Louys, J, Meloro, C, Elton, S, Ditchfield, P and Bishop, LC

Analytical framework for reconstructing heterogeneous environmental variables from mammal community structure.

http://researchonline.ljmu.ac.uk/id/eprint/247/

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Louys, J, Meloro, C, Elton, S, Ditchfield, P and Bishop, LC (2014) Analytical framework for reconstructing heterogeneous environmental variables from mammal community structure. Journal of Human Evolution, 78 (1). ISSN 1095-8606

Analytical framework for reconstructing heterogeneous environmental variables from mammal community structure

Julien Louys [a,b,]*, Carlo Meloro [b], Sarah Elton [c], Peter Ditchfield [d], Laura C. Bishop [b]

[a] *Department of Archaeology and Natural History, School of History, Culture and Languages, ANU College of Asia and the Pacific, The Australian National University, ACT 0200, Australia*

[b] *Research Centre in Evolutionary Anthropology and Palaeoecology, School of Natural Sciences and Psychology, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK*

[c] *Department of Anthropology, Durham University, Queen's Campus, Stockton, University Boulevard, Thornaby, Stockton-on-Tees TS17 6BH*

[d] *Research Laboratory for Archaeology and the History of Art, School of Archaeology, University of Oxford, Dyson Perrins Building, South Parks Road, Oxford OX1 3QY, UK*

* Corresponding author.

E-mail address: julien.louys@anu.edu.au (J. Louys).

Keywords: Palaeoenvironment; Laeotoli; Palaeoecology; Faunal community; Vegetation heterogeneity

Abstract

We test the performance of two models that use mammalian communities to reconstruct multivariate palaeoenvironments. While both models exploit the correlation between mammal communities (defined in terms of functional groups) and arboreal heterogeneity, the first uses a multiple multivariate regression of community structure and arboreal heterogeneity, while the second uses a linear regression of the principal components of each ecospace. The success of these methods means palaeoenvironment of a particular locality can be reconstructed in terms of the proportions of heavy, moderate, light, and absent tree canopy cover. The linear regression is less biased, and more precisely and accurately reconstructs heavy tree canopy cover than the multiple multivariate model. However, the multiple multivariate model performs better than the linear regression for all other canopy cover categories. Both models consistently perform better than randomly generated reconstructions. We apply both models to the palaeocommunity of the Upper Laetolil Beds, Tanzania. Our reconstructions indicate that there was very little heavy tree cover at this site (likely less than 10%), with the palaeo-landscape instead comprising a mixture of light and absent tree cover. These reconstructions help resolve the previous conflicting palaeoecological reconstructions made for this site.

Introduction

The reconstruction of past environments is one of the key objectives of palaeoecology (Louys et al., 2012), with community-based (synecological) methods being a major contributor to this effort (Andrews et al., 1979; Nesbit-Evans et al., 1981; Andrews, 1989, 1996; Kay and Madden, 1997; Reed, 1997, 1998; Kovarovic et al., 2002; Mendoza et al., 2005; Reed and Russak, 2009; Louys et al., 2009; Muldoon, 2010; Louys and Meijaard, 2010; Meloro and Kovarovic, 2013). Synecological methods use the structure of animal communities defined either ecologically or taxonomically to determine the habitats occupied. Specifically, the ecological structure of any animal community can be expressed as an *n*-dimensional space, i.e., by *n* multiple discrete variables that describe how the animals inhabit or utilise the environment or landscape in which they are found. The combination of these variables (= functional or taxonomic groups) describes the ecological space (ecospace) of the community.

The habitats that palaeoecologists seek to reconstruct are just as much multivariate ecospaces as the animal communities that inhabit them. However, in palaeosynecological analyses, environments are usually not described or reconstructed in this way. Rather, in order to facilitate comparisons between modern and fossil communities, these habitats are discretely categorised. For example, habitats can be categorised as forests, woodlands, or grasslands, and this has largely been affected by the restrictions of the multivariate methods employed (e.g., principal components analysis, principal coordinates analysis, and discriminant function analysis). By using categorisations, palaeoecologists implicitly acknowledge that these are a means of partitioning a continuous multidimensional spectrum of environmental conditions into manageable units. However, the explicit reconstruction of the multidimensional nature of the palaeoenvironment on the basis of communities has so far remained elusive (although see Cerling et al. [2011] for a different approach to reconstructing multivariate environments using stable isotopes). We contend that this has been one contributor to the description of many palaeoenvironments as ‘mosaics’ or ‘mixed’ habitats.

Alternatively, habitats and environments can be described by any number of biotic and abiotic quantitative variables. This is the cornerstone of landscape and community ecology (Jongman et al., 1995). In a recent paper, we described the multivariate relationship that exists between arboreal heterogeneity and mammal community structure (Louys et al., 2011). We analysed two separate ecospaces, mammal community and arboreal heterogeneity, both of which occupied the same geographical area. For a selection of sixty-three natural protected areas spread amongst the continents of Africa, Asia, and South and Central America, we determined the relative amounts of canopy cover of trees (arboreal heterogeneity) as well as the structure of the mammal community, in turn derived from species lists from those areas. The two separate ecospaces—one vegetational and the other faunal—were compared both between and within continents, and the relationships between them explored. A linear and significant relationship between the ecological guild of small arboreal and semi-arboreal secondary consumers and the relative proportion of continuous canopy cover was found when all continents were considered together (the categories BAS and %Heavy, respectively, see Table 1). The amount of absent tree cover was also consistently correlated with mammal community structure, especially with relative percentage of large terrestrial primary consumers.

Here, we demonstrate how this relationship can be applied to the fossil record, exploring how mammal communities can be used to retrodict arboreal heterogeneity. Although in this paper we only retrodict arboreal heterogeneity as one multidimensional biotic variable, the methods we outline could easily be expanded to include other environmental biotic and abiotic variables. We provide an outline of the techniques that could be employed, as well as exploring the limitations of this new method.

Materials and methods

Abbreviations used in this study are listed in Table 1. Ecospaces were defined from information on sixty-three natural protected areas (NPAs) in Central and South America (hereafter 'America;' *n* = 8), Africa (*n* = 23), and Asia (*n* = 32). Mammalian species lists and geographical coordinates for NPAs were taken from the Man and the Biosphere Species Database (http://ice.ucdavis.edu/mab; see Supplementary Online Material [SOM]). The locations of these protected areas are shown in Louys et al. (2011:Fig. 1). Only species lists with more than 32 species were used; this number has been suggested as the likely minimum number necessary to confidently distinguish between three discrete and broadly defined habitat types across ecosystems (namely closed, mixed, and open; Louys et al., 2009). The palaeoecology of the Upper Laetolil Beds was examined on the basis of the new method described below. The faunal list for this site was obtained from published sources (Harrison, 2011).

Ecological categories and habitat classification

Following Louys and Meijaard (2010) and Louys et al. (2011), we restricted our faunal categorisation to three ecological categories: 1) body mass, divided into small (B; 1–10 kg), medium (C; 10–45 kg), large (D; 45–180 kg), and very large (E >180 kg); 2) trophic level—primary (P) or secondary (S) consumer; 3) locomotion, either strictly terrestrial (T) or potential and strict arboreality (A; which we refer to hereafter as “arborophilic;” see below). Mammals with mean body mass less than 1 kg as well as all bats were removed, as these species sample and interact with the environment differently than larger arborophilic or terrestrial animals, and are less likely to be preserved in fossil assemblages alongside larger mammals (Damuth, 1982). Removing bats and tiny mammals follows the procedure set out in other synecological studies (e.g., Andrews et al., 1979; Kay and Madden, 1997; Rodríguez, 2004; Louys and Meijaard, 2010; Louys et al., 2011).

Our method uses a binary system when defining ecological guilds: a species either falls in one strictly defined category, or it does not. We use a very strict definition of terrestriality (Louys et al., 2011), such that animals with even implied arboreality are termed ‘arborophilic.’ For example, the porcupine (*Hystrix)* is not usually considered a tree dwelling species. However, Nowak (1999:1647) states that the porcupine “does not usually climb trees,” implying that they can climb, and hence have arboreal capabilities. Therefore we would not classify this species as strictly terrestrial (T) in our analysis, but rather potentially arboreal, i.e., arborophilic (A). This definition encompasses not only mammals that are dependent on trees for survival but also those that may potentially use trees even sporadically, such as for predator avoidance or occasional food resources. We likewise use a very strict definition of primary consumer. Ecological data for modern species were taken from Nowak (1999). This reference is comprehensive and widely available, and using it as the source of information for all taxa helps to ensure consistency of classification. We provide a list of all modern species scored according to our scheme in the Supplementary Online Material (Tab 1 of SOM spreadsheet).

The ecological categories described above were combined into fifteen discrete functional groups, and the relative proportion of each species in the community from each protected area was calculated. Our functional groups are less detailed than those used previously by other researchers (e.g., Reed, 1997, 1998; Kovarovic et al., 2002; Mendoza et al., 2005; Rodríguez et al., 2006; Reed and Russak, 2009; Meloro and Kovarovic, 2013), however they still provide accurate retrodictions (Louys and Meijaard, 2010). Moreover, these functional groups are simple, binary (for trophic group and locomotion), and almost always unequivocal. As such, fossil taxa whose palaeobiologies are largely unknown and which share no modern analogues can be assigned to functional groups with a high degree of confidence.

Vegetation structure was classified following the method described by Louys et al. (2011). Using Google Earth we captured a satellite image for each modern NPA comprising an area 25 km x 25 km (625 km^2) centred on its geographical coordinates. Each of four different types of tree cover (heavy, moderate, light, and no tree cover), or, arboreal heterogeneity (Fig. 1), was determined from these images using ArcGIS software (ESRI vs. 9.3.1). Independent signatures, recording the image properties of a defined, discrete area were calculated for each NPA. Using a maximum likelihood analysis, these signatures were subsequently used to classify the 625 km^2 area from each NPA according to the different types of tree cover. Classifications were re-examined visually to ensure maximum accuracy was achieved and independently verified by comparing our vegetation categories with the climatic data for the geographical locations of each park (Louys et al., 2011). The full list of raw proportions for both faunal and vegetation ecospaces is provided in the Supplementary Online Material, and abbreviations for each category are listed in Table 1. Statistical analyses were performed using PAST (v. 2.14; Hammer et al., 2001), Spatial Analysis in Macroecology (SAM v. 4.0; Rangel et al., 2010), and Microsoft Excel.

Standardising the ecospaces

To make the fossil communities comparable to modern communities, the raw percentage values for each of the functional groups and arboreal categories were standardised. Standardisation is a commonly employed technique in community and landscape ecology to make data more comparable and/or to ensure a better fit to statistical models (Jongman et al., 1995). In the linear regression for faunal and vegetation ecospaces the values were transformed through Principal Components Analyses (PCA) using a covariance matrix. For the multivariate regression analysis, and following Warton and Hui (2011), the logit transformation was used. It is defined by:

$$E_t = \log (E/1-E)$$

where E_t is the transformed (standardised) value, and E is the raw ecological variable expressed as a proportion (i.e., between 0 and 1). In order to deal with values equal to 0 and 1 in the transformation, Warton and Hui (2011) further suggest adding a small value, ε, to both numerator and denominator. The logit transformation used in our analyses was therefore:

$$E_t = \log ((E+ \varepsilon)/(1-E+ \varepsilon))$$

where ε is equal to the smallest proportion recorded for a functional group or vegetation category in each ecospace. For the faunal ecospace, $\varepsilon = 0.0147$, while for the vegetation ecospace, $\varepsilon = 0.0025$ (values are listed here to four decimal places, while full values were used in the analyses).

Spatial scale and autocorrelation

Many spatial data can exhibit spatial autocorrelation (SAC), in which observations close to each other in space violate a priori expectations of independence (Dormann et al., 2007). The causes of SAC are numerous, but statistical models that do not account for an environmental determinant that is itself spatially structured, and therefore causes spatial structuring in the response measured (Besag, 1974), are of particular significance to palaeoecological reconstructions. Testing for SAC is important because if observations are autocorrelated across space, statistical assumptions of independence are violated, and type I errors are inflated (Dormann et al., 2007). In order to test whether our data was subjected to SAC, we analysed the residuals of the regression between the first principal component of the faunal ecospace and the first principal component of the vegetation ecospace (Louys et al., 2011:Fig. 3b). Moran’s I was calculated for these residuals. Moran’s I is a commonly employed measure of spatial autocorrelation (Dormann et al., 2007), with values of 1 indicating perfect correlation between locations in space, and values of -1 indicating perfect dispersion. Values significantly different from 0 in our data would therefore indicate the presence of spatial autocorrelation at a given scale.

Detailed statistical examinations of differences between the global and continental datasets were presented in Louys et al. (2011), and interested readers are directed to that publication for a discussion of differences in response variables at the continental level. In short, however, one of the major findings of Louys et al. (2011) was that mammal communities exhibited convergent structure in the presence of similar environments regardless of geographical location. Furthermore, no major differences were observed between the data examined at the global level and the nested African data. Nevertheless, we argued that in order to account for potential differences between modern and fossil systems, a global dataset should be used for retrodictions whenever possible (Louys et al., 2011). In order to take into account the potentially compounding effects of examining SAC across a global dataset, however, the subset of African NPAs was also examined for Moran’s I.

The models

Here we are interested in predicting all four categories of arboreal heterogeneity (i.e., heavy, moderate, light, and no tree canopy cover). Two different models are used and compared. For the first, we use a multiple multivariate regression (MMR) between the complete standardised faunal ecospace and the standardised vegetation ecospace. This technique essentially performs four multivariate regressions simultaneously, with the overall multivariate test of significance computed using the Wilks' lambda statistic. Using this relationship, the standardised vegetation for each NPA can be calculated from the standardised fauna, and the standardisation removed by solving the logit equation for E. This yields the reconstructed vegetation values for the MMR. In the second model, we explore the relationship between the first principal component (PC) of the vegetational ecospace and the first PC of the mammal community ecospace using a linear regression model (LRM). By applying this relationship, the vegetation PC for each NPA can be calculated on the basis of the faunal PCs, and running the vegetation PCA in reverse yields the reconstructed vegetation values of the LRM. In order to evaluate the performance of the different algorithms, a third model producing entirely random values for each of the vegetation categories was run and its results compared.

Model performance

Any palaeoecological model used to reconstruct quantitative variables needs to be tested for performance; that is, how well does it reconstruct the original variables? Three performance indicators can be used to assess the efficacy of a model: bias, precision, and accuracy (Walther and Moore, 2005). Bias refers to the difference between the mean of a set of measurements or test results and the accepted reference or true value (Bainbridge, 1985), and bias may lead to a systematic under- or overestimation of true values. In our study, the bias being tested is one of measurement bias—in other words, how well the different methodologies used reconstruct the true value from which the models are derived. Two bias estimators are examined here: mean error (ME) —the mean of all the differences—and percentage overestimated (PO), the percentage of overestimations.

Precision refers to the absence of random error (Walther and Moore, 2005). Unlike bias or accuracy, it is independent of the real values and instead is derived purely from the variance produced by the measurement procedure. Three common precision measures are examined here: coefficient of variation (CV), standard deviation (SD), and inter-quartile range (IQR: the difference between the 25th and 75th quartile). To be a good predictor, the model should be as precise as the original dataset; hence, the precision measures of the real values are also listed.

Accuracy refers to how close a predicted value is to the real value (Walther and Moore, 2005). Two common accuracy measures are examined here: root mean square error (RMSE) —the square root of the mean of the squared differences between real and predicted values; and mean absolute error (MAE).

Application to the fossil record

In order to demonstrate the application of this method we applied it to a fossil community from the Pliocene of East Africa (derived from the Upper Laetolil Beds). This assemblage was chosen because it: (1) contains greater than 32 species, the suggested number of species necessary to ensure confident discrimination between habitat types (Louys et al., 2009); (2) has been variously interpreted as ‘mosaic’ or ‘mixed’ habitat (see below); and (3) is one of the best-known Pliocene hominin sites and hence is of profound interest for studies of human evolution. Furthermore, the faunal community derived from the Upper Laetolil Beds shows remarkable stasis throughout the 300 ka period of deposition, indicating little ecological diversity over this time (Su and Harrison, 2007). Hence, the faunal community preserved is very likely to be representative of the actual palaeocommunity. Fossil species were assigned to the same functional groups as modern taxa. In the case of extinct species, group assignment was based on previously published palaeobiological analyses if available. If such data were not available, group membership was inferred from the most closely related taxa, within the same genus wherever possible.

Results

Spatial autocorrelation

The plots of Moran's I show that the residuals calculated from the regression between the faunal (=X) and vegetation (Y) ecospaces were not significantly different from 0 (Table 2). Therefore, our data do not exhibit spatial autocorrelation. This result is reflected at the restricted geographical scale when only African NPAs were examined (Table 2).

The models

The MMR using all logit variables as predictors was significant (Wilk's λ = 0.09428, F = 2.41, DF = 60, 174, $p < 0.001$). The regression coefficients of the MMR are shown in Table 3. The regression between faunal and vegetation principal components (per Louys et al., 2011:Fig. 3b) yields the following relationship used in the LRM:

$$\mathrm{VegPC1} = \mathrm{FaunPC1}/0.29658$$

The summary statistics for the principal components analyses are listed in Tables 4 and 5. The performance measures for both models are shown in Table 6. Graphic representations of the relationship between real, reconstructed, and randomly generated values are shown in Figures 2 and 3.

Bias

Mean error for the %Heavy category in the model that uses randomly generated values (hereafter ‘random model’) is moderate, reflecting that real values for this category are spread throughout the spectrum from 0% to 100%; this is demonstrated by the relatively high standard deviation of real values compared to the other categories. Most values, however, are on the higher end of the scale—in other words, the majority of the NPAs examined have a significant percentage of closed canopy in the areas examined (median %Heavy = 73%). The random model thus tends to underestimate the amount of closed canopy. The MMR has a higher mean error (in the negative direction) than the random model, which is due to a tendency of this algorithm to overestimate the amount of heavy canopy cover. This effect can be seen clearly in Figure 2a, where the MMR reconstruction line sits above the real line for most NPAs. The LRM has a negligible mean error for %Heavy; however, it does tend to overestimate the amount of canopy cover, although it is not far from the desired 50%. Random mean error for both %Moderate and %Light is very high, reflecting the fact that most NPAs examined have minimal amounts of both canopy cover types. This is supported by the random overestimates for these categories, which are in the high 80% and low 90%. For the MMR algorithm, both mean errors for %Moderate and %Light are relatively low, and this model boasts the least amount of overestimation for these categories. This can be seen in Figure 2b and c; nevertheless, the overestimation is higher than desirable. The LRM again shows insignificant mean error; however, it has significantly higher rates of overestimation than MMR. Finally, the mean error for %Absent in the random model is quite negative, a result of the distribution of this category in the sample. Like %Heavy it has a high standard deviation, indicating polarisation of this category amongst the sampled NPAs. For the random model this is reflected in the percentage overestimated, showing that most parks have little %Absent; however, those that do have considerable amounts. The mean error for MMR for %Absent is relatively small, and this model tends to underestimate this category. Overall the LRM has the least biased reconstruction of this category, having insignificant minimal mean error and an overestimation value close to 50%.

Precision

The coefficients of variation, standard deviations, and inter-quartile ranges for a completely random distribution of proportions of canopy cover are approximately 60, 30, and 50–60, respectively. The real coefficient of variation for %Heavy is close to the random value, indicating that this category has a large range of canopy cover proportions. The model whose coefficient of variation is closest to the real one is the LRM, although both models are below both real and random variation. The LRM has a standard deviation closest to that of the random model for %Heavy, while the MMR approaches that of the real. The inter-quartile range for %Heavy, perhaps the most informative of the precision measures for this dataset, indicates that the real range is much higher than the random range. The model which comes closest to this value is the LRM, with the MMR range being significantly below both real and random ranges. Unlike %Heavy, the coefficient of variation for the real values of %Moderate and %Light are much higher than the random model, although standard deviations are similar. The MMR most closely approaches the coefficient of variation of the real values, while the LRM is significantly below. Both the LRM and MMR have much lower standard deviations for these two categories than either real or random values. The most precise model for %Moderate and %Absent as judged by the inter-quartile range is the MMR, although the range shown by the LRM is still significantly smaller than the random model. Finally, the MMR has precision levels closest to that of the real distribution for %Absent, although the LRM precision is significantly better than the random model.

Accuracy

The LRM has the lowest root mean square error for %Heavy and %Light; its errors for %Moderate and %Absent are slightly higher than the MMR. Both are significantly better than the random model (F = 65.17, DF = 2, 12, $p < 0.001$). The mean absolute error is again lower in the LRM for %Heavy, while for the remaining three categories the MMR is more accurate. Again, both algorithms are far superior to the random model.

Palaeoenvironment retrodiction

The assignment of Laetoli species to functional groups is shown in Table 7. The proportional representation in each functional guild is shown in Table 8. Reconstruction using the MMR suggests that the assemblage from the Upper Laetolil Beds represents an environment with 0.78% Heavy tree cover, 4.21% Moderate tree cover, 9.75% Light tree cover, and 96.51% Absent tree cover. On the other hand, reconstruction using the LRM indicates an environment with 0.52% Heavy tree cover, 7.79% Moderate tree cover, 31.70% Light tree cover, and 59.98% Absent tree cover. Both models are congruent with the limited proportion of Heavy tree cover (10.27%) determined using the relationship between BAS and %Heavy (Louys et al., 2011):

$$\%Heavy = 2.9809(BAS) - 18.124$$

Therefore, our study suggests that in an area approximately 625 km^2 around this site canopy cover was predominantly absent to light, and that dense tree cover comprised a very limited, although still present, proportion of the palaeo-landscape.

Discussion

Model performance and comparison

Both models can successfully reconstruct the heterogeneity of the vegetation on the basis of mammal community structure and are clearly more precise than the random model, although each model suffers from some loss of accuracy and precision. The mean error is lower across all categories for the LRM compared to either the random model or MMR. The MMR almost always overestimates %Heavy, while the LRM tends to underestimate %Heavy, albeit with a value close to 50%. This is similar for %Absent, with the LRM overestimating with a value close to 50%. The MMR tends to underestimate %Absent. Overall the LRM is less biased than the MMR, although the MMR shows considerably less bias than a purely random model. Neither the MMR nor the LRM could be said to be more precise than the other overall: while the LRM appears more precise for %Heavy, by the same measures the MMR is more precise for the remaining three categories.

Comparison of the models shows that whereas the LRM achieves greater precision and accuracy for %Heavy than the MMR, the MMR more accurately and precisely reconstructs %Moderate, %Light, and %Absent. The LRM has two further advantages over the MMR. In the LRM, the use of PCA ensures that the sum of the reconstructed proportions always equals 1 when the model is mapped to the vegetation ecospace. For the MMR this is not always the case. In addition, inputting more than four environmental variables into the LRM is a relatively straightforward procedure—the vegetation ecospace can be expanded to a more general environmental ecospace by adding as many biotic and abiotic variables as required, and the environmental PCA readily calculated. Finally, it is worth noting that only the first vegetation PC correlated significantly with any other faunal PCs. If variables, independent of faunal community but equally interpretable from the fossil record, could be found that significantly correlate with the second (or third) vegetation PC, this would allow its reconstruction from the geological record. Therefore, this PC could be incorporated along with PC1 into the reverse vegetation PCA calculation, adding the amount of variance explained by that PC into the model and thereby increasing its predictive power.

Some might argue that the high range of values that can be produced by the different models is not a big improvement on current palaeosynecological techniques. For example, for Laetoli the %Light tree cover ranged from 9.75% (MMR) to 31.70% (LRM), while %Absent ranged from 96.51% (MMR) to 59.98% (LRM). These ranges could be argued to be little better than reconstructing the Laetoli palaeohabitat as 'open.' However, we contend that the use of these models provides a reasonable range of values for each given arboreal category. If necessary, more weight can be given to the MMR results, given its higher accuracy and precision in these categories than the LRM. Furthermore, and beyond the use of these methods in retrodicting past environmental conditions, we predict that a major strength of the approach we outline here will be in examining quantitative habitat differences between sequentially deposited palaeocommunities. Finally, given that the models reconstruct the proportions differently, therefore giving slightly different answers, we can verify whether any given reconstruction is reasonable. Any imbalanced reconstruction that may arise due to taphonomic bias or radically different palaeocommunities should be picked up by at least one of the models, and any major discrepancies investigated further. For example, if the %Heavy category is retrodicted as 168.34%, then this could indicate either taphonomic bias or a palaeocommunity completely different to what exists today in Africa, Asia, and South and Central America. Significant differences between models (for example, if the MMR model retrodicts %Heavy at 98.76% while the LRM retrodicts this category as 0.23%) could equally indicate bias or radically different communities.

More broadly, the methods we outline in this study are complementary to the many different means of reconstructing hominin palaeoenvironments, each of which possesses unique advantages and limitations. Reconstructions of arboreal heterogeneity by our methods and the fraction woody cover using stable carbon isotopes in soils as outlined in Cerling et al. (2011) represent the most complementary approaches and allow us to examine these more or less synonymous environmental variables at different temporal and spatial scales. For example, the δ13C values in the modern soils examined by Cerling et al. (2011) reflect the amount of woody cover on a decadal timescale. In contrast, the mammalian palaeocommunities examined by synecological methods such as the ones we proposed here may represent deposition, and hence a timescale, covering tens to even hundreds of thousands of years. Autecological techniques, such as dental wear analyses, stable isotope analyses, and ecomorphology, could also be usefully applied to refine or confirm the ecological guilds we used. Alternatively, it might also be possible to apply the quantitative environmental categories described here to ecomorphological analyses seeking to reconstruct palaeoenvironments directly. Results from analyses such as ours would benefit enormously by comparisons with palaeoenvironmental data obtained from palynological studies, keeping in mind the different taxonomic, temporal, and spatial scales involved in the different methods.

Palaeoenvironment retrodiction

Laetoli, located in Tanzania at the southern edge of the eastern branch of the East African Rift Valley, comprises a succession of beds exposed along the margin of the Eyasi Plateau. The Laetolil Beds consist of two lithologic units: an upper and a lower unit. The upper unit (Upper Laetolil Beds) is composed of a series of aeolian and airfall tuffs, with the mammalian fossils recovered from this bed forming the basis of the material considered by this study. Previous palaeoenvironmental reconstructions of Laetoli have been quite contradictory. Many authors, on the basis of diverse evidence including geology, palynology, and palaeontology have suggested that the environments present during the deposition of the Laetolil beds were similar to those found today, namely arid to semi-arid grassland with patches of acacia woodland (see, for example, contributions by Hay, Bonnefille and Riollet, Gentry, Leakey, Meylan, and Watson in Leakey and Harris [1987]). Others, however, have suggested the presence of a much more significant proportion of dense bush cover and woodland, on the basis of palaeosynecology, diversity analysis, and stable carbon isotopes (see, for example, contributions by Butler, Petter, and Verdcourt in Leakey and Harris [1987]; Andrews, 1989; Reed, 1997; Su and Harrison, 2007). The most recent analyses, examining stable isotopes, mesowear, bovid ecomorphology, mammal community structures, and the bird fauna, describe the palaeoenvironment as a “vegetational mosaic with woodland, bushland, and grassland-savanna” (Harrison, 2011:12; see also Bishop et al., 2011).

Our analyses consistently suggest that the predominant vegetation type at Laetoli as represented by the large mammal fauna was absent and light tree cover. Both the LRM and MMR reconstruct small values for %Heavy, and these agree with a more simplified regression suggesting up to ~10% heavy tree cover. Given that the definition of arborophilic we use is highly inclusive, our estimates of absent or reduced tree cover at Laetoli are likely to be conservative. Hence, although clearly present, the amount of dense bush cover and woodland would have comprised only a small proportion of the palaeolandscape. Quantifying the likely proportions of tree coverage in this way refines previous reconstructions of palaeohabitat ‘mosaics’ at Laetoli.

Australopithecus afarensis survived in Pliocene East Africa for around 900,000 years and is found at Laetoli and Hadar, but also less abundantly at Dikika, Maka, Omo, Fejej, Lothagam, Tabarin, West Turkana, and Koobi Fora (reviewed in Grine et al., 2006). Palaeoenvironmental reconstructions indicate that it would have been exposed to diverse habitats, even though microwear of *A. afarensis* dental specimens from different sites and throughout the time sequence suggests dietary stasis (Grine et al., 2006). The recovery of a species from a region, or more localised area, with diverse habitats does not necessarily imply that it exploited them all. Thus, *A. afarensis* could have selectively exploited its environment (differentially preferring habitats with either heavy, light, or absent tree cover). Alternatively, it could have used its environment much more flexibly, moving and foraging through a range of different environments depending on local competition or season. Based on the observation that *A. afarensis* is found, reasonably abundantly, throughout the whole

Hadar Formation, during which palaeoenvironmental reconstructions indicate a shift from woodland to wet then dry grassland, it has been argued that the species did not favour one of these habitats in particular (Grine et al., 2006). Our reconstruction of palaeohabitat at Laetoli, another site where it is well represented, suggests that it thrived in lightly wooded to open, potentially xeric, habitats. Indeed, in contrast to palaeoenvironmental reconstructions for the slightly younger *Australopithecus africanus*, and despite forelimb adaptations indicating arboreality (Stern and Susman, 1983), the evidence for *A. afarensis* inhabiting areas of heavy tree cover in any part of its range and at any time in its tenure is equivocal (*sensu* Elton, 2008).

Conclusions

The methods we have introduced can be used by palaeoecologists to reconstruct palaeoenvironments in a multidimensional and quantifiable way. In other words, the use of these methods will allow researchers to move away from describing palaeoenvironments as ‘mosaics’ or ‘mixed habitats’ to detailing the specifics of the heterogeneity of the habitats in which the animals evolved. The methods we outline specify the abundance of particular canopy cover categories in the area inhabited by a mammal community at a fixed scale (in our case, 625 km^2). This improved insight into past ecologies has the potential to allow more detailed and analytical examination of the specific biotic and abiotic factors which lead to the modern mammalian-dominated ecosystems. It will be particularly useful when comparing changing environments through successive stratigraphic or chronological units in the geological record. However, it does not address the distribution of canopy cover on the landscape, nor does it provide any indication of the scale at which any particular feature becomes significant for organisms under investigation. Nevertheless, this research presents a solid framework for measuring abundances of environmental variables on the basis of mammal community structure, and hopefully stimulates new research and approaches in palaeoenvironmental reconstruction using the vertebrate fossil record.

Acknowledgements

We wish to thank the Leverhulme Trust for their generous support of this research (F/00 754/C). The anonymous reviewers kindly provided many detailed comments which helped improved this manuscript, and we thank them, the Associate Editor, and the Editor Mark Teaford for their valuable input.

Appendix. Supplementary material

Supplementary data related to this article can be found online at doi:xxx

References

Andrews, P. J., 1989. Palaeoecology of Laetoli. J. Hum. Evol. 18:173–181.

Andrews, P. J., 1996. Paleoecology and hominid paleoenvironments. Biol. Rev. 71:257–300.

Andrews, P. J., Lord, J. M., Nesbit-Evans, E. M., 1979. Patterns of ecological diversity in fossil and modern mammalian faunas. Biol. J. Linn. Soc. 11:177–205.

Bainbridge, T. R., 1985. The Committee on Standards: precision and bias. ASTM Standardisation News 13:44–46.

Besag, J., 1974. Spatial interaction and the statistical analysis of lattice systems. J. R. Stat. Soc. B 36:192–236.

Bishop, L. C., Plummer, T. W., Hertel, F., Kovarovic, K., 2011. Paleoenvironments of Laetoli, Tanzania as determined by antelope habitat preferences. In: Harrison, T. (Ed.), Paleontology and Geology of Laetoli: Human Evolution in Context. Springer, Dordrecht, pp. 355–366.

Cerling, T. E., Wynn, J. G., Andanje, S. A., Bird, M. I., Korir, D. K., Levin, N. E., Mace, W., Macharia, A. N., Quade, J., Remien, C. H., 2011. Woody cover and hominin environments in the past 6 million years. Nature 476:51–56.

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

Dormann, C. F., McPherson, J. M., Araujo, M. B., Bivand, R., Bolliger, J., Carl, G., Davies, R. G., Hirzel, A., Jetz, W., Kissling, W. D., Kuhn, I., Ohlemuller, R., Peres-Neto, P. R., Reineking, B., Schroder, B., Schurr, F. M., Wilson, R., 2007.

Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30:609–628.

Elton, S., 2008. The environmental context of human evolutionary history in Eurasia and Africa. J. Anat. 212:377–393.

Grine, F. E., Ungar, P. S., Teaford, M. F., El-Zaatari, S., 2006. Molar microwear in *Praeanthropus afarensis*: Evidence for dietary stasis through time and under diverse paleoecological conditions. J. Hum. Evol. 51:297–319.

Hammer, Ø., Harper, D. A. T., Ryan, P. D., 2001. PAST: Paleontological statistics software package for education and data analysis. Palaeontologia Electronica 4(1):9pp.

Harrison, T., 2011. Introduction: The Laetoli hominins and associated fauna. In: Harrison, T. (Ed.), Paleontology and geology of Laetoli: human evolution in context: volume 2. Springer, Dordrecht, pp. 1–14.

Jongman, R. H. G., ter Braak, C. J. F., van Tongeren, O. F. R., 1995. Data analysis in community and landscape ecology. Cambridge University Press, Cambridge.

Kay, R. F., Madden, R. H., 1997. Mammals and rainfall: Paleoecology of the middle Miocene at La Venta (Colombia, South America). J. Hum. Evol. 32:161–199.

Kovarovic, K., Andrews, P., Aiello, L., 2002. The palaeoecology of the Upper Ndolanya Beds at Laetoli, Tanzania. J. Hum. Evol. 43:395–418.

Leakey, M. D., Harris, J. M., 1987. Laetoli, a Pliocene site in northern Tanzania. Clarendon Press, Oxford.

Louys, J., Meijaard, E., 2010. Palaeoecology of Southeast Asian megafauna-bearing sites from the Pleistocene and a review of environmental changes in the region. J. Biogeogr. 37**:**1432–1449.

Louys, J., Meloro, C., Elton, S., Ditchfield, P., Bishop, L. C., 2011. Mammal community structure correlates with arboreal heterogeneity in faunally and geographically diverse habitats: implications for community convergence. Global Ecol. Biogeogr. 20:717–729.

Louys, J., Travouillon, K. J., Bassarova, M., Tong, H., 2009. The use of protected natural areas in palaeoecological analyses: assumptions, limitations and application. J. Archaeol. Sci. 36:2274–2288.

Louys, J., Wilkinson, D. M., Bishop, L. C., 2012. Ecology needs a palaeontological perspective. In: Louys, J. (Ed.), Paleontology in Ecology and Conservation. Springer-Verlag, Berlin, pp. 23–38.

Meloro, C., Kovarovic, K., 2013. Spatial and ecometric analyses of the Plio-Pleistocene large mammal communities of the Italian peninsula. J. Biogeogr. 40:1451–1462.

Mendoza, M., Janis, C. M., Palmqvist, P., 2005. Ecological patterns in the trophic-size structure of large mammal communities: a ‘taxon-free’ characterization. Evol. Ecol. Res. 7:505–530.

Muldoon, K. M., 2010. Paleoenvironment of Ankilitelo Cave (late Holocene, southwestern Madagascar): implications for the extinction of giant lemurs. J. Hum. Evol. 58:338–352.

Nesbit-Evans, E. M., van Couvering, J. H., Andrews, P., 1981. Palaeoecology of Miocene sites in Western Kenya. J. Hum. Evol. 10:35–48.

Nowak, R. M., 1999. Walker's mammals of the world. The Johns Hopkins University Press, London.

Rangel, T. F., Diniz-Filho, J. A. F., Bini, L. M., 2010. SAM: A comprehensive application for Spatial Analysis in Macroecology. Ecography 33:1–5.

Reed, K. E., 1997. Early hominid evolution and ecological change through the African Plio-Pleistocene. J. Hum. Evol. 32:289–322.

Reed, K. E., 1998. Using large mammal communities to examine ecological and taxonomic structure and predict vegetation in extant and extinct assemblages. Paleobiology 24:384–408.

Reed, K. E., Russak, S. M., 2009. Tracking ecological change in relation to the emergence of *Homo* near the Plio-Pleistocene boundary. In: Grine, F. E., Fleagle, J. G., Leakey, R. E. (Eds.), The first humans: origin and early evolution of the genus *Homo*. Springer, Dordrecht, pp. 159–171.

Rodríguez, J., Hortal, J., Nieto, M., 2006. An evaluation of the influence of environment and biogeography on community structure: the case of Holarctic mammals. J. Biogeogr. 33:291–303.

Stern, J. T., Susman, R. L., 1983. The locomotor anatomy of *Australopithecus afarensis*. Am. J. Phys. Anthropol. 60:279–317.

Su, D. F., Harrison, T., 2007. The paleoecology of the Upper Laetolil Beds at Laetoli: A reconsideration of the large mammal evidence. In: Bobe, R., Alesmseged, Z., Behrensmeyer, A. K. (Eds.), Hominin environments in the East African Pliocene: an assessment of the faunal evidence. Springer, Dordrecht, pp. 279–313.

Walther, B. A., Moore, J. L., 2005. The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. Ecography, 28, 815–829.

Warton, D. I., Hui, F. K. C., 2011. The arcsine is asinine: the analysis of proportions in ecology. Ecology 92:3–10.

Table 1

List of variables and their abbreviations used in this study

Ecological Guild/Vegetation heterogeneity	
Small, arborophilic primary consumer	BAP
Small, arborophilic secondary consumer	BAS
Small, terrestrial primary consumer	BTP
Small, terrestrial secondary consumer	BTS
Medium arborophilic primary consumer	CAP
Medium, arborophilic secondary consumer	CAS
Medium, terrestrial primary consumer	CTP
Medium, terrestrial secondary consumer	CTS
Large, arborophilic primary consumer	DAP
Large, arborophilic secondary consumer	DAS
Large, terrestrial primary consumer	DTP
Large, terrestrial secondary consumer	DTS
Very large, arborophilic secondary consumer	EAS
Very large, terrestrial primary consumer	ETP
Very large, terrestrial secondary consumer	ETS

Relative proportion of heavy tree cover	%Heavy
Relative proportion of moderate tree cover	%Moderate
Relative proportion of light tree cover	%Light
Relative proportion of absent tree cover	%Absent

Table 2

Summary statistics for Moran’s I calculations for spatial autocorrelation[a]

Distance Class	Count	Distance Center	Moran's I	P	I (max)	I/I(max)
Global dataset						
1	390	535.259	0.060	0.397	0.630	0.095
2	388	1425.89	0.046	0.427	0.510	0.091
3	388	2357.297	-0.129	0.070	0.452	-0.286
4	390	4319.302	-0.111	0.055	0.716	-0.155
5	388	6917.893	0.024	0.673	0.505	0.047
6	388	8449.773	-0.006	0.930	0.571	-0.011
7	390	9396.762	-0.063	0.307	0.709	-0.089
8	388	10532.784	0.028	0.668	0.735	0.038
9	388	13977.548	-0.041	0.437	0.391	-0.105
10	390	18373.225	0.048	0.256	0.264	0.184
Africa only						
1	72	541.999	0.209	0.211	0.892	0.235
2	70	1295.437	-0.273	0.101	0.858	-0.318
3	70	1946.201	-0.263	0.111	0.981	-0.268
4	70	2621.112	-0.139	0.307	1.034	-0.134
5	70	3312.529	0.180	0.176	1.069	0.169
6	70	4169.426	-0.151	0.286	0.549	-0.276
7	72	5563.796	0.062	0.538	0.607	0.102

[a] The global data and the nested African only data were computed.

Table 3

Regression coefficients of multivariate multiple regression of standardised (logit) faunal variables and vegetation variables

Regression coefficients and statistics						
		Coeff.	Std.err.	t	P	R^2
LogitHeavy	Constant	-0.55258	8.3288	-0.06635	0.94738	
	logitBAP	0.93196	0.47929	1.9445	0.057839	0.14061
	logitBAS	3.2348	1.2945	2.4989	0.016012	0.65407
	logitBTP	-0.94808	0.56158	-1.6882	0.097994	0.011624
	logitBTS	-1.5924	0.96289	-1.6537	0.10485	0.13226
	logitCAP	1.0308	0.94427	1.0916	0.28057	0.04478
	logitCAS	1.3614	0.81262	1.6753	0.10052	0.1099
	logitCTP	-0.75176	0.78724	-0.95493	0.3445	0.10554
	logitCTS	2.0864	0.70581	2.9561	0.004859	0.13468
	logitDAP	0.81934	1.8036	0.45427	0.65172	0.063889
	logitDAS	-0.23436	1.0856	-0.21587	0.83002	0.014071
	logitDTP	-1.2451	0.69987	-1.779	0.081704	0.36364
	logitDTS	-1.0248	0.6878	-1.4899	0.14293	0.31442

	logitEAS	-0.95996	0.82673	-1.1611	0.25145	0.057211
	logitETP	0.44917	0.73015	0.61518	0.54141	0.16899
	logitETS	-1.2768	1.0878	-1.1737	0.24641	0.16462
LogitMod	Constant	3.8438	8.4212	0.45644	0.65017	
	logitBAP	0.035608	0.48461	0.073478	0.94174	0.000408
	logitBAS	1.8901	1.3089	1.4441	0.15534	0.029624
	logitBTP	-0.40357	0.56781	-0.71075	0.48075	0.010281
	logitBTS	0.75669	0.97358	0.77723	0.44092	0.040029
	logitCAP	-1.8829	0.95474	-1.9722	0.054488	0.050097
	logitCAS	-1.5934	0.82164	-1.9393	0.058484	0.14294
	logitCTP	1.7213	0.79597	2.1626	0.035696	0.027819
	logitCTS	-0.76862	0.71364	-1.077	0.28696	0.023061
	logitDAP	0.93102	1.8236	0.51053	0.61207	0.013294
	logitDAS	-0.14285	1.0977	-0.13014	0.89701	0.003773
	logitDTP	0.84897	0.70764	1.1997	0.23626	0.072353
	logitDTS	1.1471	0.69543	1.6495	0.10571	0.05215
	logitEAS	0.88709	0.8359	1.0612	0.29401	0.00805

	logitETP	0.33498	0.73825	0.45375	0.6521	0.004082
	logitETS	1.3908	1.0999	1.2645	0.21229	0.07658
LogitLight	Constant	7.3979	11.355	0.65151	0.51789	
	logitBAP	0.20778	0.65343	0.31798	0.75191	0.018891
	logitBAS	0.063677	1.7648	0.036081	0.97137	0.2378
	logitBTP	0.00206	0.76563	0.002691	0.99786	0.015798
	logitBTS	0.28024	1.3127	0.21348	0.83188	0.017661
	logitCAP	-1.3122	1.2874	-1.0193	0.31327	0.012495
	logitCAS	0.050742	1.1079	0.045801	0.96366	0.062604
	logitCTP	0.76331	1.0733	0.7112	0.48047	0.006077
	logitCTS	-0.0876	0.96225	-0.09104	0.92785	0.076443
	logitDAP	2.9901	2.4589	1.216	0.23006	0.000409
	logitDAS	-0.77353	1.4801	-0.52263	0.60369	0.041657
	logitDTP	1.8364	0.95416	1.9246	0.060349	0.27996
	logitDTS	0.99485	0.93771	1.0609	0.29414	0.22906
	logitEAS	0.33272	1.1271	0.29519	0.76915	0.00433
	logitETP	-0.02622	0.99544	-0.02634	0.9791	0.098602

	logitETS	0.91223	1.4831	0.6151	0.54146	0.1193
LogitAbs	Constant	-11.514	11.9	-0.96756	0.33822	
	logitBAP	-1.0748	0.68478	-1.5696	0.12322	0.10141
	logitBAS	-4.611	1.8495	-2.4931	0.016242	0.25917
	logitBTP	0.55652	0.80235	0.69361	0.49134	0.023933
	logitBTS	0.65846	1.3757	0.47863	0.63442	0.039718
	logitCAP	1.3494	1.3491	1.0002	0.32232	0.003612
	logitCAS	-0.60173	1.161	-0.51828	0.6067	0.001202
	logitCTP	-1.1234	1.1247	-0.99877	0.32302	0.053151
	logitCTS	-1.8	1.0084	-1.785	0.080722	0.02298
	logitDAP	-3.6934	2.5769	-1.4333	0.1584	0.04742
	logitDAS	1.4498	1.5511	0.93471	0.35472	0.004148
	logitDTP	-0.7657	0.99993	-0.76575	0.44765	0.04797
	logitDTS	-0.24065	0.98269	-0.24489	0.80761	0.054228
	logitEAS	0.37836	1.1812	0.32033	0.75014	0.087651
	logitETP	-0.96637	1.0432	-0.92636	0.35899	0.05096
	logitETS	-0.29983	1.5542	-0.19291	0.84786	0.010642

Table 4

Average proportions for the functional groups (***E_f***) for all natural protected areas, and the first principal component values (eigenvalue, % variance explained, and loadings for each category, reading down) from an analysis of all faunal variables

	Average E_f	**PC1**
Eigenvalue		172.782
% variance explained		54.823
BAP	4.856351	0.13
BAS	25.64021	0.7495
BTP	5.732363	0.02811
BTS	9.228386	-0.1064
CAP	0.4186	0.0151
CAS	11.77693	0.2658
CTP	5.244462	-0.07985
CTS	9.384767	-0.2753
DAP	0.103451	0.01071
DAS	4.877133	0.02243
DTP	6.295893	-0.3584
DTS	2.167252	-0.1506

EAS	1.812107	-0.01467
ETP	11.01913	-0.3135
ETS	1.442962	0.07712

Table 5

Average proportions for the standardised vegetation categories (***E_v*)** for all natural protected areas, and the first principal component values (eigenvalue, % variance explained, and loadings for each category, reading down) from an analysis of all vegetation variables

	Average *E_v*	**PC1**
Eigenvalue		1964.39
% variance explained		66.394
%Heavy	58.30707	0.8193
%Moderate	4.954032	-0.04017
%Light	11.89459	-0.2809
%Absent	24.84431	-0.4982

Table 6

Performance indicators of the models examined[a]

	Heavy	Moderate	Light	Absent	Total
BIAS					
	Mean Error				
MMR	-19.2212	3.725819	7.630655	7.745052	-0.02992
LRM	1.04E-05	-5.1E-07	-3.6E-06	-6.3E-06	-2.4E-13
Random	9.267959	-42.5414	-43.7935	-27.9056	-26.2431
	% overestimates				
MMR	95.2381	65.07937	63.49206	38.09524	53.57143
LRM	57.14286	80.95238	71.42857	58.73016	62.69841
Random	38.09524	92.06349	87.30159	73.01587	72.61905
PRECISION					
	Coefficient of Variation				
Real	62.95141	226.6715	208.1729	118.8735	
MMR	45.38717	160.5176	155.1958	149.2625	
PCA	52.72288	30.42348	88.60734	75.24661	
Random	59.32787	56.70896	54.13668	59.81456	

	Standard deviation				
Real	36.70512	11.22938	24.76131	29.53329	
MMR	35.18788	1.971498	6.617446	25.52276	
LRM	30.74116	1.507189	10.53948	18.6945	
Random	29.09386	26.93418	30.14769	31.55212	
	Inter-quartile range				
Real	70.69381	0	7.477261	28.47749	
MMR	40.14887	1.115021	5.525482	17.26176	
LRM	66.00463	3.236101	22.62941	40.13913	
Random	56.04193	50.27096	56.28434	60.83928	
ACCURACY					
	Root mean square error				
MMR	25.87533	10.85388	22.29682	24.24292	21.6345
LRM	18.46571	10.90301	20.59138	25.34527	19.5336
Random	43.63162	52.56059	59.0943	55.26061	52.94385
	Mean absolute error				
MMR	19.74289	4.700325	10.8866	16.07606	12.85147

LRM	15.67282	7.531053	13.68161	19.51229	14.09944
Random	37.08307	45.94508	51.48273	46.77917	45.32251

[a] MMR: multivariate multiple regression; LRM: linear regression of principal components analysis.

Table 7

Species list of Upper Laetolil Beds used in this study (from Harrison, 2011), and assignment to functional groups[a]

Order	**Family**	**Taxon**	**Weight**	**Locomotor**	**Trophic**	**Functional Group**
Artiodactyla	Bovidae	"*Gazella*" *kohllarseni*	C	T	P	CTP
Artiodactyla	Bovidae	?*Raphicerus*	C	T	P	CTP
Artiodactyla	Bovidae	*Aepyceros dietrichi*	D	T	P	DTP
Artiodactyla	Bovidae	Alcelaphini large	D	T	P	DTP
Artiodactyla	Bovidae	*Brabovus nanincisus*	E	T	P	ETP
Artiodactyla	Bovidae	Cephalophini sp.	C	T	P	CTP
Artiodactyla	Bovidae	*Gazella janenschi*	C	T	P	CTP
Artiodactyla	Bovidae	*Hippotragus* sp.	E	T	P	ETP
Artiodactyla	Bovidae	*Madoqua avifluminin*	B	T	P	BTP
Artiodactyla	Bovidae	*Oryx deturi*	D	T	P	DTP
Artiodactyla	Bovidae	*Parmularius pandatus*	D	T	P	DTP
Artiodactyla	Bovidae	Reduncini sp. indet.	D	T	P	DTP
Artiodactyla	Bovidae	*Simatherium kohllarseni*	E	T	P	ETP
Artiodactyla	Bovidae	*Tragelaphus* sp.	D	T	P	DTP
Artiodactyla	Giraffidae	aff. *Giraffa jumae*	E	T	P	ETP
Artiodactyla	Giraffidae	*Giraffa stillei*	E	T	P	ETP
Artiodactyla	Giraffidae	*Sivatherium maurusium*	E	T	P	ETP
Artiodactyla	Suidae	*Kolpochoerus heseloni*	D	T	P	DTP
Artiodactyla	Suidae	*Notochoerus euilus*	D	T	P	DTP
Artiodactyla	Suidae	*Notochoerus jaegeri*	D	T	P	DTP
Artiodactyla	Suidae	*Nyanzachoerus kanamensis*	E	T	P	ETP
Artiodactyla	Suidae	*Potamochoerus afarensis*	D	T	S	DTS
Carnivora	Canidae	?*Nyctereutes barryi*	B	T	S	BTS
Carnivora	Canidae	aff. *Otocyon* sp.	B	T	S	BTS
Carnivora	Canidae	cf. *Canis* sp. A	C	T	S	CTS
Carnivora	Canidae	cf. *Canis* sp. B	C	T	S	CTS
Carnivora	Felidae	*Acinonyx* sp.	D	A	S	DAS
Carnivora	Felidae	*Caracal* sp.	C	A	S	CAS
Carnivora	Felidae	*Dinofelis petteri*	D	A	S	DAS
Carnivora	Felidae	*Homotherium* sp.	D	A	S	DAS
Carnivora	Felidae	*Panthera* sp. aff. *P. leo*	E	A	S	EAS
Carnivora	Felidae	*Panthera* sp. cf. *P. pardus*	D	A	S	DAS
Carnivora	Herpestidae	*Herpestes ichneumon*	B	A	S	BAS
Carnivora	Herpestidae	*Herpestes palaeoserengetensis*	B	A	S	BAS
Carnivora	Herpestidae	*Mungos dietrichii*	B	T	S	BTS

Carnivora	Herpestidae	*Mungos* sp. nov	B	T	S	BTS
Carnivora	Hyaenidae	?*Pachycrocuta* sp.	D	T	S	DTS
Carnivora	Hyaenidae	*Crocuta dietrichi*	D	T	S	DTS
Carnivora	Hyaenidae	*Ikelohyaena* cf. *L. abronia*	D	T	S	DTS
Carnivora	Hyaenidae	*Lycyaenops* cf. *L. silberbergi*	D	T	S	DTS
Carnivora	Hyaenidae	*Parahyaena howelli*	D	T	S	DTS
Carnivora	Mustelidae	*Mellivora* sp.	B	A	S	BAS
Carnivora	Mustelidae	*Prepoecilogale bolti*	B	A	S	BAS
Carnivora	Viverridae	*Genetta* sp.	B	A	S	BAS
Carnivora	Viverridae	*Viverra leakyi*	B	A	S	BAS
Lagomorpha	Leporidae	*Serengetilagus praecapensis*	B	T	P	BTP
Perisodactyla	Chalicotheridae	*Ancylotherium hennigi*	E	T	P	ETP
Perisodactyla	Equidae	*Eurygnathohippus* aff. *Hasumense*	D	T	P	DTP
Perisodactyla	Rhinocerotidae	*Ceratotherium efficax*	E	T	P	ETP
Perisodactyla	Rhinocerotidae	*Diceros* sp.	E	T	P	ETP
Primates	Cercopithecidae	Cercopithecoides sp.	C	A	S	CAS
Primates	Cercopithecidae	cf. *Rhinocolobus* sp.	C	A	S	CAS
Primates	Cercopithecidae	*Parapapio ado*	C	A	P	CAP
Primates	Hominindae	*Australopithecus afarensis*	D	A	S	DAS
Proboscidea	Deinotheriidae	*Deinotherium bozasi*	E	T	P	ETP
Proboscidea	Elephantidae	*Loxodonta exoptata*	E	T	P	ETP
Proboscidea	Gomphotheriidae	*Anancus ultimus*	E	T	P	ETP
Proboscidea	Stegodontidae	*Stegodon* sp. cf. *Stegodon kaisensis*	E	T	P	ETP
Rodentia	Hystricidae	*Hystrix leakeyi*	B	A	P	BAP
Rodentia	Hystricidae	*Hystrix makapanensis*	B	A	P	BAP
Rodentia	Hystricidae	*Xenohystrix crassidens*	C	T	P	CTP
Rodentia	Pedetidae	*Pedetes laetoliensis*	B	T	S	BTS
Tubulidentata	Orycteropodidae	*Orycteropus* sp.	D	T	S	DTS

[a] See text for details.

Table 8

Proportional percentage of the fauna from Upper Laeotolil Beds falling in the described functional groups

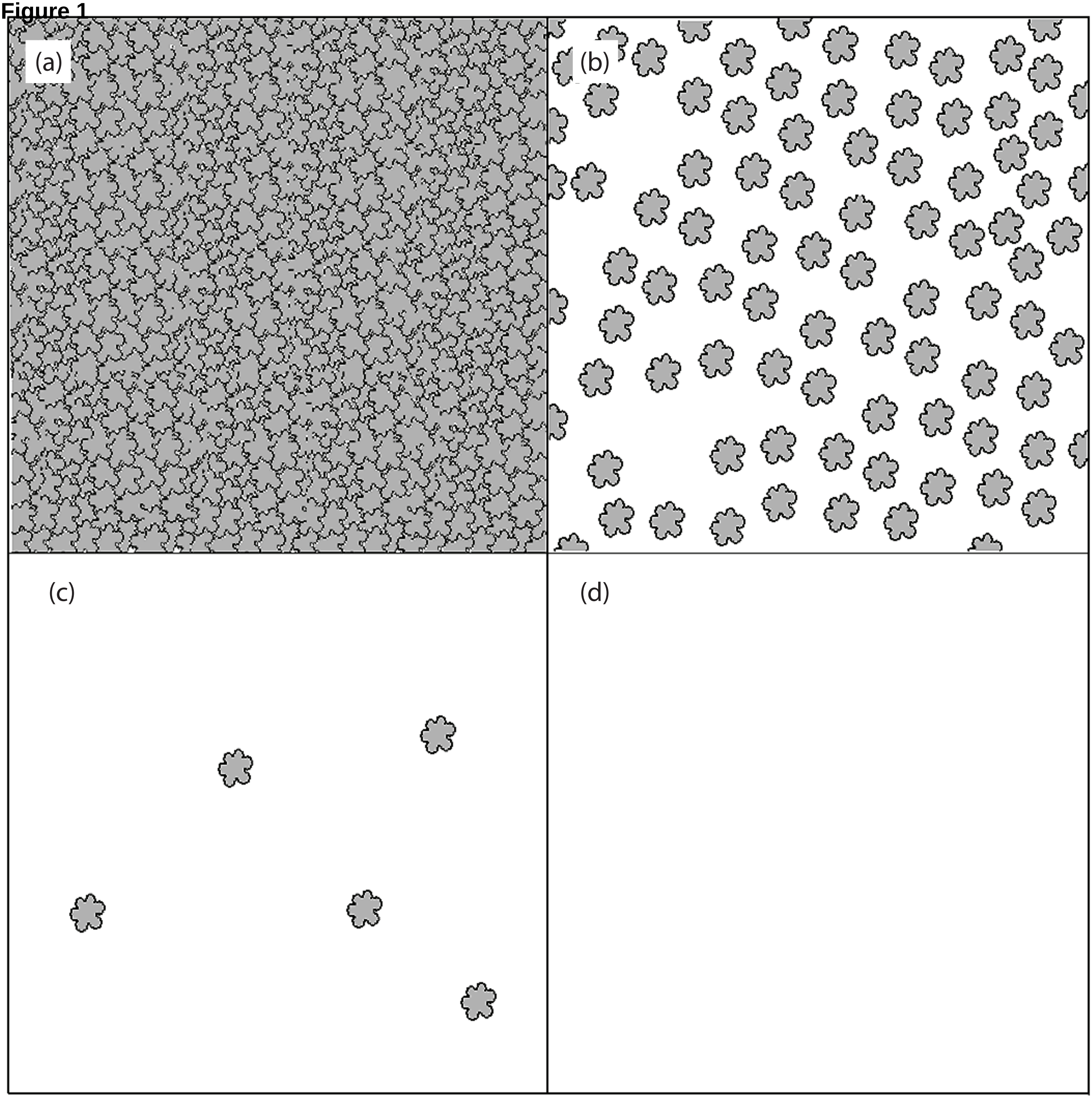
Functional Group	Number taxa	Proportion %
BAP	2	3.174603
BAS	6	9.52381
BTP	2	3.174603
BTS	5	7.936508
CAP	1	1.587302
CAS	3	4.761905
CTP	5	7.936508
CTS	2	3.174603
DAP	0	0
DAS	5	7.936508
DTP	10	15.87302
DTS	7	11.11111
EAS	1	1.587302
ETP	14	22.22222
ETS	0	0

Figure 1. Stylistic representation of arboreal heterogeneity categories as used in the analyses. (a) Heavy tree cover [%Heavy]; (b) moderate tree cover [%Moderate]; (c) light tree cover [%Light]; (d) no tree cover [%Absent]. An area of 625 km^2 for each national protected area centred on their geographical coordinates was classified according to these categories, and the proportions of each category within those areas calculated (from Louys et al., 2011).

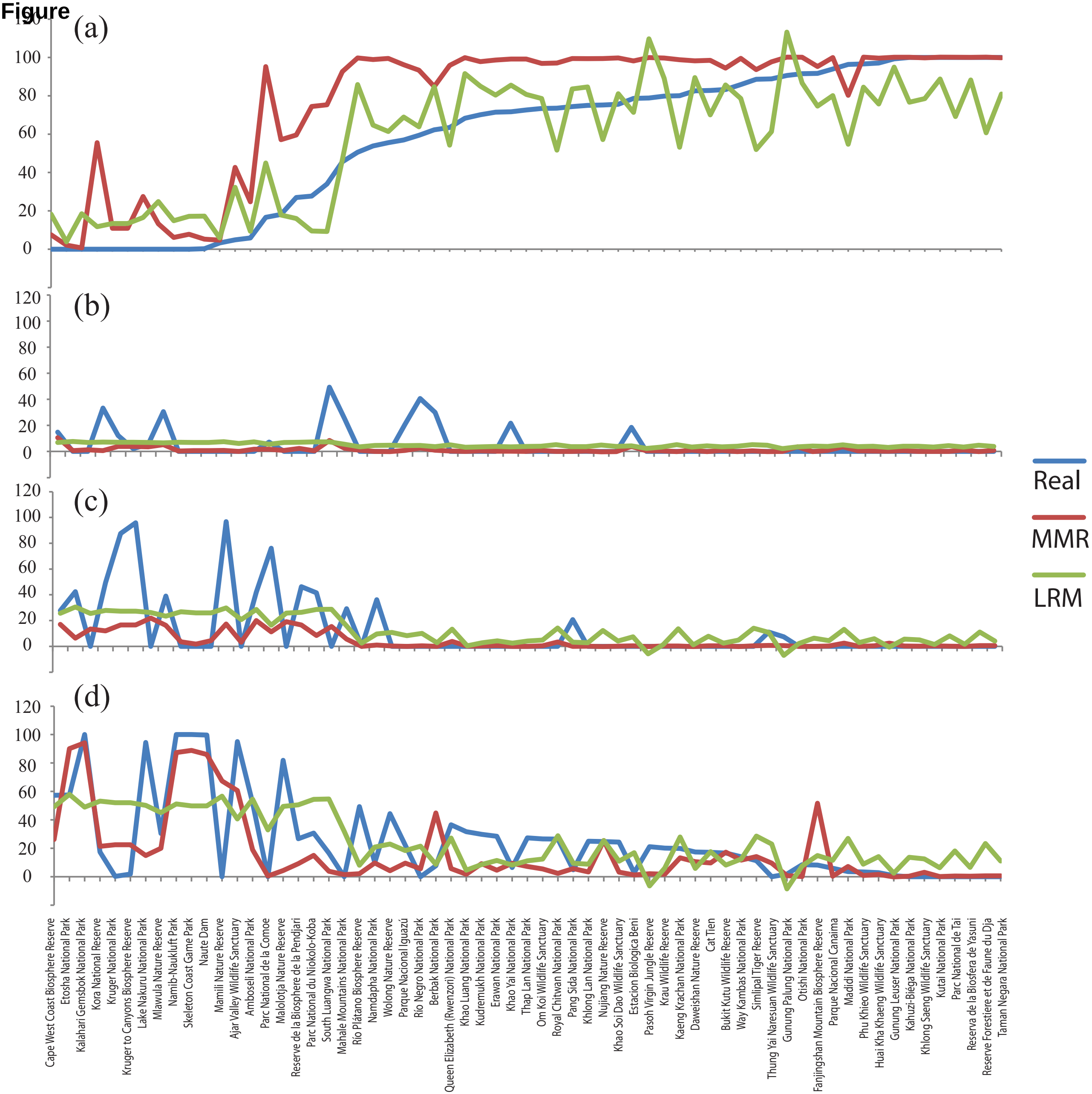
Figure 2. Comparisons between real and modelled values for each vegetation category. MMR represents the multiple multivariate regression of standardised values, LRM the linear regression of the principal components of the ecospaces. (a) Heavy tree cover [%Heavy]; (b) moderate tree cover [%Moderate]; (c) light tree cover [%Light]; (d) no tree cover [%Absent]. National protected areas are listed in order of increasing proportion of heavy tree cover over the 625 km^2 sampled.

Figure 3. Box plot of real, modelled, and random vegetation proportions. MMR represents the multiple multivariate regression of standardised values, LRM the linear regression of the principal components of the ecospaces. (a) Heavy tree cover [%Heavy]; (b) moderate tree cover [%Moderate]; (c) light tree cover [%Light]; (d) no tree cover [%Absent]. For each vegetation category, the 25–75% quartiles are drawn using the box, the median is shown with a horizontal bar inside said box, and maximum and minimum values are shown by the accompanying whiskers.

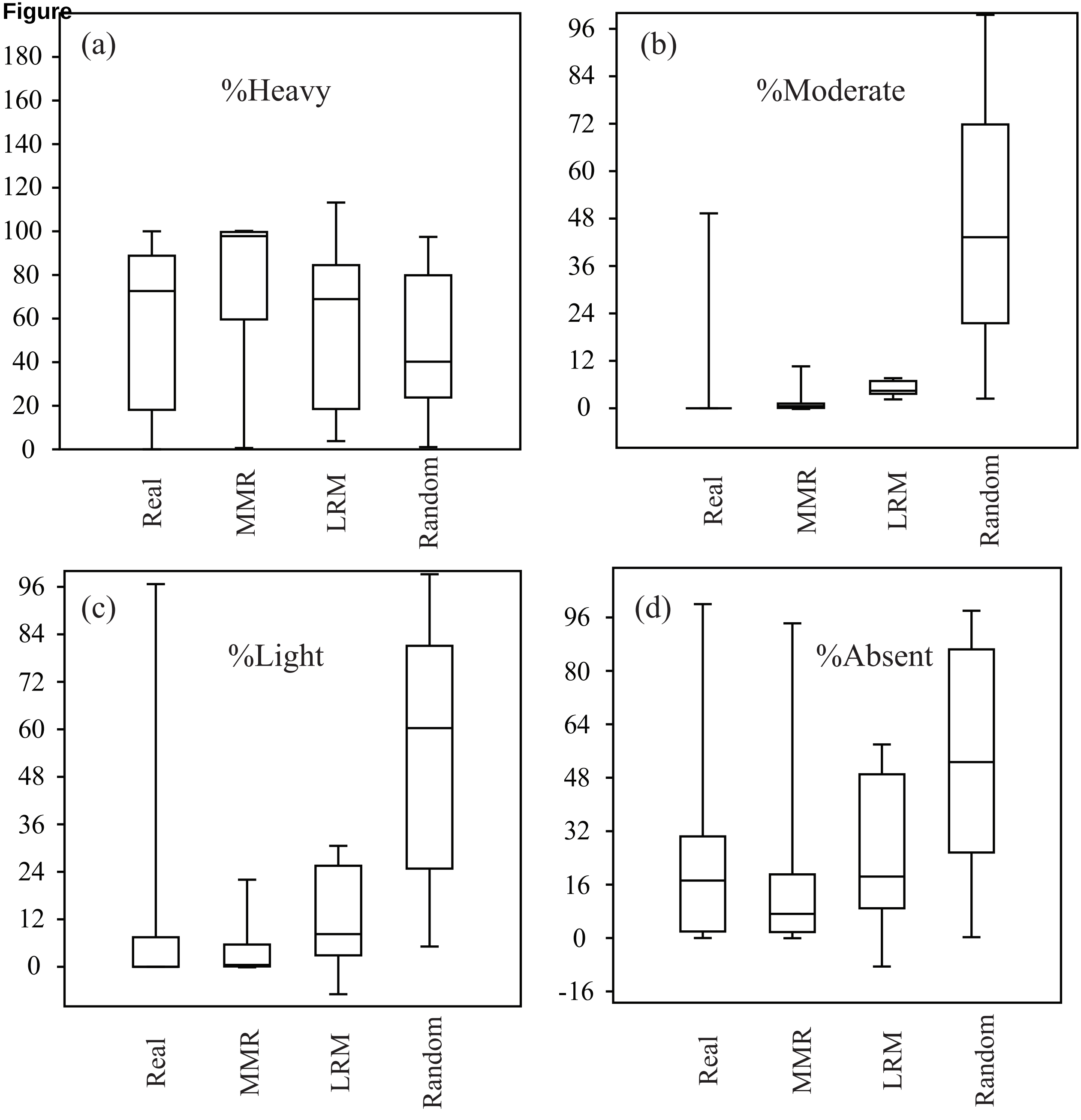
Figure 1



Figure


(a)
120
100
80
60
40
20
0
(b)
120
100
80
60
40
20
0
(c)
120
100
80
60
40
20
0
(d)
120
100
80
60
40
20
0
Real
MMR
LRM
Cape West Coast Biosphere Reserve
Etosha National Park
Kalahari Gemsbok National Park
Kora National Reserve
Kruger National Park
Kruger to Canyons Biosphere Reserve
Lake Nakuru National Park
Mlawula Nature Reserve
Namib-Naukluft Park
Skeleton Coast Game Park
Naute Dam
Mamili Nature Reserve
Ajar Valley Wildlife Sanctuary
Amboseli National Park
Parc National de la Comoe
Malolotja Nature Reserve
Reserve de la Biosphere de la Pendjari
Parc National du Niokolo-Koba
South Luangwa National Park
Mahale Mountains National Park
Río Plátano Biosphere Reserve
Namdapha National Park
Wolong Nature Reserve
Parque Nacional Iguazú
Río Negro National Park
Berbak National Park
Queen Elizabeth (Rwenzori) National Park
Khao Luang National Park
Kudremukh National Park
Erawan National Park
Khao Yai National Park
Thap Lan National Park
Om Koi Wildlife Sanctuary
Royal Chitwan National Park
Pang Sida National Park
Khlong Lan National Park
Nujiang Nature Reserve
Khao Soi Dao Wildlife Sanctuary
Estacion Biologica Beni
Pasoh Virgin Jungle Reserve
Krau Wildlife Reserve
Kaeng Krachan National Park
Daweishan Nature Reserve
Cat Tien
Bukit Kutu Wildlife Reserve
Way Kambas National Park
Similipal Tiger Reserve
Thung Yai Naresuan Wildlife Sanctuary
Gunung Palung National Park
Otishi National Park
Fanjingshan Mountain Biosphere Reserve
Parque Nacional Canaima
Madidi National Park
Phu Khieo Wildlife Sanctuary
Huai Kha Khaeng Wildlife Sanctuary
Gunung Leuser National Park
Kahuzi-Biéga National Park
Khlong Saeng Wildlife Sanctuary
Kutai National Park
Parc National de Tai
Reserva de la Biosfera de Yasuni
Reserve Forestiere et de Faune du Dja
Taman Negara National Park

Figure


(a)
%Heavy
0
20
40
60
80
100
120
140
160
180
Real
MMR
LRM
Random
(b)
%Moderate
0
12
24
36
48
60
72
84
96
Real
MMR
LRM
Random
(c)
%Light
0
12
24
36
48
60
72
84
96
Real
MMR
LRM
Random
(d)
%Absent
-16
0
16
32
48
64
80
96
Real
MMR
LRM
Random

Supplementary Material