

Supplementary Information S5 - Linear Mixed Model Effects and post-hoc ANOVA results

Definitions

mod0: “null-model “with intercept-only effect

mod1: model including a single fixed-effect term related to the set of predictive variables used in SDM's

Response variable is named “TestEvalValue” i.e., the performance score for either TSS, ROC or KAPPA

Fixed effect: “set” i.e. the name of the set of predictive variables, either: combined climate and EFA (“CLIM_EFA”), climate-only (“CLIM_only”) or EFA-only (“EFA_only”)

Random factors are: “Species” (species name: 13 in total) and “EvalMetrics” (the evaluation score type: 3 in total, TSS, ROC and KAPPA)

Model results

Model formulae used in R lmer function:

mod0: TestEvalValue ~ 1 + (1 | Species) + (1 | EvalMetrics)

mod1: TestEvalValue ~ set + (1 | Species) + (1 | EvalMetrics)

Table 1 - Scaled residuals

Minimum	1 st Quartile (25%)	Median (50%)	3 rd Quartile (75%)	Maximum
-2.436	-0.477	0.018	0.681	2.488

Figure 1 - Histogram of residuals

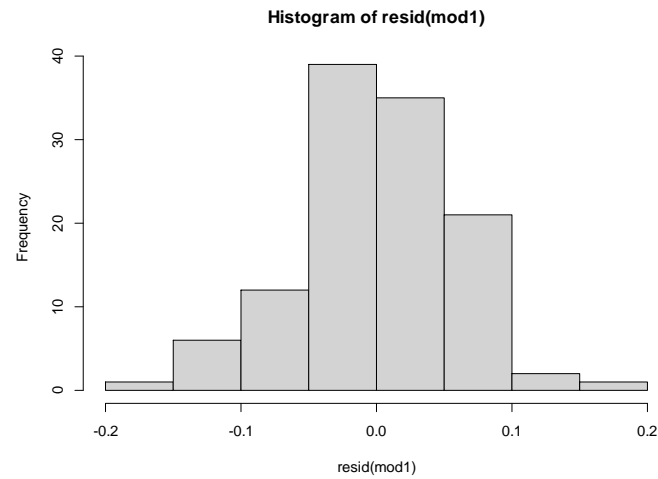


Table 2 – Random effects estimates

Groups	Name	Variance	Std.Dev.
Species	(Intercept)	0.005	0.074
EvalMetrics	(Intercept)	0.024	0.155
Residual		0.004	0.062

Number of observations: 117, groups: Species, 13; EvalMetrics, 3

Results show overall greater variation in performance scores across evaluation metrics than across species.

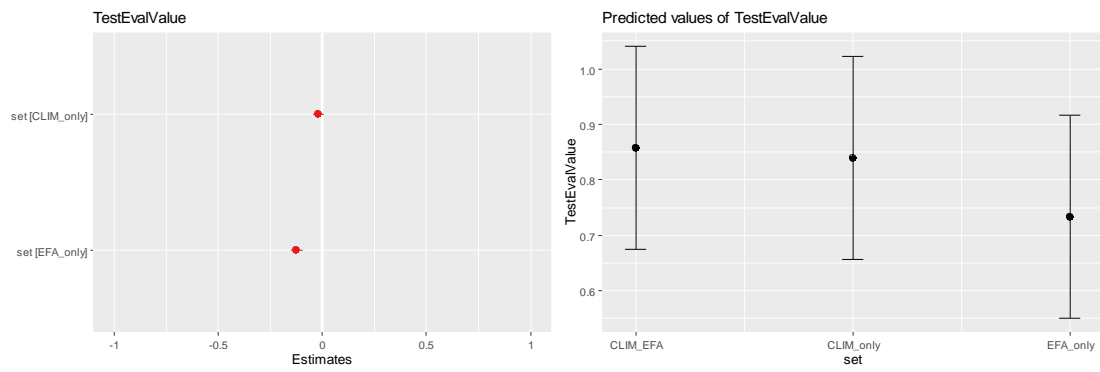
Table 3 – Fixed effects estimates

Model term	Estimate	Std. Error	t-value
(Intercept)	0.858	0.093	9.274
set [CLIM_only]	-0.018	0.015	-1.298
set [EFA_only]	-0.124	0.014	-8.715

Figure 2 – Linear mixed model results

(a)

(b)



Diagnostic: negative model coefficients (Table 3 and Fig. 2 – a) show that (despite uncertainty in coefficient estimates) the set[CLIM_only] and set[EFA_only] both have on average less performant models in relation to the combined set[CLIM_EFA]. That effect is more noticeable for the EFA_only set (Fig. 2 – b).

ANOVA results:

```
mod0 <- lmer(TestEvalValue ~ 1 + (1|Species) + (1|EvalMetrics), data = perfEvalAll)
mod1 <- lmer(TestEvalValue ~ set + (1|Species) + (1|EvalMetrics), data = perfEvalAll)
anova(mod0, mod1)
```

Table 4 – ANOVA post-hoc test results

Model	Number of parameters	AIC	BIC	Log likelihood	Deviance	Chi-square test value	Degrees of freedom	Pr(>Chisq)
<u>mod0</u>	4	-198.19	-187.14	103.09	-206.19			
<u>mod1</u>	6	-258.81	-242.24	135.41	-270.81	64.62	2	<u>$p=9.269 \times 10^{-15}$</u> (***)

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Diagnostic: the model including the “set” term (mod1) is significant at $p < 0.001$

Data used for analysis (Appendix S6)