

Low Heritability ($H^2 = 0.3$, $h^2 = 0.18$)

Diag=Diagonal, Res=Residual, R=Random Effect, SS=Single stage, UNW = Unweighted

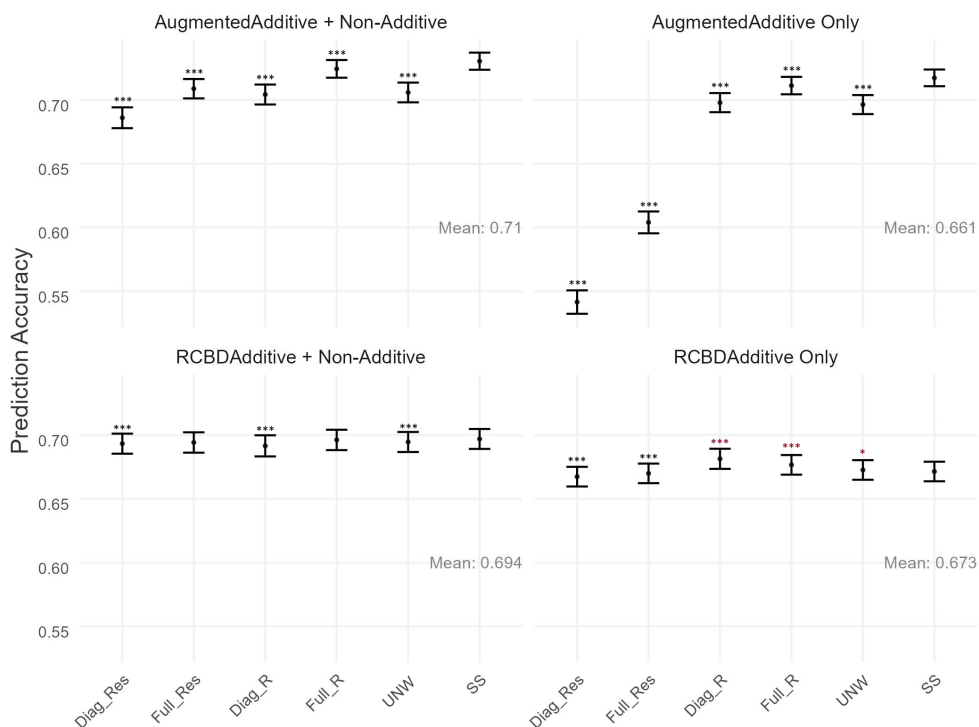


Figure S1: Overview of the average accuracy of prediction of the additive breeding values of the test set across the 500 repetitions for the different models in the four scenarios tested within the low heritability simulation (RCBD and augmented RCBD designs, models with or without a random term for non-additive genetic effects). Within each scenario, six models were tested. From left to right, they correspond to 1-4) two-stage models with the full inverse of EEV ("Full") or only its diagonal elements ("Diag") in the residuals ("Res") or a random effect ("R"); 5) non-fully-efficient two-stage model ("UNW"); 6) single-stage model (SS). The points correspond to the average accuracy and the error bars refer to the standard error of the mean. The asterisks indicate if there were significant differences between each model and the single-stage one (no asterisk: $p\text{-value} > 0.05$; one asterisk: $0.01 < p\text{-value} < 0.05$; two asterisks: $0.001 < p\text{-value} < 0.01$; three asterisks: $p\text{-value} < 0.001$). Asterisks in black were used when the model performed worse than single-stage and red color indicates that they were better.

High Heritability ($H^2 = 0.7$, $h^2 = 0.42$)

Diag=Diagonal, Res=Residual, R=Random Effect, SS=Single stage, UNW = Unweighted

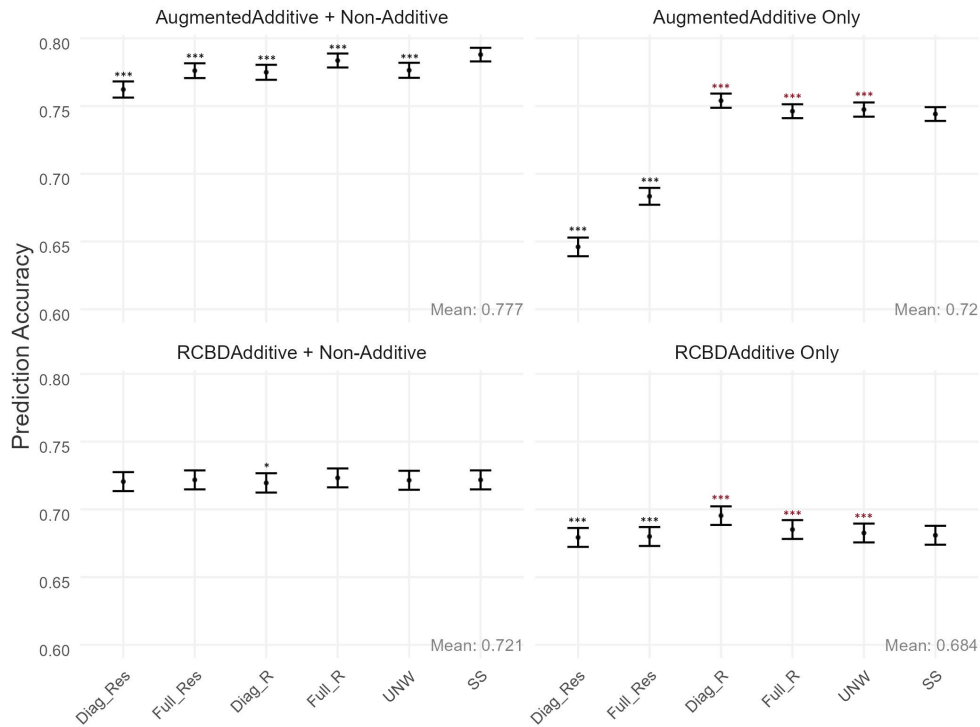


Figure S2: Overview of the average accuracy of prediction of the additive breeding values of the test set across the 500 repetitions for the different models in the four scenarios tested within the high heritability simulation (RCBD and augmented RCBD designs, models with or without a random term for non-additive genetic effects). Within each scenario, six models were tested. From left to right, they correspond to 1-4) two-stage models with the full inverse of EEV ("Full") or only its diagonal elements ("Diag") in the residuals ("Res") or a random effect ("R"); 5) unweighted two-stage model ("UNW"); 6) single-stage model (SS). The points correspond to the average accuracy and the error bars refer to the standard error of the mean. The asterisks indicate if there were significant differences between each model and the single-stage one (no asterisk: $p\text{-value} > 0.05$; one asterisk: $0.01 < p\text{-value} < 0.05$; two asterisks: $0.001 < p\text{-value} < 0.01$; three asterisks: $p\text{-value} < 0.001$). Asterisks in black were used when the model performed worse than single-stage and red color indicates that they were better.

Low Heritability ($H^2 = 0.3$, $h^2 = 0.18$), Accuracy Within Training Set

Diag=Diagonal, Res=Residual, R=Random Effect, SS=Single stage

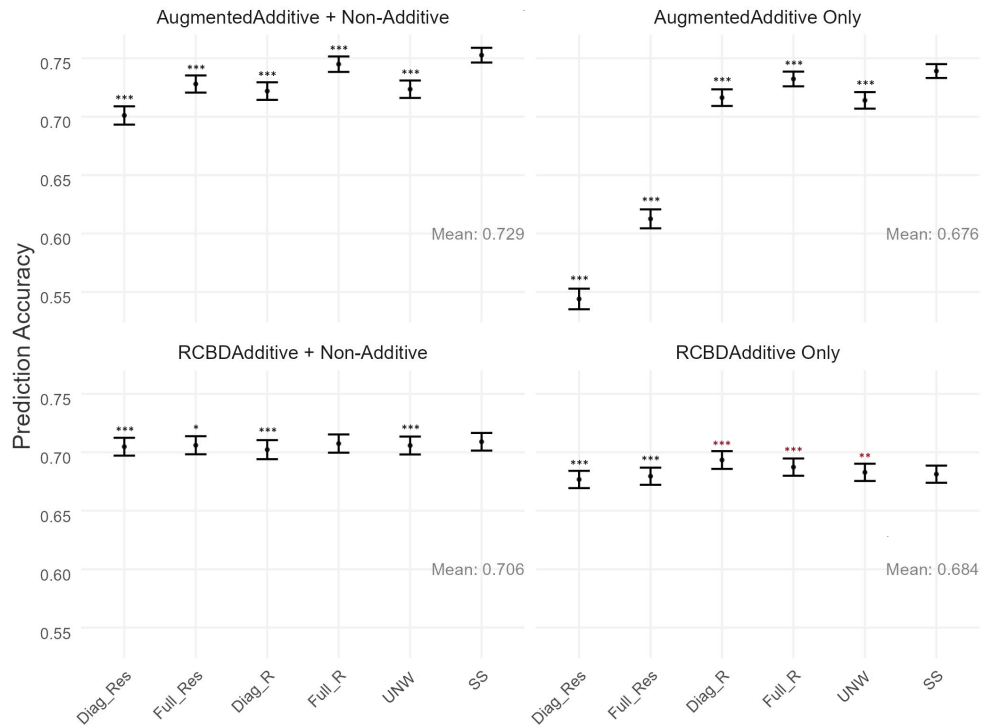


Figure S3: Overview of the average accuracy of prediction of the additive breeding values within the training set across the 500 repetitions for the different models in the four scenarios tested within the low heritability simulation (RCBD and augmented RCBD designs, models with or without a random term for non-additive genetic effects). Within each scenario, six models were tested. From left to right, they correspond to 1-4) two-stage models with the full inverse of EEV ("Full") or only its diagonal elements ("Diag") in the residuals ("Res") or a random effect ("R"); 5) unweighted two-stage model ("UNW"); 6) single-stage model (SS). The points correspond to the average accuracy and the error bars refer to the standard error of the mean. The asterisks indicate if there were significant differences between each model and the single-stage one (no asterisk: $p\text{-value} > 0.05$; one asterisk: $0.01 < p\text{-value} < 0.05$; two asterisks: $0.001 < p\text{-value} < 0.01$; three asterisks: $p\text{-value} < 0.001$). Asterisks in black were used when the model performed worse than single-stage and red color indicates that they were better.

Intermediate Heritability ($H^2 = 0.5$, $h^2 = 0.3$), Accuracy Within Training Set

Diag=Diagonal, Res=Residual, R=Random Effect, SS=Single stage

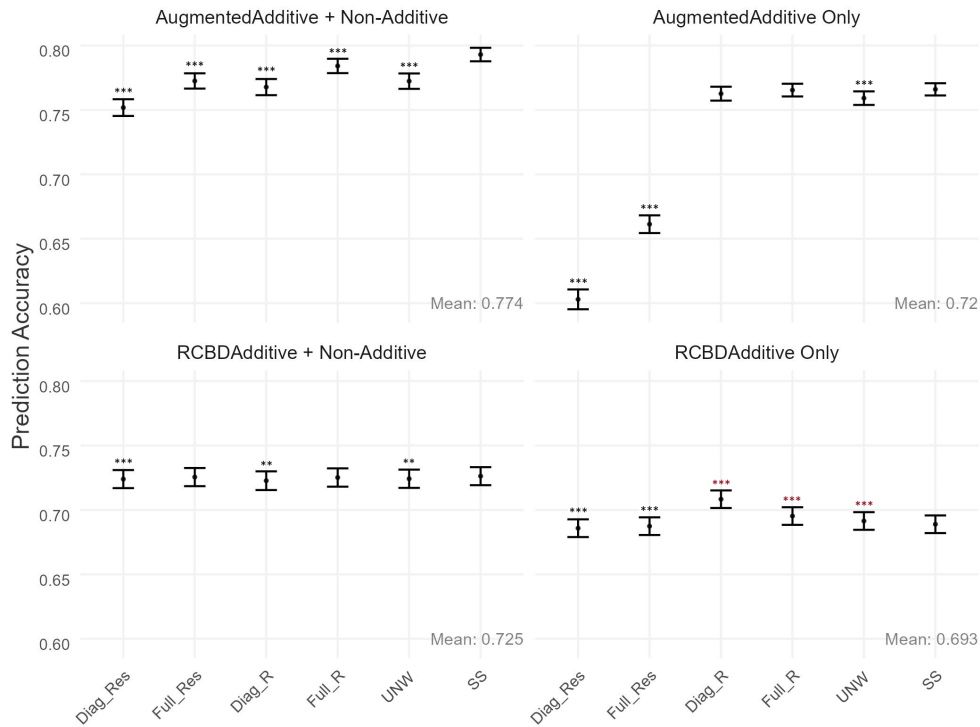


Figure S4: Overview of the average accuracy of prediction of the additive breeding values within the training set across the 500 repetitions for the different models in the four scenarios tested within the intermediate heritability simulation (RCBD and augmented RCBD designs, models with or without a random term for non-additive genetic effects). Within each scenario, six models were tested. From left to right, they correspond to 1-4) two-stage models with the full inverse of EEV ("Full") or only its diagonal elements ("Diag") in the residuals ("Res") or a random effect ("R"); 5) unweighted two-stage model ("UNW"); 6) single-stage model (SS). The points correspond to the average accuracy and the error bars refer to the standard error of the mean. The asterisks indicate if there were significant differences between each model and the single-stage one (no asterisk: $p\text{-value} > 0.05$; one asterisk: $0.01 < p\text{-value} < 0.05$; two asterisks: $0.001 < p\text{-value} < 0.01$; three asterisks: $p\text{-value} < 0.001$). Asterisks in black were used when the model performed worse than single-stage and red color indicates that they were better.

High Heritability ($H^2 = 0.7$, $h^2 = 0.42$), Accuracy Within Training Set

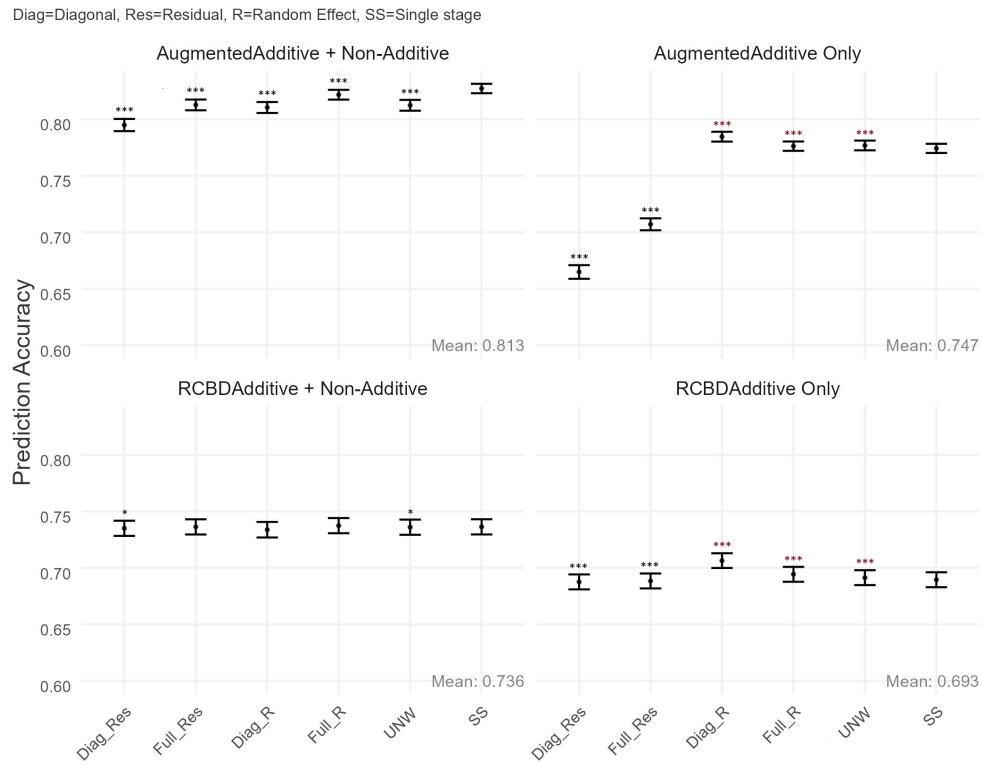


Figure S5: Overview of the average accuracy of prediction of the additive breeding values within the training set across the 500 repetitions for the different models in the four scenarios tested within the high heritability simulation (RCBD and augmented RCBD designs, models with or without a random term for non-additive genetic effects). Within each scenario, six models were tested. From left to right, they correspond to 1-4) two-stage models with the full inverse of EEV ("Full") or only its diagonal elements ("Diag") in the residuals ("Res") or a random effect ("R"); 5) unweighted two-stage model ("UNW"); 6) single-stage model (SS). The points correspond to the average accuracy and the error bars refer to the standard error of the mean. The asterisks indicate if there were significant differences between each model and the single-stage one (no asterisk: $p\text{-value} > 0.05$; one asterisk: $0.01 < p\text{-value} < 0.05$; two asterisks: $0.001 < p\text{-value} < 0.01$; three asterisks: $p\text{-value} < 0.001$). Asterisks in black were used when the model performed worse than single-stage and red color indicates that they were better.

RMSE in Variance Estimation for Low Heritability ($H^2 = 0.3$, $h^2 = 0.18$)							
		RCBD			Augmented		
		Environmental Variance	Additive Variance	Non-Additive Variance	Environmental Variance	Additive Variance	Non-Additive Variance
Additive + Non-Additive	Diag_Res	1.77	0.33	0.15	2.65	0.64	4.44
	Full_Res	1.70	0.33	0.14	1.88	0.37	1.73
	Diag_R	1.66	0.28	0.14	1.72	0.27	0.21
	Full_R	1.68	0.30	0.14	1.70	0.27	0.17
	UNW	1.74	0.32	0.15	1.76	0.29	0.21
	SS	1.69	0.32	0.14	1.69	0.28	0.16
Additive Only	Diag_Res	1.78	0.80		6.66	169.85	
	Full_Res	1.71	0.79		1.97	55.70	
	Diag_R	1.67	0.57		1.72	0.45	
	Full_R	1.68	0.70		1.70	0.52	
	UNW	1.74	0.74		1.76	0.48	
	SS	1.69	0.79		1.69	0.56	

Table S1: Root mean square error (RMSE) of the estimation of variance components for the different combinations of models and experimental designs tested in the Low Heritability scenario.

RMSE in Variance Estimation for High Heritability ($H^2 = 0.7$, $h^2 = 0.42$)							
		RCBD			Augmented		
		Environmental Variance	Additive Variance	Non-Additive Variance	Environmental Variance	Additive Variance	Non-Additive Variance
Additive + Non-Additive	Diag_Res	1.76	1.64	0.54	2.61	1.95	4.73
	Full_Res	1.70	1.64	0.54	1.87	1.61	1.83
	Diag_R	1.67	1.51	0.52	1.72	1.44	0.42
	Full_R	1.68	1.59	0.54	1.70	1.48	0.37
	UNW	1.74	1.62	0.54	1.76	1.49	0.42
	SS	1.69	1.62	0.54	1.69	1.50	0.35
Additive Only	Diag_Res	1.78	6.85		7.30	195.55	
	Full_Res	1.72	6.83		1.99	71.21	
	Diag_R	1.67	5.46		1.72	4.26	
	Full_R	1.69	6.32		1.70	5.21	
	UNW	1.74	6.53		1.76	4.55	
	SS	1.69	6.79		1.69	5.60	

Table S2: Root mean square error (RMSE) of the estimation of variance components for the different combinations of models and experimental designs tested in the High Heritability scenario.