

Template to prepare preprints and manuscripts using markdown and github actions

Timothée Poisot^{1,2,‡} Peregrin Took^{3,4} Merriadoc Brandybuck^{5,4,‡}

¹ Université de Montréal ² Québec Centre for Biodiversity Sciences ³ Inn of the Prancing Pony

⁴ Fellowship of the Ring ⁵ Green Dragon Inn

‡ These authors contributed equally to the work

Correspondance to:

Timothée Poisot — timothee.poisot@umontreal.ca

This work is released by its authors under a CC-BY 4.0 license



Last revision: *January 12, 2022*

Purpose: This template provides a series of scripts to render a markdown document into an interactive website and a series of PDFs.

Motivation: It makes collaborating on text with GitHub easier, and means that we never need to think about the output.

Internals: GitHub actions and a series of python scripts. The markdown is handled with pandoc.

1 Intro

2 Having a general solution for inferring *potential* interactions (despite the unavailability of interaction data)
3 could be the catalyst for significant breakthroughs in our ability to start thinking about species interaction
4 networks over large spatial scales. In a recent overview of the field of ecological network prediction,
5 Strydom *et al.* (2021) identified two challenges of interest to the prediction of interactions at large scales.
6 First, there is a relative scarcity of relevant data in most places globally – paradoxically, this restricts our
7 ability to infer interactions to locations where inference is perhaps the least required; second, accurate
8 predictions often demand accurate predictors, and the lack of methods that can leverage small amount of
9 data is a serious impediment to our predictive ability globally.

10 Following the definition of Dunne (2006), a metaweb is a network analogue to the regional species pool;
11 specifically, it is an inventory of all *potential* interactions between species from a spatially delimited area
12 (and so captures the γ diversity of interactions). The metaweb is, therefore, *not* a prediction of the food
13 web at a specific locale within the spatial area it covers, and will have a different structure (notably by
14 having a larger connectance; see *e.g.* Wood *et al.* 2015). These local food webs (which captures the α
15 diversity of interactions) are a subset of the metaweb's species and interactions, and have been called
16 “metaweb realizations” (Poisot *et al.* 2015). Differences between local food web and their metaweb are due
17 to chance, species abundance and co-occurrence, local environmental conditions, and local distribution of
18 functional traits, among others.

19 Because the metaweb represents the joint effect of functional, phylogenetic, and macroecological
20 processes (Morales-Castilla *et al.* 2015), it holds valuable ecological information. Specifically, it is the
21 “upper bounds” on what the composition of the local networks can be (see *e.g.* McLeod *et al.* 2021). These
22 local networks, in turn, can be reconstructed given appropriate knowledge of local species composition,
23 providing information on structure of food webs at finer spatial scales. This has been done for example for
24 tree-galler-parasitoid systems (Gravel *et al.* 2018), fish trophic interactions (Albouy *et al.* 2019), tetrapod
25 trophic interactions (O'Connor *et al.* 2020), and crop-pest networks (Grünig *et al.* 2020). Whereas the
26 original metaweb definition, and indeed most past uses of metawebs, was based on the presence/absence
27 of interactions, we focus on *probabilistic* metawebs where interactions are represented as the chance of
28 success of a Bernoulli trial (see *e.g.* Poisot *et al.* 2016); therefore, not only does our method recommend
29 interactions that may exist, it gives each interaction a score, allowing us to properly weigh them.

Yet, owing to the inherent plasticity of interactions, there have been documented instances of food webs undergoing rapid collapse/recovery cycles over short periods of time (Pedersen *et al.* 2017). The embedding of a network, in a sense, embeds its macro-evolutionary history, especially as RDPG captures ecological signal (Dalla Riva & Stouffer 2016); at this point, it is important to recall that a metaweb is intended as a catalogue of all potential interactions, which should then be filtered (Morales-Castilla *et al.* 2015). In practice (and in this instance) the reconstructed metaweb will predict interactions that are plausible based on the species' evolutionary history, however some interactions would/would not be realized due to human impact.

Dallas *et al.* (2017) suggested that most links in ecological networks may be cryptic, *i.e.* uncommon or otherwise hard to observe. This argument essentially echoes Jordano (2016): the sampling of ecological interactions is difficult because it requires first the joint observation of two species, and then the observation of their interaction. In addition, it is generally expected that weak or rare links would be more common in networks (Csermely 2004), compared to strong, persistent links; this is notably the case in food chains, wherein many weaker links are key to the stability of a system (Neutel *et al.* 2002). In the light of these observations, the results in fig. ?? are not particularly surprising: we expect to see a surge in these low-probability interactions under a model that has a good predictive accuracy. Because the predictions we generate are by design probabilistic, then one can weigh these rare links appropriately. In a sense, that most ecological interactions are elusive can call for a slightly different approach to sampling: once the common interactions are documented, the effort required in documenting each rare interaction may increase exponentially. Recent proposals suggest that machine learning algorithms, in these situations, can act as data generators (Hoffmann *et al.* 2019): in this perspective, high quality observational data can be supplemented with synthetic data coming from predictive models, which increases the volume of information available for inference. Indeed, Strydom *et al.* (2021) suggested that knowing the metaweb may render the prediction of local networks easier, because it fixes an “upper bound” on which interactions can exist; indeed, with a probabilistic metaweb, we can consider that the metaweb represents an aggregation of informative priors on the interactions.

As Herbert (1965) rightfully pointed out, “[y]ou can’t draw neat lines around planet-wide problems”; in this regard, our approach (and indeed, any inference of a metaweb at large scales) must contend with several interesting and interwoven families of problems. The first is the limit of the metaweb to embed and transfer. If the initial metaweb is too narrow in scope, notably from a taxonomic point of view, the

chances of finding another area with enough related species to make a reliable inference decreases; this would likely be indicated by large confidence intervals during ancestral character estimation, but the lack of well documented metawebs is currently preventing the development of more concrete guidelines. The question of phylogenetic relatedness and dispersal is notably true if the metaweb is assembled in an area with mostly endemic species, and as with every predictive algorithm, there is room for the application of our best ecological judgement. Conversely, the metaweb should be reliably filled, which assumes that the S^2 interactions in a pool of S species have been examined, either through literature surveys or expert elicitation. Supp. Mat. 1 provides some guidance as to the type of sampling effort that should be prioritized. While RDPG was able to maintain very high predictive power when interactions were missing, the addition of false positive interactions was immediately detected; this suggests that it may be appropriate to err on the side of “too many” interactions when constructing the initial metaweb to be transferred. The second series of problems are related to determining which area should be used to infer the new metaweb in, as this determines the species pool that must be used.

In our application, we focused on the mammals of Canada. The upside of this approach is that information at the country level is likely to be required by policy makers and stakeholders for their biodiversity assessment, as each country tends to set goals at the national level (Buxton *et al.* 2021) for which quantitative instruments are designed (Turak *et al.* 2017), with specific strategies often enacted at smaller scales (Ray *et al.* 2021). And yet, we do not really have a satisfying answer to the question of “where does a food web stop?”; the current most satisfying solutions involve examining the spatial consistency of network area relationships (see *e.g.* Galiana *et al.* 2018, 2019, 2021; Fortin *et al.* 2021), which is of course impossible in the absence of enough information about the network itself. This suggests that an *a posteriori* refinement of the results may be required, based on a downscaling of the metaweb. The final family of problems relates less to the availability of data or quantitative tools, and more to the praxis of spatial ecology. Operating under the context of national divisions, in large parts of the world, reflects nothing more than the legacy of settler colonialism. Indeed, the use of ecological data is not an apolitical act (Nost & Goldstein 2021), as data infrastructures tend to be designed to answer questions within national boundaries, and their use both draws upon and reinforces territorial statecraft; as per Machen & Nost (2021), this is particularly true when the output of “algorithmic thinking” (*e.g.* relying on machine learning to generate knowledge) can be re-used for governance (*e.g.* enacting conservation decisions at the national scale). We therefore recognize that methods such as we propose

90 operate under the framework that contributed to the ongoing biodiversity crisis (Adam 2014), reinforced
91 environmental injustice (Choudry 2013; Domínguez & Luoma 2020), and on Turtle Island especially,
92 should be replaced by Indigenous principles of land management (Eichhorn *et al.* 2019; No'kmaq *et al.*
93 2021). As we see AI/ML being increasingly mobilized to generate knowledge that is lacking for
94 conservation decisions (*e.g.* Lamba *et al.* 2019; Mosebo Fernandes *et al.* 2020), our discussion of these tools
95 need to go beyond the technical, and into the governance consequences they can have.

96 **Acknowledgements:** We acknowledge that this study was conducted on land within the traditional
97 unceded territory of the Saint Lawrence Iroquoian, Anishinabewaki, Mohawk, Huron-Wendat, and
98 Omàmiwininiwak nations. TP, TS, DC, and LP received funding from the Canadian Institute for Ecology &
99 Evolution. FB is funded by the Institute for Data Valorization (IVADO). TS, SB, and TP are funded by a
100 donation from the Courtois Foundation. CB was awarded a Mitacs Elevate Fellowship no. IT12391, in
101 partnership with fRI Research, and also acknowledges funding from Alberta Innovates and the Forest
102 Resources Improvement Association of Alberta. M-JF acknowledges funding from NSERC Discovery
103 Grant and NSERC CRC. RR is funded by New Zealand's Biological Heritage Ngā Koiora Tuku Iho
104 National Science Challenge, administered by New Zealand Ministry of Business, Innovation, and
105 Employment. BM is funded by the NSERC Alexander Graham Bell Canada Graduate Scholarship and the
106 FRQNT master's scholarship. LP acknowledges funding from NSERC Discovery Grant (NSERC
107 RGPIN-2019-05771). TP acknowledges financial support from NSERC through the Discovery Grants and
108 Discovery Accelerator Supplement programs.

109 Adam, R. (2014). *Elephant treaties: The Colonial legacy of the biodiversity crisis*. UPNE.

110 Albouy, C., Archambault, P., Appeltans, W., Araújo, M.B., Beauchesne, D., Cazelles, K., *et al.* (2019). The
111 marine fish food web is globally connected. *Nature Ecology & Evolution*, 3, 1153–1161.

112 Buxton, R.T., Bennett, J.R., Reid, A.J., Shulman, C., Cooke, S.J., Francis, C.M., *et al.* (2021). Key
113 information needs to move from knowledge to action for biodiversity conservation in Canada.
114 *Biological Conservation*, 256, 108983.

115 Choudry, A. (2013). Saving biodiversity, for whom and for what? Conservation NGOs, complicity,
116 colonialism and conquest in an era of capitalist globalization. In: *NGOization: Complicity,*
117 *contradictions and prospects*. Bloomsbury Publishing, pp. 24–44.

118 Csermely, P. (2004). Strong links are important, but weak links stabilize them. *Trends in Biochemical*

119 *Sciences*, 29, 331–334.

120 Dalla Riva, G.V. & Stouffer, D.B. (2016). Exploring the evolutionary signature of food webs' backbones
 121 using functional traits. *Oikos*, 125, 446–456.

122 Dallas, T., Park, A.W. & Drake, J.M. (2017). Predicting cryptic links in host-parasite networks. *PLOS*
 123 *Computational Biology*, 13, e1005557.

124 Domínguez, L. & Luoma, C. (2020). Decolonising Conservation Policy: How Colonial Land and
 125 Conservation Ideologies Persist and Perpetuate Indigenous Injustices at the Expense of the
 126 Environment. *Land*, 9, 65.

127 Dunne, J.A. (2006). The Network Structure of Food Webs. In: *Ecological networks: Linking structure and*
 128 *dynamics* (eds. Dunne, J.A. & Pascual, M.). Oxford University Press, pp. 27–86.

129 Eichhorn, M.P., Baker, K. & Griffiths, M. (2019). Steps towards decolonising biogeography. *Frontiers of*
 130 *Biogeography*, 12, 1–7.

131 Fortin, M.-J., Dale, M.R.T. & Brimacombe, C. (2021). Network ecology in dynamic landscapes. *Proceedings*
 132 *of the Royal Society B: Biological Sciences*, 288, rspb.2020.1889, 20201889.

133 Galiana, N., Barros, C., Braga, J., Ficetola, G.F., Maiorano, L., Thuiller, W., *et al.* (2021). The spatial scaling
 134 of food web structure across European biogeographical regions. *Ecography*, n/a.

135 Galiana, N., Hawkins, B.A. & Montoya, J.M. (2019). The geographical variation of network structure is
 136 scale dependent: Understanding the biotic specialization of hostparasitoid networks. *Ecography*, 42,
 137 1175–1187.

138 Galiana, N., Lurgi, M., Claramunt-López, B., Fortin, M.-J., Leroux, S., Cazelles, K., *et al.* (2018). The
 139 spatial scaling of species interaction networks. *Nature Ecology & Evolution*, 2, 782–790.

140 Gravel, D., Baiser, B., Dunne, J.A., Kopelke, J.-P., Martinez, N.D., Nyman, T., *et al.* (2018). Bringing Elton
 141 and Grinnell together: A quantitative framework to represent the biogeography of ecological
 142 interaction networks. *Ecography*, 0.

143 Grünig, M., Mazzi, D., Calanca, P., Karger, D.N. & Pellissier, L. (2020). Crop and forest pest metawebs shift
 144 towards increased linkage and suitability overlap under climate change. *Communications Biology*, 3,
 145 1–10.

146 Herbert, F. (1965). *Dune*. 1st edn. Chilton Book Company, Philadelphia.

147 Hoffmann, J., Bar-Sinai, Y., Lee, L.M., Andrejevic, J., Mishra, S., Rubinstein, S.M., *et al.* (2019). Machine
 148 learning in a data-limited regime: Augmenting experiments with synthetic data uncovers order in
 149 crumpled sheets. *Science Advances*, 5, eaau6792.

150 Jordano, P. (2016). Sampling networks of ecological interactions. *Functional Ecology*, 30, 1883–1893.

151 Lamba, A., Cassey, P., Segaran, R.R. & Koh, L.P. (2019). Deep learning for environmental conservation.
 152 *Current Biology*, 29, R977–R982.

153 Machen, R. & Nost, E. (2021). Thinking algorithmically: The making of hegemonic knowledge in climate
 154 governance. *Transactions of the Institute of British Geographers*, 46, 555–569.

155 McLeod, A., Leroux, S.J., Gravel, D., Chu, C., Cirtwill, A.R., Fortin, M.-J., *et al.* (2021). Sampling and
 156 asymptotic network properties of spatial multi-trophic networks. *Oikos*, n/a.

157 Morales-Castilla, I., Matias, M.G., Gravel, D. & Araújo, M.B. (2015). Inferring biotic interactions from
 158 proxies. *Trends in Ecology & Evolution*, 30, 347–356.

159 Mosebo Fernandes, A.C., Quintero Gonzalez, R., Lenihan-Clarke, M.A., Leslie Trotter, E.F. & Jokar
 160 Arsanjani, J. (2020). Machine Learning for Conservation Planning in a Changing Climate.
 161 *Sustainability*, 12, 7657.

162 Neutel, A.-M., Heesterbeek, J.A.P. & de Ruiter, P.C. (2002). Stability in Real Food Webs: Weak Links in
 163 Long Loops. *Science*, 296, 1120–1123.

164 No'kmaq, M., Marshall, A., Beazley, K.F., Hum, J., Joudry, Shalan, Papadopoulos, A., *et al.* (2021).
 165 “Awakening the sleeping giant”: Re-Indigenization principles for transforming biodiversity
 166 conservation in Canada and beyond. *FACETS*, 6, 839–869.

167 Nost, E. & Goldstein, J.E. (2021). A political ecology of data. *Environment and Planning E: Nature and*
 168 *Space*, 25148486211043503.

169 O'Connor, L.M.J., Pollock, L.J., Braga, J., Ficetola, G.F., Maiorano, L., Martinez-Almoyna, C., *et al.* (2020).
 170 Unveiling the food webs of tetrapods across Europe through the prism of the Eltonian niche. *Journal of*
 171 *Biogeography*, 47, 181–192.

172 Pedersen, E.J., Thompson, P.L., Ball, R.A., Fortin, M.-J., Gouhier, T.C., Link, H., *et al.* (2017). Signatures of
 173 the collapse and incipient recovery of an overexploited marine ecosystem. *Royal Society Open Science*,
 174 4, 170215.

- 175 PoisoT, T., Cirtwill, A.R., Cazelles, K., Gravel, D., Fortin, M.-J. & Stouffer, D.B. (2016). The structure of
176 probabilistic networks. *Methods in Ecology and Evolution*, 7, 303–312.
- 177 PoisoT, T., Stouffer, D.B. & Gravel, D. (2015). Beyond species: Why ecological interaction networks vary
178 through space and time. *Oikos*, 124, 243–251.
- 179 Ray, J.C., Grimm, J. & Olive, A. (2021). The biodiversity crisis in Canada: Failures and challenges of
180 federal and sub-national strategic and legal frameworks. *FACETS*, 6, 1044–1068.
- 181 Strydom, T., Catchen, M.D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., *et al.* (2021). A
182 roadmap towards predicting species interaction networks (across space and time). *Philosophical*
183 *Transactions of the Royal Society B: Biological Sciences*, 376, 20210063.
- 184 Turak, E., Brazill-Boast, J., Cooney, T., Drielsma, M., DelaCruz, J., Dunkerley, G., *et al.* (2017). Using the
185 essential biodiversity variables framework to measure biodiversity change at national scale. *Biological*
186 *Conservation*, SI:Measures of biodiversity, 213, 264–271.
- 187 Wood, S.A., Russell, R., Hanson, D., Williams, R.J. & Dunne, J.A. (2015). Effects of spatial scale of
188 sampling on food web structure. *Ecology and Evolution*, 5, 3769–3782.