

Draft information

version: draft-0.0

20xx

Requests to coauthors

- Should I include more networks? Easy to get some more from other published studies. Could improve the manuscript but every new data set has to be cleaned independently. This could take around 2 days per network.
- Should I increase the resolution of the bioclimatic variables? Currently working with 2.5' (~5km). Increasing to 30" (~1km) might improve the analysis but we're talking about working with ~25GB+ of climatic data instead of the 1GB used now. Unsure the benefits would outweigh the costs.
- Audrey: Could you please fill in the gaps in the rationale behind filtering out occurrences with potential low quality coordinates (page ??, line ??)
- I chose, rather arbitrarily, a mean absolute error of 10% for the sensitivity analysis. Does this choice sound appropriate to you?
- Should I remove the species overall generalism and the size of the grinellian niche from the model to simplify things. Both contribute little to the model and the story. For the former, seems to be redundant once the number of possible interacting species has been included in the model. For the later, there is little biological reasons (what I could gather from the lit.) to justify their inclusion in the model. What you think? If not, at the very least I think I would move panels a and c in Fig. 5 to the supplementary info.
- What you think would be the best way to end the paper? Thinking about what could be highlighted in the conclusion.

What would change or improve? which result kings or this?

some question as above

why not include grinell but cut number possible?

Done

- Downloaded network data from the Web of Life and cleaned it. I didn't anticipated data needed SO MUCH work! I estimated three days to do this but took roughly 3 weeks to do it properly and there is still some work to do. Quite a mess...

- Downloaded bioclimatic and habitat data
- Downloaded species occurrences. Wasn't straightforward as there is a lot of species and lots of restrictions with the GBIF API. About a week getting this to work.
- Calculated environmental niche size and the environmental suitability of species in their communities
- Modelled the normalised degree of species as a proxy of eltonian niche size using a bayesian framework

To-do

- Manually fix the remaining species names that are incorrect and were not picked up by the automatic procedure. Doing 5-10 per day, so should be done in a couple months
- Fix references. There are lots of mistakes re. author, years and publication venue particularly for those from which data was downloaded.
- Finish off discussion with a punchy conclusion.
- Fix the WAIC table. Currently shows the model without guild as unknown.
- Add Lupe's feedback to the intro. Need to better define stress and Eltonian/Grinnellian niche and don't need to say realised Eltonian niche when I really just mean number of partners

Environmental stress affects niche breadth in plant-pollinator communities

*E. Fernando Cagua*¹ (*efc29@uclive.ac.nz*)
*Audrey Lustig*² (*audrey.lustig@canterbury.ac.nz*)
*Jason M. Tylianakis*¹ (*jason.tylianakis@canterbury.ac.nz*)
*Daniel B. Stouffer*¹ (*daniel.stouffer@canterbury.ac.nz*)

~ see Alyssa's
methods

¹ Centre for Integrative Ecology, School of Biological Sciences, University of

Canterbury, Private Bag 4800, Christchurch 8041, New Zealand

² Geospatial Research Institute, University of Canterbury, Private Bag 4800,

Christchurch 8041, New Zealand

Running title: XX

Keywords: XX

Type of article: XX

Number of words: 151 in abstract; 5,025 in main text.

Number of displays: 6 figures; 1 tables; 0 text boxes.

Number of references: 86

Author for correspondence: E. Fernando Cagua (+64 20 4026 8153).

Data accessibility: Data supporting the results will be accessible in an

appropriate data repository after publication. The DOI will be included

here.

Author contributions: XX

16 Abstract

17 Evidence that the environment influences the interaction between species
18 is rapidly accumulating. However, how it happens is currently unclear as
19 environmental gradients appear to have contrasting or non-linear effects on
20 the species' trophic niche breadth depending on the environmental variable.
21 Here, we explore the relationship between the stresses imposed by the envi-
22 ronment, instead of environmental gradients directly, and niche breadth using
23 a global dataset of plant-pollinator interactions. We found that environmental
24 stress plays a significant role in determining the number of partners a species
25 interacts with, but this role is highly variable across species. In particular,
26 when faced with environmental stress, species that have a large number of
27 interactions are more likely to focus on a smaller number of, presumably
28 higher-quality, interactions. Contrastingly, the specialists that can cope
29 with increased stress are more likely to broaden their niche and engage in
30 opportunistic interactions, effectively behaving as facultative generalists.

31 Introduction

32 Species interactions are known to vary widely across space and time. There
33 are multiple examples of species that interact tightly in an ecological com-
34 munity or a particular season but not in another. It is crucial to understand
35 why and how this turnover of interactions occur because the network of
36 interactions of an ecological community is known to underpin ecosystem
37 function and stability. Despite much progress in the last decades, we are
38 still unable to predict species interactions, partially because we still have not
39 accurately identified the factors that drive species to interact under certain
40 conditions and not under another.

41 There are two main processes involved in whether two species in a community
42 interact or not. The first is an "Eltonian" niche process that is underpinned
43 by the traits of the interacting species. Compared to other kinds of trophic
44 interactions, mutualistic interactions, and plant-pollination interactions,
45 in particular, are relatively generalised. While a large proportion of ~~the~~ ^{mutualistic}
46 interactions appear to be opportunistic, a fraction of interactions shows a
47 strong signature of deep co-evolutionary history (Hutchinson, Cagua, and
48 Stouffer 2017). In contrast to opportunistic interactions with low ^{phenotypic} ~~trait~~
49 matching, the more co-evolved interactions with optimally matched traits are
50 more persistent and presumably lead to a more substantial per-capita benefit
51 for the interacting species. ^(not) The second is a neutral process underpinned by
52 the encounter rate of the interacting species. All else being equal, when the
53 encounter rate of interacting species is high, there is a higher probability ^{that they}
54 interact than if the ^{encounter} rate were lower, for example, due to species rarity.
55 Perhaps ^{is thought to} because it ^{role of the} operates at scales larger than those of the Eltonian niche
56 concept, the abiotic environment ~~is one of the factors that~~ is commonly
57 overlooked by community ecologists when explaining interaction turnover
58 (but see Gravel et al. 2018). The most evident—and best-studied—way

Why mention
Eltonian here
instead of just
niche?

This is not my
reading of the
literature so much
to be backed up
somehow

this paragraph needs more references from literature

59 through which the environment can affect the structure of ecological networks
60 is by way of patterns of community composition. At the heart of biogeography
61 is the idea that the abiotic environment can determine the “Grinnellian” niche
62 of a species; ^{the locations where a} ~~this is whether~~ the species can have a positive growth-rate when
63 rare ~~in a particular location~~. For two species to interact, they ^{clearly} first need to
64 co-occur in the same community, ^{and this} ~~which~~ can only happen if the environmental
65 conditions of ^{that} ~~the~~ community fall within the Grinnellian niche of both species.

66 Although less understood, the environment can also affect the interactions
67 themselves, after community composition has been established (Tylianakis
68 and Morris 2017). For instance, some studies show that ^{the strength of} trophic interactions,
69 like predation (McKinnon et al. 2010; Vucic-Pestic et al. 2011) and herbivory
70 (Baskett and Schemske 2018), can increase with temperature but might
71 decrease with precipitation (Pires et al. 2016), and both plant-plant (Bertness
72 and Callaway 1994) and plant-pollinator (Tur et al. 2016) interactions
73 can switch from facilitative to competitive along an elevation/temperature
74 gradient. Contrastingly, other studies have shown either no effect or non-
75 linear effects of temperature or precipitation on plant-pollinator interactions
76 (Devoto, Medan, and Montaldo 2005; Gravel et al. 2018).

77 One way to summarise the seemingly contradictory evidence of whether the
78 environment can influence whether two species interact or not is by moving
79 from gradients of temperature or precipitation to gradients of environmental
80 stress. Previous research is so far equivocal and, taken together, suggest that
81 environmental stress can have two contrasting effects on species’ Eltonian-
82 niche (Tylianakis and Morris 2017). On the one hand, it is possible that
83 when species are under environmental stress, they might be “pressured” to
84 focus on partners with which they are best adapted to interact. For instance,
85 Hoiss et al. (2012) found increased phylogenetic clustering between plants
86 and pollinators at higher altitudes; while Peralta et al. (2015) found that
87 parasitoids in plantation forest, where environmental stress was higher than

88 in native forests, were constrained to interact with hosts they were best
89 adapted to attack. Similarly, Lavandero and Tylianakis (2013) found that
90 environmental stress due to higher temperature reduced the breadth of the
91 Eltonian niche of parasitoids.

92 On the other hand, it is also possible that when species are under envi-
93 ronmental stress, they are forced to be more flexible in their interactions
94 as higher environmental stress is likely to be reflected in greater energetic
95 or reproductive costs and therefore they might not be able to sustain en-
96 counter rates with their preferred partners at sufficient levels. In line with
97 this hypothesis, Hoiss, Krauss, and Steffan-Dewenter (2015) found that the
98 specialisation of plant-pollinator networks decreased both with elevation
99 and after extreme drought events. Likewise, Pellissier et al. (2010) found a
100 positive relationship between niche breadth and environmental stress: disk-
101 or bowl-shaped blossoms (which allow a large number of potential pollinator
102 species to access pollen and nectar rewards) dominated at high altitude
103 flower communities.

104 We analyse data on plant-pollinator interactions globally and calculate the
105 environmental suitability of all species present at two or more communities
106 to. Our main aim is to leverage that information to test the two competing
107 hypotheses that relate environmental stress and Eltonian niche breadth
108 and whether there is a pattern within and across species. Specialist species
109 can become “facultative” generalists to reduce their vulnerability to the
110 absence of preferred partners (for example, when variations in climate de-
111 couple phenologies; Benadi et al. 2014). We therefore also expect that
112 as environmental stress increases species with a relatively small number of
113 partners are more likely to broaden their Eltonian niche. Species with a
114 large number of partners, on the other hand, should have a larger pool of
115 available partners and might, therefore, be more likely to narrow their niche
116 under environmental stress. Importantly, when testing these hypotheses, we

control for the potential effects of the environment in community composition and the size of the species fundamental niche, both from an Eltonian (interactions) and Grinnellian (environment) perspective.

Methods

We retrieved plant-pollinator networks from the Web of Life database (Fortuna, Ortega, and Bascompte 2014). This database contains datasets originating from 57 studies published in the primary literature between 1906 and 2016 (Kaiser-Bunbury et al. 2010, 2014; Bartomeus, Vilà, and Santamaría 2008; Arroyo, Primack, and Armesto 1982; Hattersley-Smith 1985; Barrett and Helenurm 1987; Dicks, Corbet, and Pywell 2002; Dupont, Hansen, and Olesen 2003; Elberling and Olesen 1999; Herrera 1988; Hocking 1968; Inouye and Pyke 1988; Kato et al. 1990; Medan et al. 2002; Memmott 1999; Motten 1986; Olesen, Eskildsen, and Venkatasamy 2002; Ollerton 2003; Petanidou and Vokou 1993; McMullen 1964; Mosquin 1967; Clements and Long 1923; Kevan 1970; Small 1976; Montero 2005; Stald, Valido, and Olesen 2003; Ingversen 2006; Bundgaard 2003; Bek 2006; KATO 2000; Dupont and Olesen 2009; Lundgren and Olesen 2005; Percival 1974; Philipp et al. 2006; Primack 1983; Ramirez 1989; Ramirez and Brito 1992; Schemske et al. 1978; Smith-Ramírez et al. 2005; Vázquez 2002; Robertson 1929; Ferreira Canela 2006; Kohler 2011; Lasprilla and others 2003; Sabatino 2010; Lara 2006; Las-Casas, Azevedo Júnior, and Dias Filho 2012; del Coro Arizmendi and Ornelas 1990; Abreu and Vieira 2004; Vizentin-Bugoni et al. 2016; Bezerra, Machado, and Mello 2009; INouE et al. 1990; KATO, Matsumoto, and Kato 1993; Yamazaki and Kato 2003; Gutierrez Z, Rojas-Nossa, and Stiles 2004; Kato and Miura 1993; Kakutani et al. 1990). Calculating the environmental stress of species in their community and their Eltonian niche breadth requires that the scientific names used to identify them are valid and unambiguous. This is because valid scientific names are necessary to

145 both accurately match species across studies and to obtain other information
146 from biological databases. However, a significant proportion of the data
147 obtained contain spelling mistakes, taxonomic synonyms, or non-standard
148 abbreviations. Therefore, our first step was to validate and correct scientific
149 names across studies to ensure that we matching of species across studies
150 was as accurate as possible.

151 Second, we calculated the suitability of the environment for a species in a
152 particular community as a proxy of environmental stress. We assume that the
153 environmental stress a species experience in a particular location is inversely
154 related to the suitability of the average environmental conditions in that
155 place. Our third and final step was to relate the environmental suitability
156 to the relative number of partners a species has in a community, as a proxy
157 for Eltonian niche breadth. To explore this relationship within and across
158 species, we used a multilevel Bayesian model in which we controlled for the
159 potential effects of the environment on co-occurrence.

160 **Plant-pollinator interactions**

161 All together, interaction data from the included studies included 11,231
162 unique organism names. From these 1,166 were present in more than one
163 study. From the total number of organisms, 159 were identified to the
164 subspecies or variety level, 6,759 to the species level, 1,755 to the genus level,
165 whereas the remaining 2,558 were unidentified. As the species level was the
166 most common taxonomic rank available in our interaction datasets, in all
167 further analysis, we grouped together subspecies or varieties within the same
168 species.

169 We were able to confirm the validity of 5,263 of the scientific names used
170 to identify organisms (roughly 76%). We assessed the validity of a name by
171 querying the Global Names Resolver database (<https://resolver.globalnames>).

172 org) which includes data from 98 taxonomic sources. We accessed this
173 database using the function `gnr_resolve` from the R package `taxize` 0.9.6
174 (Chamberlain and Szocs 2013; S. Chamberlain, Szoezs, et al. 2019).

175 From the remaining 1,655 names we were unable to validate, we were able to
176 identify and correct 726 that contained spelling mistakes. These spelling mis-
177 takes were corrected automatically by fuzzy matching the canonical names in
178 our data sources with those in the Global Names Resolver database. However,
179 on rare occasions, the fuzzy matching algorithm can suggest a scientific name
180 that has a similar spelling, but that corresponds to an organism in a different
181 taxonomic group, often a separate kingdom. To address this potential prob-
182 lem, we checked the taxonomic hierarchy of suggested names and confirmed
183 that it matched our expected taxon. We retrieved all taxonomic hierarchies
184 from the National Center for Biotechnology Information taxonomic database
185 (<https://www.ncbi.nlm.nih.gov/taxonomy>).

186 As species names are constantly changing, we subsequently checked for
187 possible synonyms of the canonical names in our data sources. Using data
188 from the Integrated Taxonomic Information System database ([http://www.](http://www.itis.gov)
189 [itis.gov](http://www.itis.gov)), we found synonyms and alternative names for 611 species.
190 Finding these alternative names was required for two main reasons. First,
191 because we wanted to be able to identify the cases in which the same species
192 might have been recorded with different names in various data sources. This
193 can occur not only when the canonical name has been changed but also
194 when there are widely used orthographic variants. Second, because retrieving
195 occurrence data is often only possible using the latest accepted/valid name
196 for a particular species.

197 All together, from the 1,655 names we were unable to validate, it was
198 not possible to automatically correct or find synonyms 332 of them. We
199 then manually consulted multiple online databases, chiefly Wikispecies
200 (<https://species.wikimedia.org/>), and looked for canonical names that both,

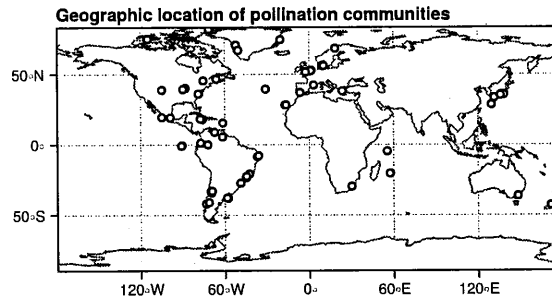


Figure 1: Worldwide distribution of pollination communities included in this study

resembled the unvalidated names and matched the geographic and taxonomic expectations. In this fashion, we were able to further correct 25 names. Most manual corrections were made on names that have been abbreviated or had more than two spelling mistakes. A complete list of manual name corrections can be seen in Table S1.

This cleaning process allowed us to match further 270 names across data sources and, by doing so, identify another 72 species that were present in more than one study. The process also allowed us to identify problematic data sources in which some names were included as both plants and pollinators. These data sources were removed from further analysis. In seven of our data sources interaction data was recorded at multiple points in time. When this was the case, we combined interaction data into one single interaction network.

The cleaning process resulted on a total of 2,555 plants and 8,406 pollinator species distributed across 73 locations around the globe (Figure 1, 2).

} this should stay

Species suitability

Our next step was to determine the habitat suitability of the species as a proxy of the environmental stress they experience in their community. As we aim to compare the trophic niche for different suitability levels, we only did this for species that were present in at least two communities. To calculate

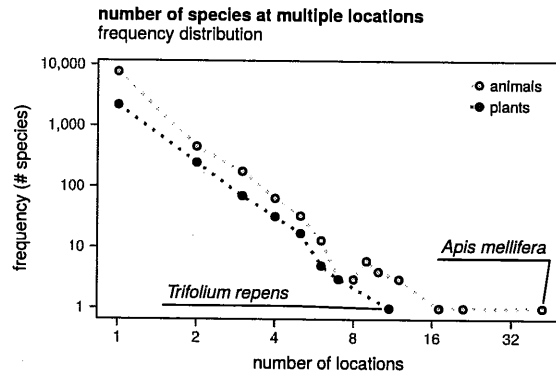


Figure 2: Frequency distribution of the number of locations in which a species is present. The most common pollinator species was *Apis mellifera*, which was sampled on 42 locations, while the most common plant species was *Trifolium repens*, which was sampled on 11 locations.

the suitability of a species in a particular location, we used a (Grinnellian) niche-factor analysis (Hirzel et al. 2002; Broennimann et al. 2012). This approach is based on the probability density function of species distribution in an environmental variable space. In a nutshell, habitats [characterised by a collection of ~~environmental variables~~] in which the species occurs most often are deemed to be suitable for the species than habitats in which the species has never been observed. This approach to estimating the habitat suitability requires two critical pieces of information. First, we require information about the occurrences of the species of interest. Second, we require information about the environmental conditions for all the locations in which the species occurs.

can't this appear later?

We retrieved occurrences from the Global Biodiversity Information Facility (GBIF; <https://www.gbif.org>) using the R package `rgbif` 0.9.6 (Chamberlain and Boettiger 2017; S. Chamberlain, Barve, et al. 2019). Specifically, for each species, we only requested occurrences for which the coordinates of the observation were available and that had no known geospatial issue in the GBIF database. Roughly, we downloaded 38.1 million occurrences for the 986 species we were interested on. These occurrences, however, contain observations of mixed quality. Therefore, we followed (ref?) and applied a

what does this mean?
Could a similar section for this part appear along with the taxonomic name part?

series of simple filters and geographic heuristics to remove those of lower quality. Specifically, we removed all occurrences with (i) a coordinate uncertainty larger than 100km; (ii) those recorded prior to 1945 (ref? or why?); (iii) those in which the number of individuals registered was either zero or more than 100 as such large counts are often erroneous (ref? or why?); and (iv) those occurrences in which the “basis of record” was not a human observation or a preserved specimen (ref? why?). We then used the R package `CoordinateCleaner` 0.9.6 (Zizka et al. 2019) and land mass and country data from Natural Earth (<https://www.naturalearthdata.com>) with a 1:10,000,000 scale to further identify and remove problematic occurrences. We removed occurrences for which their coordinates (v) fell outside the borders of the country where they were recorded; (vi) those around a country capital or the centroid of the country and province centroids; (vii) those around a biodiversity institution; and (viii) those located within oceans. Thorough this cleaning process, we removed with 7.2 million occurrences distributed accross 916 species. Finally, we integrated the occurrences from our plant-pollinator communities to the cleaned occurrences retrieved from GBIF.

We retrieved environmental data from WorldClim V2.0, which includes 19 bioclimatic variables commonly used in species distribution modelling (Fick and Hijmans 2017). We then complemented data obtained from WorldClim with data from Envirem (Title and Bemmels 2017), which includes 16 extra bioclimatic and two topographic variables. The additional set of variables from Envirem are relevant to ecological or physiological processes and as such, have the potential to improve our suitability estimation (Title and Bemmels 2018). We obtained all environmental data as rasters composed by cells of 2.5 arc-minutes. We chose this resolution because it provides a reasonable match to the locational accuracy of the species occurrences found in GBIF, particularly those that come ^{from} preserved specimens in museum

269 collections.

270 After obtaining information about species occurrence and the environment,
271 we then merged these two datasets such that the location of each occur-
272 rence was characterised by a vector with details of our 37 bioclimatic and
273 topographic variables. If a species had more than one occurrence records
274 within one of the cells of the bioclimatic raster, we only included one of
275 the occurrence records. We did this to avoid giving more weight to areas
276 with a high number of occurrences, a common scenario in occurrence records
277 collected opportunistically as the ones we use here. In this step, we removed
278 85.2% of the occurrences which resulted in a total of 4.6 million occurrences
279 used in our niche analysis.

280 A common issue of terrestrial bioclimatic datasets is that the boundaries of
281 the cells with information do not precisely match the landmass boundaries.
282 The result of this mismatch is that not all environmental variables ~~was not~~^{were}
283 available for 3,315 of the raster cells with occurrences (0.8% of the total).
284 As expected, the vast majority of these problematic cells were close to the
285 shore. To address this issue, we calculated the average value of environmental
286 variables within an 5km buffer of the centre of the cell where the variable
287 was missing and used it to approximate the value of the variable in that
288 cell. Using this procedure, we were able to fill environmental variables for
289 89.4% of the cells ~~where~~^{where} ~~was~~ they were missing. To fill ~~the~~^a remaining 350 cells, we
290 repeated the aforementioned procedure but instead ~~used~~^{used} a 10km buffer. We
291 removed from further analysis occurrences located within the 135 cells for
292 which we were unable to fill environmental variables.

293 Next, we calculate the probability density function of the species distribution
294 in environmental space. To determine the environmental space, we use
295 the first two components from a principal component analysis of the 37
296 bioclimatic variables associated with the species occurrences. Specifically
297 we use the `dudi.pca` function from the R package `ade4` 1.7.13 (Dray and

298 Dufour 2007) and center and scale all bioclimatic variables to have a mean
 299 of 0 and a unit standard deviation. We then determine the position of
 300 species occurrences in the environmental space and estimate their bivariate
 301 probability density function. We use a kernel method to estimate this density
 302 and normalise it such that it ranges between zero and one. Specifically, to
 303 calculate the probability density function we use `ecospat.grid.clim.dyn`
 304 from the R package `ecospat` 3.0 (Broennimann, Di Cola, and Guisan 2018)
 305 with a grid resolution of 200. We then determine the location in the environ-
 306 mental space of the plant-pollinator communities using the function `suprow`
 307 from `ade4`. The normalised density at that particular location corresponds
 308 to our suitability metric, which we calculate using the R package `raster`
 309 2.8.19 (Hijmans 2019). We use the kernel density method in the niche-factor
 310 analysis (Broennimann et al. 2012) rather than the distance from the mode
 311 (Hirzel et al. 2002), as it has been proposed earlier, as it has been shown to
 312 reduce the procedure's sensitivity to sampling effort and the resolution of
 313 the environmental space.

314 We used a sensitivity analysis to determine the minimum number of occur-
 315 rences that are necessary to have robust environmental suitability values
 316 in our communities. For that we used the species with most occurrences
 317 available, *Archilochus colubris*, and calculated the mean absolute error of the
 318 suitability values obtained with one thousand subsamples from the 74,791
 319 occurrences available from GBIF.

320 Data analysis

321 We then used a set of Bayesian multilevel models to evaluate the impact of
 322 environmental suitability on the size of the realised Eltonian trophic niche.
 323 We measure the size of the Eltonian niche using the normalised degree of
 324 species; ^{that} ~~this~~ is, the number of species it interacts with given the number
 325 of species in the opposite guild (Martín González, Dalsgaard, and Olesen

huh?
 if it is a probability density
 it is already normalized!

this is very
 wordy and
 it's easy to
 lose track of
 the main point

feels incomplete

2010). The normalised degree was modelled using a logit link function, and a binomial distribution in which the number of species interacts with is the number of successes, and the number of species in the opposite guild is the number of trials. We are aware that ^{whether} species interacts or not is not a Bernoulli process as species interactions are not strictly independent from each other. ^(uh) However, a binomial distribution allows us to account for the differences in species richness across communities indirectly. Importantly, however, results are qualitatively similar when we model species degree directly using a Poisson distribution and a logarithmic link function (Supp info). ← but this isn't really a test of what you state is the key caveat!

We evaluate four models to assess the relative importance of suitability ^{on...} ~~the~~ ^{first model} our baseline model ^{included} five population-level predictors and two grouping levels, species and the community. The population-level predictors in the baseline model, commonly called fixed effects, were the habitat suitability, ~~the~~ species guild (plant or ~~the~~ pollinator), ~~at the level of the~~ ^a species ^{overall} generalism, and its number of known possible partners. We included the overall level of generalism as one ^{also} ~~could~~ expect species with a large number of interactions globally to ^{also} be more likely to interact with a large number of species locally. Concretely, we estimated ~~a~~ ^{overall} species ^{level} of generalism as ^{its} ~~the~~ total ^{richness} ~~number~~ of interacting species across communities in our dataset. We also included the number of known possible partners as a predictor in our models as it allows us to control for the environmental effects on species co-occurrence. We calculate this metric by determining the number of partners with which the species is known to interact in any other community. Controlling for the number of potential partners makes our model a particularly stringent test of our environmental stress hypotheses. Often the potential and the actual number of partners is the same or very close to each other, especially for rare species present only in a few communities.

hard to follow

elaborate on this more clearly as it's a very salient point

355 As we were particularly interested in understanding whether the effect of
356 habitat suitability is conditional on the species' guild (plant or pollinator), we
357 therefore ^{also} included ^{also} guild and its interaction with suitability in the model. We ^{additionally}
358 allowed the intercept ~~of the model~~ and slope of the suitability-degree relationship
359 to vary among species. ^{This} allowed us to investigate two ^{particular} questions. First,
360 it allows us to inspect the extent to which suitability is a population or
361 a group level effect. Second, by investigating the correlation between the
362 intercept and the slope as a model parameter, it allowed us to inspect the
363 extent by which species with a small or large number of interactions respond
364 to increasing levels of environmental stress. To account for unmeasured
365 differences between communities, like sampling effort, sampling method, or
366 diversity, we calculated an intercept for each community in our study. To
367 facilitate model interpretation and convergence, we scaled all continuous
368 variables to have a mean of zero and a unit standard deviation.

369 We compared this baseline model with three alternative models in which
370 we removed ^{fixed?} one predictor at a time. To quantify the difference between
371 models, in terms of their expected out-of-sample performance, we use the
372 Wanatabe-Akaike information criterion (WAIC). All models were fitted under
373 a Bayesian framework using the R package brms 2.8.0 (Bürkner 2017, 2018)
374 as an interface for Stan (Carpenter et al. 2017). For each model, we used
375 four Markov chains of 4,000 iterations each; half of the iterations were used
376 for warmup. We used weakly informative priors for all model parameters.
377 Specifically, we used normal priors of mean zero and standard deviation ten
378 for the population-level effects and the intercepts, a half-Cauchy prior with
379 a location of zero and a scale of two for the standard deviations, and, when
380 applicable, an LKJ-correlation prior with parameter $\zeta = 1$ for the correlation
381 matrix between group-level parameters.

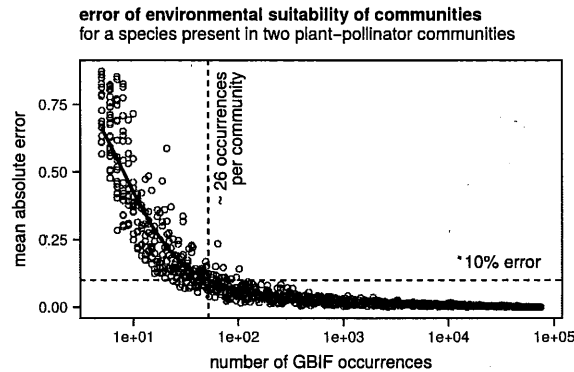


Figure 3: Sensitivity analysis of environmental suitability error. The number of independent occurrences retrieved from GBIF is inversely related to the error of environmental suitability for our plant-pollinator networks. The sensitivity analysis was performed by subsampling occurrences of *Archilochus colubris* the species in our dataset with the largest number of occurrences in GBIF, which was recorded in two of our communities.

Results

After performing our sensitivity analysis, we found that we need roughly 26, 18 independent occurrences for each community for which we calculated a suitability value in order to obtain a mean absolute error below 0.1 (Fig. 3).

We therefore removed from further analyses species for which we did not have enough occurrences to obtain robust estimates. When inspecting the suitability values of the analysed species, we found that some communities have larger overall suitability values. However, most communities included species for which habitat suitability was low and species for which it was high (Fig. 4).

We found that our models performed relatively well. The Bayesian R-squared for our baseline model was 0.91, which indicates our models were able to capture a large proportion of the variability on the data. Overall we found only weak evidence of a relationship between the environmental niche size and the normalised degree (Figure 5a). This relationship was stronger for plants than for animals. Similarly environmental suitability does not show a consistent pattern across species. Indeed, when looking at the population

the implications of this should appear in methods unless species occurrences are detailed here too

can we do marginal and conditional R^2 ?

don't you mean pollinators?

what is this?

0.1 what? units?

don't you mean suitability?

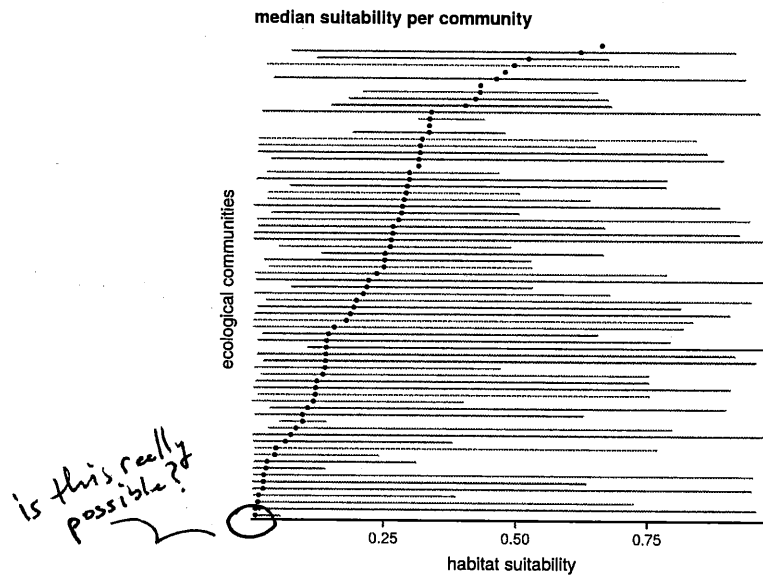


Figure 4: Median habitat suitability of communities in our dataset. Each row represents a different community and horizontal lines represent span the 2.5 and 97.5 quantiles.

level effects, suitability ^{showed} virtually no relationship with normalised degree, neither for plants or pollinators (Figure 5b). However, suitability is still an important predictor as the WAIC difference between our baseline model and ~~that~~ ^{the model} that did not include suitability was 574 ± 101 . Both among plants and pollinators for some species there is a strong negative relationship between suitability and normalised degree and some others with a strong positive relationship (Figure 6a). As expected, we found a positive relationship between ~~the~~ species generalism and normalised degree ^{ed} (5c). However our model comparisons based on WAIC show ^{ed} that including generalism in the model ^{ed} adds relatively little information once we ^d have accounted for the other variables in the model. Finally, we found a strong and positive relationship between the number of possible interactions and the number of realised interactions in the community.

The group level variation among communities was larger than that among species. The standard deviation (in the parameters scale) of the community intercepts was 1.02 [0.85, 1.23] while the standard deviation of the species

(clarify that this is across species what about separately plants and pollinators?)
unclearly written

~ didn't we say this might occlude things?

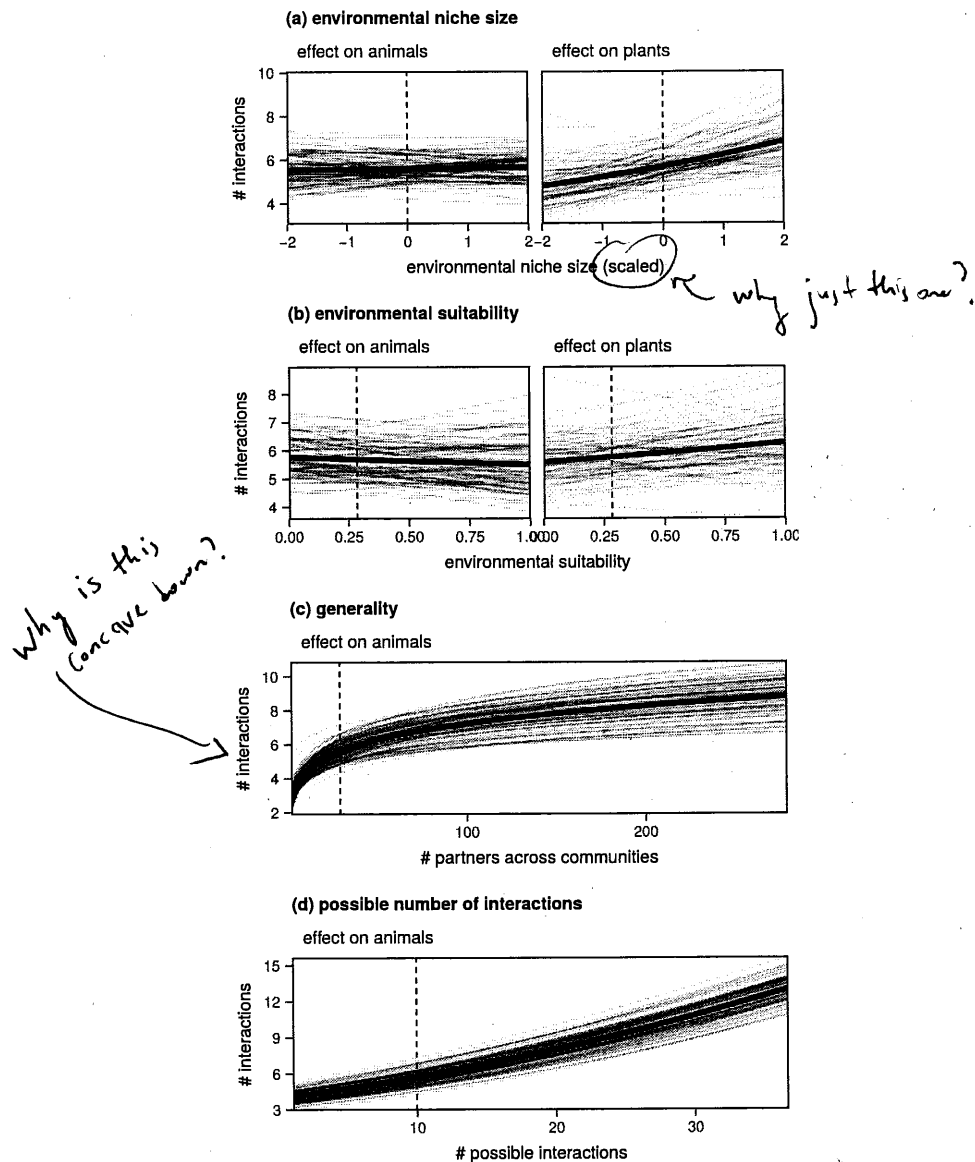


Figure 5: Conditional effects of predictors in our baseline model. The predicted values of the number of interacting species are based on a hypothetical community with 76 plants and 33 pollinators. These values correspond to the median number of species in each guild, respectively. In each panel, we condition on the mean values of all other predictors in the model, mean values for each predictor are indicated with a vertical dashed line. For model fitting, we scaled all predictors to have mean of zero and unit variance. However, with the exception of environmental niche size, here we show the unscaled predictors to facilitate interpretation. To illustrate the uncertainty around the fitted estimates we plot the fits of 100 independent draws from the posterior distribution. The thick lines indicate the mean values of the response distribution. As there was no interaction between guild and generality or the number of possible interactions, for these two predictors we only show conditional the conditional effect of pollinators.

Table 1: Comparison in out of sample predictive power of the baseline model (bold) and their alternatives. We rank models by their expected log predictive density based on their leave-one-out cross-validation information criterion (LOO). The standard error of the LOO difference provides rough guidance to the uncertainty of the model ranking. We also show the Wanatabe-Akaike information criterion (WAIC) of each model for comparison.

predictors	WAIC	SE
Suit. + Gen. + Pot. + Env.	6,167	135
Suit. + Pot. + Env.	6,157	138
Suit. + Gen. + Pot.	6,170	135
Suit. + Gen. + Env.	7,737	281
Gen. + Pot. + Env.	6,740	183
unknown	6,163	136

intercept was 0.5 [0.44, 0.57], and that of the species' suitability slope was 0.34 [0.28, 0.4] (95% credible intervals shown within square brackets). Interestingly, the slope the relationship between suitability and normalised degree was negatively correlated with the species' intercept in the model (Figure 6b). The mean correlation coefficient was -0.51 [-0.69, -0.3].

Discussion

We found that environmental stress, which we measured using habitat suitability, was an important predictor for the number of partners a species interacts with. However, this relationship was highly variable among species. So much so that, at the population level, there is no discernible pattern between stress and the number of partners. Interestingly, within species, there was a significant correlation between the number of partners a species has and the stress-partners relationship itself. Specifically, species with a large number of partners in low-stress communities were more likely to have a negative relationship and hence reduce the number of partners as stress increases. Contrastingly, species in our datasets with a small number of partners in low-stress communities were more likely to have a larger number of partners in more stressful communities.

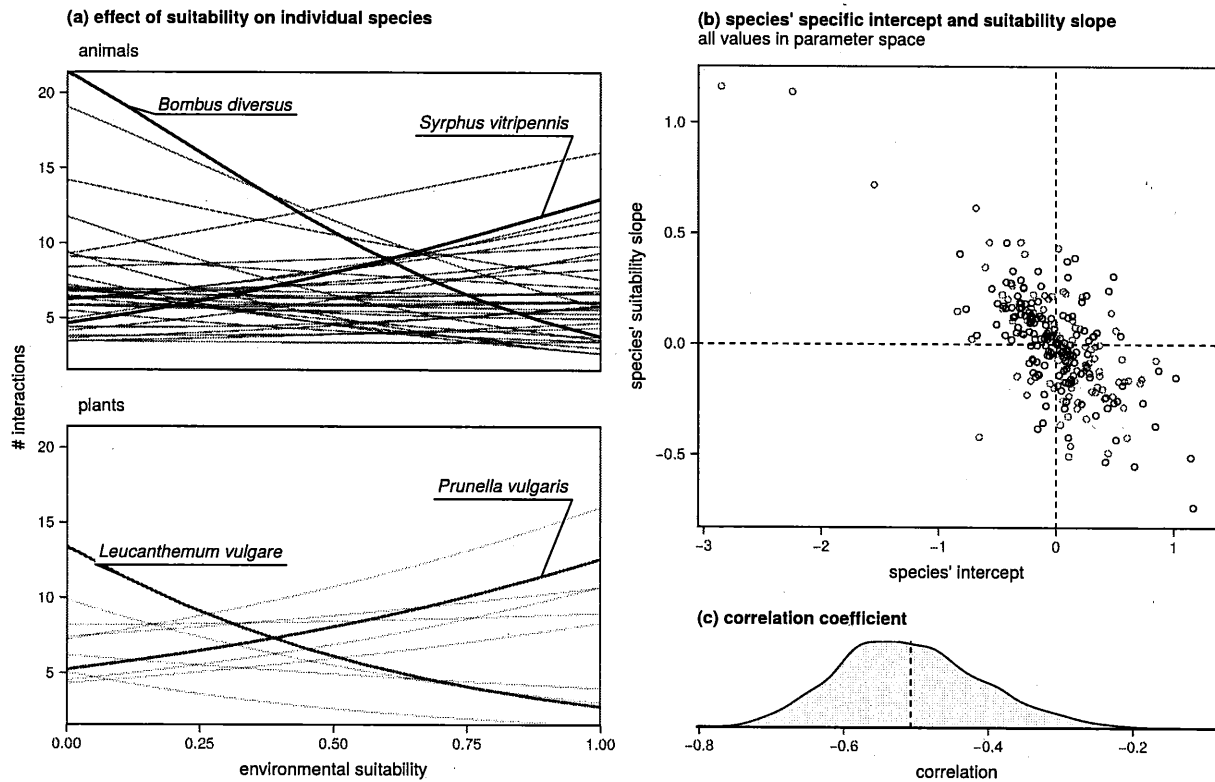


Figure 6: Species level effects of suitability. (a) Conditional effect of suitability for individual species. To facilitate visualization, we show only species for which there is suitability information in at least six communities (10 plants and 33 pollinators). As in the previous figure, fitted values assume a hypothetical community of median size. In each panel, we highlight two species for which the relationship between environmental suitability and normalised degree was particularly strong. (b) The correlation between the species' intercept and the species' slope of suitability was negatively correlated.

((?)).

what does this tell us?

should we replot the mean response?

433 Previous research supported the idea that the differences in network structure
434 along an environmental gradient were due solely to changes to community
435 composition. In other words, it showed that the differences between two
436 hypothetical communities with the same species were stochastic for the most
437 part. Here, by abstracting with multiple environmental gradients into a
438 single gradient of environmental stress, we show that this might not be
439 the case. Instead, the environment can also have a direct impact on the
440 number of partners a species interacts with and, therefore, influence network
441 structure. Our results still show that the primary channel through which
442 environment affects interaction probabilities is determining which species
443 are present or absent from a particular community. However, they also
444 show that, for a large number of species, it can have a substantial effect on
445 their realised Eltonian component of the niche. Indeed, as evidenced by the
446 predictive power of competing models, habitat suitability had a substantial
447 contribution to predicting the number of possible partners in the community.

448 Although we cannot tell with certainty how much of the effect of environmen-
449 tal stress can be attributed to either niche or neutral processes, the patterns
450 we observe are likely more related to the niche. On the one hand, it has been
451 shown that environmental models, such as our niche factor analysis, do a
452 poor job predicting species abundances across sites (Pearce and Ferrier 2001;
453 Sagarin, Gaines, and Gaylord 2006). It would be, therefore, somewhat naive
454 to assume that high habitat suitability translates directly to high species
455 abundances. On the other hand, even if species abundances are correlated
456 with suitability, we show that within a community there is a wide range
457 of suitabilities, even for the limited number of species we included in our
458 analyses. Therefore, we might expect little or no overall effect of neutral
459 processes on our response variable, as the abundance of some interaction
460 partners might increase and decrease for others. Furthermore, the fact that
461 the effects of environmental stress do not depend on whether the species is

462 a plant or a pollinator (for which there can be niche differences based on
463 individual behaviour) emphasises that co-evolutionary history, which in turn
464 shape niche processes driven by, might be driving the patterns we observe.

465 One of our most interesting results was the negative correlation between
466 the species' intercept and the slope of the habitat suitability in our model.
467 Specifically, species with a small intercept, this is, species that were likely
468 to have an above-average number of partners in a community with little
469 environmental stress, were more likely to interact with a smaller number of
470 partners as stress increased. At least for these species, the hypothesis that
471 species focus on partners with which they are best adapted to interact as
472 environmental conditions deteriorate seems to hold weight. Contrastingly,
473 species with a below-average number of partners in a low-stress environment
474 were more likely to interact with a large number of partners as stress increased.

475 At first glance, these species behave as facultative generalists when the
476 environment worsens. This is certainly true for this group of species in
477 our dataset. However, there are some nuances to this observation once we
478 take into account the biases inherent to community-level interaction data.
479 Specifically, we have no information about species for which environmental
480 conditions would, in principle, allow for a positive growth rate but are absent
481 from our datasets. A ramification of this bias is that an undetermined
482 number of species on this second group might also be more likely to focus
483 on partners with higher trait matching as conditions deteriorate. However,
484 we are not able to detect that trend because these species are not recorded
485 in community composition or interaction data. Presumably, because their
486 preferred partners are unavailable and as such, they are absent from the
487 community or too rare to be observed. Note that the correlation between
488 species' intercept and the slope is unlikely to be just a byproduct of our
489 modelling framework, as, in principle, there is nothing preventing species
490 with an above-average number of partners to interact with even more species

491 as environmental stress increases.

492 These observations have implications for understanding the assembly of
493 ecological networks and how it might be impacted by climate change, which
494 currently threatens to increase the environmental stress of species within its
495 range. Previous research suggests that flexibility of interactions in ecological
496 communities might be a rather ubiquitous phenomenon whenever environ-
497 mental conditions change in space or time (CaraDonna et al. 2017). This
498 flexibility has been shown to confer robustness and stability to the loss of
499 interaction partners (Kaiser-Bunbury et al. 2010). The current paradigm of
500 interaction turnover suggests first that trait matching is a relatively unim-
501 portant driver of interaction turnover for generalist species (CaraDonna et
502 al. 2017). The fact that species with a high number of interactions were
503 more likely to interact with fewer species as environmental stress increased
504 indicates that trait matching might still play a role in determining the in-
505 teractions of generalist species. Second, current research suggests that the
506 rewiring of the interaction network is primarily driven by generalist species
507 (Burkle, Marlin, and Knight 2013). We, however, demonstrate that at least
508 a fraction of the species that appear to be specialist in their communities
509 might be more flexible, if not more, than generalist species and, therefore,
510 might have a more significant role in network persistence than previously
511 expected. Contrastingly, specialists that are highly constrained to interact
512 with partners of high trait-matching might be more vulnerable to increased
513 climate change-induced environmental stress and habitat degradation. Taken
514 together, the fact that most species appear to be flexible interaction partners,
515 combined with the fact that those that are not flexible are less likely to per-
516 sist, might explain why specialised communities that are highly constrained
517 by trait matching (like some plant-hummingbird networks; Vizentin-Bugoni,
518 Maruyama, and Sazima 2014; Maruyama et al. 2014) might be relatively
519 rare.

520 In conclusion, we show that...

521 Acknowledgements

522 We thank Warwick Allen and Guadalupe Peralta for feedback and valuable
523 discussions. EFC acknowledges the support from the University of Can-
524 terbury Doctoral Scholarship and a New Zealand International Doctoral
525 Research Scholarship administered by New Zealand Education. DBS and
526 JMT acknowledge the support of Rutherford Discovery Fellowships (RDF-13-
527 UOC-003 and RDF-UOC-1002) and the Marsden Fund Council (UOC-1705),
528 administered by the Royal Society of New Zealand Te Apārangi.

529 References

- 530 Abreu, Carlos R M, and Milene F Vieira. 2004. "Os beija-flores e seus
531 recursos florais em um fragmento florestal de Viçosa, sudeste brasileiro."
532 *Lundiana* 5 (2): 129–34.
- 533 Arroyo, Mary T. Kalin, Richard Primack, and Juan Armesto. 1982. "Com-
534 munity sTUDIES IN POLLINATION ECOLOGY IN THE HIGH TEMPER-
535 ATE ANDES OF CENTRAL CHILE. I. POLLINATION MECHANISMS
536 AND ALTITUDINAL VARIATION." *American Journal of Botany* 69 (1):
537 82–97. <https://doi.org/10.1002/j.1537-2197.1982.tb13237.x>.
- 538 Barrett, Spencer C. H., and Kaius Helenurm. 1987. "The Reproductive Biol-
539 ogy of Boreal Forest Herbs. I. Breeding Systems and Pollination." *Canadian*
540 *Journal of Botany* 65 (10): 2036–46. <https://doi.org/10.1139/b87-278>.
- 541 Bartomeus, Ignasi, Montserrat Vilà, and LuíS Santamaría. 2008. "Contrast-
542 ing Effects of Invasive Plants in Plant-Pollinator Networks." *Oecologia* 155
543 (4): 761–70. <https://doi.org/10.1007/s00442-007-0946-1>.