Appendix 1. Stability selection

Methods

We applied stability selection (Meinshausen and Bühlmann 2010, Shah and Samworth 2013) to identify base-learners, and thus covariates, that were commonly selected in the majority of ~~100~~ random***ly drawn*** subsamples ***of size*** $\lfloor n/2 \rfloor$ of the data. ***As proposed by Shah and Samworth (2013), we used*** $B = 50$ ***complementary pairs subsamples, i.e., we randomly split the data into two halves and used both to independently fit the model. This led to 100 subsamples all together.*** We set the number of selected base-learners per boosting model (q) to 35 and established an upper bound of two for the per-family error rate (PFER; Meinshausen and Bühlmann 2010, Shah and Samworth 2013, see also Hofner et al. 2015 for details in the context of boosting) ~ ~which, given the 48 base-learners in the occupancy and count models (see Equation 1 in manuscript),~~***. This error bound*** corresponded to an upper bound of $\alpha = 0.042$ for the per-comparison error rate ***in the occupancy model and an upper bound of*** α ***= 0.021 in the count model***. ***The choice of*** q ***is rather arbitrary as long as it is large enough to incorporate all important variables in the model (Hofner et al. 2015). We used the unimodality assumption for the compuation of the error bounds in the occupancy model and, to be less conservative, we used the r-concavity assumption in the count model (Shah and Samworth 2013, Hofner et al. 2015).***

~~**Benjamin: can you check my calculation? Also, should we explain why we chose $q = 35$, or describe complementary pairs subsampling and our choice of unimodal vs. r-concavity assumptions for the two model type (occupancy vs. count)? Any other relevant details to include?**~~

BH: As you can see from the different thresholds we have a different number of base-learners in the two models. The occupancy model really has 48 base-learners but the count model has twice as many, i.e., 96 base-learners, 48 for the mean and 48 for the dispersion parameter. Furthermore, if you print any of the stabsel objects, you will find the (realized) PCER given as 0.0208 (= 1/48) for the occupancy model and 0.0188 (= 1.8 / 96) for the count model.

BH: I need to check here why the upper bound for the PFER is specified as 2 but the PFER is then given as 1.

BH: We could (and should) state the further assumptions (unimodality and r-concavity). Do you remember the reason why we use different assumptions for the two models? I added a sentence. Is this correct and do we want to keep it? Perhaps it raises more questions than it answers.

Results

Occupancy models

Given our specifications ($q = 35$; PFER upper-bound = 2, ***unimodality assumption***), only base-learners selected in all 100 subsamples (i.e., $\hat{\pi} = 1$) were identified as stable (Figure 1.1).

Count models

Given our specifications ($q = 35$; PFER upper-bound = 2, ***r-concavity assumption***), only base-learners selected in at least 90 of the 100 subsamples (i.e., $\hat{\pi} = 0.9$) were identified as stable; this threshold applies to the selection of base-learners for the conditional mean (μ) and conditional overdispersion (σ) ***simultaniously***.

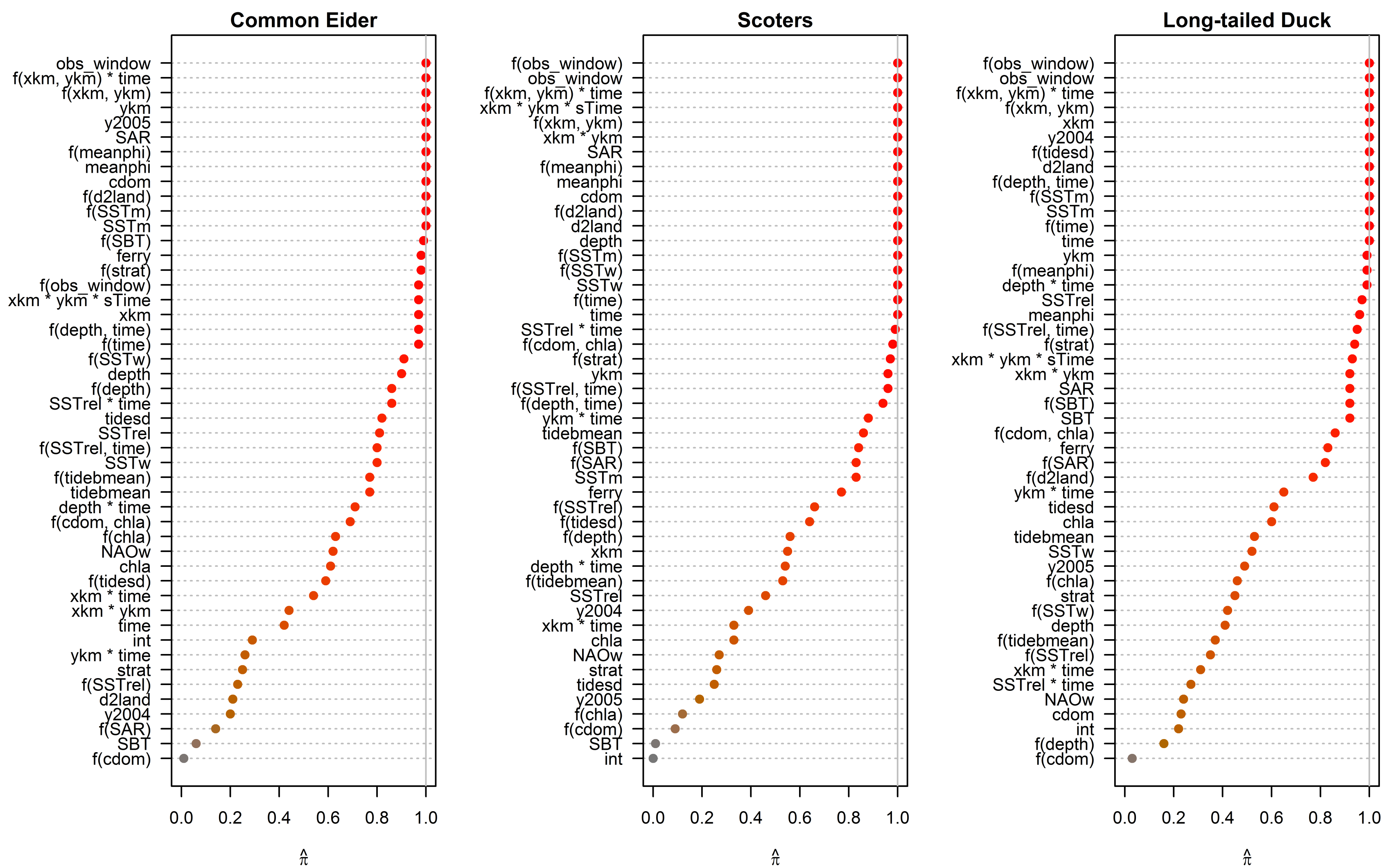

Figure 1.1 Stability selection using complementary pairs subsampling and unimodality assumption for sea duck occupancy models. The number of selected base-learners in each model run was set to $q = 35$. Base-learners with selection frequencies above the threshold ($\hat{\pi}$; vertical gray line) were considered stable with upper bound PFER $= 2$.

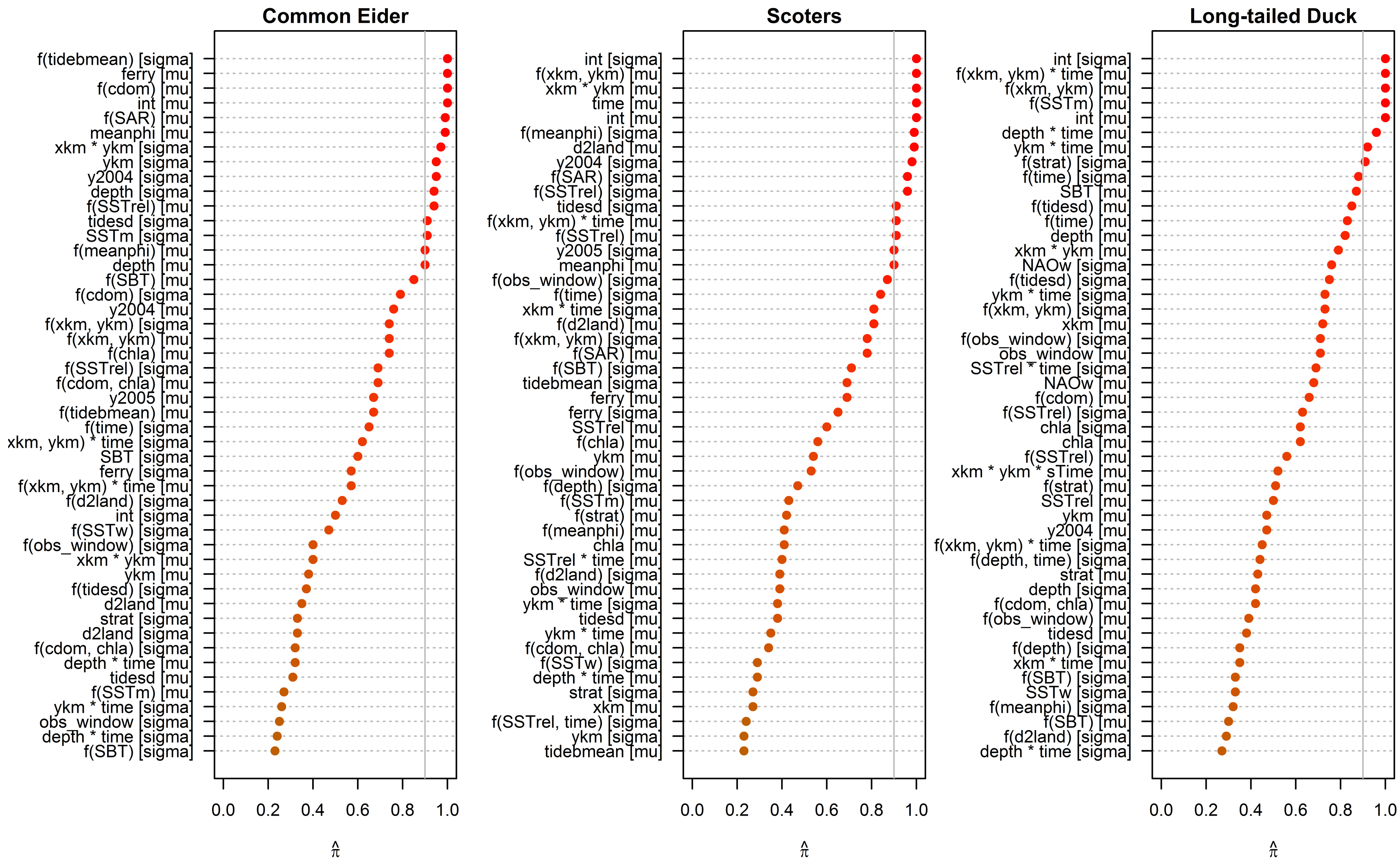

Figure 1.2 Stability selection using complementary pairs subsampling and r-concavity assumption for sea duck conditional count models. The number of selected base-learners in each model run was set to $q = 35$. Base-learners with selection frequencies above the threshold ($\hat{\pi}$; vertical gray line) were considered stable with upper bound PFER = 2. Only the top 48 base-learners are illustrated. Brackets indicate the parameter (conditional mean, *mu*, or overdispersion, *sigma*) to which the base-learner applies.

Literature cited

Hofner, B., L. Boccuto, and M. Göker. 2015. Controlling false discoveries in high-dimensional situations: Boosting with stability selection. BMC Bioinformatics 16:144.

Meinshausen, N., and P. Bühlmann. 2010. Stability selection (with discussion). Journal of the Royal Statistical Society: Series B (Statistical Methodology) 72:417–473.

Shah, R. D., and R. J. Samworth. 2013. Variable selection with error control: another look at stability selection. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 75:55–80.