Author's accepted manuscript for:

*Di Marco, M., Collen, B., Rondinini, C., & Mace, G. (2015). Historical drivers of extinction risk: using past evidence to direct future monitoring. Proceedings of the Royal Society B, 282, 20150928.*

# Historical drivers of extinction risk: using past evidence to direct future monitoring

*Di Marco, Moreno[abc1], Ben Collen[d], Carlo Rondinini[a] and Georgina M. Mace[d]*

[a]Global Mammal Assessment program, Department of Biology and Biotechnologies, Sapienza Università di Roma, Viale dell' Università 32, I-00185 Rome, Italy.

[b]ARC Centre of Excellence for Environmental Decisions, Centre for Biodiversity and Conservation Science, The University of Queensland, 4072 Brisbane, Queensland, Australia

[c]School of Geography, Planning and Environmental Management, University of Queensland, 4072 Brisbane, Queensland, Australia

[d]Centre for Biodiversity & Environment Research, Department of Genetics, Evolution & Environment, University College London, Gower Street, London WC1E 6BT UK.

**Corresponding author**

[1] Centre for Biodiversity and Conservation Science, Goddard Building, The University of Queensland, St LUCIA, QLD 4072 (Brisbane, Australia), email to moreno.dimarco@gmail.com

## Summary

Global commitments to halt biodiversity decline mean that it is essential to monitor species' extinction risk. However the work required to assess extinction risk is intensive. We demonstrate an alternative approach to monitoring extinction risk, based on the response of species to external conditions. Using retrospective IUCN Red List assessments, we classify transitions in the extinction risk of 497 mammalian carnivores and ungulates between 1975-2013. Species that moved to lower Red List categories, or remained Least Concern, were classified as "lower risk"; species that stayed in a threatened category, or moved to a higher category of risk, were classified as "higher risk". Twenty-four predictor variables were used to predict transitions, including intrinsic traits (species biology) and external conditions (human pressure, distribution state, conservation interventions). The model correctly classified up to 90% of all transitions and revealed complex interactions between variables, e.g. protected areas *vs* human impact. The most important predictors were: past extinction risk, protected area extent, geographical range size, body size, taxonomic family, human impact. Our results suggest that monitoring a targeted set of metrics, would efficiently identify species facing a higher risk, and could guide the allocation of resources between monitoring species' extinction risk and monitoring external conditions.

## Keywords

biodiversity; conservation; human threats; mammals; random forest model;

## Introduction

Despite a growing international commitment to conservation, the current biodiversity crisis is characterised by increasing human pressures and continuing decline in the status of many species and habitats [1]. Reversing this trend has become the aim of one of the ambitious Aichi biodiversity targets proposed for 2020 [2]: reducing the extinction risk of known threatened species. If this target is achieved, it will in turn have a positive synergistic effect other targets (such as the protection of forests and the maintenance of carbon stocks [3]). Progress towards meeting this global biodiversity target relies on monitoring the extinction risk of species. Over recent decades, the International Union for Conservation of Nature (IUCN) has assessed the extinction risk of more than 70,000 species of plants, vertebrates and invertebrates on the Red List of Threatened species [4]. The classification of threatened species is clearly an effective conservation tool [5], with the IUCN Red List underpinning both international policy processes [2] and research aimed at improving conservation responses [6].

However, classifying and monitoring species' extinction risk requires intensive expert effort and considerable financial resources, which is unsustainable without change in either the strategy for assessment or funding [7]. Approaches such as sampling of taxa can be used to provide short-cuts, but it remains a substantial task [8]. Overall statistics from the IUCN Red List are used for measuring the status and trends of biodiversity [1,6] and for designing global-scale strategies for conservation interventions [9]. In addition, species-specific assessments inform direct actions to address particular threats at specific times and sites, requiring a comprehensive species-level approach [10].

The extinction risk of species, assessed using the IUCN Red List criteria [11], is a consequence of their biological traits, past and current environmental conditions, direct human pressures and the interactions between these factors [12,13]. Environmental changes and pressures on species are increasing in intensity and are the main cause of current increases in extinction rates. Extinction risk modelling has been used to better represent and quantify these external drivers, which can change and intensify over a short timeframe [14,15]. Biological traits by contrast change very slowly, and determine the way in which species respond to external pressures [13]. Historical information on species' extinction risk, and the way in which risk has changed in response to known pressures, could therefore be a good way to predict future biodiversity trends, particularly when the pressures can be effectively monitored or forecast.

Di Marco *et al.* undertook a retrospective assessment of the extinction risk of the world's carnivores and ungulates between 1975 and 2008 [16] by applying the current IUCN Red List criteria [11] to historical information. Studying past trends in extinction risk can indicate the circumstances under which conservation policies and strategies are or are not successful. Retrospective assessments can also guide the interpretation of future scenarios of emerging threats, for example, inferring the likely consequences of land use change or climate change [17]. Therefore, one approach to reducing the logistical and financial constraints of constant extinction risk monitoring could be to use well-validated models, based on past trends, to predict the effect of changing external pressures on future extinction risk [18,19].

In many cases Red List categories remain stable over long periods of time, especially for the large number of species listed as Least Concern (LC) [11]. The most useful information therefore concerns those species whose extinction risk is likely to escalate. We use historical records to develop and refine models of change in extinction risk, to identify those species for which high-risk combinations of biological vulnerability and extrinsic threats occur. We use current [4] and historical [16] information on Red List categories for 497 species of mammalian carnivores and ungulates in the period 1975-2013, to represent "transitions" in species' extinction risk (Fig. 1). We classified species in two groups: "lower risk" transitions, for those species not facing a significant increase in their extinction risk over time, and "higher risk" transitions, for those species facing a significant increase in their extinction risk over time (see Methods and Table S1). This approach is not analogous to measuring ordinal transitions between Red List categories (e.g. [20]), since we deliberately highlight species that will be of greatest concern to conservation, namely those that remain at a relatively high risk of extinction over time, and those that move from lower to higher risk categories.

We acknowledge that our study species are not a representative subset of all mammals, let alone life on earth. For example, carnivores and ungulates are generally characterised by longer generation times [21] and higher risk of extinction [4] relative to other mammals. Nonetheless the high conservation attention devoted to these groups makes a perfect case for testing our analytical approach.

We predicted higher and lower extinction risk transitions for species, using a comprehensive set of variables, which represent the conditions faced by the species during the study period. Our analyses therefore mimic a hypothetical situation in which relevant biological datasets and reliable forecast environmental and conservation metrics were available in the 1970s. This would have enabled conservation planners to predict which species would be in a higher or lower risk condition over the next 40 years.

## Methods

### Obtaining extinction risk transitions

We included all species of carnivores (Carnivora), ungulates (Perissodactyla and terrestrial Cetartiodactyla) and Proboscidea (discussed below together with ungulates) currently assessed in the IUCN Red List [4]. We excluded those species identified as being historically (<1970) extinct or Data Deficient (DD). We also excluded the Saudi gazelle (*Gazella saudyia*), declared extinct in the 1980s, since we had no detailed information available for its life history traits (apart from body mass) or spatial distribution. We considered 497 species in our analyses, representing 93% of all extant species in the study groups.

We compared the most recent species' extinction risk categories assessed in the IUCN Red List [11] with a retrospective assessment for 1975 [16]. We calculated an extinction risk transition value for each species between the two time periods in terms of the number of Red List categories changed (Fig. 1). A negative transition (<0) characterised species that moved toward a lower category of risk, a stable transition (=0) characterised species that maintained the same Red List status, and a positive transition (>0) characterised species that moved toward a higher category of risk.

We considered changes in species' extinction risk over a *c.* forty-year period (1975-2013). This is a reasonable reference period for species in our study groups, as it corresponds to >10 generations for small carnivores and ~2 generations for large bodied species such as elephants and rhinos [21].

**Classifying extinction risk transitions**

Because we were most interested in species that had fared unusually badly compared to those following an average trend over the study period, we identified species with a transition value significantly higher than random, when compared to other species within the same original extinction risk category. To do this we: (i) randomly re-assigned the observed transitions across all species within each original Red List category; (ii) compared the observed transitions with the randomly assigned transitions; (iii) repeated the previous steps 10,000 times. As an example, the transition of a species moving from LC (in 1975) to NT (in 2013) was higher than a transition randomly selected from other originally LC species in ~85% of the comparisons. Species with a transition value higher than random in ≤ 5% of the comparisons were included in the "lower risk" group. Species with a transition value higher than random in > 5% of the comparisons were included in the "higher risk" group. Importantly, a species retaining the same category over the time period (net change = 0) may have a transition value higher than random if several other species in the same original category had moved to lower categories of risk (net change < 0).

The randomization resulted in two groups containing species characterised by different extinction risk trajectories (Table S1). The "lower risk" group included species that were LC throughout the study period, together with species that underwent a change from any category to a lower category of risk. The "higher risk" group included all species that underwent a change from any category to a higher category of risk, together with species that were originally threatened or near threatened and retained their category. This classification reflects the intrinsic properties of the Red List criteria, in particular the fact that remaining within the same Red List category has different implications depending upon the category. For example, a species classified as LC throughout the time period does not face any significant decline over time. In contrast, a species classified as Vulnerable (VU) throughout the time period faces a strong continuing decline in abundance (≥30%) and/or remains at a very low population size. The species in the latter case therefore has a much higher probability of extinction (≥ 10% in 100 years) [11].

**Modelling the drivers of extinction risk transition**

We modelled the probability that a species is included in the higher risk or in the lower risk group, based on its original extinction risk category and the conditions in place over the study period. Extinction risk has been shown previously to be attributable to a combination of intrinsic and extrinsic factors [13]. Following recent work [22], our model included three classes of external predictor variables and one class of intrinsic (biological) predictors (see Table 1 for a complete list and description). The external variables are intended to reflect conditions faced by the species during the study period. We measured: i) distribution state variables, such as species' range size (measured in orders of magnitude); ii) human pressure variables, such as the human influence index [23]; and iii) conservation response variables, measured as the proportional coverage and absolute extent of protected area (PAs) within species ranges (again the extent was measured as an order of magnitude). The fourth group of predictor variables reflects species life-history traits (i.e. species biology) including physical characteristics (e.g. body-size), reproductive timing (e.g. weaning age) and reproductive output (e.g. weight at birth) [24]. We used an existing dataset [25], in which multiple imputation techniques had been used to fill gaps in life-history data [26].

Obtaining measures of external predictor variables corresponding to exactly the same years as the assessment period was not always possible. Nonetheless most of these data refer to the second half of the study period (i.e. ≥ 1990s), where the highest decline in species status was observed [16]. We assumed that changes that occurred within a relevant part of the 40-year study period (especially the second half of the period) would serve as a valid approximation for the entire period. In addition, this reduces the risk of collinearity between predictor variables (including levels of habitat loss and other proxies of human pressure) and original threat status (derived from retrospective assessments of extinction risk in the 1960s-1970s). We decided to not include variables that could not reasonably be used as predictors of future extinction risk change. For example, measures related to species distribution such as biogeographical realm - while probably acting as a proxy for regional pressure levels - could not reasonably be used by conservation planners to predict future changes in extinction risk of species.

We used Random Forest modelling (RF) to estimate the probability that a species was included in the higher risk or in the lower risk group. RF modelling is a powerful tool for ecological analysis [27], and it has been successfully used to model extinction risk in mammals [28,29] and amphibians [30]. RF is a machine learning technique with a number of characteristics that make it suitable for extinction risk prediction [15], including: limited assumptions about data distributions, high classification stability and performance, and ability to cope with collinear predictors. In a recent test, RF showed the highest performance in predicting global mammal extinction risk among several machine learning methods [29]. Our model included several variables which are external to species biology (human pressures, habitat state, conservation responses), hence, in common with other studies [15], we did not include phylogenetic constrains into our analyses. However we tested whether this could influence our results by independently examining the effect of including taxonomy for predicting extinction risk [29].

We ran a full RF model, including all predictor variables, and ranked the variables according to their relative importance, i.e. their contribution to model's classification accuracy. Variable importance, as well as the classification accuracy of the model, were calculated using an automated bootstrapped cross-validation procedure (implemented within the RF routine). During each iteration of the RF model, one third of the data were left out and used to cross-validate the classification ability of the model, see [31] for additional details.

Based on the final variable importance scores, we ran a series of partial RF models, each time including one additional variable following the variables' ranked importance. First we ran the model including only the most important variable, then added the second most important variable and re-ran the model, and so on until the last variable was included. We measured the performance of each partial RF model in terms of: proportion of correctly classified species (PCC), proportion of correctly classified higher risk species (sensitivity), proportion of correctly classified lower risk species (specificity), True Skill Statistic (TSS = sensitivity + specificity -1) [32].

In order to account for the effect of including the original (1975) species Red List status in the model, we re-ran the full model after removing this variable. Because of its potential role in Red List assessments and its representation of past threat conditions [33], we also re-ran the model after removing species' range size (RangeSize). In this latter case, we also removed the variable representing extent of PA within the species range (RangeProtkm), as it has a weak positive correlation with range size ($R^2 = 0.56$). We used degraded values of both range size and PA extent, i.e. order of magnitude rather than actual values (as for previous work [33]), to better represent the availability of coarse and approximate information during the study period. Finally, we built a single conditional inference classification tree to visually represent the interaction between predictor variables.

We adopted alternative classifications of extinction risk transitions and tested the performance of our model under different formats of the response variable. First, we repeated our RF modelling using ordinal changes in Red List categories as a numeric response variable (e.g. +2 for a species moving from LC to VU; see also [20]). Second, we repeated our RF modelling after removing all species that did not change their Red List category between 1975-2013; in this case we classified the remaining species in two categories: "uplisted" for species moving to higher extinction risk categories and "downlisted", for species moving to categories of lower risk. Third, we divided species in three groups: "LC to LC", including species remaining LC throughout the study period; "downlisted", including species that underwent a downlisting in their Red List category; "higher risk", following original classification already described.

The quantification of spatial variables was performed in GRASS GIS [34]. Statistical analyses were performed in R [35] using the packages 'randomForests' [31] and 'party' [36].

## Results

Our classification of extinction risk resulted in 277 species being included in the lower risk group (55% of all species) and 220 species in the higher risk group (45% of species). The full RF model for classification of higher risk *vs* lower risk species performed well in cross-validation (Table 2): 89% of all species were correctly classified, with a sensitivity of 0.84, and a specificity of 0.93 (TSS = 0.77). After removing the Red List category in 1975 from the model (i.e. the most important predictor), 82% of the species were still correctly classified, but the ability to correctly classify higher risk transitions was reduced (sensitivity = 0.78; TSS = 0.64). Subsequent removal of range size caused further deterioration in the model performance; although 79% of species were still correctly classified, there was a substantial reduction in sensitivity and TSS (sensitivity = 0.73; TSS = 0.57).

The six most important variables in the full RF model were: Red List category in 1975, PA extent (representing conservation response), range size (representing distribution state), body size (representing biology), family (representing taxonomy) and human impact index (representing human pressure) (Fig. 2A). A sequence of partial RF models, adding one variable at a time from the most important to the least important, showed that some of the variables had a contrasting effect on sensitivity and specificity. For example adding the taxonomic family to the model substantially increased sensitivity, but reduced specificity. In contrast, adding the human influence index slightly increased both sensitivity and specificity.

The extinction risk transition of 87% of species could be correctly predicted from one variable alone (Red List category in 1975), highlighting the importance of knowing the initial condition when modelling changes in extinction risk. However this was biased toward lower risk species (specificity = 0.95 *vs* sensitivity = 0.78). Adding five additional variables did not substantially alter the overall classification ability, but improved the balance between specificity and sensitivity (Fig. 2A). Even after removing the Red List categories in 1975 from the model, the performance remained fairly good, but then several variables had to be included in order to correctly classify ~78% of the higher risk and ~86% of the lower risk species (Fig. 2B). Subsequent removal of range size required the use of >50% of all variables to achieve a sensitivity of ~73% and specificity of ~83% (Fig. S1).

A single conditional inference tree (Fig. 3), represents the interplay between correlates of extinction risk transitions. For example, species that were LC in 1975 had a much higher probability of being in the higher risk group if they had a relatively low coverage of PAs during the study period (<1,000 $km^2$) and faced a substantial increase in human population density within their range (> 30%).

When changes in Red List categories were used as an ordinal numeric response variable, the following values were observed: -3 (n=1 species), -2 (n=3), -1 (n=11), 0 (n=369), +1 (n=79), +2 (n=23), +3 (n=9), +4 (n=2). In this case the RF regression model performed poorly in terms of total variance explained (13%). The relative importance of variables in determining model performance was also different with respect to the importance measured in the transition classification model, with the 6 most important variables now being: forest cover change, family, human population change, generation length, age at first birth, proportion of protected areas (Fig S2).

When excluding species that did not undergo a change in their Red List category, our sample reduced to 15 down-listed and 113 up-listed species. The RF model then gave highly biased results in this case, due to the high class imbalance, and classified all species as being uplisted (i.e. a complete imbalance toward sensitivity). The overall classification accuracy in this case was misleadingly high (88%), as the model was unable to predict improvement in species conservation status.

When dividing species into three groups, there were 15 downlisted species, 262 LC to LC species and 220 higher risk species. Here again, the overall classification accuracy of the model was high (89%), but the predictive ability for the downlisted class was very low (only 1 correct prediction, Table S2).

**Discussion**

By focusing on extinction risk transitions, we were able to distinguish between two groups of species. The higher risk group included species that remained at high extinction risk and those whose extinction risk increased between 1970 and 2010. The lower risk group included species that remained at, or improved their status to, low extinction risk during the same period. This classification is different from the Red List status, since it identifies species that are undergoing an unusual increase in extinction risk compared to other species that started the period in the same risk category.

We included candidate predictor variables from a range of classes (see Methods) and found that a small number of variables (from different classes) can efficiently predict the extinction risk transition of ungulates and carnivores. These variables have been highlighted previously [13,28] and include initial conservation status, certain biological traits (represented by body mass), levels of human encroachment, and the degree of conservation action (represented by PA coverage). The importance of considering conservation interventions in extinction risk modelling has already been demonstrated for Australian birds [20] and for African mammals [22], and we confirm it here in a global scale analysis.

Our results show that the probability of a species being at higher risk was reduced by some adequate level of PAs coverage (one thousand $km^2$ or more; Fig. 3), while it was increased by limited PA coverage and high levels of human pressure. To a first approximation this indicates the conditions under which PAs deliver positive conservation outcomes [37]. Monitoring the progress of PA expansion and the extent of human encroachment within species ranges can therefore be strategic. Future projections of these variables may be translated into global projection of species extinction risk, and allow for a proactive planning of conservation interventions [38].

Our models  included measures of environmental change (e.g. the amount of suitable habitat for a species during the study period) and static measures of human impact (e.g. human influence index). These classes of variables were both important predictors in our model. Among general proxies of human pressures and habitat state, we also included information on levels of tree cover and tree cover change (see also [22]). While the role of these variables is probably more influential for forest-dependent than for non-forest species, it is known that habitat clearance has a contagious effect [39] and we use tree cover, a well mapped habitat feature at a global scale [40], to estimate the general condition of natural habitats within species ranges.

The extinction risk transition model performed well in cross validations, the classification ability was high for both lower risk and higher risk species. The availability of a dataset with retrospective extinction risk assessments [16] made it possible for us to validate our extinction risk model. This type of validation is common in other environmental science areas, and has been used to validate models of climate change effects on species distribution [41]. As our knowledge of past extinction risk improves, this approach could become standard practice in extinction risk modelling.

Unlike many previous studies, we did not convert IUCN Red List categories into numerical measures of extinction risk (e.g. LC to Extinct, from 0 to 5; [20,42]), or use extinction risk probabilities described in Red List Criterion E [43]. These involve assumptions about the relationship between categories and probability of extinctions that are not supported in theory or in practice [11]. We simply assumed that species in the higher risk group have higher conservation requirements than those in the lower risk group, and found that predicting ordinal changes in Red List categories (as in [20]) was substantially less efficient than predicting extinction risk transitions. We also found that excluding those species with no change in their Red List category, or assigning stable LC species to a separate group, resulted in a biased allocation of model error with downlisted species being systematically misclassified. In this case the model is unable to predict the outcome of conservation success, i.e. those situations in which the extinction risk of a species is reduced over years.

Our results on the relative importance of different predictor variables can be used to identify priorities for future data gathering. We suggest that monitoring a set of such variables over time would allow conservationists to effectively anticipate future extinction risk. The accuracy of these predictions will rest on the assumption that these variables represent the drivers of transitions in species extinction risk. Our results demonstrates that this was the case for past extinction risk transitions, but the emergence (or the exacerbation) of new threats (such as climate change) would need to be accounted for to have a robust forecasting of extinction risk [17,44]. However, this is not a weakness unique to our approach: threats to biodiversity change over time [45] and any model used to forecast extinction risk would require continuing updates and recalibration to account for emerging threats. Monitoring the emergence of new threats and the occurrence of rapid changes in external conditions will be necessary, yet even this would probably be easier than continuously assessing the extinction risk category of all species.

McCarthy et al. [20] investigated optimal investment strategies to prevent the extinction and minimise the number of threatened Australian birds, using conservation investments to model the probability of species moving between Red List categories. A similar approach could be combined with our modelling framework here, to measure the probability of undergoing a high risk transition. In this case the probability can be modelled as a function of the intrinsic and extrinsic conditions in place for the species, plus the conservation budget available. However, adequate information on global conservation expenditure for threatened species needs to be available to reliably model the relationship between investments and status change.

Our approach can provide guidance on how to allocate resources among monitoring of species extinction risk and monitoring of external conditions, it can inform the identification of key variables to be monitored. There is great potential for the application of our approach to other taxa, especially considering the increasing availability of retrospective extinction risk assessments for groups such as amphibians [46] and corals [47], and the potential to use historical information to perform retrospective assessments for other groups [16].

**Author contributions**

MDM and GMM conceived the study design; MDM performed the analyses; all authors interpreted the results, contributed to the writing, and approved the final version of the paper.

**Funding statement**

This work was supported by Fondazione Brusarosco & Società Italiana di Ecologia, through a scholarship to MDM (Premio Brusarosco 2013).

**Data accessibility statement**

Data used in this paper comes from published sources which have been appropriately cited in the Methods section.

**Acknowledgements**

We thank four anonymous reviewers for their comments. We thank Piero Visconti for providing data on habitat loss.

## References


1 Butchart, S.H.M. et al. 2010 Global Biodiversity: Indicators of Recent Declines. *Science* **328**, 1164–1168.

2 CBD 2010 *Strategic plan for Biodiversity 2011-2020*. Montreal, QC.

3 Di Marco, M., Butchart, S. H. M., Visconti, P., Buchanan, G. M., Ficetola, G. F. & Rondinini, C. 2015 Synergies and trade-offs in achieving global biodiversity targets. *Conserv. Biol,* doi:10.1111/cobi.12559

4 IUCN 2013 *IUCN Red List of Threatened Species*. Version 2013.2 (*www.iucnredlist.org*).

5 Rodrigues, A. S. L., Pilgrim, J. D., Lamoreux, J. F., Hoffmann, M. & Brooks, T. M. 2006 The value of the IUCN Red List for conservation. *Trends Ecol. Evol.* **21**, 71–76.

6 Hoffmann, M. et al. 2010 The impact of conservation on the status of the world's vertebrates. *Science* **330**, 1503–9.

7 Rondinini, C., Di Marco, M., Visconti, P., Butchart, S. & Boitani, L. 2014 Update or outdate: long term viability of the IUCN Red List. *Conserv. Lett.* **2**, 126–130.

8 Baillie, J. E. M. et al. 2008 Toward monitoring global biodiversity. *Conserv. Lett.* **1**, 18–26.

9 Rodrigues, A. S. L. et al. 2004 Global gap analysis: priority regions for expanding the global protected-area network. *Bioscience* **54**, 1092–1100.

10 Hayward, M. W. 2011 Using the IUCN Red List to determine effective conservation strategies. *Biodivers. Conserv.* **20**, 2563–2573.

11 IUCN 2001 *IUCN Red list categories and criteria, version 3.1.* IUCN Gland, Switzerland and Cambridge, UK.

12 Mace, G. M., Collar, N. J., Gaston, K. J., Hilton-Taylor, C., Akçakaya, H. R., Leader-Williams, N., Milner-Gulland, E. J. & Stuart, S. N. 2008 Quantification of extinction risk: IUCN's system for classifying threatened species. *Conserv. Biol.* **22**, 1424–1442.

13 Cardillo, M., Mace, G. M., Gittleman, J. L., Jones, K. E., Bielby, J. & Purvis, A. 2008 The predictability of extinction: biological and external correlates of decline in mammals. *Proc. R. Soc. B* **275**, 1441–1448.

14 Di Marco, M., Rondinini, C., Boitani, L. & Murray, K. A. 2013 Comparing multiple species distribution proxies and different quantifications of the human footprint map, implications for conservation. *Biol. Conserv.* **165**, 203–211.

15 Murray, K. A., Verde Arregoitia, L. D., Davidson, A., Di Marco, M. & Di Fonzo, M. M. I. 2014 Threat to the point: improving the value of comparative extinction risk analysis for conservation action. *Glob. Chang. Biol.* **20**, 483–494.

16 Di Marco, M., Boitani, L, Mallon, D., Hoffmann, M., Iacucci, A., Meijaard, E., Visconti, P., Schipper, J., Rondinini, C. 2014 A retrospective evaluation of the global decline of carnivores and ungulates. *Conserv. Biol.* **28**, 1109-1118.

17 Visconti, P. et al. 2015 Projecting global biodiversity indicators under future development scenarios. *Conserv. Lett.*, doi:10.1111/conl.2014.12159

18 Cardillo, M., Mace, G. M., Gittleman, J. L. & Purvis, A. 2006 Latent extinction risk and the future battlegrounds of mammal conservation. *Proc. Natl. Acad. Sci.* **103**, 4157–4161.

19 Davies, T. J. et al. 2008 Phylogenetic trees and the future of mammalian biodiversity. *Proc. Natl. Acad. Sci.* **105**, 11556–11563.

20 McCarthy, M.A., Thompson, C. J. & Garnett, S. T. 2008 Optimal investment in conservation of species. *J. Appl. Ecol.* **45**, 1428–1435.

21 Pacifici, M., Santini, L., Di Marco, M., Baisero, D., Francucci, L., Grottolo Marasini, G., Visconti, P. & Rondinini, C. 2013 Generation length for mammals. *Nat. Conserv.* **5**, 89–94.

22 Di Marco, M., Buchanan, G. M., Szantoi, Z., Holmgren, M., Grottolo Marasini, G., Gross, D., Tranquilli, S., Boitani, L. & Rondinini, C. 2014 Drivers of extinction risk in African mammals: the interplay of distribution state, human pressure, conservation response and species biology. *Philos. Trans. R. Soc. B* **369**, 20130198.

23 Sanderson, E. W., Jaiteh, M., Levy, M. A., Redford, K. H., Wannebo, A. V. & Woolmer, G. 2002 The human footprint and the last of the wild. *Bioscience* **52**, 891–904.

24 Bielby, J., Mace, G. M., Bininda-Emonds, O. R. P., Cardillo, M., Gittleman, J. L., Jones, K. E., Orme, C. D. L. & Purvis, A. 2007 The fast-slow continuum in mammalian life history: an empirical reevaluation. *Am. Nat.* **169**, 748–757.

25 Di Marco, M., Cardillo, M., Possingham, H. P., Wilson, K. A., Blomberg, S. P., Boitani, L. & Rondinini, C. 2012 A novel approach for global mammal extinction risk reduction. *Conserv. Lett.* **5**, 134–141.

26 Penone, C., Davidson, A., Shoemaker, K., Di Marco, M., Rondinini, C., Brooks, T., Young, B., Graham, C. & Costa, G. 2014 Imputation of missing data in life-history trait datasets: which approach performs the best? *Meth. Ecol. Evol.* **5**, 961–970.

27 Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J. & Lawler, J. J. 2007 Random forests for classification in ecology. *Ecology* **88**, 2783–2792.

28 Davidson, A. D., Hamilton, M. J., Boyer, A. G., Brown, J. H. & Ceballos, G. 2009 Multiple ecological pathways to extinction in mammals. *Proc. Natl. Acad. Sci.* **106**, 10702–10705.

29 Bland, L. M., Collen, B., Orme, C. D. L. & Bielby, J. 2015 Predicting the Conservation Status of Data-Deficient Species. *Conserv. Biol*. **29**, 250-259.

30 Murray, K.A., Rosauer, D., McCallum, H. & Skerratt, L. F. 2011 Integrating species traits with extrinsic threats: closing the gap between predicting and preventing species declines. *Proc. R. Soc. B* **278**, 1515–1523.

31 Liaw, A. & Wiener, M. 2002 The randomforest package. *R News* **2**, 18–22.

32 Allouche, O., Tsoar, A. & Kadmon, R. 2006 Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* **43**, 1223–1232.

33 Di Marco, M. & Santini, L. 2015 Human pressures predict species' geographic range size better than biological traits. *Glob. Chang. Biol*. **21**, 2169–2178.

34 GRASS Development Team 2010 *Open source GIS.Software, Version 6.4.0.* Open Source Geospatial Foundation. (*http://grass.osgeo.org*)

35 R Development CoreTeam 2011 *R: A language and environment for statistical computing.* Vienna, Austria. ( *http://www.R-project.org*)

36 Strobl, C., Hothorn, T. & Zeileis, A. 2009 Party on. *R Journ.* **1**, 14–17.

37 Geldmann, J., Barnes, M., Coad, L., Craigie, I. D., Hockings, M. & Burgess, N. D. 2013 Effectiveness of terrestrial protected areas in reducing habitat loss and population declines. *Biol. Conserv.* **161**, 230–238.

38 Rondinini, C. et al. 2011 Reconciling global mammal prioritization schemes into a strategy. *Philos. Trans. R. Soc. Lond. B.* **366**, 2722–2728.

39 Boakes, E. H., Mace, G. M., McGowan, P. J. K. & Fuller, R. a 2010 Extreme contagion in global habitat clearance. *Proc. R. Soc. B* **277**, 1081–1085.

40 Hansen, M. C. et al. 2013 High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **342**, 850–853.

41 Martínez-Meyer, E., Peterson, A. T. & Hargrove, W. W. 2004 Ecological niches as stable distributional constraints on mammal species, with implications for Pleistocene extinctions and climate change projections for biodiversity. *Glob. Ecol. Biogeogr.* **13**, 305–314.

42 Purvis, A., Gittleman, J. L., Cowlishaw, G. & Mace, G. M. 2000 Predicting extinction risk in declining species. *Proc. R. Soc. B* **267**, 1947–1952.

43 Mooers, A. Ø., Faith, D. P. & Maddison, W. P. 2008 Converting endangered species categories to probabilities of extinction for phylogenetic conservation prioritization. *PLoS One* **3**, e3700.

44 Mantyka-Pringle, C. S., Visconti, P., Di Marco, M., Martin, T. G., Rondinini, C. & Rhodes, J. R. 2015 Climate change modifies risk of global biodiversity loss due to land-cover change. *Biol. Conserv.* **187**, 103–111.

45 Mace, G. M. 2010 *Drivers of biodiversity change*. In *Trade-offs in Conservation: Deciding What to Save* (eds N. Leader-Williams R. Smith & W. Adams), pp. 349–364. Wiley-Blackwell Publishing Ltd.

46 Stuart, S. N., Chanson, J. S., Cox, N. a, Young, B. E., Rodrigues, A. S. L., Fischman, D. L. & Waller, R. W. 2004 Status and trends of amphibian declines and extinctions worldwide. *Science* **306**, 1783–1786.

47 Carpenter, K. E. et al. 2008 One-third of reef-building corals face elevated extinction risk from climate change and local impacts. *Science* **321**, 560–563.

48 Nelson, A. 2008 *Travel time to major cities: A global map of Accessibility*. Glob. Environ. Monit. Unit - Jt. Res. Cent. Ispra Italy (*http://bioval.jrc.ec.europa.eu/products/gam/index.htm*)

49 Alkemade, R., Oorschot, M., Miles, L., Nellemann, C., Bakkenes, M. & ten Brink, B. 2009 GLOBIO3: A Framework to Investigate Options for Reducing Global Terrestrial Biodiversity Loss. *Ecosystems* **12**, 374–390.

50 Visconti, P. et al. 2011 Future hotspots of terrestrial mammal loss. *Philos. Trans. R. Soc. Lond. B*. **366**, 2693–702.

51 Rondinini, C. et al. 2011 Global habitat suitability models of terrestrial mammals. *Philos. Trans. R. Soc. B* **366**, 2633–2641.

52 WCS & CIESIN 2005 *Last of the Wild Project, Version 2 (LWP-2): Global Human Footprint Dataset*. (http://sedac.ciesin.columbia.edu/data/set/wildareas-v2-human-footprint-geographic)

53 CIESIN & CIAT 2005 *Gridded Population of the World, Version 3 (GPWv3): Population Density Grid*. Palisades, NY NASA Socioecon. Data Appl. Cent.

54 CIESIN, FAO & CIAT 2005 *Gridded Population of the World, Version 3 (GPWv3): Population Count Grid, Future Estimates*. Palisades, NY NASA Socioecon. Data Appl. Cent.

55 Jones, K. E. et al. 2009 PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* **90**, 2648.

56 UNEP-WCMC 2013 *The World Database on Protected Areas (WDPA):* Cambridge, UK. (http://www.wdpa.org/)

**Table 1** Description of the variables used in the model. Variables are organised in different classes: human pressure (P), species biology (B), distribution state (D), conservation response (R). Examples of previous use of the variables for predicting extinction risk in terrestrial mammals, and the original data sources for each variables are also provided.

| Class | Variable | Description and justification | Examples | Source |
|---|---|---|---|---|
| - | Dependent variable | Extinction risk transition as described in Table S1. | | [4,16] |
| - | RL75 | Red List category in 1975, representing original species status (i.e. extinction risk at the beginning of the study period). | | [16] |
| P | Acc_50 | Travel distance from major cities (accessibility), measured as the median value of the variable within species ranges (percentiles tested: 5, 10, 20, 50). A proxy of human encroachment. | [22,29] | [48] |
| P | AOOloss | Proportional loss of suitable habitat within species ranges (1970-2010). A proxy of the main driver of mammal species decline calculated from back casts of global land cover changes, from the IMAGE integrated assessment model [49]. | [22] | [50,51] |
| P | HII_5 | Human influence index, measured as the proportion of species ranges where the variable had values larger than 5 (values tested: 5, 10, 20). A proxy of the human impact on the environment. | [22,29] | [52] |
| P | HPD90_50 | Human population density in 1990, measured as the median value the variable within species ranges (percentiles tested: 5, 10, 20, 50). A proxy of human encroachment, | [13,22,29] | [53] |
| P | PopChange | Proportional change in human population count in 1990-2010, measured as the mean value observed within species range. | | [54] |
| P | ForestCG | Proportional change in forested habitat within species ranges between 2000-2012. A proxy of natural habitat loss. | | [40] |
| B | AFB_d | Age at first birth | [24,25] | [55] |
| B | BirthW | Birth weight | [22] | [55] |
| B | BodySize | Body mass | [13,28,29] | [55] |
| B | DietBrdth | Number of dietary categories eaten by the species | [22] | [55] |
| B | InterbInt | Interbirth interval | [24] | [55] |
| B | LitPY | Litters per year | | [55] |
| B | LitSiz | Litter size | [22,24,29] | [55] |
| B | WeanAge | Weaning age | [13,24] | [55] |
| B | Fam | Taxonomic family | | [4] |
| B | Ord | Taxonomic order | [13,22] | [4] |
| B | GenLen | Generation length | [24] | [21] |
| B | HabBrdth | Number of habitat layers used by each species. | | [55] |
| D | TreeCov_50 | Median tree cover within species range in 2000 (percentiles measured: 5, 10, 20, 50). A proxy of forests state. | | [40] |
| D | Hab | Species habitat preferences, classified as: forest, grassland, shrubland, bareland, coastal or generalist (when >1 of the previous applied). | | [51] |
| D | RangeSize | Species range size, measured as an order of magnitude (e.g. 1 for ranges of 10-100 $km^2$, 2 for ranges of 100-1000 $km^2$, etc.). | [13,22,28] | [4] |
| R | RangeProt_prop | Proportion of species range covered by protected areas with an IUCN category I to IV. | [22] | [56] |
| R | RangeProtkm | Extent of protected areas within species ranges, measured as an order of magnitude (as described for "RangeSize") | | [56] |

**Table 2** Performance of the random forest models. The full model is compared with partial models, where the original species status (RL75) and the range size (RangeSize) were removed.

| Metric | Full model | RL75 removed | RL75 and RangeSize removed* |
|---|---|---|---|
| PCC† | 0.89 | 0.82 | 0.79 |
| Sensitivity | 0.84 | 0.78 | 0.73 |
| Specificity | 0.93 | 0.86 | 0.84 |
| TSS‡ | 0.77 | 0.64 | 0.57 |

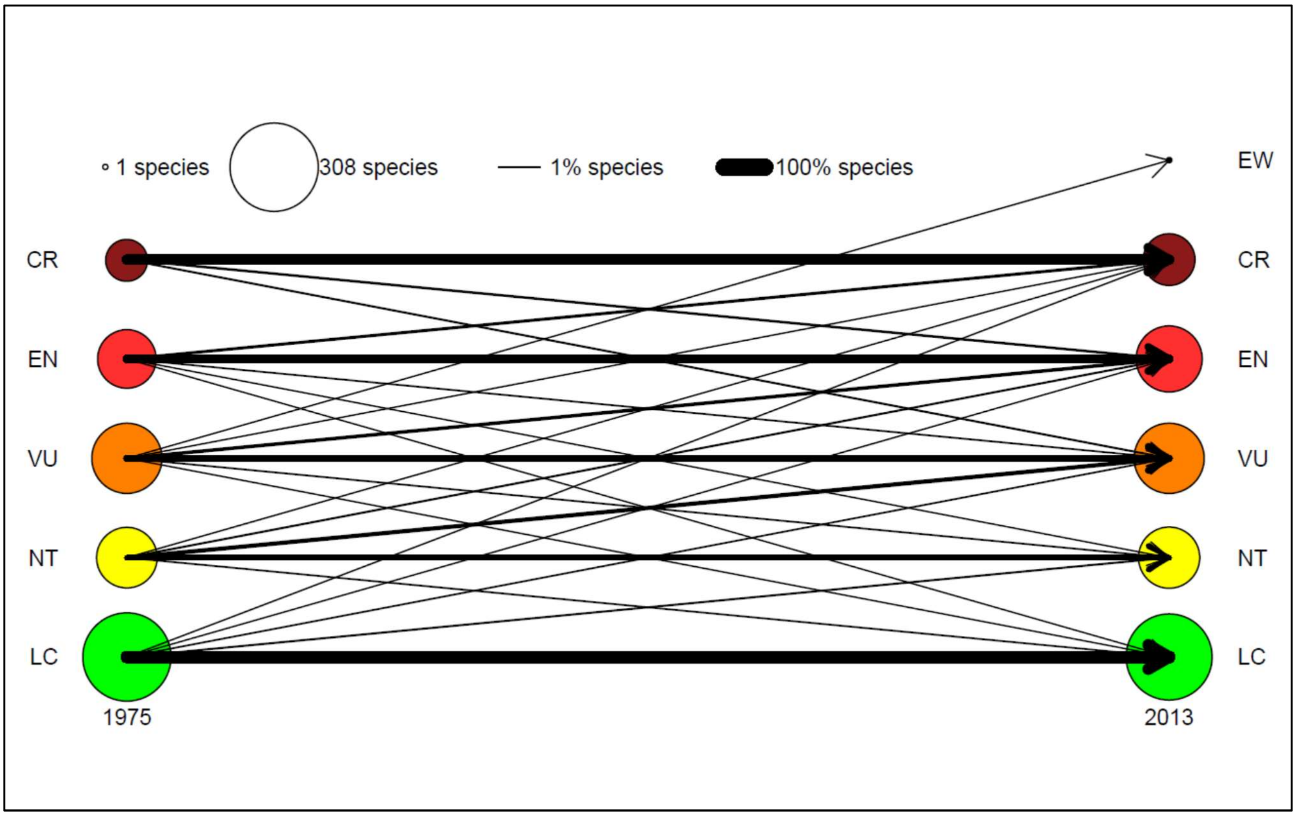
*When removing the variable RangeSize the extent of protected areas within the range was also removed, to avoid a potential surrogate effect.
†PCC, proportion of correctly classified species.
‡TSS, true skill statistics.

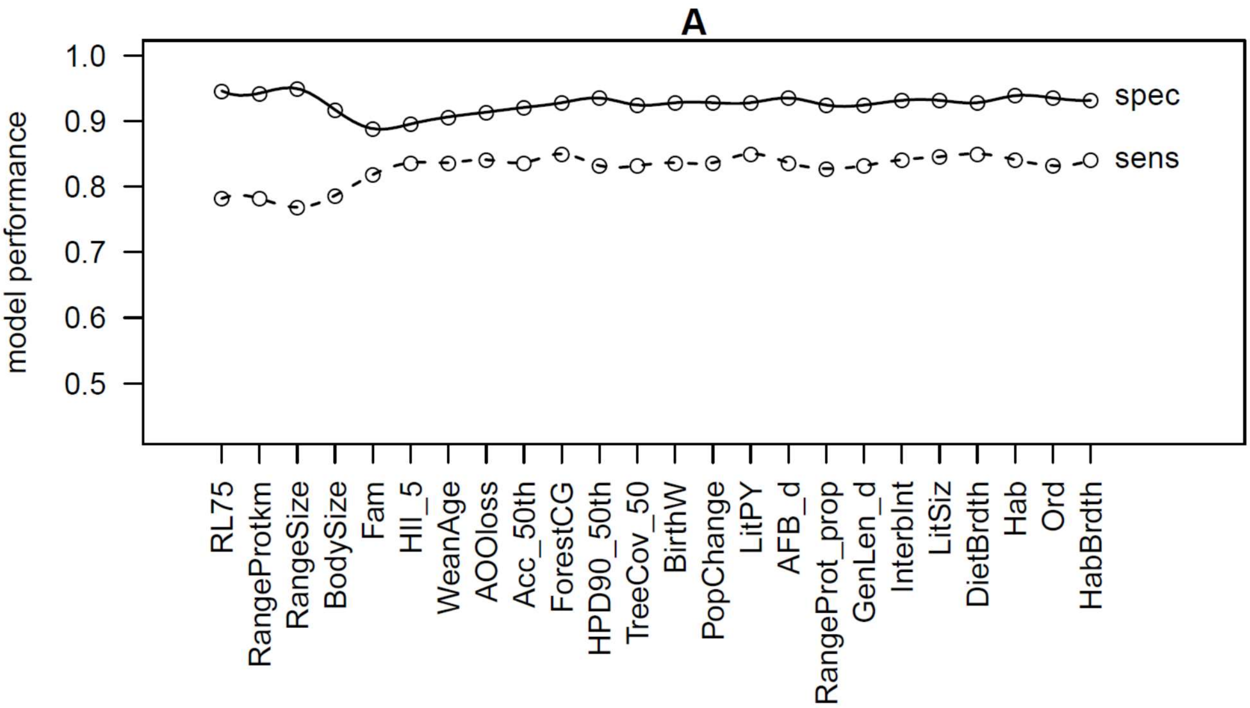
**Figure legends**

**Fig. 1** Transition of species' extinction risk categories in the period 1975-2013. The plot reports the number of species (carnivores and ungulates) in each Red List category for each time period. Circles' size is proportional to the number of species while arrows represent the proportion of species moving from an initial category to a final category (arrows' width scales with the proportion of species in the original category). Data were obtained from [4,16].

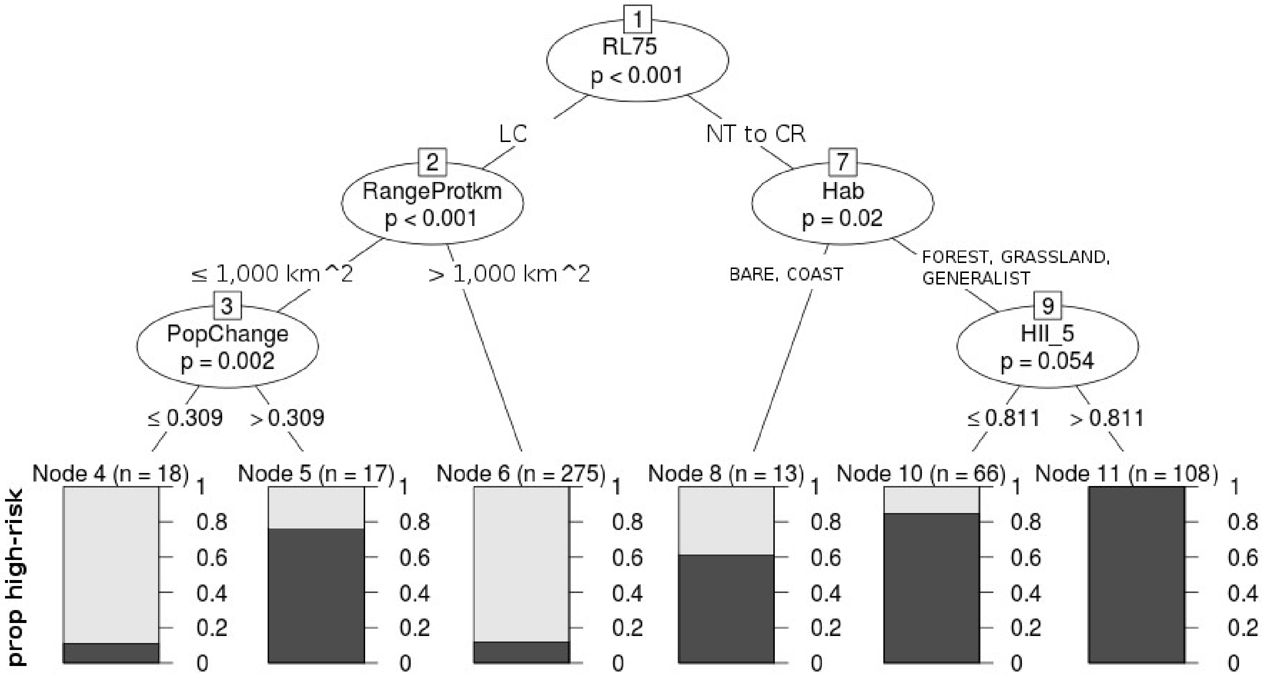
**Fig. 2** Performance of extinction risk models with an increasing number of variables, considering all variables (A) or all variables apart from original status (B). Variables are added iteratively to the models, from left to right according to their ranked importance in the original full model. Each series of symbols (y-axis) represents the specificity (spec) or sensitivity (sens) of a model that included the variables on its left or below it (x axis).

**Fig. 3** Conditional inference classification tree for extinction risk transition. Each terminal node reports (in dark grey) the proportion of higher risk species. See Table 1 for a description of the variables.


1 species
308 species
1% species
100% species
EW
CR
EN
VU
NT
LC
1975
2013


A
model performance
1.0
0.9
0.8
0.7
0.6
0.5
spec
sens
RL75
RangeProtkm
RangeSize
BodySize
Fam
HII_5
WeanAge
AOOloss
Acc_50th
ForestCG
HPD90_50th
TreeCov_50
BirthW
PopChange
LitPY
AFB_d
RangeProt_prop
GenLen_d
InterbInt
LitSiz
DietBrdth
Hab
Ord
HabBrdth

532


1
RL75
p < 0.001
LC
NT to CR
2
RangeProtkm
p < 0.001
7
Hab
p = 0.02
≤ 1,000 km^2
> 1,000 km^2
BARE, COAST
FOREST, GRASSLAND, GENERALIST
3
PopChange
p = 0.002
9
HII_5
p = 0.054
≤ 0.309
> 0.309
≤ 0.811
> 0.811
Node 4 (n = 18)
Node 5 (n = 17)
Node 6 (n = 275)
Node 8 (n = 13)
Node 10 (n = 66)
Node 11 (n = 108)
prop high-risk
1
0.8
0.6
0.4
0.2
0