



Measuring biodiversity is a critical component of addressing the drivers of Scottish, and wider, biodiversity loss. Simple measures based on surveys of the number of species within a landscape, individual species abundances, and spatio-temporal distributions are common but each only represents a facet of the biodiversity of a landscape. We carried out a literature review on Joint Species Distribution models, which are a powerful and relatively new addition to the biodiversity modelling toolbox (Pollock et al. 2014) with the potential to bring together different facets of the biodiversity picture.

Background

Joint Species Distribution models (JSDMs)

Regression-based models that predict the distribution of **multiple species simultaneously** based on:

- relationship between the presence/absence or abundance of the species
- environmental predictor variables
- interactions between the species.

Typical models are Generalised Linear Models (GLMs) with species interaction terms, and multivariate (probit) regression models with latent variables.

JSDM **limitations**: dispersion, spatial scales, dataset biases and computational cost for correlation matrices & species interactions.

More than 600 articles have been published since 2010 on the use of JSDMs in ecology, environmental science and biodiversity research.

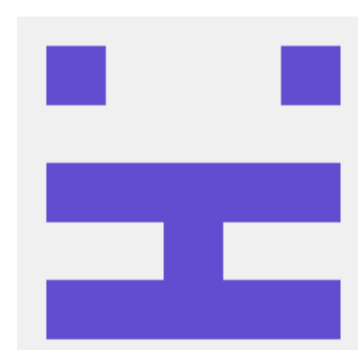
(Google Scholar; advanced search, article title to include the exact phrase "joint species distribution models", from 2010-2023)

There exist numerous R packages for JSDM, we reviewed 12 of them.

JSDM R packages

Bayesian approaches:

HMSC uses a Bayesian multivariate generalised linear model that can include species-specific trait and spatial distribution information, and phylogenetic relationships between the species.



HMSC R Package (Tikhonov et al. 2020)



jSDM uses a hierarchical Bayesian model that assumes a multivariate normal distribution for the joint species occurrence probability. However, it does not explicitly handle phylogenetic information.

jSDM R Package (Warton et al. 2015)

boral (Bayesian Ordination and Regression Analysis) uses latent variables defined through a Bayesian factor analysis with linear predictors for the latent variables. BORAL can be run both with and without environmental covariates.

Others: MistNet, gjam, BayesComm, gllvm, phyr, ppjSDM

Biases in JSDM Datasets

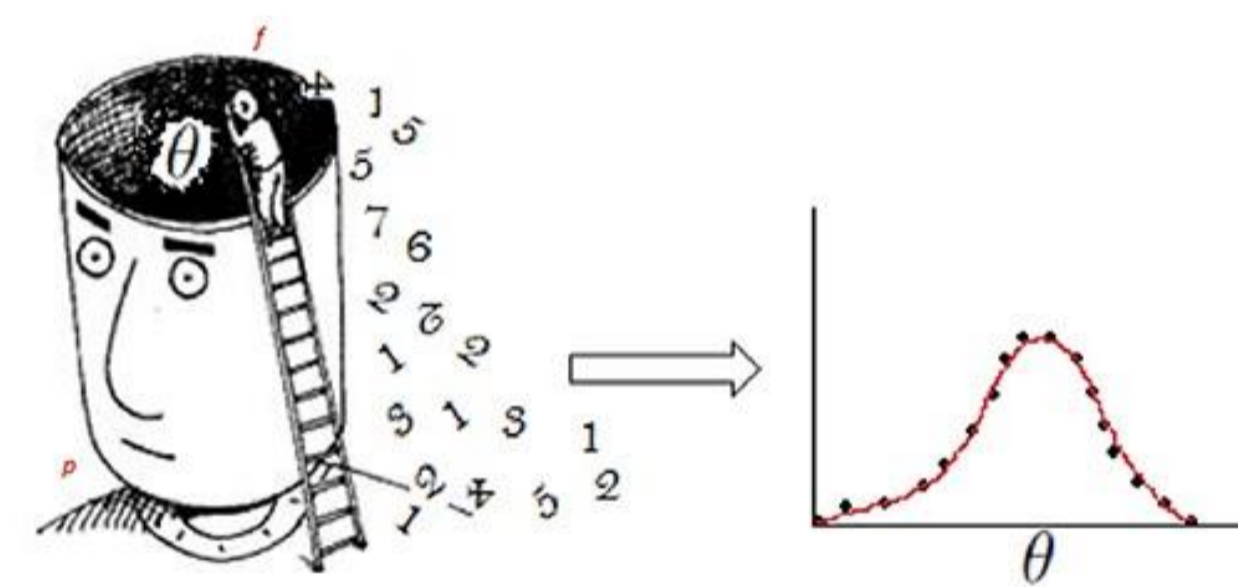
We explored several large-scale ecological datasets (scientific survey and citizen science ecology data) that may be useful to test JSDM implementation in years 2-5 of this project.

Particular interest in eDNA microbiome sequence data (National Soil Inventory Scotland) with **tens of thousands of species**.

Biases in large-scale species occurrence datasets fall into three broad categories; **taxonomic** biases (which species are recorded), **spatial** biases (where the species are recorded), and **temporal** biases (when the species was recorded).

Taxonomic biases in JSDMs may be addressed using rarefaction curves (Sanders, 1968). Some JSDM frameworks explicitly account for spatial biases using auto-correlations. Temporal biases do not strongly impact JSDM results but may introduce confounding variables.

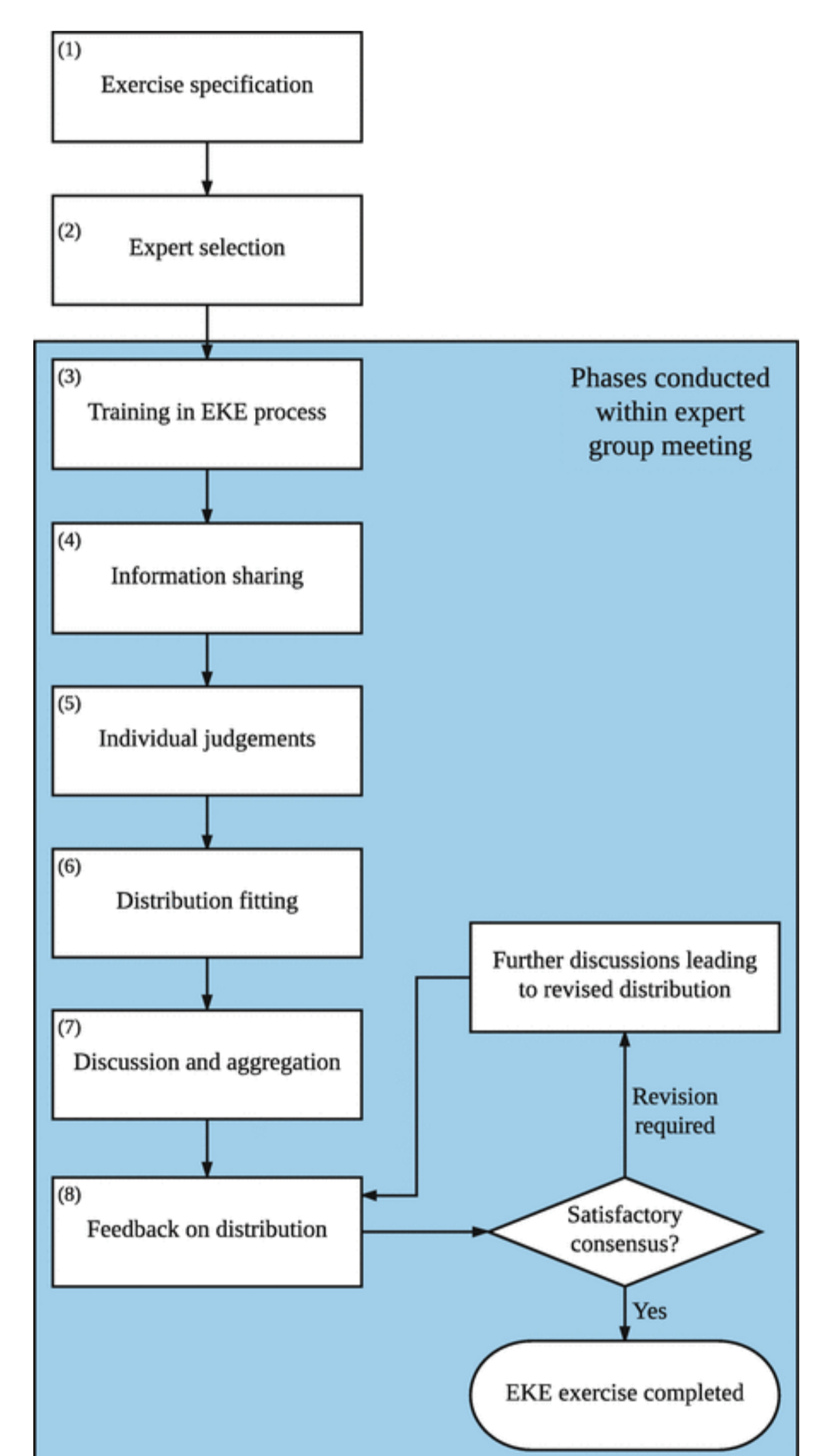
Expert Knowledge



Elicitation of expert opinion
 Source: Barrera-Causil et al. (2019)

Expert elicitation (EE) is the process of translating someone's knowledge about an uncertain quantity into a prior probability distribution in a Bayesian analysis.

EE can be used to inform the structure or size of species interactions, or to aid with informing algorithm configuration for the modelling of species distributions.



SHELF: The Sheffield Elicitation Framework
 O'Hagan et al (2006)

Future Directions

- Computation framework for simulating disparate biodiversity data sets at a range of scales, including different combinations of sampling.
- Computational framework for running a panel of existing JSDM models.
- Contrast performance of tools with different data biases and at different scales, identifying current best-practice approaches.
- Identify and test methods for mitigating the impacts of biases.

References:

Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M., Veski, P.A. and McCarthy, M.A., 2014. Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). *Methods in Ecology and Evolution*, 5(5), pp.397-406.

Tikhonov, G., Opedal, Ø.H., Abrego, N., Lehtikoinen, A., de Jonge, M.M., Oksanen, J. and Ovaskainen, O., 2020. Joint species distribution modelling with the r-package Hmsc. *Methods in ecology and evolution*, 11(3), pp.442-447.

Warton, D.I., Blanchet, F.G., O'Hara, R.B., Ovaskainen, O., Taskinen, S., Walker, S.C. & Hui, F.K.C. (2015) [So many variables: Joint modeling in community ecology](#). *Trends in Ecology & Evolution*, 30, 766-779.

Sanders, H. L. Marine Benthic Diversity: A Comparative Study. *The American Naturalist* 102, 925 (May 1968), 243-282. Publisher: The University of Chicago Press.

Barrera-Causil, C. J., Correa, J. C., & Marmolejo-Ramos, F. (2019). Experimental investigation on the elicitation of subjective distributions. *Frontiers in psychology*, 10, 862.

O'Hagan, A., Buck, C.E., Daneshkhah, A., Eiser, J.R., Garthwaite, P.H., Jenkinson, D.J., Oakley, J.E. and Rakow, T., 2006. Uncertain judgements: eliciting experts' probabilities.