Predicting predator-prey interactions in terrestrial endotherms using random forest

John Llewelyn[1,2], Giovanni Strona[3], Christopher R. Dickman[4], Aaron C. Greenville[5,6,7], Glenda M. Wardle[5], Michael S. Y. Lee[8,9], Seamus Doherty[1,2], Farzin Shabani[1,2,10], Frédérik Saltré[1,2], Corey J. A. Bradshaw[1,2]

[1]Global Ecology | *Partuyarta Ngadluku Wardli Kuu*, College of Science and Engineering, Flinders University, GPO Box 2100, Adelaide, South Australia 5001, Australia

[2]ARC Centre of Excellence for Australian Biodiversity and Heritage. EpicAustralia.org.au

[3]Faculty of Biological and Environmental Sciences, University of Helsinki, Viikinkaari 1, Biocentre 3, 00790, Helsinki, Finland

[4]Desert Ecology Research Group, School of Life and Environmental Sciences, University of Sydney, New South Wales 2006, Australia

[5]School of Life and Environmental Sciences, University of Sydney, Sydney, Australia

[6]School of Life Sciences, University of Technology Sydney, Sydney, Australia

[7]National Environmental Science Program, Threatened Species Recovery Hub, University of Sydney, Sydney, Australia

[8]College of Science and Engineering, Flinders University, GPO Box 2100, Adelaide, SA 5001, Australia

[9]Earth Sciences Section, South Australian Museum, North Terrace, Adelaide, SA 5000, Australia

[10]Department of Biological and Environmental Sciences, College of Arts and Sciences, Qatar University, PO Box 2713, Doha, Qatar

Abstract

1. Species interactions play a fundamental role in ecosystems. They affect where species can live, how their population sizes fluctuate through time, and how environmental perturbations cascade through communities. But few ecological communities have

complete data describing such interactions, which is an obstacle to understanding how ecosystems function and respond to environmental perturbations. Because it is often impractical to collect empirical data for all potential interactions in a community, various methods have been developed to infer interactions. Random forest machine learning is emerging as one of the most accurate and frequently used methods for making interaction predictions, but its performance in inferring predator-prey interactions in terrestrial vertebrates remains untested. The sensitivity of random forest performance to variation in quality of training data is also unclear.

2. We examined predator-prey interactions within and between two diverse, primarily terrestrial vertebrate classes: birds and mammals. Combining data from a global interaction dataset and a specific ecological community (Simpson Desert, Australia), we tested how well random forests predict predator-prey interactions for mammals and birds using species' ecomorphological and phylogenetic traits. We also tested how variation in training data quality affected model performance by: removing records and switching interaction records to non-interactions (false non-interactions) in the entire training dataset, or restricting these changes to records involving focal prey, focal predators, or non-focal species (focal = species for which predictions are being made).
3. We found that random forests could predict predator-prey interactions for birds and mammals using either ecomorphological or phylogenetic traits, and that these predictions were accurate even when there were no records in the training data for the focal predator or focal prey species. In contrast, false non-interactions for focal predators in the training data strongly degraded model performance.
4. Our results demonstrate that random forests can be used to identify predator-prey interactions for bird and mammal species that have few or no trophic interaction

records. Furthermore, our study provides a roadmap for predicting interactions using random forests which might help ecologists: (*i*) address knowledge gaps and explore network-related questions in data-poor situations and (*ii*) predict interactions for species invading new ecosystems.

Keywords species interactions; training data; supervised learning; ecological network; trophic interactions; food web; Simpson Desert; GloBI

Data availability Data and script available from GitHub (DOI 10.5281/zenodo.7037189).

Introduction

Species interactions are crucial to maintaining biodiversity (Thompson, 1999) because they determine how energy and nutrients flow among organisms (Allan et al., 2021; Traill et al., 2010) and affect the distribution, survival and abundance of species (Van der Putten et al., 2010). Species interactions also determine how ecological disturbances — such as those involving invasive species, habitat modification, and climate change — reverberate through ecological communities (Blois et al., 2013; Tylianakis et al., 2008). This is because environmental changes that directly alter one or a subset of species in a community can have flow-on effects for other species via interaction networks (Kaneryd et al., 2012; Morton et al., 2022). Given the need to predict and mitigate human-driven biodiversity declines, interest in modelling interaction networks for biological communities is increasing (Bohan et al., 2017; Strydom et al., 2021). However, most species interactions in most ecological communities have not been documented (Caron et al., 2022; Hortal et al., 2015; Poisot, 2022). Indeed, it is usually logistically impossible to measure how all species interact in a community — even when only one type of interaction (e.g., predator-prey) is considered (Brousseau et al., 2018;

Strona, 2022). In addition to the challenge of identifying long-established interactions, the rapid and widespread translocation of species across the planet due to migratory responses to climate change and movement by humans means we also need to identify interactions that could occur between species that have not previously had the opportunity to interact (Smith & Phillips, 2006; Valdovinos et al., 2018).

To overcome gaps in species interaction data, various methods of inferring interactions using species traits/proxies can be applied, such as quantile-regression trophic niche-space modelling (Gravel et al., 2013; Llewelyn et al., 2021), Bayesian linear modelling (Caron et al., 2022), and several machine-learning approaches (e.g., *K* nearest-neighbour—(Desjardins-Proulx et al., 2017; Pichler et al., 2020), neural networks/deep learning methods—(Fricke et al., 2022) and random forests—(Sydenham et al., 2021)). Of these, random forest machine learning is emerging as one of the most popular methods due to its ability to learn nonlinear interactions between predictor variables and outperform other inference methods (Desjardins-Proulx et al., 2017; Laigle et al., 2018; Poisot, 2022; Sydenham et al., 2021). Indeed, random forests have been used to predict various interaction types, including plant-pollinator (Ornai & Keasar, 2020; Sydenham et al., 2021), parasite-host (Kotula et al., 2020), and predator-prey interactions (Desjardins-Proulx et al., 2017; Parravicini et al., 2020). They have also been used to identify which species' traits are most important in determining whether species interact (Dellinger et al., 2019; Klomberg et al., 2022). Using training data consisting of interacting and non-interacting species, random forests can learn how ecomorphological traits, phylogenetic position, and/or taxonomy (as a proxy for phylogenetic position; Laigle et al., 2018) determine which species interact. However, questions remain regarding the ability of random forests to predict different types of interactions in different taxonomic groups, which species' data are best for inferring interactions, and how to fine-tune training datasets (Poisot, 2022).

Quantifying predator-prey interactions is central to community ecology because these interactions influence both top-down and bottom-up processes (Sinclair et al., 2003). Random forests have been applied to predict such interactions, but these applications have largely been limited to aquatic communities (Parravicini et al., 2020; Strona et al., 2021) and soil invertebrate (Desjardins-Proulx et al., 2017; Laigle et al., 2018). Indeed, the performance of random forests for inferring predator-prey interactions among terrestrial vertebrates — a group that includes many species of high conservation concern (Cox et al., 2022; Tilman et al., 2017) and of important ecological function (e.g., ecological engineers; Samson & Knopf, 1996) — has not been assessed.

If we are to model terrestrial ecosystems realistically to predict and minimise the risk of extinction cascades, reliable inference of trophic interactions is essential. However, terrestrial vertebrates are diverse structurally, phylogenetically, and functionally, potentially making the 'rules' determining interactions complex and difficult for random forests to identify (whereas just one trait — body size — largely determines trophic interactions in fish due to their gape limitation; Gravel et al., 2013; Morales-Castilla et al., 2015). Furthermore, interaction data for many terrestrial vertebrate species have been recorded opportunistically or not at all (although detailed diet data do exist for some species; Pringle & Hutchinson, 2020), potentially leading to sparse (i.e., missing most prey for most predators) and taxonomically biased datasets for training and testing random forests. Having sparse interaction records for predators in training data is problematic because random forests need to be trained with data that include both interacting (e.g., predator and prey) and non-interacting species pairs, with non-interacting species pairs usually assigned by assuming that undocumented interactions imply no interaction (i.e., pseudo-non-interactions generated from presence-only data; Pichler et al., 2020). When applied to incomplete (sparse) datasets, this process can identify undocumented but actually interacting species pairs as non-interacting pairs (but see Strona et

al., 2021). Importantly, such 'false non-interactions' in training data likely reduce the reliability of interactions inferred from random forests. Similarly, random forests trained on datasets that are taxonomically biased may perform poorly when making predictions for species not represented in the training data (Poisot et al., 2021).

We designed analyses to test whether random forest machine learning could accurately predict predator-prey interactions within and between two diverse terrestrial vertebrate classes — birds and mammals — using both a global (*GloBI*) and a single-community (Simpson Desert, Australia) predator-prey interaction dataset. We also assessed the importance of ecomorphological and phylogenetic traits for inference accuracy, as well as the sensitivity of model performance to the quality of the training data. To test sensitivity to the quality of training data, we compared two scenarios that could occur as a result of sparse training data: (1) a reduction in the number of interaction and non-interaction records used to train the models, and (2) a reduction in the number of interaction records and an increase in the number of false non-interaction records used to train the models — and we investigated how these changes affected model performance when they were restricted to different components of the training dataset (i.e., testing the effect of taxonomic bias on performance).

Our results confirm that: (*i*) random forests can predict interactions in birds and mammals, (*ii*) models trained on interaction datasets including either ecomorphological traits or phylogenetic position perform similarly well, (*iii*) models can identify the prey of predators that have no interaction records in the training data, and (*iv*) model performance is especially sensitive to false non-interactions in the training data that involve records of predatory species for which predictions are being made (i.e., focal predators). Our results highlight the unrealised potential for the application of random forests for predicting interactions among terrestrial vertebrates. Furthermore, they show to which components of the training data

random forests are most sensitive, thereby identifying the most efficient way to increase accuracy via fine-tuning of training data.

Methods

We used two interaction datasets combined with ecomorphological and phylogenetic information to test the performance of random forests for predicting predator-prey interactions in birds and mammals (Fig. S1 shows flowchart of dataset preparation). We also tested how the performance of these models was influenced by the type of species data used and the size and quality of the training data. The two interaction datasets were: (1) a dataset from the ***Glo**bal **B**iotic **I**nteractions database* (*GloBI*; globalbioticinteractions.org) supplemented with interaction data from diet studies done in Australia (Llewelyn, 2022), and (2) an ecosystem-specific dataset from the Simpson Desert in Australia that focuses on seven predators for which detailed dietary information is available (Llewelyn, 2022). We focused on birds and mammals due to the availability of detailed trait (Wilman et al., 2014) and phylogenetic (vertlife.org) information for these taxa that provided a wide range of traits for predicting interactions.

Datasets

Global data

We extracted all predator-prey interactions involving birds and mammals from *GloBI* (accessed 29/05/2020 and 26/06/2020 for birds and mammals as consumers, respectively). However, *GloBI* contains few records of species found in Oceania (including Australia) relative to other regions of the world: of the 3329 unique species combinations in the records we extracted from *GloBI*, only 109 involved a predator and prey species both found in Oceania. Because we used the *GloBI* data to train random forests for predicting interactions

of predators in the Simpson Desert (an ecosystem in Oceania and therefore lacking related records in *GloBI*; see also next section), we added interactions involving Simpson Desert species and their relatives to the *GloBI* data (Fig. S1). We added data for these species by merging two additional datasets with *GloBI*: (*i*) predator-prey records from the Simpson Desert that did not involve the seven focal predators from the testing dataset (i.e., for which we made predictions), and (*ii*) interactions of the focal Simpson Desert predator species with non-Simpson Desert prey species elsewhere in Australia (i.e., from diet studies of the focal predators outside the Simpson Desert; no interactions between the focal predators and Simpson Desert species were included in the training data). In both cases, we built these datasets by aggregating predator-prey records from diet studies (Tables S1-2). We then combined these additional interaction datasets with the *GloBI* data to create an enhanced global interaction dataset (Llewelyn, 2022).

To predict species interactions using random forests, we required a training dataset that included interacting and non-interacting pairs of species. Like most interaction datasets, our enhanced dataset did not include non-interactions i.e., since relationships are based on presence-only data. We therefore generated pseudo-non-interactions ('non-interactions' henceforth) by taking all possible two-species combinations from the species in our enhanced global interaction dataset and removing those combinations documented as 'interacting' (Fig. S1). We also removed combinations where either the 'predator' does not consume vertebrates (e.g., strictly herbivorous species could not be potential predators) or the combination involved species not found on the same continent (using the rgbif package in R; Chamberlain et al., 2022), because such non-interactions will not inform model predictions regarding known predators and which sympatric species they prey upon (i.e., the type of interactions we aim to predict in the Simpson Desert).

Random forests use explanatory variables, such as species' traits, to learn and predict associations (Desjardins-Proulx et al., 2017). We therefore added trait data for predator and prey species to the interaction and non-interaction datasets (Fig. S1). These traits included: body mass, time of activity (whether the species displays nocturnal, diurnal, and/or crepuscular activity), ability to fly, resource groups consumed (plants, vertebrates, and/or invertebrates), and percentage use of different habitat strata (Wilman et al., 2014). We added phylogenetic information by calculating eigenvectors describing a species' position in the phylogeny (eigenvector mapping; Guénard et al., 2013), a potential proxy for explanatory variables which are phylogenetically conserved. We extracted phylogenies from `VertLife` (VertLife.org), and used the `MPSEM` (Guénard & Legendre, 2022) and `ape` (Paradis & Schliep, 2019, p. 0) R packages to handle phylogenetic data and extract eigenvectors.

We also identified which species pairs in the non-interaction data included a potential prey species within the predator's preferred prey-size range (Fig. S1). This was done so that the ratio of non-interaction records from inside *versus* outside the preferred size range included in training datasets could be adjusted to optimise performance of random forests (i.e., analogous to optimising species distribution models by adjusting the ratio of pseudo-absences sampled from among versus outside observed occurrences; Barbet-Massin et al., 2012). To identify a preferred size range, we log-transformed ($\log_e$) predator and prey body masses to linearise body-size relationships (Gravel et al., 2013). We then resampled the interaction dataset to create 100 training and testing dataset pairs using all 3692 interaction records in each of the 100 pairs (75:25 interaction records in training:testing datasets). To each training set we fitted a series of upper and lower quantile regressions capturing the middle 0.99 to 0.8 quantiles at 0.01 quantile increments. We included a prey mass × predator taxonomic class interaction in the regression models to allow different body-size relationships for predatory birds and mammals (Llewelyn et al., 2021). We then combined each of the 100 testing sets

with an equal number of species pairs randomly sampled from the non-interaction set (also with body mass log-transformed), and applied the quantile regression models (i.e., the upper and lower quantile regressions combinations) built in the previous step to the testing sets. Species combinations in the testing sets that included a potential prey species from within the upper and lower quantile ranges derived from the regressions were predicted as prey, and those outside these limits were excluded as prey. We assessed model performance by comparing predictions to documented interactions (i.e., did the quantile regression models correctly identify which combinations came from the interaction and non-interaction datasets?) and used these results to calculate the true skill statistic to identify the optimal quantile regression model for assigning predator-prey interactions. A true skill statistic = 1 indicates a model that perfectly predicts interactions and non-interactions, whereas when it is 0 the model does not perform better than assigning links at random. The model capturing the middle 0.84 quantiles performed best (true skill statistic = 0.26), so we used this model to classify species pairs in the non-interaction dataset as those that included a potential prey from within the predator's preferred size range and those that included a potential prey from outside the predator's preferred size range.

Simpson Desert testing data

It is unclear how the accuracy of random forests trained and tested on data including species from many different communities (such as was the case for our enhanced global datasets) translates to model performance when applied to single ecological communities (Parravicini et al., 2020). We therefore tested model performance when applied to seven common predators from the Simpson Desert, a community we chose because it has been intensively studied for over 30 years and has detailed information on species composition and the diets of some predators. We used the species documented at established field sites in the Simpson

Desert as the species assemblage (i.e., potential prey) in our analyses, rather than all species documented in the Simpson Desert bioregion, because the former includes sympatric species that have the potential to interact whereas the latter would likely include species that seldom have the opportunity to interact.

The Simpson Desert bird and mammal assemblage from these sites includes 64 vertebrate-consuming bird and mammal species (49 birds and 15 mammals). However, the diets of many of these species have not been studied in detail. When testing the performance of random forests for predicting predator-prey interactions, we therefore restricted the Simpson Desert testing dataset to seven predatory species with detailed dietary information from studies both in the Simpson Desert and other arid regions in Australia. These seven focal predators included four birds — *Aquila audax*, *Falco berigora*, *Hamirostra melanosternon*, *Tyto alba* — and three mammals — *Canis dingo*, *Felis catus*, *Vulpes vulpes*. *Felis catus* (domestic cats) and *Vulpes vulpes* (red fox) are recent introductions to Australia (since European colonisation in 1788), *Canis dingo* (dingo) was brought to Australia by people >3000 years ago (Fillios et al., 2012), and the remaining focal predators naturally occur in Australia.

To construct the dataset, we aggregated data from published and unpublished diet studies (excluding species consumed as carrion) of the focal predators in the Simpson Desert (Fig. S1; Table S1). We also included records from diet studies done outside the Simpson Desert involving species pairs found in the Simpson Desert (i.e., we assumed that if trophic interactions occurred between the species in other communities, they would also occur in the Simpson Desert; Table S1), and records involving the focal predators and Simpson Desert species in the enhanced global interaction dataset (these records were removed from the enhanced global dataset). None of the studies we used to build the Simpson Desert interaction dataset reported non-interactions. We therefore generated non-interactions by assuming any

species pairs involving the focal predator and the species recorded from the Simpson Desert field sites (157 species) that have not been documented as predator and prey, are not predator and prey. We added the same trait and phylogenetic information to the Simpson Desert data as described for the enhanced global dataset. We did not, however, include a column indicating if non-interacting species pairs in the Simpson Desert data involved potential prey from within the predator's preferred size range because the Simpson Desert dataset was only used for testing model performance (i.e., we did not need to optimise the sampling of these non-interactions because they were not used for training random forests).

Modelling

We used the `ranger` package (Wright & Ziegler, 2017) in R to apply random forests to our data. In the first section of our study, we optimised the random forest models and tested how performance was affected by the choice of variables included in the random forests. In the second section, we tested how changes to the training dataset—including removing records and increasing the number of false non-interactions—influenced model performance (Fig. S2 shows flowchart of modelling steps).

1. Model optimisation and variables

We used the enhanced global dataset and cross-validation to optimise model parameters and compare performance of models trained using: (*i*) only ecomorphological traits, (*ii*) only phylogenetic eigenvectors, or (*iii*) both types of data (Fig. S2). We also compared performance of models that incorporated many *versus* few explanatory variables/traits to test whether more variables resulted in higher accuracy or overfitting. There were 42 variables in the many-variable ecomorphological model (21 each for predator and prey; Table S3) because there were 21 ecomorphological variables in the databases we built. To be consistent,

301 we used 42 phylogenetic eigenvectors in the many-variable phylogenetic model (the first 21
302 phylogenetic eigenvectors for predator and prey). In the few-variable models, we included 10
303 variables: the most important 10 ecomorphological variables (as determined by permutation)
304 or the first 5 phylogenetic eigenvectors for predator and prey. The many- and few-variable
305 models that incorporated both types of data simply combined the variables from the
306 respective ecomorphological and phylogenetic models (i.e., either 84 or 20 variables in total).
307 To optimise models, we divided the enhanced global dataset's interactions (n = 3692) into
308 training and testing datasets (ratio: 75:25), and combined them with non-interaction data. The
309 number of non-interaction records added to the testing datasets equalled the number of
310 interaction records in that dataset (923 of each), whereas we varied the number of non-
311 interactions in the training datasets to identify the optimal ratio (while having 2769
312 interaction records in each training dataset). We replicated the process of generating training
313 and testing datasets each time a random forest was applied in the parameter optimization and
314 cross validation steps described below.
315 We weighted non-interactions (such that their combined weight was equal to that of the
316 observed interactions) because models built with weighted non-interactions performed
317 slightly better than those built using unweighted non-interactions (Table S4; similar to results
318 from species distribution modelling – Barbet-Massin et al., 2012). We also optimised the
319 random forests in terms of: (*i*) number of non-interactions in the training data, (*ii*) proportion
320 of non-interactions from within *versus* outside the predators' preferred size ranges (as
321 classified by quantile regression; see above), (*iii*) the probability threshold for assigning
322 interactions, (*iv*) and the number of variables randomly sampled as candidates at each split in
323 the trees. We optimised these parameters by iterating through values for each parameter
324 individually, running 40 random forests (each consisting of 800 trees) at each value (on 40
325 resampled training and testing datasets) and calculating the mean true skill statistic for

326 models at each value. We identified the parameter values that led to models with the highest
327 true skill statistic as optimal. We separately optimised the 6 models that differed in terms of
328 included variables (i.e., the many- *versus* few-variable ecomorphological, phylogenetic, and
329 combined data models; Table S5 shows optimised parameter sets). We then compared
330 performance of these optimised models by applying them to 100 training and testing datasets
331 generated using the enhanced global dataset cross validation design described above, and
332 calculated the true skill statistic in each of these 100 cases.

333 After comparing performance of the six models with the enhanced global dataset, we
334 tested their performance when trained on the enhanced global dataset and applied to the
335 Simpson Desert data (Fig. S2). Again, we trained and applied 100 random forests for each of
336 the six models, and assessed their performance with the true skill statistic. We resampled
337 non-interactions for the training dataset in these 100 analyses because there were more non-
338 interactions than needed, whereas we used all 3692 interaction records in the enhanced global
339 dataset in each of the 100 training datasets.

340 We did randomization tests to assess the probability the performance of models differed
341 from one another by taking 100,000 samples from the cumulative distribution function
342 describing the performance (true skill statistic) of each of the model types and comparing
343 overlap between these samples.

344

345 *2. Quality of training dataset and model performance*

346 To test the influence of the size of the training dataset, taxonomic coverage, and prevalence
347 of false non-interactions (true interactions classified as non-interactions) on model
348 performance, we reran the analyses with the optimised ecomorphological model trained on
349 the enhanced global dataset and applied to the Simpson Desert data, but varied the number of
350 records, taxonomic coverage, and prevalence of false non-interactions in the enhanced global

351 training datasets (Fig. S2). In these analyses, we used the many-variable ecomorphological
352 model because that model performed well in the previous analyses.
353 We artificially reduced the quality of training data in eight ways to test specific
354 hypotheses. We modified the entire training dataset by (*i*) randomly removing records and (*ii*)
355 randomly switching interaction records to non-interactions (false non-interactions). We
356 hypothesised that model performance would be sensitive to removal of training data because
357 having more records helps the random forests learn associations between species traits and
358 trophic links, and we hypothesised that models would be sensitive to false non-interactions
359 because random forests would learn the wrong association between variables when taught
360 with erroneous data. We also tested what impact these two types of changes (data removal
361 and false non-interactions) had when they were restricted to subsets of the training data that
362 included either: (*iii*, *iv*) focal prey, (*v*, *vi*) focal predators, or (*vii*, *viii*) interactions where
363 neither prey nor predators were focal/Simpson Desert species. We hypothesised that model
364 performance would be especially sensitive to changes involving focal predators and prey
365 because these records involve the species for which predictions are being made, i.e., the
366 associations between these species' traits and trophic interactions should be especially
367 important for the model to learn. Furthermore, we hypothesised that model performance
368 would be sensitive—though less so—to removing records and increasing the proportion of
369 false non-interactions in non-Simpson Desert (non-focal) interactions because these
370 interactions teach the random forests trait-trophic associations that also apply to Simpson
371 Desert species.
372 For each of the eight modified training datasets described above, we reduced the size of
373 the training dataset or switched interactions to non-interaction at 1% increments relative to
374 the initial number of records in the group being modified (e.g., 1% increments of the total
375 number of interactions involving the focal predator), generating 10 training datasets at each

increment. We trained the random forests on these modified datasets, applied them to the Simpson Desert data, and calculated true skill statistics 10 times at each increment. We plotted results to make visual comparison of the sensitivity of model performance to each of the eight types of artificial variation in the quality of training data. In addition to running the models 10 times at 1% increments, we ran the models 100 times when all the relevant data were modified for each type of training-data modification and calculated the true skill statistics, allowing us to quantify and compare the impacts of each type of modification on model performance.

We also examined the mechanisms underlying changes in model performance by testing how removing records or increasing false non-interactions for the focal predator influenced the: (*i*) relative probability (or suitability) of interacting assigned by the random forest to each potential prey for each predator (i.e., did changes to the training dataset affect which species were identified as the most likely prey?), and (*ii*) mean probability assigned to potential prey for each predator (i.e., did changes to the training dataset cause an overall increase or decrease in the probabilities assigned to potential prey?). If lower model performance reflects a systematic shift in all the probabilities (upward or downward), rather than a change in which potential prey are identified as the most likely, the models could still be used to predict a predator's prey (because the rank order of probability of different prey remains similar irrespective of absolute model performance and absolute probability).

To test for changes in relative prey suitability, we measured the correlation (Pearson r) between prey probabilities assigned by the modified training models with that assigned by models trained on the full training dataset (Fig. S2). To test for changes in mean probability assigned to potential prey overall, we subtracted the mean probability assigned to potential prey for each predator as calculated by the full training data models from that calculated by the modified training dataset models (Fig. S2). For both analyses, we modified the training

401 dataset at 1% increments (as described for the model performance analyses above; 10
402 replicate training datasets/models at each increment). We also ran the models 100 times with
403 all focal predators removed and with all focal predator interactions changed to non-
404 interactions to quantify the impact of each type of modification to training data on assigned
405 relative probabilities and mean assigned probabilities. We hypothesised that removing
406 records and including more false non-interactions in the training dataset would disrupt the
407 assignment of relative suitability and reduce assigned probabilities overall.
408

409 # Results

410 **Variables and model performance**

411 Random forests trained and tested with the enhanced global data had consistently high
412 accuracy, with mean true skill statistics ranging from 0.69 to 0.77 (Fig. 1a,b). However, of
413 these models, random forests trained on few ecomorphological variables performed worse
414 than the others (Fig. 1b), such that the probability of random forests trained on few
415 ecomorphological variables performing as well as models trained on few phylogenetic
416 eigenvectors or few of both types of data was 0.001 and 0.002, respectively.

417 The models trained on the enhanced global data and applied to the Simpson Desert data
418 yielded mean true skill statistics between 0.3 and 0.39 (Figs 1-2). In both the many- and few-
419 variable models, those that used ecomorphological variables performed best (Fig. 1a,b). For
420 models trained on many variables (Fig. 1a), the probability that the ecomorphological
421 model's performance was the same as that of the phylogenetic eigenevector or the both-data-
422 type models was 0.00004 and 0.00002, respectively. For models trained on few variables
423 (Fig. 1b), the probability that the ecomorphological model's performance was the same as
424 that of the phylogenetic eigenvector model and the both-data-types model was 0.017 and
425 0.023, respectively.

426 Using few *versus* many variables caused little or no change in performance of
427 phylogenetic eigenvector or both-data-type models (Fig. 1). However, the performance of
428 ecomorphological models when applied to the enhanced global data and the Simpson Desert
429 data tended to be slightly better when they included many variables (Fig. 1) — the
430 probability that ecomorphological models with many *versus* few variables had the same
431 performance was 0.009 (enhanced global) and 0.086 (Simpson Desert).
432 Using the model that performed best on the Simpson Desert data (i.e., the many
433 ecomorphological variable model; Fig. 1), performance predicting interaction for native
434 versus recently introduced predators was similar (true skill statistics of 0.39 and 0.40 for
435 native and introduced species, respectively). The probability that model performance on
436 native predators and introduced predators was the same was 0.32.
437
438 **Effects of training dataset size, false non-interactions, and taxonomic coverage**
439 We modified the enhanced global training dataset by randomly removing records and
440 changing interaction records to false non-interactions, both to the entire training dataset as
441 well as to three subsets: (*i*) records involving the focal prey (i.e., species found in the
442 Simpson Desert), (*ii*) records involving the focal predators (i.e., the 7 Simpson Desert
443 predators for which we predicted interactions), and (*iii*) records involving species not found
444 in the Simpson Desert. Removing records reduced performance when modifying either the
445 entire training dataset (Fig. 3a; true skill statistic reduced to 0 when all records removed) or
446 the focal-predator component only (Fig. 3e; true skill statistic reduced to 0.081 ± 0.019 [mean
447 ± standard deviation] when all focal-predator records removed). However, these effects
448 became pronounced only when removing > 80% of records (Fig. 3a,e), which equates to
449 removing > 14,768 or 854 records from the entire or focal-predator components of the
450 training datasets, respectively. Removing records involving the focal prey or non-Simpson

Desert species had only weak effects on model performance (Fig. 3c,g; true skill static reduced to 0.368 ± 0.016 and 0.349 ± 0.022 for removal of focal prey and non-Simpson Desert species, respectively).

The effect of false non-interactions on model performance was stronger than that of removing records. False non-interactions assigned at random to interaction records in either the entire training dataset or the component of the training dataset involving focal predators caused more rapid declines in model performance (Fig. 3b,f; true skill static reduced to 0 and 0.016 ± 0.008 for models with all records, or all records involving focal predators changed to non-interactions, respectively), while false non-interactions in the training data involving non-Simpson Desert species caused a moderate decline in model performance (Fig. 3h; true skill static reduced to 0.243 ± 0.020). These declines were triggered when > 40% of the interaction records in the relevant group were changed to false non-interactions (Fig. 3b,f,h), which equates to 1,477 records in the entire training dataset analysis, 235 in the focal predator analysis, and 1,122 in the non-focal species analysis. False non-interactions assigned to focal prey had little effect on model performance (Fig. 3d; true skill static reduced to 0.360 ± 0.018 when all focal prey interaction records changed to false non-interactions).

Focal-predator training data and the mechanisms underlying changes in model performance

The correlation between probabilities assigned to potential prey from the full *versus* record-removed models showed a gradual decrease as more interactions involving the focal predator were removed from the training data, and this decline became more rapid when removing > 80% of records (Fig. 4a). The correlation between these probabilities reduced to 0.67 ± 0.02 when all records involving focal predators were removed from the training dataset. The mean probability assigned to potential prey for each predator was also reduced by removing records

of the focal predator from the training data and, like the correlation results, this change was small until 80% of records were removed (Fig. 4c). After removing all records, the probabilities assigned to potential prey were on average 0.24 ± 0.01 lower compared to those assigned by models trained on the full dataset.

Both the correlation and change in mean probability assigned to potential prey were more sensitive to false non-interactions involving the focal predators than to removing their records from the training dataset, showing more rapid declines (Fig. 4b,d). When changing all interactions involving the focal predators in the training data to non-interactions, the correlation between probabilities assigned to prey decreased to 0.41 ± 0.01 (Fig. 4b) and the mean probabilities assigned to potential prey dropped by 0.38 ± 0.01 (Fig. 4d).

Discussion

Our results demonstrate that random forests can predict the prey of predatory birds and mammals even when few or no interaction records involving the focal predator are included in training data (Fig. 4a). Thus, our results indicate the value of applying random forests to predict predator-prey interactions for these predators (and terrestrial vertebrates more generally), supporting their broader application to facilitate network modelling of these systems (Brousseau et al., 2018). However, our analyses highlight some limitations, showing that random forests should be applied with caution depending on the quality, quantity, and type of training data available.

We found that random forests could predict bird and mammal interactions in both our global and ecosystem-specific datasets (including for native and recently introduced predators), and that models based on ecomorphological or phylogenetic traits performed similarly well (Fig. 1) — consistent with previous studies showing that either data type can be used for predicting interactions in other taxa (Brousseau et al., 2018; Desjardins-Proulx et

al., 2017; Elmasri et al., 2020; Faisal et al., 2010; Morales-Castilla et al., 2015; Pomeranz et al., 2019). Because phylogenetic information is available for most terrestrial vertebrates (e.g., VertLife.org), ecological/morphological trait databases are increasingly common and comprehensive (e.g., Herberstein et al., 2022; Jones et al., 2009; Oliveira et al., 2017), and interaction data for many species are readily available from databases (e.g., globalbioticinteractions.org; web-of-life.es) and publications, random forests based on such data can now easily be applied to predict predator-prey interactions for most vertebrate communities. Given the importance of biotic interactions for determining species' extinction/co-extinction vulnerabilities (Strona & Bradshaw, 2018), the ease of use of random forests coupled with their flexibility in terms of inferring different types of interactions between diverse taxa (Desjardins-Proulx et al., 2017; Laigle et al., 2018; Strona et al., 2021) highlights the potential of this method to aid in the ecological network and invasive species modelling required to guide conservation decisions.

Model performance was robust to the random removal of records from the training dataset in our analyses; in fact, we had to remove > 80% of records (>14,700 individual records) before an obvious change in performance occurred (Fig. 3a). Therefore, models that use large training datasets are unlikely to be sensitive to random additions or removal of interaction and non-interaction records.

Like the random removal of records, the removal of records involving focal predators only caused a prominent decline in model performance when > 80% (Fig. 3e) of records in that group were removed (which translates into fewer — 854 — records in total needing to be removed). Conversely, the complete removal of records involving focal-prey species or non-focal species had little impact on model performance (Fig. 3c,g). Thus, declines in model performance in the random removal analyses were largely driven by the removal of focal predator records. And this reduction in model performance following removal of focal-

526 predator records was primarily due to every prey species being assigned consistently lower
527 probabilities, rather than a shuffling of prey preference (Fig. 4a,c). Because *relative* prey
528 preferences were largely maintained even when all the training records involving the focal
529 predator were removed from training data (mean Pearson $r = 0.67$; Fig. 4a), our results
530 suggest prey species could be predicted for predators that do not have any interaction data —
531 as is likely the case for understudied, rare, or extinct species. To be able to predict trophic
532 interactions for extinct species accurately has important applications in understanding past
533 extinction events (Llewelyn et al., 2021; Pires et al., 2015; Ripple & Van Valkenburgh,
534 2010), while the capacity to predict interactions for rare species could help identify cryptic
535 factors that threaten persistence (Doherty et al., 2016; Smith & Phillips, 2006). However,
536 additional work to assess the requirements of training data for predicting interactions for
537 species without interaction records is still required. For example, it is unclear how
538 phylogenetically and/or ecologically similar predators in the training data need to be to the
539 focal predator for models to be able to predict prey preferences accurately.

540 In contrast to removal, false non-interactions in the training data reduced model
541 performance more rapidly, triggering lower performance when > 40% of the interaction
542 records were changed to non-interactions. This sensitivity highlights an important issue: most
543 studies generate non-interactions by assuming that undocumented interactions imply true
544 non-interaction; however, this assumption could generate many false non-interactions,
545 especially if interaction records are sparse (i.e., if many of the predators in the training data
546 have only a small proportion of their interactions documented). To avoid generating false
547 non-interactions, one can restrict non-interactions in training datasets to species combinations
548 where the predator is too small to eat the potential prey and/or the potential predator comes
549 from a low trophic level (Strona et al., 2021). However, applying this approach could reduce
550 the capacity of random forests to identify unsuitable prey from within a predator's preferred

prey-size range or for high trophic-level species. Indeed, the optimal ratio of non-interactions sampled from within *versus* outside preferred size range in the many-variable ecomorphological model (i.e., the model that performed best in the Simpson Desert analyses) was 1:1 (Table S5), suggesting that non-interactions of suitably-size potential prey should be included in training data. Species distribution models have been assessed for the analogous problem of false and uninformative pseudo-absences, with ensuing recommendations to avoid including such absences (e.g., assignment of pseudo-absences based on geographic and climatic exclusion; Mateo et al., 2010), further highlighting the need to consider carefully how pseudo-absences and pseudo-non-interactions are generated from presence-only data (Barbet-Massin et al., 2012; Iturbide et al., 2015; Lobo et al., 2010).

Model performance was especially sensitive to false non-interactions in the training data involving focal predators. This type of change in data quality disrupted relative prey preferences and lowered probabilities assigned to interactions more extensively than did reducing the number of records involving focal predators (Fig. 4b,c), reinforcing the importance of avoiding false non-interactions when generating pseudo-non-interactions for training data; otherwise, such data can change which species are identified as most-likely prey. Although restricting the generation of non-interactions based on relative size and trophic level is one method of avoiding false non-interactions (discussed above: Strona et al., 2021), there are other strategies that could minimise the number of false non-interactions. For example, false non-interactions could be minimised by generating non-interactions on a case-by-case basis (similar to the 'target-group absences' approach used in species distribution modelling; Barbet-Massin et al., 2012), an approach shown to improve the accuracy of predator-prey interactions predicted by deep learning models (Fricke et al., 2022). We generated non-interactions for the enhanced global dataset using interaction data from different regions, studies, and sources, but this approach could generate false and

uninformative non-interactions (e.g., a species pair that were not sympatric in a study area assigned as non-interacting, but the pair are important predator and prey where they are sympatric). Instead, training data non-interactions could be generated based on individual studies or locations of recorded interactions — i.e., identify what a predator does and does not eat in a particular area. Such a nuanced approach would lead to more reliable and informative non-interactions than would be achieved by assuming that any undocumented pairs from an incomplete global database are truly not predator and prey.

Our models performed better with the enhanced global dataset than with the Simpson Desert data. This difference might indicate that the random forests learnt patterns in the global dataset that are not as effective when applied back to specific communities. For example, the random forests might have learnt from the global dataset that predators with particular traits do not eat prey with particular traits, but predators or prey with those traits might not be present in the Simpson Desert (i.e., learning that relationship did not help the random forest predict interactions in the Simpson Desert). Alternatively, random forests might also learn which predators had few or many prey records within the global dataset — another pattern that would be less effective at identifying interactions and non-interactions when applied to the Simpson Desert predators due to the similarity among focal predators in terms of number of interactions in the training data (60% of predators in the global dataset had < 5 interaction records whereas all Simpson Desert focal predators had > 10 interaction records in the global dataset). Irrespective of what caused the difference in model performance, our results indicate that performance based on global datasets might not be a strong indicator of performance when applied to single communities.

Previous studies applying random forests in aquatic and invertebrate communities tend to report higher performance than our models achieved (mean true skill statistic consistently > 0.6; Brousseau et al., 2018; Desjardins-Proulx et al., 2017; Laigle et al., 2018; Parravicini et

al., 2020; Strona et al., 2021). There are several possible explanations for this discrepancy, including: (*i*) previous studies pooling prey species into broad resource categories (Parravicini et al., 2020), (*ii*) higher-quality training data in previous studies (Laigle et al., 2018; Parravicini et al., 2020), (*iii*) previous studies including resource predictions for a wider range of trophic groups — including easily-excluded non-interactions (such as herbivores eating vertebrates; Desjardins-Proulx et al., 2017), and (*iv*) interactions might be more difficult to predict for terrestrial vertebrates than fish or soil invertebrates, although further testing is required to evaluate that possibility. Another potential means of improving the performance of random forests that neither we nor others have addressed is to include detail on intraspecific variation — including ontogenetic shifts — in traits and interactions (González-Varo & Traveset, 2016).

Random forest machine learning is a useful and flexible tool that has great potential in terms of predicting species interactions. While global interaction databases are a valuable resource for training and testing random forests, they currently include only a tiny fraction of the species interactions that occur. Furthermore, for the consumers that are included in such databases, in most cases only a small subset of their trophic interactions are documented — thereby increasing the likelihood of false-non-interactions being generated. Databases therefore need to be expanded, and a framework for generating non-interactions that avoids or minimises false non-interactions and retains informative non-interactions is required. By making such changes, the application of random forests can be further honed to infer species interactions more accurately.

Acknowledgements

We thank Phuong Chau and Tori Llewelyn for their helpful input. This work was supported by the Australian Research Council Centre of Excellence for Australian Biodiversity and Heritage (grant number CE170100015).

References

Allan, J. D., Castillo, M. M., & Capps, K. A. (2021). Energy Flow and Nutrient Cycling in Aquatic Communities. In J. D. Allan, M. M. Castillo, & K. A. Capps (Eds.), *Stream Ecology: Structure and Function of Running Waters* (pp. 357–381). Springer International Publishing. https://doi.org/10.1007/978-3-030-61286-3_12

Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: How, where and how many? *Methods in Ecology and Evolution*, *3*(2), 327–338. https://doi.org/10.1111/j.2041-210X.2011.00172.x

Blois, J. L., Zarnetske, P. L., Fitzpatrick, M. C., & Finnegan, S. (2013). Climate Change and the Past, Present, and Future of Biotic Interactions. *Science*, *341*(6145), 499–504. https://doi.org/10.1126/science.1237184

Bohan, D. A., Vacher, C., Tamaddoni-Nezhad, A., Raybould, A., Dumbrell, A. J., & Woodward, G. (2017). Next-Generation Global Biomonitoring: Large-scale, Automated Reconstruction of Ecological Networks. *Trends in Ecology & Evolution*, *32*(7), 477–487. https://doi.org/10.1016/j.tree.2017.03.001

Brousseau, P.-M., Gravel, D., & Handa, I. T. (2018). Trait matching and phylogeny as predictors of predator-prey interactions involving ground beetles. *Functional Ecology*, *32*(1), 192–202. https://doi.org/10.1111/1365-2435.12943

Caron, D., Maiorano, L., Thuiller, W., & Pollock, L. J. (2022). Addressing the Eltonian shortfall with trait-based interaction models. *Ecology Letters*, *25*(4), 889–899. https://doi.org/10.1111/ele.13966

Chamberlain, S., Barve, V., Mcglinn, D., Oldoni, D., Desmet, P., Geffert, L., & Ram, K. (2022). *rgbif: Interface to the global biodiversity information facility* (R package version 3.6.0). https://CRAN.R-project.org/package=rgbif

Cox, N., Young, B. E., Bowles, P., Fernandez, M., Marin, J., Rapacciuolo, G., Böhm, M., Brooks, T. M., Hedges, S. B., Hilton-Taylor, C., Hoffmann, M., Jenkins, R. K. B., Tognelli, M. F., Alexander, G. J., Allison, A., Ananjeva, N. B., Auliya, M., Avila, L. J., Chapple, D. G., … Xie, Y. (2022). A global reptile assessment highlights shared conservation needs of tetrapods. *Nature*, 1–6. https://doi.org/10.1038/s41586-022-04664-7

Dellinger, A. S., Chartier, M., Fernández-Fernández, D., Penneys, D. S., Alvear, M., Almeda, F., Michelangeli, F. A., Staedler, Y., Armbruster, W. S., & Schönenberger, J. (2019). Beyond buzz-pollination – departures from an adaptive plateau lead to new pollination syndromes. *New Phytologist*, *221*(2), 1136–1149. https://doi.org/10.1111/nph.15468

Desjardins-Proulx, P., Laigle, I., Poisot, T., & Gravel, D. (2017). Ecological interactions and the Netflix problem. *PeerJ*, *5*, e3644. https://doi.org/10.7717/peerj.3644

Doherty, T. S., Glen, A. S., Nimmo, D. G., Ritchie, E. G., & Dickman, C. R. (2016). Invasive predators and global biodiversity loss. *Proceedings of the National Academy of Sciences*, *113*(40), 11261–11265. https://doi.org/10.1073/pnas.1602480113

Elmasri, M., Farrell, M. J., Davies, T. J., & Stephens, D. A. (2020). A hierarchical Bayesian model for predicting ecological interactions using scaled evolutionary relationships. *The Annals of Applied Statistics*, *14*(1), 221–240. https://doi.org/10.1214/19-AOAS1296

Faisal, A., Dondelinger, F., Husmeier, D., & Beale, C. M. (2010). Inferring species interaction networks from species abundance data: A comparative evaluation of various statistical and machine learning methods. *Ecological Informatics*, *5*(6), 451–464. https://doi.org/10.1016/j.ecoinf.2010.06.005

Fillios, M., Crowther, M. S., & Letnic, M. (2012). The impact of the dingo on the thylacine in Holocene Australia. *World Archaeology*, *44*(1), 118–134. https://doi.org/10.1080/00438243.2012.646112

Fricke, E. C., Hsieh, C., Middleton, O., Gorczynski, D., Cappello, C. D., Sanisidro, O., Rowan, J., Svenning, J.-C., & Beaudrot, L. (2022). Collapse of terrestrial mammal food webs since the Late Pleistocene. *Science*, *377*(6609), 1008–1011. https://doi.org/10.1126/science.abn4012

González-Varo, J. P., & Traveset, A. (2016). The labile limits of forbidden interactions. *Trends in Ecology & Evolution*, *31*(9), 700–710. https://doi.org/10.1016/j.tree.2016.06.009

Gravel, D., Poisot, T., Albouy, C., Velez, L., & Mouillot, D. (2013). Inferring food web structure from predator-prey body size relationships. *Methods in Ecology and Evolution*, *4*(11), 1083–1090. https://doi.org/10.1111/2041-210X.12103

Guénard, G., & Legendre, P. (2022). *MPSEM: Modeling phylogenetic signals using eigenvector maps* (R package version 0.4-1). https://CRAN.R-project.org/package=MPSEM

Guénard, G., Legendre, P., & Peres-Neto, P. (2013). Phylogenetic eigenvector maps: A framework to model and predict species traits. *Methods in Ecology and Evolution*, *4*(12), 1120–1131. https://doi.org/10.1111/2041-210X.12111

Herberstein, M. E., McLean, D. J., Lowe, E., Wolff, J. O., Khan, M. K., Smith, K., Allen, A. P., Bulbert, M., Buzatto, B. A., Eldridge, M. D. B., Falster, D., Fernandez Winzer, L., Griffith, S. C., Madin, J. S., Narendra, A., Westoby, M., Whiting, M. J., Wright, I. J., & Carthey, A. J. R. (2022). AnimalTraits—A curated animal trait database for body mass, metabolic rate and brain size. *Scientific Data*, *9*(1), 265. https://doi.org/10.1038/s41597-022-01364-9

Hortal, J., de Bello, F., Diniz-Filho, J. A. F., Lewinsohn, T. M., Lobo, J. M., & Ladle, R. J. (2015). Seven Shortfalls that Beset Large-Scale Knowledge of Biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, *46*(1), 523–549. https://doi.org/10.1146/annurev-ecolsys-112414-054400

Iturbide, M., Bedia, J., Herrera, S., del Hierro, O., Pinto, M., & Gutiérrez, J. M. (2015). A framework for species distribution modelling with improved pseudo-absence generation. *Ecological Modelling*, *312*, 166–174. https://doi.org/10.1016/j.ecolmodel.2015.05.018

Jones, K. E., Bielby, J., Cardillo, M., Fritz, S. A., O'Dell, J., Orme, C. D. L., Safi, K., Sechrest, W., Boakes, E. H., Carbone, C., Connolly, C., Cutts, M. J., Foster, J. K., Grenyer, R., Habib, M., Plaster, C. A., Price, S. A., Rigby, E. A., Rist, J., … Purvis, A. (2009). PanTHERIA: A species-level database of life history, ecology, and geography of extant and recently extinct mammals: *Ecological Archives* E090-184. *Ecology*, *90*(9), 2648–2648. https://doi.org/10.1890/08-1494.1

Kaneryd, L., Borrvall, C., Berg, S., Curtsdotter, A., Eklöf, A., Hauzy, C., Jonsson, T., Münger, P., Setzer, M., Säterberg, T., & Ebenman, B. (2012). Species-rich ecosystems are vulnerable to cascading extinctions in an increasingly variable world: Extinction Cascades in a Variable World. *Ecology and Evolution*, *2*(4), 858–874. https://doi.org/10.1002/ece3.218

Klomberg, Y., Tropek, R., Mertens, J. E. J., Kobe, I. N., Hodeček, J., Raška, J., Fominka, N. T., Souto-Vilarós, D., Janečková, P., & Janeček, Š. (2022). Spatiotemporal variation in the role of floral traits in shaping tropical plant-pollinator interactions. *Ecology Letters*, *25*(4), 839–850. https://doi.org/10.1111/ele.13958

Kotula, H. J., Peralta, G., Frost, C. M., Todd, J. H., & Tylianakis, J. M. (2020). *Predicting direct and indirect non-target impacts of biocontrol agents using machine-learning approaches* [Preprint]. Ecology. https://doi.org/10.1101/2020.11.24.395095

Laigle, I., Aubin, I., Digel, C., Brose, U., Boulangeat, I., & Gravel, D. (2018). Species traits as drivers of food web structure. *Oikos*, *127*(2), 316–326. https://doi.org/10.1111/oik.04712

Llewelyn, J. (2022). Random forests for predicting predator-prey interactions in terrestrial vertebrates. *GitHub and Zenodo*. https://doi.org/10.5281/zenodo.7037189

Llewelyn, J., Strona, G., McDowell, M. C., Johnson, C. N., Peters, K. J., Stouffer, D. B., de Visser, S. N., Saltré, F., & Bradshaw, C. J. A. (2021). Sahul's megafauna were vulnerable to plant-community changes due to their position in the trophic network. *Ecography*, *2022*(1), DOI: 10.1111/ecog.06089. https://doi.org/10.1111/ecog.06089

Lobo, J. M., Jiménez-Valverde, A., & Hortal, J. (2010). The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, *33*(1), 103–114. https://doi.org/10.1111/j.1600-0587.2009.06039.x

Mateo, R. G., Croat, T. B., Felicísimo, Á. M., & Muñoz, J. (2010). Profile or group discriminative techniques? Generating reliable species distribution models using pseudo-absences and target-group absences from natural history collections: Reliability of models using pseudo-absences and target-group absences. *Diversity and Distributions*, *16*(1), 84–94. https://doi.org/10.1111/j.1472-4642.2009.00617.x

Morales-Castilla, I., Matias, M. G., Gravel, D., & Araújo, M. B. (2015). Inferring biotic interactions from proxies. *Trends in Ecology & Evolution*, *30*(6), 347–356. https://doi.org/10.1016/j.tree.2015.03.014

Morton, D. N., Keyes, A., Barner, A. K., & Dee, L. E. (2022). Merging theory and experiments to predict and understand coextinctions. *Trends in Ecology & Evolution*. https://doi.org/10.1016/j.tree.2022.06.004

Oliveira, B. F., São-Pedro, V. A., Santos-Barrera, G., Penone, C., & Costa, G. C. (2017). AmphiBIO, a global database for amphibian ecological traits. *Scientific Data*, *4*, 170123. https://doi.org/10.1038/sdata.2017.123

Ornai, A., & Keasar, T. (2020). Floral complexity traits as predictors of plant-bee interactions in a Mediterranean pollination web. *Plants*, *9*(11), 1432. https://doi.org/10.3390/plants9111432

Paradis, E., & Schliep, K. (2019). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*, 526–528.

Parravicini, V., Casey, J. M., Schiettekatte, N. M. D., Brandl, S. J., Pozas-Schacre, C., Carlot, J., Edgar, G. J., Graham, N. A. J., Harmelin-Vivien, M., Kulbicki, M., Strona, G., & Stuart-Smith, R. D. (2020). Delineating reef fish trophic guilds with global gut content data synthesis and phylogeny. *PLOS Biology*, *18*(12), e3000702. https://doi.org/10.1371/journal.pbio.3000702

Pichler, M., Boreux, V., Klein, A.-M., Schleuning, M., & Hartig, F. (2020). Machine learning algorithms to infer trait-matching and predict species interactions in ecological networks. *Methods in Ecology and Evolution*, *11*(2), 281–293. https://doi.org/10.1111/2041-210X.13329

Pires, M. M., Koch, P. L., Fariña, R. A., de Aguiar, M. A. M., dos Reis, S. F., & Guimarães, P. R. (2015). Pleistocene megafaunal interaction networks became more vulnerable after human arrival. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1814), 20151367. https://doi.org/10.1098/rspb.2015.1367

Poisot, T. (2022). Guidelines for the prediction of species interactions through binary classification. *EcoEvoRxiv*, 16.

Poisot, T., Bergeron, G., Cazelles, K., Dallas, T., Gravel, D., MacDonald, A., Mercier, B., Violet, C., & Vissault, S. (2021). Global knowledge gaps in species interaction networks data. *Journal of Biogeography*, *48*(7), 1552–1563. https://doi.org/10.1111/jbi.14127

Pomeranz, J. P. F., Thompson, R. M., Poisot, T., & Harding, J. S. (2019). Inferring predator–prey interactions in food webs. *Methods in Ecology and Evolution*, *10*(3), 356–367. https://doi.org/10.1111/2041-210X.13125

Pringle, R. M., & Hutchinson, M. C. (2020). Resolving Food-Web Structure. *Annual Review of Ecology, Evolution, and Systematics*, *51*(1), annurev-ecolsys-110218-024908. https://doi.org/10.1146/annurev-ecolsys-110218-024908

Ripple, W. J., & Van Valkenburgh, B. (2010). Linking top-down forces to the Pleistocene megafaunal extinctions. *BioScience*, *60*(7), 516–526. https://doi.org/10.1525/bio.2010.60.7.7

Samson, F. B., & Knopf, F. L. (1996). *Ecosystem Management: Selected Readings*. Springer New York. https://doi.org/10.1007/978-1-4612-4018-1

Sinclair, A. R. E., Mduma, S., & Brashares, J. S. (2003). Patterns of predation in a diverse predator–prey system. *Nature*, *425*(6955), 288–290. https://doi.org/10.1038/nature01934

Smith, J. G., & Phillips, B. L. (2006). Toxic tucker: The potential impact of Cane Toads on Australian reptiles. *Pacific Conservation Biology*, *12*(1), 40–49. https://doi.org/10.1071/pc060040

Strona, G. (2022). *Hidden Pathways to Extinction*. Springer International Publishing. https://doi.org/10.1007/978-3-030-86764-5

Strona, G., Beck, P. S. A., Cabeza, M., Fattorini, S., Guilhaumon, F., Micheli, F., Montano, S., Ovaskainen, O., Planes, S., Veech, J. A., & Parravicini, V. (2021). Ecological dependencies make remote reef fish communities most vulnerable to coral loss. *Nature Communications*, *12*(1), 7282. https://doi.org/10.1038/s41467-021-27440-z

Strona, G., & Bradshaw, C. J. A. (2018). Co-extinctions annihilate planetary life during extreme environmental change. *Scientific Reports*, *8*(1), 16724. https://doi.org/10.1038/s41598-018-35068-1

Strydom, T., Catchen, M. D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., Forero-Muñoz, N. R., Higino, G., Mercier, B., Gonzalez, A., Gravel, D., Pollock, L., & Poisot, T. (2021). A roadmap towards predicting species interaction networks (across space and time). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1837), 20210063. https://doi.org/10.1098/rstb.2021.0063

Sydenham, M. A. K., Venter, Z., Reitan, T., Rasmussen, C., Skrindo, A. B., Skoog, D. I. J., Hanevik, K.-A., Hegland, S. J., Dupont, Y. L., Nielsen, A., Chipperfield, J., & Rusch, G. (2021). MetaComNet: A random forest-based framework for making spatial predictions of plant-pollinator interactions. *Methods in Ecology and Evolution*. https://hvlopen.brage.unit.no/hvlopen-xmlui/handle/11250/2839400

Thompson, J. N. (1999). The evolution of species interactions. *Science*, *284*(5423), 2116–2118. https://doi.org/10.1126/science.284.5423.2116

Tilman, D., Clark, M., Williams, D. R., Kimmel, K., Polasky, S., & Packer, C. (2017). Future threats to biodiversity and pathways to their prevention. *Nature*, *546*(7656), 73–81. https://doi.org/10.1038/nature22900

Traill, L. W., Lim, M. L. M., Sodhi, N. S., & Bradshaw, C. J. A. (2010). Mechanisms driving change: Altered species interactions and ecosystem function through global warming. *Journal of Animal Ecology*, *79*(5), 937–947. https://doi.org/10.1111/j.1365-2656.2010.01695.x

Tylianakis, J. M., Didham, R. K., Bascompte, J., & Wardle, D. A. (2008). Global change and species interactions in terrestrial ecosystems. *Ecology Letters*, *11*(12), 1351–1363. https://doi.org/10.1111/j.1461-0248.2008.01250.x

Valdovinos, F. S., Berlow, E. L., Moisset de Espanés, P., Ramos-Jiliberto, R., Vázquez, D. P., & Martinez, N. D. (2018). Species traits and network structure predict the success and impacts of pollinator invasions. *Nature Communications*, *9*(1), 2153. https://doi.org/10.1038/s41467-018-04593-y

Van der Putten, W. H., Macel, M., & Visser, M. E. (2010). Predicting species distribution and abundance responses to climate change: Why it is essential to include biotic interactions across trophic levels. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1549), 2025–2034. https://doi.org/10.1098/rstb.2010.0037

Wilman, H., Belmaker, J., Simpson, J., de la Rosa, C., Rivadeneira, M. M., & Jetz, W. (2014). EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals: *Ecological Archives* E095-178. *Ecology*, *95*(7), 2027–2027. https://doi.org/10.1890/13-1917.1

Wright, M. N., & Ziegler, A. (2017). ranger: A fast implementation of Random Forests for high dimensional data in C++ and R. *Journal of Statistical Software*, *77*(1). https://doi.org/10.18637/jss.v077.i01

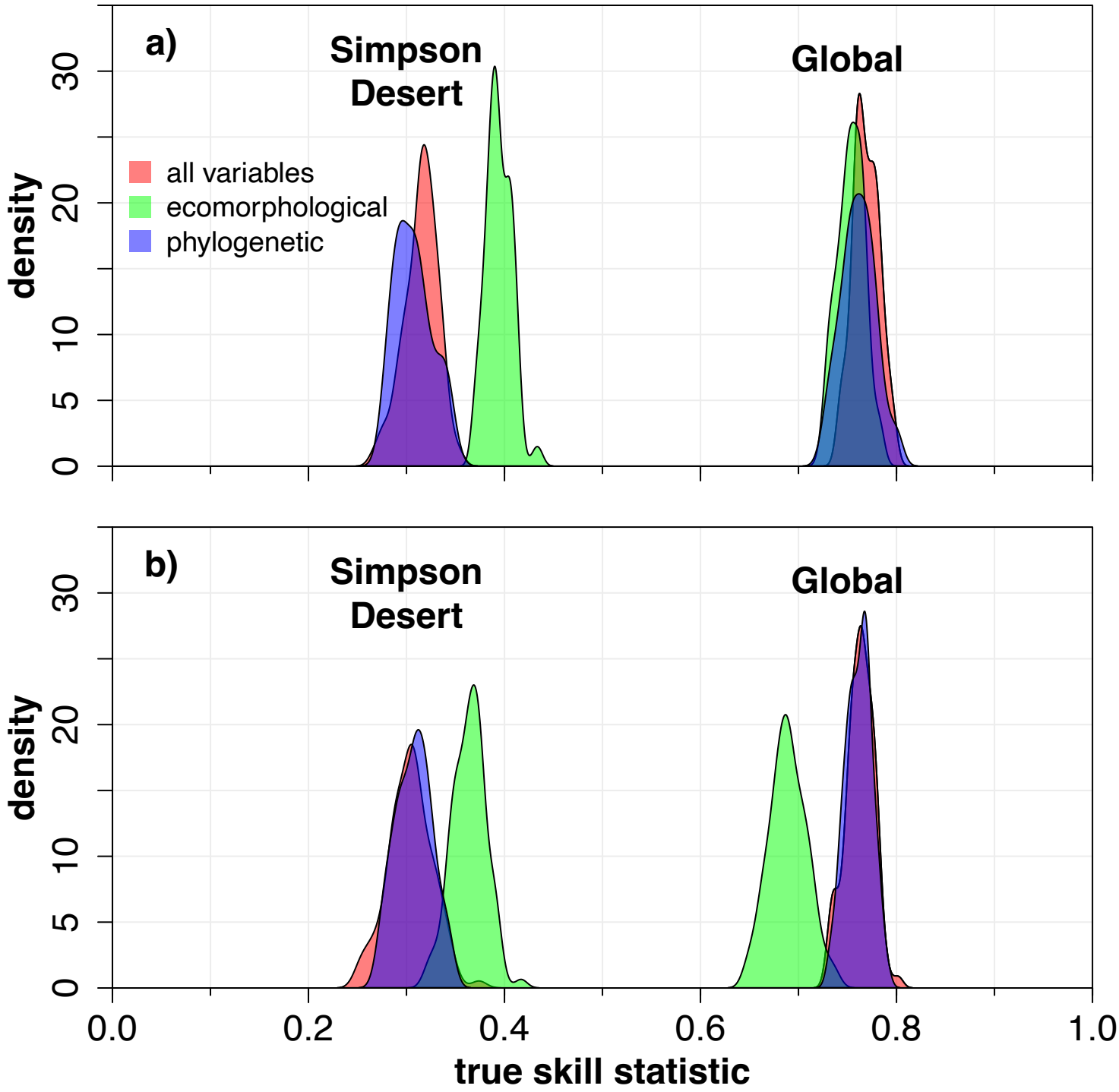

Figure 1. Density plots comparing the performance (true skill statistic) of random forest models predicting predator-prey interactions in birds and mammals. The random forests used either ecomorphological traits, phylogenetic eigenvectors, or both types of data, and were applied to a global interaction dataset and a specific ecosystem dataset (7 predators and their potential prey (sympatric species) in the Simpson Desert). Panel a) shows results from models using many variable (42, 42, and 84 variables, for ecomorphological, phylogenetic eigenvector, and both data-type models), while panel b) show results using few variables (10, 10, and 20 variables, for ecomorphological, phylogenetic eigenvector, and both data-type models).

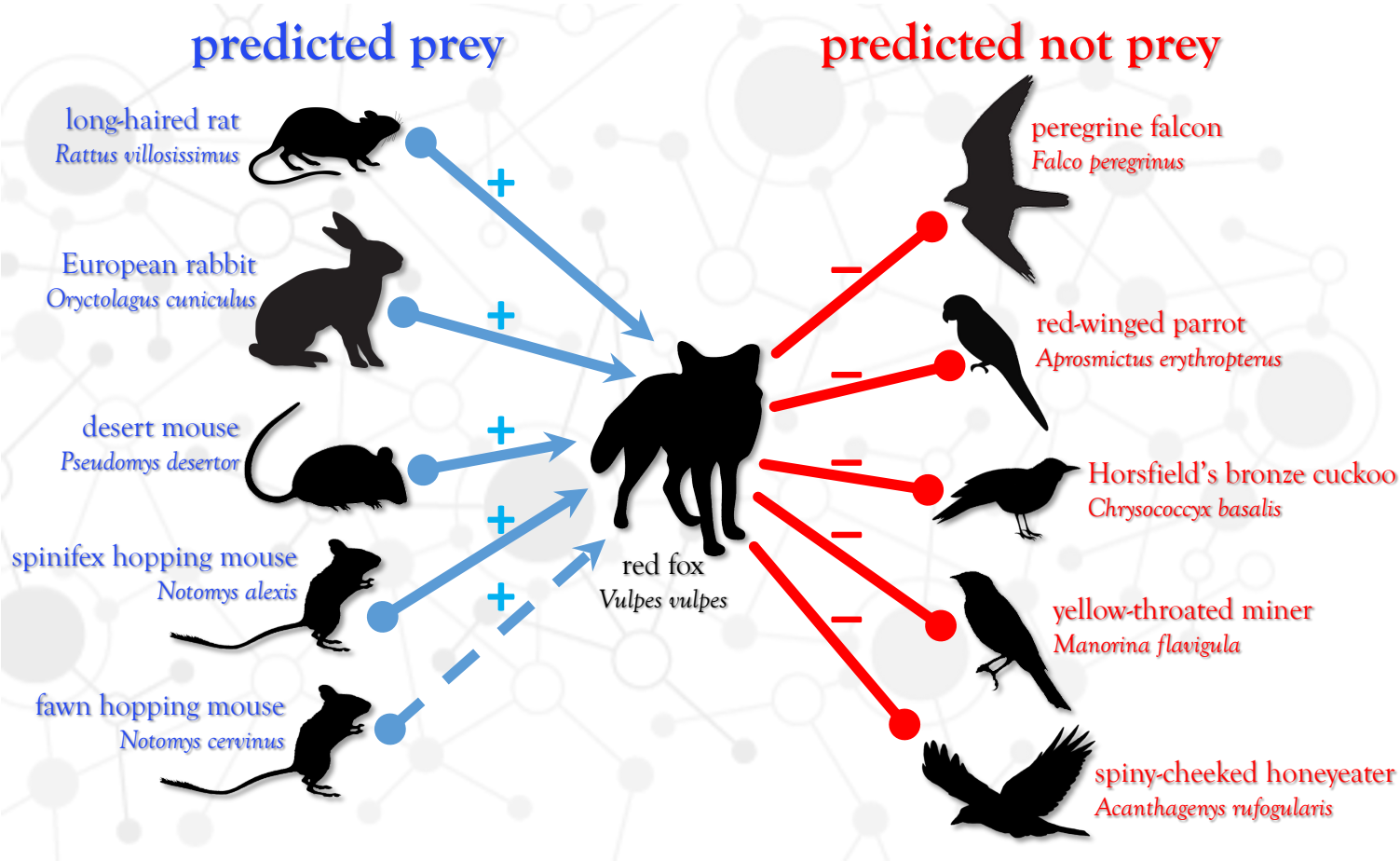

Figure 2. The five most and least likely mammalian and bird prey for the red fox (*Vulpes vulpes*) in the Simpson Desert as predicted by random forest. Solid blue arrows indicate prey that was predicted and observed, dashed blue arrows show prey that was predicted but not observed, and red lines indicate species that were not predicted or observed as prey. The five most likely prey for foxes, according to the many-variable ecomorphological model (run 100 times), were the: long-haired rat (*Rattus villosissimus*), European rabbit (*Oryctolagus cuniculus*), desert mouse (*Pseudomys desertor*), spinifex hopping mouse (*Notomys alexis*), and fawn hopping mouse (*Notomys cervinus*). The five least likely prey for foxes were the: peregrine falcon (*Falco peregrinus*), red-winged parrot (*Aprosmictus erythropterus*), Horsfield's bronze cuckoo (*Chrysococcyx basalis*), yellow-throated miner (*Manorina flavigula*), and spiny-cheeked honeyeater (*Acanthagenys rufogularis*).

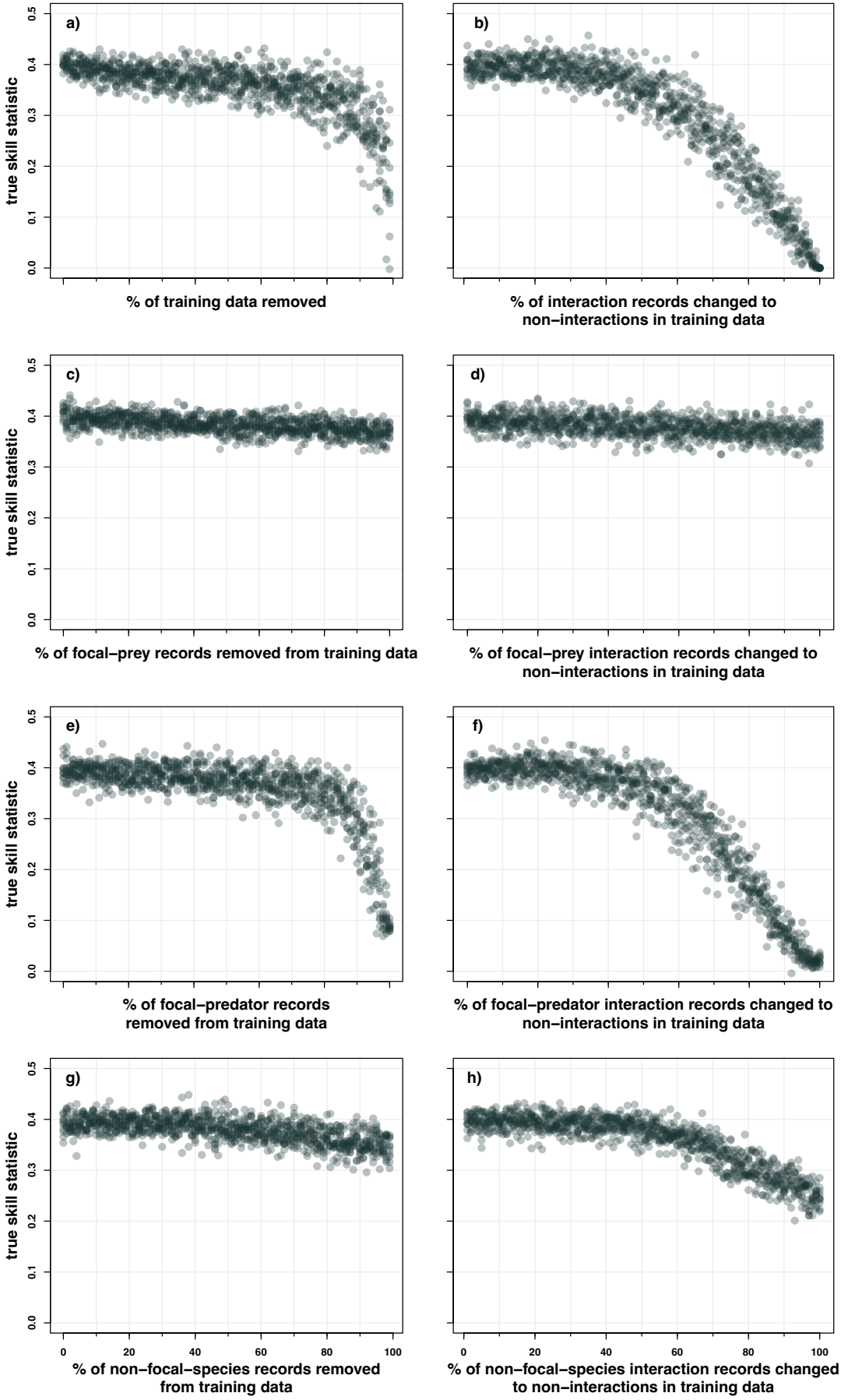

Figure 3. Effect of training-dataset quality on random forest performance predicting predator-prey interactions in birds and mammals. Each panel shows performance when a percentage of either the whole or specific taxonomic groups in the training dataset were modified. Panels a) and b) show performance when changes were made to the whole dataset, panels c) and d) show performance when changes were restricted to focal prey, panels e) and f) show performance when changes were restricted to focal predators, and panels g) and h) show performance when modifications were restricted to non-focal (non-Simpson Desert) species. Panels in the first column show the effects of removing records from the training data whereas panels in the second column show the effects of switching interaction records to non-interactions (false non-interactions). The random forests were trained on unmodified or modified global predator-prey interaction datasets and applied to seven predators and their potential prey from the Simpson Desert. Plot points indicate performance (according to the true skill statistic) when applied to the Simpson Desert predators.

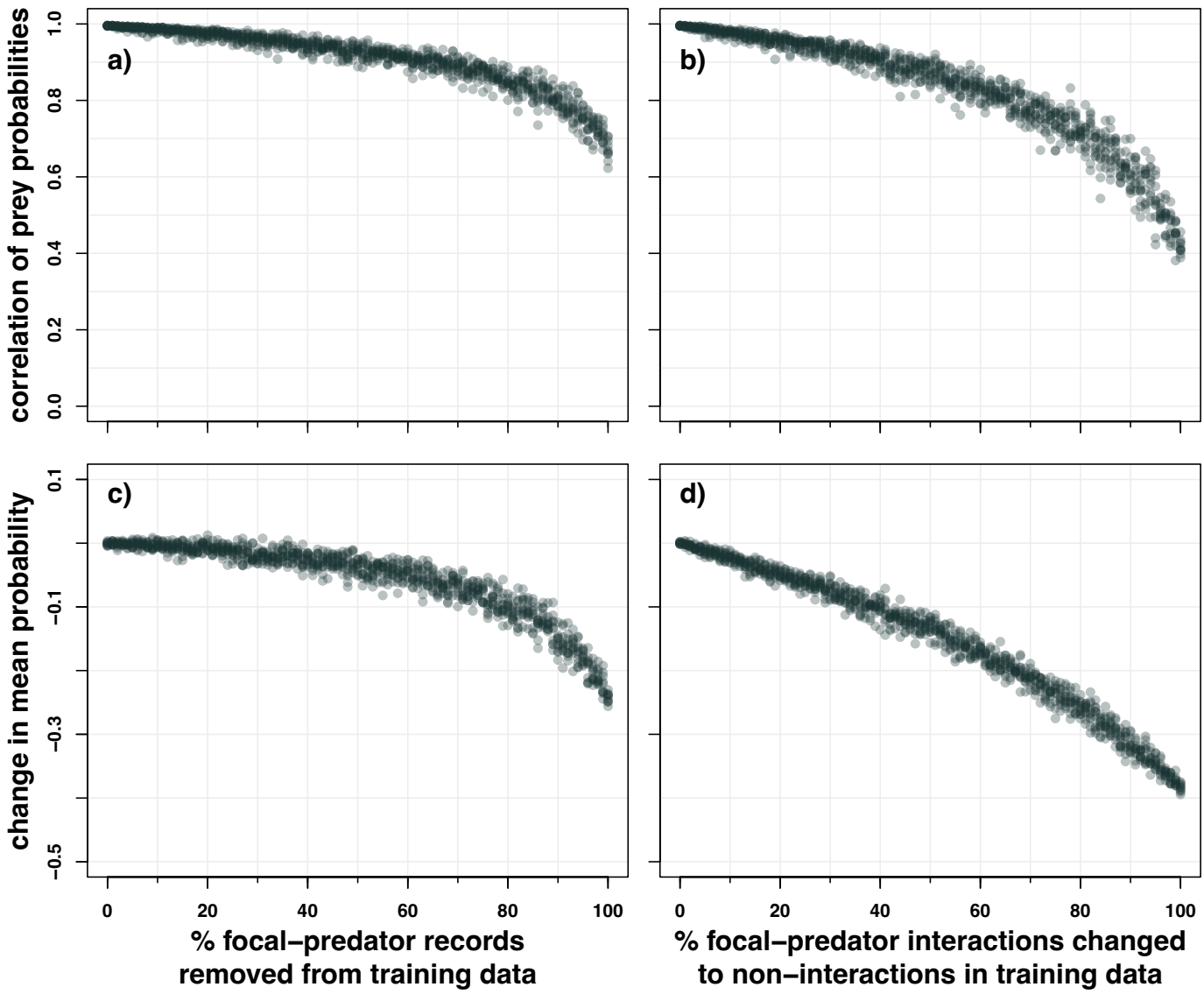

Figure 4. Effects of training dataset quality on (*i*) the predicted relative suitability of prey and (*ii*) the mean probabilities assigned to potential prey by random forests. Panels in the first column show the effects of reducing the number of interactions records involving the focal predator in the training data, whereas panels in the second column show the effect of switching interaction records involving the focal predator to non-interactions (false non-interactions). Panels a) and b) show the correlation between prey probabilities assigned by the modified-training-dataset models with that assigned by models trained on the full training dataset. Panels c) and d) show the mean probability (suitability) assigned to potential prey by the modified-training-dataset models minus that value assigned by the models trained on the complete training datasets. The random forests were trained on the modified and unmodified global predator-prey interaction datasets and applied to seven predators and their potential prey from the Simpson Desert.