

Precipitation and large herbivorous mammals I: estimates from present-day communities

J.T. Eronen^{1,2}, K. Puolamäki³, L. Liu^{1,4}, K. Lintulaakso¹, J. Damuth⁵,
C. Janis⁶ and M. Fortelius^{1,7}

¹Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland,

²Department of Computer Science, University of Helsinki, Helsinki, Finland,

³Department of Media Technology, Helsinki University of Technology, Helsinki, Finland,

⁴Laboratory of Evolutionary Systematics of Vertebrates, Chinese Academy of Science, Beijing, China,

⁵Department of Ecology, Evolution and Marine Biology, University of California, Santa Barbara,

California, USA, ⁶Department of Ecology and Evolutionary Biology, Brown University,

Providence, Rhode Island, USA and ⁷Institute of Biotechnology,

University of Helsinki, Helsinki, Finland

ABSTRACT

Question: How can mammalian community characteristics be used to estimate regional precipitation?

Data: Global distribution data of large mammals and their ecomorphology; global climate data.

Research methods: Non-linear regression-tree analysis and linear regression.

Conclusions: The methods unravelled the complex relationships between the environment and the characteristics of mammalian communities. The regression trees described here provide a reasonably accurate estimate of precipitation values for today's world. The strongest correlations are for annual precipitation versus diet ($R^2 = 0.665$), precipitation versus tooth crown height ($R^2 = 0.658$), and precipitation versus diet and tooth crown height combined ($R^2 = 0.742$)

Keywords: climate, community structure, decision trees, herbivorous mammals, hysodonty, precipitation.

INTRODUCTION

There is a growing need to provide numerical estimates of environmental variables for the present and the past. This is especially true for climate and ecosystem models, which need such estimates as input for boundary conditions. As climatic modelling has moved deeper into the historical past and our data coverage of older time periods has improved, the need

Correspondence: J.T. Eronen, Department of Geosciences and Geography, University of Helsinki, PO Box 64, FI-00014 Helsinki, Finland. e-mail: jussi.t.eronen@helsinki.fi

Consult the copyright statement on the inside front cover for non-commercial copying policies.

to estimate environmental conditions for pre-Quaternary times has become more urgent. Furthermore, there is a growing need to estimate biotic and abiotic variables for past times to validate the growing number of results from paleoclimate models.

Although there are many different methodologies for reconstructing past environmental conditions based on plant data [e.g. Leaf Margin Analysis (Wolfe, 1979); Climate Leaf Analysis Multivariate Program (Wolfe, 1993); Coexistence Approach (Mosbrugger and Utescher, 1997); Climate Amplitude Method (Fauquette *et al.*, 1998)], data from large fossil mammals, which are both spatially and temporally well-sampled, have not been extensively used for quantitative estimates of past environmental conditions. Large mammals are an important component of the biosphere, and are increasingly threatened by human activities. There is a rich literature on the relationships between large mammals (especially herbivores) and a multitude of biotic and abiotic variables (see below).

It is known that plant productivity and quality are influenced by precipitation and nutrient availability (Reich *et al.*, 1997; Walker and Langridge, 1997). Not surprisingly, herbivore community structure and distributions are in turn affected by plant quality and productivity (e.g. Du Toit *et al.*, 1989; Belovsky, 1997; Owen-Smith, 2002). Previous studies have linked rainfall and animal biomass (Rosenzweig, 1968; Coe *et al.*, 1976), primary productivity and herbivore community biomass (McNaughton *et al.*, 1989), rainfall and species richness of herbivores (Rosenzweig, 1995; Danell *et al.*, 1996), and environmental controls of large herbivore biomass (Olf *et al.*, 2002). These relationships are usually made through the food-source of mammals, and therefore herbivores are the trophic group of choice in such studies.

A number of researchers have shown a relationship between aspects of large mammal herbivore communities and local environmental conditions (e.g. Bell, 1971; Jarman, 1974; Gordon and Illius, 1989; Caughley and Gunn, 1993; Clutton-Brock *et al.*, 1997; Mysterud *et al.*, 2001; Owen-Smith, 2002). More specifically, we have been able to show a relationship between the average value of tooth crown height (hypsodonty) in a herbivore community and the precipitation level of the community's local habitat (Damuth and Fortelius, 2001; J. Damuth, unpublished data). Damuth and Fortelius (2001) introduced and Fortelius *et al.* (2002) applied the use of mean hypsodonty to estimate past precipitation values from fossil molar teeth of large herbivorous mammals. This method has also been applied by Fortelius *et al.* (2003, 2006), Eronen and Rook (2004), and Eronen (2006). Hypsodonty-based estimates have also been used in several studies, including Jernvall and Fortelius (2004) and Eronen *et al.* (2009). Damuth *et al.* (2002) introduced and Janis *et al.* (2004) used a technique called 'per species mean hypsodonty', where mean hypsodonty was related to species diversity (the number of species). The method we introduce here also includes species diversity, although the implementation is different. Here we present a more detailed study on how these aspects of mammalian communities can be used to estimate precipitation.

Tooth crown height (specifically of the cheek teeth used for mastication) is a measure of dental durability. Mammals have only a single adult dentition (i.e. only one set of replacement teeth). Thus, if a herbivore consumes abrasive vegetation, the teeth will be liable to be worn down before the end of the natural lifespan, with repercussions on reproductive output, and hence fitness. To extend the useful life of teeth, the dental materials themselves cannot be made much more durable, so the height of the tooth crown must increase (making the tooth hypsodont). Hypsodonty has evolved multiple times among herbivorous mammals (Janis and Fortelius, 1988). Brachyodont, or low crowned, teeth are seen in herbivores that eat relatively non-abrasive food such as soft browse (leaves of dicotyledonous plants) in a relatively grit-free environment. A greater degree of hypsodonty

indicates a diet that is more abrasive. This usually means a diet containing greater amounts of grass, but other plants in more arid areas may also contain abrasive dust and grit with similar effects (see Janis and Fortelius, 1988).

Following Janis and Fortelius (1988) and Fortelius *et al.* (2002, 2003, 2006), we argue that hypsodonty is fundamentally an adaptive response to increasing demands for wear tolerance and functional durability of the dentition. Thus, hypsodonty is related to habitats that are open and contain arid-adapted vegetation, whose available plant material is more fibrous and abrasive (Van Valen, 1960; Fortelius, 1985; Janis and Fortelius, 1988; Solounias *et al.*, 1994; Fortelius and Solounias, 2000; C. Janis, unpublished data). In short, we can expect details reflecting regional ecology to be recorded in the herbivore dental morphology (Fortelius and Hokkanen, 2001; Fortelius *et al.*, 2002; Jernvall and Fortelius, 2002). In fact, we contend that communities containing primarily hypsodont herbivores imply a local condition of the consumable vegetation that might be termed 'generalized water stress'. Herein, we will quantify this relationship.

METHODS AND MATERIALS

Precipitation variables were obtained from Hijmans *et al.* (2005; available online at <http://www.worldclim.org>). The data consist of global climate layers in the 10-min resolution ($18.6 \times 18.6 = 344 \text{ km}^2$ at the equator) version of the data set. Climate values were converted to a gridded format with a resolution of 0.5° latitude and longitude for grid cells ($\sim 55 \text{ km}$ at the equator, 720×360 cells) to match the mammal data. For climate analysis, we used the following bioclimatic (BIOCLIM) variables that are contained in the WorldClim data set: (1) total annual precipitation; (2) precipitation of the wettest month; (3) precipitation of the driest month; (4) precipitation seasonality, coefficient of variation; (5) precipitation of the wettest quarter; (6) precipitation of the driest quarter; (7) precipitation of the warmest quarter; (8) precipitation of the coldest quarter. Precipitation data are in millimetres.

We selected the orders Artiodactyla, Perissodactyla, and Primates for our investigation. These orders encompass the majority of large herbivorous land mammal species. We excluded the two elephant species (order Proboscidea) from our analysis, but this had a negligible effect on the results (see Discussion below). We also excluded carnivorous omnivores and carnivores because their relationship to climate is more complex than that of herbivores.

The geographic ranges of the mammalian species were derived from the World Wildlife Fund's species distributions (WWF WildFinder; <http://www.worldwildlife.org/science/>). The ranges are based on occurrences lists for the ecoregion divisions of Olson *et al.* (2001). Ecoregions are contiguous regions across which environmental conditions are similar, as are the fauna and flora. The ranges from this data set are estimated by identifying the ecoregions where the species is present and then making the assumption that the species distribution is ecoregion-wide (i.e. present throughout that entire ecoregion). The ranges used here are thus likely overestimated, covering a larger geographic region than is actually inhabited by the species; however, this is not an issue for this study because the range of climatic conditions associated with the species is not affected by these extensions (since the additional areas belong to the same ecoregions, which have by definition the same climate), and because the ranges estimated here are only slightly different from ranges reported in the literature. [For further details, see Olson *et al.* (2001).]

We converted the WildFinder data to a grid format with a resolution of 0.5° latitude and longitude for grid cells ($\sim 55 \text{ km}$ at the equator, 720×360 cells) and recorded all the species

present within each cell. When more than one ecoregion was included in the area of a cell, we chose the ecoregion with the highest proportion of area within the cell. We excluded the Indo-Pacific area because we lack data for most species there. We also excluded Australasia because the herbivores there are marsupials, for which we lack good data. After this conversion, we exported all the grid cells associated with each species to a matrix format where one occurrence of a species in a grid cell is represented by one row.

For ecomorphology, we used the following variables for each species: body mass, body length, diet, and tooth crown height. We gathered the data from published literature sources (Janis, 1988, 1995; Janis and Fortelius, 1988; Nowak, 1991; Wilson and Ruff, 1999; Fortelius and Solounias, 2000; Smith *et al.*, 2003; Wilson and Reeder, 2005; Myers *et al.*, 2007; NatureServe, 2007), and complemented these with unpublished data (from M. Fortelius, C. Janis, J.T. Eronen and L. Liu). For the body mass variable, we used the mean recorded body mass (kg), and for body length, the total mean snout-to-vent length (cm). For diet, we used the following categories: herbivore, frugivore, and omnivore. We further split the herbivore group into grazer, mixed feeder, and browser (based on Hoffmann and Stewart, 1972; Hoffmann, 1989). For tooth crown height, we used three different categories: low crowned (brachyodont), medium high crowned (mesodont), and high crowned (hypsodont), as described in Fortelius *et al.* (2002).

Our aim here is to predict a precipitation-related quantity in a grid cell with a set of covariates specific to that cell. We need regression analysis. After preliminary analyses, we decided to concentrate on three precipitation-related climate variables: annual precipitation, precipitation of the wettest quarter, and precipitation of the driest quarter. These are easiest to interpret and use, and they give the strongest correlations with ecomorphological variables. As covariates or predictor variables, we used the average body size, average body mass, a set of variables related to tooth crown height, and a set of variables related to diet. We also report the results for the combination of tooth crown variables and diet variables. The covariates used in this paper are reported in detail in Table 1.

We used both linear regressions and regression trees (Breiman *et al.*, 1984) to predict the annual precipitation (mm). Both are multivariate methods – that is, they can handle several covariates at the same time. We used the implementation of linear regression in GNU R (R Development Core Team, 2009), and the regression tree library *rpart* in GNU R, for all computations (Therneau *et al.*, 2009). The *rpart* library uses cross-validation and other procedures to prune the trees in order to avoid over-fitting the data [see Venables and Ripley (2002) and Breiman *et al.* (1984) for discussion and references about how to construct a regression tree and about the cross-validation procedures]. Note that for linear regressions, one can use stepwise model selection to prune out unnecessary parameters and avoid over-fitting the data (Venables and Ripley, 2002). For example, cladistic analysis uses linear regressions with stepwise model selection. However, unlike linear models, regression trees can express complex non-linear relationships between the covariates.

We used only those grid cells with at least two recorded values of tooth crown height and diet relevant for the analysis. There are 53,054 grid cells and 584,341 taxon occurrences in our data set. The mean number of species per locality is thus 11.014 (= 584,341/53,054). We tested all of the precipitation-related BIOCLIM variables against our mammal ecomorphological variables with all combinations (not reported here).

Linear regression and the regression tree both minimize the squared prediction error between the actual and the predicted values of annual precipitation. We can measure the discrepancy between the actual and predicted values, because we know both the actual value (from the WorldClim data set) and the value predicted by the regression. We use

Table 1. Variables, specific to a grid cell, used as covariates in a regression to estimate the average annual precipitation, precipitation of the wettest quarter, and the precipitation of the driest quarter (all in mm) in the grid cell

Name of set of covariates	Covariates specific to a grid cell
Body size	SVLENGTH (mean animal estimated snout-to-vent length in centimetres, without tusk or tail, in a grid cell)
Body mass	MASS (mean animal body mass in kilograms in a grid cell)
Tooth crown height	NHYP (number of species with hypsodonty data), NHYP1 (number of low crowned species), NHYP2 (number of medium high crowned species), NHYP3 (number of high crowned taxa), pHYP1 (fraction of low crowned species), pHYP2 (fraction of medium high crowned species), pHYP3 (fraction of high crowned taxa)
Diet	NDIET (number of species with diet data), DIETB (number of browsers), DIETF (number of fruit eaters), DIETG (number of grazers), DIETM (number of mixed feeders), DIETO (number of omnivores), pDIETB (fraction of browsers), pDIETF (fraction of fruit eaters), pDIETG (fraction of grazers), pDIETM (fraction of mixed feeders), pDIETO (fraction of omnivores)
Diet + tooth crown height	The ‘tooth crown height’ and ‘diet’ variables listed above, respectively

the R^2 measure to assess the error. The R^2 statistic is commonly interpreted to be the proportion of variance explained by the regression.

To facilitate the use of the regression tree method to resolve environmental proxies, we describe, in the Results section, the step-by-step process of using the tree. Our example is the tree that uses tooth crown height to estimate annual precipitation (Fig. 1). We also provide other regression trees that can be used to estimate annual, wettest, and driest quarter precipitation based on different sets of characteristics (see online Appendix 1: evolutionary-ecology.com/data/2538A1.pdf), together with a description of the variables necessary to use the trees. The regression tree can be visualized as a binary tree (hence the name; for an example, see Fig. 1). To use the tree, traverse it from its root node at the top to a leaf node at the bottom. At each node, choose the left or right branches on the tree based on the covariates: if the condition associated with the node is true, take the left branch, otherwise follow the right branch (see Results for further description). The regression tree has an advantage over a linear regression if the response to the covariates is non-linear, as is often the case in biology. Another advantage of the regression tree is that it can identify context-dependent associations among multiple correlated covariate variables; regression trees do not, for example, assume that the covariates are independent, as is the case with linear regression. Moreover, the same covariate variables can occur several times at different levels of the tree. Finally, a regression tree can be represented in an easily understandable graphical format, allowing ready interpretation of the results. See online Appendix 2 (evolutionary-ecology.com/data/2538A2.pdf) for further discussion on the comparison of linear regression and regression trees, and Legendre and Legendre (1998) for further discussion of linear regression in ecology.

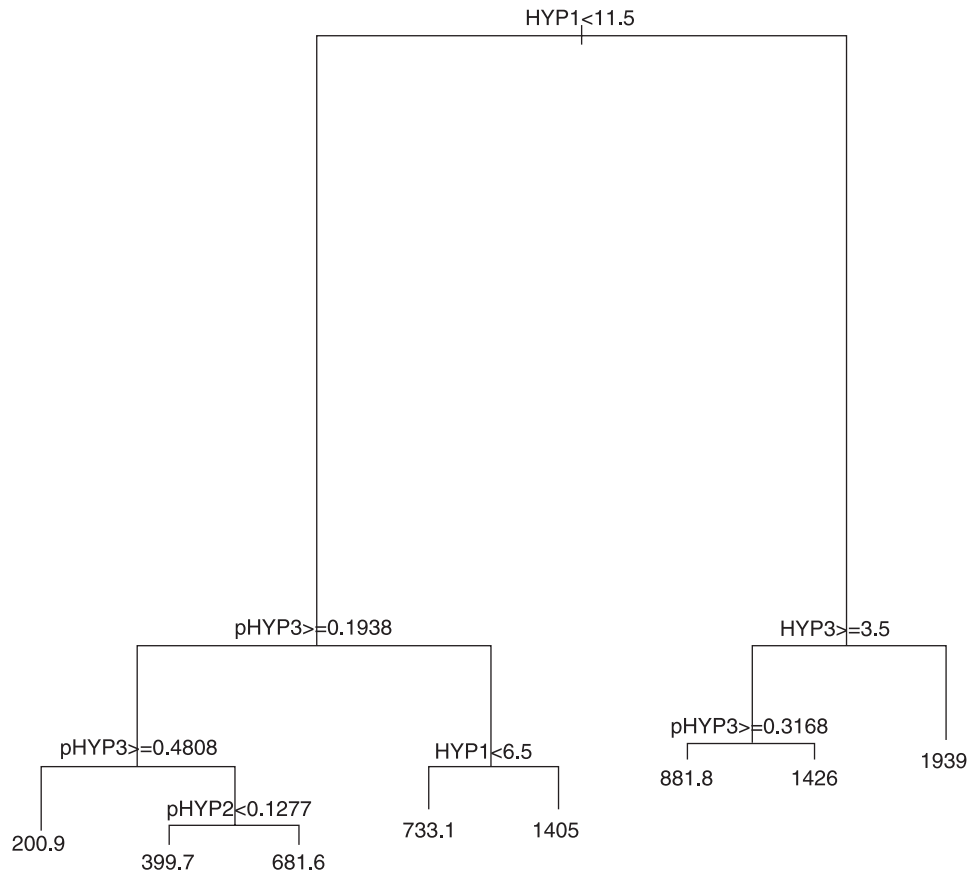
bio12 ~ H

Fig. 1. Decision tree for annual precipitation using hypsodonty alone as regressor (see online Appendix 1 for other decision trees generated: evolutionary-ecology.com/data/2538A1.pdf).

We have ignored spatial autocorrelation and trends in our analysis. [See *Diniz-Filho et al. (2003)* and *Hawkins et al. (2007)* for discussion about the effects of ignoring spatial autocorrelation in regression analysis.] In other words, we have ignored the fact that estimated values (e.g. mean annual precipitation) in nearby grid cells may be correlated. In so doing, we may have overestimated the statistical significance of our results, or over-fitted the regression tree model. Consequently, the resulting regression tree – while more accurate for the present-day mammal and climatic data used to fashion the regression tree structure – may be too detailed. Taking the autocorrelation carefully into account could, in principle, result in smaller regression trees in which some of the branches near the leaf nodes would have been pruned out. However, our regression trees have only 8–12 leaf nodes, meaning that each leaf node describes a large number of grid cells, on average about 5000 cells; hence, we are confident that potential over-fitting is not a serious issue here. Even if there were some over-fitting, the regression tree would still fit present-day data well, as shown by the R^2 values. [For previous use of regression trees in ecology, see, for example, *Roff and*

Roff (2003), Sankaran *et al.* (2005), Jones *et al.* (2006), and Davidson *et al.* (2009).] Most relevant to our study is Mendoza's (2007) use of decision trees to generate dietary groupings of herbivorous mammals according to ecomorphological species traits.

To visualize the results, we used thematic mapping and grid interpolation in MapInfo 8.5 with the following settings: IDW-interpolation, 50 km grid size, 100 km search radius, 100 grid border. The interpolated maps were then imported as grids to Vertical Mapper, where they were assigned to classes and then contoured to connect the areas with similar values. Inflection values vary between different variables. (See the individual figure legends.)

RESULTS

To estimate the global relationship of precipitation with mammalian characteristics, we used both linear regression and regression-tree methods. The best single correlations for annual precipitation using linear regression are for diet [$R^2 = 0.658$ (mean error = 384 mm)] and tooth crown height [$R^2 = 0.581$ (mean error = 388 mm)]. As single regressors, both body mass and body size have a weak correlation with precipitation ($R^2 = 0.109$ and 0.389 , respectively) (Table 2). The best combination for linear correlation is diet and tooth crown height [$R^2 = 0.705$ (mean error = 337 mm)]. When diet is used to estimate precipitation values, the strongest correlations are for annual precipitation ($R^2 = 0.658$) and wettest quarter precipitation ($R^2 = 0.600$) (Table 2). The driest quarter has a weak correlation ($R^2 = 0.350$). Tooth crown height gives a similar range of correlations as diet (annual precipitation = 0.581 , wettest quarter = 0.499 , driest quarter = 0.334) (Table 2). Our regression tree uses both the proportions and absolute numbers of species in a representative taxon to estimate precipitation. The method uses decimals in absolute numbers (such as 6.5), which is just a technicality of the methodology. Of course, in reality we cannot have

Table 2. Results of the linear regression and regression-tree analysis

Variable	Annual precipitation		Driest quarter precipitation		Wettest quarter precipitation	
	Linear	Tree	Linear	Tree	Linear	Tree
Body size	0.389 (518.33)	0.566 (437.10)	0.110 (87.09)	0.442 (68.93)	0.381 (232.75)	0.496 (209.89)
Body mass	0.109 (625.89)	0.444 (494.57)	0.084 (88.37)	0.416 (70.52)	0.071 (285.06)	0.374 (234.07)
Diet	0.658 (388.05)	0.665 (384.11)	0.350 (74.42)	0.546 (62.23)	0.600 (187.05)	0.659 (172.57)
Tooth crown height	0.581 (429.31)	0.658 (387.98)	0.334 (75.31)	0.486 (66.20)	0.499 (209.38)	0.663 (171.69)
Diet + tooth crown height	0.705 (360.02)	0.742 (337.04)	0.446 (68.71)	0.548 (62.03)	0.649 (175.09)	0.700 (162.02)

Note: The values comparing diet alone and hypsodonty alone are in **bold**. The table shows the R^2 values and the respective standard deviations (in parentheses). The regression trees are shown in Fig. 1 and online Appendix 1 (evolutionary-ecology.com/data/2538A1.pdf).

half of a grazer in a community; for example, >6.5 should be understood as ‘more than six’. In our example, we try to estimate the mean annual precipitation of each grid cell based on the characters of the herbivorous mammal community in that grid cell.

The regression tree method returns larger correlation values, suggesting that the mammal characteristics used have a non-linear relationship with precipitation. The correlation values are larger when we use the full set of absolute and relative numbers than when we use only relative numbers. The strongest correlations are for annual precipitation (diet = 0.665, tooth crown height = 0.658, diet and tooth crown height = 0.742) and wettest quarter precipitation (diet = 0.659, tooth crown height = 0.663, diet and tooth crown height = 0.700) (Table 2). The fact that hypsodonty performs almost as well as diet for the annual and wettest quarter precipitation estimation suggests that both variables may capture much of the same environmental signal.

Now, we describe the step-by-step process of using the tree. Our example is the tree that uses tooth crown height to estimate annual precipitation (Fig. 1). The tree starts from the root at the top and splits into two at each node according to the covariates in question. The vertical length of the line between the nodes illustrates the fraction of variance that is explained by the covariate in question. The tree should be read down as far as one of the leaf nodes, thus providing an estimate of the mean annual precipitation based on the particulars of the community in question.

The regression tree for annual precipitation using only tooth crown height information as the estimator yields an accuracy of 0.658 and a mean error of 388 mm (Table 2). The first and most important determinant is the absolute number of brachydont species (Fig. 1). If the number of brachydont species (HYP1) is less than 11.5, we go to the left. If it is more than 11.5, we go to the right. On the second level, the split uses the number and proportion of hypsodont species (pHYP3 and HYP3). Looking at the left branch of the tree, if the proportion of hypsodont species (pHYP3) is larger than 19.38%, the determining factor is again the proportion of hypsodont species (pHYP3). If it is larger than 48.08%, then the precipitation estimate is very low (200.9 mm). If it is small (less than 48.08%), then the proportion of medium high crowned species (pHYP2) determines whether the precipitation is low (399.7 mm) or medium (681.6 mm). Looking at the other branch on the right, we see that if the proportion of hypsodont species (pHYP3) is less than 19.38%, then it is the number of brachydont species (NHYP1) that determines whether the precipitation is medium (733.1 mm) or high (1405 mm). On the right side of the tree, when the number of brachydont species is large, the second level split is determined by the number of hypsodont species (HYP3). If it is large (more than 3), then the proportion of hypsodont species determines whether the precipitation is medium (881.8 mm) or high (1426 mm). If the number of hypsodont species (HYP3) is less than 3, then the precipitation estimate is very high (1939 mm). The height of the branches in the visualization correspond to the variance explained; for example, in the tree in Fig. 1, the choice at the root node (whether HYP1 is above or below 11.5) already explains most of the variation in the data. Relative to this first choice, the other choices, or branches of the tree, account only for fine-tuning of the result.

Consider the following. Cell 10.25 N, 11.75 E, located in West Sudanian Savanna (WWF ecoregion code AT0772, for a description of ecological conditions and fauna, see: http://www.worldwildlife.org/wildworld/profiles/terrestrial/at/at0772_full.html), has the following faunal community structure: number of species (NHYP) = 23; number of brachydont species (HYP1) = 11; number of mesodont species (HYP2) = 3; number of hypsodont species (HYP3) = 9; proportion of brachydont species (pHYP1) = 0.478;

proportion of mesodont species ($pHYP2$) = 0.13; proportion of hypsodont species ($pHYP3$) = 0.39. (Online Appendix 3 has the species list; evolutionary-ecology.com/data/2538A3.pdf.) Using the hypsodonty predictive tree (Fig. 1), we will now estimate the annual precipitation for this grid cell. We start from the root (the upmost node). Because the fauna has fewer than 12 brachydont species ($HYP1 < 11.5$), we continue to the left side in the regression tree. The next node is $pHYP3 > 0.1938$. Because the proportion of hypsodont species ($pHYP3$) in this cell is 0.39, we continue to the left. The next node is $pHYP3 > 0.4808$. As $pHYP3$ in this cell is 0.39, we go right. The next node is $pHYP2 < 0.1277$. In this cell, the proportion of mesodont species ($pHYP2$) is 0.13, so we continue to the right. We arrive at the value of 681.6 mm annual precipitation. The actual annual precipitation (from WorldClim data set) for this grid cell is 743 mm.

To test the geographical resolution of our precipitation estimates, we estimated the mean annual precipitation values for the present-day world using: hypsodonty alone (Fig. 2A); diet alone (Fig. 2B; for regression tree: evolutionary-ecology.com/data/2538A1.pdf); hypsodonty and diet in combination (Fig. 2C and evolutionary-ecology.com/data/2538A1.pdf). For comparison, we also show the actual precipitation map calibrated to the same ranges (Fig. 2D). Plotting the variables on a map makes it easier to understand the environmental context of the splitting sequence in the regression tree. To this end, we also show the number of species with low crowned teeth (Fig. 3A), the proportion of frugivores (Fig. 3B), and the proportion of mixed feeders (Fig. 3C) within the grid cells. These plots show how the variables resolve the spatial pattern, and what regions they resolve best.

DISCUSSION

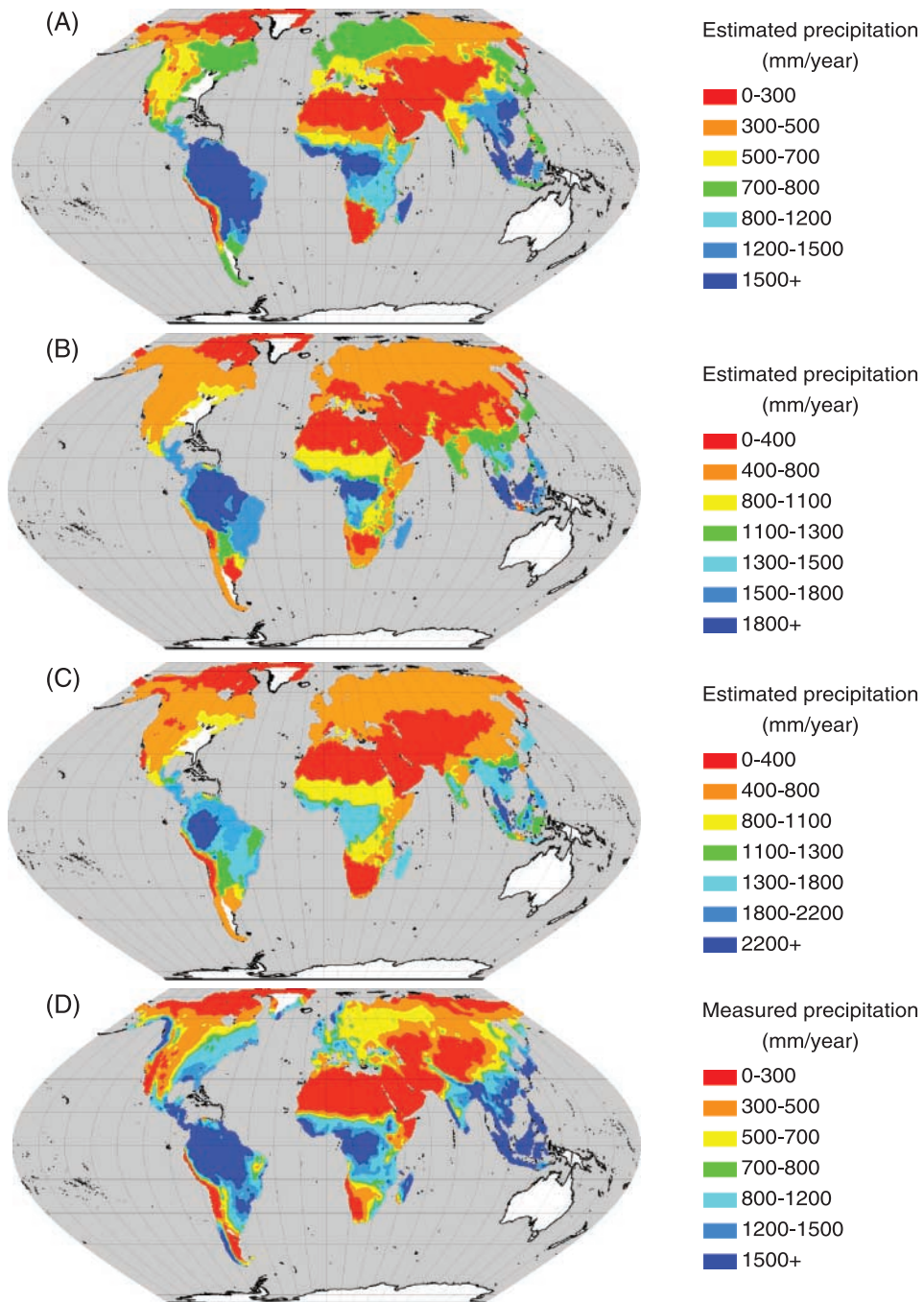
The regression tree appears to be a good way to estimate rainfall because it allows for non-linear responses and it takes the correlations between the input variables into account. Our results show that a non-linear algorithm estimates the precipitation better than does linear correlation. So the community structure of mammals may well respond non-linearly to changes in environmental variation.

Mammal communities reflect the vegetational habitat, which in turn reflects patterns of rainfall. The causal chain between rainfall and animal abundances goes through grass and foliage production, which is approximately linearly related to rainfall (Huxman *et al.*, 2004). The slope of the rainfall–production relationship, which can be expressed as WUE (water use efficiency), is a function of soil nutrient availability (Huxman *et al.*, 2004). In more moist regions, animal productivity is limited not by the gross availability of fodder but by its protein (nitrogen) content (Ellery *et al.*, 1996). Le Houerou (1984) and Le Houerou *et al.* (1988) related rain use efficiency (RUE) to productivity in arid environments. Huxman *et al.* (2004) demonstrated that RUE decreases as rainfall increases. They also related the rainfall to annual net primary production (ANPP) in world biomes. The slope of the ANPP–rainfall relationship decreases in mesic environments and increases in more arid environments (Huxman *et al.*, 2004). This means that production in the arid environments is more rainfall limited, whereas in more humid environments the production is temperature or nutrient limited. In high northern latitudes, primary production is controlled by temperature and light more than by precipitation (Churkina and Running, 1998; Nemani *et al.*, 2003), and there, the relationship between precipitation and mammalian community traits becomes less clear. The same applies to the extremely humid areas where rainfall does not limit primary production. In Figure 2 of

Huxman *et al.* (2004), one can see that in regions with precipitation of over 2000 mm per year, production increases very little when precipitation increases. In our regression tree (Fig. 1), we have only one (or a few; see evolutionary-ecology.com/data/2538A1.pdf) precipitation values in excess of 1500 mm. So our estimation method saturates and mammalian communities cannot be confidently used to estimate rainfall above ~2000 mm. Note, however, that this is a very high level of rainfall in present-day habitats. Most present-day habitats with rainfall exceeding 1500 mm are tropical forests and those over 2000 mm are tropical rainforests. Our regression trees can, in principle, differentiate between tropical forest and tropical rainforest. Beyond saying a place is a tropical rainforest, however, we cannot estimate its precipitation.

The estimated precipitation values presented here have quite a large standard error (see Table 2). This is in part caused by known exceptions to the general trend that result from re-invasion of humid habitats by hypsodont forms. In today's mammal fauna, the main examples of re-invasion are grazing bovid species that range into tropical rainforest (e.g. *Bos gaurus*, *Bos javanicus*, *Syncerus caffer*), and the elephants (*Elephas* and *Loxodonta*), which are highly mobile habitat generalists. We excluded the elephants from the analysis because their large home ranges and broad spectrum of habitats and climate conditions mean that they would contribute mostly noise to our analysis. We retained the bovinds, however, because their grazing diet is in harmony with their hypsodonty despite their atypical environmental setting. (The rainforest grasses on which they feed represent an interesting topic for future research that we cannot pursue here.) The exclusion of the two species of elephant from our regressions had only a small effect on the results. They would have contributed to just one further split in the extremely high precipitation estimate, a split that we are not able to resolve well in any case because of the saturation effect discussed above. However, despite these complications, the estimated annual precipitation values based on regression trees using hypsodonty, diet, and hypsodonty plus diet (Fig. 2A–C) depict the actual annual precipitation pattern (Fig. 2D) remarkably well. Africa in particular, excluding Madagascar, is extremely well estimated, and the Eurasian continent is relatively well estimated. The New World is also relatively well estimated, but there are more differences between estimated and actual values than in Africa or Eurasia. Below we discuss some possible reasons for the errors of estimation.

Fig. 2. (A) Predicted annual precipitation based on decision tree analysis using hypsodonty alone as regressor. The precipitation values have been classified to seven classes, reflecting the distribution of prediction outcome from Fig. 1. (B) Predicted annual precipitation based on decision tree for annual precipitation using diet alone as regressor. The precipitation values have been classified to seven classes, reflecting the distribution of prediction outcome from Appendix 1, Tree 1. Note that the classification is slightly different from that in (A). The main difference to (A) is a slight increase in the resolution of the high precipitation range (over 1500 mm). The difference to the pattern of (A) is that the areas of high and extremely high precipitation in the tropics are spatially better differentiated than they are in (A). (C) Predicted annual precipitation based on decision tree for annual precipitation using diet and hypsodonty in combination as regressors. The precipitation values have been classified to seven classes, reflecting the distribution of prediction outcome from Appendix 1, Tree 2. Note that the classification is slightly different from that in (A) and (B). The high precipitation range is better resolved than in (A) and (B) (over 1500 mm range), and the high precipitation areas in the tropics are spatially slightly better differentiated than they are in (A) and (B). (D) Measured annual precipitation from WorldClim data set calibrated to the same classes as (A).



The likely reason for the poor estimates in Madagascar is that it has almost no terrestrial herbivores. That appears to be an historical accident. There never were any large terrestrial herbivores on Madagascar, and most of the large terrestrial and arboreal species have gone

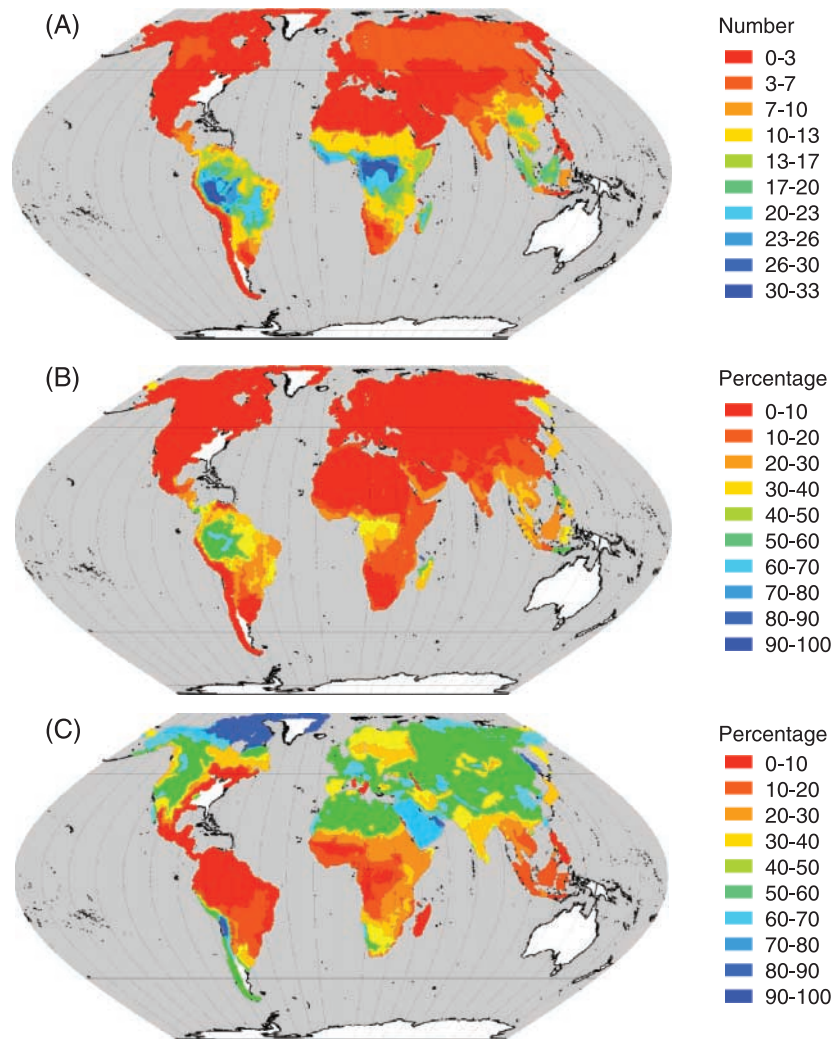


Fig. 3. (A) The number of species with low crowned teeth within grid cells. (B) The proportion of frugivores within grid cells. (C) The proportion of mixed feeders within grid cells.

extinct during the last 10,000 years due to excessive human hunting and human-induced environmental change (Burney *et al.*, 2004). All Madagascan primates have low crowned teeth, as is the case for all primates, and their diet alone is not enough to capture regional environmental variation in Madagascar. The proportion of frugivores (Fig. 3B) does capture some environmental variation, but this variable by itself is not adequate.

In Eurasia, the main differences between the maps (Fig. 2A–C vs. 2D) are in the alpine regions of Europe and the highlands of central India. Neither the hypsodonty nor the diet values capture the precipitation variation in the Alps. This is probably because our mammal data are too coarse to reveal the orographically driven precipitation increase in the Alps or on the western coast of Norway. Comparing Fig. 2A and Fig. 2B reveals that the diet-based

and hypsodonty-based prediction both capture some of the variation on the Indian subcontinent, although the regional variation is somewhat different. Their combined effect is seen in Fig. 2C. The orographically driven Indian monsoon in the South Himalayan region (Fig. 2D) is not captured by our estimation method. On the other hand, both hypsodonty- and diet-based estimators, separately and combined, captured the Asian monsoon precipitation regime to a fair extent (Fig. 2A–C). Perhaps our method is too coarse to detect orographically driven changes in precipitation. A likely reason is that mammals are mobile and track the seasonal variation in mountains, which causes mammal species to be recorded in multiple adjacent ecoregions.

The New World exhibits more differences between estimated and actual precipitation values than does the Old World. Part of the reason may be the large-mammal extinctions at the end of the Pleistocene (Koch and Barnosky, 2006). The wetter southeast of North America is not represented in our maps at all because of lack of ecomorphological data in these regions (not enough species with diet and tooth crown data). On the other hand, the drier, western areas of North America are represented in all of the estimation maps (Fig. 2A–C), as well as its more humid northeastern corner. In South America, tropical rainforest, with its extremely high precipitation, covers large areas. This is especially troubling for hypsodonty estimation (Fig. 2A). Yet the combined hypsodonty and diet estimation captures much of the precipitation variation in South America despite the existence of some problematical areas.

We expected that the substantial end-Pleistocene global loss of large mammal diversity (Koch and Barnosky, 2006) that has not been regenerated (Barnosky, 2008) would bias our results in some way. As shown above, there are some small-scale discrepancies, but much of the regional and all of the global scale patterns are well represented in our predictions. This is further confirmed by predictions for Africa. Africa was the continent least affected by end-Pleistocene megafaunal extinctions (Koch and Barnosky, 2006), and indeed Africa does show the best correspondence between predicted and actual rainfall patterns. This is significant for the potential use of this method to estimate climatic conditions of the past, as the African continent has the best available approximation of the mammal community structures that could be thought to represent pre-anthropogenic conditions.

CONCLUSIONS

Estimation methods based on regression trees offer powerful tools to unravel the complex relationships between the environment and the characteristics of mammalian communities. We developed a method to estimate precipitation from a set of ecomorphological characteristics of large herbivorous mammal communities. The regression trees described above all give reasonably accurate estimates of precipitation values for today's world, and thus we have some confidence that our methodology can be applied to paleocommunities to estimate past environmental conditions. The applicability of different regression trees to paleontological material is severely limited by the little information that is available for living species, such as direct observations of diet. Therefore, the trees that are preferable for reconstruction of past conditions are those that successfully use information derived directly from preserved morphology that correlates with diet, such as hypsodonty.

ACKNOWLEDGEMENTS

This work was supported by the Academy of Finland (M.F.), the Kone Foundation (J.T.E., L.L.), and the Emil Aaltonen Foundation (K.L.). We thank Arne Micheels, Torsten Utescher, Volker Mosbrugger, Louis Francois, and Angela Bruch for discussions on different proxy methods and climate modelling. We thank Jukka Jernvall for discussions on mammal ecomorphology and community structure and Susanna Sova for advice on Malagasy faunas. Finally, we thank Michael Rosenzweig for his thorough review and editing of the paper, which improved it considerably.

REFERENCES

- Barnosky, A.D. 2008. Megafauna biomass tradeoff as a driver of Quaternary and future extinctions. *Proc. Natl. Acad. Sci. USA*, **105**(suppl. 1): 11543–11548.
- Bell, 1971. A grazing system in Serengeti. *Sci. Am.*, **225**: 86–93.
- Belovsky, G.E. 1997. Optimal foraging and community structure: the allometry of herbivore food selection and competition. *Evol. Ecol.*, **11**: 641–672.
- Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. 1984. *Classification and Regression Trees* [software]. Belmont, CA: Wadsworth.
- Burney, D.A., Burney, L.P., Godfrey, L.R., Jungers, W.L., Goodman, S.M., Wright, H.T. *et al.* 2004. A chronology for late prehistoric Madagascar. *J. Human Evol.*, **47**: 25–63.
- Caughley, G. and Gunn, A. 1993. Dynamics of large herbivores in deserts: kangaroos and caribou. *Oikos*, **67**: 47–55.
- Churkina, G. and Running, S.W. 1998. Contrasting climatic controls on the estimated productivity of global terrestrial biomes. *Ecosystems*, **1**: 206–215.
- Clutton-Brock, T.H., Illius, A.W., Wilson, K., Grenfell, B.T., MacColl, A.D.C. and Albon, S.D. 1997. Stability and instability in ungulate populations: an empirical analysis. *Am. Nat.*, **49**: 195–219.
- Coe, M.J., Cumming, D.H. and Phillipson, J. 1976. Biomass and production of large African herbivores in relation to rainfall and primary production. *Oecologia*, **22**: 341–354.
- Damuth, J. and Fortelius, M. 2001. Reconstructing mean annual precipitation, based on mammalian dental morphology and local species richness. In *EEDEN Plenary Workshop on Late Miocene to Early Pliocene Environments and Ecosystems* (J. Agustí and O. Oms, eds.), pp. 23–24. Sabadell, Spain: EEDEN Programme, European Science Foundation.
- Damuth, J., Fortelius, M., Andrews, P., Badgley, C., Hadly, E.A., Hixon, S. *et al.* 2002. Reconstructing mean annual precipitation based on mammalian dental morphology and local species richness. *J. Vert. Paleontol.*, **22**(suppl.): 48A.
- Danell, K., Lundberg, P. and Niemelä, P. 1996. Species richness in mammalian herbivores: patterns in the boreal zone. *Ecography*, **19**: 404–409.
- Davidson, A.D., Hamilton, M.J., Boyer, A.G., Brown, J.H. and Ceballos, G. 2009. Multiple ecological pathways to extinction in mammals. *Proc. Natl. Acad. Sci. USA*, **106**: 10702–10705.
- Diniz-Filho, J.A.F., Bini, L.M. and Hawkins, B.A. 2003. Spatial autocorrelation and red herrings in geographical ecology. *Global Ecol. Biogeogr.*, **12**: 53–64.
- Du Toit, J.T. and Owen-Smith, N. 1989. Body size, population metabolism, and habitat specialization among large African herbivores. *Am. Nat.*, **133**: 736–740.
- Ellery, W.N., Scholes, R.J. and Scholes, M.C. 1996. The distribution of sweetveld and sourveld in South Africa's grassland biome in relation to environmental factors. *Afr. J. Range Forage Sci.*, **12**: 38–45.
- Eronen, J.T. 2006. Eurasian Neogene large herbivorous mammals and climate. *Acta Zool. Fenn.*, **216**: 1–72.
- Eronen, J.T. and Rook, L. 2004. The Mio-Pliocene European primate fossil record: dynamics and habitat tracking. *J. Human Evol.*, **47**: 323–341.

- Eronen, J.T., Mirzaie Ataabadi, M., Micheels, A., Karne, A., Bernor, R.L. and Fortelius, M. 2009. Distribution history and climatic controls of the Late Miocene Pikermian chronofauna. *Proc. Natl. Acad. Sci. USA*, **106**: 11867–11871.
- Fauquette, S., Guiot, J. and Suc, J.-P. 1998. A method for climatic reconstruction of the Mediterranean Pliocene using pollen data. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, **144**: 183–201.
- Fortelius, M. 1985. Ungulate cheek teeth: developmental, functional, and evolutionary interrelations. *Acta Zool. Fenn.*, **180**: 1–76.
- Fortelius, M. and Hokkanen, A. 2001. The trophic context of hominoid occurrence in the later Miocene of western Eurasia – a primate-free view. In *Phylogeny of the Neogene Hominoid Primates of Eurasia* (L. De Bonis, G. Koufos and A. Andrews, eds.), pp. 19–47. Cambridge: Cambridge University Press.
- Fortelius, M. and Solounias, N. 2000. Functional characterization of ungulate molars using the abrasion–attrition wear gradient: a new method for reconstructing paleodiets. *Am. Mus. Nov.*, **3301**: 1–36.
- Fortelius, M., Eronen, J.T., Jernvall, J., Liu, L., Pushkina, D., Rinne, J. *et al.* 2002. Fossil mammals resolve regional patterns of Eurasian climate change during 20 million years. *Evol. Ecol. Res.*, **4**: 1005–1016.
- Fortelius, M., Eronen, J.T., Liu, L.P., Pushkina, D., Tesakov, A., Vislobokova, I. *et al.* 2003. Continental-scale hypsodonty patterns, climatic paleobiogeography and dispersal of Eurasian Neogene large mammal herbivores. In *Distribution and Migration of Tertiary Mammals in Eurasia: A Volume in Honour of Hans De Bruijn* (J.W.F. Reumer and W. Wessels, eds.). *DEINSEA*, **10**: 1–11.
- Fortelius, M., Eronen, J.T., Liu, L., Pushkina, D., Tesakov, A., Vislobokova, I. *et al.* 2006. Late Miocene and Pliocene large land mammals and climatic changes in Eurasia. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, **238**: 219–227.
- Gordon, I.J. and Illius, A.W. 1989. Resource partitioning by ungulates on the Isle of Rhum. *Oecologia*, **79**: 383–389.
- Hawkins, B.A., Diniz-Filho, J.A.F., Bini, L.M., De Marco, P. and Blackburn, T.M. 2007. Red herrings revisited: spatial autocorrelation and parameter estimation in geographical ecology. *Ecography*, **30**: 375–384.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. and Jarvis, A. 2005. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.*, **25**: 1965–1978.
- Hoffmann, R.R. 1989. Evolutionary steps of ecophysiological adaptation and diversification of ruminants: a comparative view of their digestive system. *Oecologia*, **78**: 443–457.
- Hoffmann, R.R. and Stewart, D.R.M. 1972. Grazer or browser: classification based on the stomach structure and feeding habits of East African ruminants. *Mammalia*, **36**: 226–240.
- Huxman, T.E., Smith, M.D., Fay, P.A., Knapp, A.K., Shaw, M.R., Lolk, M.E. *et al.* 2004. Convergence across biomes to a common rain-use efficiency. *Nature*, **429**: 651–654.
- Janis, C.M. 1988. An estimation of tooth volume and hypsodonty indices in ungulate mammals, and the correlation of these factors with dietary preferences. In *Teeth Revisited: Proceedings of the VIIIth International Symposium on Dental Morphology* (D.E. Russell, J.-P. Santoro and D. Sigogneau-Russell, eds.), pp. 367–387. Paris: Mémoires de Musée d'Histoire Naturelle, series C.
- Janis, C.M. 1995. Correlations between craniodental anatomy and feeding in ungulates: reciprocal illumination between living and fossil taxa. In *Functional Morphology in Vertebrate Paleontology* (J.J. Thomason, ed.), pp. 76–98. Cambridge: Cambridge University Press.
- Janis, C. and Fortelius, M. 1988. On the means whereby mammals achieve increased functional durability of their dentitions, with special reference to limiting factors. *Biol. Rev. (Camb.)*, **63**: 197–230.
- Janis, C.M., Damuth, J. and Theodor, J.M. 2004. The species richness of Miocene browsers, and implications for habitat type and primary productivity in the North American grassland biome. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, **207**: 371–398.

- Jarman, P.J. (1974). The social organisation of antelope in relation to their ecology. *Behaviour*, **48**: 215–267.
- Jernvall, J. and Fortelius, M. 2002. Common mammals drive the evolutionary increase of hypsodonty in the Neogene. *Nature*, **417**: 538–540.
- Jernvall, J. and Fortelius, M. 2004. Maintenance of trophic structure in fossil mammal communities: site occupancy and taxon resilience. *Am. Nat.*, **164**: 614–624.
- Jones, M.J., Fielding, A. and Sullivan, M. 2006. Analysing extinction risk in parrots using decision trees. *Biodivers. Conserv.*, **15**: 1993–2007.
- Koch, P.L. and Barnosky, A.D. 2006. Late Quaternary extinctions: state of the debate. *Annu. Rev. Ecol. Evol. Syst.*, **37**: 215–250.
- Legendre, P. and Legendre, R. 1998. *Numerical Ecology*, 2nd edn. Amsterdam: Elsevier.
- Le Houerou, H.N. 1984. Rain use efficiency: a unifying concept in arid-land ecology. *J. Arid Environ.*, **7**: 213–247.
- Le Houerou, H.N., Bingham, R.L. and Skerbek, W. 1988. Relationship between the variability of primary production and the variability of annual precipitation in world arid lands. *J. Arid Environ.*, **15**: 1–18.
- McNaughton, S.J., Oesterheld, M., Frank, D.A. and Williams, K.J. 1989. Ecosystem-level patterns of primary productivity and herbivory in terrestrial habitats. *Nature*, **341**: 142–144.
- Mendoza, M. 2007. Decision trees: a machine learning methodology for characterizing morphological patterns resulting from ecological adaptations. In *Automated Taxon Identification in Systematics: Theory, Approaches and Applications* (N. MacLeod, ed.), pp. 261–276. London: Taylor & Francis.
- Mosbrugger, V. and Utescher, T. 1997. The coexistence approach – a method for quantitative reconstructions of Tertiary terrestrial palaeoclimate data using plant fossils. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, **134**: 61–86.
- Myers, P., Espinosa, R., Parr, C.S., Jones, T., Hammond, G.S. and Dewey, T.A. 2007. *Animal diversity* (available at: <http://animaldiversity.ummz.umich.edu/site/index.html>).
- Mysterud, A., Stenseth, N.C., Yoccoz, N.G., Langvatn, R. and Steinheim, G. 2001. Nonlinear effects of large-scale climatic variability on wild and domestic herbivores. *Nature*, **410**: 1096–1099.
- NatureServe. 2007. *NatureServe Explorer: An Online Encyclopedia of Life*, version 6.2. Arlington, VA: NatureServe (<http://www.natureserve.org/explorer>).
- Nemani, R.R., Keeling, C.D., Hashimoto, H., Jolly, W.M., Piper, S.C., Tucker, C.J. *et al.* 2003. Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science*, **300**: 1560–1563.
- Nowak, R.M. 1991. *Walkers' Mammals of the World*. Baltimore, MD: Johns Hopkins University Press.
- Olf, H., Ritchie, M. and Prins, H.H.T. 2002. Global environmental controls of diversity in large herbivores. *Nature*, **415**: 901–904.
- Olson, D.M., Dinerstein, D., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C. *et al.* 2001. Terrestrial ecoregions of the world: a new map of life on earth. *BioScience*, **51**: 933–938.
- Owen-Smith, N. 2002. *Adaptive Herbivore Ecology: From Resources to Populations in Variable Environments*. Cambridge: Cambridge University Press.
- R Development Core Team. 2009. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing (<http://www.R-project.org>).
- Reich, P.B., Walters, M.B. and Ellsworth, D.S. 1997. From tropics to tundra: global convergence in plant functioning. *Proc. Natl. Acad. Sci. USA*, **94**: 13730–13734.
- Roff, D.A. and Roff, R.J. 2003. Of rats and Maoris: a novel method for the analysis of patterns of extinction in the New Zealand avifauna before European contact. *Evol. Ecol. Res.*, **5**: 759–779.
- Rosenzweig, M. 1968. Net primary productivity of terrestrial communities: prediction from climatological data. *Am. Nat.*, **102**: 67–74.

- Rosenzweig, M.L. 1995. *Species Diversity in Space and Time*. Cambridge: Cambridge University Press.
- Sankaran, M., Hanan, N.P., Scholes, R.J., Ratnam, J., Augustine, D.J., Cade, B.S. *et al.* 2005. Determinants of woody cover in African savannas. *Nature*, **438**: 846–849.
- Smith, F.A., Lyons, S.K., Ernest, S.K.M., Jones, K.E., Kaufman, D.M., Dayan, T. *et al.* 2003. Body mass of late Quaternary mammals. *Ecology*, **84**: 3403.
- Solounias, N., Fortelius, M. and Freeman, P. 1994. Molar wear rates in ruminants: a new approach. *Ann. Zool. Fenn.*, **31**: 219–227.
- Therneau, T.M., Atkinson, B. and Ripley, B. 2009. *rpart library*, version 3.1-45. [computer program] (<http://cran.r-project.org/web/packages/rpart/index.html>).
- Van Valen, L. 1960. A functional index of hypsodonty. *Evolution*, **14**: 531–532.
- Venables, W.M. and Ripley, B.D. 2002. *Modern Applied Statistics with S*, 4th edn. New York: Springer.
- Walker, B.H. and Langridge, J.L. 1997. Predicting savanna vegetation structure on the basis of plant available moisture (PAM) and plant available nutrients (PAN): a case study from Australia. *J. Biogeogr.*, **24**: 813–825.
- Wilson, D.E. and Reeder, D.M. 2005. *Mammal Species of the World*. Baltimore, MD: Johns Hopkins University Press.
- Wilson, D.E. and Ruff, S. 1999. *The Smithsonian Book of North American Mammals*. Washington, DC: Smithsonian Institution.
- Wolfe, J.A. 1979. Temperature parameters of humid and mesic forests of eastern Asia and relation to forests of other regions of the northern hemisphere and Australasia. *US Geol. Surv. Prof. Pap.*, **1106**: 1–37.
- Wolfe, J.A. 1993. A method for obtaining climate parameters from leaf assemblages, *US Geol. Surv. Bull.*, **204**: 1–71.

