Supplement 1: Overview of published papers that used the data published in this article, or a derivative of it

Topic	**Overview of content**	**Reference**
Calibration	POC (presence-only calibration) plots – a measure for exploring calibration of presence-only model predictions.	Phillips, S.J., Elith, J., 2010. POC-plots: calibrating species distribution models with presence-only data. Ecology 91, 2476–2484.
Error in location	Evaluated how spatial error in occurrence records influenced performance of 10 SDM techniques 40 species and 4 regions. While spatial error reduced model performance in 3 regions two SDM techniques, boosted regression trees and maximum entropy were most robust to spatial error.	Graham, C.H., Elith, J., Hijmans, R.J., Guisan, A., Peterson, A.T., Loiselle, B.A., 2008. The influence of spatial errors in species occurrence data used in distribution models. Journal of Applied Ecology 45, 239–247.
Grain size	Determined how a 10-fold change in environmental grain size influenced model performance based on 10 SDM techniques, 50 species and 5 regions. There was a weak trend towards poorer model performance as grain size increased, especially toward greater sample sizes, though the result was not consistent across all species and regions. Sessile plant species in datasets of greatest locational accuracy were the most affected, but the ranking of modelling techniques was not affected by grain change.	Guisan, A., Graham, C.H., Elith, J., Huettmann, F., NCEAS Species Distribution Modelling Group, ., 2007. Sensitivity of predictive species distribution models to change in grain size: insights from a multi-models experiment across five continents. Diversity and Distributions 13, 332–340.
Imbalance in presence vs background sample sizes	For 226 species across 6 regions and making use of the PA dataset to evaluate calibration, this paper applies robust Bayesian decision theory to develop weighting approaches (of P vs BG data) that improve log loss on test data.	Dudík, M., Phillips, S.J., 2009. Generative and Discriminative Learning with Unknown Labeling Bias, in: Koller, D., Schuurmans, D., Bengio, Y., Bottou, L. (Eds.), Advances in Neural Information Processing Systems 21. Curran Associates, Inc., pp. 401–408.
Maxent	Evaluated and further developed the species distribution modeling Maxent. Using the 226 species and 6 regions from the NCEAS dataset several recommendations and improvements were introduced including: modeling of complex relationships using "hinge" features, providing logistic output, showing that default parameter settings generally performed as well as highly tuned models.	Phillips, S.J., Dudík, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31, 161–175.
Maxent open source release	Open source release of Maxent, and release of a new R package to implement Maxent as an infinitely weighted logistic regression	Phillips, S.J., Anderson, R.P., Dudík, M., Schapire, R.E., Blair, M.E., 2017. Opening the black box: an open-source release of Maxent. Ecography 40, 887–893.
Model comparison	Applied 16 modelling approaches to the data supplied here, analysing predictive performance at evaluation sites using three statistical measures. Results analysed in a generalised linear mixed model (GLMM), exploring effects of method, region and species on performance. Distance between sites in modelling and evaluation data analysed for both geographic and environmental space.	Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.McC., Peterson, A.T., Phillips, S.J., Richardson, K.S., Scachetti-Pereira, R., Schapire, R.E., Soberón, J., Williams, S., Wisz, M.S., Zimmermann, N.E., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29, 129–151.

Topic	Overview of content	Reference
Multivariate adaptive regression splines	Compared generalised additive models (GAMs) with multivariate adaptive regression splines (MARS) fitted to either individual species, or using multi-response models which use information across species to determine modelled responses. Tested two approaches to background selection. Results (AUC and predictive deviance) analysed in a generalised linear mixed model (GLMM), exploring effects of method, region and species on performance. Provided a tutorial for MARS.	Elith, J., Leathwick, J.R., 2007. Predicting species distributions from museum and herbarium records using multiresponse models fitted with multivariate adaptive regression splines. Diversity and Distributions 13, 165–175.
Null models	Provided appropriate null models to deal with spatial sorting bias and calibration	Hijmans, R.J., 2012. Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. Ecology 93, 679–688
Sample bias	A comprehensive background to problems with bias in presence-only samples, with details of how to deal with it. By replacing a random background sample with occurrence locations from a set of species collected in the same region (i.e., a target group) both occurrence and target-group background data have similar biases. The target background approach increased model performance, especially when there is a strong bias in the data.	Phillips, S.J., Dudík, M., Elith, J., Graham, C.H., Lehmann, A., Leathwick, J., Ferrier, S., 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecological Applications 19, 181–197.
Sample bias	Tested new methods for dealing with biased presence-only records, using Maxent. Evaluated methods on simulated data and also the NCEAS data. This includes the theoretical development behind the target-group background method of reference 9.	Dudík, M., Phillips, S.J., Schapire, R.E., 2006. Correcting sample selection bias in maximum entropy density estimation, in: Weiss, Y., Schölkopf, B., Platt, J.C. (Eds.), Advances in Neural Information Processing Systems 18. MIT Press, pp. 323–330.
Sample size	Assessed the effect of three sample sizes (100, 30, and 10 records per species) for 46 species across 12 modelling techniques in six regions. Model performance decreased with decreasing sample size, together with increased variability across species and between models. MaxEnt was the modelling technique performing best across sample sizes, but no technique performed constantly well at the smallest one. Most advanced techniques accounting for predictor interactions and allowing complex responses (e.g. boosted regression trees) performed better at large sample sizes but not at the smallest ones.	Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., Guisan, A., NCEAS Predicting Species Distributions Working Group, 2008. Effects of sample size on the performance of species distribution models. Diversity and Distributions 14, 763–773.
Trees in Switzerland	Determined how various factors of the study design – modeling techniques, location error, grain, and sample size - affected predictions of 30 tree species in Switzerland, and if species traits could explain differences in model performance. Results revealed a larger variance across species than across modelling techniques, with a same ranking of modelling techniques as in reference 2. Location error and sample size tended to reduce model performance, whereas grain had little or no effect on average. Slow-growing late-successional species proved easier to predict than fast-growing early-successional species.	Guisan, A., Zimmermann, N.E., Elith, J., Graham, C.H., Phillips, S., Peterson, A.T., 2007. What matters for predicting the occurences of trees: techniques, data, or species' characteristics? Ecological Monographs 77, 615–530.