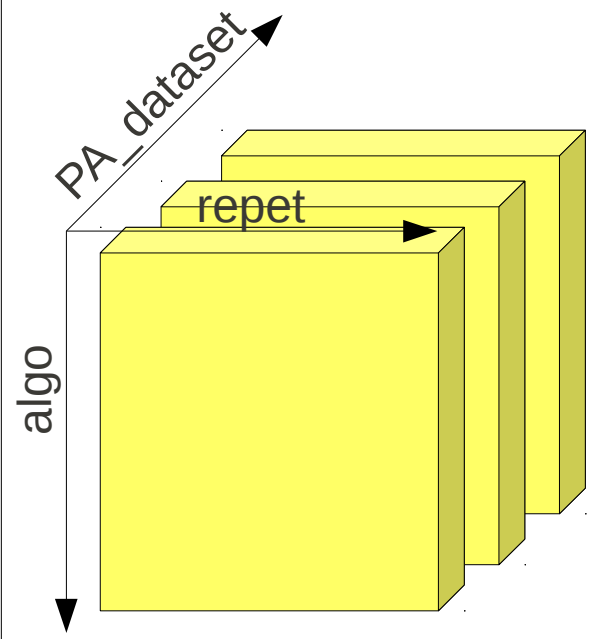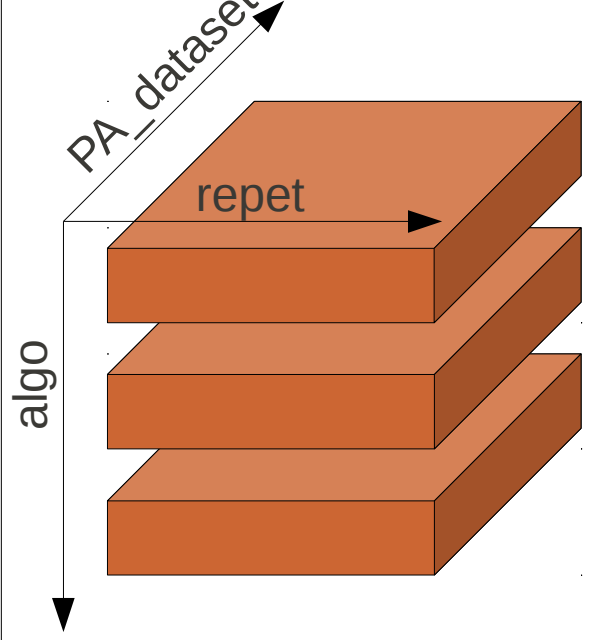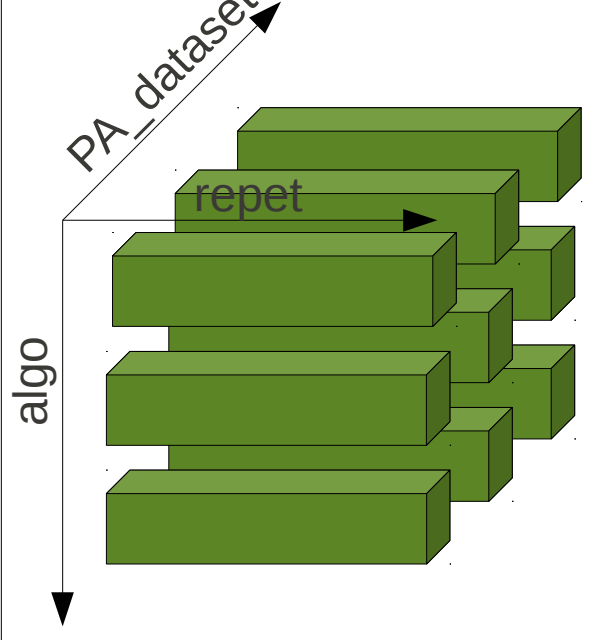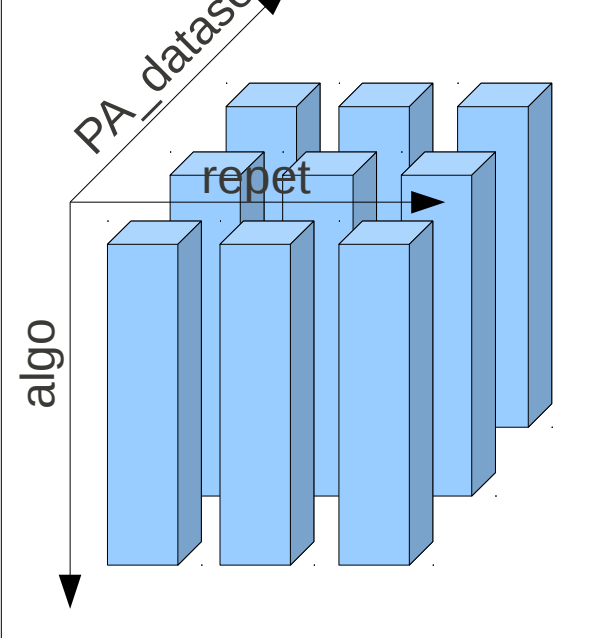# Ensemble Modelling : the different available ways to build ensemble forecasts (em.by argument setting)

The following table shows how models built at the **BIOMOD_Modeling(…)** step can be combined to build ensemble models. These ensembles are controlled by the **em.by** parameter of the **BIOMOD_EnsembleModeling(...)** function.

Several ensemble modelling algorithms can be set up (e.g. mean of probabilities, median of probabilities, confidence intervals, committee averaging... see ?**BIOMOD_EnsembleModeling**)

All ensemble models are evaluated according to selected evaluation metrics and could be projected later using **BIOMOD_EnsembleForecasting(...)** function.

Models can be assembled following 3 dimensions : algo (e.g. GLM, GAM...), PA_dataset (e.g. PA1, PA2... or AllData if no pseudo-absences selection were carried out) and repetition (full, rep1, rep2...). If one of these dimension is 1 (e.g. If no pseudo absences selection done then PA_dataset will be only 'AllData', if no repetition done then repet will be only 'full'...), some ensemble models can be equivalent.

Graphical representation	'em.by' value	Remarks
	'all'	• Output: 1 ensemble model. • Evaluation: Ensemble models are evaluated on the evaluation data if some have been entered at BIOMOD_FormatingData(...) step. If not, the union of all PA_dataset points is used as reference. In the latter case, all 'NA' points (pseudo-absences) are considered as absences. This approach is not perfectly fair because some points used for individual models calibration are re-used for the ensemble models' evaluation.
	'PA_dataset'	• Output: 1 ensemble model is built for each pseudo-absence dataset • Evaluation: Ensemble models are evaluated on the evaluation data if some have been entered at BIOMOD_FormatingData(...) step. If not the whole corresponding PA_dataset is taken as reference.
	'algo'	• Output: 1 ensemble model is built for each algorithm • Evaluation: Ensemble models are evaluated on the evaluation data if some have been entered at BIOMOD_FormatingData(...) step. If not, the union of all PA_dataset points is used as reference. In the latter case, all 'NA' points (pseudo-absences) are considered as absences. This approach is not perfectly fair because some points used for individual models calibration are re-used for ensemble models' evaluation.
	'PA_dataset+algo'	• Output: 1 ensemble model is built for each combination algorithm x pseudo absence selection (then the ensemble is made through the repetitions only). • Evaluation: Ensemble models are evaluated on the evaluation data if some have been entered at BIOMOD_FormatingData(...) step. If not the whole corresponding PA_dataset is taken as reference.
	'PA_dataset+repet'	• Output: 1 ensemble model is built for each combination of pseudo absence selection x repetitions (then the ensemble is made through the algorithms only). • Evaluation: Ensemble models are evaluated on the evaluation data if some have been entered at BIOMOD_FormatingData(...) step. If not, the union of all PA_dataset points is used as reference. In the latest case, all 'NA' points (pseudo-absences) are considered as absences. This approach is not perfectly fair because some points used for individual models calibration are re-used for ensemble models' evaluation.