Ecography 37: 1133–1138, 2014
doi: 10.1111/ecog.01294

Subject Editor: Signe Normand. Editor-in-Chief: Jens-Christian Svenning. Accepted 16 September 2014

Ecological niche models of mammalian glacial refugia show consistent bias

Edward Byrd Davis, Jenny L. McGuire and John D. Orcutt

E. B. Davis (edavis@uoregon.edu), Dept of Geological Sciences, Univ. of Oregon, Eugene, OR 97403, USA, and Museum of Natural and Cultural History, Univ. of Oregon, Eugene, OR 97403, USA. – J. L. McGuire, School of Environmental and Forest Sciences, Univ. of Washington, Seattle, WA 98195-2100, USA. – J. D. Orcutt, Dept of Geology, Cornell College, Mt Vernon, IA 52314, USA.

Ecological niche models (ENMs) are crucial tools for anticipating range shifts driven by climate change. As hypotheses of future biotic change, they can be difficult to test using independent data. The fossil record is the best way to assess the ability of ENMs to correctly predict range shifts because it provides empirical ranges under novel climate conditions. We tested the performance of ENMs using fossil distributions from the Last Glacial Maximum (LGM, ~21 000 yr ago). We compared hindcast ENM LGM distribution hypotheses for five species of small mammals, drawn from the published literature, to the known LGM fossil record for those species and found a consistent southern prediction bias in the ENMs. This bias urges caution in interpreting future range predictions, and we suggest that the Pleistocene and Holocene fossil record should be used as an additional resource for calibrating niche modelling for conservation planning.

As climates change, organisms respond by evolving in place or shifting geographic ranges to follow optimal habitat (Holt 1990, Vrba 1993). The need to anticipate range shifts in the face of current, rapid climate change has provoked development of ecological niche models (ENMs) to predict these shifts (Hijmans and Graham 2006, Pearson and Raxworthy 2007, Waltari and Guralnick 2009). ENMs commonly use the climatic components of modern distributions to predict future distributions under global warming scenarios. Despite the important potential utility of ENMs, their predictive power outside modern conditions cannot be tested absent alternative climate regimes or costly experimental physiology studies. Fossil and paleoclimatic records allow such tests to be conducted by hindcasting range shifts during past intervals of climate change that can then be compared to empirical fossil occurrence data.

Using the fossil record to test ENMs under alternative climate regimes is a growing area of inquiry in conservation paleontology (Varela et al. 2011). These studies have important implications for mammalian conservation biology because, since they first began to be conducted a decade ago, their results have reflected a changing understanding of the apparent accuracy of hindcasting. In looking at many small mammal species, Martínez-Meyer et al. (2004) suggested that ecological niches have been relatively stable since the Last Glacial Maximum (LGM, 21–18 ka; Petit et al. 1999) and that ENMs should therefore be able to accurately project species distributions through time. However, Guralnick and Pearman (2010) found a low rate of overlap between hindcast and actual LGM ranges in their study of alpine and flatland species of rodents, with alpine species showing higher accuracy than flatland species. Similarly, McGuire and Davis (2013) found that ENMs did not accurately hindcast past populations in 3 of 5 closely related *Microtus* species; however, these species were in close geographic and trophic proximity to one another with great potential for interspecific competition to affect niche stability through time. Several papers have also demonstrated inaccuracies in hindcast tree and other plant projections, indicating a potentially multi-trophic problem (Pearman et al. 2008, Guralnick and Pearman 2010, Williams et al. 2013).

The conflict between the initial findings of Martínez-Meyer et al. (2004) and the subsequent studies has profound implications for transference: using ENMs to predict or hindcast geographic ranges under alternative climate scenarios. An important goal of conservation paleontology must be to determine whether transference is feasible (Hadly and Barnosky 2009, Dawson et al. 2011, Dietl and Flessa 2011, McGuire and Davis 2014). Conservation paleontology must also work to test the effectiveness of this kind of transference, exploring the limitations of the technique and whether some evolutionary and/or ecological groups are better modeled than others. Here, we examine ENM accuracy under alternate climates for a group of mammalian species selected from studies of LGM refugia. We build upon the published ENMs of Waltari et al. (2007), hindcasts of refugia for several species of small mammals during the LGM. We compare the range predictions of these models to the LGM mammal fossil record and test whether they correctly hindcast fossil occurrences.

Material and methods

We are testing the effectiveness of standard ENM predictions under alternate climates. We did not create new ENMs; instead, we used peer-reviewed, published LGM hindcasts created by Waltari et al. (2007) using a standard combination of Garp (Stockwell and Peters 1999) and MaxEnt (Phillips et al. 2006) models run on two LGM reconstructions: community climate system model (CCSM; Collins et al. 2006) and model for interdisciplinary research on climate (MIROC; Hasumi and Emori 2004).

We tested these hindcasts by overlaying the LGM fossil record, limiting our analysis to the five mammal species in both the LGM study and the FAUNMAP II database (Graham and Lundelius 2010): *Blarina brevicauda*, *Glaucomys sabrinus*, *Glaucomys volans*, *Martes americana*, and *Myodes gapperi*. These species span the USA and Canada, providing a good test of the hindcasts. The FAUNMAP II database (Graham and Lundelius 2010) contains published North American (USA + Canada) Plio-Pleistocene fossil mammal occurrences.

We included FAUNMAP II sites with evidence of ages within the LGM or the prior cold interval, back to 40 ka (Petit et al. 1999). This study interval, 40 to 17 ka, was chosen to optimize sample size while limiting temperature deviations from the LGM. Our study interval includes temperatures between 9.4 and 4.8°C below mean recent value (Petit et al. 1999, data available at <ftp://ftp.ncdc.noaa.gov/pub/data/paleo/icecore/antarctica/vostok/deutnat.txt>), with 90% of the interval within 3.3°C of the LGM value. Of course, the climate models focus on the LGM, so any temporal window must include sites from warmer temperatures than the model. Sites confidently within the range 40–17 ka are red points in Fig. 1. We also included sites that do not have high-precision dating but are likely within the cold interval according to biostratigraphic or other dating calculations. These sites have uncertain minimum age estimates and may come from times immediately after the LGM (yellow points in Fig. 1). Any sites that do fall outside of this interval might be expected to indicate species ranges that differ from those predicted by the ENMs due to the rapid climatic and environmental changes associated with deglaciation (Petit et al. 1999). However, for each species one or more sites that disagree with the hindcast are confidently within or before the LGM (Table 1).

To assess the extent to which the hindcasts appropriately predicted fossil localities, we quantitatively evaluated the number of true and false positives and negatives (categories a, b, c, d of Fielding and Bell 1997). True positives (a) are LGM localities within a hindcast that correctly do contain the species. False positives (b) are LGM localities within a hindcast not containing the species, i.e. sites the hindcast

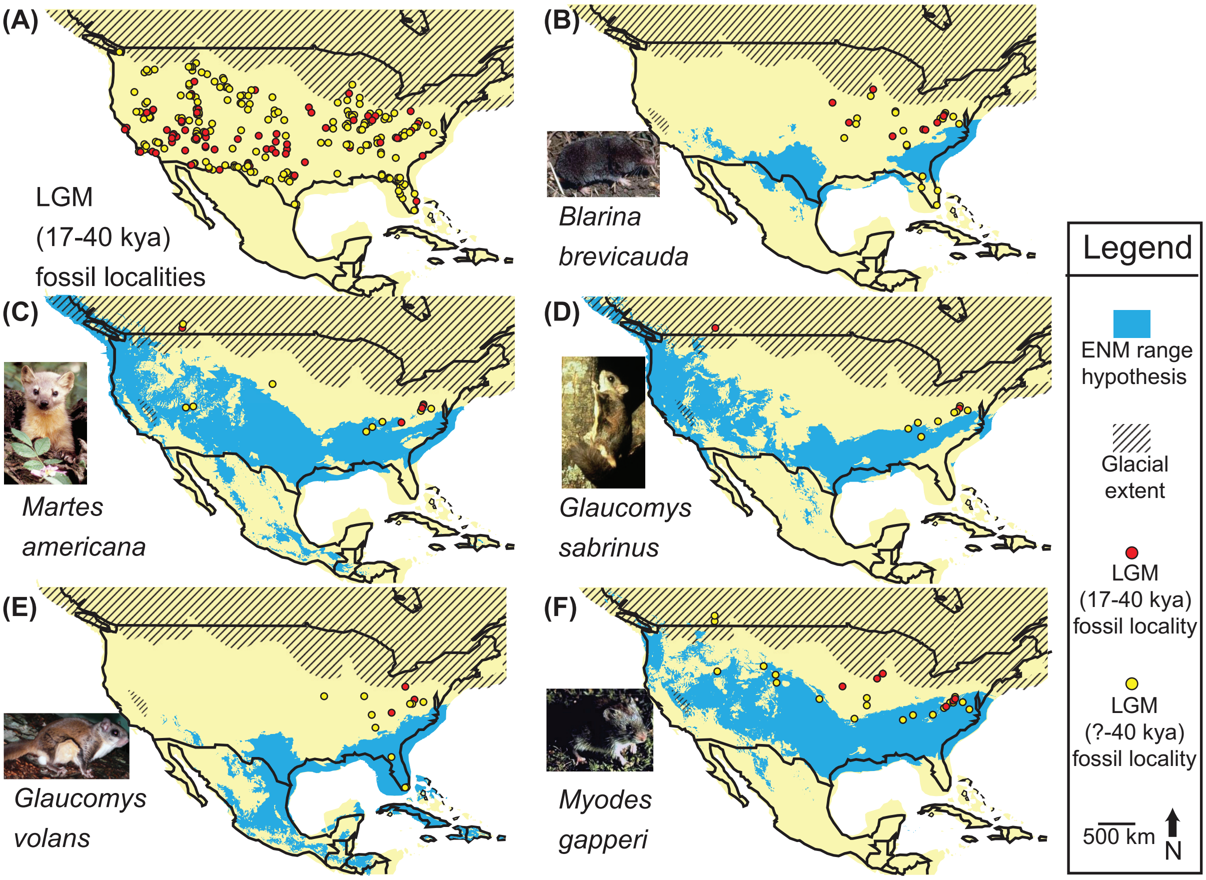

Figure 1. Paleodistribution maps for the five mammal species under consideration. Points highlight fossil occurrences, blue-shaded areas indicate ENM hindcasts, hashed area indicates extent of glacial ice. Red points are from the high confidence window (both maximum and minimum ages within 40–17 ka) and yellow points are added in the inclusive window (only maximum age estimate within 40–17 ka). Panel (A) shows the distribution of all LGM fossil sites in the conterminous USA from the FAUNMAP II database.

Table 1. Occurrence data for inclusive window (sites with a maximum age 40–17 ka). a: true positives. b: false positives. c: false negatives. d: true negatives. o: observed correct predictions (a + d). e: expected correct predictions, given the null model of random assignment. e_m: expected maximum chance value, a model assigning absence to all sites. p_1: one-tailed probability of o, compared to e. p_2: one-tailed probability of o, compared to e_m. N.S. indicates non-significance when $o < e$ or $o < e_m$. $N = 749$ for this analysis.

Species	a	b	c	d	o	e	e_m	p_1	p_2
Blarina brevicauda	1	29	36	683	684	679	712	0.25	N.S.
Martes americana	6	104	15	624	630	708	728	N.S.	N.S.
Glaucomys sabrinus	2	98	14	635	637	718	733	N.S.	N.S.
Glaucomys volans	3	26	13	707	710	718	733	N.S.	N.S.
Myodes gapperi	17	110	16	606	623	686	716	N.S.	N.S.

indicates should contain the species but do not. False negatives (c) are LGM localities outside of a hindcast that do contain the species, i.e. sites the hindcast indicates should not contain the species but do. True negatives (d) are LGM localities outside of a hindcast that correctly do not contain the species. Using the *z*-test of Huberty (1994), as directed by Fielding and Bell (1997), we tested the distribution of these categories with reference to the expected success rate of both a purely random model (Table 1) and the trivial model that assigns all fossil localities to absences (Table 2). Note that the equation for the *z*-statistic in Fielding and Bell (1997) is not formatted as clearly as in Huberty (1994) or Huberty and Olejnik (2006). The density of LGM fossil sites (Fig. 1A) indicates that, except for the central Gulf coast and middle Atlantic coast, areas of absence are likely true absences from the fossil record and not a consequence of poor sampling, so the *z*-test is an appropriate tool for testing effectiveness of these hindcasts.

Results

In general, the LGM models predict refugia to the south of observed fossil occurrences (Fig. 1). Additionally, none of these hindcasts is significantly better than randomly assigning presence, with four of the five actually making fewer correct predictions than the null model (Table 1). Strikingly, none of the hindcasts makes more correct predictions than simply assuming absence at all sites (Table 2). The hindcasts show no clear relationship between positive predictive power (PPP) and diet or habitat (Table 3).

Blarina brevicauda, the northern short-tailed shrew, is an insectivore found in the northeastern USA and southeastern Canada (George et al. 1986). The fossil distribution disagrees with the ENM (Fig. 1B, Table 3): 97% of localities are outside the hindcast (100% high confidence window). The hindcast for *B. brevicauda* is the only one to make more correct predictions than expected by the null model, but is not significantly so (Table 1). Counter to the hindcast, fossils occur in Florida, the Great Plains, the Appalachians, and the Great Lakes region but not Texas.

Table 2. Occurrence data for high confidence window (sites with both maximum and minimum ages 40–17 ka). a, b, c, d, o, e, e_m, p_1, p_2 as in Table 1. $N = 354$ for this analysis.

Species	a	b	c	d	o	e	e_m	p_1	p_2
Blarina brevicauda	0	5	9	340	340	336	345	0.19	N.S.
Martes americana	1	34	4	315	316	344	349	N.S.	N.S.
Glaucomys sabrinus	0	27	1	326	326	352	353	N.S.	N.S.
Glaucomys volans	0	5	3	346	346	348	351	N.S.	N.S.
Myodes gapperi	5	27	3	319	324	338	346	N.S.	N.S.

Table 3. Summary of ecology and predictive power. Ecological data from species accounts cited in text. Diet: c = carnivore, h = herbivore, i = insectivore. Hab.: habitat preference, m = montane, f = flatland. Specialist?: whether or not the species is considered a habitat and/or dietary specialist. PPP: positive predictive power, the proportion of correctly predicted occurrences. NPP: negative predictive power, the proportion of correctly predicted absences.

				Inclusive		High conf.	
Species	Diet	Hab.	Specialist?	PPP	NPP	PPP	NPP
Blarina brevicauda	i, c	f	N	0.03	0.95	0.00	0.97
Martes americana	c	m	N	0.05	0.98	0.03	0.99
Glaucomys sabrinus	h	m	N	0.02	0.98	0.00	1.00
Glaucomys volans	h, i	f	N	0.10	0.98	0.00	0.99
Myodes gapperi	h, i	f, m	N	0.13	0.97	0.16	0.99

Martes americana, the American marten, is a carnivore found in Canada and the mountain west of the USA, with historic range in the Great Lakes and Ohio Valley (Clark et al. 1987). The fossils show a distinct deviation from the hindcast (Fig. 1C, Table 3); 71% of fossil localities fall outside the hindcast (80% high confidence window). The hindcast shows a southern and coastal prediction unsupported by the fossil record.

Glaucomys sabrinus and *Glaucomys volans* are closely related species of flying squirrel. *Glaucomys sabrinus*, the larger northern flying squirrel, is found across Canada and, in the USA, in the northeast, northwest, and mountainous areas (Wells-Gosling and Heaney 1984). The hindcast encompasses much of North America, but 88% (100% high confidence window – only one point) of fossil sites fall outside this range (Fig. 1D, Table 3). Critically, no LGM fossils are west of the Mississippi, suggesting that hindcasts in the western and central USA are in error. *Glaucomys volans*, the smaller southern flying squirrel, is found across the eastern USA from the Atlantic to the Great Plains (Dolan and Carter 1977). The fossils disagree with the hindcast: 81% (100% high confidence window) are north of the hindcast, with none in the northern Gulf Coast (Fig. 1E, Table 3).

Myodes gapperi, Gapper's red-backed vole, is an omnivorous rodent found across southern Canada and in the Rocky Mountains, Great Plains, Atlantic coast, and Appalachians of the USA (Merrit 1981). This hindcast has the best success rate, with only 48% (38% high confidence window) of fossil sites north of the hindcast (Fig. 1F, Table 3).

Discussion

Our results clearly show important mismatches between the fossil record and hindcasts of LGM ranges: all of the ENMs incorrectly reconstruct refugia further south and inland than known LGM occurrences (Fig. 1). The majority of these reconstructions hindcast ranges in Texas and the western Gulf Coast, suggesting that the climate of that region is an attractor for the ENMs. The overall pattern observed in fossils is one of less actual range change relative to modern distributions than predicted, and fossils of most of the taxa are found closer to the ice front than was predicted by the ENMs, concurrent with findings in LGM vegetation ranges (Jackson et al. 2000) and other fossil mammals (Guralnick and Pearman 2010). The northern location of these North American LGM fossil occurrences also aligns with the growing support for northern LGM refugia in European ENM studies (Fløjgaard et al. 2009, Vega et al. 2010, Svenning et al. 2011).

The density of southern fossil localities (Fig. 1A) makes the ubiquity of southern predictions in LGM models of great concern. The majority of these southern sites come from our inclusive window, so there is a possibility that some of the overprediction is a consequence of sites from the warming interval after the LGM. Clearly, improvement in dating these sites should be a priority for conservation paleobiologists because of their importance for understanding the dynamics of glacial refugia. The question of LGM refugia in these and other fossil mammals will not be completely resolved until this age control improves; however, our conceptual and statistical approach offers unique insights into the hindcast method for informing transference in ENMs. It is important to remember that these fossil localities are not samples of single specimens: they are samples of an entire population of animals. Each fossil occurrence represents a large number of actual specimens, often numbering in the 100s to 1000s. One consequence to this sampling that has been overlooked in previous studies to consider LGM hindcasts (Martínez-Meyer et al. 2004, Guralnick and Pearman 2010, McGuire and Davis 2013) is the importance of false positives (class b in Table 1, 2) in evaluating the success of hindcasts. Hindcast areas that include a large number of fossil sites at which the species in question has not been found have very poor positive predictive power (Table 3). In previous work, the success of hindcasts has only been measured by their sensitivity, the rate of correct prediction of known fossil occurrences; we recommend that conservation paleobiogeography make full use of the fossil record and evaluate hindcasts for their overall correct prediction rate.

We propose that the southern LGM refugia predicted by these ENM reconstructions are mismatches with the true LGM ranges and arise either from insufficiencies in the modelling methods, inaccuracies in the reconstructed paleoclimate surface, or fallacy of the assumption that temperature and precipitation are the main factors limiting mammalian distributions. As shown in Table 1, across the five species we have sampled, there are on the order of 300 false positives – sites that occur within the refugia predicted by ENM hindcasts that do not contain the species included in this study. Again, these localities contain typical LGM faunas, not unusual in their preservational mode, and provide an adequate sample size to make a systematic geographic bias highly improbable. Additionally, we consistently find false negatives (class c in Table 1, 2) in localities that are north of the ENM hindcasts. These northern sites might have been captured by the models if the thresholds used by Waltari et al. (2007), >5 for GARP and 10 for MaxENT, were relaxed, allowing a larger hindcast area by the ENMs. If thresholds are set too high, they might be expected to hindcast ranges only in regions in which a species is relatively common. However, lower-threshold hindcasts would also contain a set of false negatives greater than the current number. Consequently, lower thresholds would be unlikely to make the models significantly better than the null model. The thresholds used by Waltari et al. (2007) are reasonable, based on ENMs of extant taxa (Pearson and Raxworthy 2007), and at any rate cannot account for the preponderance of southern range reconstructions: if these species were truly common in the hindcast regions, we would expect to see proportionately higher occurrences in those regions. It is clear the hindcasts are reconstructing ranges with a southern bias.

Our findings are in agreement with those of Guralnick and Pearman (2010), McGuire and Davis (2013), and Williams et al. (2013), but differ importantly from those of Martínez-Meyer et al. (2004). The latter study found widespread evidence of niche stability, which was interpreted as a high level of ENM predictive power across alternative climate regimes. There are several important recent changes in approach that could explain this discrepancy between recent LGM hindcast studies and the seminal Martínez-Meyer et al. (2004) study. First are differences in LGM sample size. Our current study includes only mammals that are well-represented in the LGM fossil record; no taxon examined by Martínez-Meyer et al. (2004) is present in more than 8 unique localities, while our locality sample sizes range from 16 to 37 (average 25) in the inclusive window. Simlarly, Guralnick and Pearman (2010) included no species with fewer than 12 LGM occurrences, with an average of 18. The sample size of McGuire and Davis (2013) was much smaller, because it was taxonomically focused. Further, as we have discussed, we included in our analysis sites from which fossils of the appropriate age have been uncovered but that have not yielded fossils of the taxon in question. In this way, our study serves as a test of ENM predictive power, instead of a test of niche conservatism, as in Martínez-Meyer et al. (2004).

The second important difference between the approach of Martínez-Meyer et al. (2004) and more recent studies relates to the temporal window of the LGM. Martínez-Meyer et al. (2004) included specimens from localities dated from between 20.5 and 14.5 ka. Guralnick and Pearman (2010) used a window that emphasized the full glacial, 20 to 15 ka, but allowed sites with bracketing dates from the interval 35 to 10 ka. Our window (40–17 ka) is less constrained than either of these studies, but, as explained in the Material and methods, we structured our window to encompass conditions within 4.6°C of the LGM value according to Petit et al. (1999). These different approaches produce different sample sizes, of course, but also sample from warmer climates closer to modern conditions, with temperatures extending to 5.6°C above the LGM for Martínez-Meyer et al. (2004) and 8.6°C above the LGM for the bracketing interval of Guralnick and Pearman (2010).

Finally, Martínez-Meyer et al. (2004) determined their significance cutoff using a Monte Carlo method to subsample modern assemblages, but this approach may not address the problem of projecting ENMs onto alternative climate regimes. The fossil record of the LGM does not simply consist of presence data: it includes a robust set of collections that can be used to explicitly test for over-prediction.

Ultimately, our approach focused on the technical success of transference of ENMs, and not on the strict niche conservation question of Martínez-Meyer et al. (2004). So, while niches may be conserved over this interval (but see Maguire 2013), we suggest that our methods provide a more direct test of the practicality of ENM methods for predicting new ranges. Our results agree with those of Williams et al. (2013) who, with an approach complementary to our hindcasting, found that a hypothetical Ice Age ecologist charged with preserving biodiversity in the face of deglaciation following the most recent glacial cycle would not be able to establish effective preserves based on ENM predictions for the distributions of tree species.

While we find southern over-prediction in both flatland and montane species (Table 3), Guralnick and Pearman (2010) found a southern bias only in flatland species, with the notable exception of *Microtus montanus*, which showed analogous southern over-prediction to our results (Fig. 13.1B in Guralnick and Pearman 2010). They interpreted their result as either a lag in geographic range change in response to the slower rate of environmental change across open land in comparison to across elevational gradients, or in situ adaptation or broad ecological plasticity in the flatland species. We find this pattern in both flatland and montane species (Table 3), suggesting to us that southern over-prediction is a broader problem rather than just a flatland problem. Our finding allows us to eliminate one of the hypotheses Guralnick and Pearman (2010) offered to explain the pattern, the idea that latitudinal gradients in seasonality have a greater control of the distribution of flatland species than montane species. Instead, we hypothesize that the variation in southern over-prediction observed by Guralnick and Pearman (2010) is structured by species-specific sampling of climate distributions and not by flatland vs montane habitat: their montane group contained many habitat specialists (*Thomomys talpoides*, *Marmota flaviventris*, *Phenacomys intermedius*), and their flatland group contained many generalists (*Tamias striatus*, *Microtus pinetorum*, *Microtus ochrogaster*). See Mammalian Species accounts for details on the ecology of these species (Frase and Hoffmann 1980, Smolen 1981, Snyder 1982, McAllister and Hoffmann 1988, Stalling 1990, Verts and Carraway 1999). Our study, conversely, samples only generalists (Table 3). Ultimately, the problem could be one of generalists vs specialists, warranting further investigation. Future work analyzing LGM niche reconstructions for another continent, e.g. the European models of Fløjgaard et al. (2009), would provide new lines of evidence for untangling this problem, which may also be influenced by a continent's unique regional connectivity and/or history of glaciation.

In the end, we propose four possible causes for these patterns of biased prediction: 1) modern distributions may not reflect the full range of environmental conditions in which a species can survive and, taken alone, serve as poor predictors of their potential distributions under other climate regimes. 2) The environmental tolerances of mammalian species can evolve fast enough to have changed since the LGM, so modern distributions are poor predictors of deeper time distributions, but may still be good predictors of shallow time responses to climate change. The overlap in LGM ranges between two congeners whose ranges do not currently overlap – *G. sabrinus* and *G. volans* – suggests that this may be the case; the cooler climate of the LGM may have decreased the importance of competitive exclusion within *Glaucomys*. 3) The problem lies with the general circulation models (GCMs) used to reconstruct LGM climate, so the reconstructed ranges are biased southwards because of incorrect temperature and/or precipitation values near the continental glaciers. Other workers have suggested problems with GCM reconstruction of climate regimes near glaciers (Hyde and Peltier 1993, Jackson et al. 2000, McGuire and Davis 2013), so this problem must be addressed before the others can be considered. 4) The models used to hindcast ranges are based on correlations between climatic variables and occurrence data in modern ecosystems. Both GARP and MaxENT overparamaterize these correlations, making them powerful tools for estimating ranges of modern taxa. However, the changing distribution and relationships between climatic variables mean that modern ecosystems may not be appropriate analogs for alternative climatic regimes, such as those that existed during the LGM and that may result from future warming.

Future modelling efforts must incorporate information about the biogeographic history of a region as well as interactions between competitors, predators, vegetation, and human land use within modern ecosystems to properly predict the reorganization of landscape-scale species distributions (Zurrell et al. 2009, Lavergne et al. 2010, Dawson et al. 2011, Fløjgaard et al. 2011). It is also clear that the possibility of large fundamental niches that are poorly sampled by modern distributions cannot be rectified without adding data from other climate regimes than those that currently exist. Future work must include reference to fossil occurrences and paleoenvironments, finding new methods that optimize ENMs in the context of both modern and fossil climate data (Varela et al. 2011). As GCMs improve, providing transient models of changing glacial climate (Ganopolski et al. 2010, Abe-Ouchi et al. 2013), conservation paleontologists can leverage these climate hindcasts to test models of species distributions in the context of ever-improving age control of fossil sites. Palaeontological data provide an essential resource for conservation biology, not just for testing hypotheses based on modern data, but also for providing the largest possible dataset for making informed conservation policy decisions (Hadly and Barnosky 2009).

Acknowledgements – Portions of this work were completed while JLM was a fellow at NESCent, funded by NSF #EF-0905606. We wish to thank S. Normand, P. D. Polly, R. Guralnick, A. Barnosky, and C. Marshall for constructive comments on earlier versions of this manuscript. We also thank members of the Hopkins Lab at UO and the Barnosky Lab at UCB for critical discussion in the creation of this work. All fossil data included in this analysis are available from the FAUNMAP II database, online at <www.ucmp.berkeley.edu/faunmap/index.html>.

References

Abe-Ouchi, A. et al. 2013. Insolation-driven 100,000-year glacial cycles and hysteresis of ice-sheet volume. – Nature 500: 190–193.

Clark, T. W. et al. 1987. *Martes americana*. – Mamm. Species 289: 1–8.

Collins, W. D. et al. 2006. The community climate system model version 3 (CCSM3). – J. Clim. 19: 2122–2143.

Dawson, T. P. et al. 2011. Beyond predictions: biodiversity conservation in a changing climate. – Science 332: 53–58.

Dietl, G. P. and Flessa, K. W. 2011. Conservation paleobiology: putting the dead to work. – Trends Ecol. Evol. 26: 30–37.

Dolan, P. G. and Carter, D. C. 1977. *Glaucomys volans*. – Mamm. Species 78: 1–6.

Fielding, A. H. and Bell, J. F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. – Environ. Conserv. 24: 38–49.

Fløjgaard, C. et al. 2009. Ice age distributions of European small mammals: insights from species distribution modelling. – J. Biogeogr. 36: 1152–1163.

Fløjgaard, C. et al. 2011. Deconstructing the mammal species richness pattern in Europe – towards an understanding of the relative importance of climate, biogeographic history, habitat heterogeneity and humans. – Global Ecol. Biogeogr. 20: 218–230.

Frase, B. A. and Hoffmann, R. S. 1980. *Marmota flaviventris*. – Mamm. Species 135: 1–8.

Ganopolski, A. et al. 2010. Simulation of the last glacial cycle with a coupled climate ice-sheet model of intermediate complexity. – Clim. Past 6: 229–244.

George, S. B. et al. 1986. *Blarina brevicauda*. – Mamm. Species 261: 1–9.

Graham, R. W. and Lundelius Jr, E. L. 2010. FAUNMAP II: new data for North America with a temporal extension for the Blancan, Irvingtonian and early Rancholabrean, ver. 1.0. – <www.ucmp.berkeley.edu/faunmap/index.html>.

Guralnick, R. and Pearman, P. B. 2010. Using species occurrence databases to determine niche dynamics of montane and lowland species since the Last Glacial Maximum. – In: Spehn, E. M. and Korner, C. (eds), Data mining for global trends in mountain biodiversity. CRC Press, pp. 125–135.

Hadly, E. A. and Barnosky, A. D. 2009. Vertebrate fossils and the future of conservation biology. – In: Dietl, G. P. and Flessa, K. W. (eds), Conservation paleobiology: using the past to manage the future. Yale Univ. Press, pp. 39–59.

Hasumi, H. and Emori, S. 2004. K-1 coupled GCM (MIROC) description. – Center for Climate System Research, Univ. of Tokyo.

Hijmans, R. J. and Graham, C. H. 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. – Global Change Biol. 12: 2272–2281.

Holt, R. D. 1990. The microevolutionary consequences of climate change. – Trends Ecol. Evol. 5: 311–315.

Huberty, C. J. 1994. Applied discriminant analysis. – Wiley.

Huberty, C. J. and Olejnik, S. 2006. Applied Manova and discriminant analysis. – Wiley.

Hyde, W. T. and Peltier, W. R. 1993. Effect of altered boundary conditions on GCM studies of the climate of the Last Glacial Maximum. – Geophys. Res. Lett. 20: 939–942.

Jackson, S. T. et al. 2000. Vegetation and environment in eastern North America during the Last Glacial Maximum. – Quat. Sci. Rev. 19: 489–508.

Lavergne, S. et al. 2010. Biodiversity and climate change: integrating evolutionary and ecological responses of species and communities. – Annu. Rev. Ecol. Evol. Syst. 41: 321–350.

Maguire, K. C. 2013. Testing for ecological niche stability of mammalian species from the Last Glacial Maximum to present. – J. Vertebr. Paleontol. Prog. Abstracts 169A.

Martínez-Meyer, E. et al. 2004. Ecological niches as stable distributional constraints on mammal species, with implications for Pleistocene extinctions and climate change projections for biodiversity. – Global Ecol. Biogeogr. 13: 305–314.

McAllister, J. A. and Hoffmann, R. S. 1988. *Phenacomys intermedius*. – Mamm. Species 305: 1–8.

McGuire, J. L. and Davis, E. B. 2013. Using the palaeontological record of *Microtus* to test species distribution models and reveal responses to climate change. – J. Biogeogr. 40: 1490–1500.

McGuire, J. L. and Davis, E. B. 2014. Conservation paleobiogeography: the past, present, and future of ecosystem function. – Ecography 37: 1092–1094.

Merrit, J. F. 1981. *Clethrionomys gapperi*. – Mamm. Species 146: 1–9.

Pearman, P. B. et al. 2008. Prediction of plant species distributions across six millennia. – Ecol. Lett. 11: 357–369.

Pearson, R. G. and Raxworthy, C. J. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. – J. Biogeogr. 34: 102–117.

Petit, J. R. et al. 1999. Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. – Nature 399: 429–346.

Phillips, S. J. et al. 2006. Maximum entropy modelling of species geographic distributions. – Ecol. Model. 190: 231–259.

Smolen, M. J. 1981. *Microtus pinetorum*. – Mamm. Species 147: 1–7.

Snyder, D. P. 1982. *Tamias striatus*. – Mamm. Species 168: 1–8.

Stalling, D. T. 1990. *Microtus ochrogaster*. – Mamm. Species 355: 1–9.

Stockwell, D. R. B. and Peters, D. P. 1999. The GARP modelling system: problems and solutions to automated spatial prediction. – Int. J. Geogr. Inform. Syst. 13: 143–158.

Svenning, J.-C. et al. 2011. Applications of species distribution modeling to paleobiology. – Quat. Sci. Rev. 30: 2930–2947.

Varela, S. et al. 2011. Using species distribution models in paleobiogeography: a matter of data, predictors and concepts. – Palaeogeogr. Palaeoclimatol. Palaeoecol. 310: 451–463.

Vega, R. et al. 2010. Northern glacial refugia for the pygmy shrew *Sorex minutus* in Europe revealed by phylogeographic analyses and species distribution modelling. – Ecography 33: 260–271.

Verts, B. J. and Carraway, L. N. 1999. *Thomomys talpoides*. – Mamm. Species 618: 1–11.

Vrba, E. S. 1993. Turnover-pulses, the Red Queen, and related topics. – Am. J. Sci. 293A: 418–452.

Waltari, E. and Guralnick, R. P. 2009. Ecological niche modelling of montane mammals in the Great Basin, North America: examining past and present connectivity of species across basins and ranges. – J. Biogeogr. 36: 148–161.

Waltari, E. et al. 2007. Locating Pleistocene Refugia: comparing phylogeographic and ecological niche model predictions. – PLoS One 2: e563.

Wells-Gosling, N. and Heaney, L. R. 1984. *Glaucomys sabrinus*. – Mamm. Species 229: 1–8.

Williams, J. W. et al. 2013. The ice age ecologist: testing methods for reserve prioritization during the last global warming. – Global Ecol. Biogeogr. 22: 289–301.

Zurrell, D. et al. 2009. Static species distribution models in dynamically changing systems: how good can predictions really be? – Ecogeography 32: 733–744.