

Trophic interaction models predict interactions across regions, not food web properties.

Table of Content:

Appendix S1 – Description of the food webs: Brief description of how the four food webs were collated, and how we extracted predator-prey interactions.

Appendix S2 – Trait coverage and imputation: Additional information on the coverage of the trait data and on the fit of the MissForest algorithm used for data imputation

Appendix S3 – Description of the structure of the predictive model: We detail the structure of the predictive model (response variable, predictors, transformations, priors).

Appendix S4 - Description of species' roles and food web-level properties: We describe each species role and food web-level metrics, how we calculated them, and their relation to functioning.

Appendix S5 - Boosted regression tree methods and results: Brief description of the method used to fit and test the boosted regression trees and the associated main results.

Appendix S6 - Convergence of the predictive models: Distribution of rank-normalized potential scale reduction factor on split chains (Rhat) and rank plots of the population-level effects.

Appendix S7 - Predicting species interactions extended results: Results of the predicting species interactions with other evaluation metrics, and results of the linear models between model performance and distances estimates.

Appendix S8 - Predicting species roles extended results: Additional results for the prediction of species' role (linear model R² and coefficients).

Appendix S9 - Predicting food web properties extended results: Additional results for the prediction of food web results.

Appendix S10 – Results with alternative Serengeti food web.

Appendix S11 – References: References for the supplemental material

Appendix S1: Description of the food webs

Europe food web: The food web comprising European tetrapods was assembled by Maiorano et al. (2020). It contains the potential trophic interactions between all terrestrial vertebrates in Europe based on data from atlases, field guides, published papers, and expert knowledge. Maiorano et al. (2020) originally reported interactions between predators and their prey at different life-stage (eggs, larvae/young, or adult). In this study, we focused on interactions between predators and adult prey only. We combined the layer of the continental European bioregions (version 3) European Environment Agency (<https://www.eea.europa.eu/data-and-maps/data/biogeographical-regions-europe-3>) to define its spatial domain, which covers an area of $11 \times 10^6 \text{ km}^2$.

Pyrenees food web: The Pyrenees vertebrate food web was assembled independently from the European food web by Lurgi et al. (2012) who compiled all known interactions between vertebrate species in the alpine ecosystem of the southeastern Pyrenees mountain range. Data was collated from atlases, existing databases, published scientific articles and reports, and expert knowledge. This network is described at the species level and for this study, we excluded fish species. We defined the spatial domain of the food web to be of the European Mountain Area layer provided by the European Environment Agency (<https://www.eea.europa.eu/data-and-maps/data/european-mountain-areas>), covering an area of $50 \times 10^3 \text{ km}^2$.

Northern Québec and Labrador food web: The trophic tundra food web of Labrador and northern Quebec (Canada) was compiled by Berteaux et al. (2018). The Northern Québec and Labrador food web documents the presence (absence) of trophic relationships between all birds, mammals, and groups of plants (e.g., moss, ferns, grass) and arthropods (e.g., Lepidoptera, Diptera, Arachnida) whose distribution overlap Labrador and Quebec north of 50°N (Nunavik region). Trophic interactions were collated from a literature review, data from existing databases, and expert knowledge. For our study, we focused on the bird and mammal species (i.e., the vertebrates) of this network. We clipped at Latitude 50°N the polygon of the Québec province and removed Newfoundland from the Newfoundland and Labrador province layer provided by the Open Canada data portal (<https://open.canada.ca/data/en/dataset/a883eb14-0c0e-45c4-b8c4-b54c4a819edb>) to define its spatial domain, which covers $1.3 \times 10^6 \text{ km}^2$.

Serengeti food web: The Serengeti food web compiles all interactions between resource categories (e.g., seeds, shrubs, fruits and nectar), invertebrate groups (e.g., thrips, bees, moths, crickets), and vertebrate species in the Serengeti National Park (Tanzania) and was assembled from a literature review (de Visser et al., 2011). The nodes of the original food webs are resource categories, or trophic groups including one or more invertebrate groups and vertebrate species. In this study, we focused on vertebrate species, and assumed that species within a trophic group share the same predator and prey species. We used the Serengeti ecosystem layer provided by Serengeti GIS & Data Center (<https://serengetidata.weebly.com/boundaries.html>) to define its spatial domain, which covers $36 \times 10^3 \text{ km}^2$.

Appendix S2: Trait coverage and imputation

We imputed missing trait data using the *missForest* function from the *missForest* package in R (Stekhoven, 2013; Stekhoven & Buehlmann, 2012).

For each variable with missing values, the function fit a random forest using the observed values to predict missing values. The imputation procedure is repeated until a stopping criterion is met. Like for other bootstrap aggregating (bagging) algorithm, the mean of the predictions made by the individual trees of the random forest is used for continuous variables and the mode for categorical variables.

We ran the function for birds, mammals, amphibians, and reptiles separately. Each time, we included species that occur at least one of the food webs (we ran *missForest* for all birds, all amphibians, all mammals, and all reptiles across the four food webs). We used all traits available (see following tables), including trait not used for interaction predictions, and the order, family, and genus of the species in the *missForest* function. This approach allows to (1) make sure a species present at different location as the same trait values, (2) take into account the large functional differences across class, (3) maximize the available data for the trait imputation within class. Not all traits were included in the predictive models to reduce collinearity among predictors (e.g., body length and body mass; age of maturity and longevity). The traits used in predictive models are bolded in the following tables.

Here, we document the trait coverage and imputation performance for each class. The traits in bold are the one used for interaction prediction (the other traits are not used in the following analyses). The performance metric is the normalized root squared mean squared error (NRMSE) for the continuous missing values, and proportion of falsely classified (PFC) for the categorical missing values. For both performance metrics, the closer to 0, the better the fit. The performance metrics are defined as follow:

$$NRMSE = \sqrt{\frac{mean((X_{true} - X_{imp})^2)}{var(X_{true})}}$$

where X_{true} is the complete continuous trait matrix, X_{imp} is the imputed continuous trait matrix.

PFC is the proportion of falsely classified categorical traits.

Amphibians (118 species):

Trait	Variable type	Coverage (%)	Mean (SD)
Trophic level (3 categories)	Categorical	62.7	NA
Diel Activity (2 categories)	Categorical	68.6	NA
Forest (Binary)	Categorical	88.1	NA
Savanna (Binary)	Categorical	88.1	NA
Shrubland (Binary)	Categorical	88.1	NA
Grassland (Binary)	Categorical	88.1	NA
Wetland (Binary)	Categorical	88.1	NA
Rocky areas (Binary)	Categorical	88.1	NA
Cave and subterranean (Binary)	Categorical	88.1	NA
Desert (Binary)	Categorical	88.1	NA
Marine (Binary)	Categorical	88.1	NA
Marine/Intertidal or coastal/supratidal (Binary)	Categorical	88.1	NA
Artificial (Binary)	Categorical	88.1	NA
Introduced vegetation (Binary)	Categorical	88.1	NA
Habitat breadth	Continuous	88.1	10.92 (7.92)
Body length (mm)	Continuous	91.5	103.11 (56.2)
Body mass (g)	Continuous	55.1	44.1 (174)
Age of Maturity (day)	Continuous	59.3	876.6 (396)
Maximum longevity (day)	Continuous	45.8	5442 (2848)
Clutch size (number of offspring)	Continuous	77.1	1980 (3899)
Adult SVL (cm)	Continuous	0*	NA
Mean longevity (day)	Continuous	0*	NA
Generation length (day)	Continuous	0*	NA

*When a trait is missing for all species, the trait is not used in the missForest algorithm

Error rate:

NRMSE: 0.64

PFC: 0.08

Birds (714 species):

Trait	Variable type	Coverage (%)	Mean (SD)
Trophic level (3 categories)	Categorical	81.7	NA
Diel Activity (2 categories)	Categorical	95.6	NA
Forest (Binary)	Categorical	97.8	NA
Savanna (Binary)	Categorical	97.8	NA
Shrubland (Binary)	Categorical	97.8	NA
Grassland (Binary)	Categorical	97.8	NA
Wetland (Binary)	Categorical	97.8	NA
Rocky areas (Binary)	Categorical	97.8	NA
Cave and subterranean (Binary)	Categorical	97.8	NA
Desert (Binary)	Categorical	97.8	NA
Marine (Binary)	Categorical	97.8	NA
Marine/Intertidal or coastal/supratidal (Binary)	Categorical	97.8	NA
Artificial (Binary)	Categorical	97.8	NA
Introduced vegetation (Binary)	Categorical	97.8	NA
Habitat breadth	Continuous	97.8	12.75 (11.6)
Body mass (g)	Continuous	99.2	715 (4343)
Age of Maturity (day)	Continuous	67.8	645 (448)
Maximum longevity (day)	Continuous	70.2	6871 (4602)
Clutch size (number of offspring)	Continuous	98.6	4.33 (2.12)
Adult SVL (cm)	Continuous	30.1	48.2 (24.7)
Mean longevity (day)	Continuous	70.2	5697 (3433)
Generation length (day)	Continuous	98.6	2466 (1515)

Error rate:

NRMSE: 0.49

PFC: 0.09

Mammals (450 species):

Trait	Variable type	Coverage (%)	Mean (SD)
Trophic level (3 categories)	Categorical	95.8	NA
Diel Activity (2 categories)	Categorical	94.9	NA
Forest (Binary)	Categorical	91.6	NA
Savanna (Binary)	Categorical	91.6	NA
Shrubland (Binary)	Categorical	91.6	NA
Grassland (Binary)	Categorical	91.6	NA
Wetland (Binary)	Categorical	91.6	NA
Rocky areas (Binary)	Categorical	91.6	NA
Cave and subterranean (Binary)	Categorical	91.6	NA
Desert (Binary)	Categorical	91.6	NA
Marine (Binary)	Categorical	91.6	NA
Marine/Intertidal or coastal/supratidal (Binary)	Categorical	91.6	NA
Artificial (Binary)	Categorical	91.6	NA
Introduced vegetation (Binary)	Categorical	91.6	NA
Habitat breadth	Continuous	91.6	5.87 (5.32)
Body mass (g)	Continuous	96.4	34958 (235221)
Age of Maturity (day)	Continuous	63.8	460 (563)
Maximum longevity (day)	Continuous	75.3	4784 (4440)
Clutch size (number of offspring)	Continuous	86.2	3.33 (2.00)
Adult SVL (cm)	Continuous	80.0	52.3 (78.3)
Mean longevity (day)	Continuous	75.3	4410 (4237)
Generation length (day)	Continuous	95.6	1535 (1154)

Error rate:

NRMSE: 0.71

PFC: 0.12

Reptiles (283 species):

Trait	Variable type	Coverage (%)	Mean (SD)
Trophic level (3 categories)	Categorical	68.2	NA
Diel Activity (2 categories)	Categorical	76.0	NA
Forest (Binary)	Categorical	64.0	NA
Savanna (Binary)	Categorical	64.0	NA
Shrubland (Binary)	Categorical	64.0	NA
Grassland (Binary)	Categorical	64.0	NA
Wetland (Binary)	Categorical	64.0	NA
Rocky areas (Binary)	Categorical	64.0	NA
Cave and subterranean (Binary)	Categorical	64.0	NA
Desert (Binary)	Categorical	64.0	NA
Marine (Binary)	Categorical	64.0	NA
Marine/Intertidal or coastal/supratidal (Binary)	Categorical	64.0	NA
Artificial (Binary)	Categorical	64.0	NA
Introduced vegetation (Binary)	Categorical	64.0	NA
Habitat breadth	Continuous	64.0	5.29 (2.98)
Body mass (g)	Continuous	95.8	2559 (29423)
Age of Maturity (day)	Continuous	42.4	975 (990)
Maximum longevity (day)	Continuous	64.0	5350 (6840)
Clutch size (number of offspring)	Continuous	83.4	7.35 (8.68)
Adult SVL (cm)	Continuous	44.5	10.16 (12.92)
Mean longevity (day)	Continuous	40.6	4128 (5070)
Generation length (day)	Continuous	0*	NA

*When a trait is missing for all species, the trait is not used in the missForest algorithm

Error rate:

NRMSE: 0.78

PFC: 0.11

Appendix S3: Description of the structure of the predictive model

The predictive models we use is a Bayesian hierarchical generalized linear model with 13 trait-based predictor variables with varying intercepts and slopes for the order of the predator. All traits were extracted from Etard et al. (2020). All traits are described in the original paper, but here, we describe briefly each derived variable used in the predictive models:

- Population-level effects:
 - Trophic level of the predator: Categorical with 3 levels - Herbivore, Omnivore, and Carnivore. Coded as 2 binary variables (Omnivore and Carnivore).
 - Habitat Breadth of the predator: Quantitative - Number of habitats the predator is known to use based on the second level of the IUCN habitat data.
 - Body mass of the predator: Quantitative - Log-transformed body mass of the predator
 - Longevity of the predator: Quantitative - Life span of an individual of the predator species. Species that live longer can be expected to be predators has they favour survival over reproduction.
 - Clutch Size of the predator: Quantitative - Number of offspring or eggs per clutch/litter by the predator. Species that have less offsprings can be expected to be predators has they favour survival over reproduction.
 - Trophic level of the prey: Categorical with 3 levels - Herbivore, Omnivore, and Carnivore. Coded as 2 binary variables (Omnivore and Carnivore)
 - Habitat Breadth of the prey: Quantitative - Number of habitats the prey is known to use based on the second level of the IUCN habitat data.
 - Body mass of the prey: Quantitative - Log-transformed body mass of the prey
 - Longevity of the prey: Quantitative - Life span of an individual of the prey species. Species that have shorter life can be expected to be prey has they favour reproduction over survival.
 - Clutch Size of the prey: Quantitative - Number of offspring or eggs per clutch/litter by the prey. Species that have more offsprings can be expected to be prey has they favour reproduction over survival.
 - Activity Time match: Categorical binary – The predator and the prey are active during the same time of the day (1) or not (0). The original activity time data for the prey and the predator is a binary variable with possible values nocturnal and non-nocturnal (diurnal, crepuscular, or cathemeral)
 - Habitat match: Quantitative – Jaccard similarity between the habitat use of the predator and the prey. The original habitat data are 12 binary variables: Forest, Savanna, Shrubland, Grassland, Wetland, Rocky areas, Cave and subterranean, Desert, Marine, Marine/Intertidal or coastal/supratidal, Artificial, Introduced vegetation
 - Body mass match: The squared difference between the log-transformed body mass of the predator and the log-transformed body mass of the prey.

We centered and scaled the quantitative predictors by subtracting each value by the mean and dividing by twice the standard deviation. The mean and standard deviation of each predictor was computed over the entire set of species (over all food webs). The predictors that we scaled are: the log-transformed habitat breadth of the predator and the prey, the log-transformed body mass of the predator and the prey, the log-transformed longevity of the predator and the prey, the log-transformed clutch/litter size of the predator and the prey, the habitat match, and the body mass match.

- Group-level effects (intercept and slopes):

- Order of the predator
 - Northern Québec and Labrador food web: 21 levels
 - Europe: 36 levels
 - Pyrenees: 23 levels
 - Serengeti: 41 levels

- Model:

$$L_{ij} \sim Bernoulli(p_{ij}),$$

$$\text{logit}(p_{ij}) = \alpha + \alpha_{\text{predator}[j]} + \sum_{k=1}^{13} (\beta_k \times T_k + \beta_{k,\text{predator}[j]} \times T_k)$$

- L_{ij} is the interaction between predator j and prey i .
- P_{ij} is the probability of interaction between predator j and prey i
- α and β are the fixed intercept and slopes
- $\alpha_{\text{predator}[j]}$ and $\beta_{k,\text{predator}[j]}$ are the order-specific intercept and slopes for the order of predator j
- T_k are the trait-based predictors.

- Priors:

$$\alpha, \beta \sim Normal(0, 1)$$

$$\alpha_{\text{predator}[j]}, \beta_{k,\text{predator}[j]} \sim Normal(\mu, \sigma)$$

$$\mu \sim Normal(0, 1)$$

$$\sigma \sim HalfCauchy(0, 5)$$

- Number of iterations:

- Warmup: 2000 iterations
- Inference: 2000 iterations

- Training:

- 70% of all interactions in each food web and a equal number of non-interactions (random undersampling): Northern Québec and Labrador food web (769 interactions and non-interactions), Europe (40 422 interactions and non-interactions), Pyrenees (582 interactions and non-interactions), Serengeti (7727 interactions and non-interactions). Random undersampling is a strategy to handle imbalanced dataset where one class of the response variable (non-interactions) is a lot more prevalent than the other class (interactions). Imbalanced dataset can influence and hinder training of some predictive models.

- Validation:

- For within food web predictions (i.e., trophic interaction predictions in the food web used for training): We used the remaining 30% of the interactions of each food webs and a

number of non-interactions that makes the prevalence of interactions in the validation dataset equal to the prevalence of interactions in the entire food web.

- For across food web predictions (i.e., trophic interaction predictions in food webs other than the one used for training): We used the entire food web (all interactions and non-interactions).

- Model formula:

```
brms_form <- bf(interaction ~ 1 + (Omnivore.predator + Carnivore.predator +
Habitat_breadth.predator + BM.predator + Longevity.predator + ClutchSize.predator + Omnivore.prey
+ Carnivore.prey + Habitat_breadth.prey + BM.prey + Longevity.prey + ClutchSize.prey +
ActivityTime.match + Habitat.match + BM.match) + (1 + (Omnivore.predator + Carnivore.predator +
Habitat_breadth.predator + BM.predator + Longevity.predator + ClutchSize.predator + Omnivore.prey
+ Carnivore.prey + Habitat_breadth.prey + BM.prey + Longevity.prey + ClutchSize.prey +
ActivityTime.match + Habitat.match + BM.match) || Order.predator), family = bernoulli())
```

Appendix S4: Description of species roles and food web-level properties

Species roles:

For more information on what each species role metrics represent and aim to describe, we recommend the reviews made by Cirtwill et al. (2018) and Delmas et al. (2019). Here, we very briefly describe what these metrics are, how they are related to the functioning and the stability of food webs:

- **Number of prey:** We calculated the number of prey of a species as the number of interactions for which the species is the predator (in-degree). It measures how generalized the species is. The loss of species with many prey are more likely to trigger top-down cascades (Curtsdotter et al., 2011). In contrast, specialist predators are more vulnerable to co-extinction. In-degree and out-degree quantify the direct influence of species on other species.
- **Number of predators:** We calculated the number of predators of a species as the number of interactions for which the species is the prey (out-degree). It measures the vulnerability of the prey. The loss of species with many predators are more likely to trigger bottom-up cascades (Curtsdotter et al., 2011). In-degree and out-degree quantify the direct influence of species on other species.
- **Betweenness:** The betweenness centrality of a species is the number of times the shortest path linking any two species in the network goes through the focal species (Freeman, 1977). Betweenness quantify the number of food chains a species takes part in, and thus how it affects energy flows (Cirtwill et al., 2018). It is considered one important measures of species topological importance (Jordán, 2009). We calculated betweenness using the function *centr_betw* in the R package *igraph* (Csardi & Nepusz, 2006).
- **Closeness:** Closeness centrality measures how far a species is to any other species. It is therefore a more “global” measure of centrality than betweenness and quantify how “efficiently” the change in this species will spread through the food web. We used the function *centr_clo* in the R package *igraph* (Csardi & Nepusz 2006).
- **Eigenvector centrality:** Eigenvector centrality uses the eigenvector of the largest eigenvalue of the adjacency matrix describing the food web. It is another “global” measure of centrality that aims to describe species importance in the food web. Eigenvector centrality is related to the contribution of species to network stability (Allesina & Pascual, 2009). We used the function *centr_eigen* of the R package *igraph* (Csardi & Nepusz 2006).
- **Trophic level:** Trophic level describe the “vertical” position (primary consumer, secondary consumer, etc.) of a species in a food web. We calculated trophic level using the function *TrophInd* of the package *NetIndices* package (Kones et al., 2009). The function calculates trophic level as 1 + the weighted average of the trophic levels of its food items. Top predators and basal species are expected to have larger effect on the community through top-down and bottom-up cascades (Dyer & Letourneau, 2003).
- **Omnivory:** Omnivory is defined as the variety of trophic levels a species feed on. Omnivory has the potential to stabilize food webs by mitigating trophic cascades, although this effect is debated (Bascompte et al., 2005). We calculated omnivory using the function *TrophInd* of the package *NetIndices* package (Kones et al. 2009).

- **Within-module degree:** Within-module degree quantify how a species is connected within its module (Guimerà & Amaral, 2005). Modules (groups of species interacting more together than with other species) are first defined using a spin-glass model and simulated annealing using the function *cluster_spinglass* from the package *igraph* (Csardi & Nepusz 2006). We used the function *calc_topological_roles* of the R package *multiweb* (Saravia 2022) to calculate within-module degree. Species with high within-module degree are important in the stability and functioning of their module.
- **Participation coefficient:** Participation coefficient quantify how much a species interact with species in other module (Guimerà & Amaral 2005). A participation coefficient equal to 0 defines a species solely interacting within its module, whereas 1 defines a species interacting with species with all other modules equally. Species with high participation coefficient connects module together, and therefore, are important in the cohesion of the entire food web. We used the function *calc_topological_roles* of the R package *multiweb* (Saravia, 2022) to calculate participation coefficient.
- **Motif profile:** The motif profile of a species is the frequency of which a species occupies a position within each motif (Milo et al., 2002). We used 3-species motifs, which are the 13 different arrangements 3 species can take. For 3-species motifs, there are 30 different position a species can take. The motif profile has been proposed to be a more complete description of the Eltonian niche of a species (Cirtwill et al., 2018; Stouffer et al., 2012). Some motifs such as “linear chain”, “intraguild predation”, “direct competition”, and “apparent competition” influence the stability of food web (Borrelli, 2015; Stouffer & Bascompte, 2010).

Food web properties:

- **Connectance:** Connectance is defined as the proportion of interactions relative to the total number of possible interactions. Here, since we have a directed network with self-loop allowed, it is calculated as the number of interactions (L) divided by the squared the number of species (S^2). Connectance relates to community sensitivity to perturbations and influence its dynamics (Dunne et al., 2002; Vieira & Almeida-Neto, 2015).
- **Mean and maximum trophic level:** The mean and maximum trophic level are calculated as the average and highest trophic levels among species within a food web. These measures “vertical” diversity, describing on how many trophic levels energy flows through in the food web. Vertical diversity interacts with horizontal diversity (within trophic level diversity) to influence community stability (Duffy et al., 2007; Zhao et al., 2019).
- **Motifs distribution:** The motif distribution is the frequency of motifs (Milo et al. 2002). Here, we focused on the 13 possible 3-species motifs. As mentioned in species role section, some motifs such as “linear chain”, “intraguild predation”, “direct competition”, and “apparent competition” influence the stability of food web (Borrelli, 2015; Stouffer & Bascompte, 2010). We used the function *motif* of the package *igraph* (Csardi & Nepusz 2006) to get the motif distribution of food webs.
- **Food web diameter:** The diameter is the longest of all shortest paths between any two species in the food web. The diameter influence how “efficiently” a perturbation spread thought the

entire food web. We calculated food web diameter using the function *diameter* of the package *igraph* (Csardi & Nepusz 2006).

- **Number of clusters:** The number of clusters (i.e., modules) are the number of non-overlapping highly connected species. We determined the number of clusters using the function *cluster_springlass* from the package *igraph* (Csardi & Nepusz 2006). Perturbations spread more efficiently within a cluster than across clusters. So, a higher number of clusters should increase resistance to perturbations.
- **Modularity:** Modularity quantify how food webs are divided into modules (groups of species interacting more together than with other species). Modularity increases food web persistence (Stouffer & Bascompte, 2011). We determined modules/clusters using the function *cluster_springlass* from the package *igraph* (Csardi & Nepusz 2006).

Appendix S5: Boosted regression tree methods and results

Methods

To make sure that the boosted regression trees (BRT) use the same amount of information than the Bayesian hierarchical generalized linear model, we used the set of 13 trait-based predictor variables described in Appendix S3. We also added a 14th predictor, the order of the predator as predictors, as this trait is used as a group-level effects in the Bayesian model. The predictors were centered and scaled as described in Appendix S3. The training and validation sets are also the exact same as the ones to train and test the Bayesian models.

We used the packages gbm (Greenwell et al., 2019) and dismo (Hijmans et al., 2017) in R to fit the BRT on each of the training set. We set the tree complexity to 5, the learning rate to 0.01, and the bag fraction to 0.5 for each model.

Results

Here, we present the main results of the manuscript:

Predicting trophic interactions:

Table S5.1: Area under the receiver operating curve (AUC) and area under the precision-recall-gain curve (AUPRG) of each BRT model predicting every food web. Values in parenthesis are the AUC and AUPRG obtained by the Bayesian models.

Model	Europe	Pyrenees	North Québec and Labrador	Serengeti
AUC				
Europe	0.98 (0.96)	0.91 (0.89)	0.89 (0.9)	0.82 (0.8)
Pyrenees	0.86 (0.86)	0.96 (0.95)	0.88 (0.85)	0.76 (0.79)
North Québec And Labrador	0.83 (0.82)	0.89 (0.9)	0.96 (0.95)	0.74 (0.75)
Serengeti	0.81 (0.85)	0.69 (0.78)	0.73 (0.77)	0.84 (0.92)
AUPRG				
Europe	0.99 (0.96)	0.81 (0.63)	0.91 (0.91)	0.80 (0.84)
Pyrenees	0.83 (0.9)	0.97 (0.95)	0.88 (0.94)	0.63 (0.84)
North Québec And Labrador	0.87 (0.88)	0.93 (0.9)	0.98 (0.93)	0.68 (0.8)
Serengeti	0.73 (0.84)	0.64 (0.73)	0.73 (0.77)	0.85 (0.93)

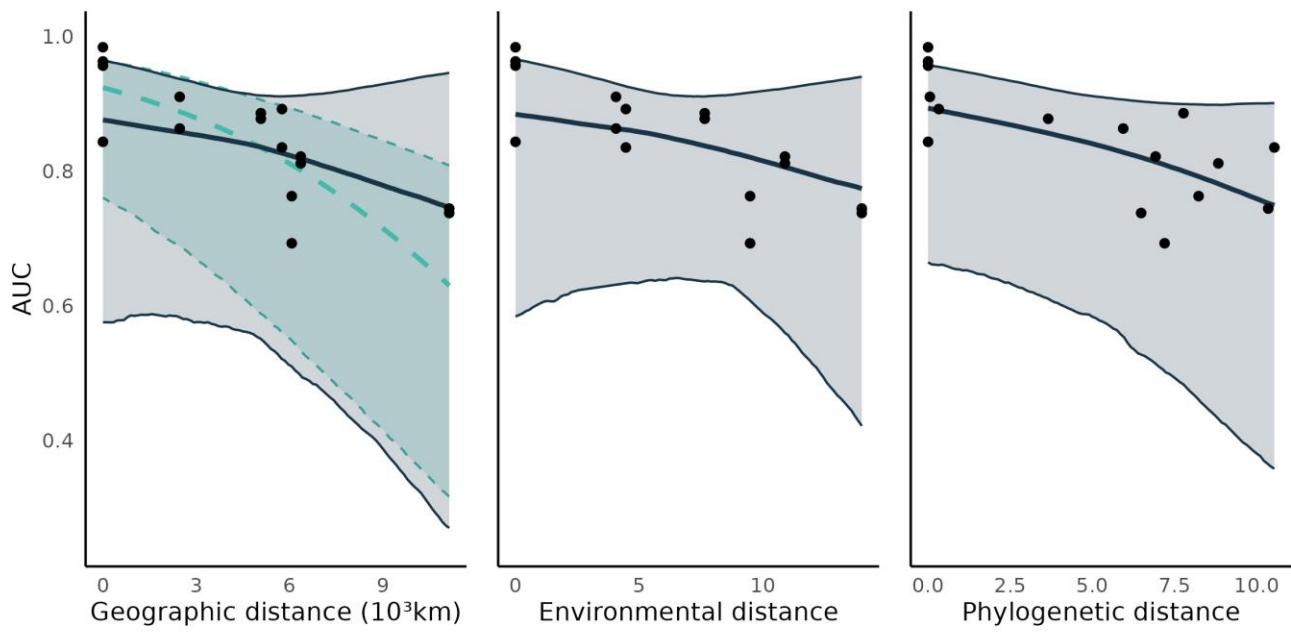


Figure S5.1: Transferability of BRT models. Each point is the predictive performance (AUC) of a BRT-food web predicted combination (16 combinations; Table S5.1). The trend lines are the median effects with their 95% credible interval constructed with the posterior predictive distribution of geographic, environmental, and phylogenetic distances on predictive performance. Shown are the total (turquoise; -0.63, CrI 95% = [-0.28, -0.96]) and partial (dark blue; -0.28, CrI 95% = [-1.12, 0.59]) effects of geographic distance, and the direct (controlling for geographic distance) effects of environmental (-0.27, CrI 95% = [-1.12, 0.59]) and phylogenetic distances (-0.41, CrI 95% = [-1.07, 0.20]).

Predicting species trophic role:

Similar to the Bayesian hierarchical models, BRTs failed to accurately predict species trophic roles (Figure S5.2, compared to Figure 4). The number of prey, within-module degree, participation coefficient, and the frequency of most motif positions were slightly better predicted by the bayesian hierarchical models, especially for within food web predictions. The remaining trophic roles were similarly well predicted by the BRTs and the Bayesian hierarchical models.

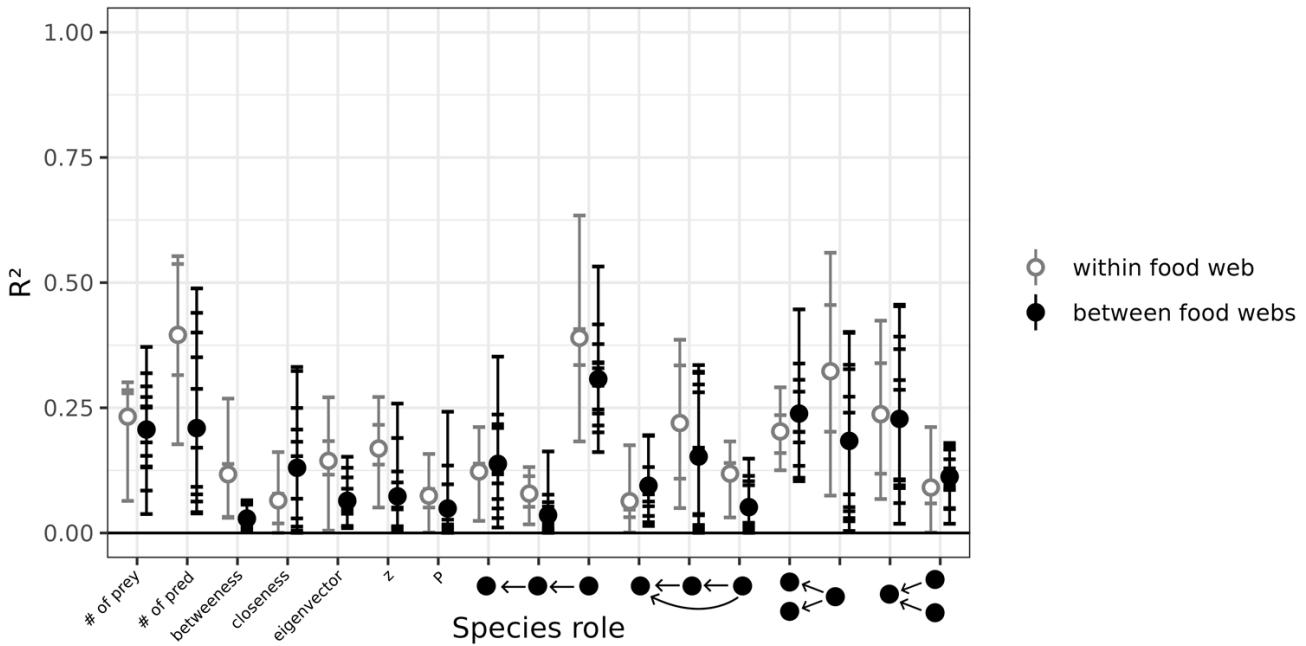


Figure S5.2: Performance of boosted regression tree models to predict species' role measured as the proportion of the variance in trophic positions explained by the models (R^2). From left to right, the figure shows the number of prey, number of predators, betweenness, closeness, eigenvector centrality, within-module degree (z), participation coefficient (P), times the species is at a specific position in the linear chain (predator – consumer - resource), intraguild predation (omnivore – consumer - resource), direct competition (2 consumers – 1 resource), and apparent competition motifs (1 consumer – 2 resources). The short horizontal lines highlight the R^2 for each role, model, and food web predicted combination. Grey open dots are the mean R^2 for within food web predictions. Full black dots are the mean R^2 for between food web predictions.

Predicting food web properties:

Overall, some food web properties are better predicted by the boosted regression trees than by the Bayesian models (Figure S5.3, compared to Figure 5). Connectance is especially much less overpredicted (relative error of 0.6 instead of 1.10), and the frequency of intraguild predation motifs is also much better predicted. However, we failed calculate mean and maximum trophic levels. This is likely due to the prediction of many trophic cycles (e.g., two species feeding on each other), which can create problem when defining trophic levels. The other properties predicted by the BRT are similarly biased to the properties predicted by the Bayesian models.

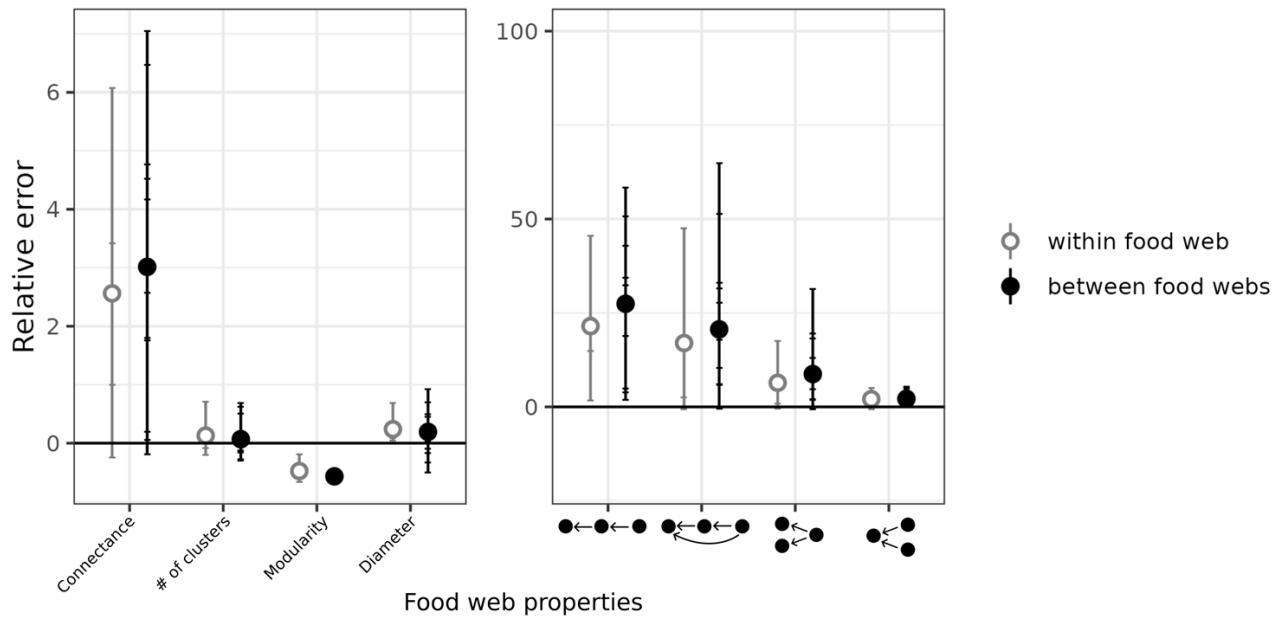


Figure S5.3: Relative error of global food web properties predicted by the boosted regression trees. From left to right, the figure shows the relative error for connectance, number of clusters, modularity, diameter, number of linear chain (predator – consumer - resource), intraguild predation (omnivore – consumer - resource), direct competition (2 consumers – 1 resource), and apparent competition (1 consumer – 2 resources) motifs . Short horizontal lines are the relative errors for each property, model, and food web predicted combination. Grey open dots are the mean relative errors for within food web predictions and full black dots are the mean relative errors for between food web predictions.

Appendix S6: Convergence of the predictive models

We evaluated convergence and adequacy of the four predictive models using rank plots, posterior predictive checks, and the rank-normalized potential scale reduction factor on split chains.

Rank trace plots are similar to trace plots and are used to assess the mixing of the chains (Gelman et al., 2013). With trace rank plots, we check if the MCMC sampler explores the parameter space adequately. For each iteration (x axis), the parameter varies (y axis), and the chains are ranked from 1 (lowest value of the parameter) to 4 (highest value of the parameter). In the ideal case, we want that the chains to change rank more or less randomly (no chain staying at a given rank more than others). Conversely, if a given chain stays in the same rank for many iterations in a row or is more often at the same rank than other chains, it is sign of a problem with mixing. Here, all the rank plots suggest adequate mixing of the chains (Figure S6.1, S6.4, S6.7, S6.10).

Posterior predictive checks compare the distribution of the observed data to the distribution of the posterior predictive distribution (Gelman et al., 2013). A discrepancy in distributions can be caused by a model missfit. A model that fits correctly should replicate the observed distribution of interactions/non interactions (counts on the y axis, and 0 and 1 on the x axis) of each “group” of the model (in our models, we have varying parameters for the order of the predator, so an histogram for each predator order). A problematic model would fail to replicate the distributions. Here, the four models replicate the distributions of the observed data (Figure S6.2, S6.5, S6.8, S6.11).

Finally, the rank-normalized potential scale reduction factor on split chains (Rhat) is similar to the potential scale reduction factor, and is a summary statistics evaluating model convergence (Vehtari et al., 2021). Rhat close to 1 indicate a correctly converged model, and a Rhat larger than 1 ($Rhat > 1.1$) would indicate that the model did not converge correctly. Here, we have a large number of Rhat (one for each slope and intercept of each predator order), so we are plotting the histogram of Rhat. Each histogram is suggesting the adequate convergence of the models (Figure S6.3, S6.6, S6.9, S6.12).

Northern Québec and Labrador model:

- Trace rank plots of the population-level effects:

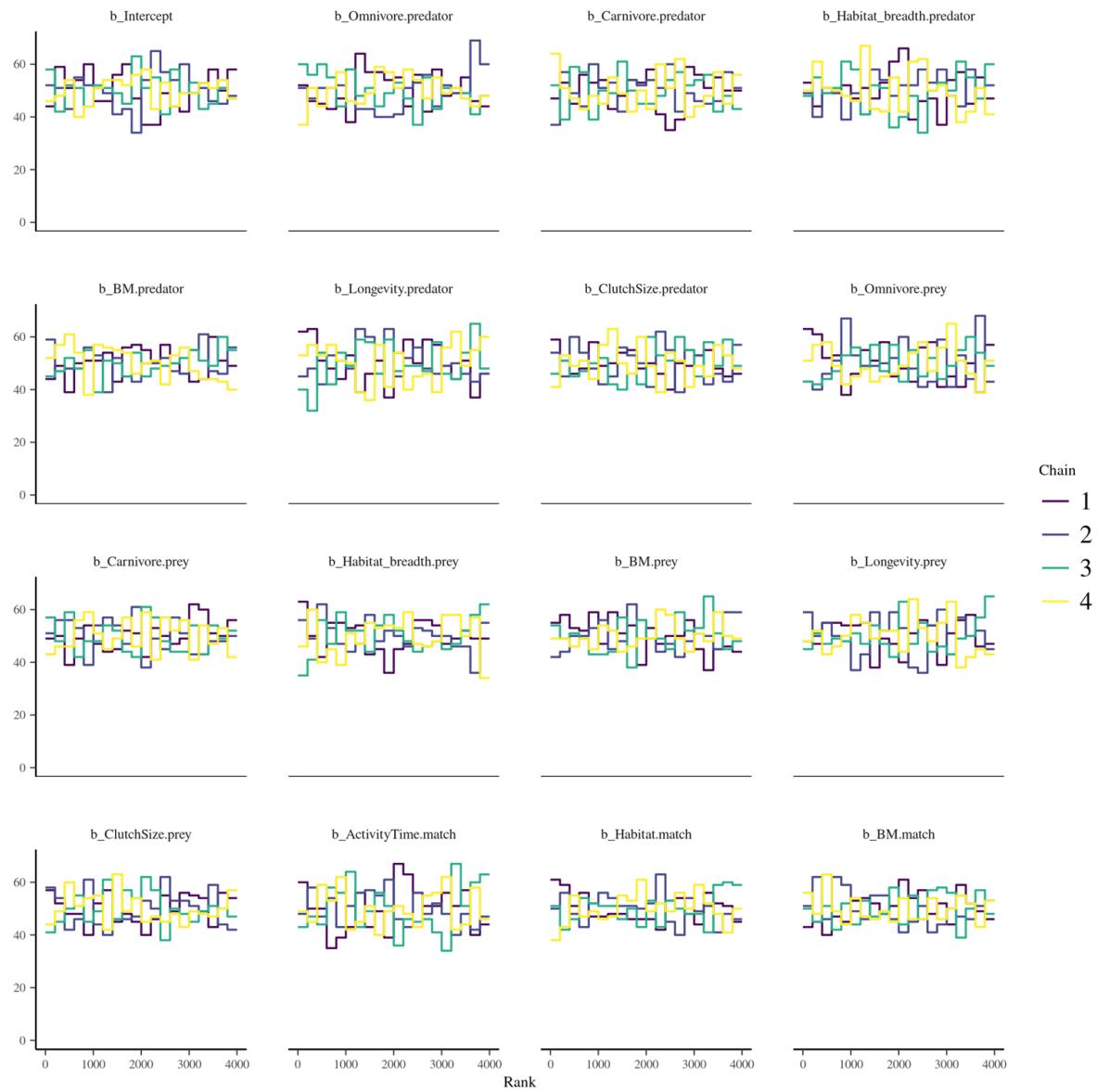


Figure S6.1: Trace rank plots of the population-level effects for the model calibrated on the Northern Québec and Labrador food web.

- Posterior predictive checks:

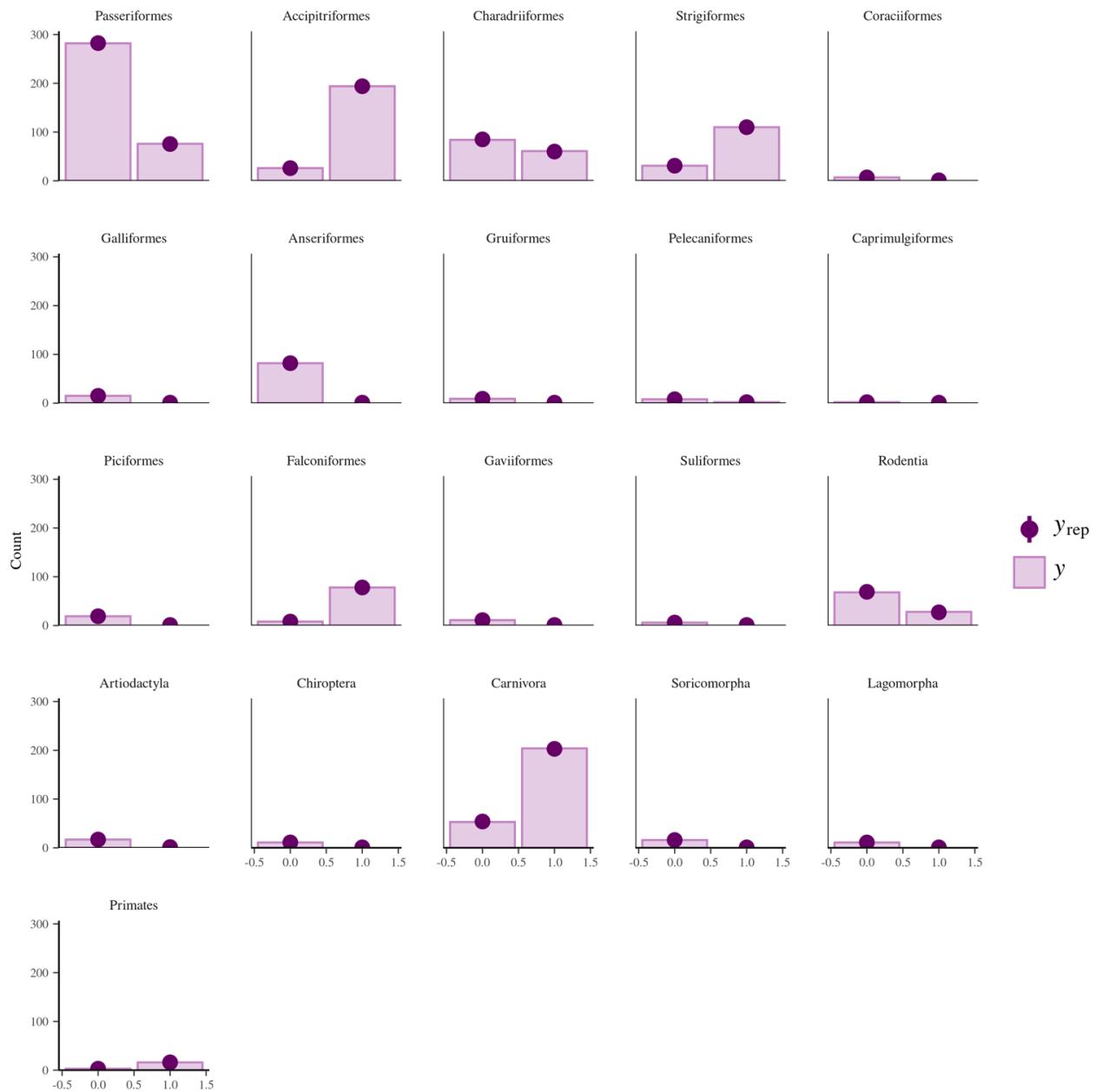


Figure S6.2: Comparison of the posterior predictive distribution (y_{rep} ; purple points) and the observed distribution (y ; purple bars) of the model calibrated on the Northern Québec and Labrador food web.

- Rank-normalized potential scale reduction factors on split chains:

**Distribution of rank-normalized potential
scale reduction factors for the
Northern Québec and Labrador model**

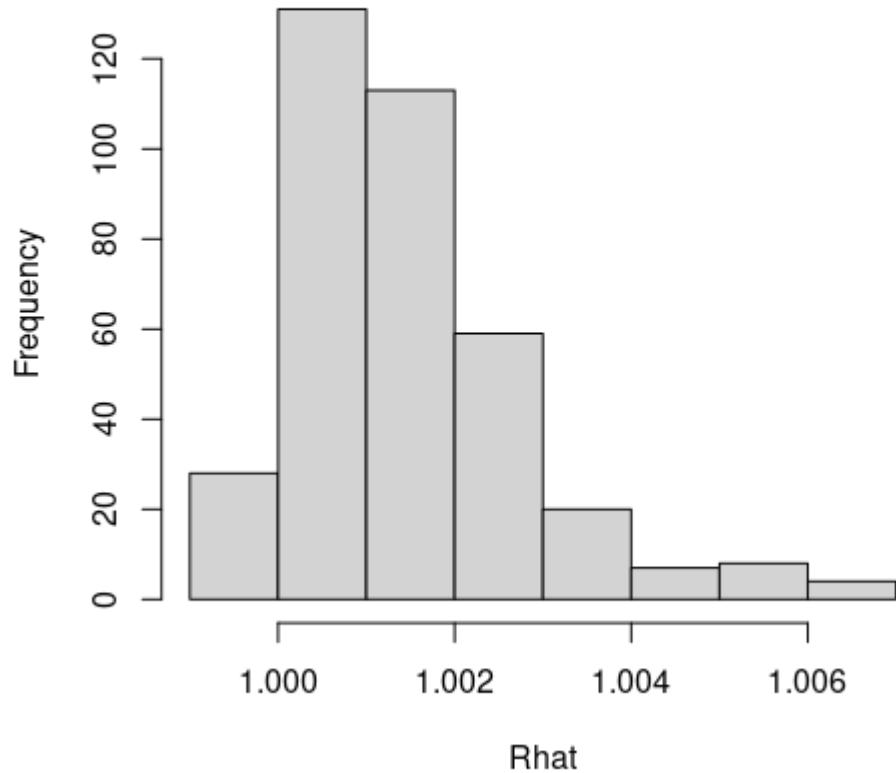


Figure S6.3: Distribution of the rank-normalized potential scale reduction factors on split chains for the Northern Québec and Labrador food web

Europe model:

- Trace rank plots of the population-level effects:

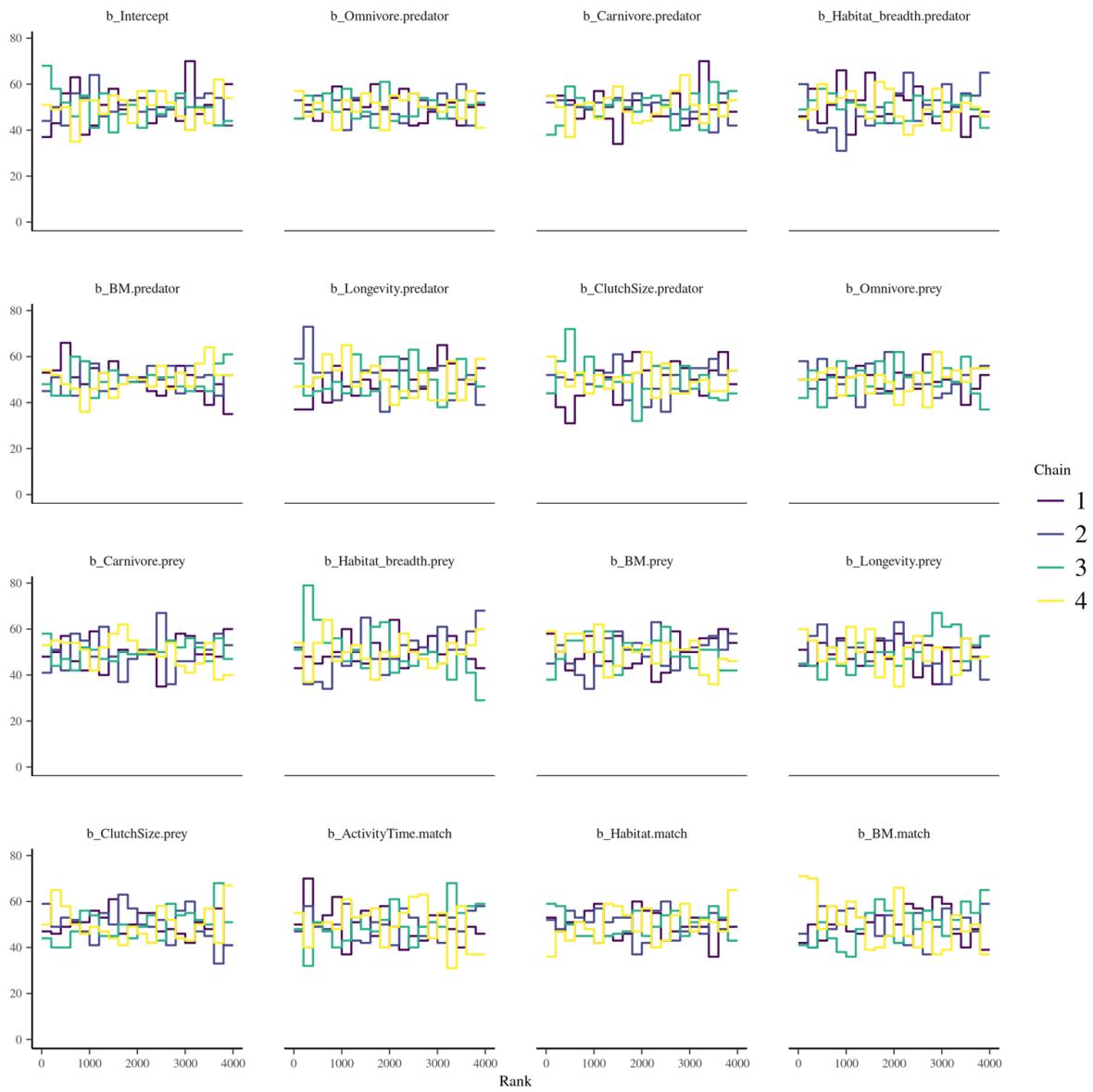


Figure S6.4: Trace rank plots of the population-level effects for the model calibrated on the European food web of terrestrial vertebrates.

- Posterior predictive checks:

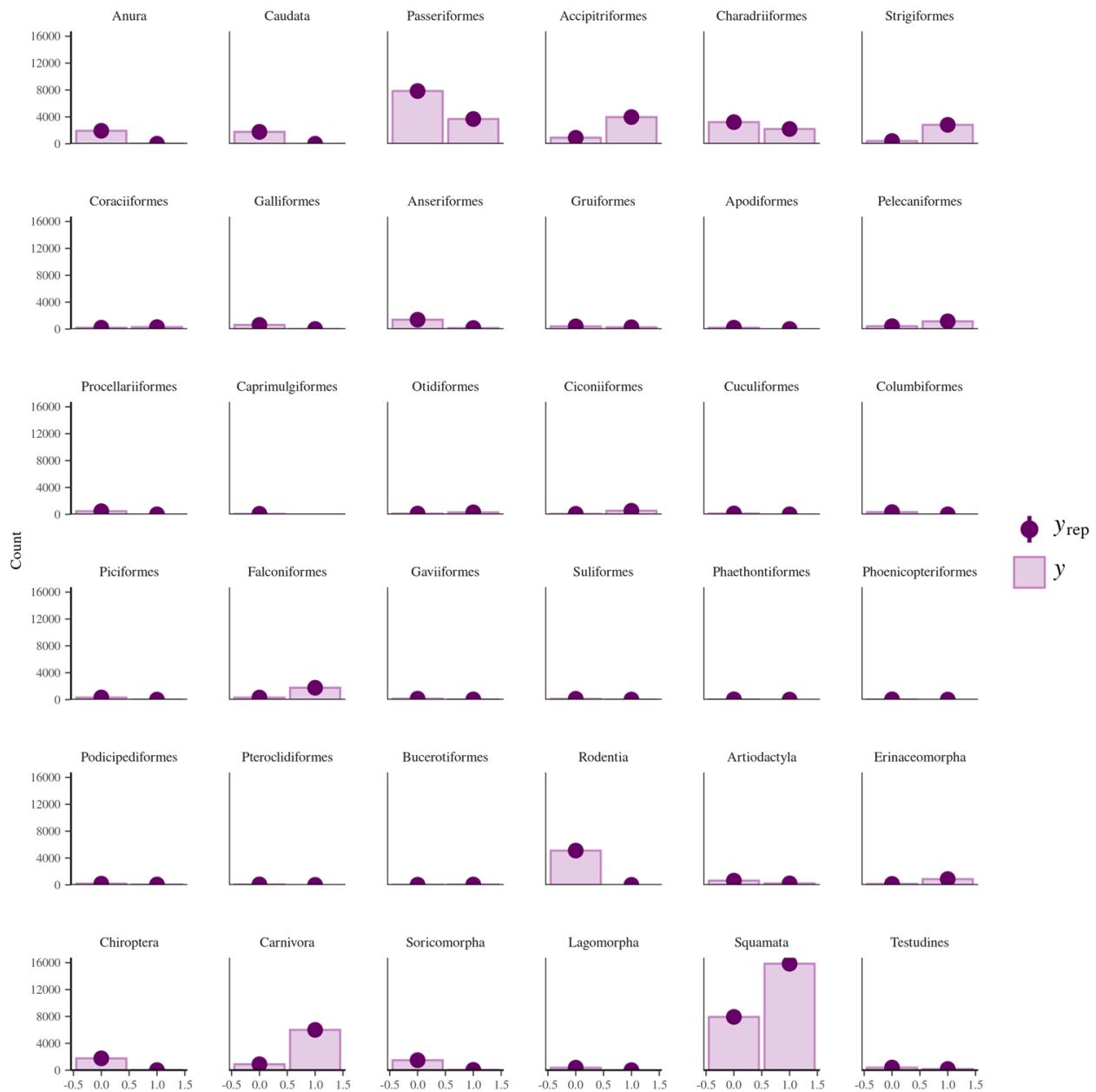


Figure S6.5: Comparison of the posterior predictive distribution (y_{rep} ; purple points) and the observed distribution (y ; purple bars) of the model calibrated on the food web of European terrestrial vertebrates.

- Potential scale reduction factors:

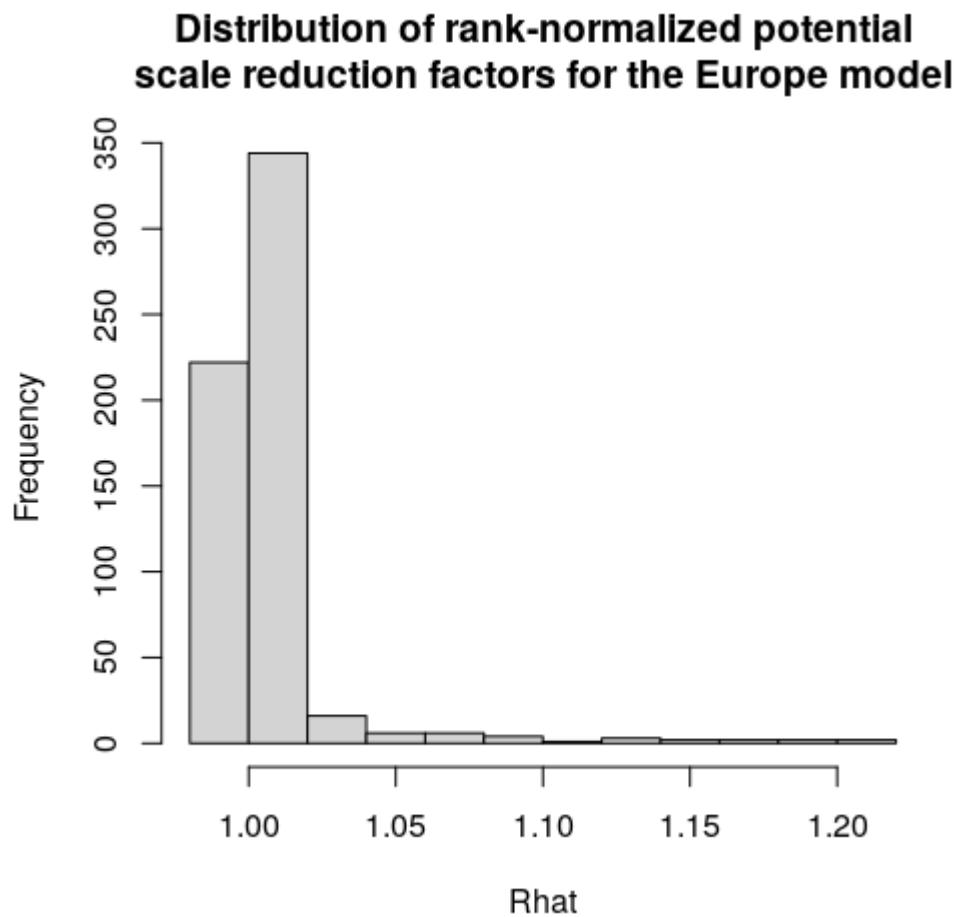


Figure S6.6: Distribution of the rank-normalized potential scale reduction factors on split chains for the food web of European terrestrial vertebrates.

Pyrenees model:

- Rank plots of the population-level effects:

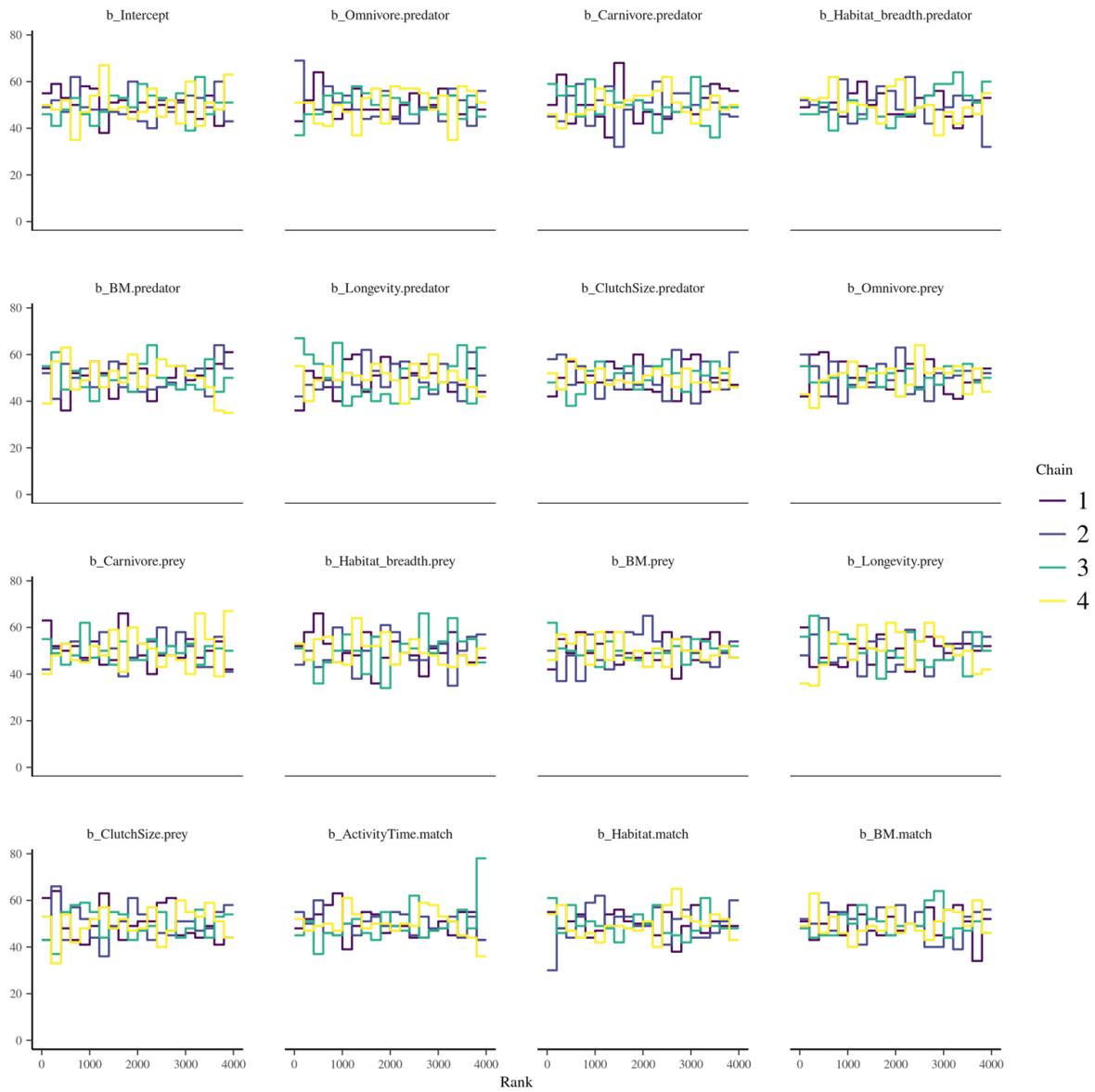


Figure S6.7: Trace rank plots of the population-level effects for the model calibrated on the Pyrenees food web.

- Posterior predictive checks:

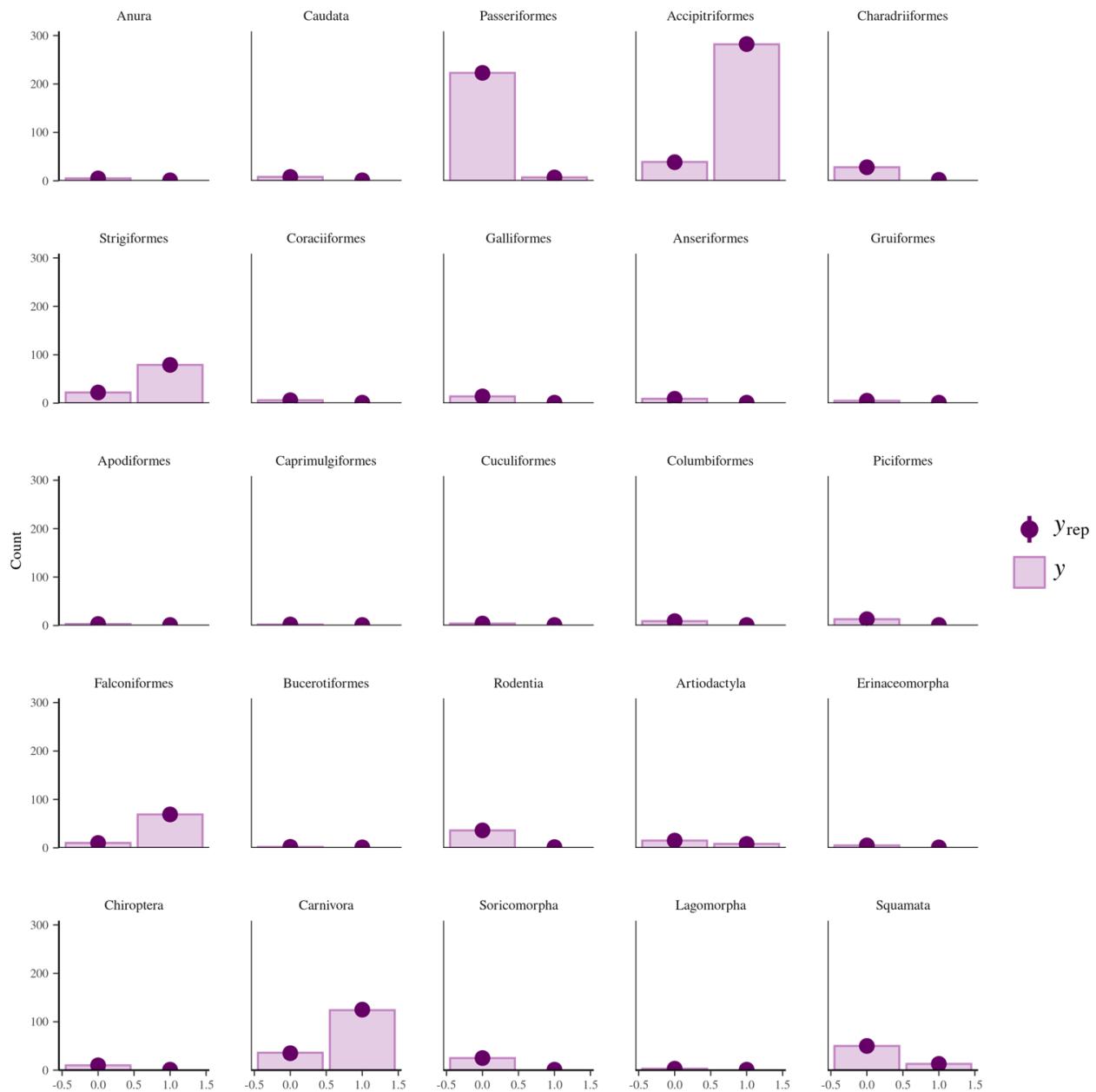


Figure S6.8: Comparison of the posterior predictive distribution (y_{rep} ; purple points) and the observed distribution (y ; purple bars) of the model calibrated on the Pyrenees food web.

- Potential scale reduction factors:

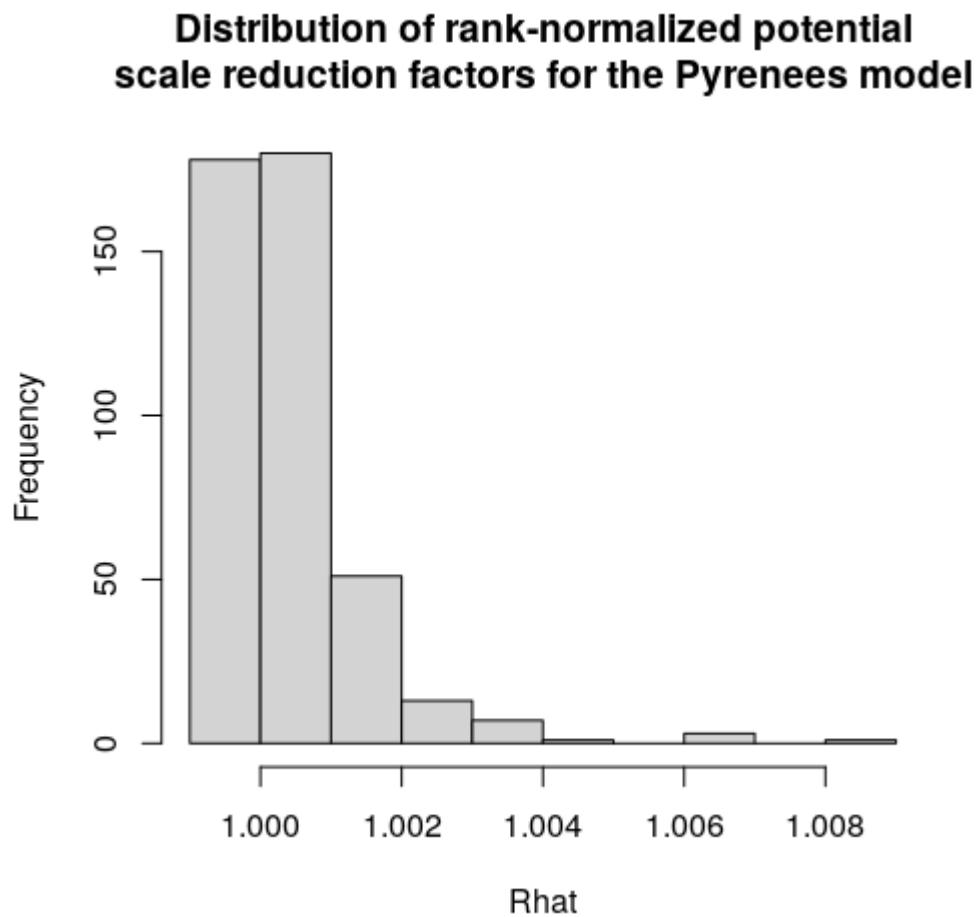


Figure S6.9: Distribution of the rank-normalized potential scale reduction factors on split chains for the Pyrenees food web

Serengeti model:

- Rank plots of the population-level effects:

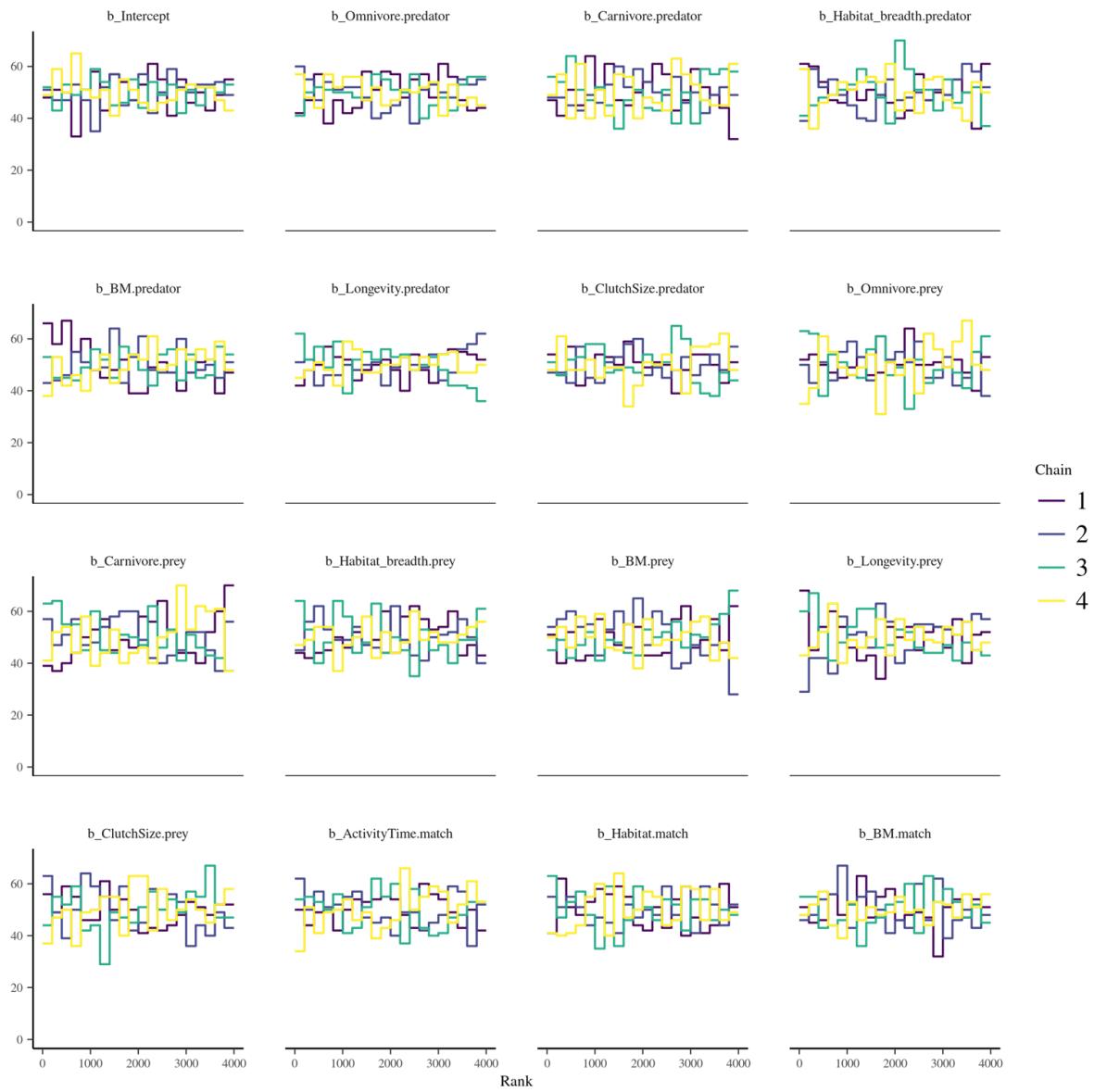


Figure S6.10: Trace rank plots of the population-level effects for the model calibrated on the Serengeti food web.

- Posterior predictive checks:



Figure S6.11: Comparison of the posterior predictive distribution (y_{rep} ; purple points) and the observed distribution (y ; purple bars) of the model calibrated on the Serengeti food web.

- Potential scale reduction factors:

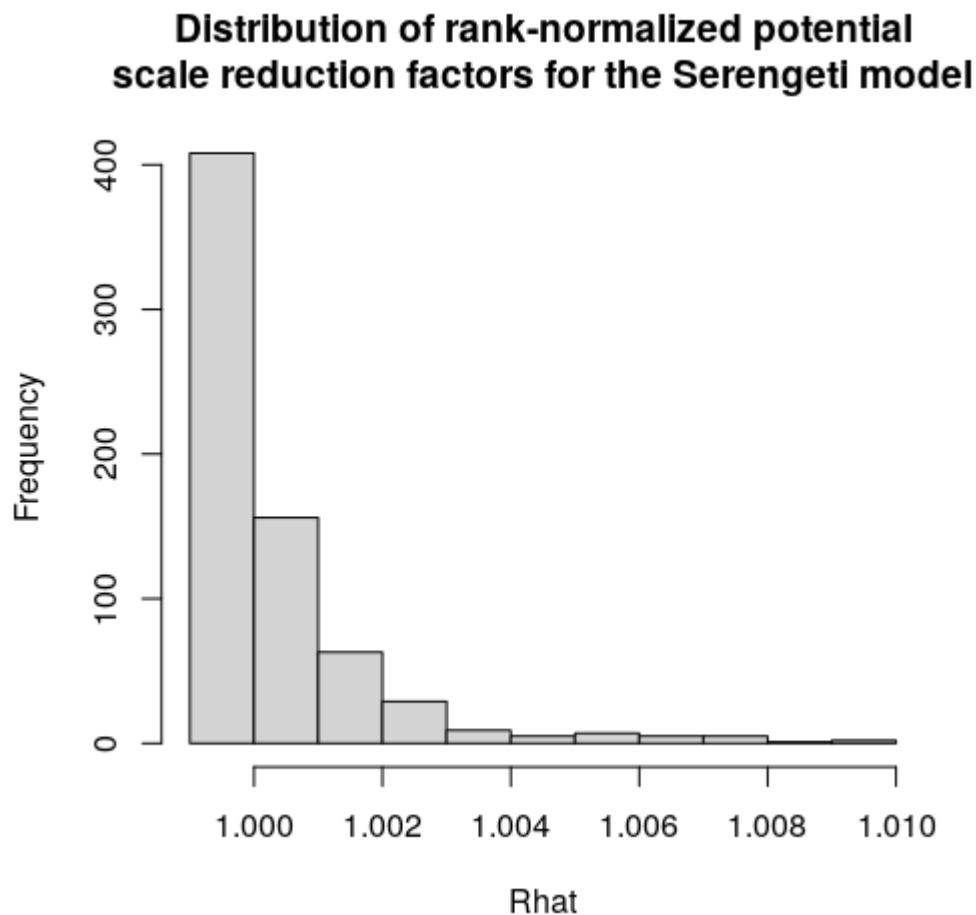


Figure S6.12: Distribution of the rank-normalized potential scale reduction factors on split chains for the Serengeti food web

Appendix S7: Predicting species interactions extended results.

In this appendix, we present additional results on model performance. More precisely, we report additional performance metrics: the area under the precision-recall-gain curve, true positive rate, true negative rate, positive predictive value, and negative predictive value.

Like the area under the receiver operating characteristic curve (AUC), the **area under the precision-recall-gain curve (auprg)** is a threshold-independent metric that calculates the area under a curve. However, instead of using trade-off between false positive and true positive rate across thresholds, auprg uses the trade-off between positive predictive value (also called precision) and the true positive rate (also called recall) across thresholds. The auprg transform the precision-recall curve into an harmonic scale to fix some issues with the standard precision-recall curve (e.g., non-universality of baseline, non-linear interpolation). auprg varies between 0 for random predictions and 1 for perfect predictions. For more information, we suggest referring to Flach & Kull (2015).

True positive rate (tpr) is the proportion of observed positives (interactions) predicted as positives (interactions). Instead of applying an arbitrary threshold on the predicted probabilities of interactions, we calculated the tpr for 100 samples of the posterior distribution.

True negative rate (tnr) is the proportion of observed negatives (non-interactions) predicted as negatives (non-interactions). Instead of applying an arbitrary threshold on the predicted probabilities of interactions, we calculated tnr for 100 samples of the posterior distribution.

Positive predictive value (ppv) is the proportion of predicted positives (interactions) that are observed positives (interactions). It measures how confident we can be that a predicted interaction is an actual interaction. Instead of applying an arbitrary threshold on the predicted probabilities of interactions, we calculated ppv for 100 samples of the posterior distribution.

Negative predictive value (npv) is the proportion of predicted negatives (non-interactions) that are observed negatives (non-interactions). It measures how confident we can be that a predicted non-interaction is an actual non-interaction. Instead of applying an arbitrary threshold on the predicted probabilities of interactions, we calculated the true negative rate for 100 samples of the posterior distribution.

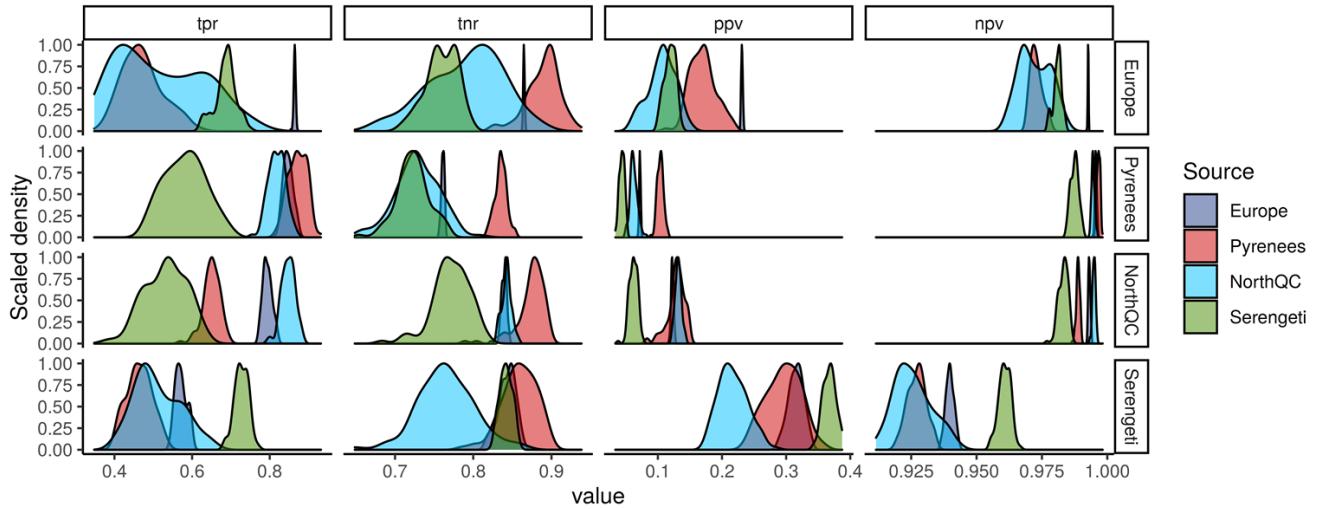


Figure S7.1: Scaled density distribution of true positive rate (tpr), true negative rate (tnr), positive predictive value (ppv), and negative predictive value (npv) of each model (Source) predicting every food webs (rows). tpr is the proportion of correctly predicted interactions, tnr is the proportion of correctly predicted non-interactions, ppv is the proportion of predicted interactions that are observed interactions, and npv is the proportion of predicted non-interactions that are observed non-interactions. We calculated tpr, tnr, ppv, and npv with 100 samples of the posterior distribution of the entire food webs.

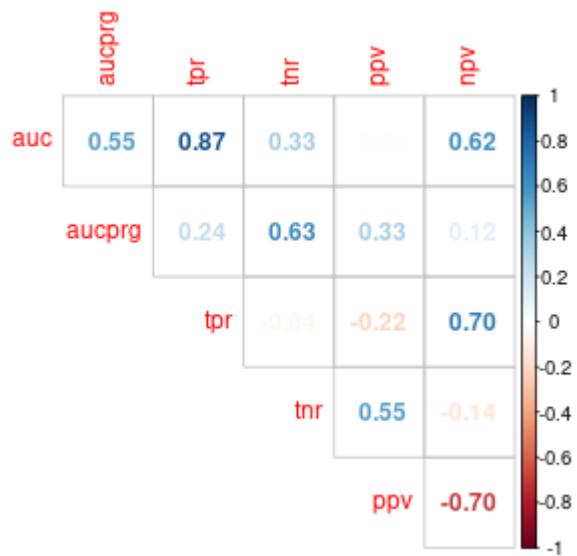


Figure S7.2: Correlation between performance metrics for predicting the interactions of the entire food webs. *auc* is the area under the receiver operating curve, *aucpr* is the area under the precision-recall curve, *tpr* is the true positive rate, *tnr* is the true negative rate, *ppv* is the positive predictive value, and *npv* is the negative predictive value

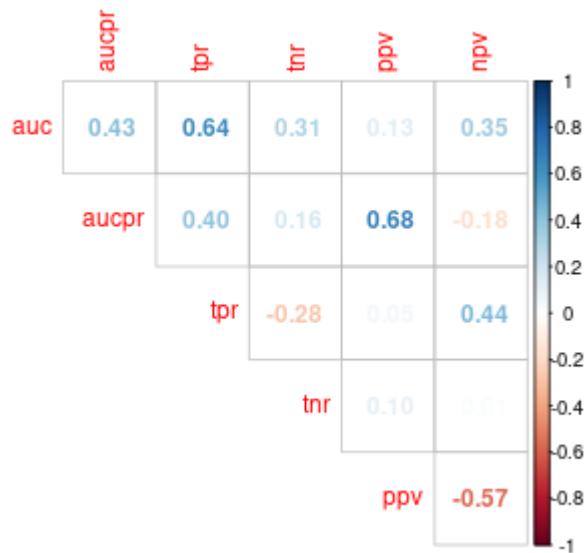


Figure S7.3: Correlation between performance metrics for predicting the interactions of each species within a food web. See caption of figure S7.2 for abbreviations.

Model performance in relationship with distance:

Here, we present the outputs of the model relating model performance to geographic, phylogenetic, and environmental distances:

- Model 1: Total effect of geographic distance on AUC
 $\text{logitauc} \sim \text{geo.dist_sc} + (1|\text{Model}) + (1|\text{Food web})$
fixed effects:
Intercept: Estimate = 1.83, 95% CrI = [1.24, 2.26]
 geo.dist_sc : Estimate = -0.57, 95% CrI = [-0.78, -0.36]
random effect:
 $\text{sd}(1|\text{Model})$: Estimate = 0.25, 95% CrI = [0.01, 0.91]
 $\text{sd}(1|\text{Food web})$: Estimate = 0.27, 95% CrI = [0.01, 1.11]
- Model 2: Partial effect of geographic distance on AUC
 $\text{logitauc} \sim \text{geo.dist_sc} + \text{phylo.dist_sc} + \text{env.dist_sc} + (1|\text{Model}) + (1|\text{Food web})$
fixed effects:
Intercept: Estimate = 1.78, 95% CrI = [0.70, 2.25]
 geo.dist_sc : Estimate = -0.26, 95% CrI = [-0.87, 0.35]
 phylo.dist_sc : Estimate = -0.29, 95% CrI = [-0.74, 0.13]
 env.dist_sc : Estimate = -0.10, 95% CrI = [-0.74, 0.60]
random effect:
 $\text{sd}(1|\text{Model})$: Estimate = 0.39, 95% CrI = [0.01, 2.06]
 $\text{sd}(1|\text{Food web})$: Estimate = 0.24, 95% CrI = [0.01, 1.01]
- Model 3: Effect of phylogenetic distance on AUC after controlling for geographic distance
 $\text{logitauc} \sim \text{phylo.dist_sc} + \text{geo.dist_sc} + (1|\text{Source}) + (1|\text{Target})$
fixed effects:
Intercept: Estimate = 1.82, 95% CrI = [1.13, 2.24]
 geo.dist_sc : Estimate = -0.34, 95% CrI = [-0.68, 0.01]
 phylo.dist_sc : Estimate = -0.31, 95% CrI = [-0.69, 0.06]
random effect:
 $\text{sd}(1|\text{Model})$: Estimate = 0.30, 95% CrI = [0.01, 1.07]
 $\text{sd}(1|\text{Food web})$: Estimate = 0.24, 95% CrI = [0.01, 0.98]
- Model 4: Effect of environmental distance on AUC after controlling for geographic distance
 $\text{logitauc} \sim \text{env.dist_sc} + \text{geo.dist_sc} + (1|\text{Source}) + (1|\text{Target})$
fixed effects:
Intercept: Estimate = 1.83, 95% CrI = [1.12, 2.24]
 geo.dist_sc : Estimate = -0.29, 95% CrI = [-0.90, 0.31]
 env.dist_sc : Estimate = -0.30, 95% CrI = [-0.90, 0.32]
random effect:
 $\text{sd}(1|\text{Model})$: Estimate = 0.23, 95% CrI = [0.01, 0.92]
 $\text{sd}(1|\text{Food web})$: Estimate = 0.25, 95% CrI = [0.01, 1.12]

Model performance in relationship with species class:

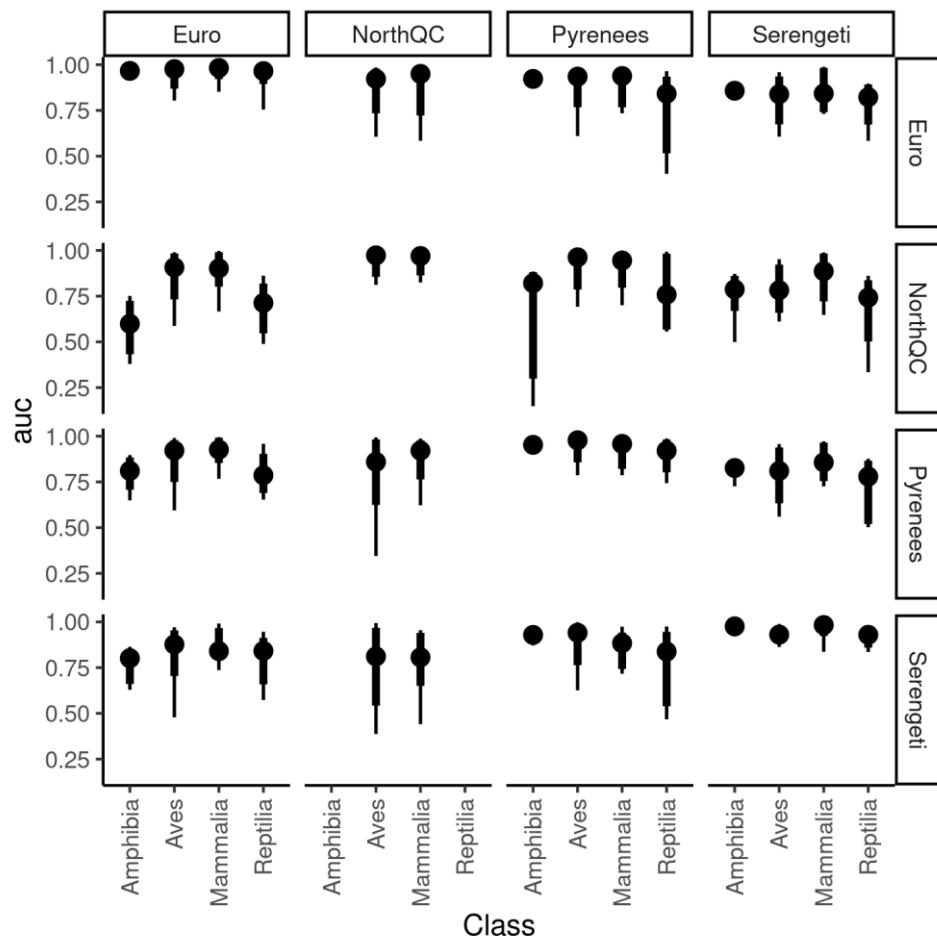


Figure S7.4: Class specific performance of predicting species interaction for each combination of model (rows) and food web predicted (columns). The point intervals represent the median, the 80%, and 95% intervals of auc across species within each class.

Appendix S8: Predicting species role extended results

For each of the combination of species role metric considered, model and food web predicted, we fitted a simple linear regression using OLS through the function `lm()` in R. The response variable is the predicted values, and the predictor variables are the observed values. We extracted the slopes, intercept, and R^2 of each regression. R^2 measures the correlation between the predicted and observed values. The intercept measures the average biases: intercept > 0 represent average overestimation, intercept < 0 means average underestimation. The slope measures the homogenization or heterogenization of the values: slope > 1 means species have more variable values in the predictions, slope < 1 means that species have less variable values in the predictions. The following figures show the R^2 (Figure S8.1) and the fitted line of the linear regression (Figure S8.2-20) to ease the interpretation of the results.

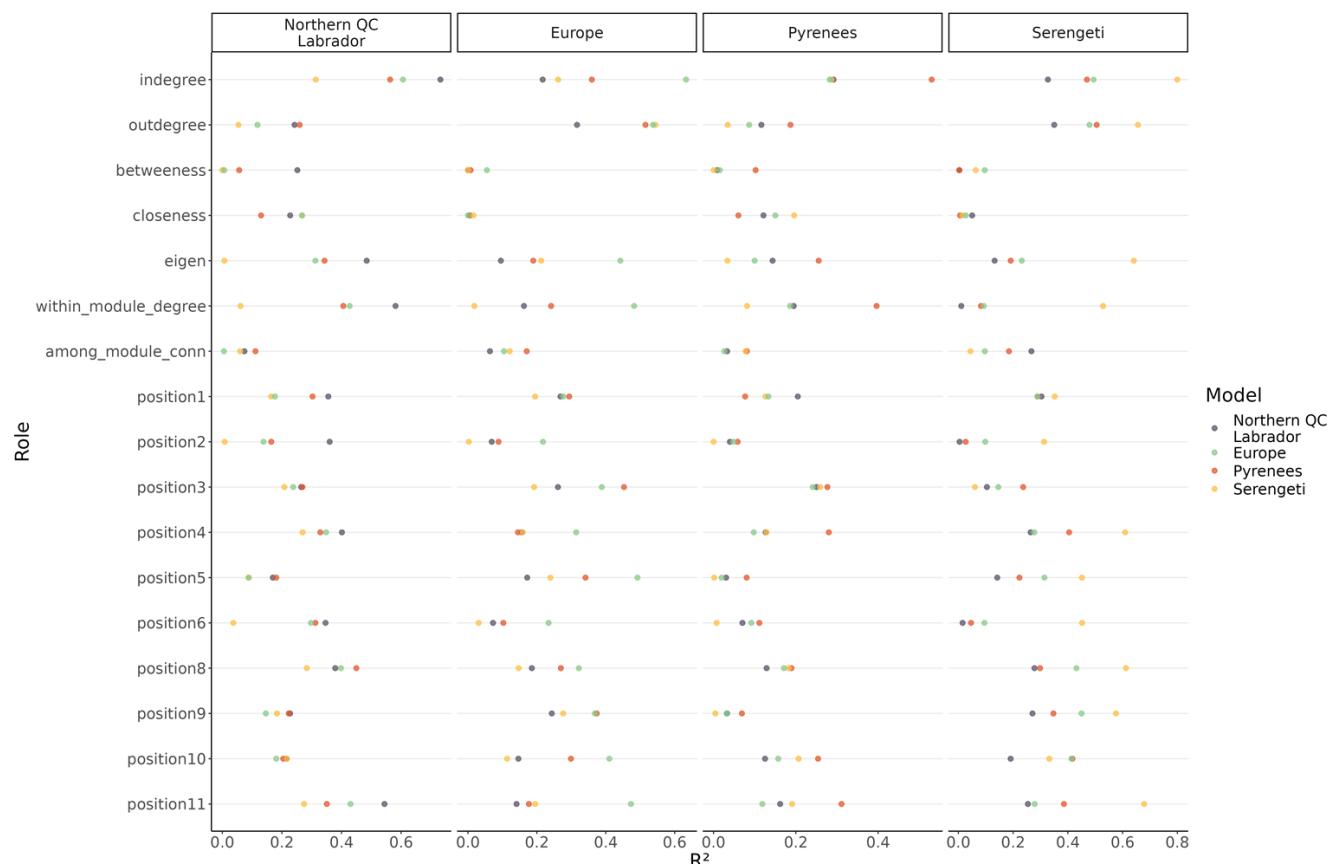


Figure S8.1: Predictive performance of each model (colors) to predict species role for every food web (column). Performance is measured with the coefficient of determination (R^2). From top to bottom: Number of prey, number of predator, betweenness, closeness, eigenvector centrality, within-module degree, participation coefficient, frequency of predator position in a linear chain motif, consumer position in a linear chain motif, resource in a linear chain motif, omnivore in a intraguild predation motif, consumer in a intraguild predation motif, resource in a intraguild predation motif, predation in a direct competition motif, prey in a direct competition motif, predator in a apparent competition motif, and prey in a apparent competition motif.

Number of prey

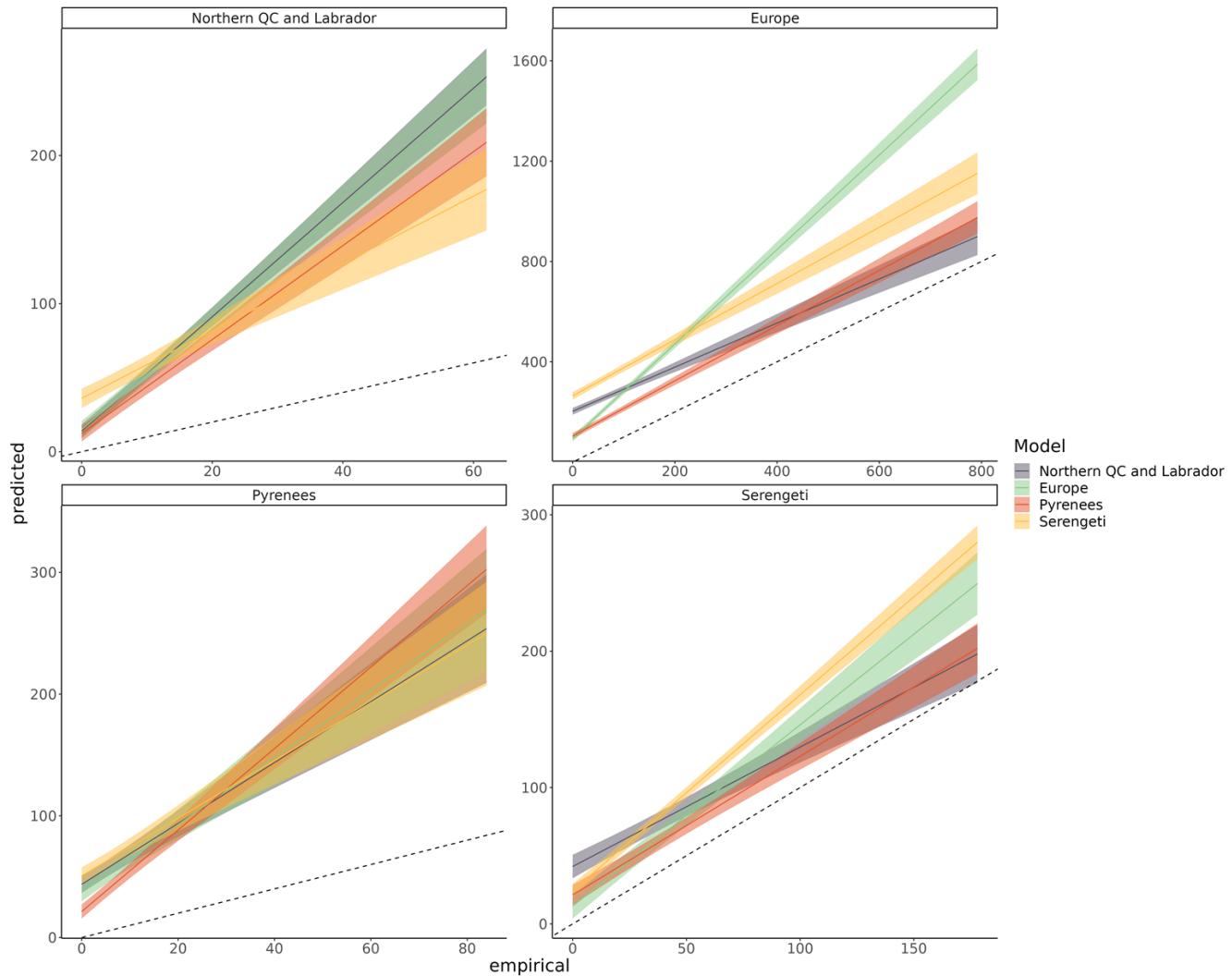


Figure S8.2: Linear regression comparing the predicted number of prey and empirical number of prey of species for each model predicting every food webs. The main bias is that the number of prey is overpredicted (intercept > 0). In many cases, the bias is more important for generalist predators ($slope > 1$). Interestingly, the Serengeti model tend to better predict generalist species, but less specialist species (compared to other models). Its overall performance is still much lower for predicting the number of prey in the Northern Quebec food web (Figure S8.1) given there are many more specialists than generalists.

Number of predators

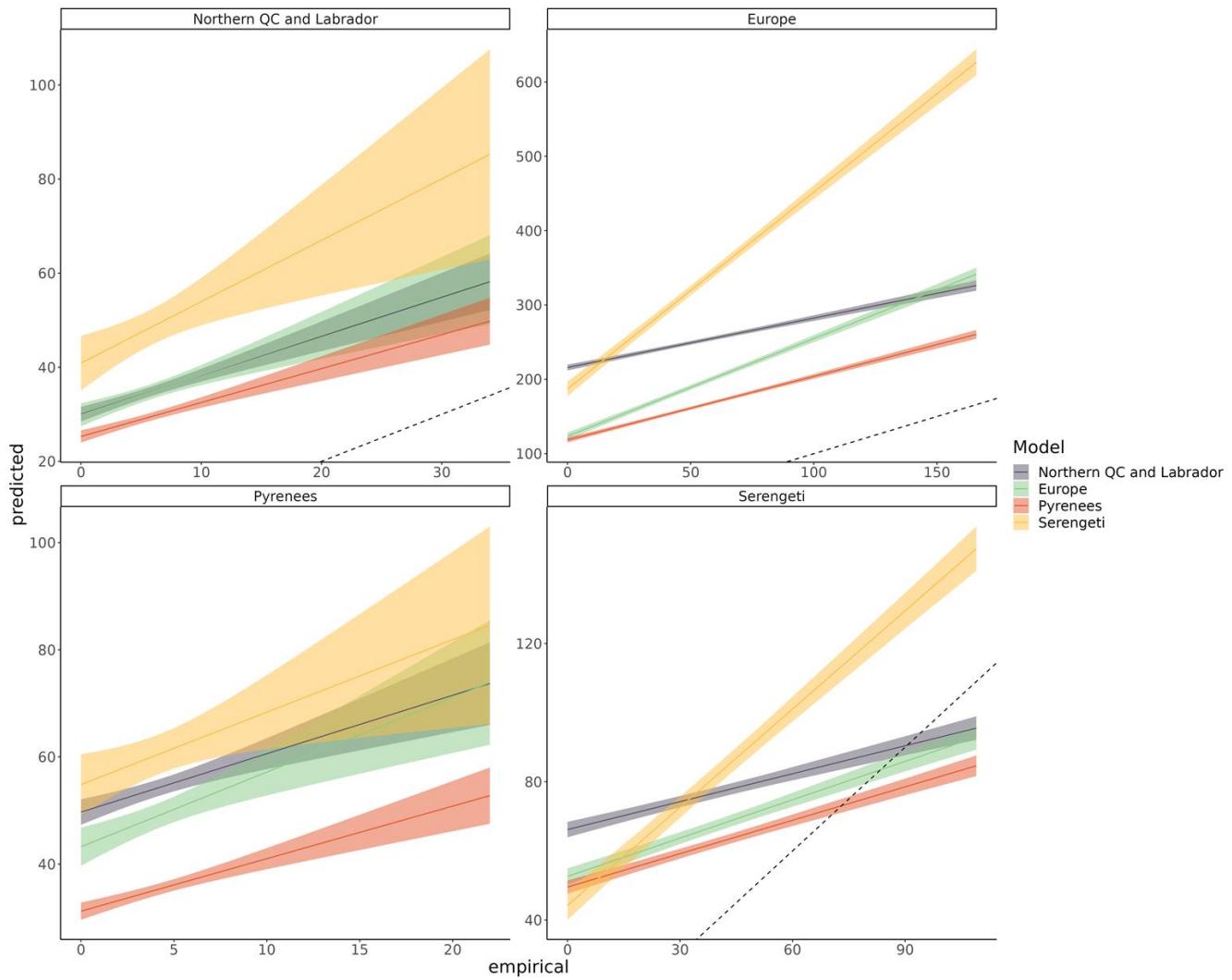


Figure S8.3: Linear regression comparing the predicted number of predators and empirical number of predators of species for each model predicting every food webs. The main bias is that the number of prey is overpredicted (intercept > 0). In many cases, the bias is more important for generalist predators (slope < 1).

Betweenness

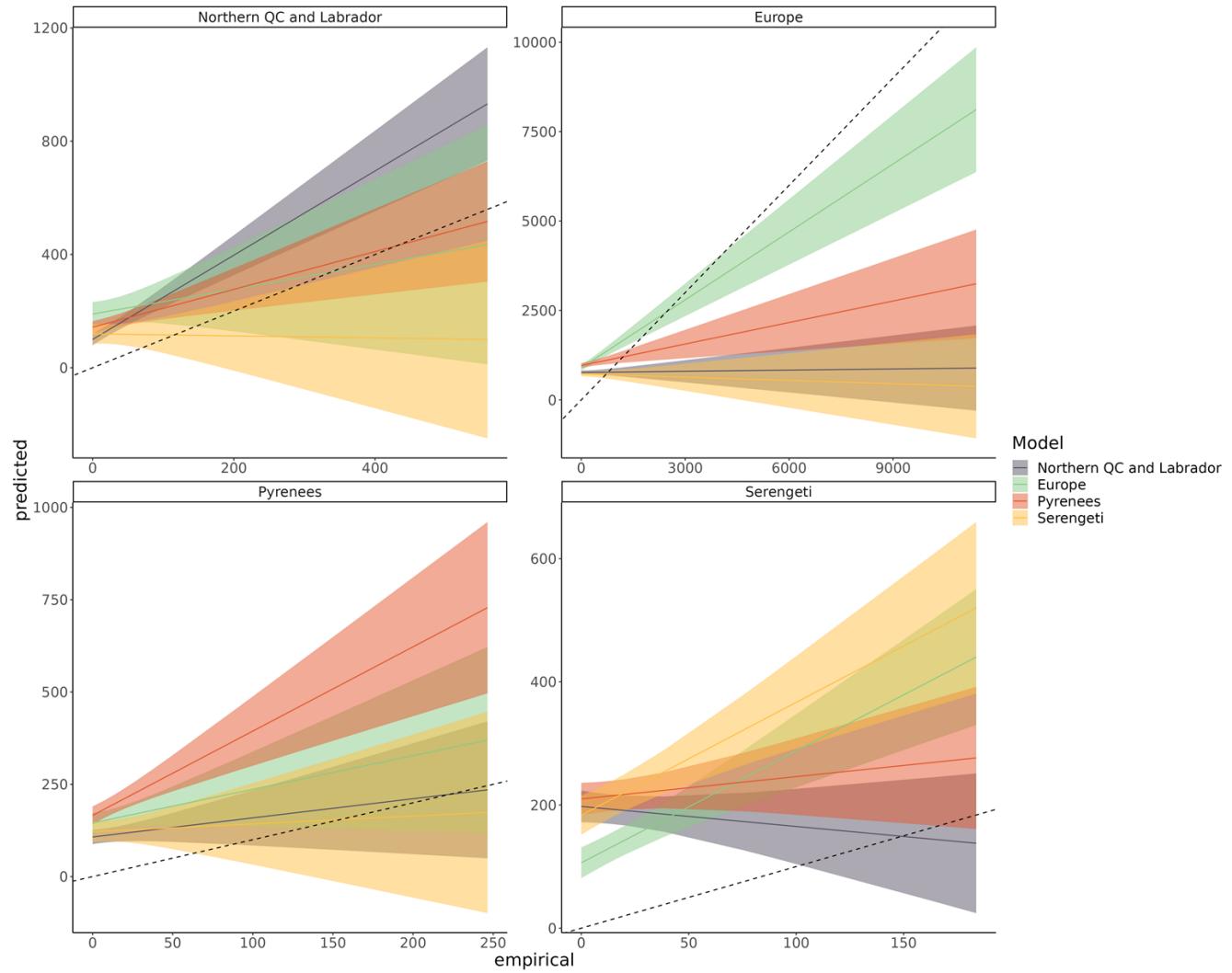


Figure S8.4: Linear regression comparing the predicted betweenness centrality and empirical betweenness centrality of species for each model predicting every food webs. The main bias is that predicted betweenness is more homogeneous across species (slope < 1; overpredicted for low end of the range, and underpredicted for high end of the range).

Closeness

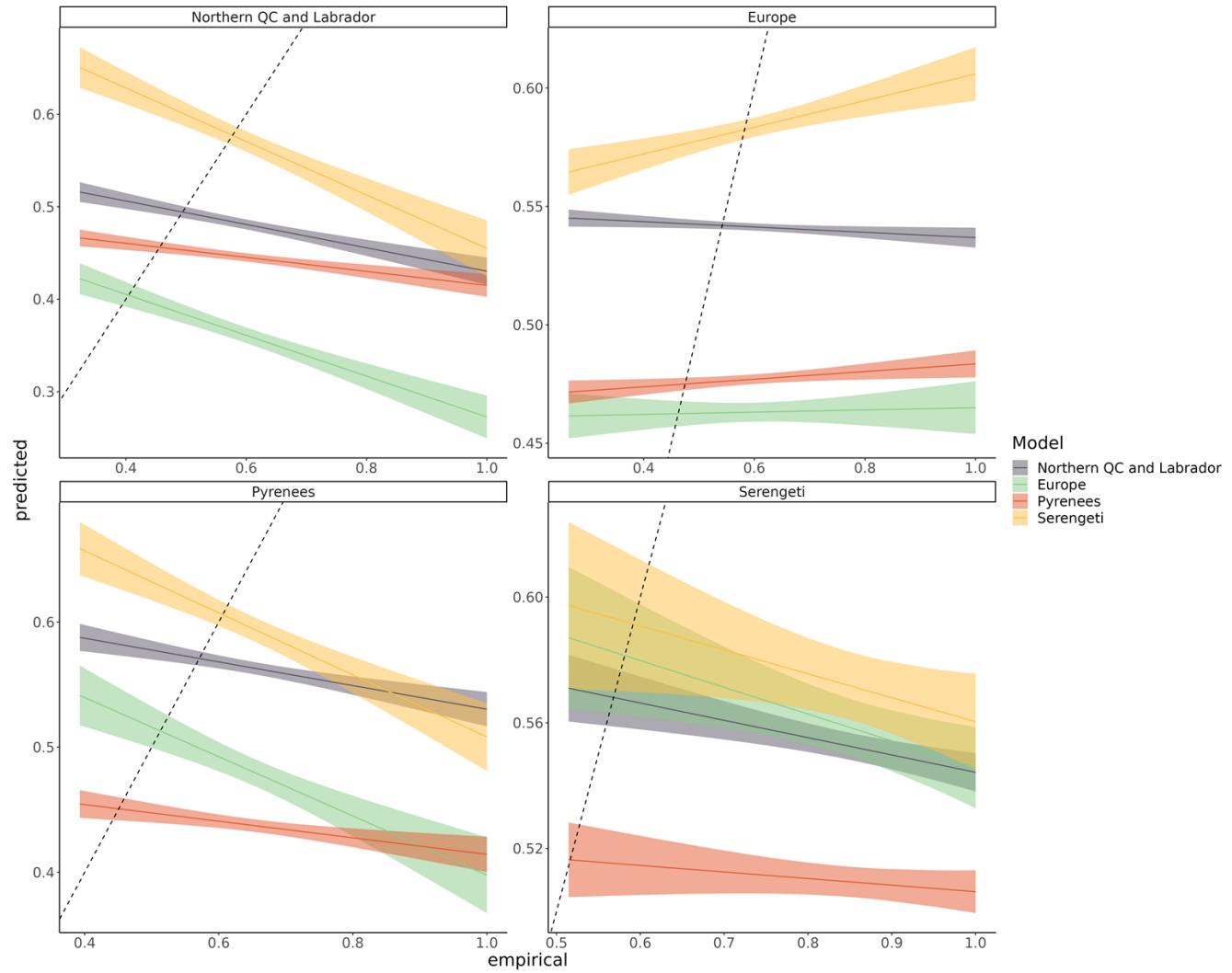


Figure S8.5: Linear regression comparing the predicted closeness centrality and empirical closeness centrality of species for each model predicting every food webs. The main bias is that predicted closeness is more homogeneous across species (slope < 1; overpredicted for low end of the range, and underpredicted for high end of the range).

Eigenvector centrality

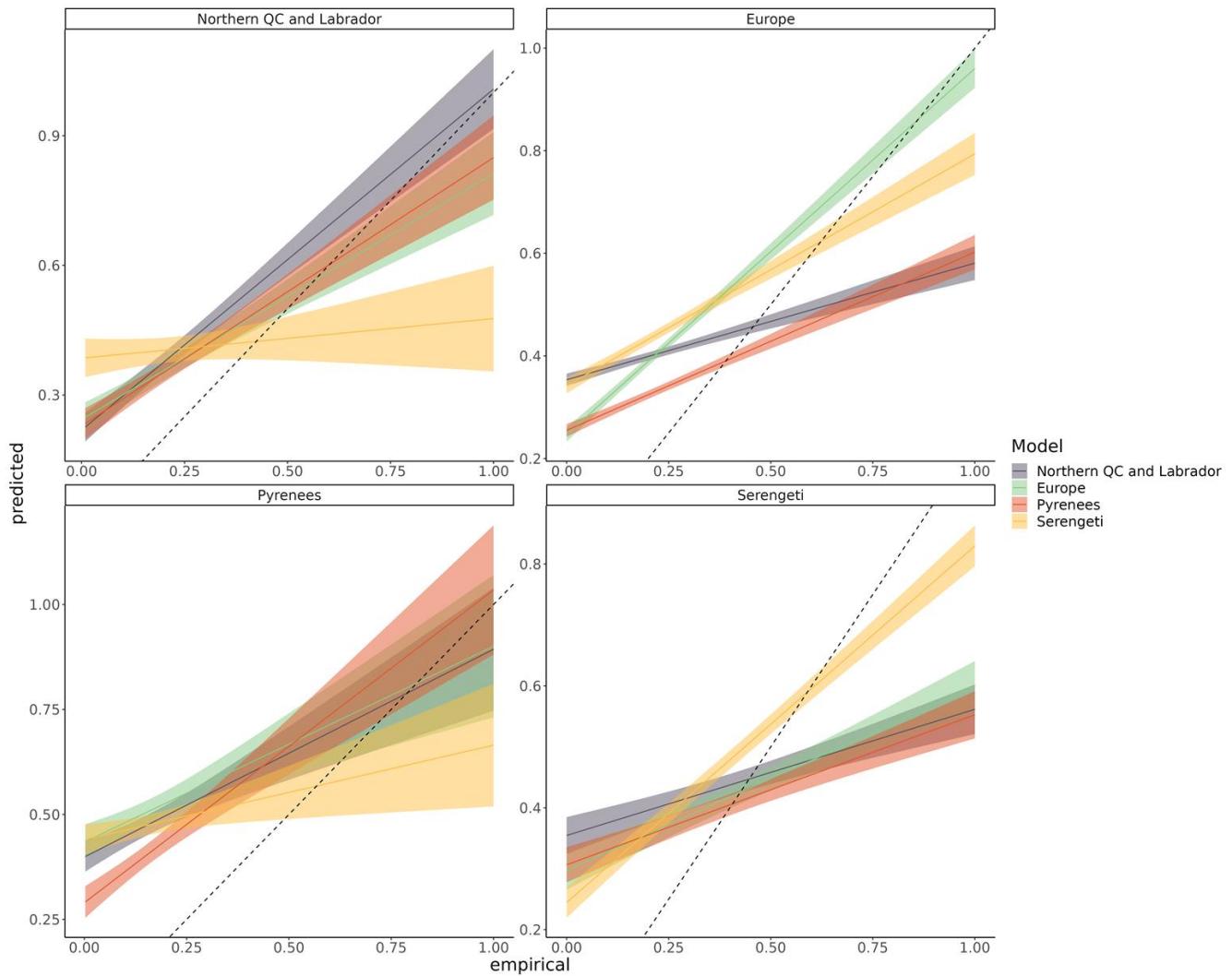


Figure S8.6: Linear regression comparing the predicted eigenvector centrality and empirical eigenvector centrality of species for each model predicting every food webs. The main bias is that predicted eigenvector centrality is more homogeneous across species (slope < 1; overpredicted for low end of the range, and underpredicted for high end of the range).

Trophic Level

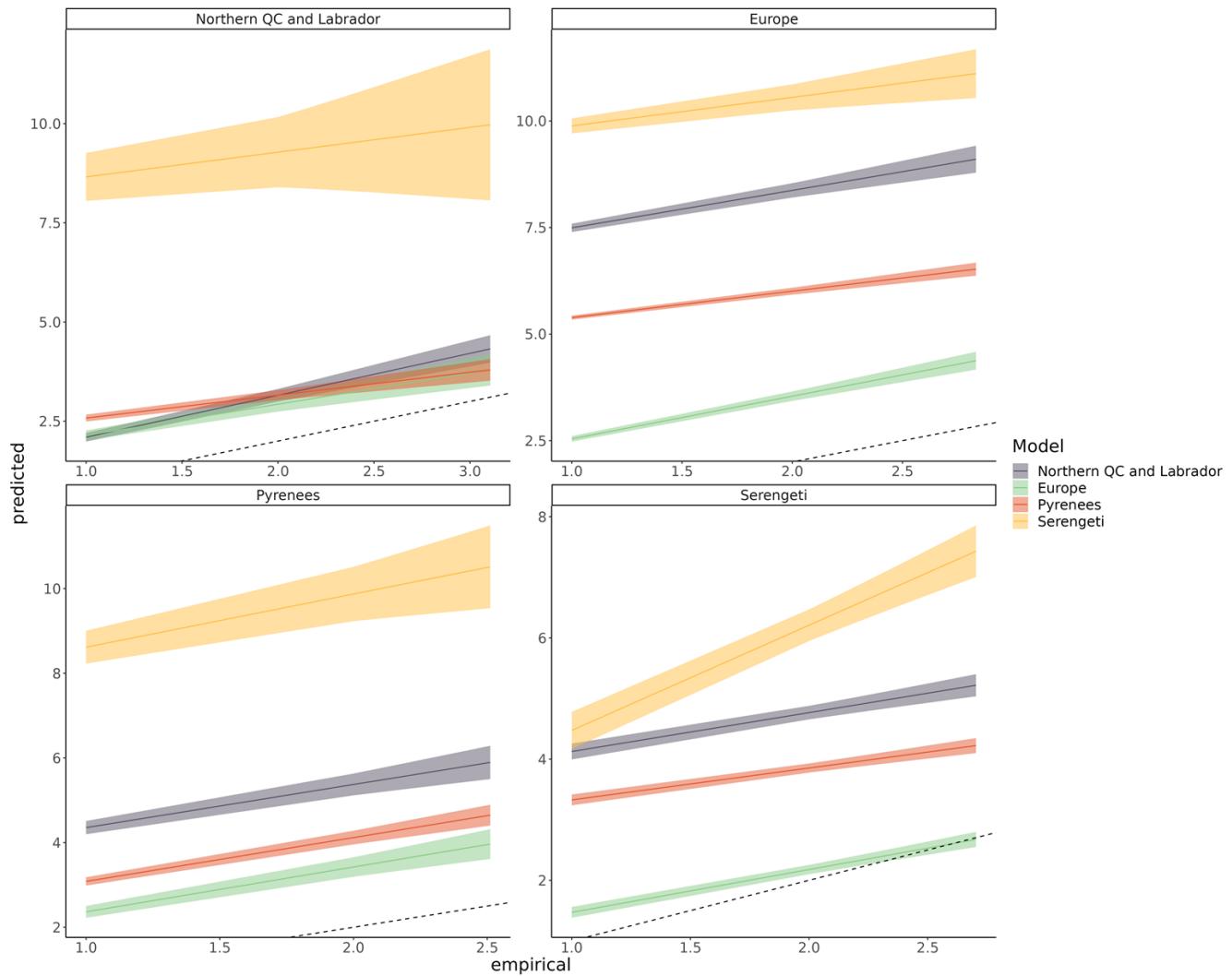


Figure S8.7: Linear regression comparing the predicted trophic level and empirical trophic level of species for each model predicting every food webs. The main bias is that predicted the trophic level of most species is overpredicted across the range (intercept > 0).

Omnivory

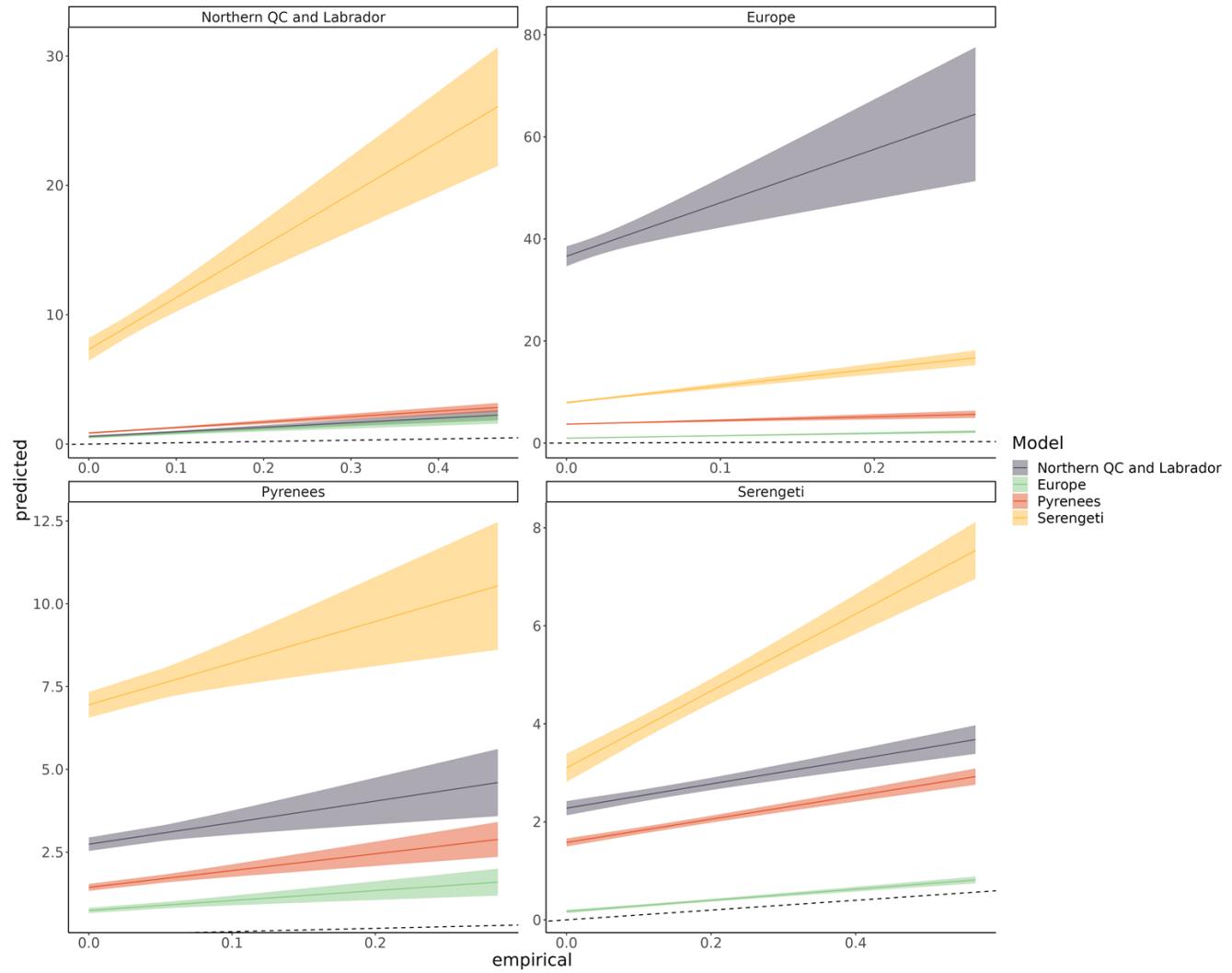


Figure S8.8: Linear regression comparing the predicted omnivory index and empirical omnivory index of species for each model predicting every food webs. The main bias is that predicted the level of omnivory of most species is overpredicted across the range (intercept > 0).

Within-module degree

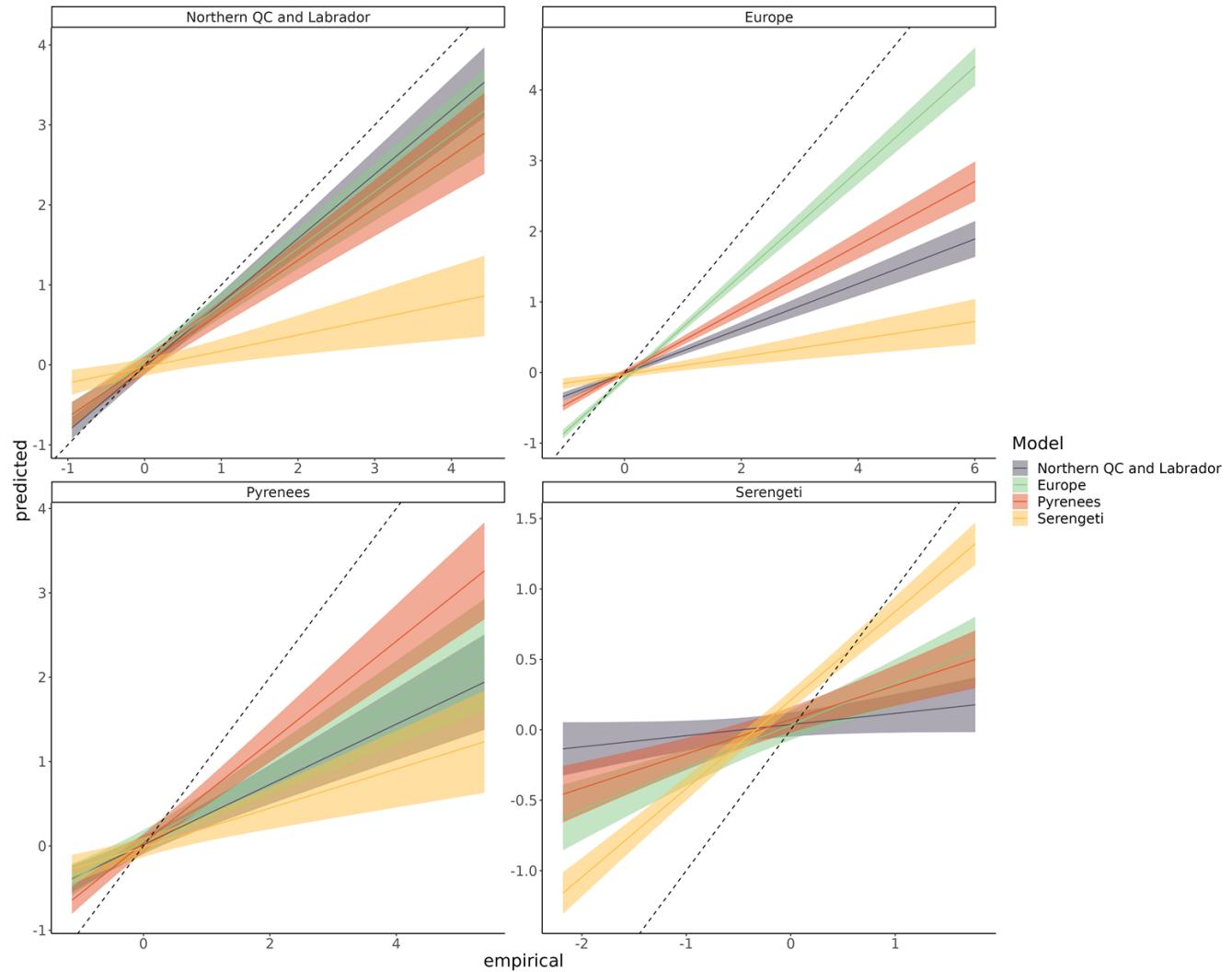


Figure S8.9: Linear regression comparing the predicted within-module degree and empirical within-module degree of species for each model predicting every food webs. The main bias is that predicted within-module degree is more homogeneous across species (slope < 1; overpredicted for low end of the range, and underpredicted for high end of the range).

Participation coefficient

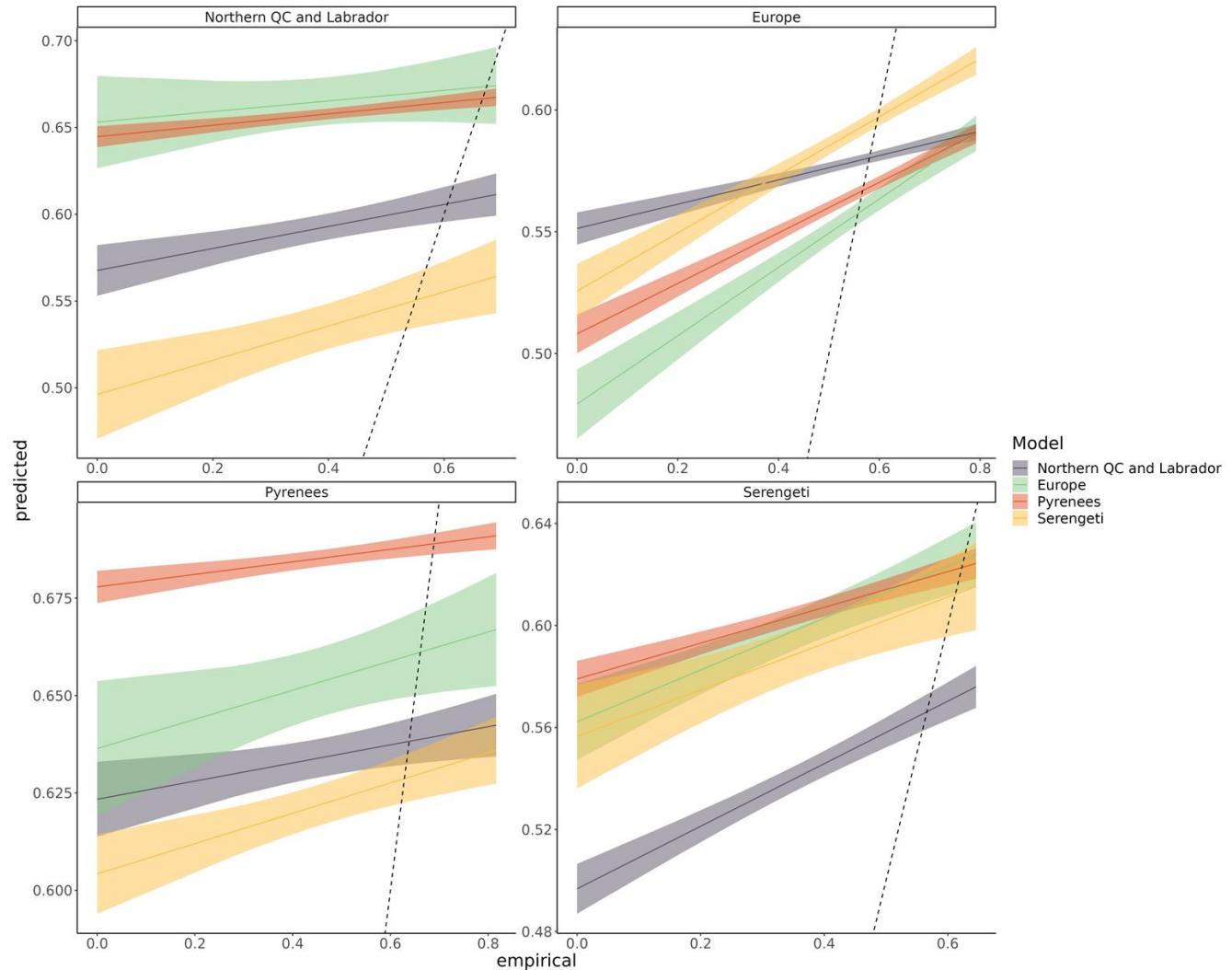


Figure S8.10: Linear regression comparing the predicted participation coefficient and empirical participation coefficient of species for each model predicting every food webs. The main bias is that predicted participation coefficient is more homogeneous across species (slope < 1; overpredicted for low end of the range, and underpredicted for high end of the range).

Frequency of the top predator position in the linear food chain motif

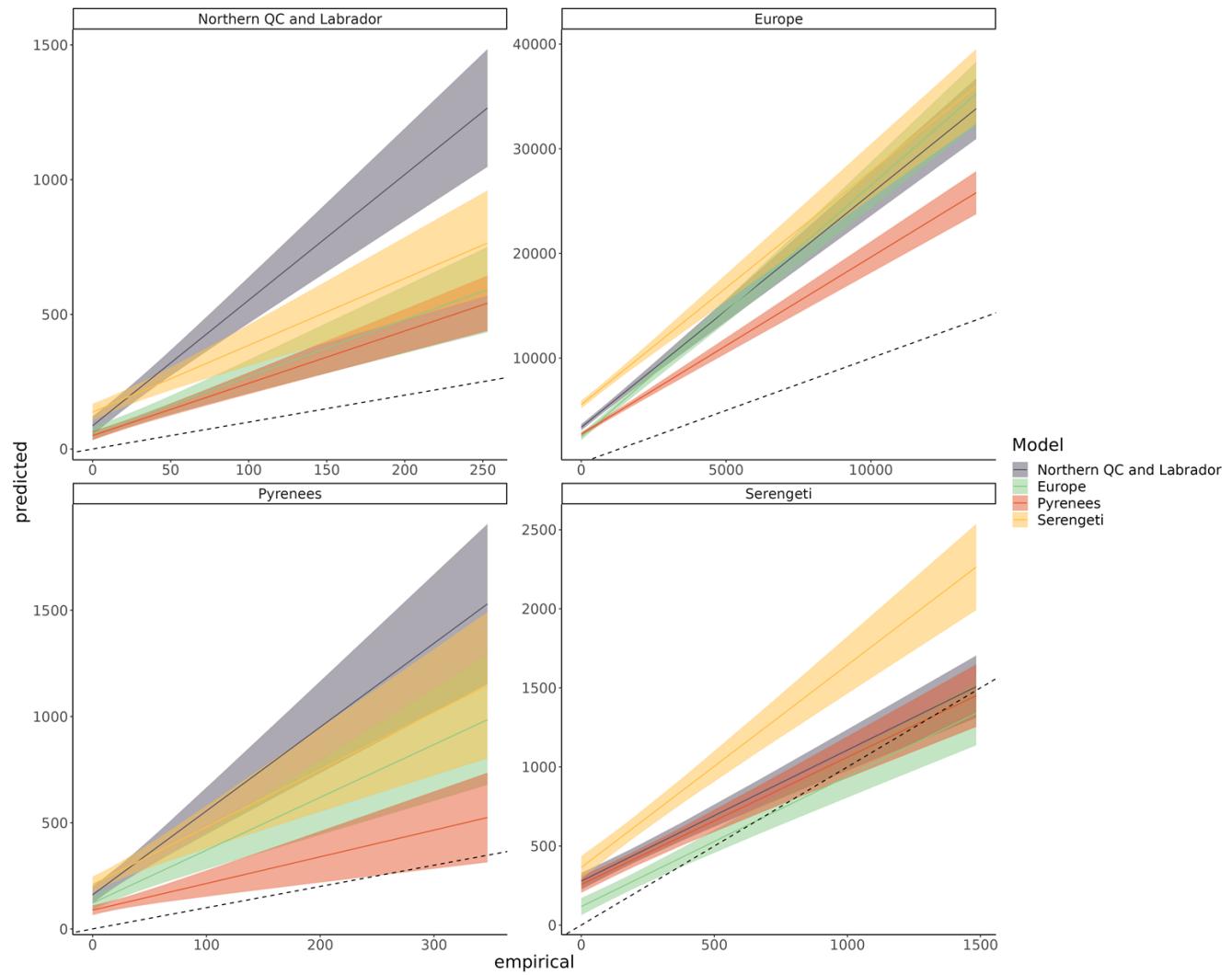


Figure S8.11: Linear regression comparing the predicted frequency of the top predator position in the linear food chain motif and empirical frequency of the top predator position in the linear food chain motif of species for each model predicting every food webs. The main bias is that predicted the frequency of top predator position of most species is overpredicted across the range (intercept > 0).

Frequency of the consumer position in the linear food chain motif

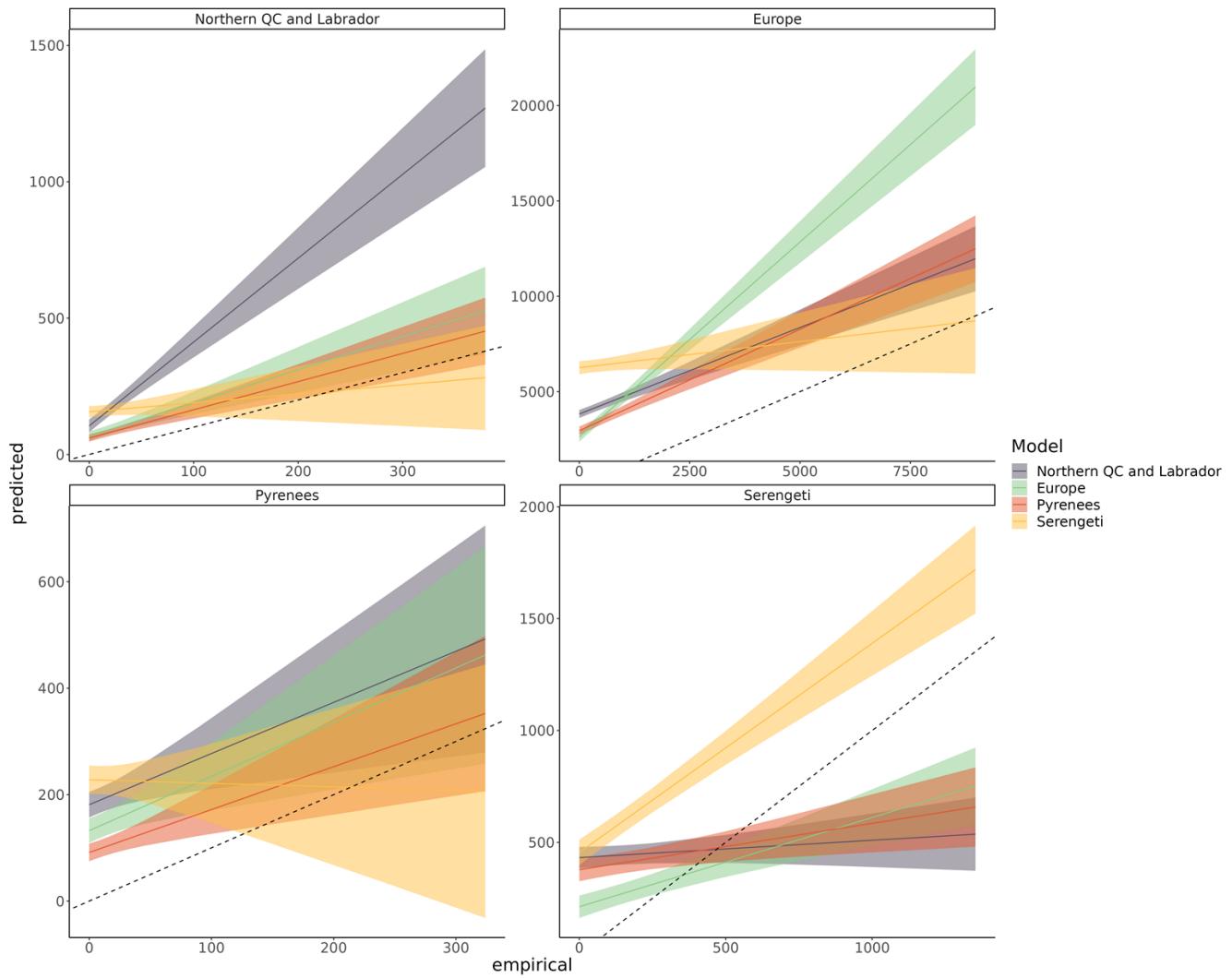


Figure S8.12: Linear regression comparing the predicted frequency of the consumer position in the linear food chain motif and empirical frequency of the consumer position in the linear food chain motif of species for each model predicting every food webs. The main bias is that predicted the frequency of consumer position of most species is overpredicted across the range (intercept > 0).

Frequency of the resource position in the linear food chain motif

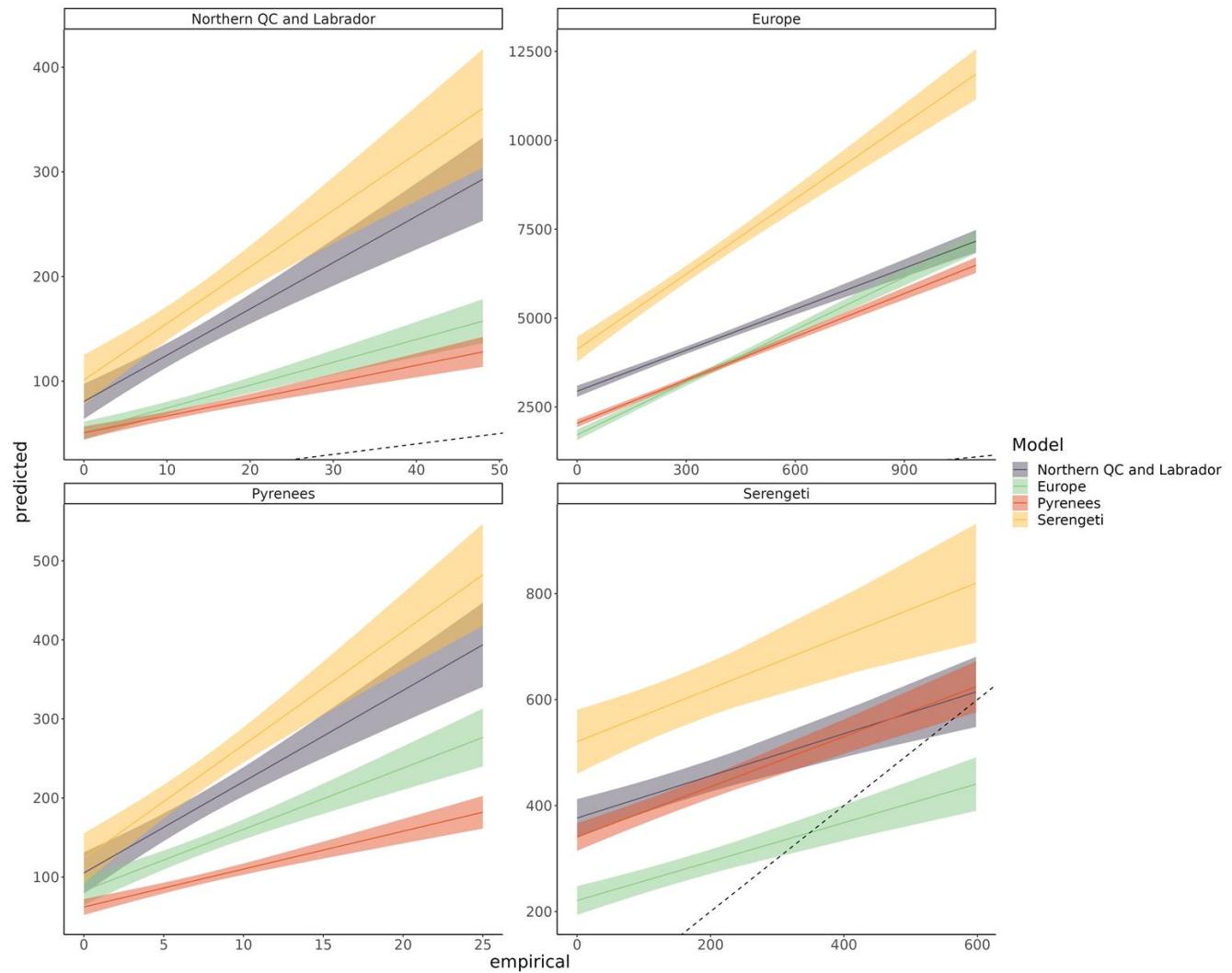


Figure S8.13: Linear regression comparing the predicted frequency of the resource position in the linear food chain motif and empirical frequency of the resource position in the linear food chain motif of species for each model predicting every food webs. The main bias is that predicted the frequency of resource position of most species is overpredicted across the range (intercept > 0).

Frequency of the omnivore position in the intraguild predation motif

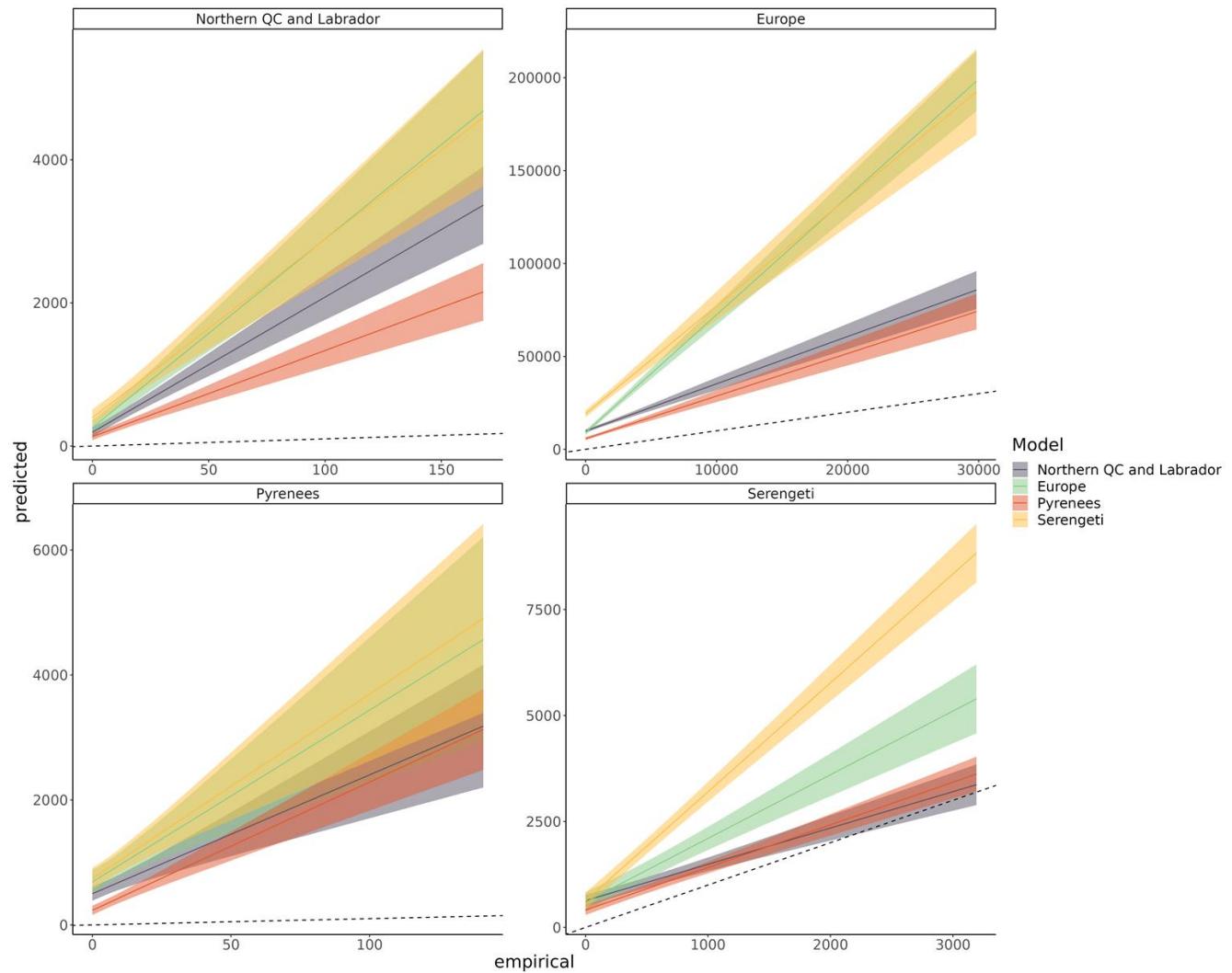


Figure S8.14: Linear regression comparing the predicted frequency of the omnivore position in the intraguild predation motif and empirical frequency of the omnivore position in the intraguild predation motif of species for each model predicting every food webs. The main bias is that predicted the frequency of omnivore position of most species is overpredicted across the range (intercept > 0).

Frequency of the consumer position in the intraguild predation motif

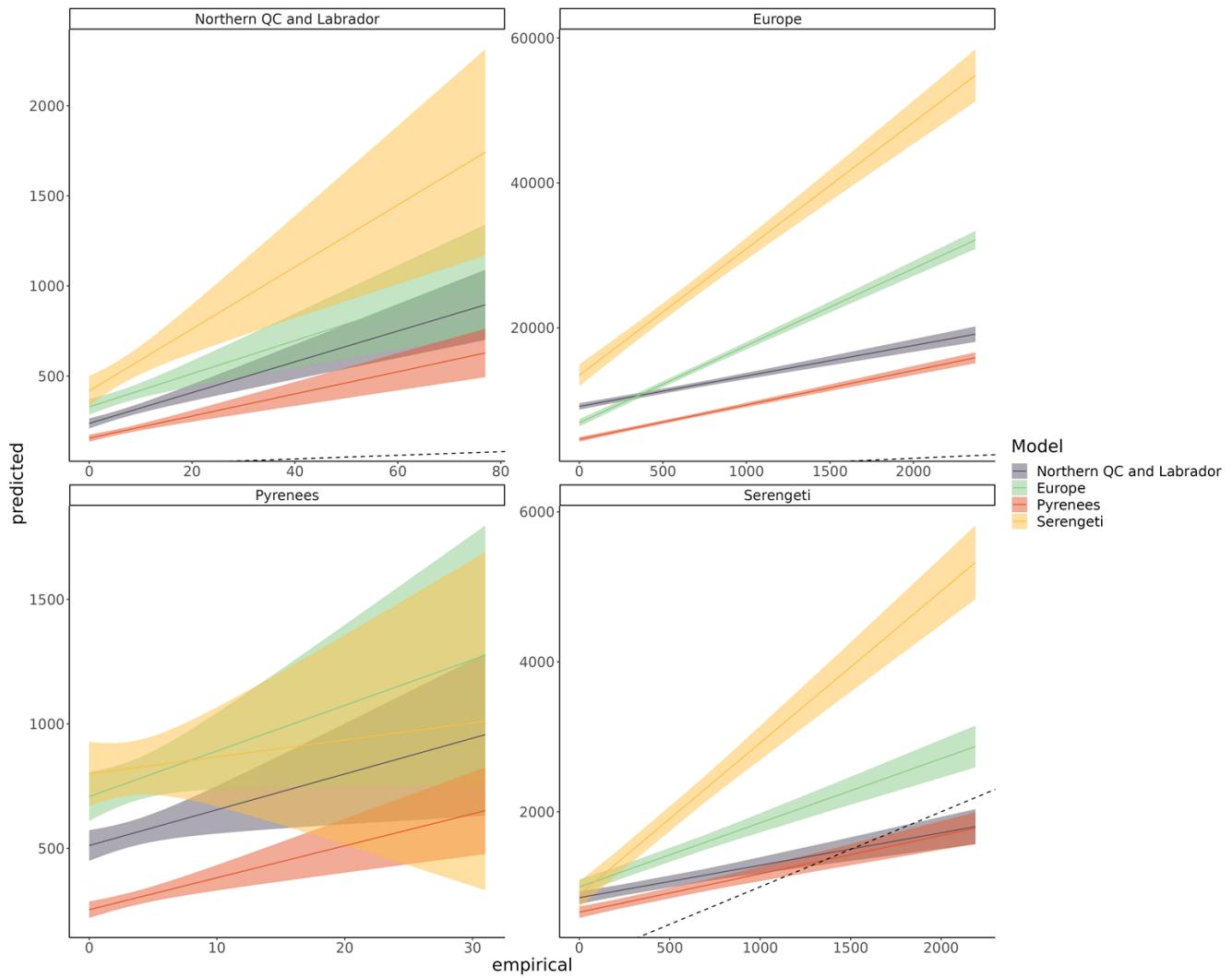


Figure S8.15: Linear regression comparing the predicted frequency of the consumer position in the intraguild predation motif and empirical frequency of the consumer position in the intraguild predation motif of species for each model predicting every food webs. The main bias is that predicted the frequency of consumer position of most species is overpredicted across the range (intercept > 0).

Frequency of the resource position in the intraguild predation motif

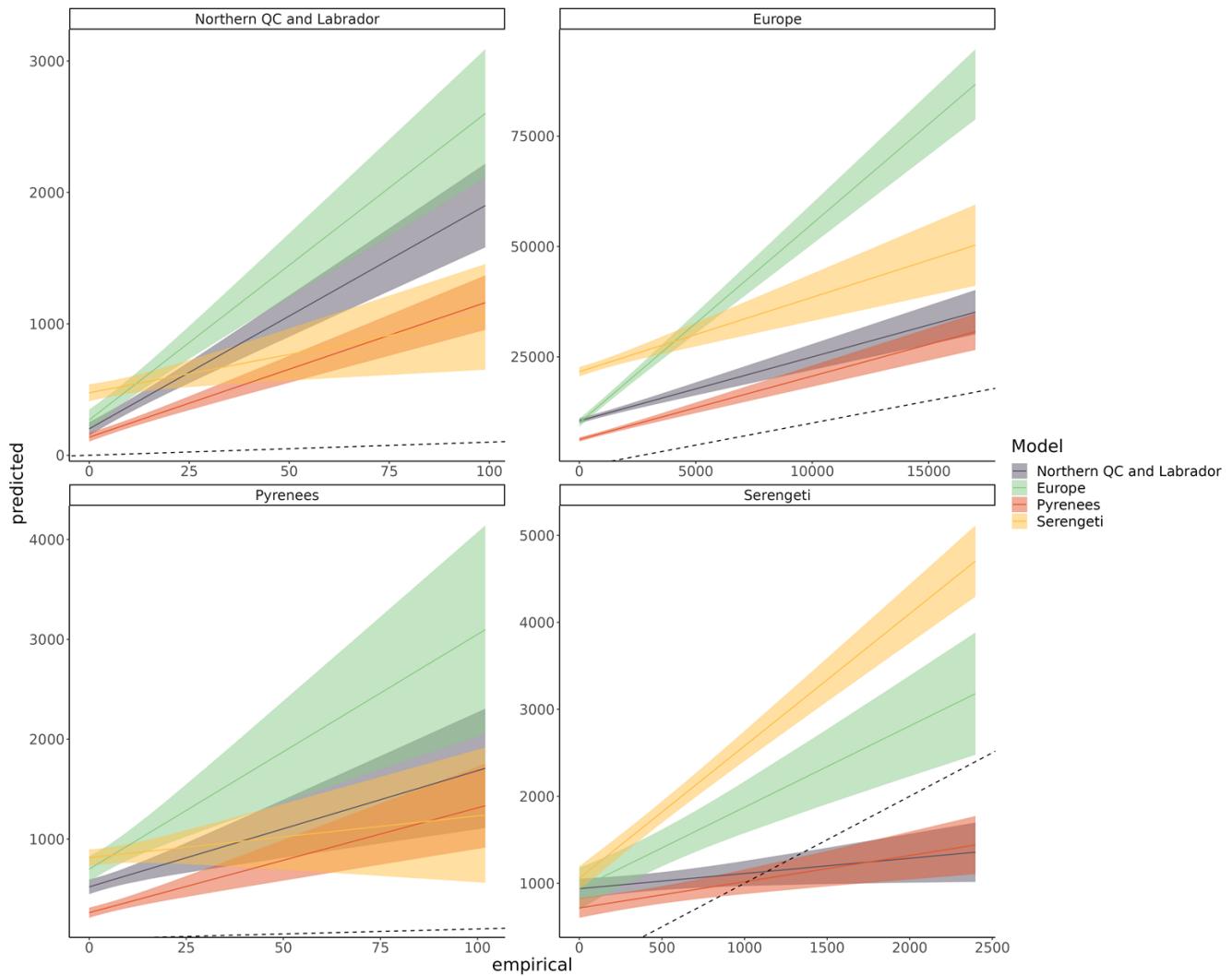


Figure S8.16: Linear regression comparing the predicted frequency of the resource position in the intraguild predation motif and empirical frequency of the resource position in the intraguild predation motif of species for each model predicting every food webs. The main bias is that predicted the frequency of resource position of most species is overpredicted across the range (intercept > 0).

Frequency of the consumer position in the direct competition motif

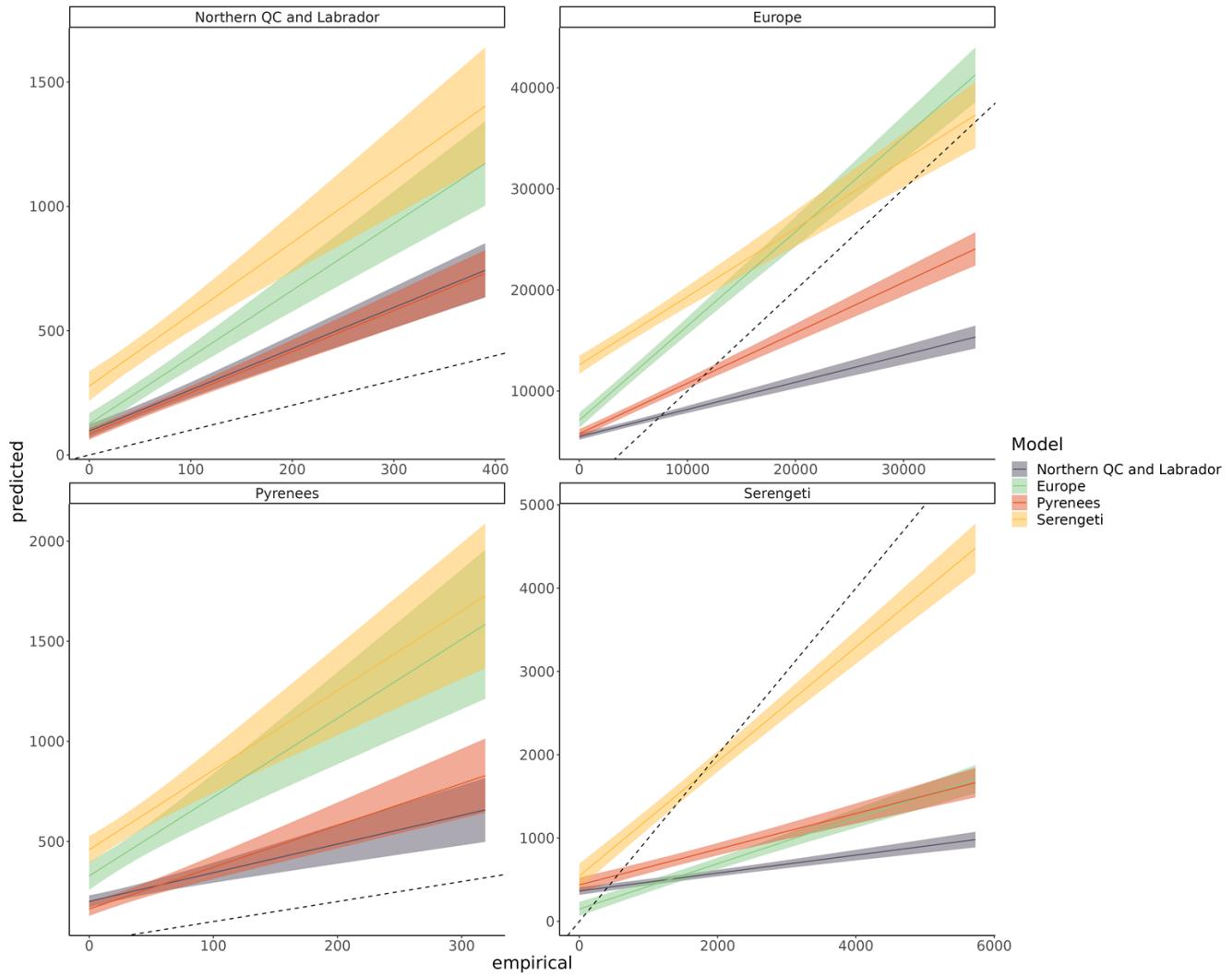


Figure S8.17: Linear regression comparing the predicted frequency of the consumer position in the direct competition motif and empirical frequency of the consumer position in the direct competition motif of species for each model predicting every food webs. The main bias is that predicted the frequency of consumer position of most species is overpredicted across the range (intercept > 0). In the Serengeti food web, the frequency of consumer position is more homogeneous across species (slope < 1)

Frequency of the resource position in the direct competition motif

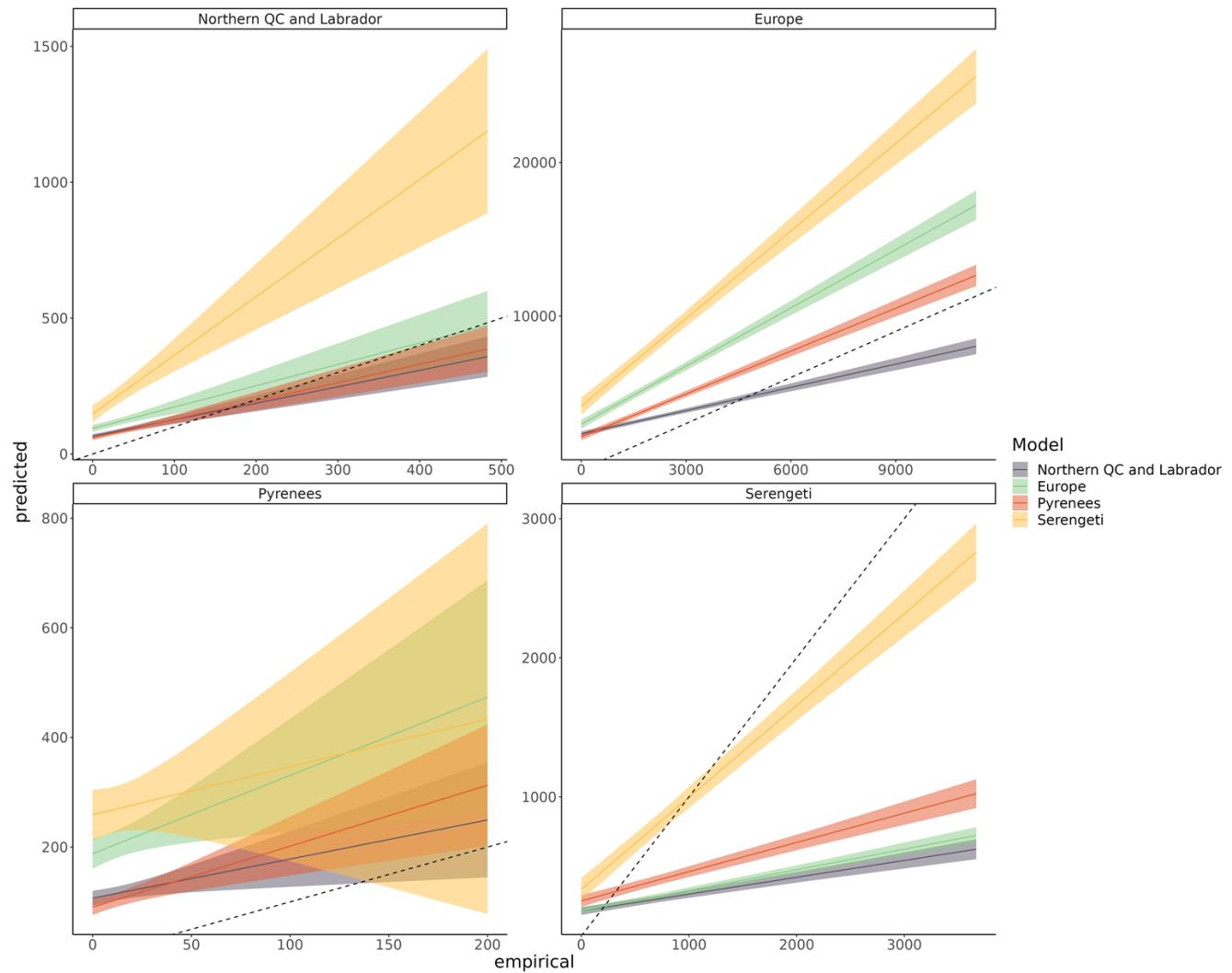


Figure S8.18: Linear regression comparing the predicted frequency of the resource position in the direct competition motif and empirical frequency of the resource position in the direct competition motif of species for each model predicting every food webs. The main bias is that predicted the frequency of resource position of most species is overpredicted across the range (intercept > 0). In the Serengeti food web, the frequency of resource position is more homogeneous across species (slope < 1)

Frequency of the consumer position in the apparent competition motif

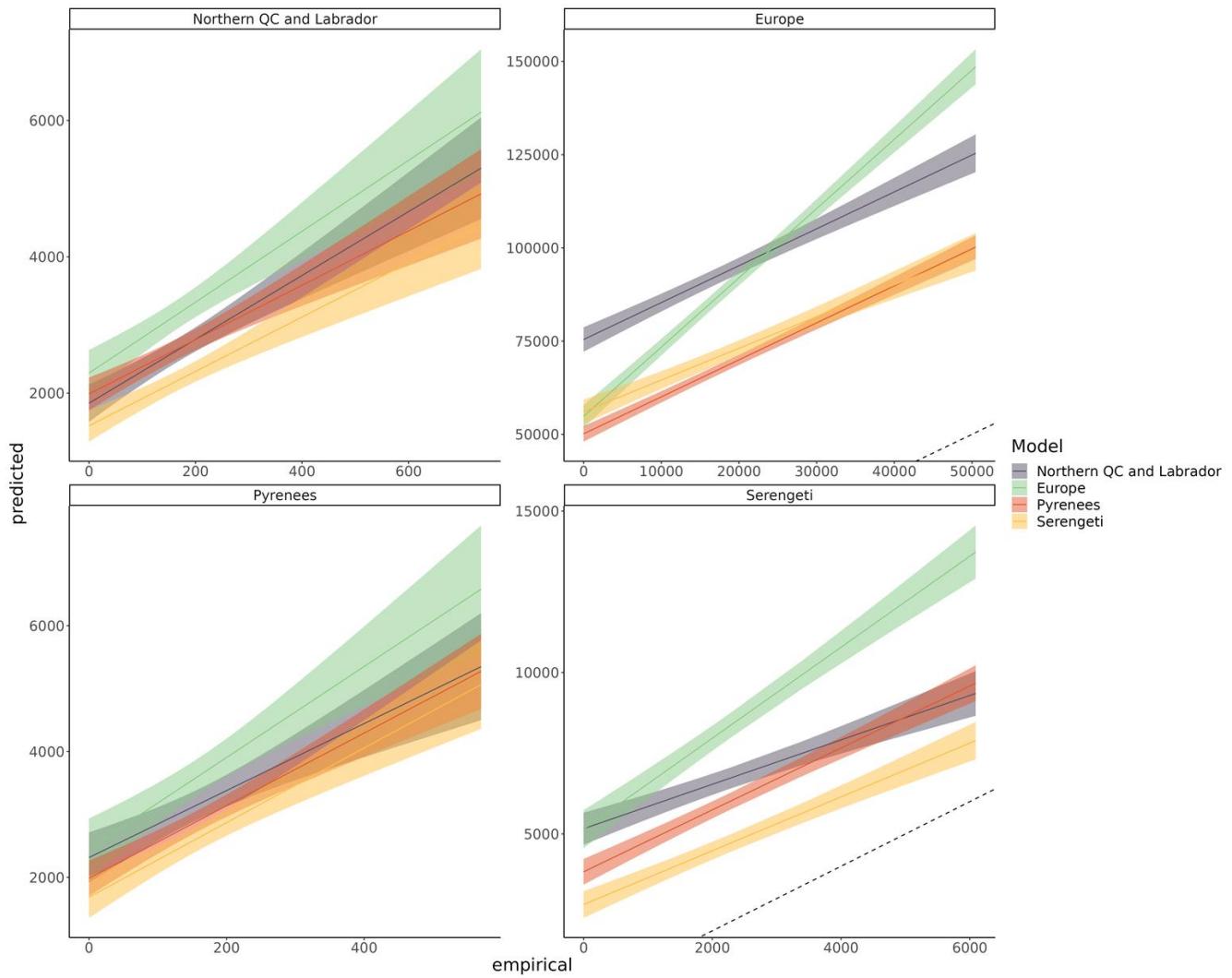


Figure S8.19: Linear regression comparing the predicted frequency of the consumer position in the apparent competition motif and empirical frequency of the consumer position in the apparent competition motif of species for each model predicting every food webs. The main bias is that predicted the frequency of consumer position of most species is overpredicted across the range (intercept > 0).

Frequency of the resource position in the apparent competition motif

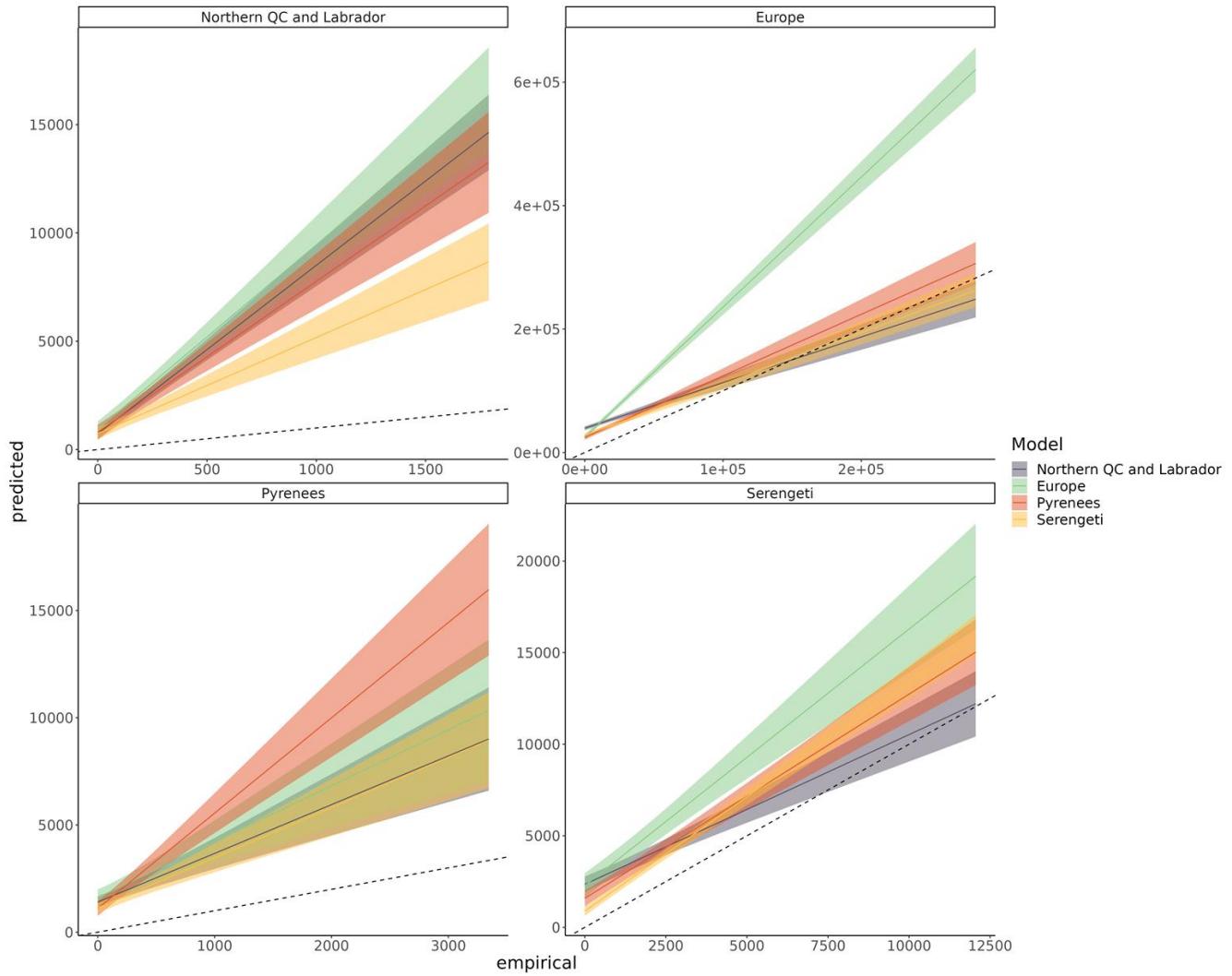


Figure S8.20: Linear regression comparing the predicted frequency of the resource position in the apparent competition motif and empirical frequency of the resource position in the apparent competition motif of species for each model predicting every food webs. The main bias is that predicted the frequency of resource position of most species is overpredicted across the range (intercept > 0).

Appendix S9: Predicting food web properties extended results

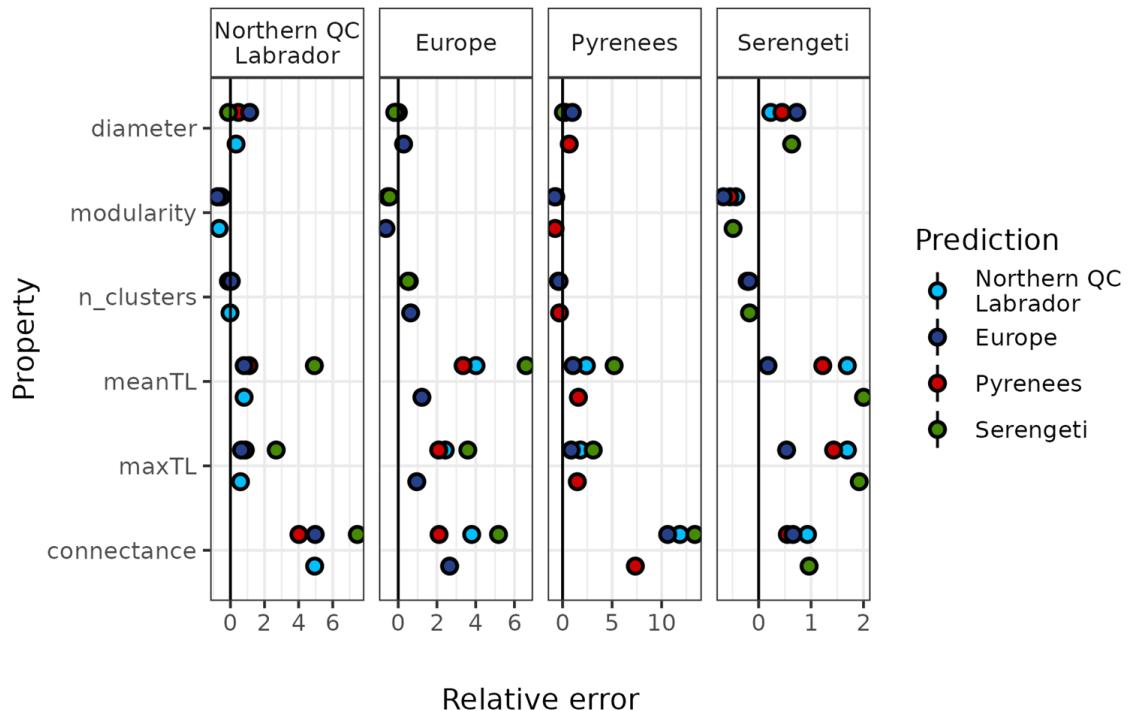


Figure S9.1: The relative error of the predicted food web properties. Relative error is the difference between the predicted and empirical estimates divided by the empirical estimate. From top to bottom food web diameter, modularity, number of clusters, mean trophic level, maximum trophic level, and connectance.

Appendix S10: Results with alternative Serengeti food web and without the Serengeti

Methods

The nodes in the original Serengeti food web are originally trophic groups. In the main manuscript, we assume that all species within a trophic group share all the same prey and predators. We made this decision based on our experience: such aggregation is not that far from what is oftentimes done by experts when filling gaps in the other food webs (e.g., “all fox-like species could eat all small mammals”). To make sure this assumption does not qualitatively influence our findings, we replicated our results without including the Serengeti food web, and by keeping the nodes of the Serengeti food web as trophic groups. To get the predictors of each node, we used the mean (for continuous predictors) and the mode (for categorical predictors) across species within each trophic group, ignoring within-group trait variations.

Results with the Serengeti trophic groups as nodes:

Predicting species interactions

Table S10.1: Area under the receiver operating curve (AUC) and area under the precision-recall-gain curve (AUPRG) of each food web model predicting every food web. Values in red are better predicted using the Serengeti food web used in the main manuscript (species as nodes), whereas values in blue are better predicted using the Serengeti with trophic groups as nodes.

Model AUC	Food web			
	Europe	Pyrenees	North Québec and Labrador	Serengeti
Europe	0.96	0.89	0.9	0.87
Pyrenees	0.87	0.95	0.86	0.85
North Québec And Labrador	0.81	0.91	0.95	0.86
Serengeti	0.7	0.8	0.7	0.89
AUPRG				
Europe	0.97	0.66	0.86	0.87
Pyrenees	0.9	0.98	0.95	0.92
North Québec And Labrador	0.88	0.91	0.96	0.94
Serengeti	0.72	0.79	0.81	0.88

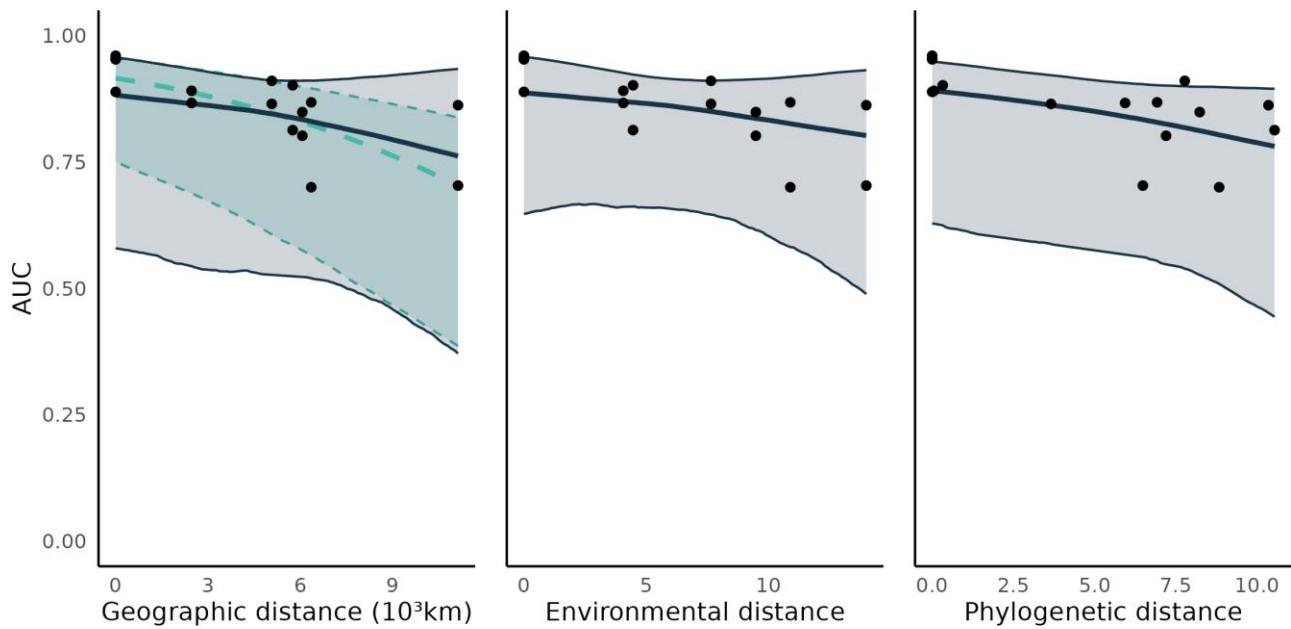


Figure S10.1: Transferability of predictive models. Each point is the predictive performance (AUC) of a model-food web predicted combination (16 combinations; Table S10.1). The trend lines are the median effects with their 95% credible interval constructed with the posterior predictive distribution of geographic, environmental, and phylogenetic distances on predictive performance. Shown are the total (turquoise; -0.49 (95% CrI = [-0.73, -0.24])) and partial (dark blue; -0.25 (95% CrI = [-0.91, 0.44])) effects of geographic distance, and the direct (controlling for geographic distance) effects of environmental (-0.20 (95% CrI = [-0.90, 0.48])) and phylogenetic distances (-0.33 (95% CrI = [-0.77, 0.11])).

In the alternative version of the Serengeti food web, nodes are groups of species. So, we did not replicate the results for species specific performance because we judged unreasonable to compare the predictability of interactions to phylogenetic distance and functional distance for nodes in the Serengeti to other food webs.

Overall, the results are similar to what we obtained using the version of the Serengeti food web we used in the main manuscript. The Serengeti food web with trophic group as nodes is better predicted by all models, but models calibrated on it predicted less well the other food webs (Table S10.1). A possible reason for this difference is that, by using a coarser species resolution, we loose small variations in the trait-interaction relationships making the interactions between trophic groups less noisy, but we loose information useful to predict species-species interactions in other food webs.

Predicting species' roles

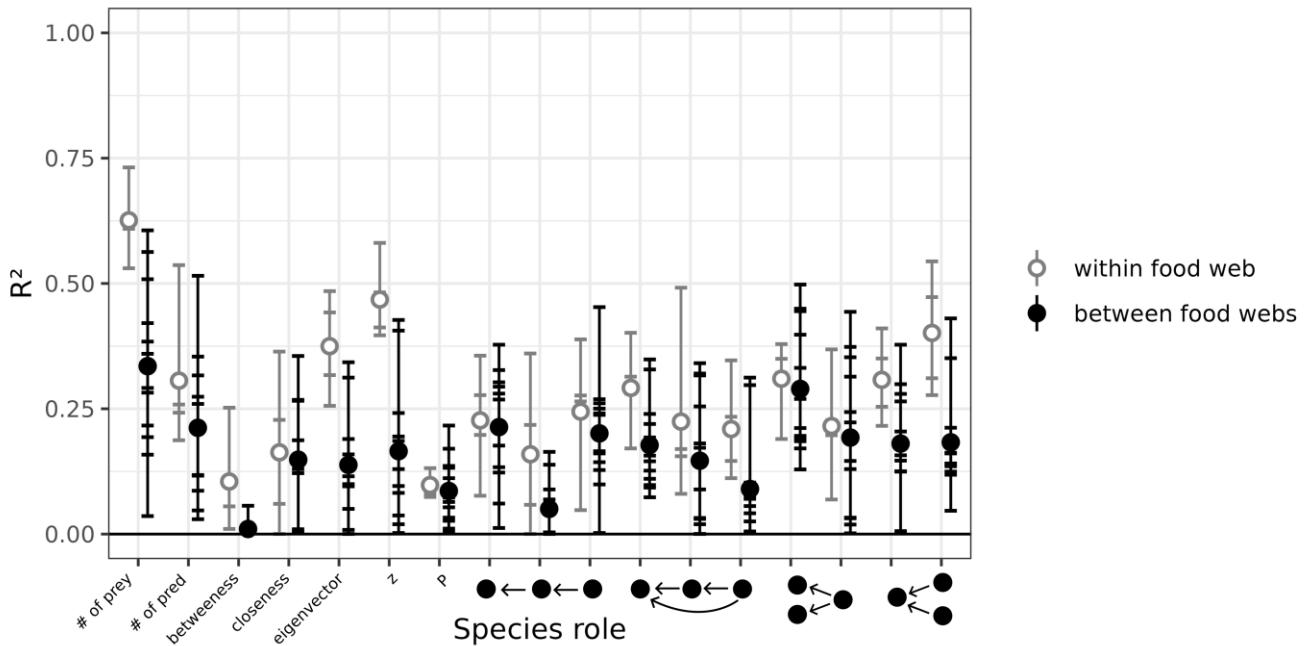


Figure S10.2: Performance of predictive models to predict species' role measured as the proportion of the variance in trophic positions explained by the models (R^2). From left to right, the figure shows the number of prey, number of predators, betweenness, closeness, eigenvector centrality, within-module degree (z), participation coefficient (P), times the species is at a specific position in the linear chain (predator – consumer - resource), intraguild predation (omnivore – consumer - resource), direct competition (2 consumers – 1 resource), and apparent competition motifs (1 consumer – 2 resources). The short horizontal lines highlight the R^2 for each role, model, and food web predicted combination. Grey open dots are the mean R^2 for within food web predictions. Full black dots are the mean R^2 for between food web predictions.

Predicting food web properties

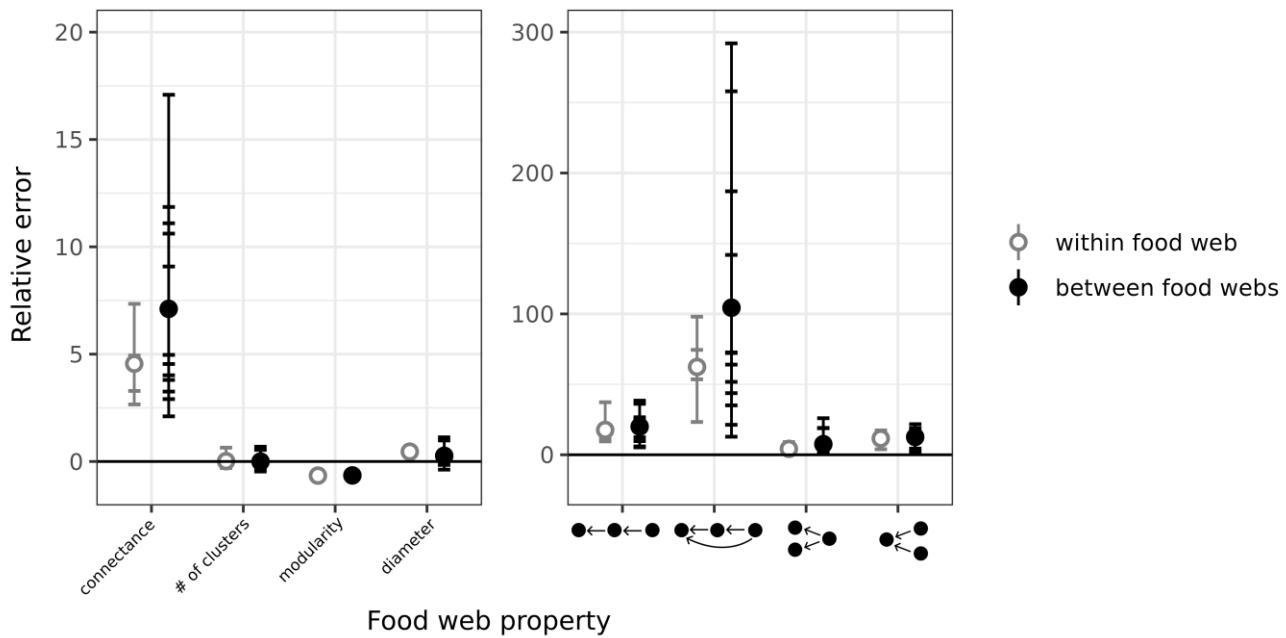


Figure S10.3: Relative error of global food web properties predicted by the boosted regression trees. From left to right, the figure shows the relative error for connectance, number of clusters, modularity, diameter, number of linear chain (predator – consumer - resource), intraguild predation (omnivore – consumer - resource), direct competition (2 consumers – 1 resource), and apparent competition (1 consumer – 2 resources) motifs. Short horizontal lines are the relative errors for each property, model, and food web predicted combination. Grey open dots are the mean relative errors for within food web predictions and full black dots are the mean relative errors for between food web predictions.

Overall, for both species' roles and food web properties, we do not find any differences in our results. Species' roles are still not well predicted, and predicted food webs are more connected and less modular. The results for trophic levels is not shown here because the European model predicted cycles in the Serengeti food web, interacting with how trophic levels are calculated.

Results after excluding the Serengeti food web

Predicting species interactions

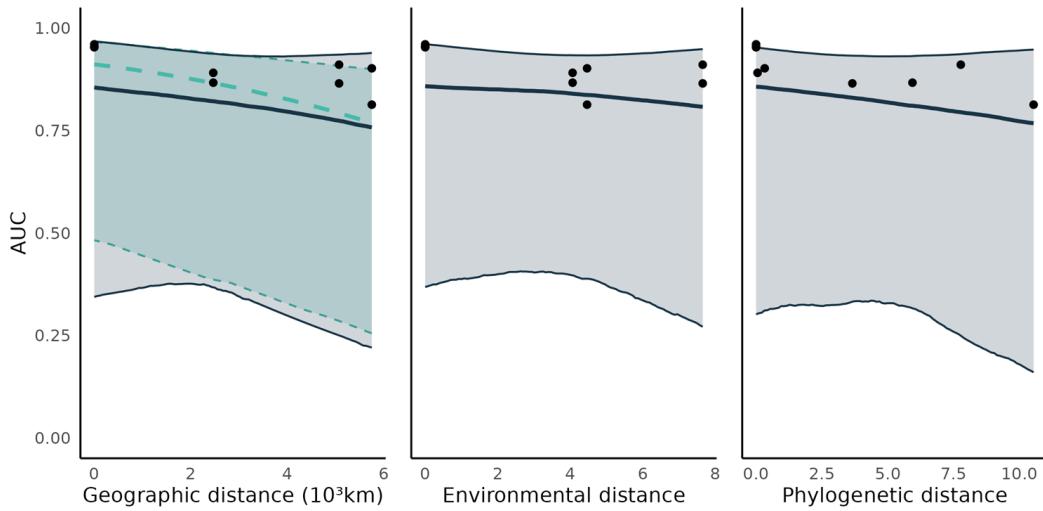


Figure S10.4: Transferability of predictive models. Each point is the predictive performance (AUC) of a model-food web predicted combination (12 combinations). The trend lines are the median effects with their 95% credible interval constructed with the posterior predictive distribution of geographic, environmental, and phylogenetic distances on predictive performance. Shown are the total (turquoise; -0.48 (95% CrI = [-0.90, -0.03])) and partial (dark blue; -0.28 (95% CrI = [-1.17, 0.65])) effects of geographic distance, and the direct (controlling for geographic distance) effects of environmental (-0.13 (95% CrI = [-0.94, 0.62])) and phylogenetic distances (-0.23 (95% CrI = [-1.08, 0.62])).

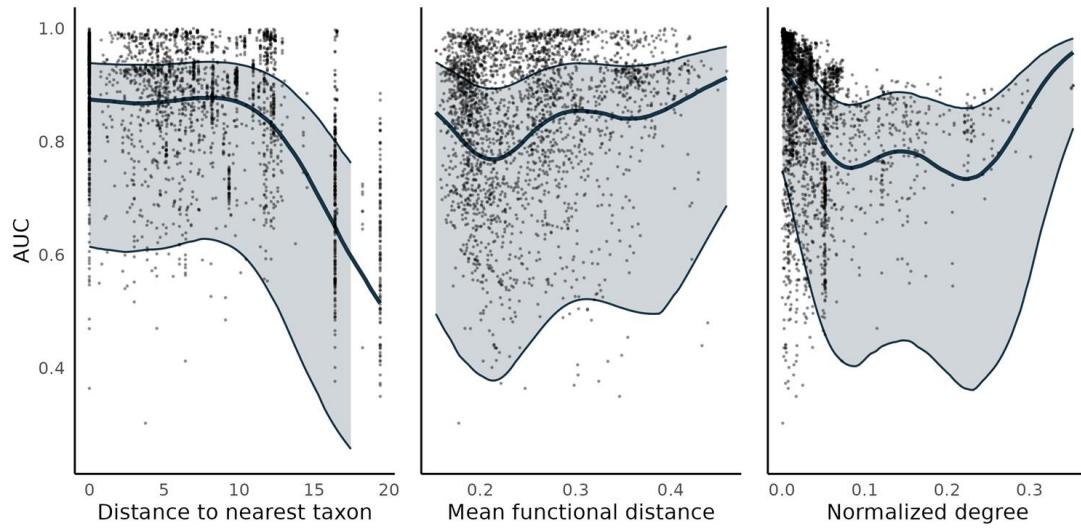


Figure S10.5: Model predictive performance for the species interactions in alternative food webs. Each point shows the model performance in predicting the prey and predators of a single species using area under the receiver operating curve (AUC) as a metric. For each panel, the thick line is the median effect while the envelope is the 95% credible interval on predictive performance built using the posterior predictive distribution of (a) distance to the nearest taxon, (b) mean functional pairwise distance, and (c) normalized degree on predictive performance.

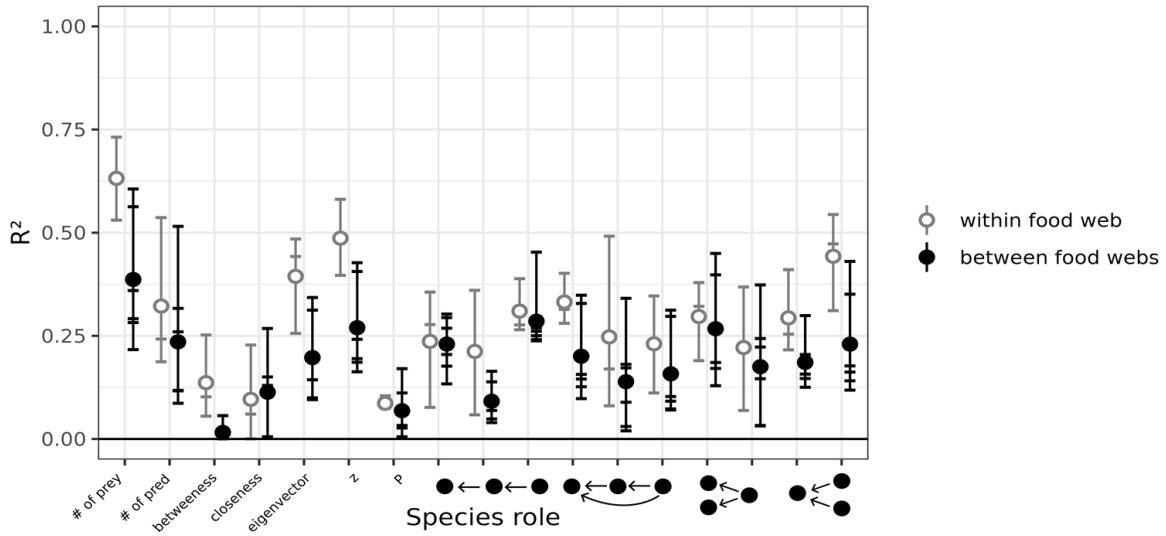


Figure S10.6: Predicting species' role measured as the proportion of the variance in trophic positions explained by the models (R^2). From left to right, the figure shows the number of prey, number of predators, betweenness, closeness, eigenvector centrality, within-module degree (z), participation coefficient (P), times the species is at a specific position in the linear chain (predator – consumer - resource), intraguild predation (omnivore – consumer - resource), direct competition (2 consumers – 1 resource), and apparent competition motifs (1 consumer – 2 resources).

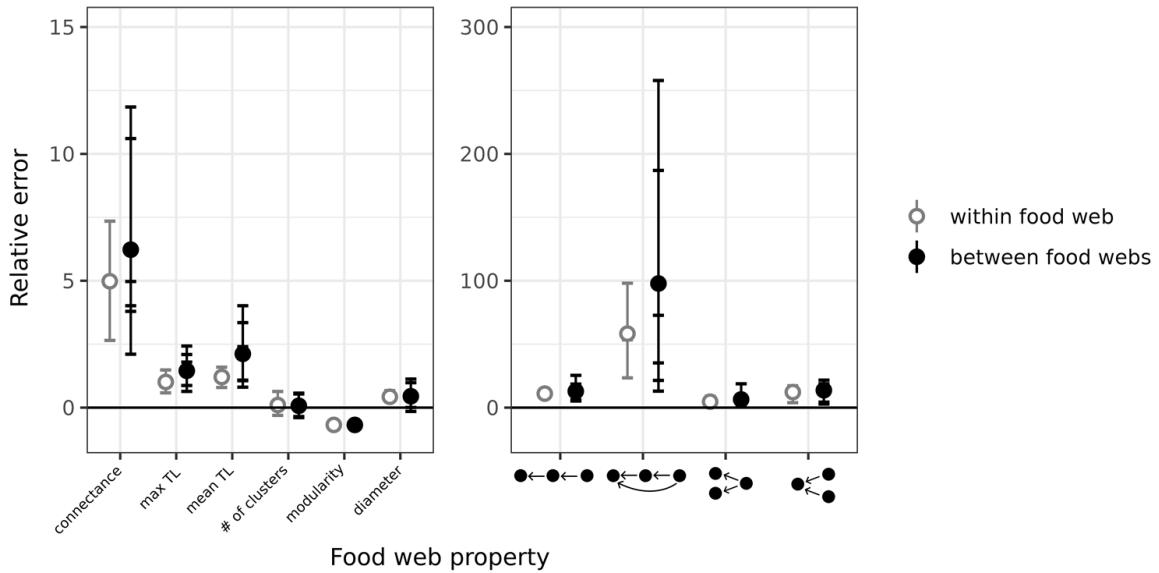


Figure S10.7: Relative error of global food web properties predicted by the boosted regression trees. From left to right, the figure shows the relative error for connectance, number of clusters, modularity, diameter, number of linear chain (predator – consumer - resource), intraguild predation (omnivore – consumer - resource), direct competition (2 consumers – 1 resource), and apparent competition (1 consumer – 2 resources) motifs.

Overall, the results are similar to what we obtain when keeping the Serengeti food web. The effects of environmental and phylogenetic distances are slightly weaker, and credible intervals are overall larger, which is expected as we reduce sample sizes.

Appendix S11 – References

- Allesina, S., & Pascual, M. (2009). Googling Food Webs: Can an Eigenvector Measure Species' Importance for Coextinctions? *PLOS Computational Biology*, 5(9), e1000494.
<https://doi.org/10.1371/journal.pcbi.1000494>
- Bascompte, J., Melián, C. J., & Sala, E. (2005). Interaction strength combinations and the overfishing of a marine food web. *Proceedings of the National Academy of Sciences*, 102(15), 5443–5447.
<https://doi.org/10.1073/pnas.0501562102>
- Berteaux, D., Ropars, P., & Casajus, N. (2018). Toundra Nunavik: Matrice des relations trophiques entre espèces du Labrador et du Québec nordique, v. 1.0 (1980-2010). *Nordicana D36*.
<https://doi.org/10.5885/45555CE-DA1FF11FA4254703>
- Borrelli, J. J. (2015). Selection against instability: Stable subgraphs are most frequent in empirical food webs. *Oikos*, 124(12), 1583–1588. <https://doi.org/10.1111/oik.02176>
- Cirtwill, A. R., Dalla Riva, G. V., Gaiarsa, M. P., Bimler, M. D., Cagua, E. F., Coux, C., & Dehling, D. M. (2018). A review of species role concepts in food webs. *Food Webs*, 16, e00093.
<https://doi.org/10.1016/j.fooweb.2018.e00093>
- Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, 1695.
- Curtsdotter, A., Binzer, A., Brose, U., de Castro, F., Ebenman, B., Eklöf, A., Riede, J. O., Thierry, A., & Rall, B. C. (2011). Robustness to secondary extinctions: Comparing trait-based sequential deletions in static and dynamic food webs. *Basic and Applied Ecology*, 12(7), 571–580.
<https://doi.org/10.1016/j.baae.2011.09.008>
- de Visser, S. N., Freymann, B. P., & Olff, H. (2011). The Serengeti food web: Empirical quantification and analysis of topological changes under increasing human impact. *Journal of Animal Ecology*, 80(2), 484–494. <https://doi.org/10.1111/j.1365-2656.2010.01787.x>
- Delmas, E., Besson, M., Brice, M.-H., Burkle, L. A., Riva, G. V. D., Fortin, M.-J., Gravel, D., Guimarães, P. R., Hembry, D. H., Newman, E. A., Olesen, J. M., Pires, M. M., Yeakel, J. D., & Poisot, T. (2019). Analysing ecological networks of species interactions. *Biological Reviews*, 94(1), 16–36. <https://doi.org/10.1111/brv.12433>

- Duffy, J. E., Cardinale, B. J., France, K. E., McIntyre, P. B., Thébaud, E., & Loreau, M. (2007). The functional role of biodiversity in ecosystems: Incorporating trophic complexity. *Ecology Letters*, 10(6), 522–538. <https://doi.org/10.1111/j.1461-0248.2007.01037.x>
- Dunne, J. A., Williams, R. J., & Martinez, N. D. (2002). Food-web structure and network theory: The role of connectance and size. *Proceedings of the National Academy of Sciences*, 99(20), 12917–12922. <https://doi.org/10.1073/pnas.192407699>
- Dyer, L. A., & Letourneau, D. (2003). Top-down and bottom-up diversity cascades in detrital vs. Living food webs. *Ecology Letters*, 6(1), 60–68. <https://doi.org/10.1046/j.1461-0248.2003.00398.x>
- Etard, A., Morrill, S., & Newbold, T. (2020). Global gaps in trait data for terrestrial vertebrates. *Global Ecology and Biogeography*, 29(12), 2143–2158. <https://doi.org/10.1111/geb.13184>
- Flach, P., & Kull, M. (2015). Precision-Recall-Gain Curves: PR Analysis Done Right. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (Vol. 28). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2015/file/33e8075e9970de0cfea955af4644bb2-Paper.pdf
- Freeman, L. C. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry*, 40(1), 35–41. <https://doi.org/10.2307/3033543>
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis* (0 ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/b16018>
- Greenwell, B., Boehmke, B., Cunningham, J., Developers, G., & Greenwell, M. B. (2019). Package ‘gbm.’ *R Package Version*, 2(5).
- Guimerà, R., & Amaral, L. A. N. (2005). Cartography of complex networks: Modules and universal roles. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(02), P02001. <https://doi.org/10.1088/1742-5468/2005/02/P02001>
- Hijmans, R. J., Phillips, S., Leathwick, J., Elith, J., & Hijmans, M. R. J. (2017). Package ‘dismo.’ *Circles*, 9(1), 1–68.

- Jordán, F. (2009). Keystone species and food webs. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1524), 1733–1741. <https://doi.org/10.1098/rstb.2008.0335>
- Kones, J. K., Soetaert, K., Oevelen, D. van, & Owino, J. (2009). Are network indices robust indicators of food web functioning? A Monte Carlo approach. *Ecological Modelling*, 220, 370–382. <http://dx.doi.org/10.1016/j.ecolmodel.2008.10.012>
- Lurgi, M., López, B. C., & Montoya, J. M. (2012). Climate change impacts on body size and food web structure on mountain ecosystems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1605), 3050–3057. <https://doi.org/10.1098/rstb.2012.0239>
- Maiorano, L., Montemaggiore, A., Ficetola, G. F., O'Connor, L., & Thuiller, W. (2020). TETRA-EU 1.0: A species-level trophic metaweb of European tetrapods. *Global Ecology and Biogeography*, geb.13138. <https://doi.org/10.1111/geb.13138>
- Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. (2002). Network Motifs: Simple Building Blocks of Complex Networks. *Science*, 298(5594), 824–827. <https://doi.org/10.1126/science.298.5594.824>
- Saravia, L. A. (2022). *multiweb: Ecological network analysis includind multiplex networks*. <https://github.com/lasaravia/multiweb>
- Stekhoven, D. J. (2013). *missForest: Nonparametric missing value imputation using random forest* [Manual].
- Stekhoven, D. J., & Buehlmann, P. (2012). MissForest—Non-parametric missing value imputation for mixed-type data. *Bioinformatics (Oxford, England)*, 28(1), 112–118.
- Stouffer, D. B., & Bascompte, J. (2010). Understanding food-web persistence from local to global scales. *Ecology Letters*, 13(2), 154–161. <https://doi.org/10.1111/j.1461-0248.2009.01407.x>
- Stouffer, D. B., & Bascompte, J. (2011). Compartmentalization increases food-web persistence. *Proceedings of the National Academy of Sciences*, 108(9), 3648–3652. <https://doi.org/10.1073/pnas.1014353108>
- Stouffer, D. B., Sales-Pardo, M., Sirer, M. I., & Bascompte, J. (2012). Evolutionary Conservation of Species' Roles in Food Webs. *Science*, 335(6075), 1489–1492. <https://doi.org/10.1126/science.1216556>

- Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P.-C. (2021). Rank-Normalization, Folding, and Localization: An Improved R⁺ for Assessing Convergence of MCMC (with Discussion). *Bayesian Analysis*, 16(2), 667–718. <https://doi.org/10.1214/20-BA1221>
- Vieira, M. C., & Almeida-Neto, M. (2015). A simple stochastic model for complex coextinctions in mutualistic networks: Robustness decreases with connectance. *Ecology Letters*, 18(2), 144–152. <https://doi.org/10.1111/ele.12394>
- Zhao, Q., Brink, P. J. V. den, Carpentier, C., Wang, Y. X. G., Rodríguez-Sánchez, P., Xu, C., Vollbrecht, S., Gillissen, F., Vollebregt, M., Wang, S., & Laender, F. D. (2019). Horizontal and vertical diversity jointly shape food web stability against small and large perturbations. *Ecology Letters*, 22(7), 1152–1162. <https://doi.org/10.1111/ele.13282>