

Global Ecology and Biogeography

A
Mac

**Trophic interaction models predict interactions across
ecosystems, not food webs**

Journal:	<i>Global Ecology and Biogeography</i>
Manuscript ID	GEB-2023-0032
Manuscript Type:	Research Article
Keywords:	Ecological predictions, Food web, Model transferability, Terrestrial vertebrates, Trait matching, Trophic interactions

SCHOLARONE™
Manuscripts

1
2 1 **Trophic interaction models predict interactions across ecosystems, not food webs.**
3
4 2
5
6 3 **Running title:** Interaction models transferability
7
8
9
10 4
11
12 5 **Abstract**
13
14 6 **Aim:** Trophic interactions are central to our understanding of essential ecosystem functions as
15
16 7 well as their stability. Predicting these interactions has become increasingly common due to the
17
18 8 lack of empirical data on trophic interactions for most taxa in most ecosystems. We aim to
19
20 9 determine whether and how accurately we can extrapolate to new communities both in terms of
21
22 10 pairwise predator-prey interactions and higher level food web attributes (i.e., species position,
23
24 11 food web-level properties).
25
26
27
28 12 **Location:** Canada, Europe, Tanzania.
29
30
31 13 **Time period:** Current.
32
33
34 14 **Major taxa studied:** Terrestrial vertebrates
35
36 15 **Methods:** We use a trait-based model of pairwise trophic interactions, calibrated independently
37
38 16 on four different terrestrial vertebrate food webs (Canadian tundra, Serengeti, alpine south-
39
40 17 eastern Pyrenees, and entire Europe) and assess the ability of each calibrated instance of the
41
42 18 model to predict alternative food webs. We test how well predictions recover individual
43
44 19 predator-prey interactions as well as higher level food web properties across geographical
45
46 20 locations.
47
48
49 21 **Results:** We find that, given enough phylogenetic and environmental similarities between food
50
51 22 webs, trait-based models predict most interactions and their absence correctly ($AUC > 0.82$),
52
53 23 even across highly contrasting environments. However, network metrics were less well-predicted
54
55
56
57
58
59
60

1
2
3 24 than single interactions by our models. Predicted food webs were more connected, less modular,
4
5 25 and had higher mean trophic levels than observed.
6
7

8 26 **Main conclusions:** Theory predicts that the variability observed in food webs can be explained
9
10 27 by differences in trait distributions and trait-matching relationships. Trait-based models can
11
12 28 predict potential interactions amongst species in an ecosystem when calibrated using food web
13
14 29 data from reasonably similar ecosystems. This suggests that food webs vary spatially primarily
15
16 30 through changes in trait distributions. These models, however, are less good at predicting system
17
18 31 level food web properties. We thus highlight the need for methodological advances to
19
20 32 simultaneously address trophic interactions and the structure of food webs across time and space.
21
22
23
24 33
25
26 34 **Key Words:** ecological predictions; food web; model transferability; terrestrial vertebrates; trait
27
28 35 matching; trophic interactions

1
2
3 **36 Introduction**
4

5 37 Ecosystem functions (e.g., energy flows and material cycling) and community stability depend
6 38 on the trophic relationships that link species within a community (Harvey et al., 2017). Despite
7
8 39 the importance of food webs for understanding ecosystem structure and dynamics, recognized
9 40 over the last 80 years (Lindeman, 1942), we still face major challenges to develop accurate
10
11 41 descriptions of natural food webs. The lack of trophic interactions data across most locations and
12
13 42 taxa (Poisot et al., 2021) lies at the core of the fundamental technical and practical challenges in
14
15 43 food web ecology. Observing interactions is more challenging than observing species because
16
17 44 two individuals need to be simultaneously detected while interacting (Jordano, 2016). This
18
19 45 challenge is exacerbated by the fact that the number of possible interactions in food webs
20
21 46 increases quadratically with the number of species, making the potential set of observations to be
22
23 47 made extremely large. Determining all possible interactions among species within a food web is
24
25 48 thus difficult even in species-poor ecosystems, which calls for a systematic approach to predict
26
27 49 links.
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

50
51 51 A first step towards constructing food webs is to focus on predator-prey relationships because
52 52 they are the most commonly recorded type of ecological interaction, and have been shown to
53 53 respond to a predictable set of neutral and niche processes (Morales-Castilla et al., 2015). A
54 54 neutral model for trophic interactions suggests that the probability and strength of interactions
55 55 depend only onto the co-occurrences and abundances of species (Canard et al., 2012). However,
56 56 recent studies have shown that even if co-occurrence is a requirement for species to interact, it is
57 57 not evidence for realized trophic interactions (Blanchet et al., 2020). We must thus be careful
58 58 when making assertions about trophic interactions based solely on co-occurrence data. Niche

1
2
3 59 theory predicts that the matching between the functional traits of predators and those of their
4 60 prey (e.g., smaller predators eat smaller prey) should improve predictions by identifying
5 61 interactions that are feasible (Morales-Castilla et al. 2015). Trait-based models have been used to
6 62 predict food webs in freshwater streams (Pomeranz et al., 2019), marine fishes (Albouy et al.,
7 63 2014), terrestrial systems comprising vertebrate (Caron et al., 2022; Fricke et al., 2022) and
8 64 invertebrate species (Laigle et al., 2018), among others. Phylogenetic relationships are also
9 65 informative since interactions and species role (i.e., species positions in the food web) tend to be
10 66 evolutionary conserved (Gómez et al., 2010; Stouffer et al., 2012). Strydom et al. (2021) recently
11 67 used a mammal phylogeny to map latent traits extracted from the European mammalian food
12 68 web to predict its Canadian counterpart. Using this approach, the authors were able to recover
13 69 90% of known trophic interactions among Canadian mammals without any prior information on
14 70 the food web.

15
16
17 71
18
19 72 It is, however still unclear how well predictive models of trophic interactions can transfer
20 73 knowledge across different regions. In Strydom et al. (2021), it is hard to know how inferences
21 74 would transfer to more contrasting environments. Europe and Canada share similar bioclimatic
22 75 conditions and, despite the few species common to both regions, more than half of Canadian
23 76 mammals have congeneric species in Europe. It is also unclear how well trophic interaction
24 77 models can predict the underlying structure of entire food webs emerging from individual trophic
25 78 interactions. Food web structure encompasses system-level properties such as connectivity or
26 79 number of trophic levels, as well as species traits such as number of prey or their position within
27 80 the food web. Most studies are aimed at predicting either the properties of food webs (e.g.,
28 81 Williams and Martinez 2008) or their interactions (e.g., Laigle et al. 2018; Pomeranz et al. 2019).

1
2
3 82 They rarely assess how well predicted interactions can recover food web structure. Despite being
4
5 83 able to recover most trophic interactions among European tetrapods, Caron et al. (2022) found
6
7 84 that food web connectance was systematically overpredicted across Europe by 2-4 times.
8
9

10 85 However, the authors did not explore how other food web properties (e.g., maximum trophic
11
12 86 level, modularity) or species positions (e.g., trophic level, centrality) were predicted.
13
14 87

15
16
17 88 Here, we test whether predictive models calibrated using a network of predator-prey interactions
18
19 89 (i.e. a food web) from one geographical region of the world can reliably predict interactions,
20
21 90 species role, and food web structure in other ecosystems across the world. Using terrestrial
22
23 91 vertebrate food webs from Europe, the Pyrenees, Northern Québec and Labrador, and the
24
25 92 Serengeti, we develop trait-based Bayesian hierarchical models to study prediction efficiency of
26
27 93 species interactions and food web properties across regions. We aim to (1) identify the factors
28
29 94 (i.e., geographic, environmental, functional, or phylogenetic distances) influencing the
30
31 95 transferability of models across ecosystems and geographical regions, (2) predict trophic roles
32
33 96 across species and (3) quantify the predicted food web properties. We expect more similar food
34
35 97 webs will yield better transferability and, based on previous results, expect the prediction of
36
37 98 pairwise interactions to be superior to that of food web properties.
38
39
40
41
42 99

43
44
45 100 **Methods**
46
47 101 *Food web data*
48
49
50 102 We extracted predator-prey interactions among terrestrial vertebrates (mammals, birds,
51
52 103 amphibians and reptiles) from four well-resolved food webs (Table 1, Figure 1): the European
53
54 104 food webs of tetrapods (Maiorano et al., 2020), the Pyrenees vertebrate food webs (Lurgi et al.,
55
56
57
58
59
60

1
2
3 105 2012), the Northern Québec and Labrador food web (Berteaux et al., 2018), and the Serengeti
4
5 106 food web (de Visser et al., 2011). All four food webs are compiled from literature review and
6
7 107 completed by expert knowledge. The four food webs document the predator-prey interactions
8
9 108 (i.e., the predator could feed on the prey species) between all terrestrial vertebrates. Trophic
10
11 109 interaction is a binary variable where 0 defines the absence of predator-prey interaction and 1 the
12
13 110 presence of predator-prey interaction between two species. The nodes in the original Serengeti
14
15 111 food web are trophic groups including one or more vertebrate species. In this study, we assumed
16
17 112 that species within a trophic group share the same predator and prey species. We describe each
18
19 113 food web in Appendix S1.

20
21 114
22
23
24 115 The species composition of the four food webs are different (Table 1). There are no amphibians
25
26 116 or reptiles in Northern Québec and Labrador, the Pyrenees food web is dominated by birds
27
28 117 (67%) and mammals (23%), with very few reptiles (8%), the European food web has a
29
30 118 comparable number of reptiles (21%) and mammals (25%), and almost half (46%) of the
31
32 119 Serengeti food web are mammals. The Europe, Pyrenees, and Northern Québec and Labrador
33
34 120 food webs all have comparable mean trophic levels (between 1.24 and 1.3) and connectance
35
36 121 (between 0.02 and 0.05) compared to the Serengeti food web (mean trophic level: 1.61;
37
38 122 connectance: 0.12). In Europe, the Pyrenees and Northern Québec and Labrador most species are
39
40 123 basal species (e.g., insectivores, herbivores, piscivores), whereas many more species feed on
41
42 124 terrestrial vertebrates (non-basal species) in the Serengeti.

43 125
44
45
46
47
48
49
50
51
52 126 *Trait data*
53
54
55
56
57
58
59
60

1
2
3 127 We extracted terrestrial vertebrate species ecological traits from the database compiled by Etard
4
5 128 et al. (2020). This dataset combines species-level information from large freely available
6
7 129 secondary trait databases (e.g., EltonTraits: Wilman et al. (2014), AmphiBIO: Oliveira et al.
8
9 130 (2017)). Overall, the database includes traits for 6 990 amphibian, 11 634 bird, 5 380 mammal,
10
11 131 and 10 612 reptile species. We extracted the body mass (mean: 11 kg; range: 0.001–4 220 kg),
12
13 132 longevity (mean: 5 885 days; range: 91–46 386 days), litter or clutch size (mean: 131 offsprings;
14
15 133 range: 1–20 000 offsprings), habitat breadth (number of habitats a species uses, using level 2 of
16
17 134 the IUCN Habitat Classification Scheme; mean: 10 habitats, range: 1–90 habitats), trophic level
18
19 135 (3 levels: herbivore, omnivore, carnivore), activity time (2 categories: nocturnal, non-nocturnal)
20
21 136 and habitat use (12 categories: forest, savanna, shrubland, grassland, wetland, rocky
22
23 137 areas/cave/subterranean, desert, marine, marine/intertidal or coastal/supratidal, artificial,
24
25 138 introduced vegetation) for all species considered in our study.

30
31 139
32
33 140 To match species in the trait databases to the species in the food webs, we standardized their
34
35 141 names following the taxonomic backbone of the Global Biodiversity Information Facility (GBIF)
36
37 142 using the function *name_backbone* from the package *rgbif* (Chamberlain et al., 2022) in the R
38
39 143 statistical language (R Core Team, 2022). In each food web, we excluded species for which no
40
41 144 taxonomic information or none of the traits were available. For species that have one or more of
42
43 145 the traits documented, we imputed missing traits with the MissForest algorithm using the
44
45 146 *missForest* R package (Stekhoven & Buehlmann, 2012) for amphibians, birds, mammals and
47
48 147 reptiles separately. MissForest uses random forests to iteratively predict missing data from the
49
50 148 known data. Each random forest uses a different trait as response variable and the remaining
51
52 149 traits as predictors.

1
2
3 150
4

5 151 *Phylogeny data*
6

7 152 We used phylogenetic data to measure how model transferability was influenced by phylogenetic
8 relatedness. We used published global phylogenies for birds (Jetz et al. 2012), amphibians (Jetz
9 153 and Pyron 2018), squamates (Tonini et al., 2016), turtles (Thomson et al., 2021), and mammals
10 154 (Upham et al., 2019). All five phylogenies are dated, were built from molecular data, and
11 155 delivered as a posterior distribution of trees. We sampled 100 trees from the posterior of each
12 156 phylogeny and calculated the mean cophenetic distance from these samples between all species
13 157 of the four food webs. Following Letten & Cornwell (2015), we square root transformed
14 158 cophenetic distances to better represent functional dissimilarity.
15 159
16 160

17 161 *Predictive models*
18

19 162 We calibrated a Bayesian hierarchical generalized linear model on each of the four food webs
20 163 (Figure 1a). The response data are trophic interactions we modelled as Bernoulli distributed.
21 164 Because Caron et al. (2022) found that trait-interaction relationships vary between predator
22 165 groups, we added varying slope and intercept terms for the taxonomic order of predators. For
23 166 each model, we randomly drew 30% of the data for validation to keep the prevalence of trophic
24 167 interaction in the validation subset equal to the prevalence of the entire food web. We used all
25 168 predator-prey interactions of the remaining 70% of the data and an equal number of absence of
26 169 interactions for calibration.
27
28 170 $L_{ij} \sim \text{Bernouilli}(p_{ij}),$
29
30 171 $\text{logit}(p_{ij}) = \alpha + \alpha_{\text{predator}[j]} + \sum_{k=1}^{13} (\beta_k \times T_k + \beta_{k,\text{predator}[j]} \times T_k),$
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 172 where L_{ij} is the occurrence of interaction between predator j and prey i , p_{ij} is the associated
4 probability of interaction, α and β are the fixed intercept and slopes of the linear model, and
5
6 173 $\alpha_{predator[j]}$ and $\beta_{k, predator[j]}$ are the random intercepts and slopes for the order of predator j .
7
8 174
9
10 175

12
13 176 We used 13 trait-based predictors in the models: five predator traits (trophic level, body mass,
14 habitat breadth, longevity, and clutch size), five prey traits (trophic level, body mass, habitat
15 breadth, longevity, and clutch size), and three trait-match predictors (match in activity time,
16 habitat use and body mass). Since the activity time trait is binary (nocturnal and non-nocturnal)
17 in Etard et al. (2020), activity time match is also a binary variable where a 1 means the predator
18 and the prey share the same activity time trait. Habitat match is the Jaccard similarity index
19 between the habitat used by the prey and the predator across the 12 habitat categories, which
20 takes into account overlap in habitat and specialization to shared habitats. Body mass match is
21 the squared difference between the log-transformed body mass of the prey and the predator,
22 because we expect predators to eat prey within a given body mass interval (small enough to be
23 handled, but large enough to be profitable. We log-transformed body mass, longevity and clutch
24 size, and scaled each continuous predictor (after transformation) by subtracting out the mean and
25 dividing by two times the standard deviation so coefficients of continuous predictors are
26 comparable to unscaled binary predictors (Gelman, 2008).

44
45 190
46
47 191 Fixed effects were drawn from a normal distribution with a mean of 0 and a standard deviation
48
49 192 of 1. Random effects were drawn from normal distributions. We used a normal distribution with
50
51 193 a mean of 0 and a standard deviation of 1 as prior for the mean and a Half-Cauchy distribution
52
53 194 with a scale parameter of 5 as prior for the standard deviation of the random effects:
54
55

1
2
3 195 $\alpha, \beta \sim Normal(0,1),$
4
5 196 $\alpha_{predator[j]}, \beta_{k, predator[j]} \sim Normal(\mu, \sigma),$
6
7 197 $\mu \sim Normal(0,1),$
8
9
10 198 $\sigma \sim HalfCauchy(0,5),$
11
12
13 199 Model parameters were estimated with Hamiltonian Monte Carlo (Neal, 2011). To fit the
14 models, we ran four chains, each with 2000 warm-up iterations, followed by 2000 iterations for
15 inference. We diagnosed convergence and adequacy with rank plots, posterior predictive checks,
16 and we calculated the rank-normalized potential scale reduction factor on split chains for all runs
17 (Vehtari et al. 2021; Appendix S5). We conducted the analyses using Stan (Carpenter et al.,
18
19 203 2017) through the package *brms* in R (Bürkner 2017).
20
21
22
23
24
25
26 205
27
28
29 206 *Predicting species interactions*
30
31
32 207 To quantify model transferability, we used each model to predict the food web fitted with that
33 model and that of the other regions considered (Figure 1a). For each possible predator-prey pair,
34 we extracted the mean of the posterior as the probability of interaction, and the standard
35 deviation of the posterior as the uncertainty around the predicted probability of interaction. When
36 predicting the food web on which the model was calibrated (i.e., within food web predictions),
37 we compared the predicted interaction probabilities to the validation subset of the food web.
38
39
40 211 When predicting food webs other than the one on which the model was calibrated (i.e., between
41 food web predictions), we compared the predicted interaction probabilities to the entire empirical
42 food web. We measured performance with the area under the receiver operating characteristic
43 curve (AUC; Hanley and McNeil 1982). AUC varies from 0 to 1 where 0.5 indicates that the
44 model failed to rank interactions higher than absences of interactions (i.e., random predictions),
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 218 and 1 indicates that the model systematically ranked interactions higher than non-interactions
4
5 219 (i.e., perfect predictions). We also measured the area under the precision-recall curve (Davis &
6
7 220 Goadrich, 2006), and directly used posterior draws to get distributions for the true positive rate,
8
9 221 true negative rate, positive predictive value and negative predictive value (Appendix S6).

10 222
11
12

13 223 To explore factors influencing model transferability, we assessed the performance of models to
14
15 224 predict each food web using three distance measures: geographic distance, environmental
16
17 225 dissimilarity, and phylogenetic relatedness. We measured geographic distance as the Euclidean
18
19 226 distance between the polygon centroid delimiting the spatial domain of each food web.
20
21 227 Environmental dissimilarity was quantified using all 19 bioclimatic variables in WorldClim
22
23 228 (Hijmans, 2021). We randomly drew 500 points within each polygon corresponding to the spatial
24
25 229 domain of our food webs, extracted bioclimatic data for these points, and used the mean of each
26
27 230 bioclimatic variable to calculate the bioclimatic centroid of each food web. We calculated
28
29 231 environmental distance as the Euclidean distance between the food web bioclimatic centroids.
30
31 232 We repeated this step 10 times (10 sets of 500 random points) to make sure the environmental
32
33 233 distance estimates are robust to random sampling (Appendix S4). To measure phylogenetic
34
35 234 relatedness, we used the mean cophenetic distance to the nearest taxon of every species in the
36
37 235 predicted food web. We averaged over all species in the predicted food web the cophenetic
38
39 236 distance to the phylogenetically closest species in the food web used for calibration. This
40
41 237 approach measures the amount of evolutionary history in the predicted food web undocumented
42
43 238 by the food web on which the model was calibrated.

44 239
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 240 Another test of the factors influencing model transferability focuses on analyzing predictive
4 performance at the species level. For each combination of predicted food web models (i.e.,
5 curves in Figure 1b), we measured how accurately the set of prey and predators of each species
6 were predicted also using the AUC. We modelled species-specific performance in terms of how
7 connected the focal species is and how distinct the focal species is to the species pool used to
8 calibrate the predictive model. To do this, we used species normalized degree (number of
9 interactions divided by the maximum possible number of interactions), the functional mean
10 pairwise distance (Mouchet et al., 2010) and distance to nearest taxon (Tucker et al., 2017).
11
12 247 Functional mean pairwise distance is the average Gower distance (Gower, 1971) between the
13 focal species and all species in the food web used for calibration. To calculate Gower distances,
14 we used all traits available in Etard (2020) through the function *funct.dist* from the R package
15 mFD (Magneville et al., 2022). Distance to nearest taxon was quantified as the cophenetic
16 distance between the focal species and the closest relative in the species in the food web used for
17 model calibration.
18
19 254
20
21 255 *Predicting species' functional roles*
22
23 256 Next, we were interested in how well species functional roles were predicted by our models. The
24 functional role of a species is determined by its position in the food web (Cirtwill et al., 2018),
25 which we quantified using five metrics related to the species centrality within the food webs
26 (number of prey, number of predators, betweenness, closeness, eigenvector centrality), two
27 metrics related to their trophic position (trophic levels and omnivory), two module-based metrics
28 (within-module degree and participation coefficient; Guimerà & Amaral, 2005), and the motif
29

1
2
3 262 profile of each species (Stouffer et al., 2012). We detail each metric, their relation to functional
4 role, and how they were calculated in Appendix S3.
5
6 264
7
8
9

10 265 To measure how well species roles were predicted, we compared each species position metric in
11 empirical food webs to the species position in predicted food webs. For predicted food webs, we
12 measured each role metric on 100 samples of the posterior distribution of the entire food web
13 with the mean as the best point estimate for the metric and the standard deviation as measure of
14 uncertainty. For each combination of model, predicted food web, and species role metric, we
15 fitted a linear regression between the predicted position and the empirical position and used the
16 coefficient of determination (R^2). We also explored prediction biases using the simple linear
17 models' coefficients. We expect an intercept of zero for perfect predictions of species role, and
18 deviation from zero would suggest systematic bias across the range of the role metric. We expect
19 a slope of one for perfect predictions. A slope less than one would suggest that the role metric of
20 species at the lower range are overpredicted, whereas the role metric of species at the upper
21 range are underpredicted (i.e., more homogeneous role across species than in the empirical food
22 web). A slope greater than one would suggest the opposite (i.e., more heterogeneous role across
23 species than in the empirical food web).
24
25 279
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

280 *Predicting food web properties*

281 Finally, we investigated how well the global properties of food webs were predicted. We selected
282 a range of metrics commonly used to quantify food web structure and which have been shown to
283 influence food web functioning and stability: connectance, mean trophic level, maximum trophic
284 level, motifs distribution, food web diameter, number of clusters, and modularity (Borrelli, 2015;
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
5510
5511
5512
5513
5514
5515
5516
5517
5518
5519
5520
5521
5522
5523
5524
5525
5526
5527
5528
5529
5530
5531
5532
5533
5534
5535
5536
5537
5538
5539
5540
5541
5542
5543
5544
5545
5546
5547
5548
5549
5550
5551
5552
5553
5554
5555
5556
5557
5558
5559
55510
55511
55512
55513
55514
55515
55516
55517
55518
55519
55520
55521
55522
55523
55524
55525
55526
55527
55528
55529
55530
55531
55532
55533
55534
55535
55536
55537
55538
55539
55540
55541
55542
55543
55544
55545
55546
55547
55548
55549
55550
55551
55552
55553
55554
55555
55556
55557
55558
55559
55560
55561
55562
55563
55564
55565
55566
55567
55568
55569
55570
55571
55572
55573
55574
55575
55576
55577
55578
55579
55580
55581
55582
55583
55584
55585
55586
55587
55588
55589
55590
55591
55592
55593
55594
55595
55596
55597
55598
55599
555100
555101
555102
555103
555104
555105
555106
555107
555108
555109
555110
555111
555112
555113
555114
555115
555116
555117
555118
555119
555120
555121
555122
555123
555124
555125
555126
555127
555128
555129
555130
555131
555132
555133
555134
555135
555136
555137
555138
555139
555140
555141
555142
555143
555144
555145
555146
555147
555148
555149
555150
555151
555152
555153
555154
555155
555156
555157
555158
555159
555160
555161
555162
555163
555164
555165
555166
555167
555168
555169
555170
555171
555172
555173
555174
555175
555176
555177
555178
555179
555180
555181
555182
555183
555184
555185
555186
555187
555188
555189
555190
555191
555192
555193
555194
555195
555196
555197
555198
555199
555200
555201
555202
555203
555204
555205
555206
555207
555208
555209
555210
555211
555212
555213
555214
555215
555216
555217
555218
555219
555220
555221
555222
555223
555224
555225
555226
555227
555228
555229
555230
555231
555232
555233
555234
555235
555236
555237
555238
555239
555240
555241
555242
555243
555244
555245
555246
555247
555248
555249
555250
555251
555252
555253
555254
555255
555256
555257
555258
555259
555260
555261
555262
555263
555264
555265
555266
555267
555268
555269
555270
555271
555272
555273
555274
555275
555276
555277
555278
555279
555280
555281
555282
555283
555284
555285
555286
555287
555288
555289
555290
555291
555292
555293
555294
555295
555296
555297
555298
555299
555300
555301
555302
555303
555304
555305
555306
555307
555308
555309
555310
555311
555312
555313
555314
555315
555316
555317
555318
555319
555320
555321
555322
555323
555324
555325
555326
555327
555328
555329
555330
555331
555332
555333
555334
555335
555336
555337
555338
555339
555340
555341
555342
555343
555344
555345
555346
555347
555348
555349
555350
555351
555352
555353
555354
555355
555356
555357
555358
555359
555360
555361
555362
555363
555364
555365
555366
555367
555368
555369
555370
555371
555372
555373
555374
555375
555376
555377
555378
555379
555380
555381
555382
555383
555384
555385
555386
555387
555388
555389
555390
555391
555392
555393
555394
555395
555396
555397
555398
555399
555400
555401
555402
555403
555404
555405
555406
555407
555408
555409
555410
555411
555412
555413
555414
555415
555416
555417
555418
555419
555420
555421
555422
555423
555424
555425
555426
555427
555428
555429
555430
555431
555432
555433
555434
555435
555436
555437
555438
555439
555440
555441
555442
555443
555444
555445
555446
555447
555448
555449
555450
555451
555452
555453
555454
555455
555456
555457
555458
555459
555460
555461
555462
555463
555464
555465
555466
555467
555468
555469
555470
555471
555472
555473
555474
555475
555476
555477
555478
555479
555480
555481
555482
555483
555484
555485
555486
555487
555488
555489
555490
555491
555492
555493
555494
555495
555496
555497
555498
555499
555500
555501
555502
555503
555504
555505
555506
555507
555508
555509
555510
555511
555512
555513
555514
555515
555516
555517
555518
555519
555520
555521
555522
555523
555524
555525
555526
555527
555528
555529
555530
555531
555532
555533
555534
555535
555536
555537
555538
555539
555540
555541
555542
555543
555544
555545
555546
555547
555548
555549
555550
555551
555552
555553
555554
555555
555556
555557
555558
555559
555560
555561
555562
555563
555564
555565
555566
555567
555568
555569
555570
555571
555572
555573
555574
555575
555576
555577
555578
555579
555580
555581
555582
555583
555584
555585
555586
555587
555588
555589
555590
555591
555592
555593
555594
555595
555596
555597
555598
555599
5555100
5555101
5555102
5555103
5555104
5555105
5555106
5555107
5555108
5555109
5555110
5555111
5555112
5555113
5555114
5555115
5555116
5555117
5555118
5555119
5555120
5555121
5555122
5555123
5555124
5555125
5555126
5555127
5555128
5555129
5555130
5555131
5555132
5555133
5555134
5555135
5555136
5555137
5555138
5555139
5555140
5555141
5555142
5555143
5555144
5555145
5555146
5555147
5555148
5555149
5555150
5555151
5555152
5555153
5555154
5555155
5555156
5555157
5555158
5555159
5555160
5555161
5555162
5555163
5555164
5555165
5555166
5555167
5555168
5555169
5555170
5555171
5555172
5555173
5555174
5555175
5555176
5555177
5555178
5555179
5555180
5555181
5555182
5555183
5555184
5555185
5555186
5555187
5555188
5555189
5555190
5555191
5555192
5555193
5555194
5555195
5555196
5555197
5555198
5555199
5555200
5555201
5555202
5555203
5555204
5555205
5555206
5555207
5555208
5555209
5555210
5555211
5555212
5555213
5555214
5555215
5555216
5555217
5555218
5555219
5555220
5555221
5555222
5555223
5555224
5555225
5555226
5555227
5555228
5555229
5555230
5555231
5555232
5555233
5555234
5555235
5555236
5555237
5555238
5555239
5555240
5555241
5555242
5555243
5555244
5555245
5555246
5555247
5555248
5555249
5555250
5555251
5555252
5555253
5555254
5555255
5555256
5555257
5555258
5555259
5555260
5555261
5555262
5555263
5555264
5555265
5555266
5555267
5555268
5555269
5555270
5555271
5555272
5555273
5555274
5555275
5555276
5555277
5555278
5555279
5555280
5555281
5555282
5555283
5555284
5555285
5555286
5555287
5555288
5555289
5555290
5555291
5555292
5555293
5555294
5555295
5555296
5555297
5555298
5555299
5555300
5555301
5555302
5555303
5555304
5555305
5555306
5555307
5555308
5555309
5555310
5555311
5555312
5555313
5555314
5555315
5555316
5555317
5555318
5555319
5555320
5555321
5555322
5555323
5555324
5555325
5555326
5555327
5555328
5555329
5555330
5555331
5555332
5555333
5555334
5555335
5555336
5555337
5555338
5555339
5555340
5555341
5555342
5555343
5555344
5555345
5555346
5555347
5555348
5555349
5555350
5555351
5555352
5555353
5555354
5555355
5555356
5555357
5555358
5555359
5555360
5555361
5555362
5555363
5555364
5555365
5555366
5555367
5555368
5555369
5555370
5555371
5555372
5555373
5555374
5555375
5555376
5555377
5555378
5555379
5555380
5555381
5555382
5555383
5555384
5555385
5555386
5555387
5555388
5555389
5555390
5555391
5555392
5555393
5555394
5555395
5555396
5555397
5555398
5555399
5555400
5555401
5555402
5555403
5555404
5555405
5555406
5555407
5555408
5555409
5555410
5555411
5555412
5555413
5555414
5555415
5555416
5555417
5555418
5555419
5555420
5555421
5555422
5555423
5555424
5555425
5555426
5555427
5555428
5555429
5555430
5555431
5555432
5555433
5555434
5555435
5555436
5555437
5555438
5555439
5555440
5555441
5555442
5555443
5555444
5555445
5555446
5555447
5555448
5555449
5555450
5555451
5555452
5555453
5555454
5555455
5555456
5555457
5555458
5555459
5555460
5555461
5555462
5555463
5555464
5555465
5555466
5555467
5555468
5555469
5555470
5555471
5555472
5555473
5555474
5555475
5555476
5555477
5555478
5555479
5555480
5555481
5555482
5555483
5555484
5555485
5555486
5555487
5555488
5555489
5555490
5555491
5555492
5555493
5555494
5555495
5555496
5555497
5555498
5555499
5555500
5555501
5555502
5555503
5555504
5555505
5555506
5555507
5555508
5555509
5555510
5555511
5555512
5555513
5555514
5555515
5555516
5555517
5555518
5555519
5555520
5555521
5555522
5555523
5555524
5555525
5555526
5555527
5555528
5555529
5555530
5555531
5555532
5555533
5555534
5555535
5555536
5555537
5555538
5555539
5555540
5555541
5555542
5555543
5555544
5555545
5555546
5555547
5555548
5555549
5555550
5555551
5555552
5555553
5555554
5555555
5555556
5555557
5555558
5555559
55555510
55555511
55555512
55555513
55555514
55555515
55555516
55555517
55555518
55555519
55555520
5555

1
2
3 285 Vermaat et al., 2009). As for species position, we evaluated these properties on the empirical
4
5 286 food webs and compared them to the properties predicted using the mean of 100 samples of the
6
7 287 posterior food web prediction. We detail each metric, their relation to food web function, and
8
9 288 how they were calculated in Appendix S3.
10
11 289

12
13 290 **Results**
14
15

16 291 *Predicting trophic interactions*
17
18

19 292 For all food webs, trophic interactions were better predicted by the model calibrated on the same
20
21 293 food web (within food web predictions) than by models calibrated on other food webs (between
22
23 294 food web predictions; Table 2). For within food web predictions, AUC varied between 0.92 and
24
25 295 0.96. Model performance was also good ($AUC > 0.82$) for transfer between the Europe,
26
27 Pyrenees, and Northern Québec and Labrador food webs. Models did not transfer as well from
28
29 and to the Serengeti food web, but performance was still good ($AUC > 0.75$). The area under the
30
31 298 precision-recall curve, true positive rate, true negative rate, positive predictive value, and
32
33 299 negative predictive value were all positively correlated with AUC and showed the same overall
34
35 300 pattern (Appendix S6).
36
37
38

39 301
40
41 302 To explore the factors influencing the transferability of interaction models, we modelled their
42
43 303 performance relative to the geographic, environmental, and phylogenetic distances between the
44
45 304 calibrated and predicted food web (Figure 2). Overall, performance tended to decrease with
46
47 305 environmental (direct effect estimate: -0.52, 95%CrI = [-0.89, -0.11]) and phylogenetic distance
48
49 306 (direct effect estimate: -0.39, 95%CrI = [-0.78, -0.03]). Performance also decreased with
50
51 307 geographic distance (total effect estimate: -0.54, 95%CrI = [-0.79, -0.30]), but this effect
52
53
54
55

1
2
3 308 disappeared after controlling for phylogenetic and environmental distances (direct effect
4
5 309 estimate: -0.07, 95%CrI = [-0.48, 0.34]).
6
7 310
8
9
10 311 We also identified the species for which the interactions were incorrectly predicted between food
11
12 312 web. We modelled species-specific performance to the phylogenetic and functional distance
13
14 313 between the focal species and the species pool used to train the models (Figure 3). As expected,
15
16 314 species with phylogenetically close relatives in the species pool used for calibration were, on
17
18 315 average, better predicted than distant relatives (Figure 3a). Species-specific performance slightly
19
20 316 decreased at low and intermediate phylogenetic distances, and then dropped significantly at large
21
22 317 distances. Models calibrated and predicted across classes (e.g. mammals to amphibians) had the
23
24 318 lowest performance, as expected. This situation only occurs with the Northern Québec and
25
26 319 Labrador model predicting the other three food webs as there are no amphibians or reptiles in
27
28 320 Northern Québec and Labrador. Surprisingly, predictive performance remained qualitatively
29
30 321 unchanged by functional distance (Figure 3b). We also found that interactions of specialist (i.e.,
31
32 322 species with few interactions) and generalist species (i.e., species with many interactions) were,
33
34 323 on average, better predicted than interactions of species of intermediate specialization (Figure
35
36 324 3c).
37
38
39
40
41
42 325
43
44
45 326 *Predicting species functional role*
46
47 327 We found significant variation in how well species trophic positions were predicted across
48
49 328 models (Figure 4). Species roles were slightly better predicted by within food web predictions
50
51 329 than by between food web predictions. Interestingly, some measures of centrality (betweenness
52
53
54
55
56
57
58
59
60

1
2
3 330 and closeness) were not well predicted, whereas others (number of prey and predators,
4
5 331 eigenvector centrality) were relatively well predicted.
6
7 332
8
9
10 333 We also found important biases in the predictions of species roles when we fitted linear
11
12 334 regressions between species roles in predicted food webs to their roles in the empirical food
13
14 335 webs. The number of prey, predators, trophic level, omnivory, and the frequency of motif
15
16 336 positions tended to be similarly overestimated across species (slopes close to 1, and intercept
17
18 337 greater than 0). For other measures such as eigenvector centrality, betweenness, closeness, and
19
20 338 module-based roles, species had more similar values between predicted and empirical food webs
21
22
23 339 (Appendix S7, slopes less than 1).
24
25
26 340
27
28
29 341 *Predicting food web properties*
30
31 342 The majority of food web properties were not well predicted by our interaction models (Figure
32
33 343 5). Connectance, mean and max trophic levels, and the frequency of most motifs were
34
35 344 overpredicted, whereas modularity was slightly underpredicted. Mean and variance of relative
36
37 345 errors were greater for between- compared to within-food web predictions. In general, predicted
38
39 346 food webs were more connected, displayed a higher frequency of most motifs, and were less
40
41 347 modular and with less basal species, which increased the trophic level of most species.
42
43
44
45 348
46
47 349 **Discussion**
48
49
50 350 Predictive models of trophic interactions have recently become central in filling knowledge gaps
51
52 351 about how predator-prey interactions vary across space and time. Here, we showed that trait-
53
54 352 based trophic interaction models can predict interactions across ecosystems. We found that,
55
56
57
58
59
60

1
2
3 353 given enough phylogenetic and environmental similarities between the system on which the
4
5 354 model is calibrated and the system for which the predictions are made, models predicted most
6
7 355 interactions reasonably well. Our results suggest that, for terrestrial vertebrate food webs, trait
8
9 356 relationships driving interactions appear to be relatively general even in highly contrasting
10
11 357 environments. Although models were successful at predicting interactions, they were less
12
13 358 successful at predicting higher-level food web properties. We found systematic biases in the
14
15 359 species position and food web properties predictions. Biases varied across metrics, but overall,
16
17 360 the predicted food webs were more highly connected, less modular, had more trophic levels, with
18
19 361 species within them being more homogeneously connected than their observed counterparts.
20
21
22 362 These higher-level properties of food webs were especially poorly predicted when making
23
24 363 between food web predictions (i.e., knowledge transfer).
25
26
27 364
28
29
30 365 The trait-matching framework of trophic niche theory assumes that variation across food webs
31
32 366 arises through differences in the trait-matching rules driving interactions and the distribution of
33
34 367 traits in different systems (Gravel et al., 2016). Our ability to predict food web interactions
35
36 368 across contrasting environments partially agrees with this framework by suggesting that spatial
37
38 369 food web variation is mainly driven by changes in the distribution of functional traits, and less so
39
40 370 by the variation of trait-matching relationships. These results are in line with previous research
41
42 371 finding generalities made on trait-interactions relationships across European bioregions (Caron et
43
44 372 al., 2022), the predator-prey body-size ratios within habitat, predator, and prey types (Brose et
45
46 373 al., 2006), and the trait-interactions relationships in soil invertebrates across three forest areas in
47
48 374 Germany (Laigle et al. 2018). Given that trait-matching rules driving species interactions seem
49
50 375 general in reasonably similar environments, it should be possible to use data from well-studied
51
52
53
54
55

1
2
3 376 areas to predict interactions in areas we know very little about or forecast (and hindcast) food
4
5 377 webs given new trait distributions.
6
7
8 378
9
10

11 379 We also highlighted a major limitation of trophic interaction models: trophic interaction models
12
13 380 predict interactions, not food webs. Even if most interactions and absences of interactions are
14
15 381 well predicted (high true positive and negative rates), there are many more absences of
16
17 382 interactions to predict than presences in real food webs (low connectance). This might explain
18
19 383 why our models systematically overpredicted the number of interactions (number of prey and
20
21 384 predators) across species (Appendix S7). This bias probably propagated through the food webs,
22
23
24 385 explaining why the centrality of species was more evenly distributed, and why the predicted food
25
26 386 webs were more connected, less modular, and with higher trophic levels.
27
28
29 387
30
31

32 388 The structural properties of food webs (i.e., connectance, number of trophic levels, modularity)
33
34 389 influence the stability, invasibility, and productivity of ecosystems (Duffy et al., 2007; Lurgi et
35
36 390 al., 2014; Wang & Brose, 2018), whereas the position of species within food webs determine
37
38 391 their functional role, and can help identify keystone species and prevent cascading effects of
39
40 392 extinction (Cirtwill et al., 2018; Estes et al., 2011). Here, predicting individual links failed to
41
42 393 predict higher-level properties, suggesting that there are constraints acting on the structures of
43
44 394 food webs that trophic interaction models cannot capture. A consequence of such constraints is
45
46 395 the spatial and temporal variations of ecological network, which have gained a lot of interest
47
48 396 recently (Baiser et al., 2019; Gravel et al., 2019). Because sampling interactions at large scales is
49
50 397 difficult, predictions by interaction models could help investigate the variation of interactions
51
52
53 398 and network structures simultaneously (e.g., Albouy et al., 2014), which would be possible only
54
55
56
57
58
59
60

1
2
3 399 if the biases in predicted network structure are constant across the gradient of interest. Given our
4
5 400 results, testing the assumption that the bias is constant should be necessary to robustly measure
6
7 401 the variation in network structure.
8
9
10 402

11
12 403 Correcting biases in higher-level property predictions presents an opportunity to improve species
13
14 404 interaction predictions. In species distribution models, methods have been developed to harness
15
16 405 biases in higher-level properties (e.g., species richness) to correct distribution predictions (e.g.,
17
18 406 Leung et al., 2019). These models correct systematic biases in predictions similar to those we
19
20 found in our study. Therefore, methods that would combine predictions of interactions and
21
22 networks have the potential to provide better food web predictions. Structural food web models
23
24 409 can predict the probability distribution of many food web properties (Williams & Martinez,
25
26 410 2008). Information on probable food web structure could be used to correct posterior predictive
27
28 411 distributions of species interactions.
29
30
31 412

32
33 413 Our study suffers from a few limitations that, if overcome, could move us closer to a
34
35 414 comprehensive framework for ecological interactions and networks prediction. First, our study
36
37 415 relied on terrestrial vertebrates. Although there is no clear *a priori* reason not to be applicable in
38
39 other systems, our results are unlikely to be general across all taxa and types of interactions. We
40
41 416 are not aware of another other test of interaction model transferability, but it would be interesting
42
43 417 to investigate if our results hold for systems where trait-matching relationships are stronger or
44
45 418 weaker. The exclusion of non-vertebrates (e.g., plants, invertebrates, parasites) also influence the
46
47 419 empirical and predicted species role and food web properties. For example, the first trophic
48
49 420 levels in our food webs were not primary producers, but species not feeding on terrestrial
50
51 421 levels in our food webs were not primary producers, but species not feeding on terrestrial
52
53
54
55

1
2
3 422 vertebrates (e.g., herbivores, invertivores). However, extension of trait-matching models to also
4 423 include invertebrates can be achieved with a coarser resolution (Li et al., 2023) or with additional
5 traits (Laigle et al., 2018). Second, the food webs we used were potential and binary food webs.
6
7 424 Trait-matching models predict the probabilities that a species could eat another species given
8 they are encountering each other. Additional data, such as co-occurrence and abundance data, are
9 needed to make predictions of realized and quantitative interactions. Finally, due to the scarcity
10 of food web data, we only had four food webs to work with. This means we only had four sets of
11 within food web predictions, and 12 sets of between food web predictions. This explains the
12 large uncertainty for some of our results (e.g., Figure 2). Despite the low sample size, we still
13 detected significant relationship between model transferability and geographic, phylogenetic, and
14 environmental distances.

15
16 433
17
18 434 Overall, we found that trait-based interaction models can transfer knowledge relatively well
19 given enough phylogenetic and environmental similarities between systems. They are useful for
20 producing initial expectations even in contrasting ecosystems. This, and other recent research,
21 suggests there are fundamental trait-based constraints on trophic interactions that are
22 generalizable to some extent. The nature of this constraint also appears to result from species
23 traits rather than the ‘match’ between a predator and its prey, which has interesting ties to the
24 broader question of how traits vary between ecosystems and along gradients (Gravel et al.,
25 2016). Future research could better link subfields within trait-based ecology for a comprehensive
26 understanding of how species relate to their environment and to other species, and to enable
27 better predictions of the responses of species and ecosystems to threats and global change.

28 444
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 445 References
4
5

- 6 Albouy, C., Velez, L., Coll, M., Colloca, F., Loc'h, F. L., Mouillot, D., & Gravel, D. (2014).
7 From projected species distribution to food-web structure under climate change. *Global
8 Change Biology*, 20(3), 730–741. <https://doi.org/10.1111/gcb.12467>
- 9
10 Baiser, B., Gravel, D., Cirtwill, A. R., Dunne, J. A., Fahimipour, A. K., Gilarranz, L. J.,
11 Grochow, J. A., Li, D., Martinez, N. D., McGrew, A., Poisot, T., Romanuk, T. N.,
12 Stouffer, D. B., Trotta, L. B., Valdovinos, F. S., Williams, R. J., Wood, S. A., & Yeakel,
13 J. D. (2019). Ecogeographical rules and the macroecology of food webs. *Global Ecology
14 and Biogeography*, 28(9), 1204–1218. <https://doi.org/10.1111/geb.12925>
- 15
16 Berteaux, D., Ropars, P., & Casajus, N. (2018). Toundra Nunavik: Matrice des relations
17 trophiques entre espèces du Labrador et du Québec nordique, v. 1.0 (1980-2010).
18
19 *Nordicana D36*. <https://doi.org/10.5885/45555CE-DA1FF11FA4254703>
- 20
21 Blanchet, F. G., Cazelles, K., & Gravel, D. (2020). Co-occurrence is not evidence of ecological
22 interactions. *Ecology Letters*, 23(7), 1050–1063. <https://doi.org/10.1111/ele.13525>
- 23
24 Borrelli, J. J. (2015). Selection against instability: Stable subgraphs are most frequent in
25 empirical food webs. *Oikos*, 124(12), 1583–1588. <https://doi.org/10.1111/oik.02176>
- 26
27 Brose, U., Jonsson, T., Berlow, E. L., Warren, P., Banasek-Richter, C., Bersier, L.-F., Blanchard,
28 J. L., Brey, T., Carpenter, S. R., Blandenier, M.-F. C., Cushing, L., Dawah, H. A., Dell,
29 T., Edwards, F., Harper-Smith, S., Jacob, U., Ledger, M. E., Martinez, N. D., Memmott,
30 J., ... Cohen, J. E. (2006). Consumer–Resource Body-Size Relationships in Natural Food
31 Webs. *Ecology*, 87(10), 2411–2417. [https://doi.org/10.1890/0012-9658\(2006\)87\[2411:CBRINF\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[2411:CBRINF]2.0.CO;2)
- 32
33 Bürkner, P.-C. (2017). **brms**: An *R* Package for Bayesian Multilevel Models Using *Stan*. *Journal
34 of Statistical Software*, 80(1). <https://doi.org/10.18637/jss.v080.i01>
- 35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3
4 Canard, E., Mouquet, N., Marescot, L., Gaston, K. J., Gravel, D., & Mouillot, D. (2012).
5
6 Emergence of Structural Patterns in Neutral Trophic Networks. *PLOS ONE*, 7(8),
7
8 e38295. <https://doi.org/10.1371/journal.pone.0038295>
9
10 Caron, D., Maiorano, L., Thuiller, W., & Pollock, L. J. (2022). Addressing the Eltonian shortfall
11 with trait-based interaction models. *Ecology Letters*, ele.13966.
12
13 <https://doi.org/10.1111/ele.13966>
14
15 Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker,
16 M., Guo, J., Li, P., & Riddell, A. (2017). *Stan: A Probabilistic Programming Language*.
17
18 *Journal of Statistical Software*, 76(1). <https://doi.org/10.18637/jss.v076.i01>
19
20 Chamberlain, S., Barve, V., Mcglinn, D., Oldoni, D., Desmet, P., Geffert, L., & Ram, K. (2022).
21
22 *rgbif: Interface to the global biodiversity information facility API [Manual]*.
23
24 <https://CRAN.R-project.org/package=rgbif>
25
26 Cirtwill, A. R., Dalla Riva, G. V., Gaiarsa, M. P., Bimler, M. D., Cagua, E. F., Coux, C., &
27 Dehling, D. M. (2018). A review of species role concepts in food webs. *Food Webs*, 16,
28 e00093. <https://doi.org/10.1016/j.fooweb.2018.e00093>
29
30 Davis, J., & Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves.
31
32 *Proceedings of the 23rd International Conference on Machine Learning - ICML '06*,
33 233–240. <https://doi.org/10.1145/1143844.1143874>
34
35 de Visser, S. N., Freymann, B. P., & Olff, H. (2011). The Serengeti food web: Empirical
36 quantification and analysis of topological changes under increasing human impact.
37
38 *Journal of Animal Ecology*, 80(2), 484–494. <https://doi.org/10.1111/j.1365-2656.2010.01787.x>
39
40 Duffy, J. E., Cardinale, B. J., France, K. E., McIntyre, P. B., Thébault, E., & Loreau, M. (2007).
41
42 The functional role of biodiversity in ecosystems: Incorporating trophic complexity.
43
44 *Ecology Letters*, 10(6), 522–538. <https://doi.org/10.1111/j.1461-0248.2007.01037.x>

- 1
2
3
4 Estes, J. A., Terborgh, J., Brashares, J. S., Power, M. E., Berger, J., Bond, W. J., Carpenter, S.
5 R., Essington, T. E., Holt, R. D., Jackson, J. B. C., Marquis, R. J., Oksanen, L., Oksanen,
6 T., Paine, R. T., Pikitch, E. K., Ripple, W. J., Sandin, S. A., Scheffer, M., Schoener, T.
7 W., ... Wardle, D. A. (2011). Trophic Downgrading of Planet Earth. *Science*, 333(6040),
8 301–306. <https://doi.org/10.1126/science.1205106>
- 9
10 Etard, A., Morrill, S., & Newbold, T. (2020). Global gaps in trait data for terrestrial vertebrates.
11
12 *Global Ecology and Biogeography*, 29(12), 2143–2158.
13
14 <https://doi.org/10.1111/geb.13184>
- 15 Fricke, E. C., Hsieh, C., Middleton, O., Gorczynski, D., Cappello, C. D., Sanisidro, O., Rowan,
16 J., Svenning, J.-C., & Beaudrot, L. (2022). Collapse of terrestrial mammal food webs
17 since the Late Pleistocene. *Science*, 377(6609), 1008–1011.
18
19 <https://doi.org/10.1126/science.abn4012>
- 20 Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics*
21
22 in *Medicine*, 27(15), 2865–2873. <https://doi.org/10.1002/sim.3107>
- 23 Gómez, J. M., Verdú, M., & Perfectti, F. (2010). Ecological interactions are evolutionarily
24 conserved across the entire tree of life. *Nature*, 465(7300), Art. 7300.
25
26 <https://doi.org/10.1038/nature09113>
- 27 Gower, J. C. (1971). A General Coefficient of Similarity and Some of Its Properties. *Biometrics*,
28
29 27(4), 857–871. <https://doi.org/10.2307/2528823>
- 30 Gravel, D., Albouy, C., & Thuiller, W. (2016). The meaning of functional trait composition of
31 food webs for ecosystem functioning. *Philosophical Transactions of the Royal Society B:*
32
33 *Biological Sciences*, 371(1694), 20150268. <https://doi.org/10.1098/rstb.2015.0268>
- 34 Gravel, D., Baiser, B., Dunne, J. A., Kopelke, J.-P., Martinez, N. D., Nyman, T., Poisot, T.,
35 Stouffer, D. B., Tylianakis, J. M., Wood, S. A., & Roslin, T. (2019). Bringing Elton and
36 Grinnell together: A quantitative framework to represent the biogeography of ecological
37 interaction networks. *Ecography*, 42(3), 401–415. <https://doi.org/10.1111/ecog.04006>
- 38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3
4 Guimerà, R., & Amaral, L. A. N. (2005). Cartography of complex networks: Modules and
5 universal roles. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(02),
6 P02001. <https://doi.org/10.1088/1742-5468/2005/02/P02001>
- 7
8 Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver
9 operating characteristic (ROC) curve. *Radiology*, 143(1), 29–36.
10
11 <https://doi.org/10.1148/radiology.143.1.7063747>
- 12
13 Harvey, E., Gounand, I., Ward, C. L., & Altermatt, F. (2017). Bridging ecology and
14 conservation: From ecological networks to ecosystem function. *Journal of Applied
15 Ecology*, 54(2), 371–379. <https://doi.org/10.1111/1365-2664.12769>
- 16
17 Hijmans, R. J. (2021). *raster: Geographic data analysis and modeling* [Manual].
18
19 <https://CRAN.R-project.org/package=raster>
- 20
21 Jetz, W., & Pyron, R. A. (2018). The interplay of past diversification and evolutionary isolation
22 with present imperilment across the amphibian tree of life. *Nature Ecology & Evolution*,
23 2(5), Art. 5. <https://doi.org/10.1038/s41559-018-0515-5>
- 24
25 Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., & Mooers, A. O. (2012). The global diversity
26 of birds in space and time. *Nature*, 491(7424), Art. 7424.
27
28 <https://doi.org/10.1038/nature11631>
- 29
30 Jordano, P. (2016). Sampling networks of ecological interactions. *Functional Ecology*, 30(12),
31 1883–1893. <https://doi.org/10.1111/1365-2435.12763>
- 32
33 Laigle, I., Aubin, I., Digel, C., Brose, U., Boulangeat, I., & Gravel, D. (2018). Species traits as
34 drivers of food web structure. *Oikos*, 127(2), 316–326. <https://doi.org/10.1111/oik.04712>
- 35
36 Letten, A. D., & Cornwell, W. K. (2015). Trees, branches and (square) roots: Why evolutionary
37 relatedness is not linearly related to functional distance. *Methods in Ecology and
38 Evolution*, 6(4), 439–444. <https://doi.org/10.1111/2041-210X.12237>
- 39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3
4 Leung, B., Hudgins, E. J., Potapova, A., & Ruiz-Jaen, M. C. (2019). A new baseline for
5 countrywide α -diversity and species distributions: Illustration using >6,000 plant species
6 in Panama. *Ecological Applications*, 29(3), e01866. <https://doi.org/10.1002/eap.1866>
- 7
8
9
10 Li, J., Luo, M., Wang, S., Gauzens, B., Hirt, M. R., Rosenbaum, B., & Brose, U. (2023). A size-
11 constrained feeding-niche model distinguishes predation patterns between aquatic and
12 terrestrial food webs. *Ecology Letters*, 26(1), 76–86. <https://doi.org/10.1111/ele.14134>
- 13
14 Lindeman, R. L. (1942). The Trophic-Dynamic Aspect of Ecology. *Ecology*, 23(4), 399–417.
15
16 https://doi.org/10.2307/1930126
- 17
18 Lurgi, M., Galiana, N., López, B. C., Joppa, L. N., & Montoya, J. M. (2014). Network
19 complexity and species traits mediate the effects of biological invasions on dynamic food
20 webs. *Frontiers in Ecology and Evolution*, 2.
21
22 https://www.frontiersin.org/articles/10.3389/fevo.2014.00036
- 23
24 Lurgi, M., López, B. C., & Montoya, J. M. (2012). Climate change impacts on body size and
25 food web structure on mountain ecosystems. *Philosophical Transactions of the Royal
26 Society B: Biological Sciences*, 367(1605), 3050–3057.
27
28 https://doi.org/10.1098/rstb.2012.0239
- 29
30 Magneville, C., Loiseau, N., Albouy, C., Casajus, N., Claverie, T., Escalas, A., Leprieur, F.,
31 Maire, E., Mouillot, D., & Villéger, S. (2022). mFD: An R package to compute and
32 illustrate the multiple facets of functional diversity. *Ecography*, 2022(1).
33
34 https://doi.org/10.1111/ecog.05904
- 35
36 Maiorano, L., Montemaggiori, A., Ficetola, G. F., O'Connor, L., & Thuiller, W. (2020).
37
38 TETRA-EU 1.0: A species-level trophic metaweb of European tetrapods. *Global Ecology
39 and Biogeography*, geb.13138. <https://doi.org/10.1111/geb.13138>
- 40
41 Morales-Castilla, I., Matias, M. G., Gravel, D., & Araújo, M. B. (2015). Inferring biotic
42 interactions from proxies. *Trends in Ecology & Evolution*, 30(6), 347–356.
43
44 https://doi.org/10.1016/j.tree.2015.03.014
- 45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- Mouchet, M. A., Villéger, S., Mason, N. W. H., & Mouillot, D. (2010). Functional diversity measures: An overview of their redundancy and their ability to discriminate community assembly rules. *Functional Ecology*, 24(4), 867–876. <https://doi.org/10.1111/j.1365-2435.2010.01695.x>
- Neal, R. M. (2011). MCMC Using Hamiltonian Dynamics. In *Handbook of Markov Chain Monte Carlo* (pp. 113–162). CRC Press. <https://doi.org/10.1201/b10905-7>
- 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(1), Art. 1. <https://doi.org/10.1038/sdata.2017.123>
- Poisot, T., Bergeron, G., Cazelles, K., Dallas, T., Gravel, D., MacDonald, A., Mercier, B., Violet, C., Vissault, S., & Chapman, D. (2021). Global knowledge gaps in species interaction networks data. *Journal of Biogeography*, jbi.14127. <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>
- R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- 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., 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>
- Strydom, T., Bouskila, S., Banville, F., Barros, C., Caron, D., Farrell, M. J., Fortin, M.-J., Hemming, V., Mercier, B., Pollock, L., Runghen, R., Riva, G. V. D., & Poisot, T. (2021).

- 1
2
3
4 *Food web reconstruction through phylogenetic transfer of low-rank network*
5 *representation*. EcoEvoRxiv. <https://doi.org/10.32942/osf.io/y7sdz>
- 6
7
8 Thomson, R. C., Spinks, P. Q., & Shaffer, H. B. (2021). A global phylogeny of turtles reveals a
9 burst of climate-associated diversification on continental margins. *Proceedings of the*
10 *National Academy of Sciences*, 118(7), e2012215118.
11
12 <https://doi.org/10.1073/pnas.2012215118>
- 13
14 Tonini, J. F. R., Beard, K. H., Ferreira, R. B., Jetz, W., & Pyron, R. A. (2016). Fully-sampled
15 phylogenies of squamates reveal evolutionary patterns in threat status. *Biological*
16 *Conservation*, 204, 23–31. <https://doi.org/10.1016/j.biocon.2016.03.039>
- 17
18 Tucker, C. M., Cadotte, M. W., Carvalho, S. B., Davies, T. J., Ferrier, S., Fritz, S. A., Grenyer,
19 R., Helmus, M. R., Jin, L. S., Mooers, A. O., Pavoine, S., Purschke, O., Redding, D. W.,
20 Rosauer, D. F., Winter, M., & Mazel, F. (2017). A guide to phylogenetic metrics for
21 conservation, community ecology and macroecology. *Biological Reviews*, 92(2), 698–
22 715. <https://doi.org/10.1111/brv.12252>
- 23
24 Upham, N. S., Esselstyn, J. A., & Jetz, W. (2019). Inferring the mammal tree: Species-level sets
25 of phylogenies for questions in ecology, evolution, and conservation. *PLOS Biology*,
26 17(12), e3000494. <https://doi.org/10.1371/journal.pbio.3000494>
- 27
28 Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P.-C. (2021). Rank-
29 Normalization, Folding, and Localization: An Improved R^+ for Assessing Convergence of
30 MCMC (with Discussion). *Bayesian Analysis*, 16(2), 667–718.
31
32 <https://doi.org/10.1214/20-BA1221>
- 33
34 Vermaat, J. E., Dunne, J. A., & Gilbert, A. J. (2009). Major dimensions in food-web structure
35 properties. *Ecology*, 90(1), 278–282. <https://doi.org/10.1890/07-0978.1>
- 36
37 Wang, S., & Brose, U. (2018). Biodiversity and ecosystem functioning in food webs: The
38 vertical diversity hypothesis. *Ecology Letters*, 21(1), 9–20.
39
40 <https://doi.org/10.1111/ele.12865>

1
2
3
4 Williams, R. J., & Martinez, N. D. (2008). Success and Its Limits among Structural Models of
5 Complex Food Webs. *Journal of Animal Ecology*, 77(3), 512–519.
6
7

8 Wilman, H., Belmaker, J., Simpson, J., Rosa, C. de la, Rivadeneira, M. M., & Jetz, W. (2014).
9 EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals.
10
11 *Ecology*, 95(7), 2027–2027. <https://doi.org/10.1890/13-1917.1>
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

1

2

3 **446 Data Accessibility Statement**

4

5 447 All data sources have been previously published. All data used in the analyses and relevant code

6

7 448 will be archived on an Open Science Framework repository.

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

For Peer Review

1
2
3 449
4
5

6 *Table 1: Summary of properties of the food webs used in this study after excluding species for*
7 *which no match was found in the GBIF Backbone Taxonomy or was not documented in the trait*
8 *database. Trait coverage is the percentage of traits documented in Etard et al. (2020) before*
9 *imputing missing traits.*

	Europe	Pyrenees	North Québec and Labrador	Serengeti
No. of species	1135	196	200	298
Prop of amphibians	0.09	0.02	0	0.06
Prop of birds	0.45	0.67	0.78	0.35
Prop of mammals	0.25	0.23	0.22	0.46
Prop of reptiles	0.21	0.08	0	0.13
No. of interactions	57 746	831	1 098	11 038
Connectance	0.05	0.02	0.03	0.12
Mean trophic level	1.24	1.25	1.3	1.61
Trait coverage (%)	83.9	92.1	96.3	81.9

25 451
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

452

Table 2: Food webs are better predicted by their own calibrated model. Area under the receiver operating curve (AUC) of each model predicting every food web.

		Food web			
		Europe	Pyrenees	North Québec and Labrador	Serengeti
Model					
	Europe	0.96	0.89	0.9	0.8
	Pyrenees	0.86	0.95	0.85	0.79
	North Québec and Labrador	0.82	0.9	0.95	0.75
	Serengeti	0.85	0.78	0.77	0.92

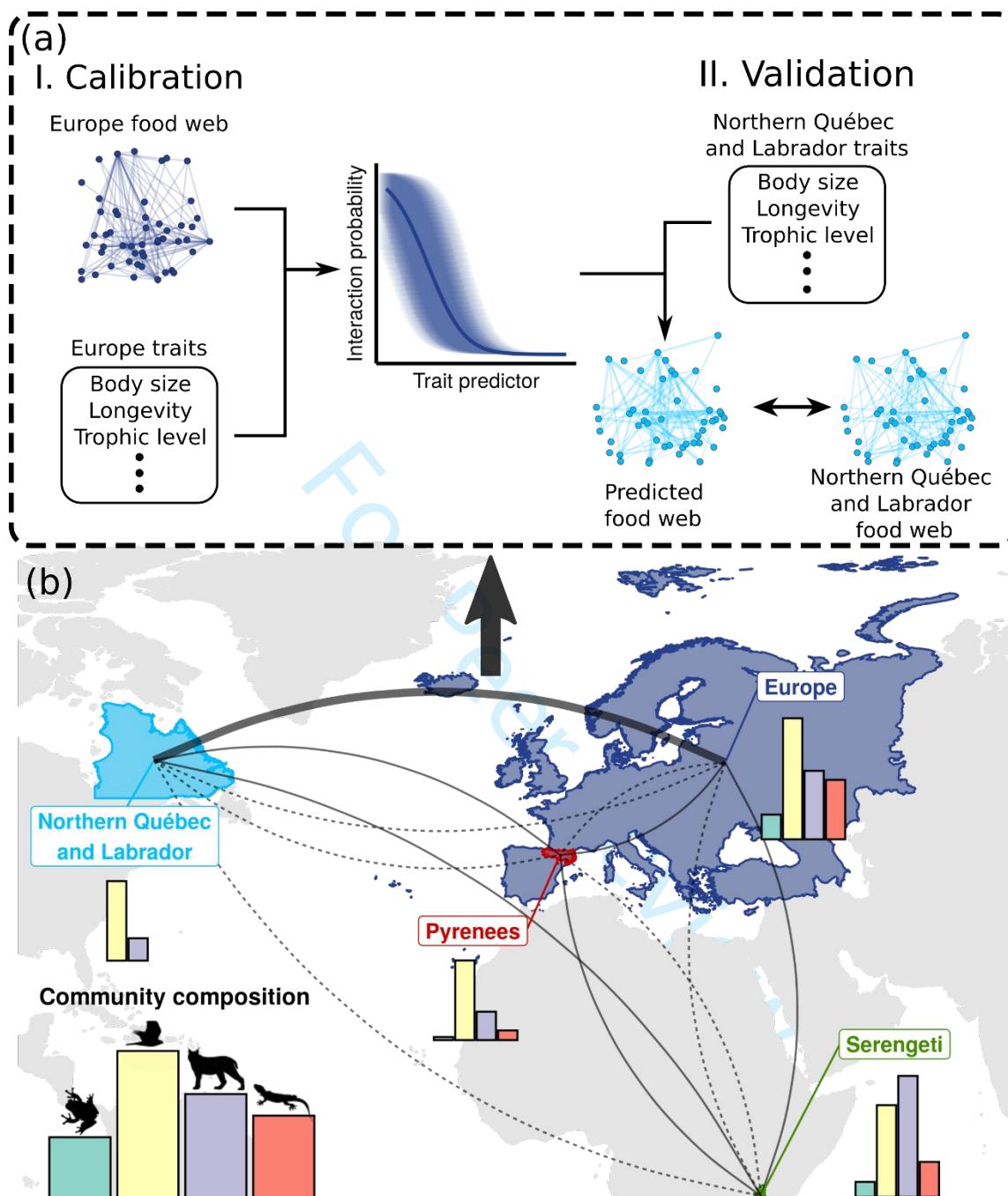


Figure 1: Trophic interactions model transferability analysis workflow. We calibrated trophic interaction model using each food web considered in this study separately (panel a.I). We validated the four models on a validation subset for within food web predictions, and the entire food webs for between food web predictions (panel a.II). Panel (a) shows an example workflow for the between food web predictions. Panel (b) maps the spatial domain of each food web. The 12 curves are the model food web pairs for between food web predictions. The within food web predictions are not shown in the figure. Bar plots are the proportion of amphibians (green), birds (yellow), mammals (purple), and reptiles (red) in each food web.

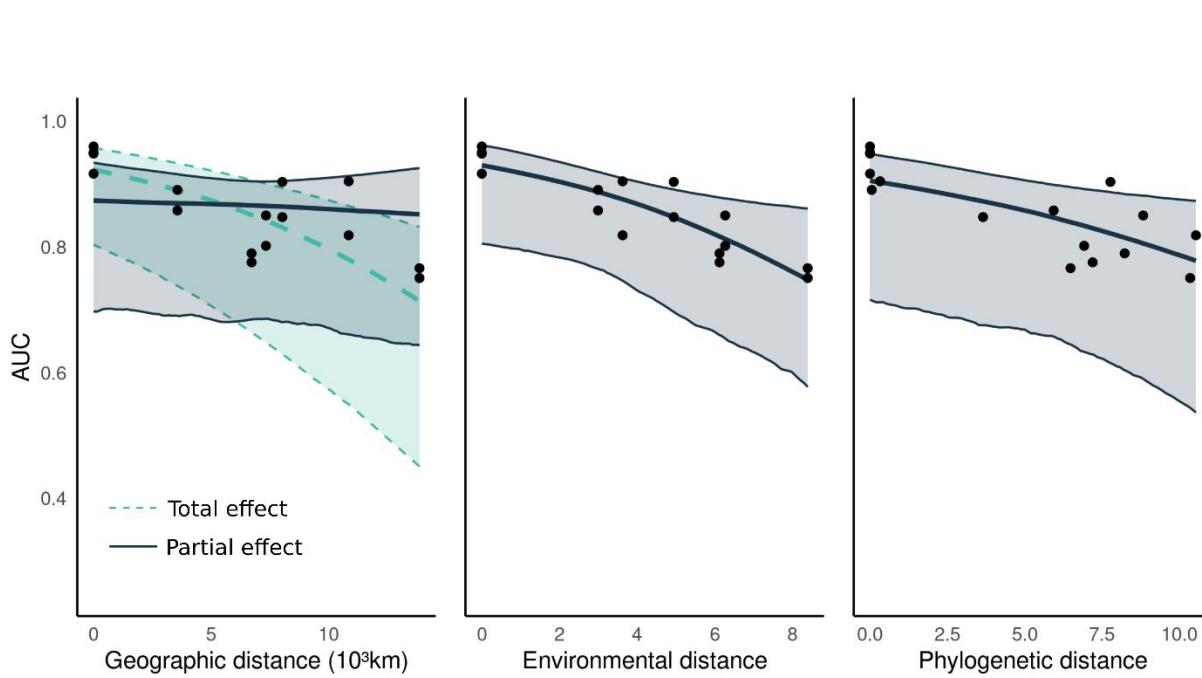


Figure 2: Transferability of predictive models. Points are the predictive performance (AUC) of the sixteen combination of model-food web prediction (Table 2). The trend lines are the median effects with their 95% credible interval constructed the posterior predictive distribution of geographic, environmental, and phylogenetic distances on predictive performance. Shown are the total (turquoise) and partial (dark blue) effects of geographic distance (controlling for environmental and phylogenetic distances), and the direct (controlling for geographic distance) effects of environmental and phylogenetic distances.

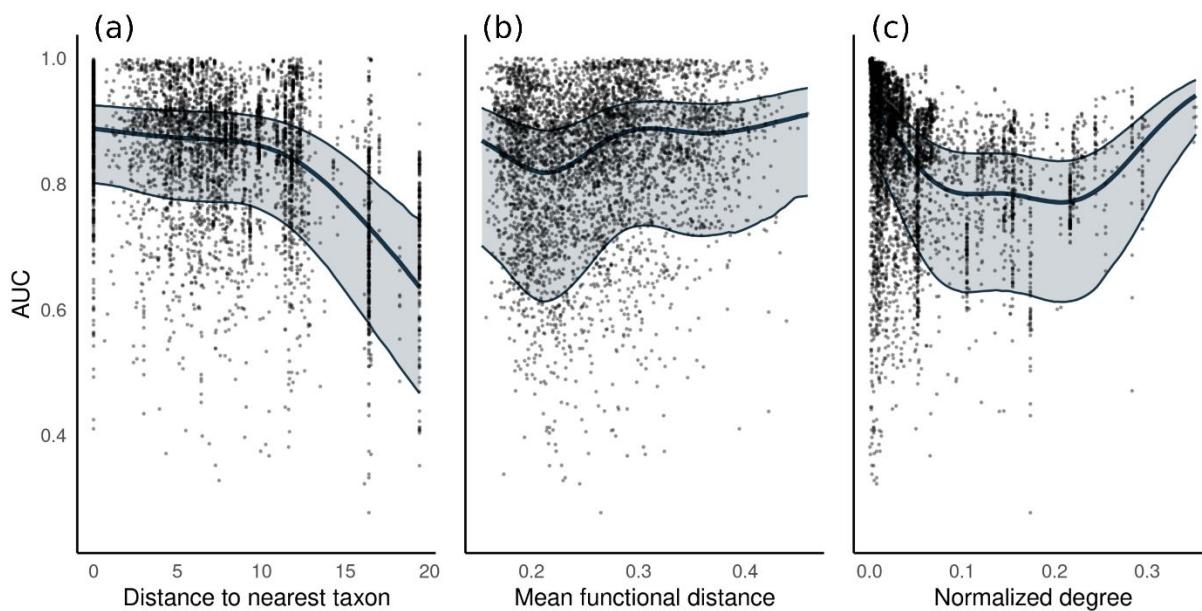


Figure 3: Predicting species interactions. Performance of the model calibrated on each food web to predict the interactions of species in the other food webs. Each point is the performance to predict the prey and predators of a single species. The trend lines are the median effects with their 95% credible interval constructed from the posterior predictive distribution of (a) distance to the nearest taxon, (b) mean functional pairwise distance, and (c) normalized degree on predictive performance.

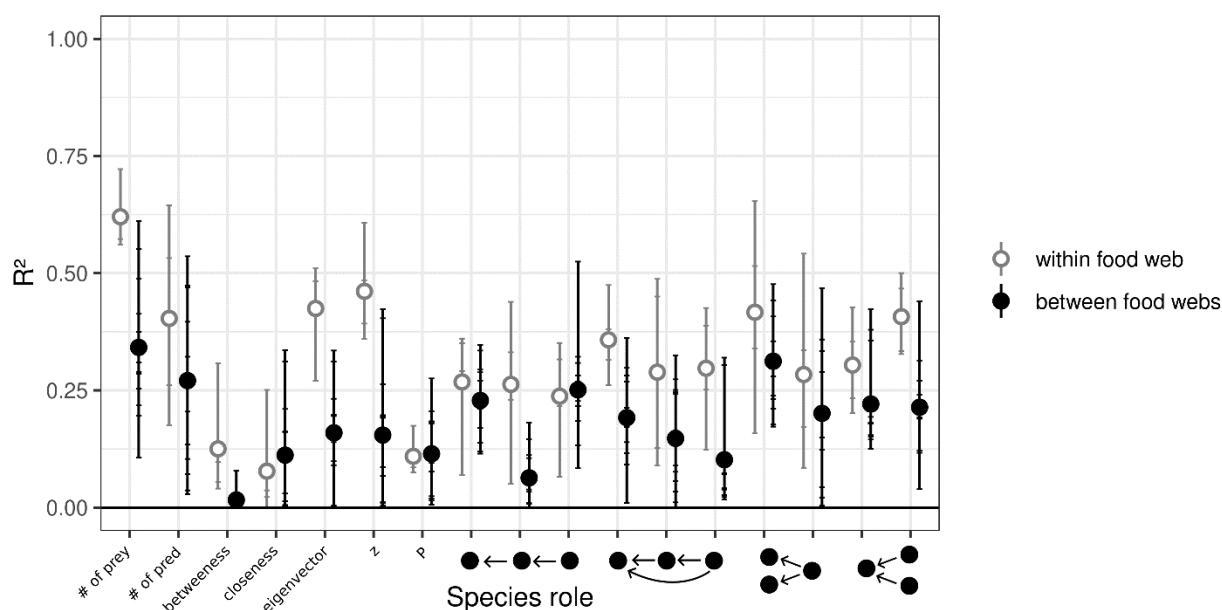


Figure 4: Predicting species functional role. Performance of models to predict species functional role measured as the proportion of the variance in trophic positions explained by the models (R^2). From left to right, the figure shows the number of prey, number of predators, betweenness, closeness, eigenvector 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 are the R^2 for each role, model, and food web predicted combination. Grey open dots are the mean R^2 with for within food web predictions. Full black dots are the mean R^2 for across food web predictions.

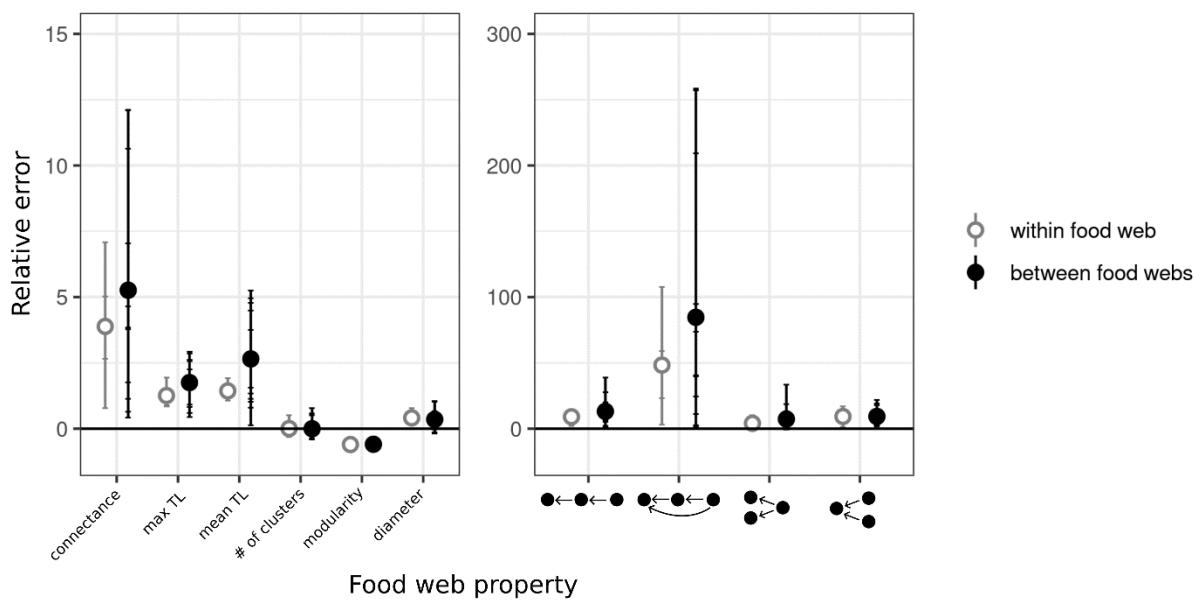


Figure 5: Prediction error of global food web properties. The relative error of the predicted food web properties. Relative error is the difference between the predicted and empirical estimates divided by the empirical estimate. From left to right, the figure shows the relative error for connectance, maximum trophic level, mean trophic level, 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 motifs (1 consumer – 2 resources). 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. The relative errors were presented in two panels to highlight the important difference in relative error between network properties and species roles

1
2
3 **Trophic interaction models predict interactions across**
4 **ecosystems, not food webs.**
5
6
7

8 **Table of Content:**
9

10 **Appendix S1 – Description of the food webs:** Brief description of how the four food web were
11 collated, and how we extracted predator-prey interactions
12
13

14 **Appendix S2 – Description of the structure of the predictive model:** We detail the structure of the
15 predictive model (response variable, predictors, transformations, priors).
16
17

18 **Appendix S3 - Description of species functional roles and food web-level properties:** We describe
19 each functional role and food web-level metrics, how we calculated them, and their relation to
20 functioning.
21
22

23 **Appendix S4 - Environmental distance sensitivity analysis to random sampling:** We test the effect
24 of the random sampling for the environmental distance estimate.
25
26

27 **Appendix S5 - Convergence of the predictive models:** Distribution of rank-normalized potential
28 scale reduction factor on split chains (Rhat) and rank plots of the population-level effects.
29
30

31 **Appendix S6 - Predicting species interactions extended results:** Results of the predicting species
32 interactions with other evaluation metrics, and results of the linear models between model performance
33 and distances estimates.
34
35

36 **Appendix S7 - Predicting functional roles extended results:** Additional results for the prediction of
37 species functional role (linear model RZ and coefficients).
38
39

40 **Appendix S8 - Predicting food web properties extended results:** Additional results for the prediction
41 of food web results.
42
43

44 **Appendix S9 – References:** References for the supplemental material
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1 2 Appendix 1: Description of the food webs 3

4 *Europe food web*: The food web comprising European tetrapods was assembled by Maiorano et al.
5 (2020). It contains the potential trophic interactions between all terrestrial vertebrates in Europe based
6 on data from atlases, field guides, published papers, and expert knowledge. Maiorano et al. (2020)
7 originally reported interactions between predators and their prey at different life-stage (eggs,
8 larvae/young, or adult). In this study, we focused on interactions between predators and adult prey only.
9

10 *Pyrenees food web*: The Pyrenees vertebrate food web was assembled independently from the
11 European food web by Lurgi et al. (2012) who compiled all known interactions between vertebrate
12 species in the alpine ecosystem of the southeastern Pyrenees mountain range. Data was collated from
13 atlases, existing databases, published scientific articles and reports, and expert knowledge. This
14 network is described at the species level and for this study, we excluded fish species.
15

16 *Northern Québec and Labrador food web*: The trophic tundra food web of Labrador and northern
17 Quebec (Canada) was compiled by Berteaux et al. (2018). The Northern Québec and Labrador food
18 web documents the presence (absence) of trophic relationships between all birds, mammals, and groups
19 of plants (e.g., moss, ferns, grass) and arthropods (e.g., Lepidoptera, Diptera, Arachnida) whose
20 distribution overlap Labrador and Quebec north of 50°N (Nunavik region). Trophic interactions were
21 collated from a literature review, data from existing databases, and expert knowledge. For our study, we
22 focused on the bird and mammal species (i.e., the vertebrates) of this network.
23

24 *Serengeti food web*: The Serengeti food web compiles all interactions between resource categories (e.g.,
25 seeds, shrubs, fruits and nectar), invertebrate groups (e.g., thrips, bees, moths, crickets), and vertebrate
26 species in the Serengeti National Park (Tanzania) and was assembled from a literature review (de
27 Visser et al., 2011). The nodes of the original food webs are resource categories, or trophic groups
28 including one or more invertebrate groups and vertebrate species. In this study, we focused on
29 vertebrate species, and assumed that species within a trophic group share the same predator and prey
30 species.
31

1 2 Appendix 2: Description of the structure of the predictive model 3

4
5 The predictive models we use is a Bayesian hierarchical generalized linear model with 13 trait-based
6 predictor variables with varying intercepts and slopes for the order of the predator. All traits were
7 extracted from Etard et al. (2020). Gaps in the database were imputed using MissForest (Stekhoven,
8 2013; Stekhoven & Buehlmann, 2012). All traits are described in the original paper, but here, we
9 describe briefly each derived variable used in the predictive models:
10

- 11 · Population-level effects:
 - 12 o Trophic level of the predator: Categorical with 3 levels - Herbivore, Omnivore, and
13 Carnivore. Coded as 2 binary variables (Omnivore and Carnivore).
 - 14 o Habitat Breadth of the predator: Quantitative - Number of habitats the predator is known to
15 use based on the second level of the IUCN habitat data.
 - 16 o Body mass of the predator: Quantitative - Log-transformed body mass of the predator
 - 17 o Longevity of the predator: Quantitative - Life span of an individual of the predator species.
 - 18 o Clutch Size of the predator: Quantitative - Number of offspring or eggs per clutch/litter by
19 the predator.
 - 20 o Trophic level of the prey: Categorical with 3 levels - Herbivore, Omnivore, and Carnivore.
21 Coded as 2 binary variables (Omnivore and Carnivore)
 - 22 o Habitat Breadth of the prey: Quantitative - Number of habitats the prey is known to use
23 based on the second level of the IUCN habitat data.
 - 24 o Body mass of the prey: Quantitative - Log-transformed body mass of the prey
 - 25 o Longevity of the prey: Quantitative - Life span of an individual of the prey species.
 - 26 o Clutch Size of the prey: Quantitative - Number of offspring or eggs per clutch/litter by the
27 prey.
 - 28 o Activity Time match: Categorical binary – The predator and the prey are active during the
29 same time of the day (1) or not (0). The original activity time data for the prey and the
30 predator is a binary variable with possible values nocturnal and non-nocturnal (diurnal,
31 crepuscular, or catemeral)
 - 32 o Habitat match: Quantitative – Jaccard similarity between the habitat use of the predator and
33 the prey. The original habitat data are 12 binary variables: Forest, Savanna, Shrubland,
34 Grassland, Wetland, Rocky areas, Cave and subterranean, Desert, Marine, Marine/Intertidal
35 or coastal/supratidal, Artificial, Introduced vegetation
 - 36 o Body mass match: The squared difference between the log-transformed body mass of the
37 predator and the log-transformed body mass of the prey.

41
42 For each food web, we centered and scaled the quantitative predictors by subtracting each value by the
43 mean and dividing by twice the standard deviation. The predictors that we scaled are: the log-
44 transformed habitat breadth of the predator and the prey, the log-transformed body mass of the predator
45 and the prey, the log-transformed longevity of the predator and the prey, the log-transformed
46 clutch/litter size of the predator and the prey, the habitat match, and the body mass match.
47

- 48 · Group-level effects (intercept and slopes):
 - 49 o Order of the predator
 - 50 □ Northern Québec and Labrador food web: 21 levels
 - 51 □ Europe: 36 levels
 - 52 □ Pyrenees: 23 levels
 - 53 □ Serengeti: 41 levels

- 1 · Model:

2 $L_{ij} \sim Bernouilli(p_{ij}),$

3 $logit(p_{ij}) = \alpha + \alpha_{predator[j]} + \sum_{k=1}^{13} (\beta_k \times T_k + \beta_{k,predator[j]} \times T_k),$

- 4
5
6
7
8
9
10
11
 - 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
 - $\alpha_{predator[j]}$ and $\beta_{predator[j]}$ are the order-specific intercept and slopes for the order of predator j
 - T_k are the trait-based predictors.

- 12 · Priors:

13 $\alpha, \beta \sim Normal(0, 1)$

14 $\alpha_{predator[j]}, \beta_{k,predator[j]} \sim Normal(\mu, \sigma)$

15 $\mu \sim Normal(0, 1)$

16 $\sigma \sim HalfCauchy(0, 5)$

- 17 · Number of iterations:

- 18
 - Warmup: 2000 iterations
 - Inference: 2000 iterations

- 19 · Calibration data:

- 20
 - 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 calibration of some predictive models.

- 21 · Validation:

- 22
 - For within food web predictions (i.e., trophic interaction predictions in the food web used for calibration): 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. Some measures like the area under the precision-recall curve (auprc) are influenced by prevalence. Controlling for prevalence allows comparison of within food web predictions and across food webs predictions for a given food web.
 - For across food web predictions (i.e., trophic interaction predictions in food webs other than the one used for calibration): We used the entire food web (all interactions and non-interactions).

```
1      . Model formula:  
2  
3 brms_form <- bf(interaction ~ 1 + Omnivore.predator +  
4 Carnivore.predator + Habitat_breadth.predator + BM.predator +  
5 Longevity.predator + Clutchsize.predator + Omnivore.prey +  
6 Carnivore.prey + Habitat_breadth.prey + BM.prey + Longevity.prey +  
7 ClutchSize.prey + ActivityTime.match + Habitat.match + BM.match) +  
8 (1 + (Omnivore.predator + Carnivore.predator +  
9 Habitat_breadth.predator + BM.predator + Longevity.predator +  
10 ClutchSize.predator + Omnivore.prey + Carnivore.prey +  
11 Habitat_breadth.prey + BM.prey + Longevity.prey + ClutchSize.prey +  
12 ActivityTime.match + Habitat.match + BM.match) || Order.predator),  
13 family = bernoulli())  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60
```

1 2 Appendix 3: Description of species functional roles and food web-level properties 3

4 Species functional roles: 5

6 For more information on what each species role metrics represent and aim to describe, we recommend
7 the reviews made by Cirtwill et al. (2018) and Delmas et al. (2019). Here, we very briefly describe
8 what these metrics are, how they are related to the functioning and the stability of food webs:
9

- 10 · **Number of prey:** We calculated the number of prey of a species as the number of interactions
11 for which the species is the predator (in-degree). It measures how generalized the species is.
12 The loss of species with many prey are more likely to trigger top-down cascades (Curtsdotter et
13 al., 2011). In contrast, specialist predators are more vulnerable to co-extinction. In-degree and
14 out-degree quantify the direct influence of species on other species.
15
- 16 · **Number of predators:** We calculated the number of predators of a species as the number of
17 interactions for which the species is the prey (out-degree). It measures the vulnerability of the
18 prey. The loss of species with many predators are more likely to trigger bottom-up cascades
19 (Curtsdotter et al., 2011). In-degree and out-degree quantify the direct influence of species on
20 other species.
21
- 22 · **Betweenness:** The betweenness centrality of a species is the number of times the shortest path
23 linking any two species in the network goes through the focal species (Freeman, 1977).
24 Betweenness quantify the number of food chains a species takes part in, and thus how it affects
25 energy flows (Cirtwill et al., 2018). It is considered one important measures of species
26 topological importance (Jordán, 2009). We calculated betweenness using the function
27 *centr_betw* in the R package *igraph* (Csardi & Nepusz, 2006).
28
- 29 · **Closeness:** Closeness centrality measures how far a species is to any other species. It is
30 therefore a more “global” measure of centrality than betweenness and quantify how
31 “efficiently” the change in this species will spread through the food web. We used the function
32 *centr_clo* in the R package *igraph* (Csardi & Nepusz 2006).
33
- 34 · **Eigenvector centrality:** Eigenvector centrality uses the eigenvector of the largest eigenvalue of
35 the adjacency matrix describing the food web. It is another “global” measure of centrality that
36 aims to describe species importance in the food web. Eigenvector centrality is related to the
37 contribution of species to network stability (Allesina & Pascual, 2009). We used the function
38 *centr_eigen* of the R package *igraph* (Csardi & Nepusz 2006).
39
- 40 · **Trophic level:** Trophic level describe the “vertical” position (primary consumer, secondary
41 consumer, etc.) of a species in a food web. We calculated trophic level using the function
42 *TrophInd* of the package *NetIndices* package (Kones et al., 2009). The function calculate trophic
43 level as $1 + \text{the weighted average of the trophic levels of its food items}$. Top predators and basal
44 species are expected to have larger effect on the community through top-down and bottom-up
45 cascades (Dyer & Letourneau, 2003).
46
- 47 · **Omnivory:** Omnivory is defined as the variety of trophic levels a species feed on. Omnivory
48 has the potential to stabilize food webs by mitigating trophic cascades, although this effect is
49 debated (Bascompte et al., 2005). We calculated omnivory using the function *TrophInd* of the
50 package *NetIndices* package (Kones et al. 2009).
51

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

22 **Food web properties:**

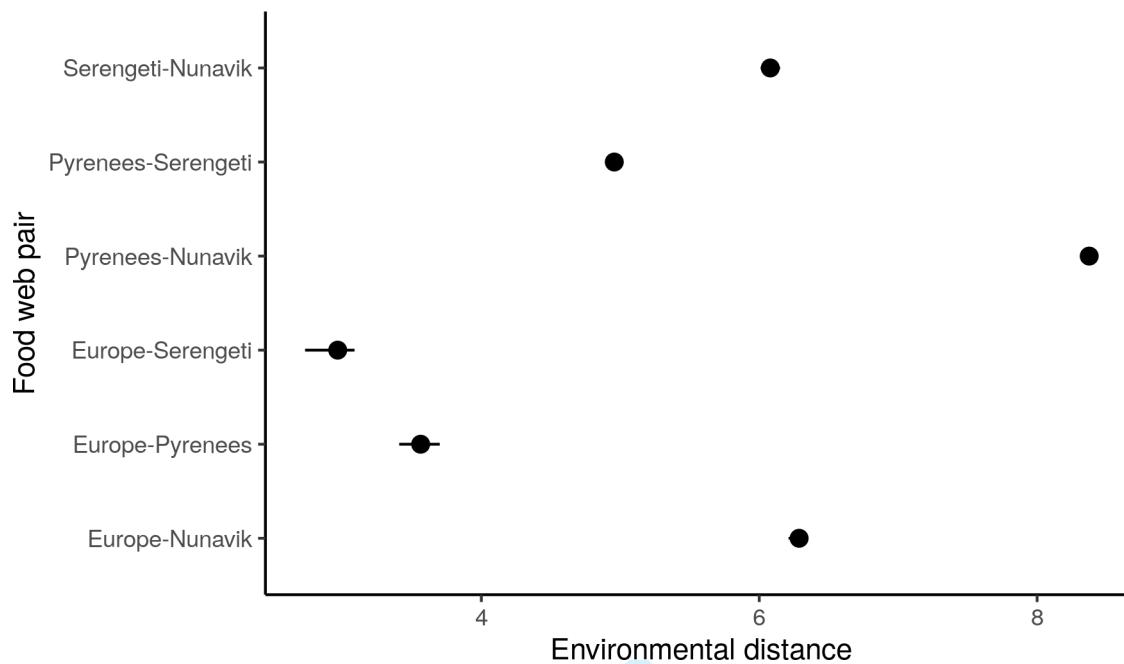
- 23 · **Connectance:** Connectance is defined as the proportion of interactions relative to the total
24 number of possible interactions. Here, since we have a directed network with self-loop allowed,
25 it is calculated as the number of interactions (L) divided by the squared the number of species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

- 1
- 2 · **Number of clusters:** The number of clusters (i.e., modules) are the number of non-overlapping
 - 3 highly connected species. We determined the number of clusters using the function
 - 4 *cluster_springlass* from the package *igraph* (Csardi & Nepusz 2006). Perturbations spread more
 - 5 efficiently within a cluster than across clusters. So, an higher number of clusters should increase
 - 6 resistance to perturbations.
- 7
- 8 · **Modularity:** Modularity quantify how food webs are divided into modules (groups of species
 - 9 interacting more together than with other species). Modularity increases food web persistence
 - 10 (Stouffer & Bascompte, 2011). We determined modules/clusters using the function
 - 11 *cluster_springlass* from the package *igraph* (Csardi & Nepusz 2006).
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- 26
- 27
- 28
- 29
- 30
- 31
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60

For Peer Review

1 2 Appendix 4: Environmental distance sensitivity analysis to random sampling 3

4
5 We sampled 10 times 500 points within the spatial domain of each food web. We extracted the 19
6 bioclimatic variables from WordClim (Hijmans, 2021). We then calculated environmental distance as
7 the Euclidean distance between the food web bioclimatic centroids (mean of each bioclimatic
8 variables) for the 10 iterations.
9



31 *Figure S4.1: Environmental distance estimates between the food webs included in our*
32 *study. We repeated the random sampling of 500 spatial points within the domain of each*
33 *food web 10 times and calculated the environmental distances between food webs for*
34 *each sampling sets. The points are the mean environmental distances and the error bar*
35 *are the range of estimates across sampling sets.*

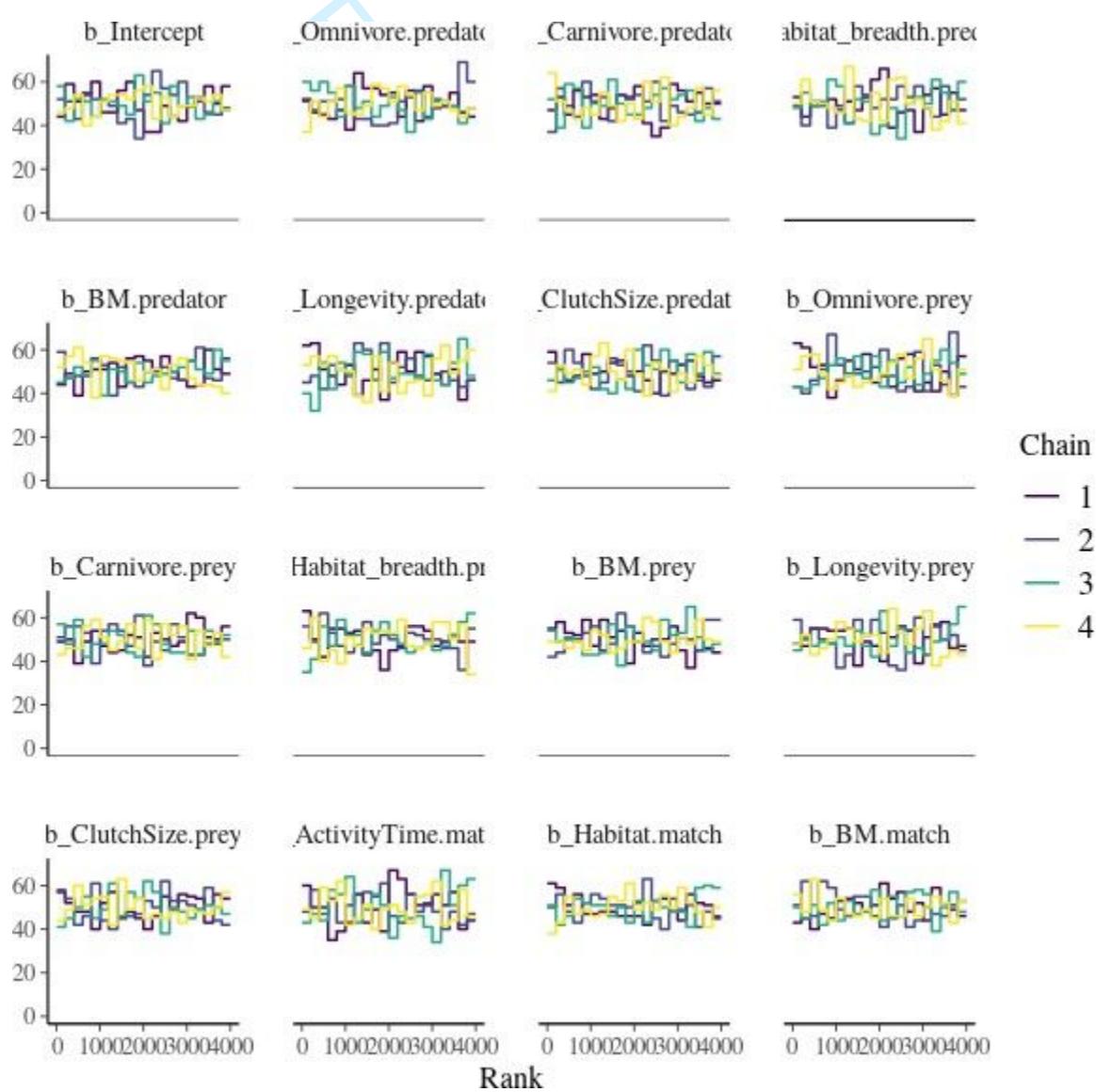
36
37 The within food web pair variation is much lower than the across pairs distances. We conclude that
38 sampling 500 random points is big enough to get an accurate estimate of the environmental distances
39 between food webs.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1 2 Appendix 5: Convergence of the predictive models 3

4
5 We evaluated convergence and adequacy of the four predictive models using rank plots, posterior
6 predictive checks, and the rank-normalized potential scale reduction factor on split chains. Rank plots
7 are similar to trace plots and are used to assess the mixing of the chains (Gelman et al., 2013). Posterior
8 predictive checks compare the distribution of the observed data to the distribution of the posterior
9 predictive distribution (Gelman et al., 2013). A discrepancy in distributions can be caused by a model
10 missfit. Finally, the rank-normalized potential scale reduction factor on split chains (Rhat) is similar to
11 the potential scale reduction factor, and is a summary statistics evaluating model convergence (Vehtari
12 et al., 2021). Rhat close to 1 indicate a correctly converged model.
13

14
15 Northern Québec and Labrador model:
16

- 17 · Rank plots of the population-level effects:



56
57 *Figure S5.1: Rank plots of the population-level effects for the model calibrated on the*
58 *Northern Québec and Labrador food web.*

1
2 · Posterior predictive checks:
3

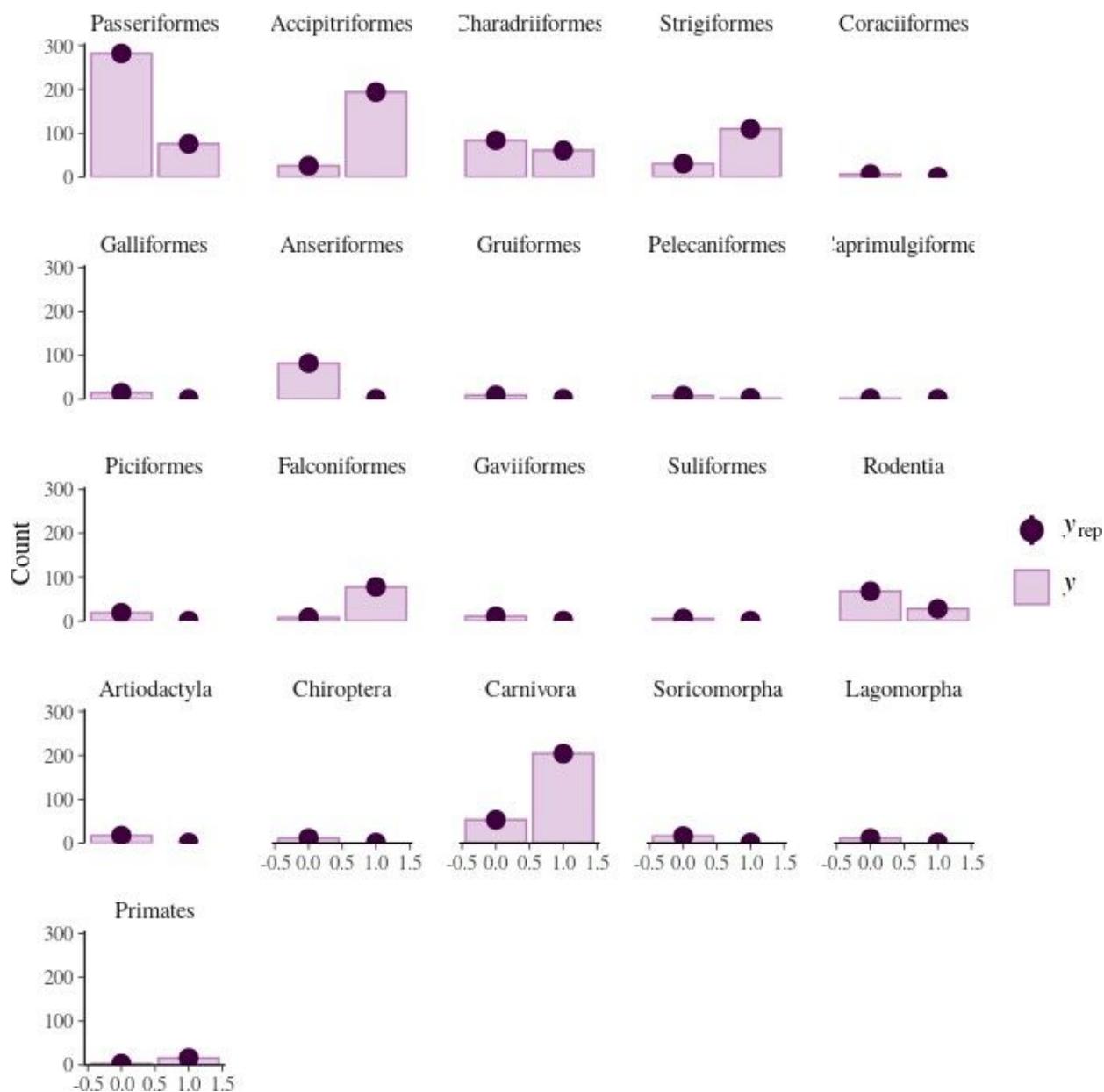


Figure S5.2: Comparison of the posterior predictive distribution (y_{rep} ; purple points) and the observed distribution (y ; purple bars) of the model calibrated on the Northern Québec and Labrador food web.

- 1
2 · Rank-normalized potential scale reduction factors on split chains:
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

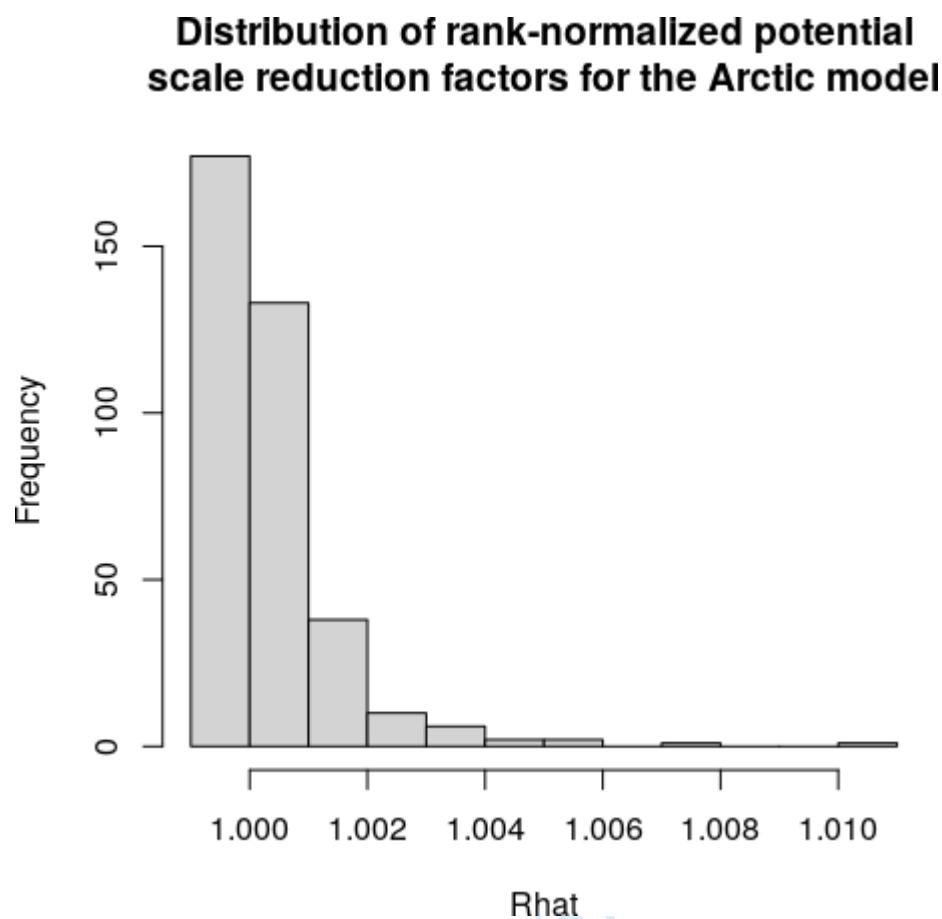
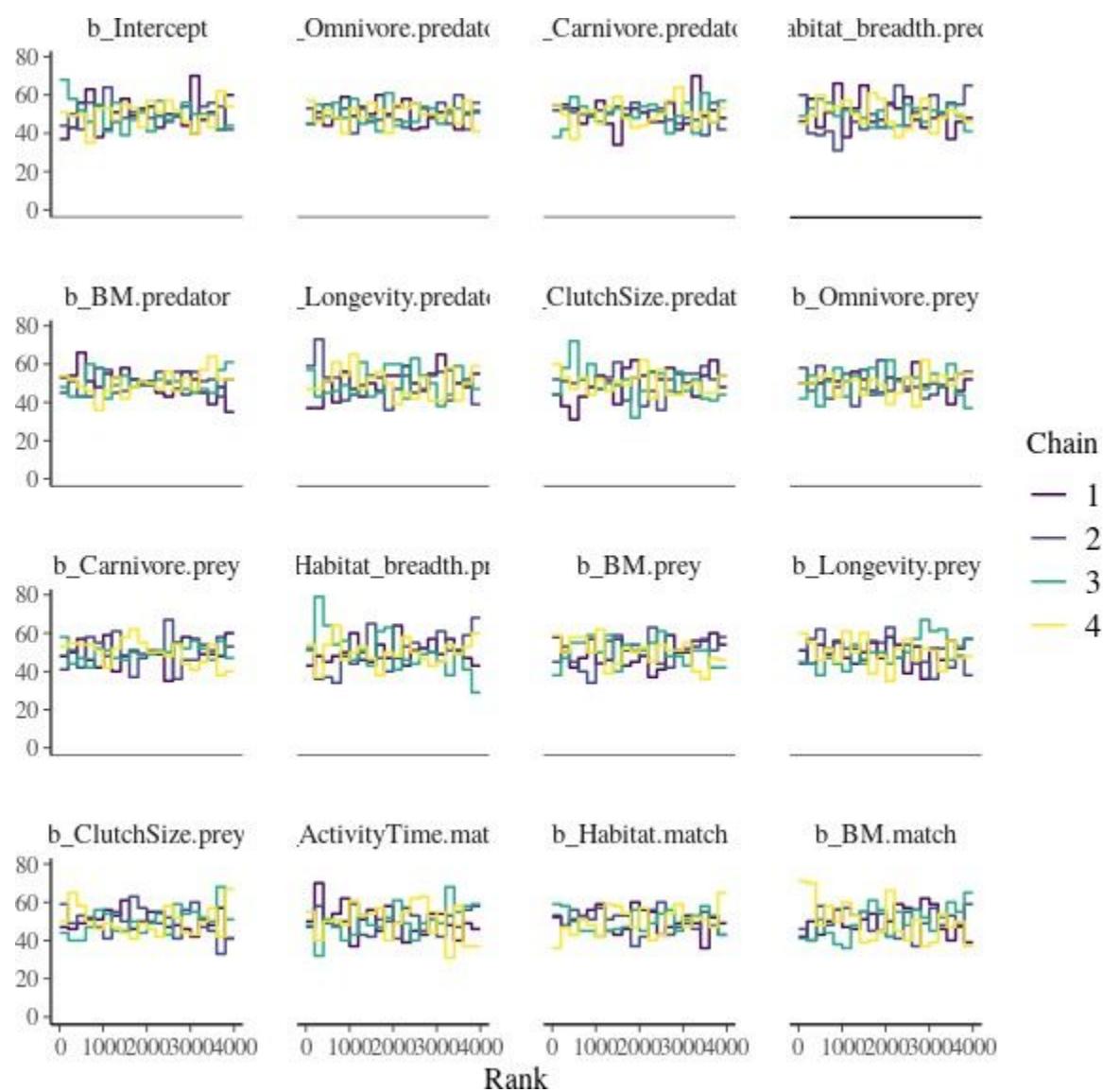


Figure S5.3: Distribution of the rank-normalized potential scale reduction factors on split chains for the Northern Québec and Labrador food web

1
2 Europe model:
3
4

- Rank plots of the population-level effects:



42 *Figure S5.4: Rank plots of the population-level effects for the model calibrated on the*
43 *European food web of terrestrial vertebrates.*

1
2 · Posterior predictive checks:
3

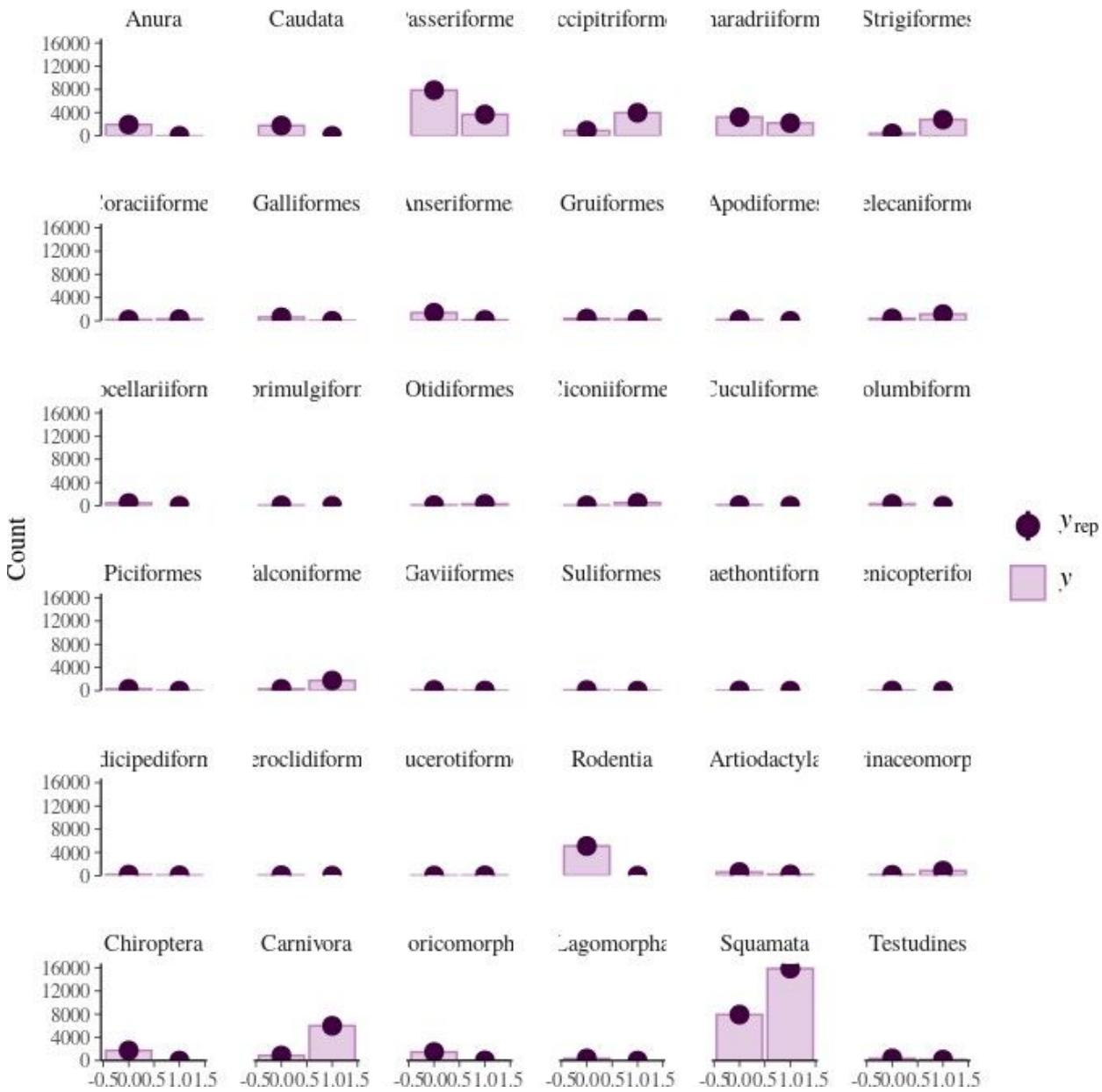


Figure S5.5: Comparison of the posterior predictive distribution (y_{rep} ; purple points) and the observed distribution (y ; purple bars) of the model calibrated on the food web of European terrestrial vertebrates.

- 1
2 · Potential scale reduction factors:
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

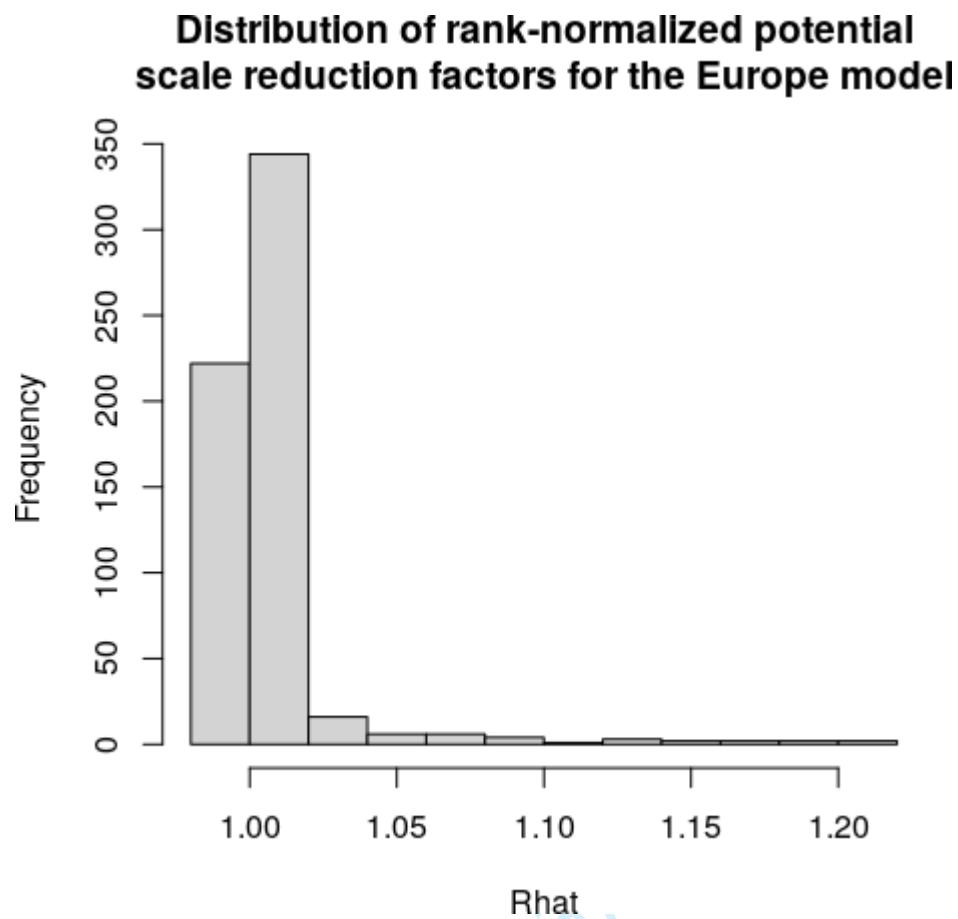
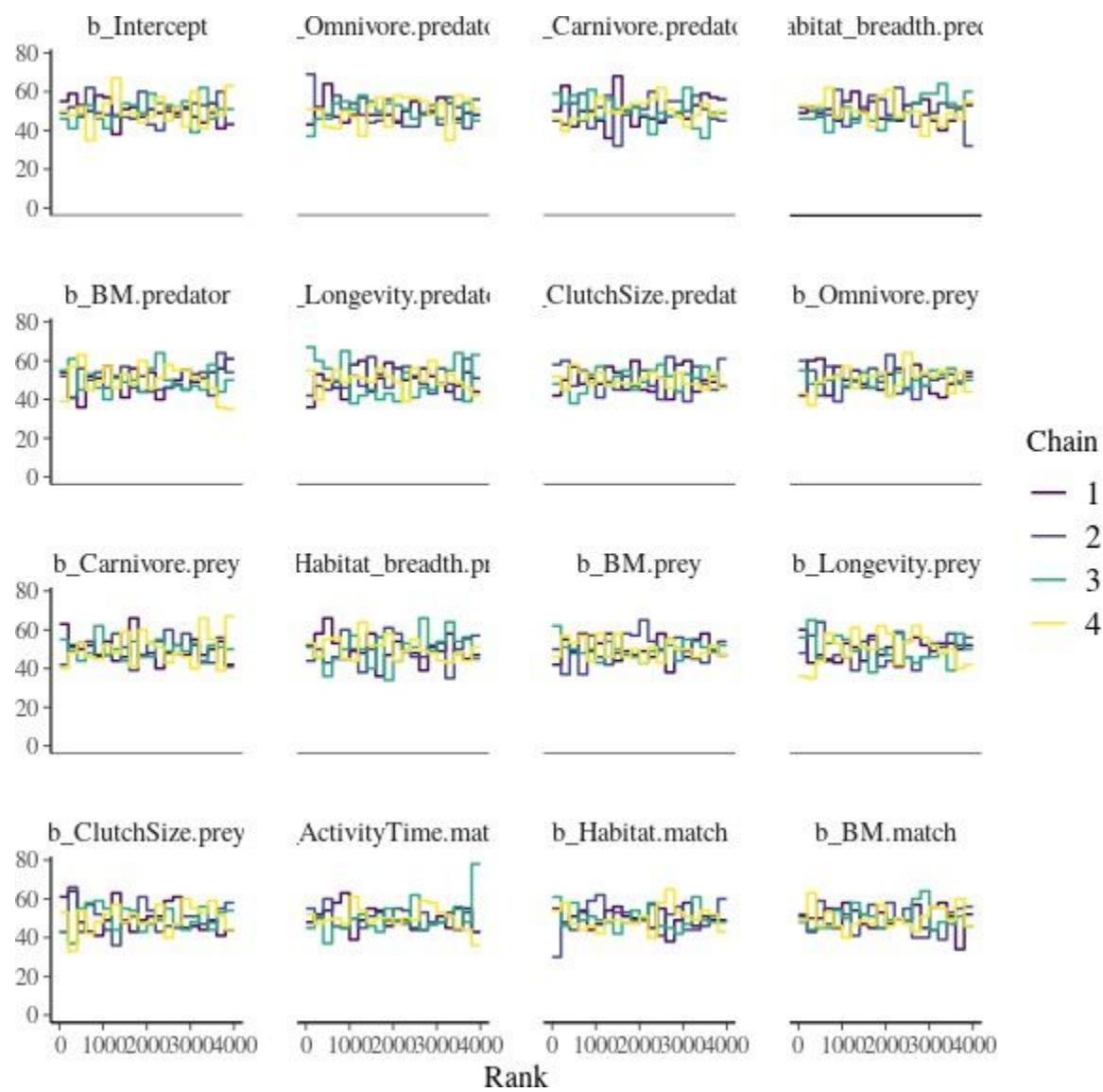


Figure S5.6: Distribution of the rank-normalized potential scale reduction factors on split chains for the food web of European terrestrial vertebrates.

1
2 Pyrenees model:
3
4

- Rank plots of the population-level effects:



43 Figure S5.7: Rank plots of the population-level effects for the model calibrated on the
44 Pyrenees food web.
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2 · Posterior predictive checks:
3

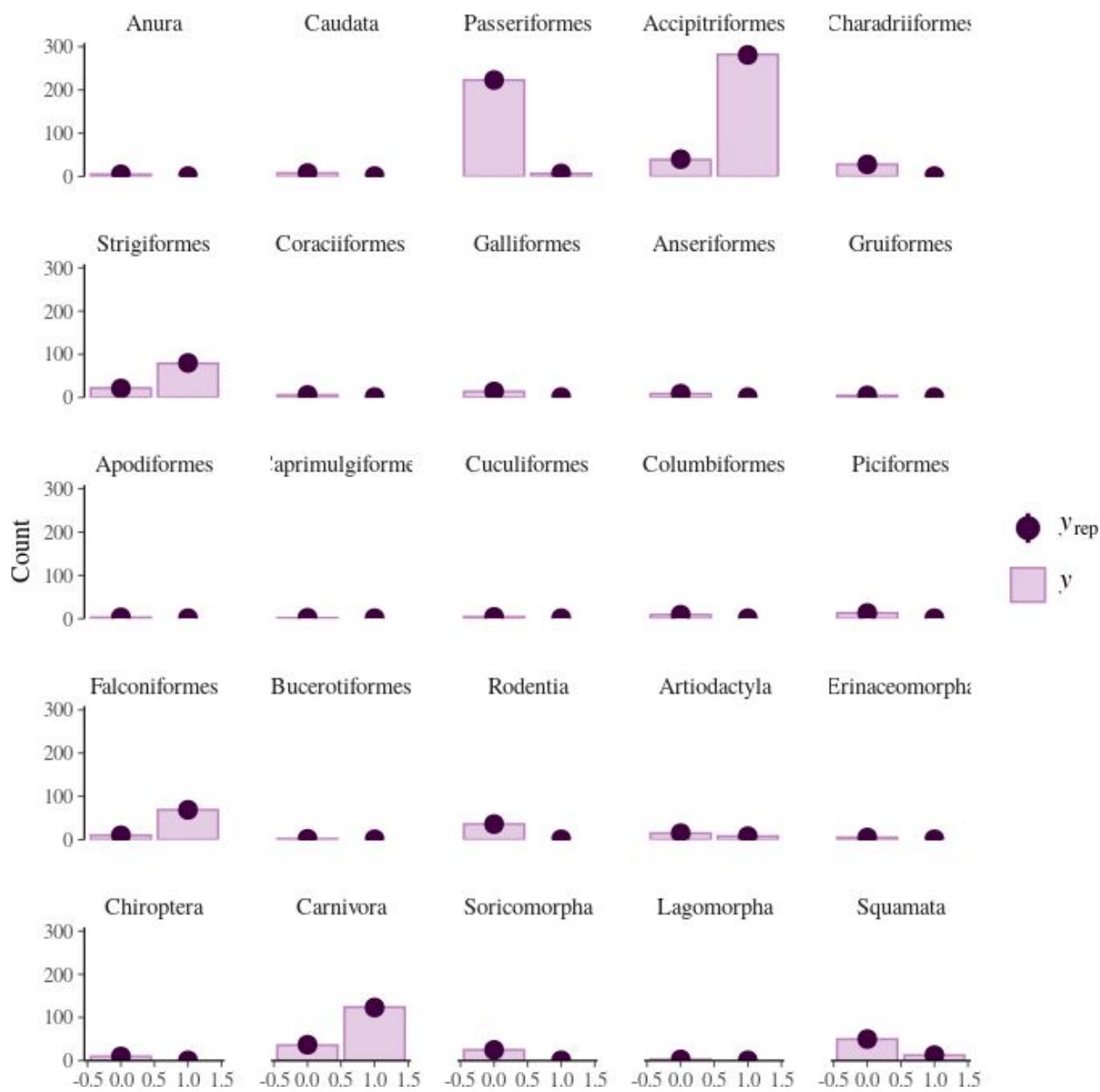


Figure S5.8: Comparison of the posterior predictive distribution (y_{rep} ; purple points) and the observed distribution (y ; purple bars) of the model calibrated on the Pyrenees food web.

- 1
2 · Potential scale reduction factors:
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

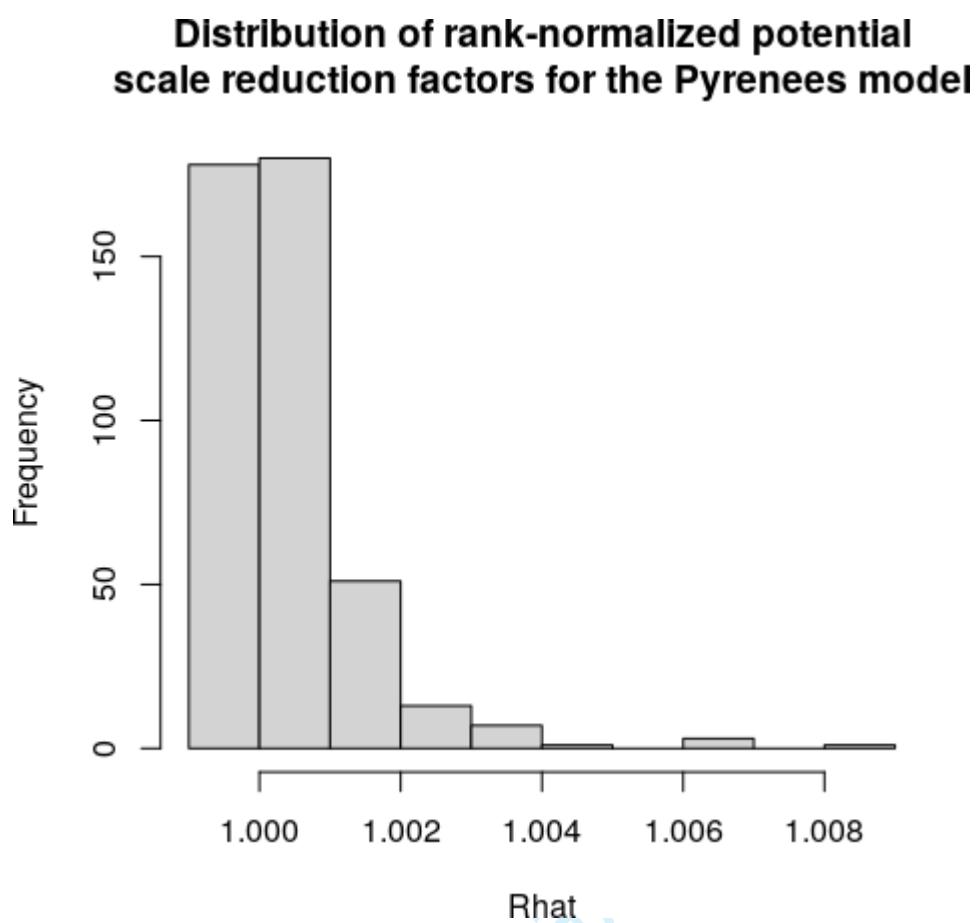
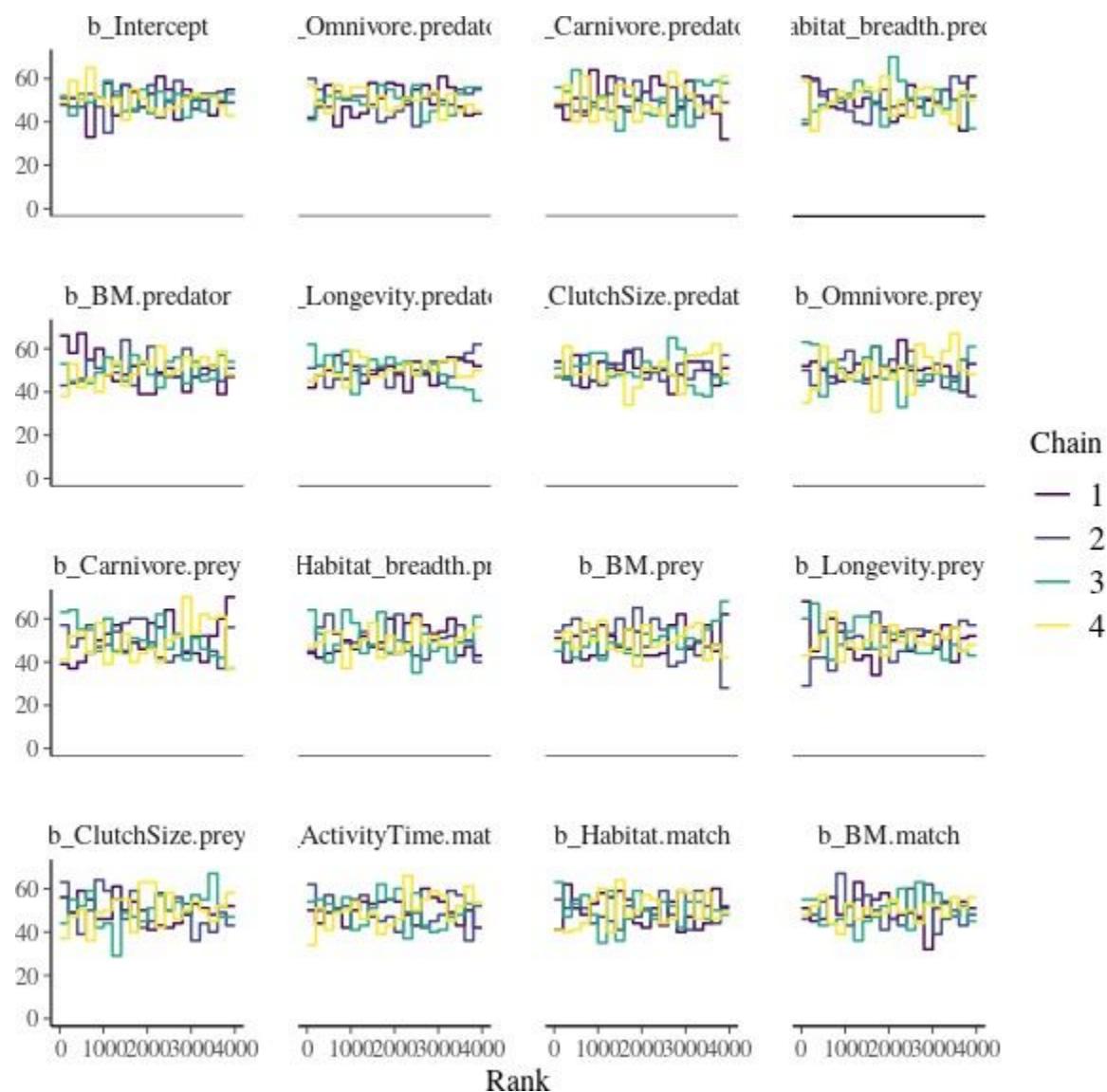


Figure S5.9: Distribution of the rank-normalized potential scale reduction factors on split chains for the Pyrenees food web

1
2 *Serengeti model:*

- 3 · Rank plots of the population-level effects:
-
- 4

41
42 *Figure S5.10: Rank plots of the population-level effects for the model calibrated on the*
43 *Serengeti food web.*

1
2 · Posterior predictive checks:
3

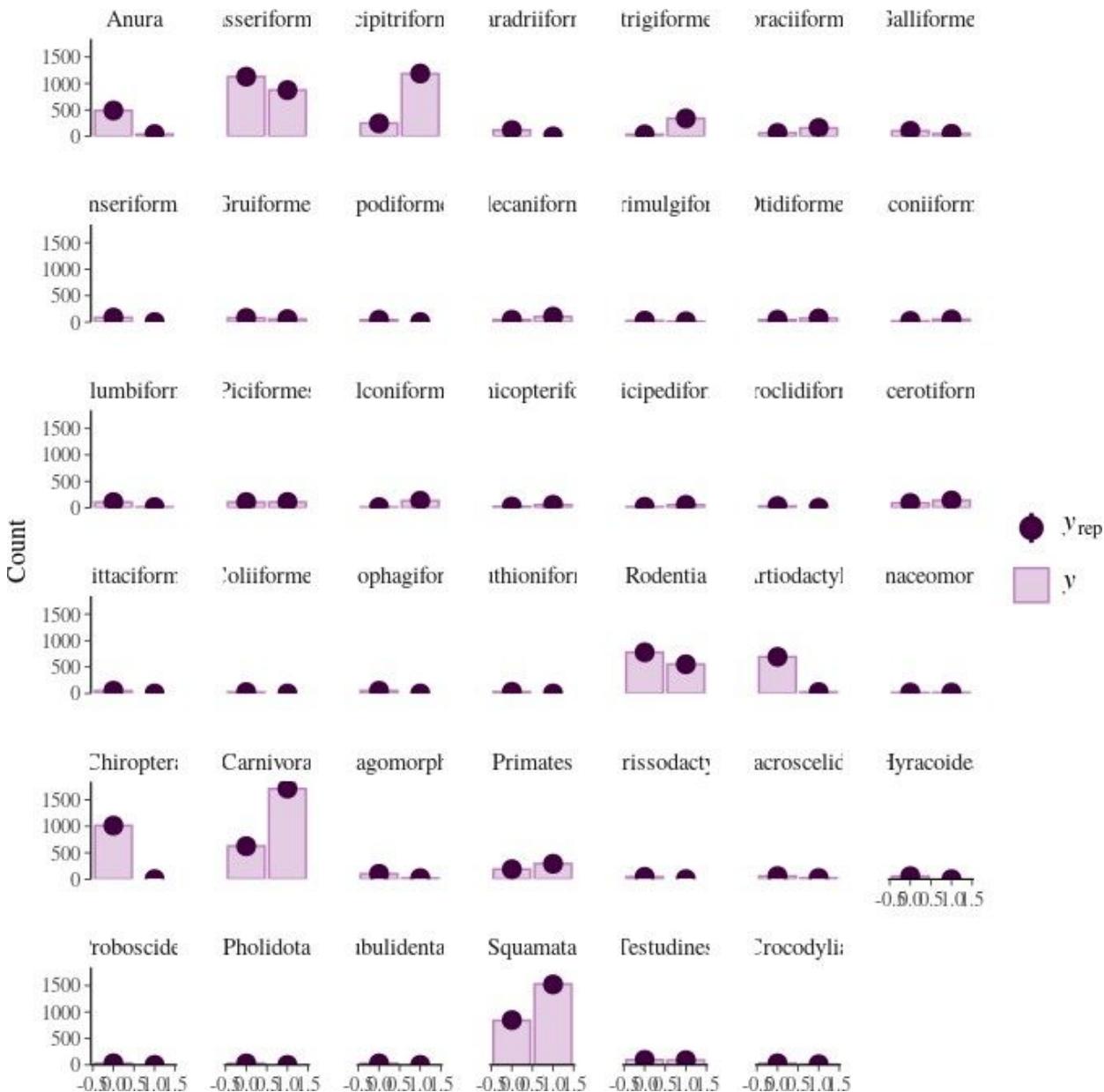
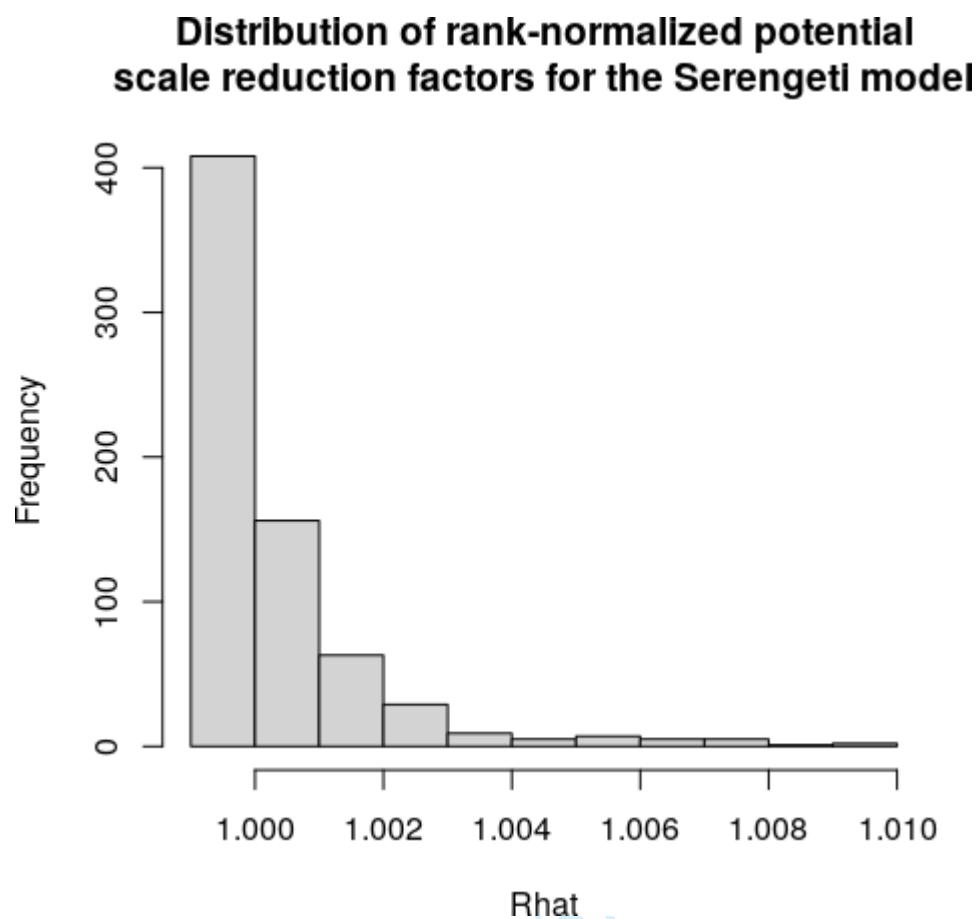


Figure S5.11: Comparison of the posterior predictive distribution (y_{rep} ; purple points) and the observed distribution (y ; purple bars) of the model calibrated on the Serengeti food web.

- 1
2 · Potential scale reduction factors:
3
4
5
6

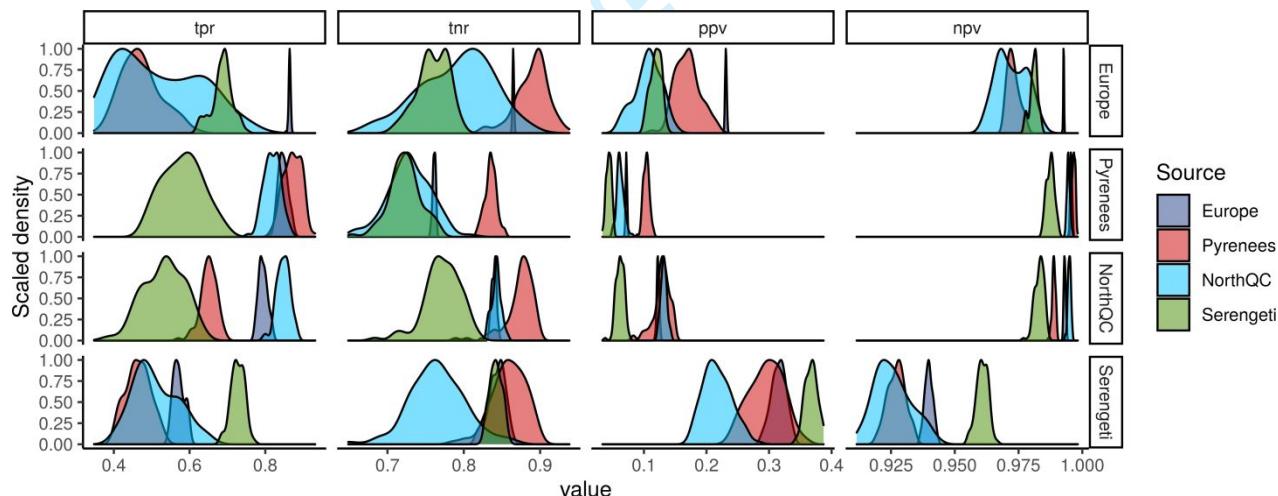


34
35 *Figure S5.12: Distribution of the rank-normalized potential scale reduction*
36 *factors on split chains for the Serengeti food web*

1 2 Appendix 6: Predicting species interactions extended results 3 4

5
6 *Table S6.1: Area under the precision-recall curve (auprc) of each model*
 7 *predicting every food webs. The baseline for auprc (value for random*
 8 *predictions) depends on the prevalence of interactions in the validation*
 9 *dataset (Europe: 0.05; Pyrenees: 0.02; North Québec and Labrador:*
 10 *0.03; Serengeti: 0.12). Consequently, model performance can be*
 11 *compared for a given predicted food web (within a column) but not*
 12 *across food webs (across columns).*

Model	Food web			
	Europe	Pyrenees	North Québec and Labrador	Serengeti
Europe	0.47	0.09	0.19	0.37
Pyrenees	0.23	0.22	0.2	0.37
North Québec and Labrador	0.18	0.15	0.27	0.32
Serengeti	0.18	0.06	0.08	0.56



45 *Figure S6.1: Scaled density distribution of true positive rate (tpr), true negative rate (tnr), positive*
 46 *predictive value (ppv), and negative predictive value (npv) of each model (Source) predicting every*
 47 *food webs (rows). tpr is the proportion of correctly predicted interactions, tnr is the proportion of*
 48 *correctly predicted non-interactions, ppv is the proportion of predicted interactions that are observed*
 49 *interactions, and npv is the proportion of predicted non-interactions that are observed non-*
 50 *interactions. We calculated tpr, tnr, ppv, and npv with 100 samples of the posterior distribution of the*
 51 *entire food webs.*

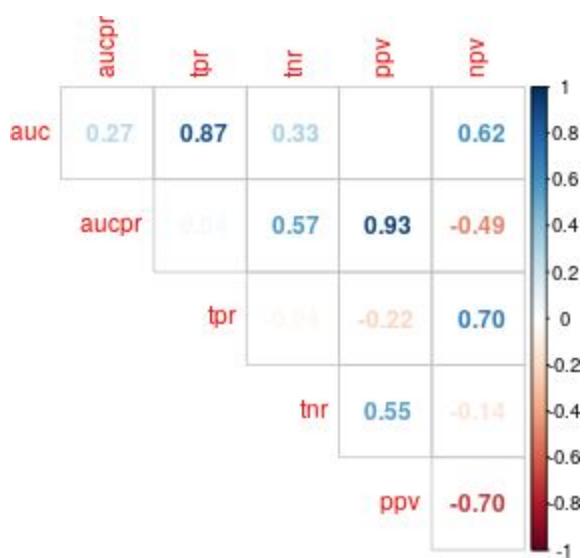


Figure S6.2: Correlation between performance metrics for predicting the interactions of the entire food webs. auc is the area under the receiver operating curve, aucpr is the area under the precision-recall curve, tpr is the true positive rate, tnr is the true negative rate, ppv is the positive predictive value, and npv is the negative predictive value

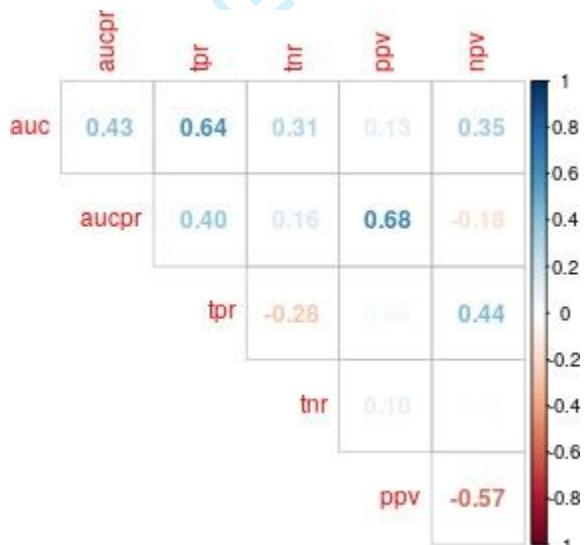


Figure S6.3: Correlation between performance metrics for predicting the interactions of each species within a food web. See caption of figure S6.2 for abbreviations.

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 = 1.73, 95% CrI = [0.71, 2.30]
 geo.dist_sc : Estimate = -0.54, 95% CrI = [-0.79, -0.30]
random effect:
 $\text{sd}(1|\text{Model})$: Estimate = 0.38, 95% CrI = [0.01, 1.67]
 $\text{sd}(1|\text{Food web})$: Estimate = 0.41, 95% CrI = [0.01, 1.160]
- 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 = 1.84, 95% CrI = [1.17, 2.26]
 geo.dist_sc : Estimate = -0.07, 95% CrI = [-0.48, 0.34]
 phylo.dist_sc : Estimate = -0.20, 95% CrI = [-0.63, 0.20]
 env.dist_sc : Estimate = -0.36, 95% CrI = [-0.84, 0.17]
random effect:
 $\text{sd}(1|\text{Model})$: Estimate = 0.25, 95% CrI = [0.01, 0.99]
 $\text{sd}(1|\text{Food web})$: Estimate = 0.25, 95% CrI = [0.01, 1.10]
- 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 = 1.75, 95% CrI = [0.88, 2.29]
 geo.dist_sc : Estimate = -0.26, 95% CrI = [-0.59, 0.08]
 phylo.dist_sc : Estimate = -0.39, 95% CrI = [-0.78, -0.03]
random effect:
 $\text{sd}(1|\text{Model})$: Estimate = 0.40, 95% CrI = [0.02, 1.44]
 $\text{sd}(1|\text{Food web})$: Estimate = 0.31, 95% CrI = [0.01, 1.24]
- 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 = 1.69, 95% CrI = [0.40, 2.23]
 geo.dist_sc : Estimate = -0.08, 95% CrI = [-0.49, 0.29]
 env.dist_sc : Estimate = -0.52, 95% CrI = [-0.89, -0.11]
random effect:
 $\text{sd}(1|\text{Model})$: Estimate = 0.23, 95% CrI = [0.01, 1.13]
 $\text{sd}(1|\text{Food web})$: Estimate = 0.35, 95% CrI = [0.01, 1.51]

Appendix 7: Predicting species role extended results

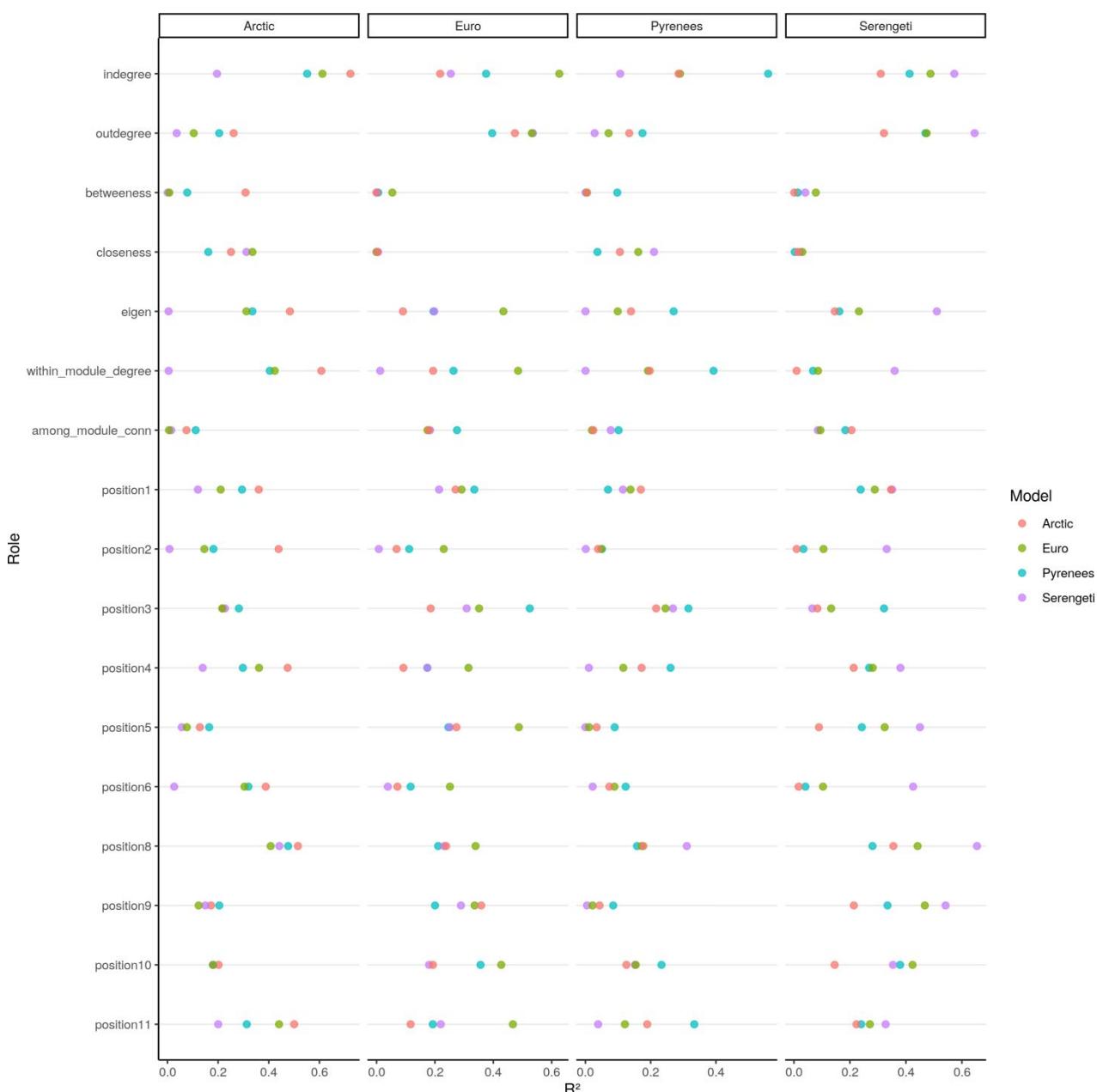


Figure S7.1: Predictive performance of each model (colors) to predict species role for every food webs (column). Performance is measured with the coefficient of determination (R^2). Arctic refers to the Northern Québec and Labrador food web. From top to bottom: Number of prey, number of predator, betweenness, 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.

1
2
3 Number of prey
4
5
6

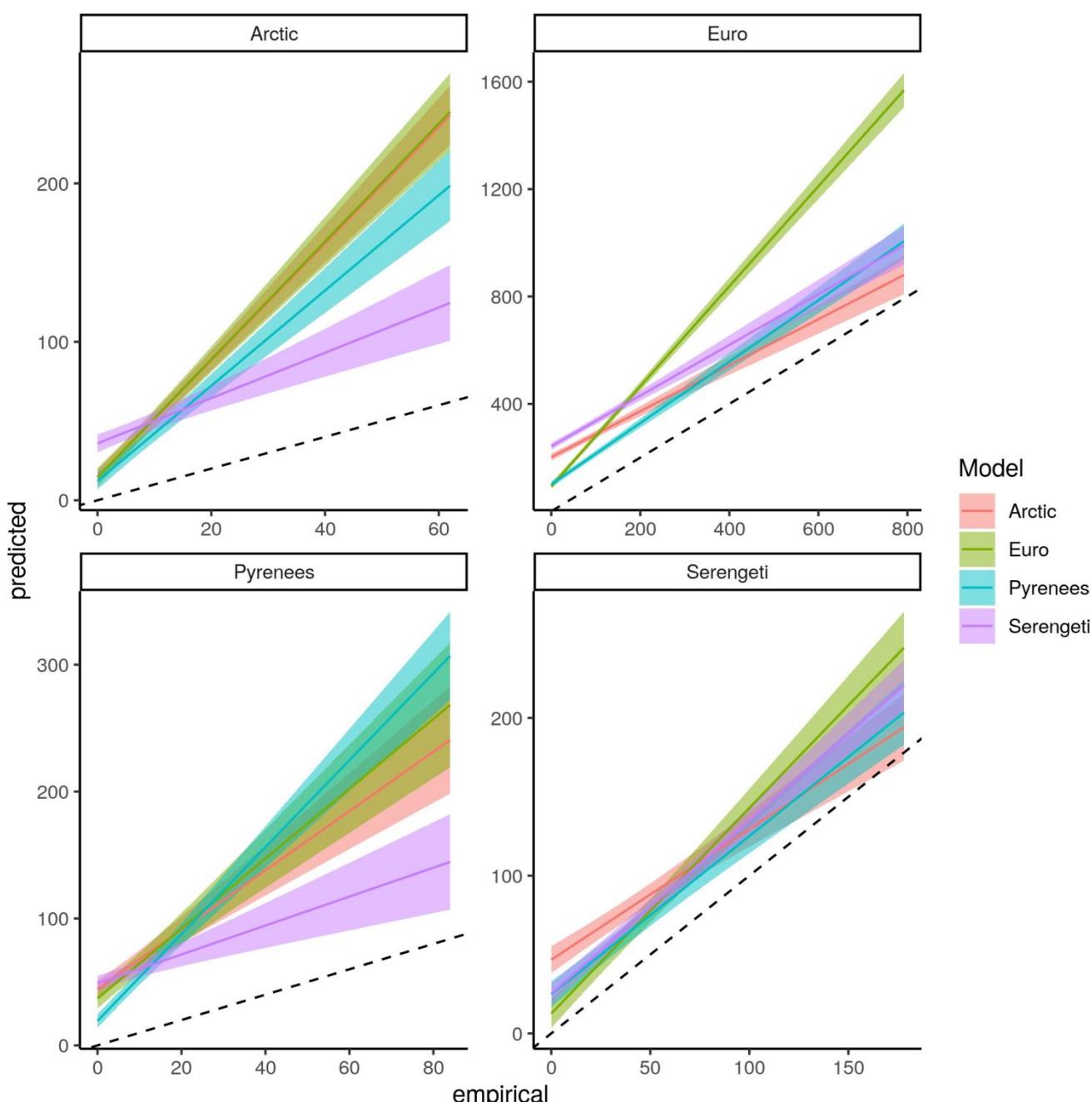


Figure S7.2: Linear regression comparing the predicted number of prey and empirical number of prey of species for each model predicting every food webs.

Number of predators

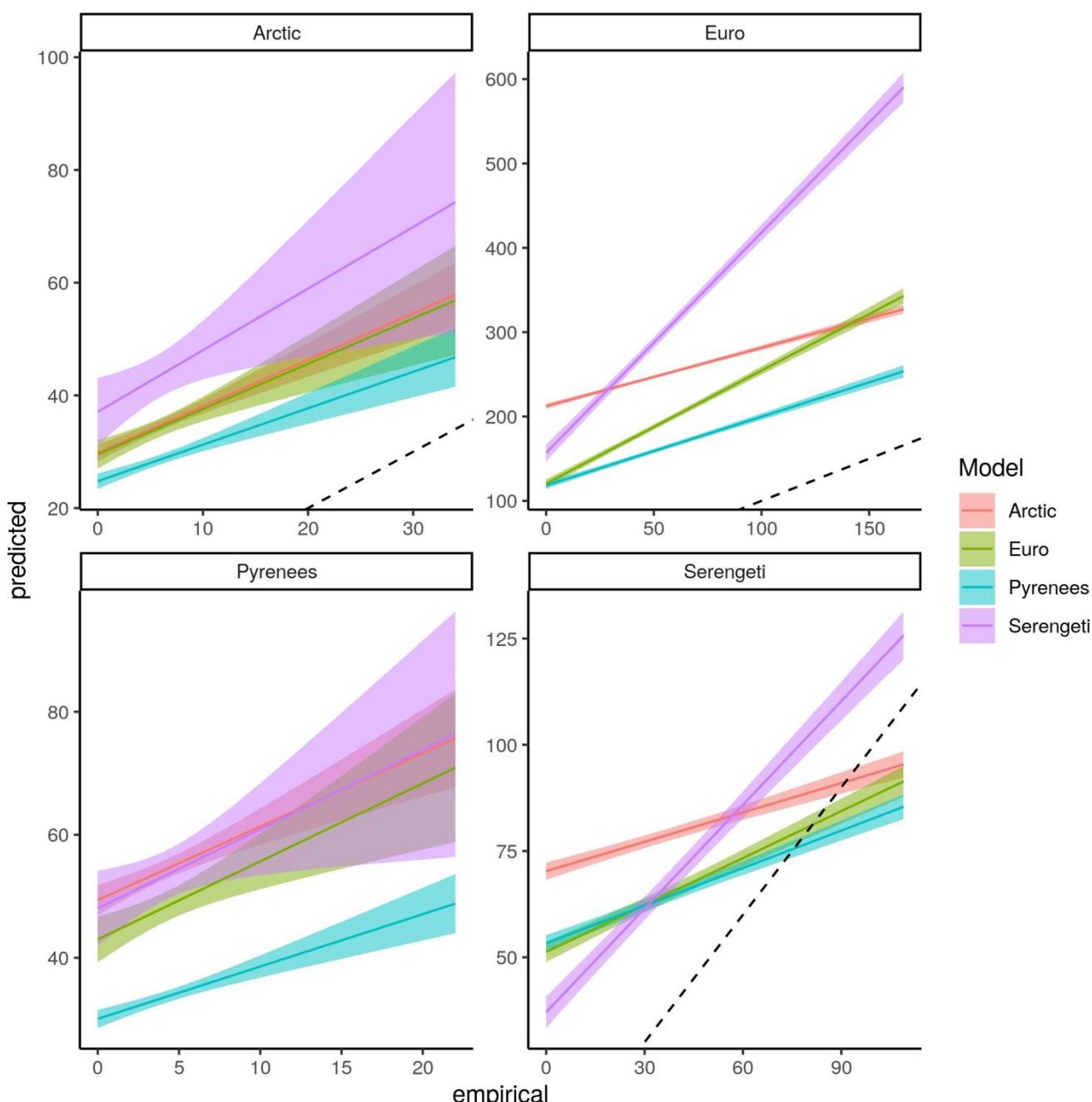
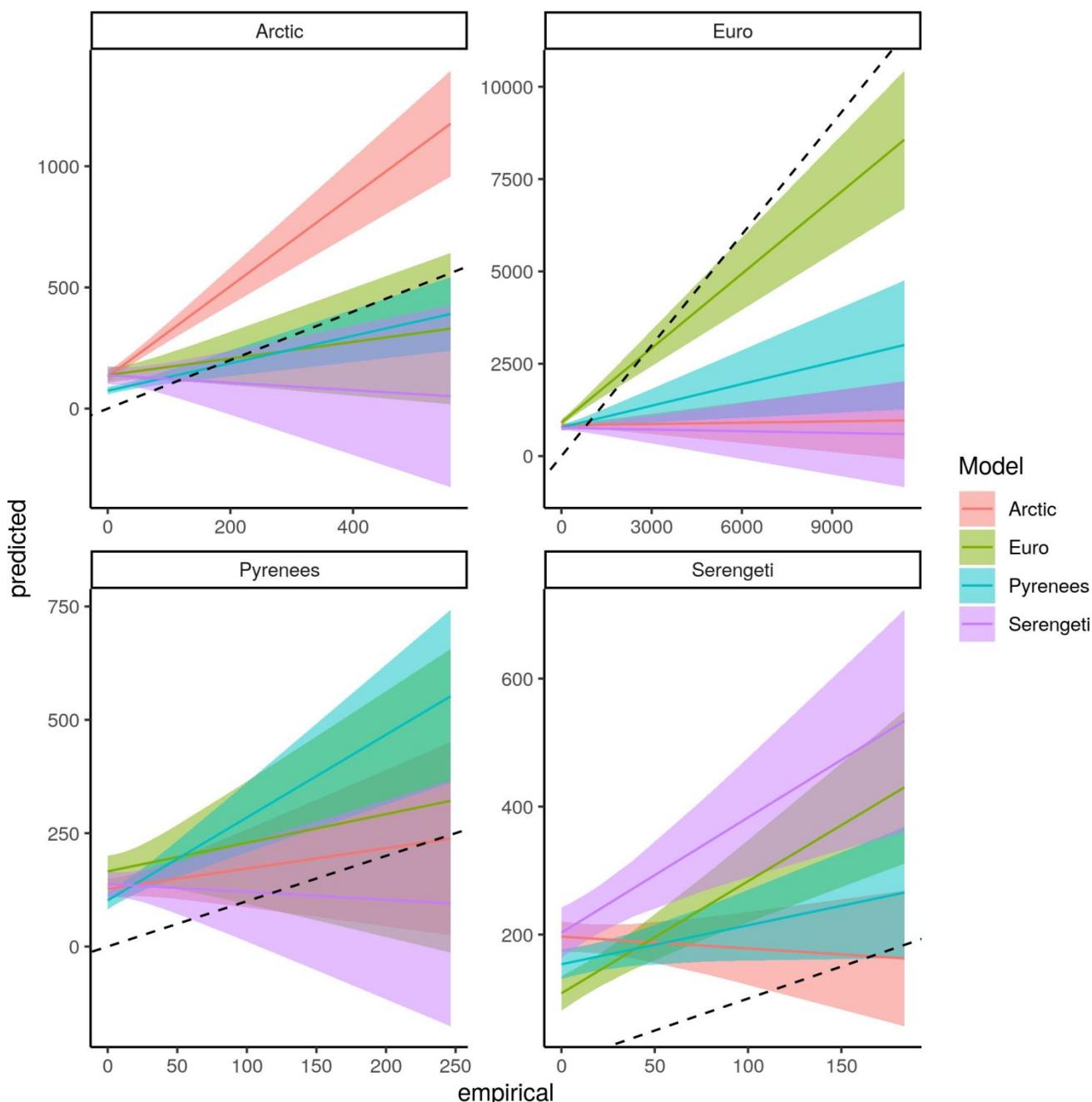


Figure S7.3: Linear regression comparing the predicted number of predators and empirical number of predators of species for each model predicting every food webs.

1
2 Betweenness
3



45 Figure S7.4: Linear regression comparing the predicted betweenness centrality and empirical
46 betweenness centrality of species for each model predicting every food webs.
47

Closeness

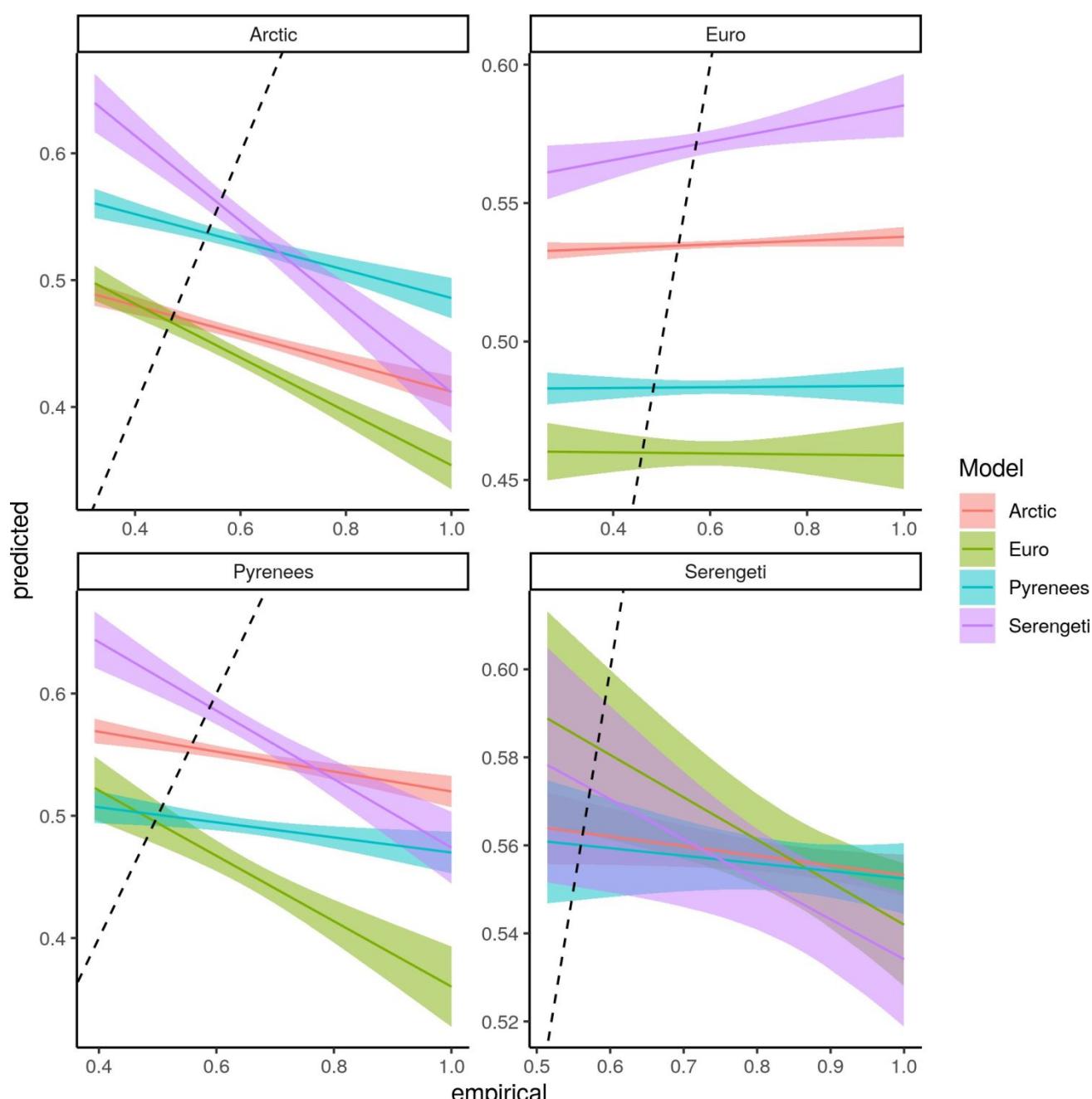


Figure S7.5: Linear regression comparing the predicted closeness centrality and empirical closeness centrality of species for each model predicting every food webs.

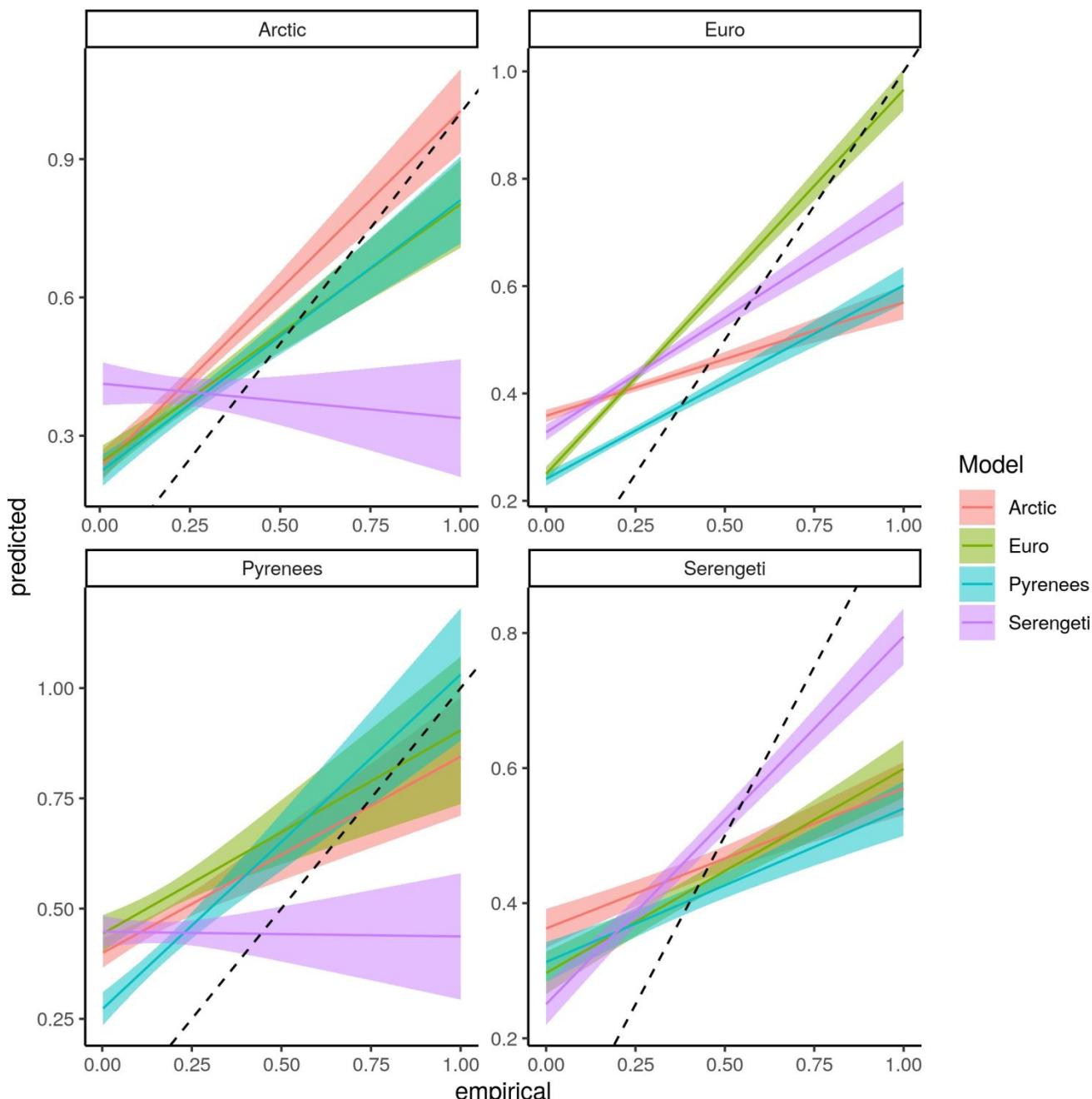
1
2 Eigenvector centrality
3

Figure S7.6: Linear regression comparing the predicted eigenvector centrality and empirical eigenvector centrality of species for each model predicting every food webs.

Trophic Level

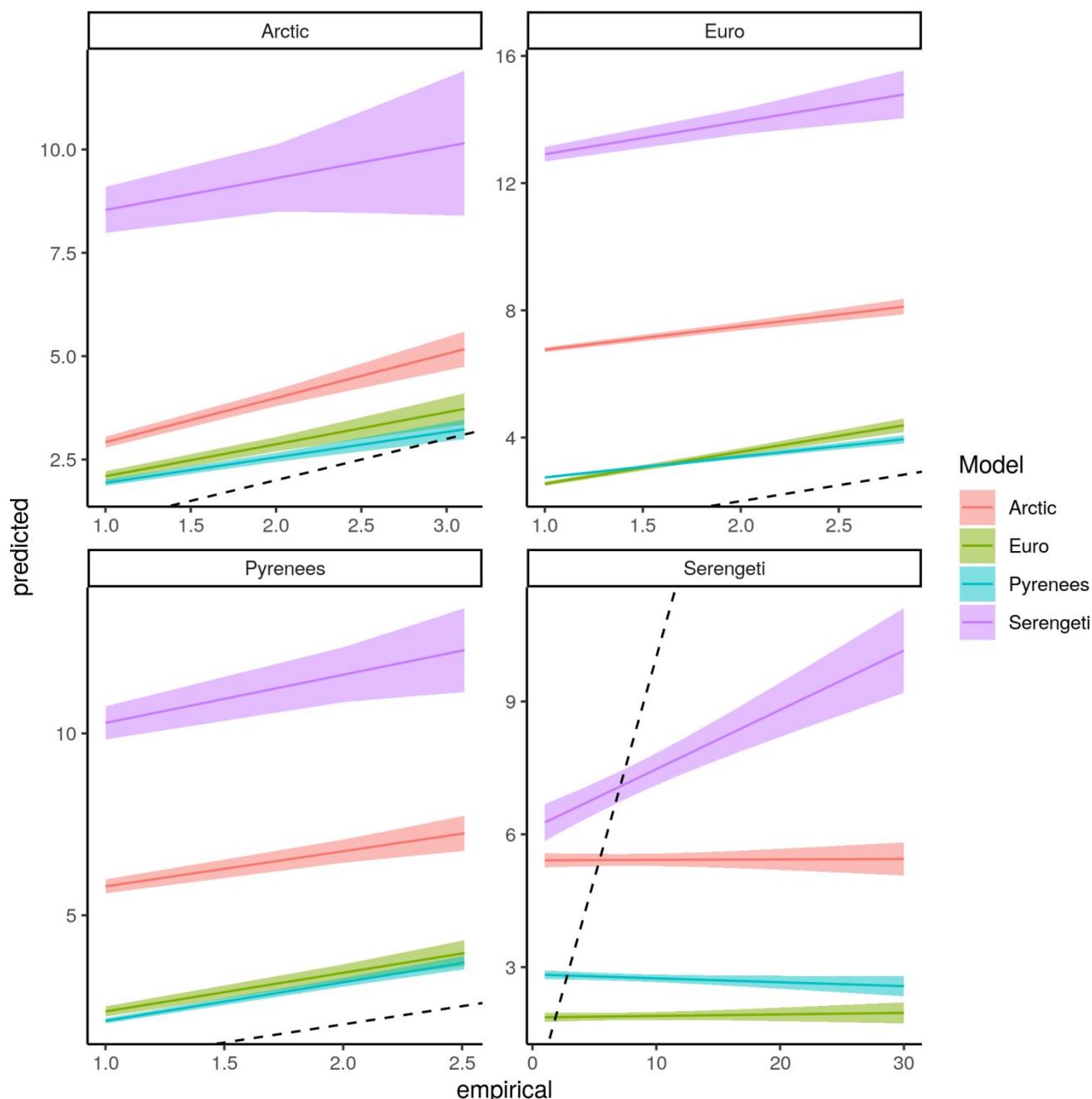


Figure S7.7: Linear regression comparing the predicted trophic level and empirical trophic level of species for each model predicting every food webs.

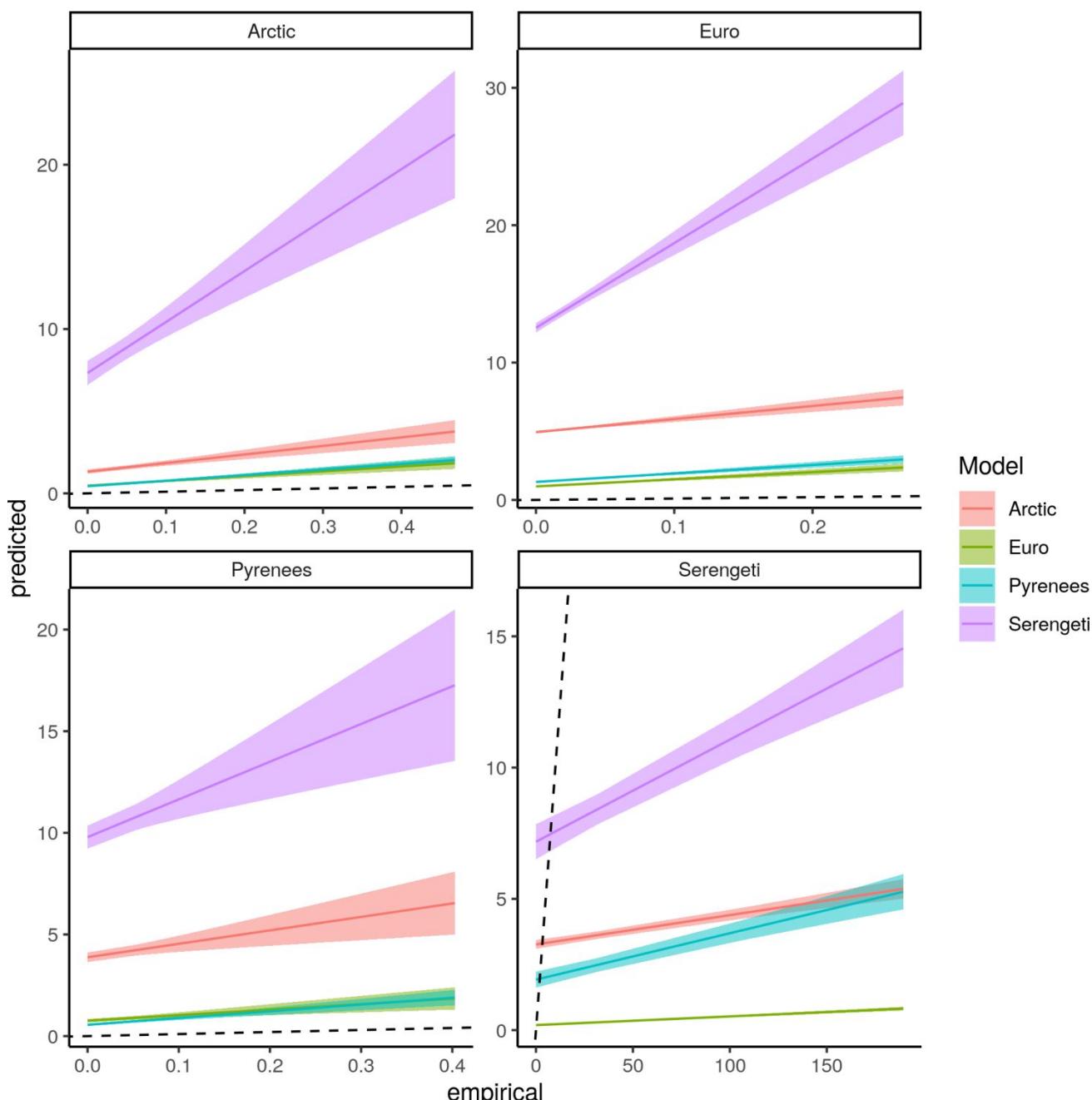
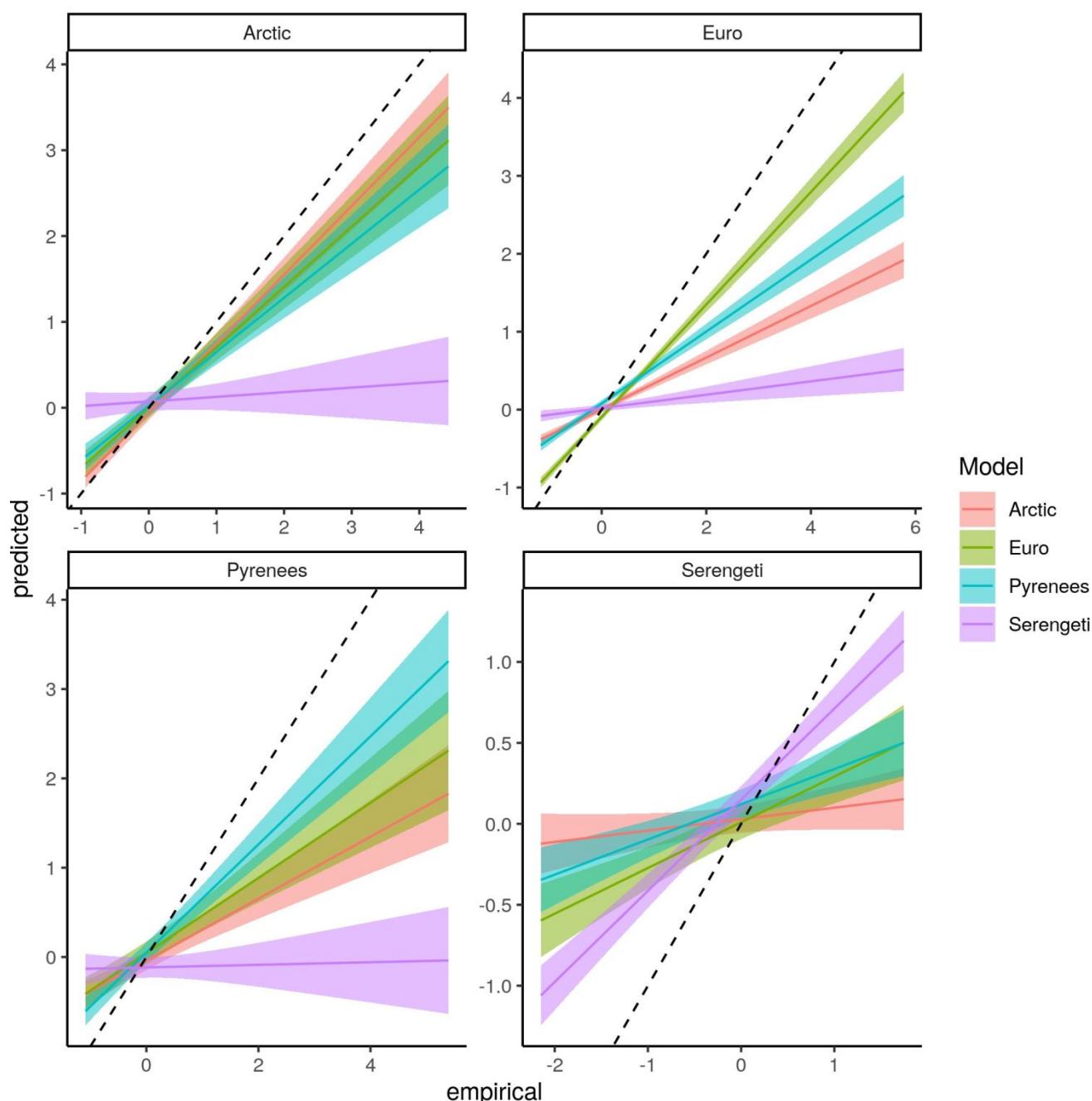
1
2 Omnivory
3

Figure S7.8: Linear regression comparing the predicted omnivory index and empirical omnivory index of species for each model predicting every food webs.

1
2 Within-module degree
3

45 Figure S7.9: Linear regression comparing the predicted within-module degree and empirical within-
46 module degree of species for each model predicting every food webs.
47

Participation coefficient

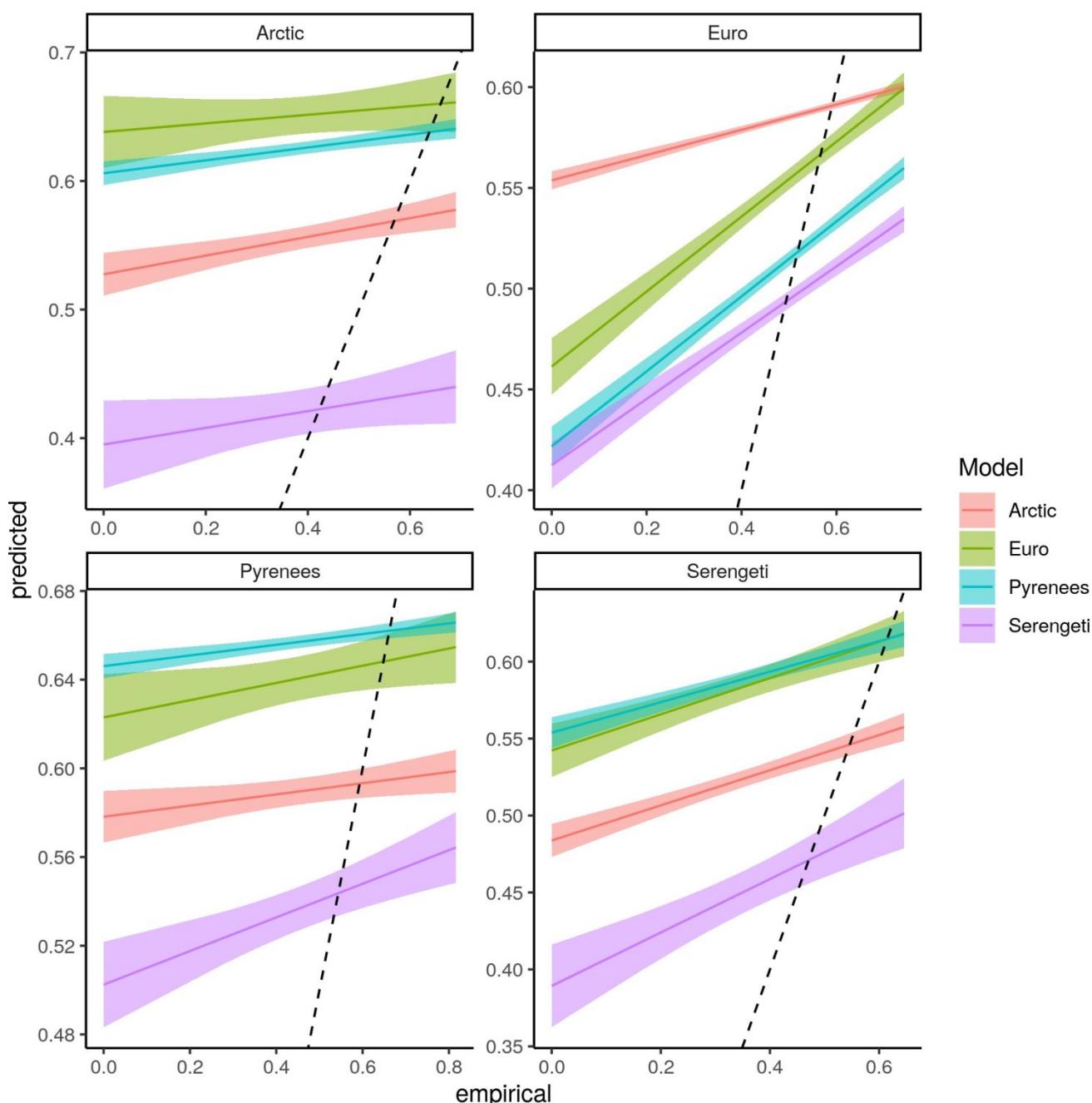


Figure S7.10: Linear regression comparing the predicted participation coefficient and empirical participation coefficient of species for each model predicting every food webs.

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

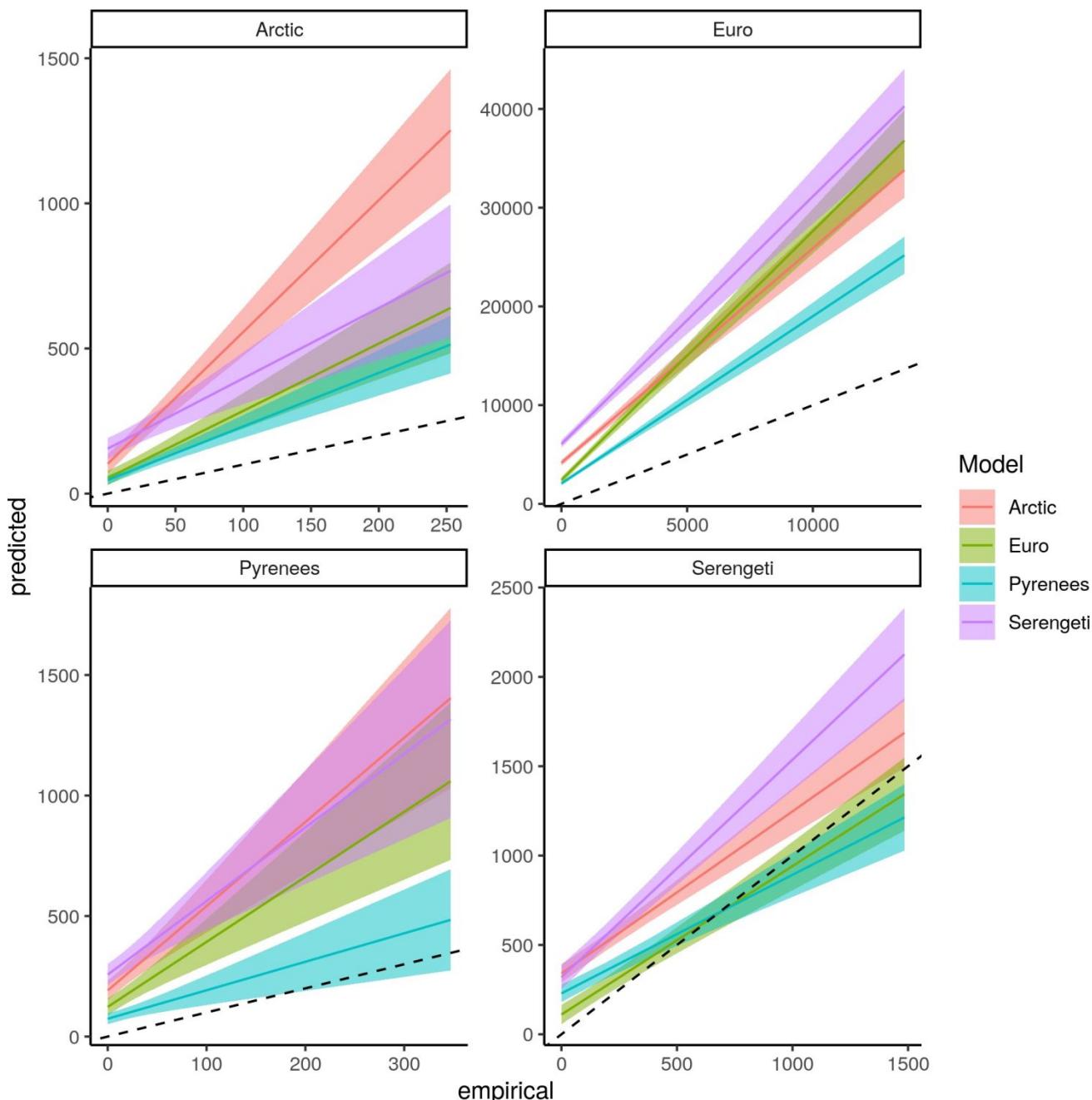


Figure S7.11: Linear regression comparing the predicted frequency of the top predator position in the linear food chain motif and empirical frequency of the top predator position in the linear food chain motif of species for each model predicting every food webs.

Frequency of the consumer position in the linear food chain motif

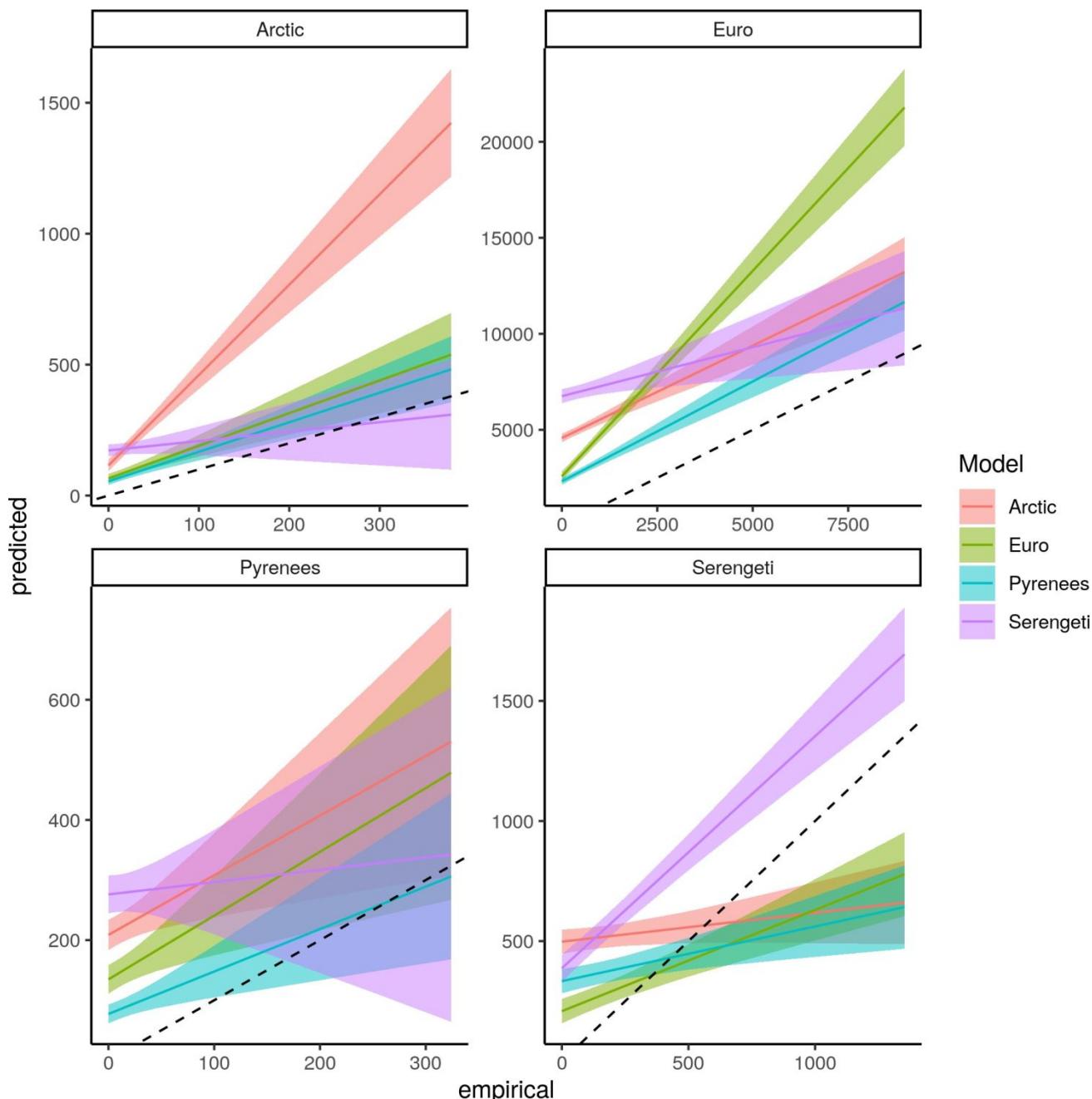


Figure S7.12: Linear regression comparing the predicted frequency of the consumer position in the linear food chain motif and empirical frequency of the consumer position in the linear food chain motif of species for each model predicting every food webs.

Frequency of the resource position in the linear food chain motif

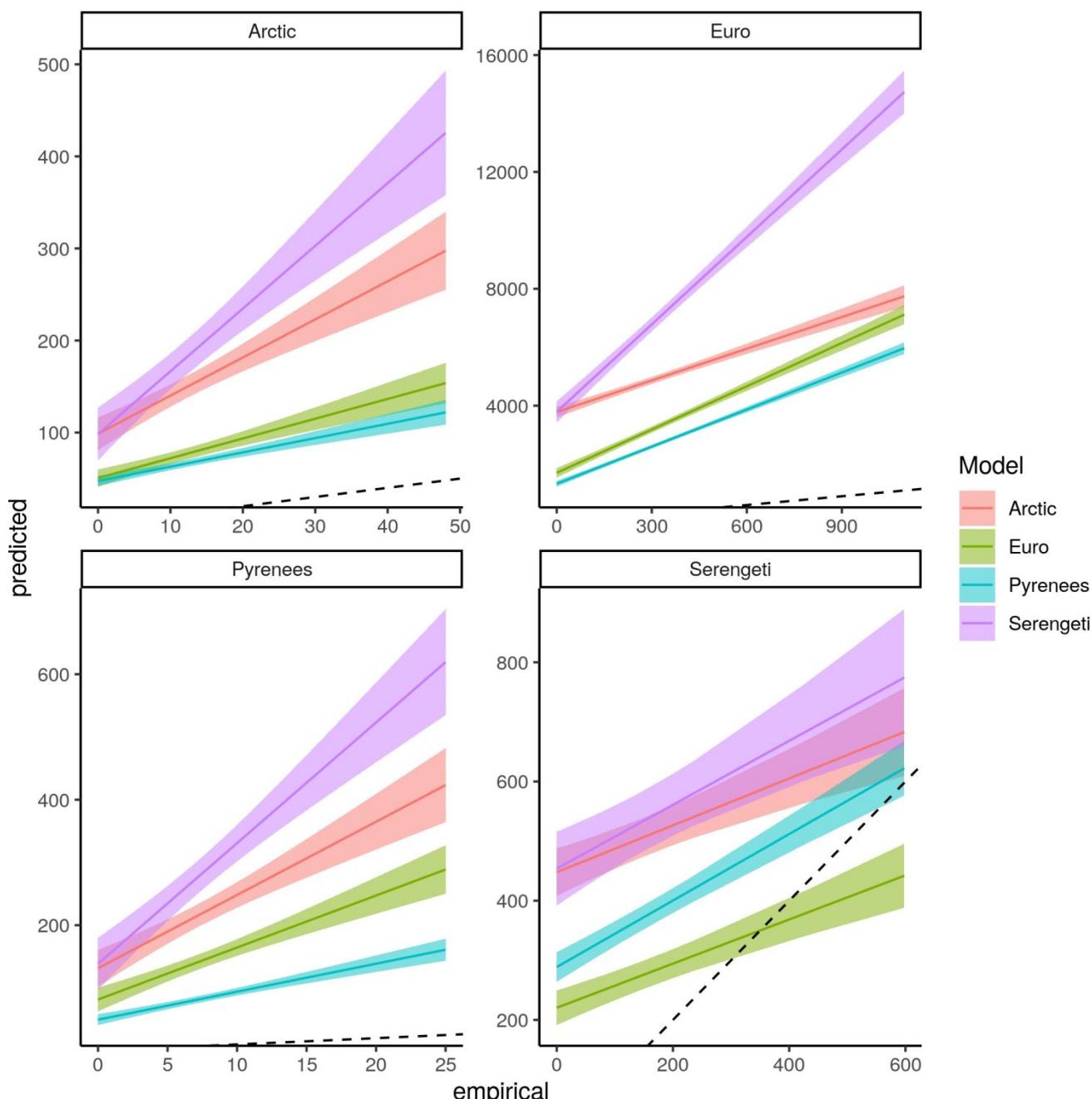


Figure S7.13: Linear regression comparing the predicted frequency of the resource position in the linear food chain motif and empirical frequency of the resource position in the linear food chain motif of species for each model predicting every food webs.

Frequency of the omnivore position in the intraguild predation motif

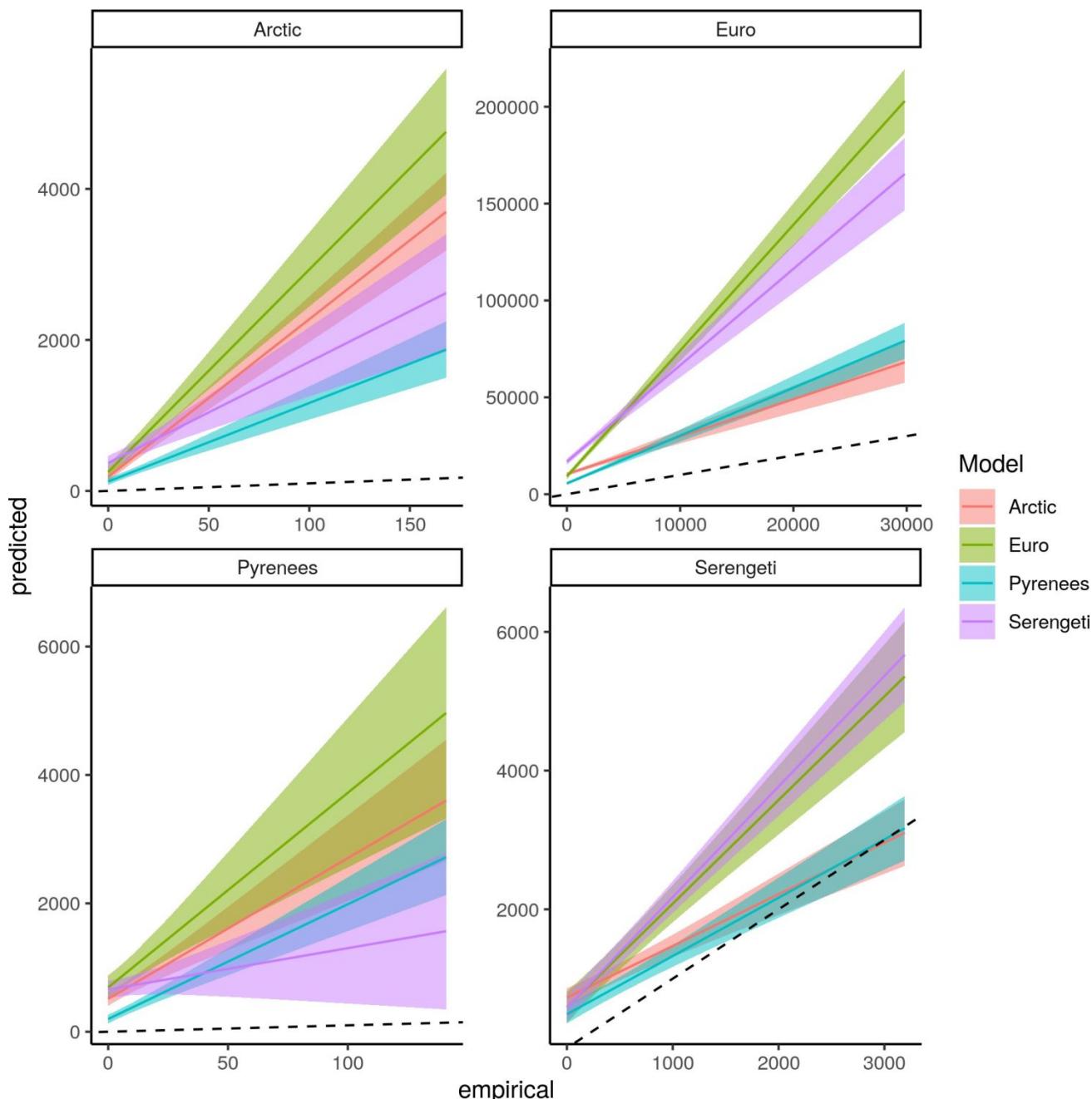


Figure S7.14: Linear regression comparing the predicted frequency of the omnivore position in the intraguild predation motif and empirical frequency of the omnivore position in the intraguild predation motif of species for each model predicting every food webs.

Frequency of the consumer position in the intraguild predation motif

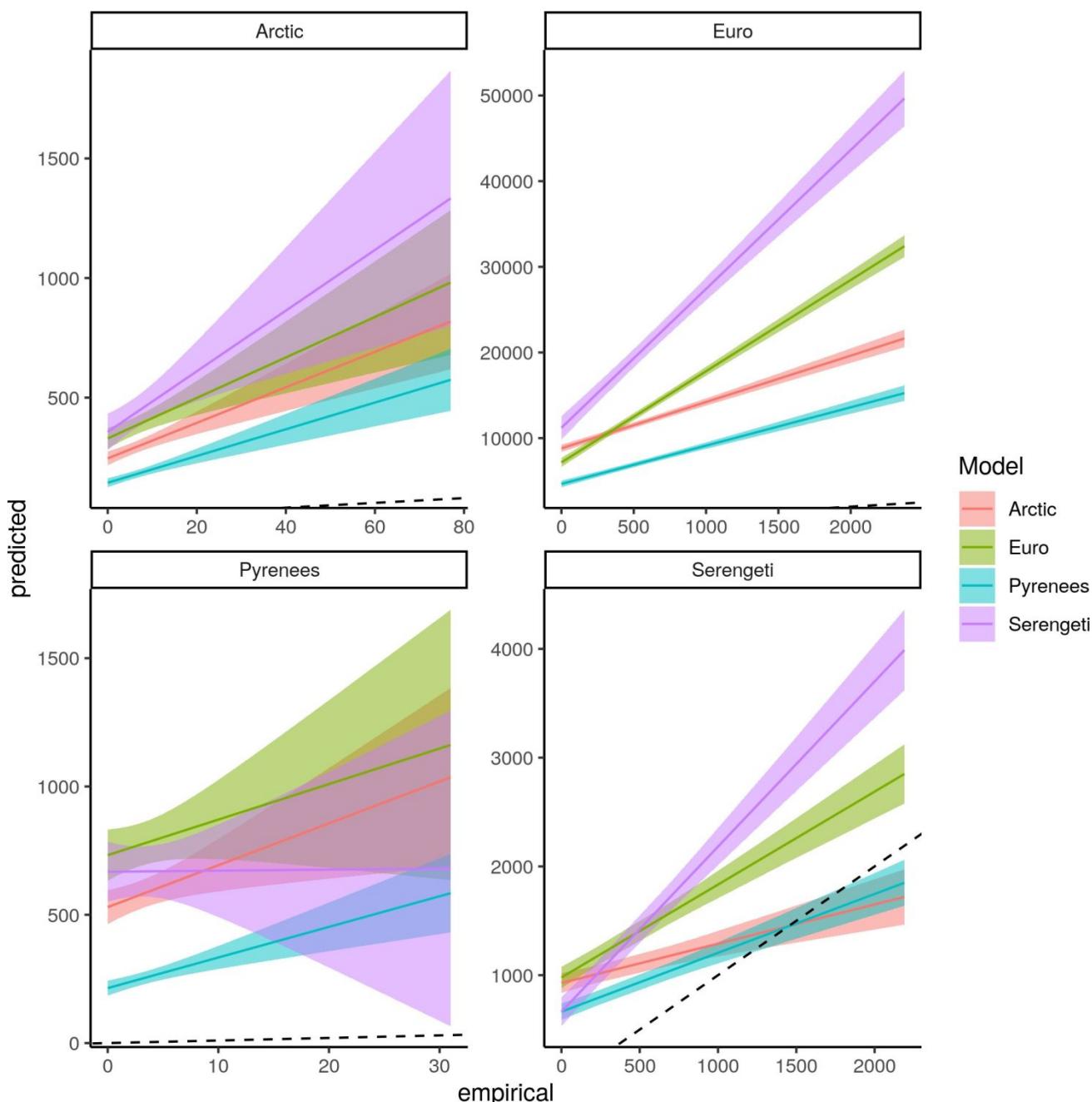


Figure S7.15: Linear regression comparing the predicted frequency of the consumer position in the intraguild predation motif and empirical frequency of the consumer position in the intraguild predation motif of species for each model predicting every food webs.

Frequency of the resource position in the intraguild predation motif

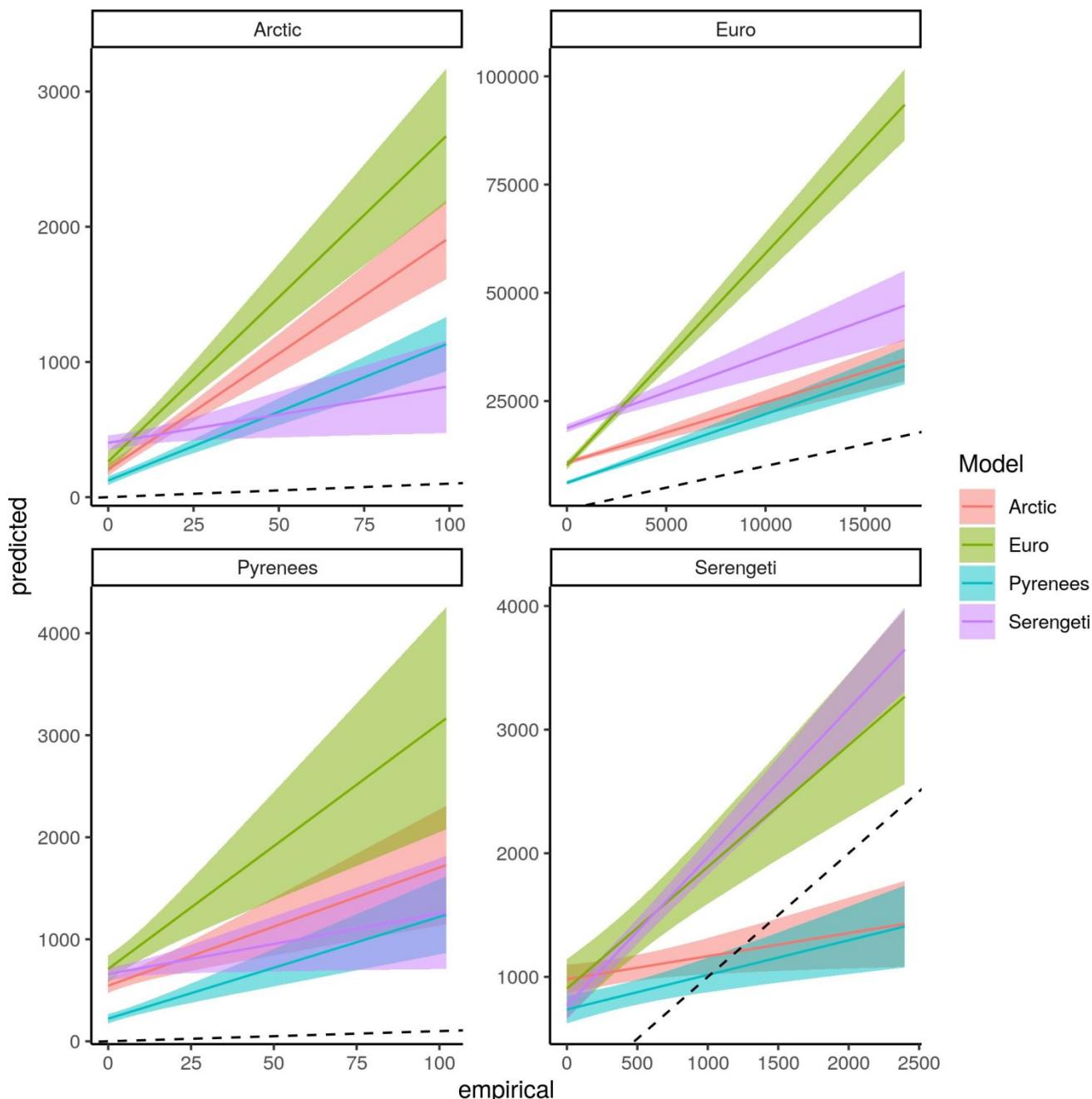


Figure S7.16: Linear regression comparing the predicted frequency of the resource position in the intraguild predation motif and empirical frequency of the resource position in the intraguild predation motif of species for each model predicting every food webs.

Frequency of the consumer position in the direct competition motif

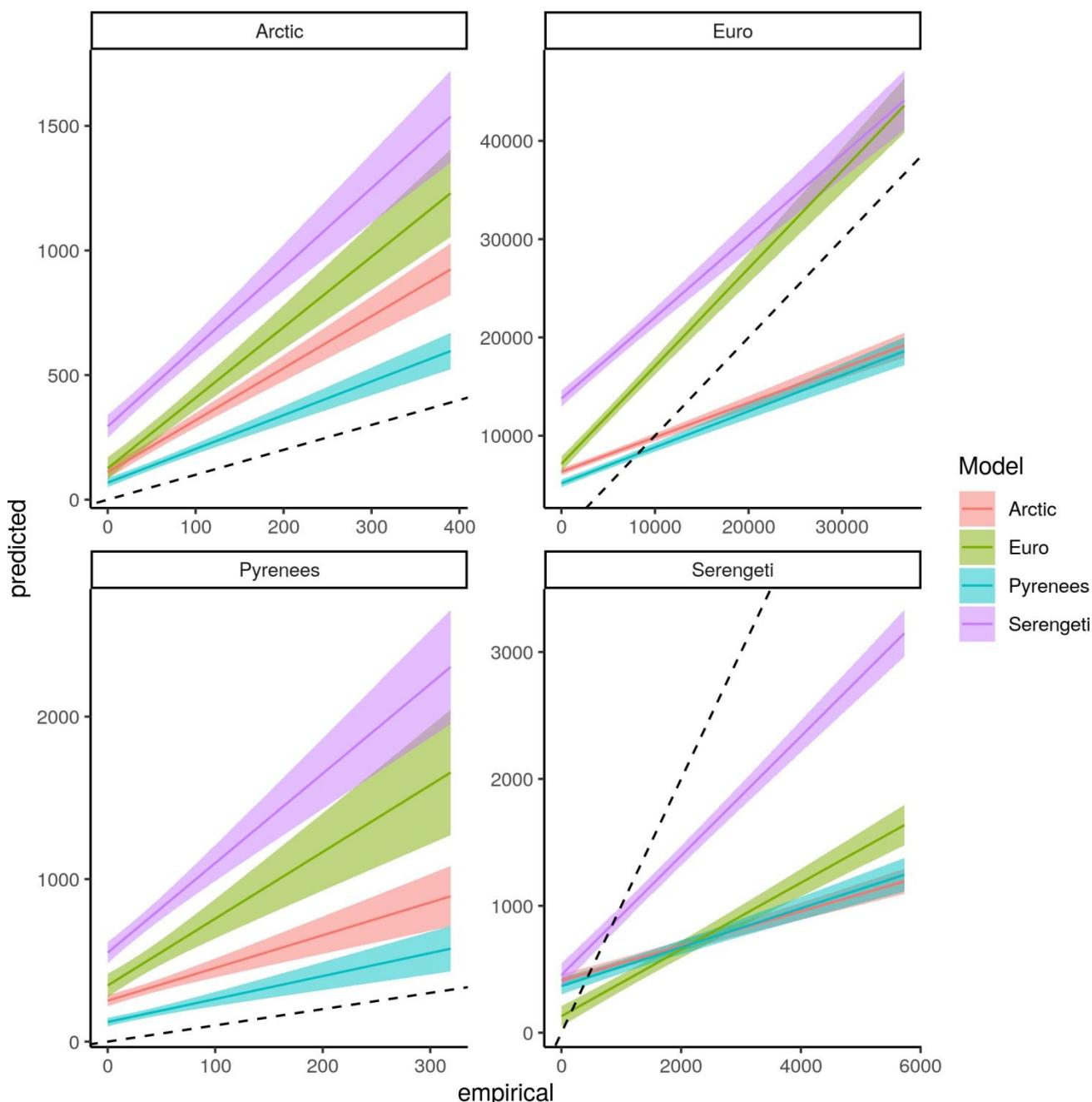
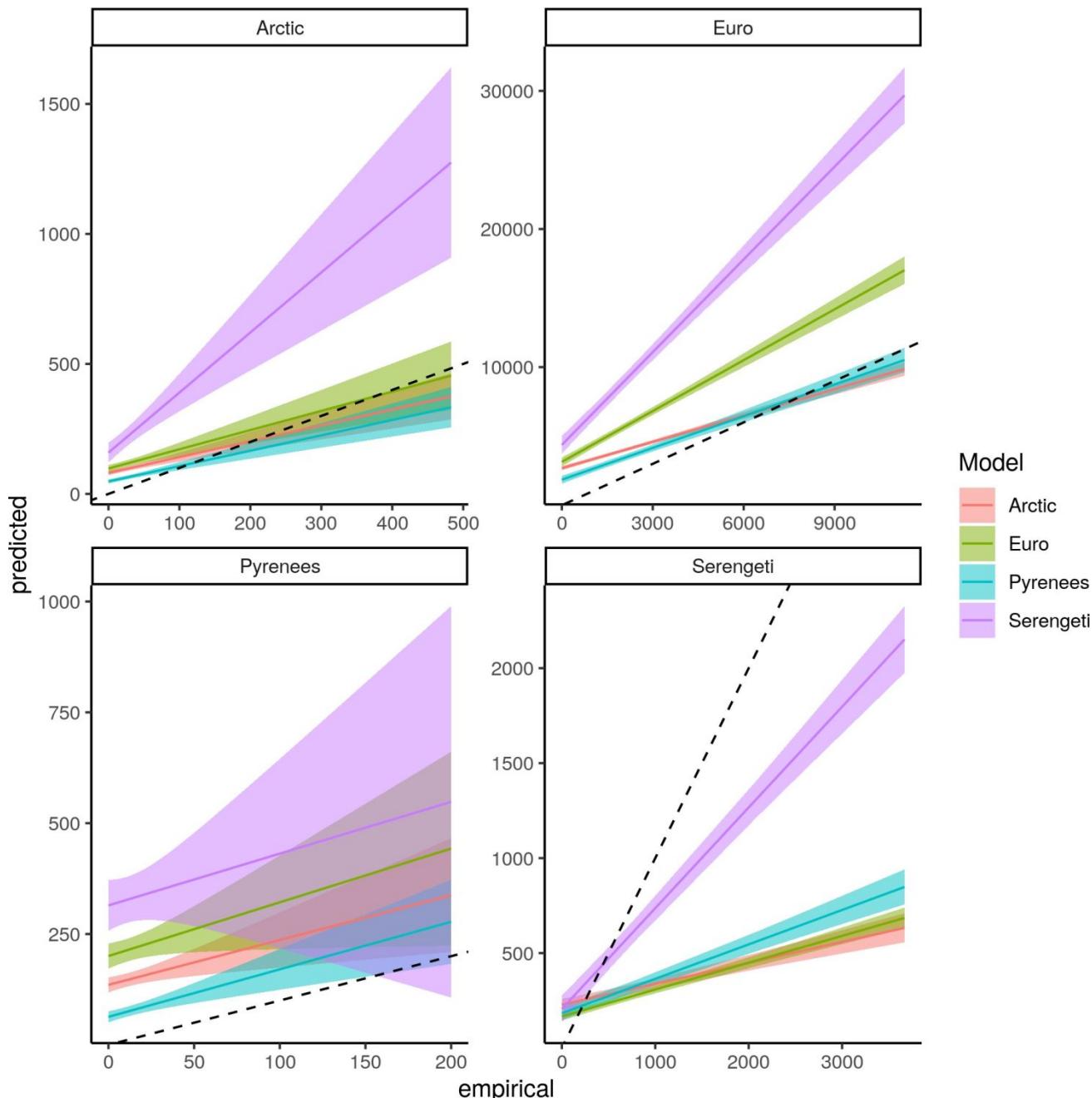


Figure S7.17: Linear regression comparing the predicted frequency of the consumer position in the direct competition motif and empirical frequency of the consumer position in the direct competition motif of species for each model predicting every food webs.

1
2 Frequency of the resource position in the direct competition motif
3



45 Figure S7.18: Linear regression comparing the predicted frequency of the resource position in the
46 direct competition motif and empirical frequency of the resource position in the direct competition
47 motif of species for each model predicting every food webs.
48

Frequency of the consumer position in the apparent competition motif

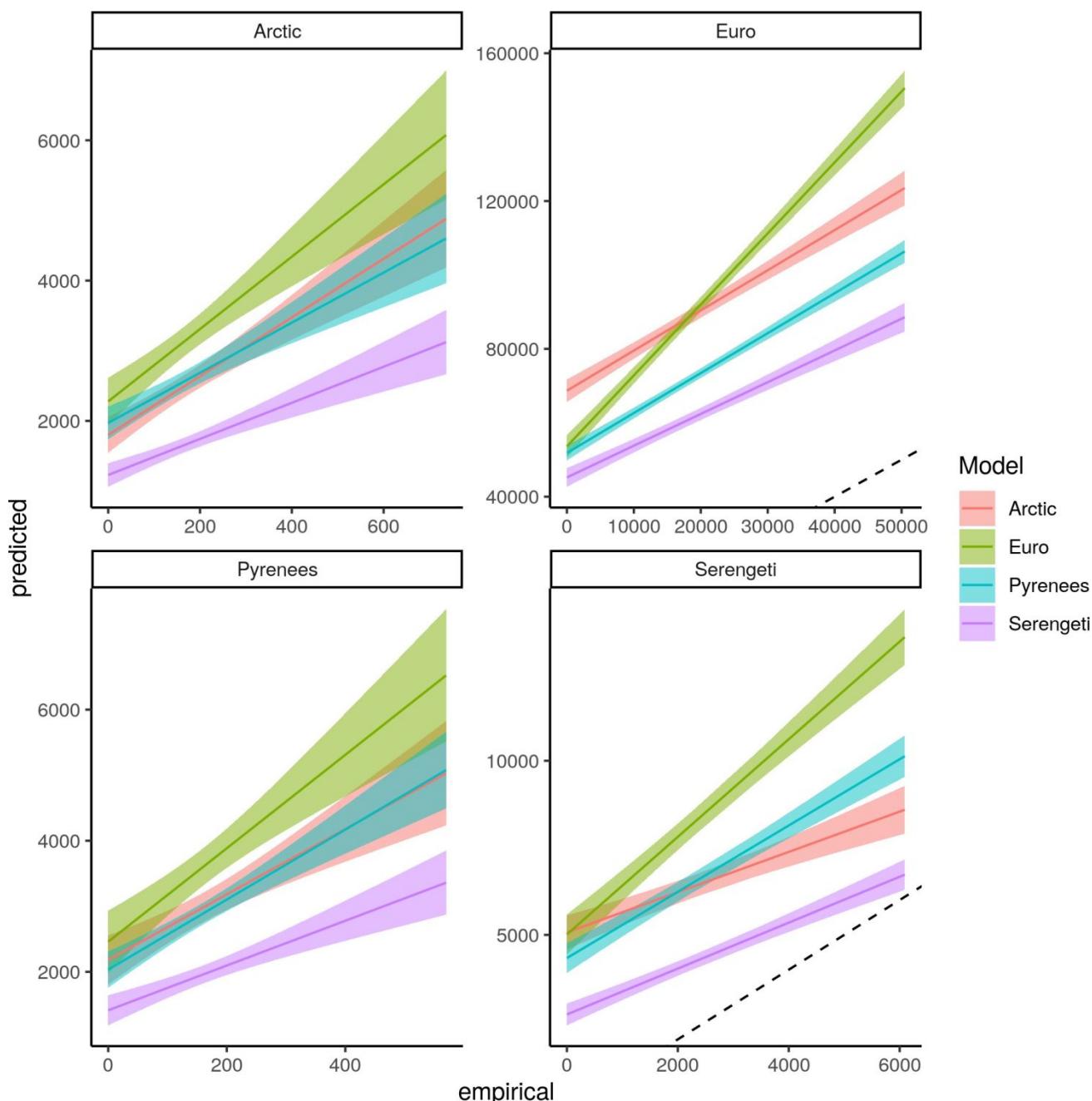


Figure S7.19: Linear regression comparing the predicted frequency of the consumer position in the apparent competition motif and empirical frequency of the consumer position in the apparent competition motif of species for each model predicting every food webs.

Frequency of the resource position in the apparent competition motif

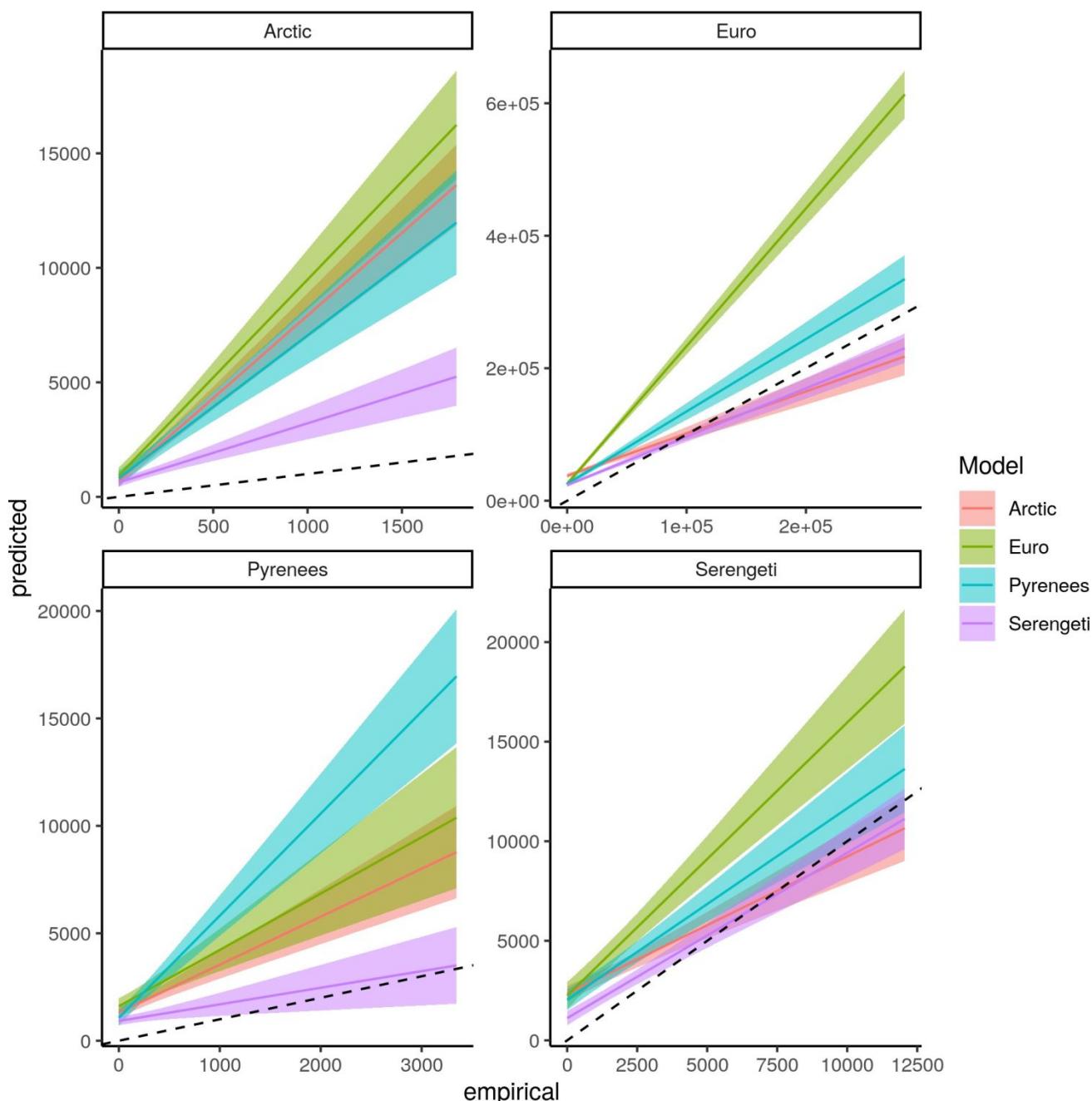


Figure S7.20: Linear regression comparing the predicted frequency of the resource position in the apparent competition motif and empirical frequency of the resource position in the apparent competition motif of species for each model predicting every food webs.

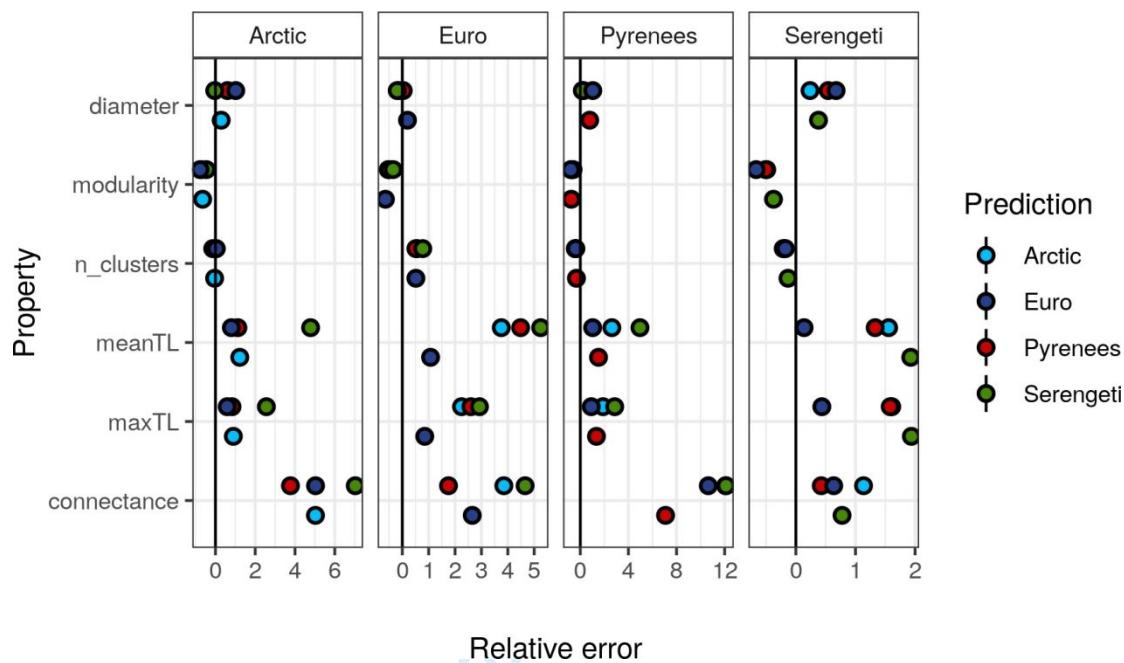
Appendix 8: Predicting food web properties extended results

Figure S8.1: 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. Arctic refers to the Northern Québec and Labrador food web

Appendix S9 – 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>

- 1
2 Duffy, J. E., Cardinale, B. J., France, K. E., McIntyre, P. B., Thébaud, E., & Loreau, M. (2007). The
3 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>
- 4
5 Dunne, J. A., Williams, R. J., & Martinez, N. D. (2002). Food-web structure and network theory: The
6 role of connectance and size. *Proceedings of the National Academy of Sciences*, 99(20), 12917–
7 12922. <https://doi.org/10.1073/pnas.192407699>
- 8
9 Dyer, L. A., & Letourneau, D. (2003). Top-down and bottom-up diversity cascades in detrital vs.
10 Living food webs. *Ecology Letters*, 6(1), 60–68. <https://doi.org/10.1046/j.1461-0248.2003.00398.x>
- 11
12 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>
- 13
14 Freeman, L. C. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry*, 40(1), 35–
15 41. <https://doi.org/10.2307/3033543>
- 16
17 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>
- 18
19 Guimerà, R., & Amaral, L. A. N. (2005). Cartography of complex networks: Modules and universal
20 roles. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(02), P02001.
21 <https://doi.org/10.1088/1742-5468/2005/02/P02001>
- 22
23 Hijmans, R. J. (2021). *raster: Geographic data analysis and modeling* [Manual]. <https://CRAN.R-project.org/package=raster>
- 24
25 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>
- 26
27 Kones, J. K., Soetaert, K., Oevelen, D. van, & Owino, J. (2009). Are network indices robust indicators
28 of food web functioning? A Monte Carlo approach. *Ecological Modelling*, 220, 370–382. <http://dx.doi.org/10.1016/j.ecolmodel.2008.10.012>
- 29
30 Lurgi, M., López, B. C., & Montoya, J. M. (2012). Climate change impacts on body size and food web
31 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>

- 1
2 Maiorano, L., Montemaggiori, A., Ficetola, G. F., O'Connor, L., & Thuiller, W. (2020). TETRA EU
3
4 1.0: A species level trophic metaweb of European tetrapods. *Global Ecology and*
5
6 *Biogeography*, geb.13138. <https://doi.org/10.1111/geb.13138>
- 7
8 Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. (2002). Network Motifs:
9
10 Simple Building Blocks of Complex Networks. *Science*, 298(5594), 824–827.
11
12 <https://doi.org/10.1126/science.298.5594.824>
- 13
14 Saravia, L. A. (2022). *multiweb: Ecological network analysis includind multiplex networks.*
15
16 <https://github.com/lasaravia/multiweb>
- 17
18 Stekhoven, D. J. (2013). *missForest: Nonparametric missing value imputation using random forest*
19
20 [Manual].
- 21
22 Stekhoven, D. J., & Buehlmann, P. (2012). MissForest—Non-parametric missing value imputation for
23
24 mixed-type data. *Bioinformatics (Oxford, England)*, 28(1), 112–118.
- 25
26 Stouffer, D. B., & Bascompte, J. (2010). Understanding food-web persistence from local to global
27
28 scales. *Ecology Letters*, 13(2), 154–161. <https://doi.org/10.1111/j.1461-0248.2009.01407.x>
- 29
30 Stouffer, D. B., & Bascompte, J. (2011). Compartmentalization increases food-web persistence.
31
32 *Proceedings of the National Academy of Sciences*, 108(9), 3648–3652.
33
34 <https://doi.org/10.1073/pnas.1014353108>
- 35
36 Stouffer, D. B., Sales-Pardo, M., Sirer, M. I., & Bascompte, J. (2012). Evolutionary Conservation of
37
38 Species' Roles in Food Webs. *Science*, 335(6075), 1489–1492.
39
40 <https://doi.org/10.1126/science.1216556>
- 41
42 Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P.-C. (2021). Rank-Normalization,
43
44 Folding, and Localization: An Improved $R^{\hat{}}$ for Assessing Convergence of MCMC (with
45
46 Discussion). *Bayesian Analysis*, 16(2), 667–718. <https://doi.org/10.1214/20-BA1221>
- 47
48 Vieira, M. C., & Almeida-Neto, M. (2015). A simple stochastic model for complex coextinctions in
49
50 mutualistic networks: Robustness decreases with connectance. *Ecology Letters*, 18(2), 144–152.
51
52 <https://doi.org/10.1111/ele.12394>
- 53
54 Zhao, Q., Brink, P. J. V. den, Carpentier, C., Wang, Y. X. G., Rodríguez Sánchez, P., Xu, C.,
55
56 Vollbrecht, S., Gillissen, F., Vollebregt, M., Wang, S., & Laender, F. D. (2019). Horizontal and
57
58
59
60

1
2 vertical diversity jointly shape food web stability against small and large perturbations. *Ecology*
3
4 *Letters*, 22(7), 1152–1162. <https://doi.org/10.1111/ele.13282>
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review