

# Modelling extinction risk in multispecies data sets: phylogenetically independent contrasts versus decision trees

J. Bielby · M. Cardillo · N. Cooper · A. Purvis

Received: 4 November 2008 / Accepted: 10 August 2009  
© Springer Science+Business Media B.V. 2009

**Abstract** Many recent studies of extinction risk have attempted to determine what differences exist between threatened and non-threatened species. One potential problem in such studies is that species-level data may contain phylogenetic non-independence. However, the use of phylogenetic comparative methods (PCM) to account for non-independence remains controversial, and some recent studies of extinction have recommended other methods that do not account for phylogenetic non-independence, notably decision trees (DTs). Here we perform a systematic comparison of techniques, comparing the performance of PCM regression models with corresponding non-phylogenetic regressions and DTs over different clades and response variables. We found that predictions were broadly consistent among techniques, but that predictive precision varied across techniques with PCM regression and DTs performing best. Additionally, despite their inability to account for phylogenetic non-independence, DTs were useful in highlighting interaction terms for inclusion in the PCM regression models. We discuss the implications of these findings for future comparative studies of extinction risk.

**Keywords** Comparative analyses · Conservation · Decision trees · Extinction risk · Non-independent data · Phylogenetic comparative methods

---

**Electronic supplementary material** The online version of this article (doi:[10.1007/s10531-009-9709-0](https://doi.org/10.1007/s10531-009-9709-0)) contains supplementary material, which is available to authorized users.

---

J. Bielby (✉)

Institute of Zoology, The Zoological Society of London, Regent's Park, London NW1 4RY, UK  
e-mail: [jon.bielby@ioz.ac.uk](mailto:jon.bielby@ioz.ac.uk)

J. Bielby · N. Cooper · A. Purvis

Department of Biological Sciences, Imperial College London, Silwood Park Campus, Ascot SL5 7PY, UK

M. Cardillo

Centre for Macroevolution and Macroecology, Research School of Biology, Australian National University, Canberra, ACT 0200, Australia

## Abbreviations

DTs Decision trees

PCM Phylogenetic comparative methods

TIPS Comparative analyses using species (the 'tips' of phylogenetic tree branches) as independent data-points

## Introduction

Given ongoing rates of biodiversity loss, there is a real need for ecology and conservation to become more predictive in order to make predictions that practitioners or policy-makers can use (Sutherland 2006). One area in which predictive models have become more common in recent years is in the examination of species' susceptibility to the processes causing increased levels of extinction risk. In particular, there has been an increased interest in how inherent traits, such as species life-history and ecology, and external factors, such as human population pressure, interact to produce the current pattern of species risk. As a result the number of studies testing proposed correlates of extinction risk has grown considerably, incorporating a wide range of taxonomic groups (Cardillo et al. 2005a; Cooper et al. 2008; Koh et al. 2004; Laurance 1991; Owens and Bennett 2000; Reed and Shine 2002; Sullivan et al. 2000) at a range of scales (Fisher and Owens 2004; Purvis et al. 2005). Such studies have two main aims. First, they can discriminate among competing hypotheses about which intrinsic traits of species' biology predispose species to decline in the face of anthropogenic impacts. Second, quantitative models linking biological traits to extinction risk can be used to predict which species may face a high risk of extinction in the future, and therefore inform conservation management options (Cardillo et al. 2006, 2004; Jones et al. 2006; Sullivan et al. 2006).

How should comparative data on threatened and non-threatened species be analysed? Although extinction risk itself is not an evolved trait (Cardillo et al. 2005b; Putland 2005), it often has a phylogenetic signal (Bennett and Owens 1997; Bielby et al. 2006; McKinney 1997; Purvis et al. 2000a; Russell et al. 1998) as do many of the proposed predictor variables (Freckleton et al. 2002; Purvis et al. 2005). Ignoring this phylogenetic signal and treating species as independent is likely to result in pseudoreplication (Felsenstein 1985) which may render subsequent analyses statistically invalid, with an elevated Type I error rate (Harvey and Pagel 1991; although Type II errors may also occur, Gittleman and Luh 1992). In order to avoid such errors, the use of phylogenetically independent contrasts (Felsenstein 1985) has become widespread when testing proposed correlates of extinction risk (Fisher and Owens 2004; Purvis 2008), while other similar methods such as generalised estimating equations (Paradis and Claude 2002) have also been advocated.

Although their use is widespread, there are questions about the suitability of phylogenetic comparative methods (PCM) in studies of extinction risk. The phylogenetic signal in extinction risk is not necessarily as strong as assumed by the most commonly used approach, the independent contrasts method which generates contrasts as differences between sister clades and uses these as the basis for analysis (e.g. Collen et al. 2006; Purvis et al. 2005; see also Cardillo et al. 2005a, b). The correction for non-independence may therefore be too severe, potentially leading to 'over-correction' (Ricklefs and Starck 1996), with comparisons between very closely related species having undue influence on subsequent analyses (Purvis 2008). While some studies have incorporated data-led branch

length adjustments in order to ameliorate over-correction (e.g. Cardillo et al. 2005a; Halsey et al. 2006; Stuart-Fox et al. 2007) these may not remove the problem entirely. Further error may be introduced into PCM analyses as a result of inaccuracies in the phylogenetic tree, branch lengths, and violations of assumptions made regarding the mode of evolution (Ives et al. 2007; Ricklefs and Starck 1996; Symonds 2002) limiting the benefits of using PCM. Additionally, the sheer number of candidate predictor variables included in models of extinction risk may potentially remove the need to account for phylogeny: if a model contains most of the important variables that produce non-independence in the error term of the model, the need to correct for phylogenetic relationships can be greatly reduced (Grafen 1989). In reality, however, most models explain less than 50% of the variation in extinction risk.

Although categorical explanatory and response variables are frequently included in analyses (Fisher and Owens 2004), most PCM approaches are designed to deal with continuous traits. Some methods (e.g. Grafen 1989) permit both continuous and discrete explanatory variables but very few approaches, aside from the one used in the analyses presented here (e.g. Paradis and Claude 2002), are intended for binary response variables (e.g., threatened vs. not threatened) that may be important in studies of extinction risk.

The early focus in PCM was testing whether an association between two variables was significant when pseudoreplication was removed (Felsenstein 1985; Ridley 1983), more than investigating the form of any relationship (e.g. nonlinearity) or estimating parameters. Extinction risk models are often complex, with many hypothesised predictors potentially interacting, and parameter estimation is now more commonly of interest as PCM are being used to build predictive models with the aim of, for example, prioritising conservation actions (Cardillo et al. 2006). Although phylogenetic and nonphylogenetic approaches estimate the same underlying parameters (Pagel 1993), measurement error of the traits used has a greater effect on phylogenetic than nonphylogenetic comparative studies (Harmon and Losos 2005; Purvis and Webster 1999; Ricklefs and Starck 1996). Additionally, most implementations of PCM assume linearity of relationships and do not lend themselves to straightforward testing or relaxation of this assumption. Although complexities such as non-linearities and interactions between explanatory variables can be included in the environment of a model incorporating PCM (e.g. Cardillo et al. 2005a; Cooper et al. 2008), more commonly they are not (Quader et al. 2004). Visualising such complexities when using contrasts is more difficult than with raw species data, which may partly explain why they are not more often included in analyses. The addition of non-linearities and interaction terms alongside numerous predictor variables may also lead to 'parameter proliferation' and further reduce the power of the analyses in question (Crawley 2002).

Decision trees (hereafter DTs) provide an alternative approach for building complex predictive models that may be more capable of dealing with some of these obstacles (Breiman et al. 1984; Crawley 2002; De'ath and Fabricius 2000). DTs are statistical models that highlight which explanatory variables (continuous or categorical) are the most important in explaining variation in a specified response variable (continuous or categorical). DTs are built by repeatedly splitting the data into two subsets. At each split, or branch, the data is partitioned on the basis of whether it falls above or below a threshold value of a selected predictor variable. The predictor variable and threshold value are selected that explain most deviance in the response variable, making the two subsets each as homogenous as possible. The process is continued until further splits explain no further deviance or the terminal node reaches a minimum size (i.e., number of cases). In explaining variation in the response, each variable may be used once, more than once, or not at all; this allows identification of variables that interact, have nonlinear effects, or are

not included in the model. In recent years the utility of DTs has been reflected in the number of publications using them for a range of purposes, such as modelling climatic niches (e.g. Araujo et al. 2005) and species distributions (e.g. Elith et al. 2006). It has recently been argued that DTs are an ideal way to model extinction risk (Sullivan et al. 2006) and conservation status (Jones et al. 2006; Koh et al. 2004) but, while they are able to highlight important interactions, they do not account for the phylogenetic structure in comparative data. Additionally, the structure of DTs may sometimes be unstable with the importance of variables and the order of those variables changing as a result of small changes in the available data. Such changes can make it difficult to identify which traits are most important in terms of predicting a species' risk of extinction, although predictive accuracy and precision may not be negatively affected.

If PCMs and DTs are both considered as useful tools in modelling extinction risk, it is important to clarify how the two approaches perform relative to each other and more traditional cross-species (hereafter TIPS) regression models when analysing trait associations with extinction risk. Although Sullivan et al. (2006) advocated the use of DTs, they did not perform suitable comparable analyses using independent contrasts or an alternative PCM. Here, we analyse extinction risk in both amphibians and mammals, and Rapid Decline (a binary response variable) in amphibians, using PCM, TIPS regression models and DTs. The use of different clades, kinds of response variable, and phylogenetic hypotheses of differing certainty allow us to compare how the modelling techniques perform under a range of conditions with the aim of evaluating the strengths of each in building predictive models of extinction risk. Specifically we aimed to address the following four questions:

1. Do DTs identify interactions and non-linearities of explanatory variables?
2. Are the results of DTs and TIPS regressions affected by the possible non-independence of species level data?
3. Are the predictions made regarding species' susceptibility to extinction risk consistent across modelling techniques?
4. How precise are the predictions made by the different techniques?

By highlighting strengths and weaknesses of all three approaches and evaluating how similar their results are, we hope to be able to make recommendations for future analyses of extinction risk.

## Methods

### Analyses

We took PCM analyses of extinction risk from three recent completed studies and applied TIPS and DT to them. Following the original analyses (Cardillo et al. 2008; Cooper et al. 2008), we modelled extinction risk in two clades (mammals and frogs) and four separate orders within mammals, using both PCM and TIPS regression models, and DTs. In both mammalian and anuran analyses our measure of extinction risk were the IUCN Red List categories (IUCN 2004) converted to a continuous index ranging from 0 to 5, 0 corresponding to "Least Concern", 5 to "Extinct in the Wild" or "Extinct" (Cardillo et al. 2005a; Cooper et al. 2008; Purvis et al. 2005). We excluded threatened species not listed under criterion A of the Red List from our analyses of extinction risk to avoid circularity (Purvis et al. 2000b).

**Table 1** Description of the comparisons made and optimal tree size (defined by the number of terminal nodes) for each as determined by tenfold cross validation

Comparison reference	Comparison	Type of variable	Species included/total data available	Optimal tree size
A	Extinction risk in frogs	Continuous	125/367	2
B	RD vs. non-RD threatened frogs	Binary	112/359	3
C	Enigmatic RD vs. non-Enigmatic RD threatened frogs	Binary	327/359	4
D	Bd+ RD species vs. Bd+ non-RD species of frogs	Binary	101/102	3
E	Bd+ RD species vs. Bd– RD species frogs	Binary	84/227	4
F	Extinction risk in mammals	Continuous	400/3,989	10
G	Extinction risk in bats	Continuous	268/837	12
H	Extinction risk in marsupials	Continuous	48/231	5
I	Extinction risk in primates	Continuous	98/206	13
J	Extinction risk in rodents	Continuous	1,213/1,775	11

Additionally, we modelled a binary variable, Rapid Decline (hereafter “RD”), in anurans. RD species are those that have experienced a genuine increase in IUCN Red list category since 1980. In addition to comparing all RD species with non-RD threatened species, we analysed susceptibility to RD as a result of different threatening processes: “enigmatic” threats (Stuart et al. 2004), chytridiomycosis, and infection by the pathogen *Batrachochytrium dendrobatidis* (hereafter “Bd”) among RD species. The range of comparisons made allowed us to evaluate the utility of DTs, phylogenetic comparative method (PCM), and TIPS regression models in a number of different clades, using phylogenetic trees of different resolution, with different types of response variables (i.e. binary and continuous). The exact details of the analyses performed are described in Table 1.

## Predictor variables

### *Anurans*

Biological data on frogs species came from the dataset underlying two previous analyses (Bielby et al. 2008; Cooper et al. 2008). Full details of data collection, and variables examined are provided in those papers. Appendix A contains further details of variables, variable transformation, method of collection and sources used.

The final anuran dataset (available in Appendix B) contained 553 frog species, representing 32 of 33 families from all six continents on which amphibians are present. The number of species in the dataset was restricted largely by the lack of biological and life history data available.

### *Mammals*

Biological data on mammalian species came from the “PanTheria” database, a dataset of ecological and life-history traits for 4,030 mammalian species, from over 3,300 published literature sources. Full details of data collection and quality checking are provided

in Jones et al. (2009). Mammalian geographic distributions were used to calculate range size and for summarizing human impact and environmental conditions within each species' range. Details of the predictor variables included in the mammalian analyses are in Cardillo et al. (2008).

### Phylogenetic and non-phylogenetic regression models

#### *Modelling extinction risk in anurans (comparison A)*

In frogs, the relationship between extinction risk and the predictor variables outlined in Appendix A were explored using phylogenetically independent contrasts as outlined in Cooper et al. (2008). A minimum adequate model (MAM) was built by including all variables in the model and deleting those with the highest  $P$ -value until all remaining terms were significant at  $P \leq 0.05$ .

#### *Modelling RD in anurans (comparisons B–E)*

Phylogenetic models of frog susceptibility to Rapid Decline were built using generalised estimating equations (GEEs; Paradis and Claude 2002), as described in Bielby et al. (2008).

For all of the frog analyses, we used a recently published genus-level amphibian phylogeny (Frost et al. 2006) with taxonomy as a surrogate for phylogeny below the genus level. As we did not have branch lengths for all species, we set branch lengths to equal 1 unit.

#### *Modelling extinction risk in mammals and mammalian orders (comparisons F–J)*

In mammals as a whole (comparison F) and each of four clades (comparisons G–J) extinction risk was modelled using phylogenetically independent contrasts using a dated, composite supertree phylogeny of 4,510 mammalian species (Bininda-Emonds et al. 2007), and following the protocol of Cardillo et al. (2008). MAMs were built by starting with a full set of predictor variables, including additive linear effects only and followed the model building procedures described in Cardillo et al. (2004) and Purvis et al. (2000b). For all analyses using PCM, equivalent TIPS models were built using the same model-building algorithms. Contrasts were calculated after transforming phylogenetic branch lengths by raising them to a power ( $\kappa$ ), with the value of  $\kappa$  optimized for each variable to minimize the correlation between absolute scaled contrasts and their standard deviations (Garland et al. 1992).

#### *Tree models*

DTs were first of all grown to a maximal size, that is, until further splits result in no decrease in the residual deviance or terminal nodes reach a minimum size ( $n = 5$ ), and then pruned to an optimum that is determined by cross-validation techniques. However, as a result of the high occurrence and the uneven distribution of missing values of predictor variables among cases, we first needed to reduce the impact of missing values on the tree-building process and the results obtained. One possibility for reducing the number of missing values would be to use multiple imputation methods (Little and Rubin 2002).

However, using such methods on species-level comparative datasets may result in a loss of phylogenetic structure in the dataset (see Cardillo et al. 2008). For the purposes of these analyses, which aimed to compare results of DTs and PCMs, with specific focus on the impact of phylogenetic structure within the data, we felt that inputting data would risk further complicating the comparisons we made. The default setting is for DTs is to leave out any cases with missing values, which may result in a tree built on a subset of the data that is limited in size by an unimportant predictor variable (i.e. one that does not explain variance in the response variable). We maximised the sample size of cases included in the DT by adopting the following procedure. A maximal tree was grown from the full set of predictor variables. Of the predictor variables not included in the maximal tree, the sample-limiting variable was removed from the dataset. This reduced the number of cases with missing values, and increased the number of cases entering the tree building process. The full-sized tree was then rebuilt with the new, reduced data subset. Once again the unimportant predictor most limiting sample size was dropped from the tree building subset. This process was continued until no further changes to the full tree occurred. Although the method used to minimize the number of missing values may under some circumstances affect the constituent variables, it should not affect the predictive accuracy of the model. As the focus of these analyses was to assess predictive accuracy and consistency among methods, we felt that the methods used were suited to the task.

The tree size (i.e. the number of terminal nodes) was then optimised using tenfold cross-validation conducted on each tree size from the maximum to the minimum number of terminal nodes. For each of 500 iterations, the dataset was split into 10 subsets, each of which in turn was excluded from the tree-building process, while the remaining nine were used to model the data. The excluded subset was then used to validate the model by comparing the predicted fitted values with observed values of the response, yielding an estimated error (i.e. squared errors; SE) for that subset. The 500 iterations of the tenfold validation process therefore led to 5,000 values of SE for a tree of the specified size, which were used to calculate a mean squared error (MSE) and associated standard error. The optimal tree size was the smallest tree within one standard error of the tree size with the lowest MSE. Once the optimal tree size was known, a tree of that size was built using the dataset obtained in the sample size maximising process (De'ath and Fabricius 2000). All DTs were built using the tree function (package = tree) of the statistical language R (R Development Core Team 2009), and all cross-validation was performed using scripting written in R.

### *Consistency of model predictions*

The fitted values obtained via each technique were not constrained to the integers that formed the response variables for that comparison (e.g. 0 or 1 in comparisons B–E) but could take the form of any decimal value within the limit of that comparison. To assess the consistency of the predictions made by different modelling techniques, we performed pairwise Pearson's correlations of fitted values obtained using each technique. The correlations were limited to those species that were common to the two techniques being compared.

### *Precision of model predictions*

Although strong correlations between fitted values may indicate that predictions are consistent across techniques, they do not indicate how precise these predictions were. In order

to determine whether the prediction precision varied significantly among techniques the residual values obtained from the models were squared and compared using Kruskal–Wallis tests with modelling technique as a factor, the null hypothesis being that no significant differences existed. In the case of a significant difference among techniques, post-hoc non-parametric multiple comparison tests were conducted to determine where those differences lay. Squared residual values (squared error) were a suitable measure of predictive precision as they quantify the difference between the predicted and observed extinction risk of a given species. All Kruskal–Wallis and subsequent post-hoc tests were conducted using the `kruskal.test` and `npmc` (library = `npmc`) functions of the statistical language R (R Development Core Team 2009).

#### *Phylogenetic structure in model residuals*

The squared residual values obtained from each model were analysed using ANOVA with taxonomic Family as a factor, as a simple test of whether a phylogenetic structure was present. The presence of phylogenetic structure in the model residuals suggests that the predictions made varied in precision among sub-clades at or below the Family level.

#### *Complexities in the data*

In order to determine whether DTs were useful in outlining complexities in the data we compared interactions and non-linearities included as significant predictors in the regression models to those outlined in the DTs. This allowed a qualitative comparison of complexities in the data. Additionally, any important interactions or non-linearities highlighted by DT models were added as terms in the relevant PCM regression model, with the aim of seeing whether they were significant predictors of extinction risk, and whether they improved the predictive precision judged by the MSE. In order to fit interaction terms, the single predictors within the term were also included in the model.

## **Results**

### Optimal tree size

The optimal DT size obtained through tenfold cross-validation for each comparison is given in Table 1. The size of optimum trees varied greatly ranging from only two terminal nodes for comparison A to 13 terminal nodes for comparison I. The proportion of species included in each tree also varied for both amphibian and mammalian models as a result of missing values in the dataset. Details of each optimal DT model in terms of number of nodes and number of constituent species are included in Table 1; the total data available highlighted in the table is the number of species meeting the criteria for inclusion in that analysis (e.g. after all non-RD species were removed from the dataset there were a total of 227 species for inclusion in comparison E).

### Consistency of model predictions

The consistency of the predictions made by all the three techniques was strong and the fitted values were always significantly positively correlated among methods (see Table 2

**Table 2** Summary of technique performance in terms of consistency, precision, and presence of a phylogenetic signal

Comparison	Modelling technique	Correlation coefficient of fitted values			Mean squared error	Phylogenetic signal in residuals ( <i>F</i> )
		PCM	TIPS	Tree		
A	PCM	–	0.96***	0.78***	0.17 (A)	5.03***
	TIPS	–	–	0.84***	0.83 (C)	2.47***
	DT	–	–	–	0.76 (B)	1.45
B	PCM	–	0.58***	0.66***	0.29 (B)	20.43***
	TIPS	–	–	0.53***	0.42 (C)	4.26***
	DT	–	–	–	0.22 (A)	3.39***
C	PCM	–	0.40***	0.20*	0.26 (B)	6.84***
	TIPS	–	–	0.77***	0.37 (C)	4.66***
	DT	–	–	–	0.23 (A)	6.12***
D	PCM	–	0.97***	0.87***	0.06 (A)	1.52
	TIPS	–	–	0.86***	0.29 (B)	0.51
	DT	–	–	–	0.11 (A)	1.396
E	PCM	–	0.48***	0.43***	0.14 (A)	2.25***
	TIPS	–	–	0.71***	0.40 (B)	0.58
	DT	–	–	–	0.22 (A)	1.05
F	PCM	–	0.81***	0.66***	0.12 (A)	2.52***
	TIPS	–	–	0.70***	0.42 (B)	1.55***
	DT	–	–	–	0.24 (A)	1.89***
G	PCM	–	0.72***	0.55***	0.20 (A)	1.59 <sup>a</sup>
	TIPS	–	–	0.35***	0.42 (C)	1.56 <sup>a</sup>
	DT	–	–	–	0.21 (B)	1.19
H	PCM	–	0.67***	0.63***	0.01 (A)	3.53***
	TIPS	–	–	0.75***	0.22 (B)	1.209
	DT	–	–	–	0.13 (A)	1.52
I	PCM	–	0.90***	0.75***	0.17 (A)	3.38***
	TIPS	–	–	0.75***	0.66 (B)	1.65 <sup>a</sup>
	DT	–	–	–	0.27 (A)	1.85
J	PCM	–	0.71***	0.39***	0.39 (B)	3.47***
	TIPS	–	–	0.45***	0.32 (B)	2.00 <sup>a</sup>
	DT	–	–	–	0.28 (A)	2.99***

<sup>a</sup>  $P = 0.05\text{--}0.1$ ; \*  $P < 0.05$ ; \*\*  $P < 0.01$ ; \*\*\*  $P < 0.001$ 

for summary, Appendix C for full details). However, there was variation in the strength of the correlations among comparisons. In two of the four analyses with a binary response (comparisons C and E), there were marked differences in predictive consistency between methods. In these comparisons the consistency was lower between PCM and the other methods (see Table 2) than between TIPS regression and DTs, suggesting that non-phylogenetic approaches made more consistent predictions. However, when response variables were continuous (A, F–J) there was generally a stronger correlation between PCM and TIPS regression than between DTs and the other methods (Table 2).

## Precision of the model predictions

Kruskal–Wallis tests on the model residuals showed that there were significant differences in the predictive abilities of modelling techniques within a comparison. Results of non-parametric multiple comparison tests show where these differences lay (see Table 2 for summary, Appendix D for full details). In three comparisons, DTs predicted level of risk more precisely than PCM (comparisons B, C, and J), whereas PCM was more precise in two comparisons (A and G). DTs and PCM performed equally well in five comparisons (D, E, F, H, and I). In comparisons with a binary response (B–E) DT performed particularly well, having a predictive precision that was better or on a par with PCM (Table 2). In two of the six comparisons with a continuous response variable, PCM made predictions that were more precise than DT (comparisons A and G, but not F, H, I or J), whereas DT was more precise in one comparison (J). TIPS regressions had significantly higher residual values, and therefore lower predictive precision, than PCM in all but one comparison (comparison J).

## Phylogenetic signal in model residual values

Analysis of residuals from PCM models suggest the presence of a strong phylogenetic effect at the Family level in eight of the ten comparisons made (A, B, C, E, F, H, I, and J; see Table 2 for summary, Appendix E for full details). In one further comparison (G), model residuals had a phylogenetic signal with a *P*-value of between 0.05 and 0.1. The residuals of four TIPS regression models (comparisons A, B, C, and F) contained a significant phylogenetic signal at the Family level. Similarly, the residuals of the DTs in four of the comparisons contained a phylogenetic signal (comparisons B, C, F, and J). In TIPS regression, comparisons G, I and J had a phylogenetic signal that had a *P*-value of <0.1. Model residuals from all three techniques of comparison D lacked a phylogenetic signal. However, the small number of degrees of freedom in this and some other comparisons (e.g. E, H, and I) means that these analyses may have very low power to detect a phylogenetic pattern.

## Non-linear relationships

In PCM regression models, 11 non-linear relationships were identified as significant predictors of extinction risk in five comparisons (Comparison A: geographic range<sup>2</sup>, geographic range<sup>3</sup>, temperature<sup>2</sup> and temperature<sup>3</sup>; Comparison F: HPD<sup>2</sup> and Population density<sup>2</sup>; Comparison G: adult body mass<sup>2</sup>, geographic range<sup>2</sup>, and HPD of the 5th percentile<sup>2</sup>; Comparison H: geographic range<sup>2</sup>; Comparison I: HPD of the 5th percentile<sup>2</sup>). Fewer (six) significant non-linear relationships were found in TIPS regression models (Comparison A: geographic range<sup>2</sup>, geographic range<sup>3</sup>, temperature<sup>2</sup>, temperature<sup>3</sup>, snout-vent length<sup>2</sup> and snout-vent<sup>3</sup>). None of the non-linear relationships outlined in PCM or TIPS regression models were obvious from the corresponding DTs.

## Interactions between predictor variables in regression models

In the PCM regression models, four significant interaction terms in three comparisons were found to be significant predictors of extinction risk (Comparison F: geographic range:HPD, and population density:HPD; Comparison G: geographic range:Actual Evapotranspiration;

**Table 3** Interaction terms added to the PCM regression models based on the corresponding DT

Comparison	Interaction term	Significant as a term?	Original predictive accuracy (MSE)	Predictive accuracy (MSE) after the addition of interaction term
A	NA	NA	NA	NA
B	Aquatic lifestage and NPP	No	0.29	0.22
C	HPD and isothermality	No	0.23	0.262
C	Isothermality and Temp	No	0.23	0.247
D	Geographic range and altitude	No	0.06	0.07
E	Clutch size and HPD	No	0.14	0.21
E	Clutch size and altitude	No	0.14	0.22
F	Neonatal body mass and litters per year	Yes	0.12	0.02
G	Adult body mass and geographic range	Yes	0.2	0.08
H	NA	NA	NA	NA
I	Adult body mass and geographic range	Yes	0.17	0.07
J	Geographic range and AET	No	0.09	0.23

In comparison A and H there were no informative interactions in the optimal tree to add to the PCM model

Comparison I: adult body mass: absolute latitude). No significant interaction terms were found in the TIPS regression models.

#### Inclusion of interactions from DTs

The interaction terms added to the PCM regression model on the basis of the DTs are presented in Table 3. None of the interaction terms added to models of extinction risk/RD in anurans were significant predictors, although in comparison B the predictive precision of the model was improved by the addition of the term. In contrast, the interaction terms added to three of the four mammalian clade models (comparisons F, G, I, but not J) were significant predictors of extinction risk and improved the predictive precision of the models.

## Discussion

The analyses of extinction risk and RD presented here were used to assess the utility of three different approaches: PCM, TIPS regression, and DTs. Although there has been discussion surrounding the use of these three techniques (Harvey and Rambaut 1998; Putland 2005; Sullivan et al. 2006), these analyses represent the first systematic comparison of the methods, and have important implications for future studies.

Generally, both PCM and DTs performed well in terms of the precision of the predictions of a species' extinction risk. DTs were more precise than PCM in three of the ten comparisons, while PCM were more precise in two. TIPS analyses clearly performed less well than the other techniques: in addition to their inability to account for phylogenetic non-independence, and higher rate of type I and type II error (Harvey and Rambaut 1998)

TIPS regression methods are less precise in their predictions than alternative modelling techniques (either PCM or DT). We therefore cannot recommend the use of TIPS regression models in modelling extinction risk.

Although both PCM and DT performed well in terms of precision and consistency, some of the differences between the results suggest that DTs are not well suited to modelling continuous, rather than binary, response variables. The limited range of predictions available via the DT approach may explain some of the technique's relatively poor performance compared to PCM in comparisons with a continuous response variable. In analyses of extinction risk, the observed values of the response variable ranged from 0 to 5, as did the fitted values obtained via PCM and TIPS regression. However, with DT the fitted values were restricted to a finite number of discrete values defined by the terminal nodes of that DT. The lower consistency in predictions made by DTs versus the regression techniques may in part be explained by the correlation of continuous and discrete fitted values in these comparisons. Likewise, the limited number of fitted values obtained via DTs compared to PCM reduced the precision with which DTs resulted in larger residual values in comparisons with continuous response variables leading to a reduced precision compared to PCM. Although in comparisons with a binary response variable DT fitted values are also limited to a finite number of discrete values, the range of possible fitted values is much lower, making this artefact less likely to dominate the differences in consistency and precision among techniques. Furthermore, the continuous response used in our analyses only ranges from 0 to 5; perhaps with a larger range there may be more of a discrepancy in the predictive precision of DTs and PCM regressions.

In addition to their relatively strong predictive ability, on the basis of our results DTs may be a useful tool in identifying informative interactions in datasets (De'ath and Fabricius 2000; Jones et al. 2006; Sullivan et al. 2006). Non-linear relationships and interactions among variables are rarely investigated in macroecological studies, and yet may be particularly important in some instances, including analyses of extinction risk (e.g. Cardillo et al. 2005a). Our results suggest that DTs could be a useful tool in identifying non-linear relationships and interactions that should be included as terms in the PCM regression model-building process. The inclusion of interaction terms in PCMs on the basis of their appearance in DT resulted in interactions remaining as significant terms which improved predictive ability in three of five of our analyses of mammalian extinction risk (comparisons F, G, and I). The inclusion and impact of these interactions, highlights how insightful DT models may be in informing PCMs with regards complexities in the dataset. The fact that DTs did not identify all significant interactions and nonlinearities suggests that they should not be used as the sole method for finding such complexities, but that they could usefully supplement existing model building algorithms.

One potential weakness that DT, or any other non-PCM, may have in modelling extinction risk is their inability to account for phylogenetic non-independence, a subject that has generated much discussion (Cardillo et al. 2005b; Harvey and Rambaut 1998; Putland 2005; Sullivan et al. 2006). The strong performance of PCM in the analyses presented here suggests that accounting for phylogeny does allow precise predictions of species' conservation status. Additionally, the residuals obtained from PCM analyses suggest the presence of a phylogenetic signal among taxonomic Families within a comparison, a signal that was not obvious from the corresponding TIPS and DT analyses. The existence of a phylogenetic signal in PCM analyses is perhaps not surprising. The strong signal suggests that the relationships between explanatory and response variables varied among subclades within a comparison. By design, PCM models reduce the influence of clade-specific relationships when the overall model is fitted, which in the presence of

strong among-clade differences would result in the observed structure in the residual values (see Purvis et al. 2000b). Any variation in regression parameters among subtaxa may be important in conservation planning as they highlight the importance and utility of focussed taxonomic studies if predictive models are to be used for directing applied management or policy (Cardillo et al. 2008; Fisher and Owens 2004).

So, in the context of modelling extinction risk, what implications do our results have? They suggest that, although the predictions of all three methods are consistent, there are important differences between them. The strong predictive ability of PCM and their ability to reduce the influence of clade specific relationships in the data, provide support for the use of PCM techniques as the mainstay of future efforts model extinction risk. However, the strong predictive precision of DTs, in conjunction with their ability to highlight important interactions within the data, suggests that they can also play an important role in this field. We suggest that DTs may be used as a first step in identifying terms for inclusion in a suitable method of PCM in future efforts to predict species susceptibility to extinction.

**Acknowledgments** The authors would like to thank Andrew King, Amber Teacher and two anonymous reviewers for useful comments on the manuscript. This work was conducted thanks to NERC studentship NER/S/A/2004/12987.

## References

- Araujo MB, Whittaker RJ, Ladle RJ, Erhard M (2005) Reducing uncertainty in projections of extinction risk from climate change. *Glob Ecol Biogeogr* 14:529–538
- Bennett PM, Owens IPF (1997) Variation in extinction risk among birds: chance or evolutionary predisposition? *Proc R Soc Lond B Biol Sci* 264:401–408
- Bielby J, Cunningham AA, Purvis A (2006) Taxonomic selectivity in amphibians: ignorance, geography or biology? *Anim Conserv* 9:135–143
- Bielby J, Cooper N, Cunningham AA, Garner TWJ, Purvis A (2008) Predicting rapid declines in the world's frogs. *Conser Lett* 2:82–90
- Bininda-Emonds ORP, Cardillo M, Jones KE, MacPhee RDE, Beck RMD, Grenyer R, Price SA et al (2007) The delayed rise of present-day mammals. *Nature* 446:507–512
- Breiman L, Friedman JH, Olshen RA, Stone CG (1984) Classification and regression trees. Wadsworth International Group., Belmont, California
- Cardillo M, Purvis A, Sechrest W, Gittleman JL, Bielby J, Mace GM (2004) Human population density and extinction risk in the world's carnivores. *PLoS Biol* 2:909–914
- Cardillo M, Mace GM, Jones KE, Bielby J, Bininda-Emonds ORP, Sechrest W, Orme CDL et al (2005a) Multiple causes of high extinction risk in large mammal species. *Science* 309:1239–1241
- Cardillo M, Mace GMM, Purvis A (2005b) Response: problems of studying extinction risks. *Science* 310:1277–1278
- Cardillo M, Mace GM, Gittleman JL, Purvis A (2006) Latent extinction risk and future battlegrounds of mammal conservation. *Proc Natl Acad Sci USA* 103:4157–4161
- Cardillo M, Mace GM, Gittleman JL, Jones KE, Bielby J, Purvis A (2008) The predictability of extinction: biological and external correlates of decline in mammals. In: *Proceedings of the royal society B-biological sciences* (in review)
- Collen B, Bykova E, Ling S, Milner-Gulland EJ, Purvis A (2006) Extinction risk: a comparative analysis of central Asian vertebrates. *Biodivers Conserv* 15:1859–1871
- Cooper N, Bielby J, Thomas G, Purvis A (2008) Macroecology and extinction risk correlates of frogs. *Glob Ecol Biogeogr* 17:211–221
- Crawley MJ (2002) Statistical computing—an introduction to data analysis using S-plus. Wiley, Chichester
- De'ath G, Fabricius KE (2000) Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology* 81:3178–3192
- Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ et al (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129–151

- Felsenstein J (1985) Phylogenies and the comparative method. *Am Nat* 125:1–15
- Fisher DO, Owens IPF (2004) The comparative method in conservation biology. *Trends Ecol Evol* 19:391–398
- Freckleton RP, Harvey PH, Pagel M (2002) Phylogenetic analysis and comparative data: a test and review of evidence. *Am Nat* 160:712–726
- Frost DR, Grant T, Faivovich J, Bain RH, Hass A, Haddad CFB, De Sa RO et al (2006) The amphibian tree of life. *Bull Am Mus Nat Hist* 297:1–370
- Garland T, Harvey PH, Ives AR (1992) Procedures for the analysis of comparative data using phylogenetically independent contrasts. *Syst Biol* 41:18–32
- Gittleman JL, Luh HK (1992) On comparing comparative methods. *Annu Rev Ecol Syst* 23:383–404
- Grafen A (1989) The phylogenetic regression. *Philos Trans R Soc Lond B Biol Sci* 326:119–157
- Halsey LG, Butler PJ, Blackburn TM (2006) A phylogenetic analysis of the allometry of diving. *Am Nat* 167:276–287
- Harmon LJ, Losos JB (2005) The effect of intraspecific sample size on Type I and Type II error rates in comparative studies. *Evolution Int J org Evolution* 59:2705–2710
- Harvey PH, Pagel M (1991) *The comparative method in evolutionary biology*. Oxford University Press, Oxford
- Harvey PH, Rambaut A (1998) Phylogenetic extinction rates and comparative methodology. *Proc R Soc Lond B Biol Sci* 265:1691–1696
- IUCN (2004) 2004 IUCN Red List of Threatened Species
- Ives AR, Midford PE, Garland T (2007) Within-species variation and measurement error in phylogenetic comparative methods. *Syst Biol* 56:252–270
- Jones MJ, Fielding A, Sullivan M (2006) Analysing extinction risk in parrots using decision trees. *Biol Conserv* 15:1993–2007
- Jones KE, Bielby J, Cardillo M, Fritz SA, O'Dell J, Orme CDL, Safi K, Sechrest W, Boakes EH, Carbone C, Connolly C, Cutts MJ, Foster JK, Grenyer R, Habib M, Plaster CA, Price SA, Rigby EA, Rist J, Teacher A, Bininda-Emonds ORP, Gittleman JL, Mace GM, Purvis A (2009) PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* 90:2648
- Koh LP, Sodhi NS, Brook BW (2004) Ecological proneness of extinction proneness in tropical butterflies. *Conserv Biol* 18:1571–1578
- Laurance WF (1991) Ecological correlates of extinction proneness in Australian tropical rain forest mammals. *Conserv Biol* 5:79–89
- Little RJA, Rubin DB (2002) *Statistical analysis with missing data*. Wiley, Hoboken
- McKinney ML (1997) Extinction vulnerability and selectivity: combining ecological and paleontological views. *Annu Rev Ecol Syst* 28:495–516
- Owens IPF, Bennett PM (2000) Ecological basis of extinction risk in birds: habitat loss versus human persecution and introduced predators. *Proc Natl Acad Sci USA* 97:12144–12148
- Pagel M (1993) Seeking the evolutionary regression coefficient—an analysis of what comparative methods measure. *J Theor Biol* 164:191–205
- Paradis E, Claude J (2002) Analysis of comparative data using generalized estimating equations. *J Theor Biol* 218:175–185
- Purvis A (2008) Phylogenetic approaches to the study of extinction. *Ann Rev Ecol Evol Syst* 39:301–319
- Purvis A, Webster AJ (1999) Phylogenetically independent comparisons and primate phylogeny. In: Lee PC (ed) *Comparative primate socioecology*. Cambridge University Press, Cambridge, pp 44–70
- Purvis A, Agapow PM, Gittleman JL, Mace GM (2000a) Nonrandom extinction and the loss of evolutionary history. *Science* 288:328–330
- Purvis A, Gittleman JL, Cowlishaw G, Mace GM (2000b) Predicting extinction risk in declining species. *Proc R Soc Lond B Biol Sci* 267:1947–1952
- Purvis A, Cardillo M, Grenyer R, Collen B (2005) Correlates of extinction risk: phylogeny, biology, threat and scale. In: Purvis A, Brooks TM, Gittleman JL (eds) *Phylogeny and conservation*. Cambridge University Press, Cambridge
- Putland D (2005) Problems of studying extinction risks. *Science* 310:1277
- Quader S, Isvaran K, Hale RE, Miner BG, Seavy NE (2004) Nonlinear relationships and phylogenetically independent contrasts. *J Evol Biol* 17:709–715
- R Development Core Team (2009) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN: 3-900051-07-0, <http://www.R-project.org>
- Reed RN, Shine R (2002) Lying in wait for extinction: ecological correlates of conservation status among Australian Elapid Snakes. *Conserv Biol* 16:451–461

- Ricklefs RE, Starck JM (1996) Applications of phylogenetically independent contrasts: a mixed progress report. *Oikos* 77:167–172
- Ridley M (1983) The explanation of organic diversity: the comparative method and adaptations for mating. Oxford University Press, Oxford
- Russell GJ, Brooks TM, McKinney MM, Anderson CG (1998) Present and future taxonomic selectivity in bird and mammal extinctions. *Conserv Biol* 12:1365–1376
- Stuart SN, Chanson JS, Cox NA, Young BE, Rodrigues ASL, Fischman DL, Waller RW (2004) Status and trends of amphibian declines and extinctions worldwide. *Science* 306:1783–1786
- Stuart-Fox D, Moussalli A, Whiting MJ (2007) Natural selection on social signals: signal efficacy and the evolution of chameleon display coloration. *Am Nat* 170:916–930
- Sullivan MS, Gilbert F, Rotheray G, Croasdale S, Jones M (2000) Comparative analyses of correlates of Red data book status: a case study using European hoverflies (Diptera:Syrphidae). *Anim Conserv* 3:91–95
- Sullivan M, Jones M, Lee DC, Marsden SJ, Fielding AH, Young EV (2006) A comparison of predictive methods in extinction risk studies: contrasts and decision trees. *Biol Conserv* 15:1977–1991
- Sutherland WJ (2006) Predicting the ecological consequences of environmental change: a review of the methods. *J Appl Ecol* 43:599–616
- Symonds MRE (2002) The effects of topological inaccuracy in evolutionary trees on the phylogenetic comparative. *Syst Biol* 51:541–555