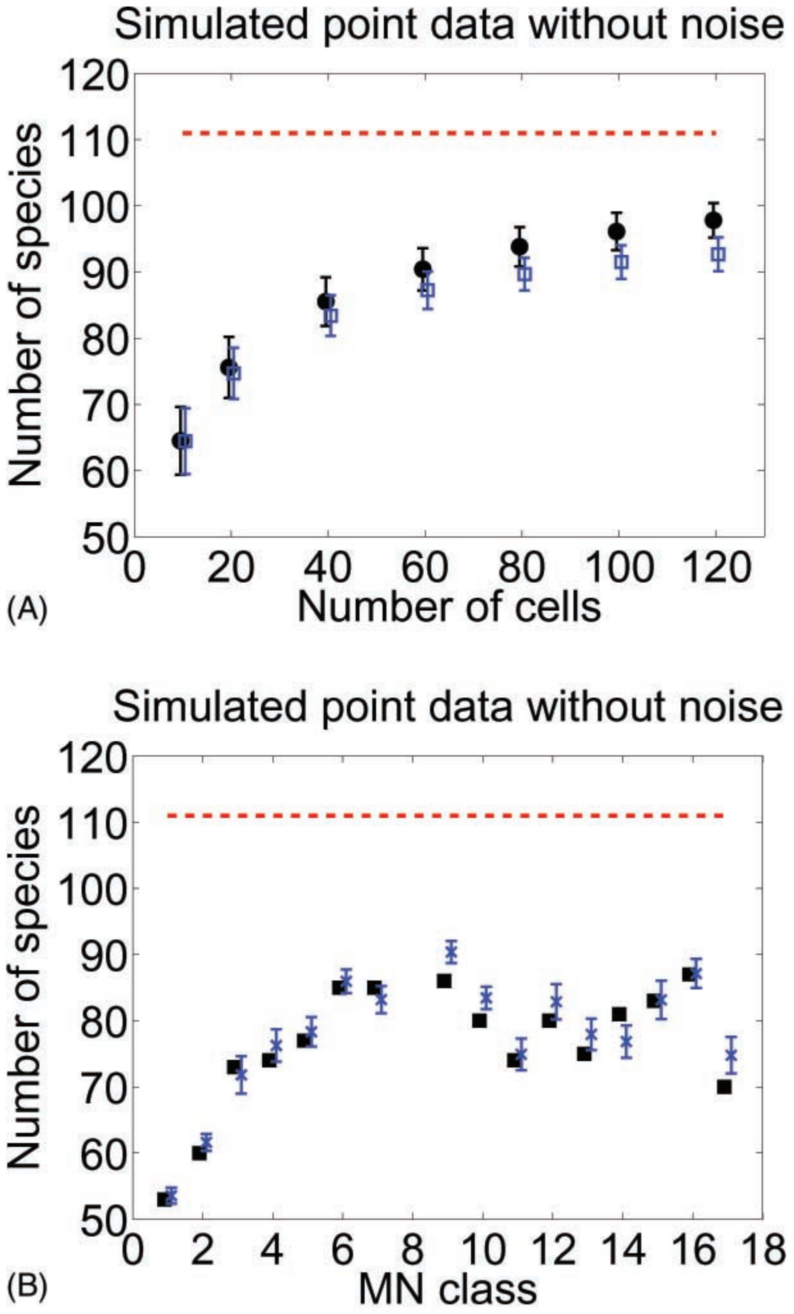

Fig. 3. Number of species estimated by modern grid data indicated by dashed line, and point data indicated by different markers. On average, 68% of the original diversity is predicted by the point data when there is no noise in the presence information. (A) Sequence of increased sampling. Random locations (solid circles) and random locations containing NOW localities (open squares). To separate overlapping symbols, there is an offset of 1.0 between the x-values of the two samplings. (B) NOW localities (solid squares) and perturbed NOW localities (×). To separate overlapping symbols, there is an offset of 0.2 between the x-values of the two samplings.

Table 1. Estimates of species count, occupancy, percentage of herbivorous mammals, mean hypsodonty, correlation of hypsodonty and longitude, and correlation of diversity and longitude of the modern EMMA data predicted by fossil localities from each MN-zone

		Number of species		Occupancy		Percentage of herbivores		Hypsodonty		Hypsodonty–longitude correlation		Diversity–longitude correlation	
MN-zone	Number of locations	Without noise	With noise	Without noise	With noise	Without noise	With noise	Without noise	With noise	Without noise	With noise	Without noise	With noise
1	14	53	46	0.355	0.315	0.698	0.693	1.998	1.974	−0.351	0.376	0.656	0.254
2	39	60	58	0.371	0.274	0.700	0.682	2.015	2.132	−0.567	−0.433	0.546	0.389
3	44	73	69	0.323	0.258	0.701	0.671	1.993	2.030	−0.636	−0.447	0.246	0.104
4	58	74	62	0.339	0.194	0.695	0.684	1.990	1.916	−0.468	−0.225	0.565	0.622
5	82	77	71	0.347	0.274	0.713	0.702	1.965	2.033	−0.406	−0.228	0.399	0.288
6	88	85	81	0.339	0.282	0.706	0.701	1.912	1.928	−0.437	−0.204	0.158	−0.011
7	72	85	76	0.347	0.250	0.705	0.677	1.921	1.917	−0.644	−0.253	0.251	0.295
9	116	86	80	0.363	0.250	0.713	0.718	1.880	1.928	−0.614	−0.405	0.121	0.222
10	54	80	71	0.403	0.218	0.714	0.708	1.893	1.982	−0.774	−0.413	0.053	0.117
11	31	74	70	0.379	0.202	0.721	0.658	1.874	2.050	−0.788	−0.670	0.183	0.286
12	72	80	69	0.347	0.169	0.722	0.706	1.962	1.986	−0.564	−0.530	0.137	0.195
13	39	75	66	0.298	0.161	0.743	0.655	2.017	2.268	−0.603	−0.674	0.594	0.579
14	29	81	68	0.371	0.161	0.736	0.676	1.962	2.045	−0.748	−0.637	0.192	0.136
15	32	83	68	0.323	0.210	0.708	0.660	1.890	2.068	−0.663	−0.402	0.239	0.304
16	58	87	78	0.298	0.145	0.734	0.746	1.876	1.928	−0.362	−0.000	−0.175	−0.142
17	23	70	63	0.339	0.234	0.701	0.719	2.007	1.930	−0.635	−0.532	−0.173	−0.034
EMMA	1528	111	108	0.331	0.234	0.716	0.717	1.914	1.933	−0.373	−0.177	0.241	0.196

cells than for the NOW locations, presumably due to the clumped nature of the spatial distribution of NOW localities. Clumping results in a more than average degree of pseudo-replication between localities sampled. The probability of observing a species based on point-like location data remains at about 60%. This is in line with previous studies of representativeness of good-quality fossil datasets, especially the study of Alba *et al.* (2001), who estimated a recovery of about 75% of species in the exceptionally good Spanish record of fossil land mammals, which makes up a major part of the NOW data.

Whereas the point-like data corresponding to fossil localities capture only about 60% of the original number of species, they are much better at predicting the structure (fractional characteristics) of mammal communities. Occupancy (commonness) of the mammal species was analysed by testing how large a fraction (percentage) of the species occurs in at least 20% of the sites. The results show that the point-like data without noise provide a good estimate of the fraction of common species, but noise reduces the accuracy of the estimate (Table 1). On average, about 22% of species are estimated to occur in at least 20% of the localities even by the worst case (NOW locations with noise), whereas the actual value for the complete grid data is about 33%. Without noise, the estimate for occupancy given by fossil localities is 35% on average.

Randomly sampled grid cells containing fossil localities slightly overestimate occupancy, whereas random sampling over the whole grid results in slight underestimates of occupancy (Fig. 4A). The clumped nature of fossil locations can be seen as an explanation for the higher estimate of occupancy when sampled over grid cells with fossil localities.

With noise, both cases estimate about the same occupancy (22%) as average NOW locations (Fig. 4B). However, the fall in occupancy when noise is added also has a significant effect on the full EMMA data: with noise only 23% of species occur in at least 20% of the grid cells (Table 1). When noise is added there are likely to be fewer species in individual grid cells, and thus the probability that a species occurs often is decreased, explaining the drop in occupancy. Thus, samples of random grid cells and of fossil localities both give very good estimates of occupancy when compared with the full noisy data.

The point-like data seem to be especially good at estimating trophic structure (the fraction of herbivores) in the mammal communities and structural derivatives such as mean hypsodonty values of herbivorous mammals.

Without noise, the fraction of herbivores estimated by random locations or by random locations containing NOW localities is very close to the actual fraction seen in the EMMA data (72%) – even with very small sample sizes. The estimates for random locations are not affected by noise (Fig. 5A), whereas with random samples over locations containing fossil sites the estimated fraction of herbivores becomes smaller (70%) (Fig. 5B).

The fraction of herbivores estimated by localities per MN zone is also close to the actual fraction in the EMMA data with a relatively small variation (from 65% to 75%) even with noise (Table 1). This is in line with previous studies suggesting that community structure is robustly preserved in fossil data in general (Damuth, 1982) and in the NOW data in particular (Jernvall and Fortelius, 2004).

Mean hypsodonty (tooth crown height) of the cheek teeth of herbivorous mammals is an important measure of mammal community structure. It is directly connected with properties of local vegetation and thus indirectly with climate, especially precipitation (Fortelius *et al.*, 2002; Eronen, 2006; Eronen *et al.*, 2009, 2010a, 2010b). Mean hypsodonty estimated using NOW locations (2.01) is on average very similar to that indicated by the complete EMMA data (1.91), even with noise added (Table 1). Random sampling over grid cells containing

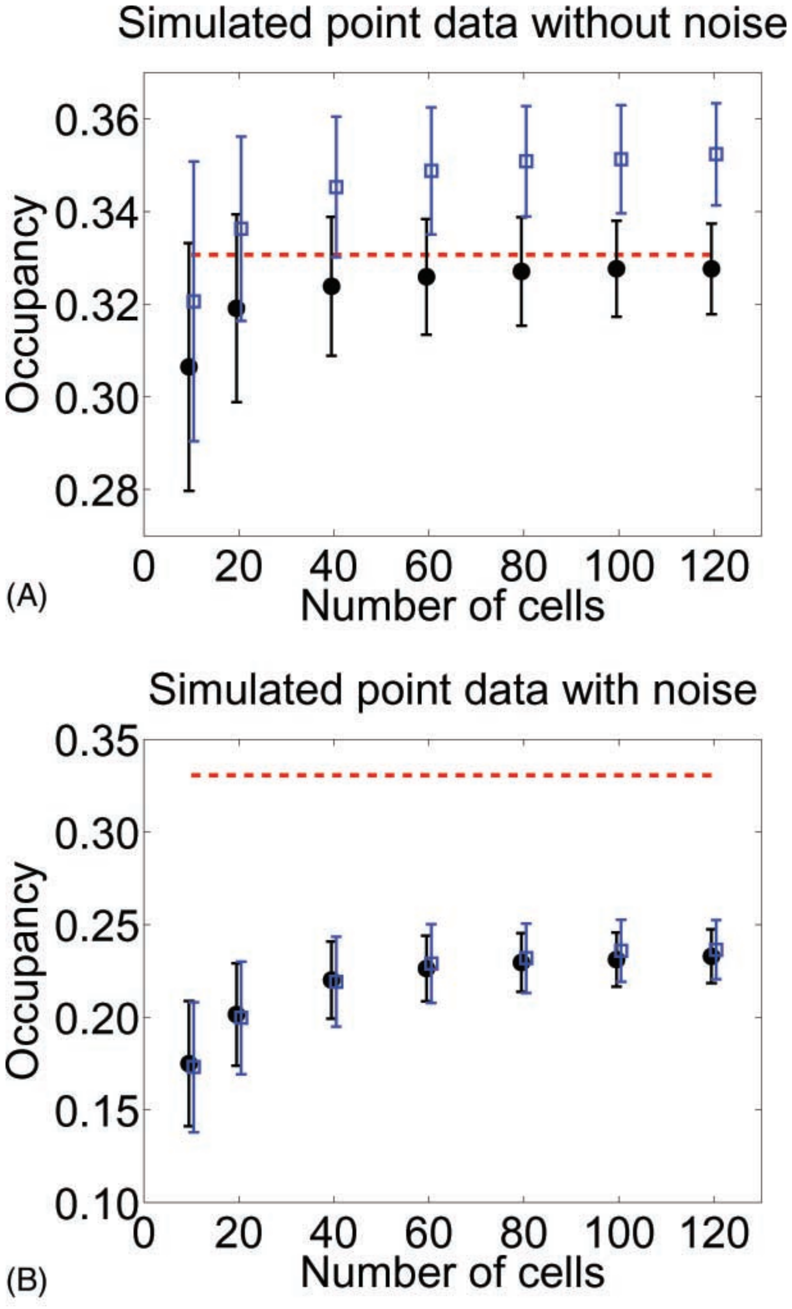

Fig. 4. Occupancy (commonness) of the species as a fraction of the species occurring in at least 20% of the sites estimated by modern grid data indicated by dashed line, and point data indicated by solid circles (random locations) and open squares (random locations containing fossil localities). To separate overlapping symbols, there is an offset of 1.0 between the x-values of the two samplings. (A) Without noise; (B) with noise.

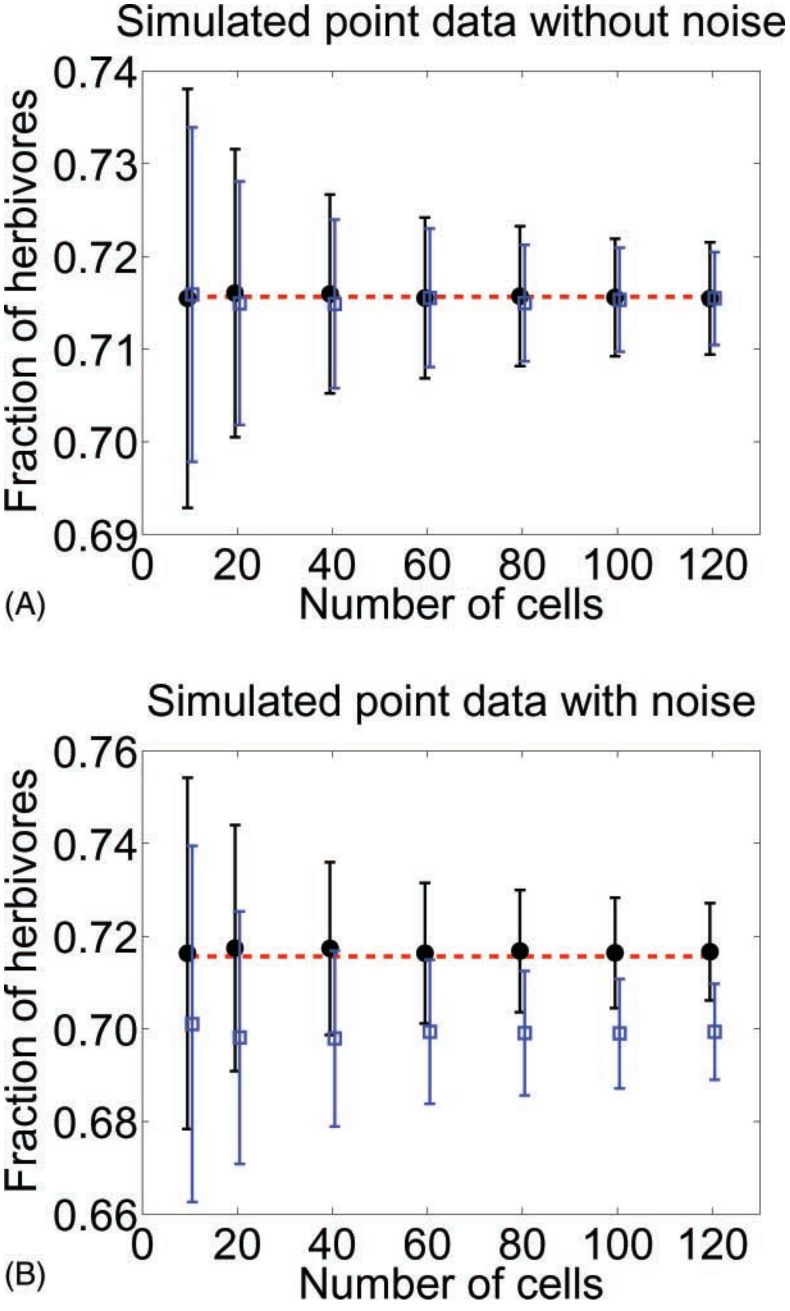

Fig. 5. Fraction of herbivores estimated by modern grid data indicated by dashed line, and point data indicated by solid circles (random locations) and open squares (random locations containing fossil localities). To separate overlapping symbols, there is an offset of 1.0 between the x-values of the two samplings. (A) Without noise; (B) with noise.

fossil localities results in similar estimates. And randomly sampling grid cells over the whole grid gives a mean hypsodonty estimate very close to the actual value with and without noise – even with the smallest sample size (10 grid cells). Random samples without repetitions over the grid cells containing NOW locations behave as random samples over the whole grid, i.e. the estimate of mean hypsodonty is very close to the actual value with and without noise.

On average, NOW locations show the same correlation of diversity with longitude as the complete grid data (0.24), both without (0.25) and with noise (0.23). Even though the correlation is the same on average, the standard deviation is large (0.25 without noise and 0.21 with noise). On average, MN zones 1–7 show a stronger correlation than the correlation over the complete grid, and MN zones 9–17 show a much weaker correlation than over the complete grid (Fig. 6). These differences can be explained by the spatial distribution of sites on different MN zones: on average, the locations from MN zones 1–7 are farther west than those from MN zones 9–17.

The randomly sampled grid cells show about the same correlation of diversity with longitude as the complete grid. But cells randomly sampled from grid cells containing fossil locations show a significantly weaker correlation (both with and without noise). This suggests that the spatial distribution of localities has a significant effect on the patterns observed.

The correlation between hypsodonty and longitude shown by NOW localities per MN units is between −0.35 and −0.79 (standard deviation 0.14) without noise and between +0.38 and −0.67 (standard deviation 0.27) with noise, whereas it is −0.37 for the complete grid data. Without noise, the randomly sampled grid cells showed about the same correlation between hypsodonty and longitude as the correlation over the complete grid, whereas grid cells randomly sampled over fossil locations showed a stronger correlation.

With noise, both randomly sampled grid cells and randomly sampled NOW localities showed a weaker correlation than is seen in the complete data. However, the correlation between hypsodonty and longitude is affected by noise even in the complete data: with noise, the correlation for the whole EMMA drops to −0.18 (Table 1). The estimate of correlation for random samples is close to the correlation seen for the full EMMA data with noise added, whereas the estimate for random samples over fossil localities decreases about the same amount again. The effect of noise is the same in the complete data and in random samples (both over the whole grid and grid cells containing fossil localities). But samples based on NOW locations show higher estimates for correlation (both with and without noise).

The amplification of the geographic gradient observed for the NOW locations is most probably a result of their clumped spatial distribution, which results in a gradient appearing more as a threshold. These results suggest that the spatial distribution of localities has a significant but not decisive effect on the patterns observed.

For random samples (both over the whole grid and over the fossil localities), the number of cells sampled has a clear effect on the extent to which the number of species in the entire grid is revealed. But for trophic structure, hypsodonty, and occupancy, the information is revealed to a high degree even with the smallest sample sizes (for instance, 40 grid cells out of 1528 possible is less than a 3% sample). A larger sample size results in a smaller standard deviation, however. For ‘NOW locations per MN zone’, the sample size has a larger effect on the estimates: the MN zones with most localities (MN 5, MN 6, and MN 9) tend, overall, to give less variable estimates than MN zones with fewer localities.

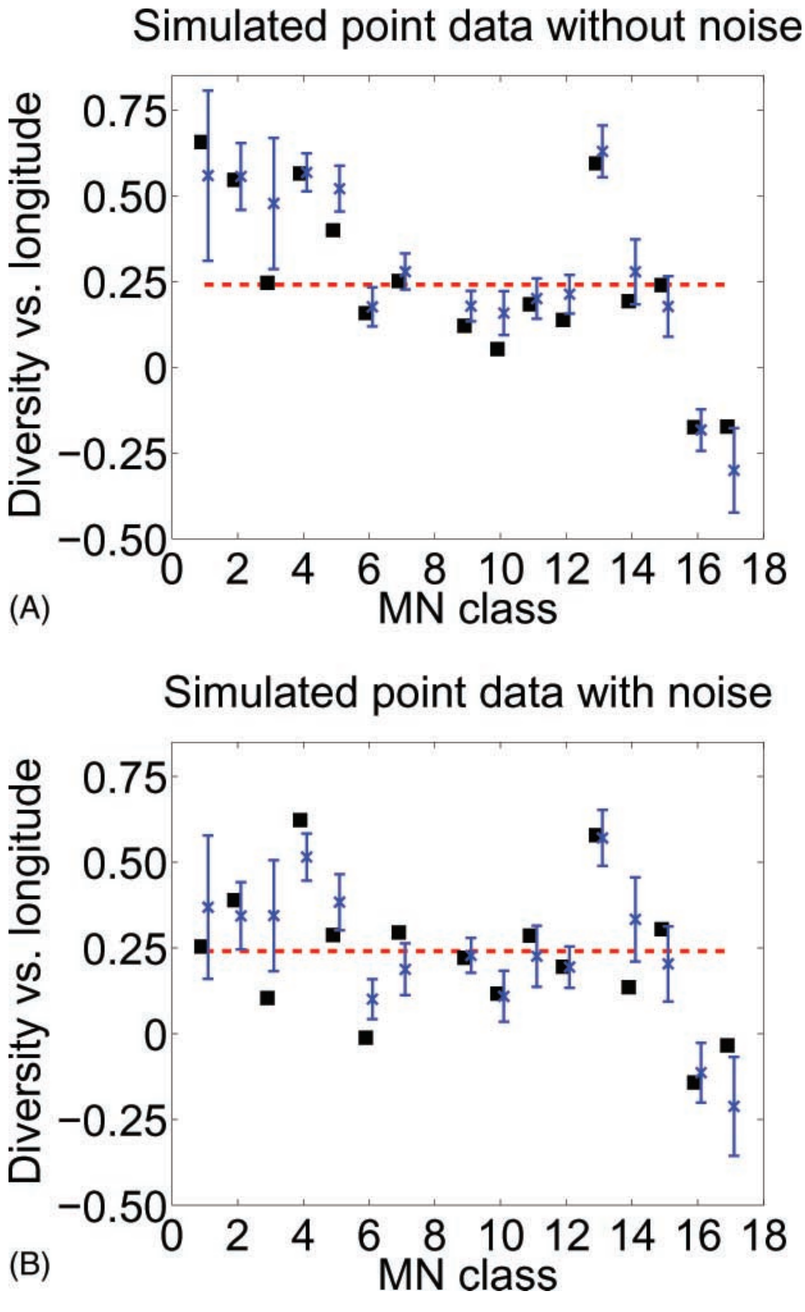

Fig. 6. Correlation between diversity and longitude shown by modern grid data indicated by dashed line, and point data indicated by solid squares (NOW localities) and crosses (perturbed NOW localities). To separate overlapping symbols, there is an offset of 0.2 between the x-values of the two samplings. (A) Without noise; (B) with noise.

Spatial patterns are clearly more sensitive to sampling than estimates for the whole area. Random samples capture the same information about spatial patterns as the complete data, but the effect of the clumped nature of fossil localities can be seen in samples based on fossil localities. For instance, for random samples over NOW locations, the correlation between hypsodonty and longitude is overestimated, whereas the correlation between diversity and longitude is underestimated.

Overall, the estimates of characteristics obtained for point-like data are rather robust with respect to noise. Noise has a visible effect on estimates of occupancy and on the correlation between hypsodonty and longitude. However, a similar effect of noise is observed also in the complete data (Table 1).

CONCLUSIONS

It appears that the point-like data provided by grid cells containing fossil localities of the MN-zones of Eurasia do provide reasonably good estimates of the characteristics of modern mammal communities and even an approximately 60% appearance of the total diversity of mammal species. This result is promising for future studies of fossil mammal communities because it indicates that fossil localities, despite their incomplete spatial coverage and noisy quality, can be used, with relatively good precision, to estimate properties of the real regional mammal faunas. The main caveat arising from our analysis concerns the effects that a spatially uneven distribution of localities may have on the analysis. Among the effects that could arise from a spatially clumped locality distribution are exaggeration of spatial gradients and a lower-than-expected effect of locality count on the number of species. The latter phenomenon might lead to secondary complications such as overcorrection for sampling effect. In addition to fossil data, our results have implications for all point-like datasets, such as data collected from game parks and nature reserves. We suggest that the kind of analysis presented here could be usefully applied to other datasets, both as a test of the generality of our results and as a simple means to assess the representativeness of the data.

ACKNOWLEDGEMENTS

We thank the Societas Europaea Mammalogica and Tony Mitchell-Jones for providing the distributional data used to prepare the *Atlas of European Mammals*. This study was supported by the Academy of Finland.

REFERENCES

Alba, D.M., Agustí, J. and Moyà-Solà, S. 2001. Completeness of the mammalian fossil record in the Iberian Neogene. *Paleobiology*, **27**: 79–83.

Alroy, J., Marshall, C.R., Bambach, R.K., Bezusko, K., Foote, M., Fürsich, F.T. *et al.* 2001. Effects of sampling standardization on estimates of Phanerozoic marine diversification. *Proc. Natl. Acad. Sci. USA*, **98**: 6261–6266.

Alroy, J., Aberhan, M., Bottjer, D.J., Foote, M., Fürsich, F.T., Harries, P.J. *et al.* 2008. Phanerozoic trends in the global diversity of marine invertebrates. *Science*, **321**: 97–100.

Benton, M.J. 1994. Paleontological data and identifying mass extinctions. *Trends Ecol. Evol.*, **9**: 181–185.

Benton, M.J. and Simms, M.J. 1995. Testing the marine and continental fossil records. *Geology*, **23**: 601–604.

Damuth, J. 1982. Analysis of the preservation of community structure in assemblages of fossil mammals. *Paleobiology*, **8**: 434–446.

Eronen, J.T. 2006. Eurasian Neogene large herbivorous mammals and climate. *Acta Zool. Fenn.*, **216**: 1–72.

Eronen, J.T., Mirzaie Ataabadi, M., Micheels, A., Karme, A., Bernor, R.L. and Fortelius, M. 2009. Distribution history and climatic controls of the Late Miocene Pikermian chronofauna. *Proc. Natl. Acad. Sci. USA*, **106**: 11867–11871.

Eronen, J.T., Puolamäki, K., Liu, L., Lintulaakso, K., Damuth, J., Janis, C. *et al.* 2010a. Precipitation and large herbivorous mammals, part I: estimates from present-day communities. *Evol. Ecol. Res.*, **12**: 217–233.

Eronen, J.T., Puolamäki, K., Liu, L., Lintulaakso, K., Damuth, J., Janis, C. *et al.* 2010b. Precipitation and large herbivorous mammals, part II: application to fossil data. *Evol. Ecol. Res.*, **12**: 235–248.

Foote, M. 1997. Estimating taxonomic durations and preservation probability. *Paleobiology*, **23**: 278–300.

Foote, M. and Sepkoski, J.J., Jr. 1999. Absolute measures of the completeness of the fossil record. *Nature*, **398**: 415–417.

Fortelius, M., coordinator. 2010. *Neogene of the Old World Database of Fossil Mammals (NOW)* (available at: http://www.helsinki.fi/science/now/).

Fortelius, M., Werdelin, L., Andrews, P., Bernor, R.L., Gentry, A., Humphrey, W. *et al.* 1996. Provinciality, diversity, turnover and paleoecology in land mammal faunas of the later Miocene of western Eurasia. In *The Evolution of Western Eurasian Neogene Mammal Faunas* (R. Bernor, V. Fahlbusch and W. Mittmann, eds.), pp. 414–448. New York: Columbia University Press.

Fortelius, M., Eronen, J., Jernvall, J., Liu, L., Pushkina, D., Rinne, J. *et al.* 2002. Fossil mammals resolve regional patterns of Eurasian climate change over 20 million years. *Evol. Ecol. Res.*, **4**: 1005–1016.

Fortelius, M., Eronen, J.T., Liu, L., Pushkina, D., Tesakov, A., Vislobokova, I.A. *et al.* 2006. Late Miocene and Pliocene large land mammals and climatic changes in Eurasia. *Palaeogeogr., Palaeoclimatol., Palaeoecol.*, **238**: 219–227.

Heikinheimo, H., Fortelius, M., Eronen, J.T. and Mannila, H. 2007. Biogeography of European land mammals shows environmentally distinct and spatially coherent clusters. *J. Biogeogr.*, **34**: 1053–1064.

Jernvall, J. and Fortelius, M. 2002. Common mammals drive the evolutionary increase of hypsodonty in the Neogene. *Nature*, **417**: 538–540.

Jernvall, J. and Fortelius, M. 2004. Maintenance of trophic structure in fossil mammal communities: site occupancy and taxon resilience. *Am. Nat.*, **164**: 614–624.

Liu, L., Eronen, J.T. and Fortelius, M. 2009. Significant mid-latitude aridity in the middle Miocene of East Asia. *Palaeogeogr., Palaeoclimatol., Palaeoecol.*, **279**: 201–206.

Mein, P. 1975. Résultats du Groupe de Travail des Vertébrés. In Report on Activity of R.C.M.N.S Working Groups – 6 (J. Senes, ed.). *Congress of the Regional Committee of Mediterranean Neogene Stratigraphy, Proceedings*, **1**: 78–1 (Bratislava).

Mein, P. 1989. Updating of MN Zones. In *European Neogene Mammal Chronology* (E.H. Lindsay, V. Fahlbusch and P. Mein, eds.), pp. 73–90. NATO ASI Series, A #180. New York: Plenum Press.

Mitchell-Jones, A.J., Amori, G., Bogdanowicz, W., Krystufek, B., Reijnders, P.J.H., Spitzenberger, F. *et al.* 1999. *The Atlas of European Mammals*. London: Academic Press.

Peláez-Campomanes, P. and van der Meulen, A.J. 2009. Diversity of mammals in the Neogene of Europe: comparing data quality of large and small mammals in the NOW database. Hellenic J. *Geosci.*, **44**: 105–115.

Steininger, F.F. 1999. Chronostratigraphy, geochronology and biochronology of the Miocene 'European Land Mammal Mega-Zones' (ELMMZ) and the Miocene 'Mammal-Zones (MN-Zones)'. In *The Miocene Land Mammals of Europe* (G.E. Rössner and K. Heissig, eds.), pp. 9–24. Munich: Verlag Dr. Friedrich Pfeil.

Valentine, J.W. 1989. How good was the fossil record? Clues from the Californian Pleistocene. *Paleobiology*, **15**: 83–94.