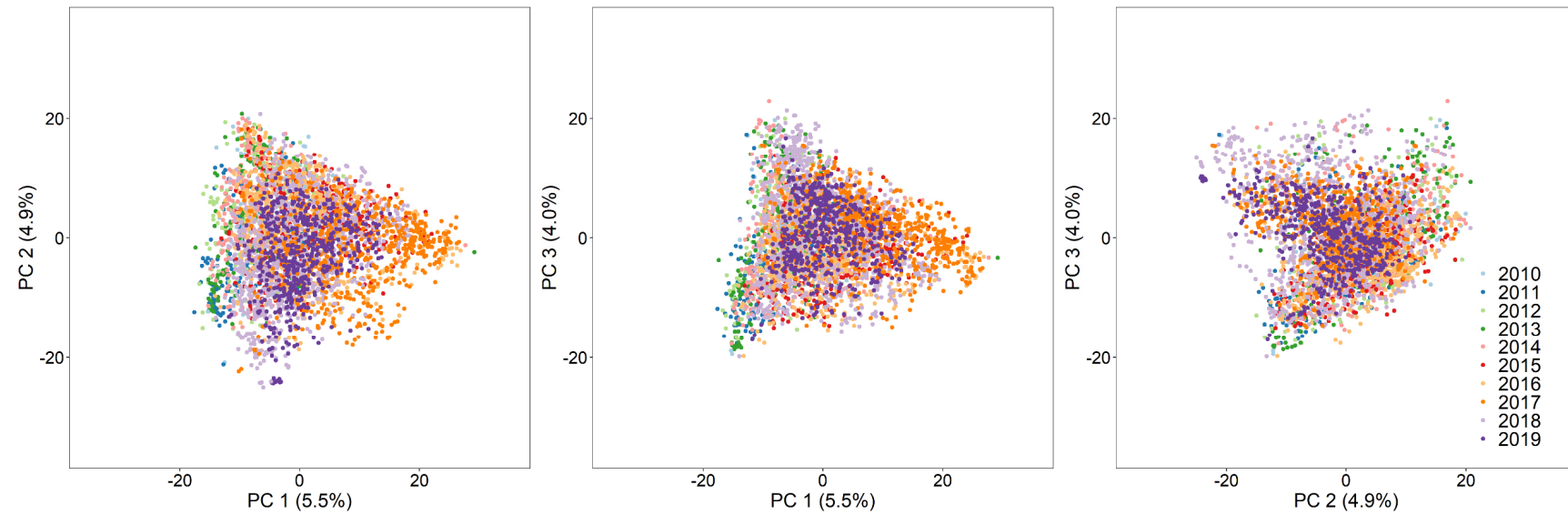**Supplementary material**

Article Title: Multi-year dynamics of single-step genomic prediction in an applied wheat breeding program

Journal: Agronomy (ISSN 2073-4395; CODEN: ABSGGL)
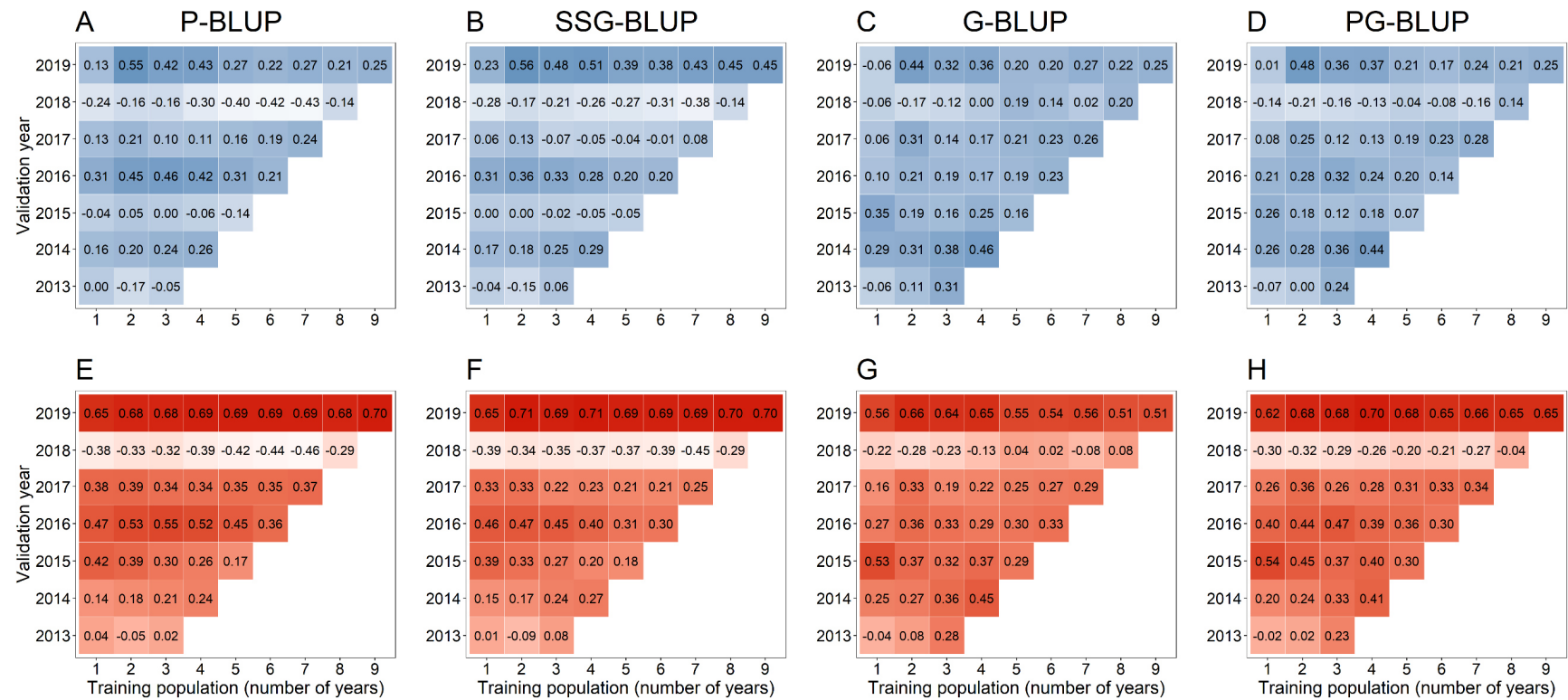
Authors: Sebastian Michel, Franziska Löschenberger, Ellen Sparry, Christian Ametz, Hermann Bürstmayr

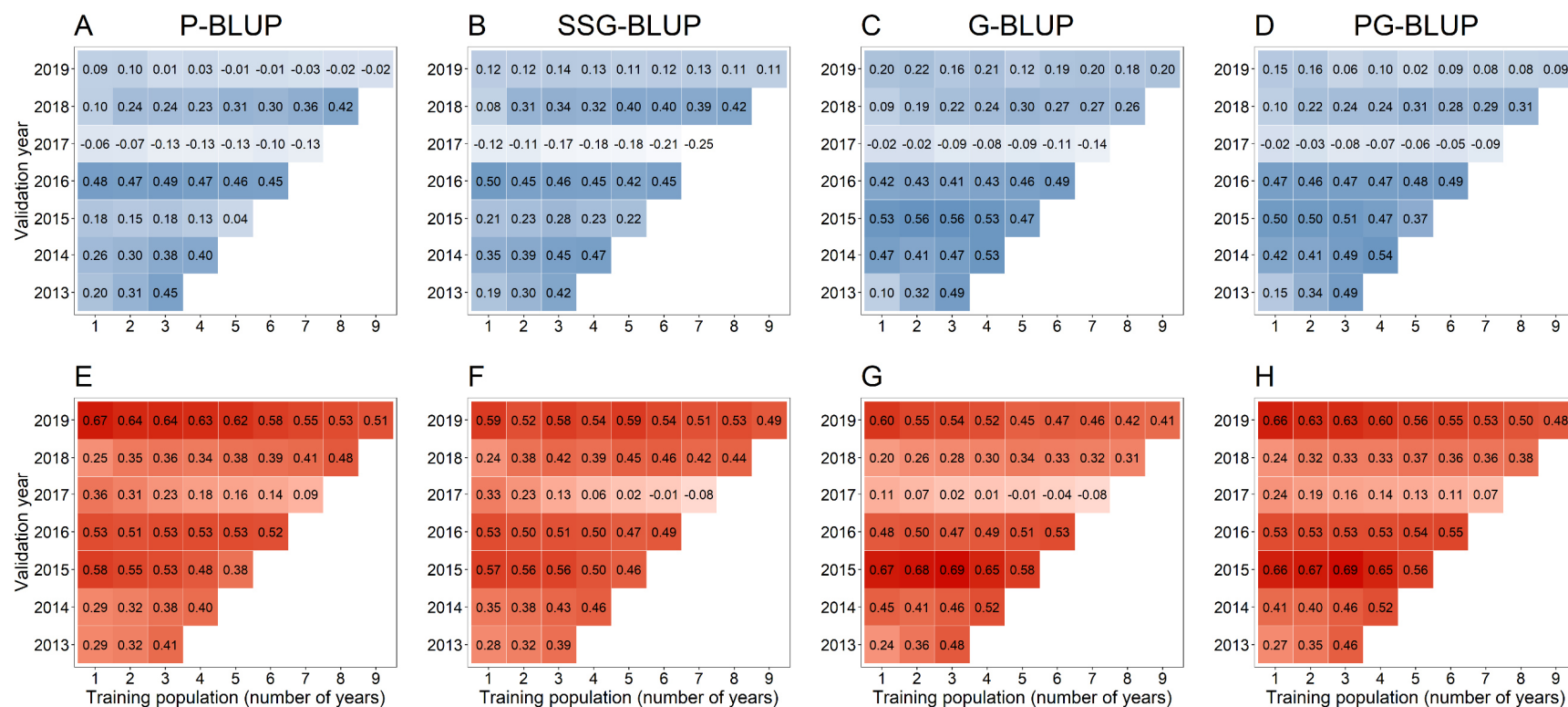**Name, affiliation, and email of corresponding author:**

Sebastian Michel
Department for Agrobiotechnology (IFA-Tulln)
Institute for Biotechnology in Plant Production
University of Natural Resources and Life Sciences, Vienna (BOKU)
Konrad-Lorenz-Str. 20, 3430 Tulln, Austria
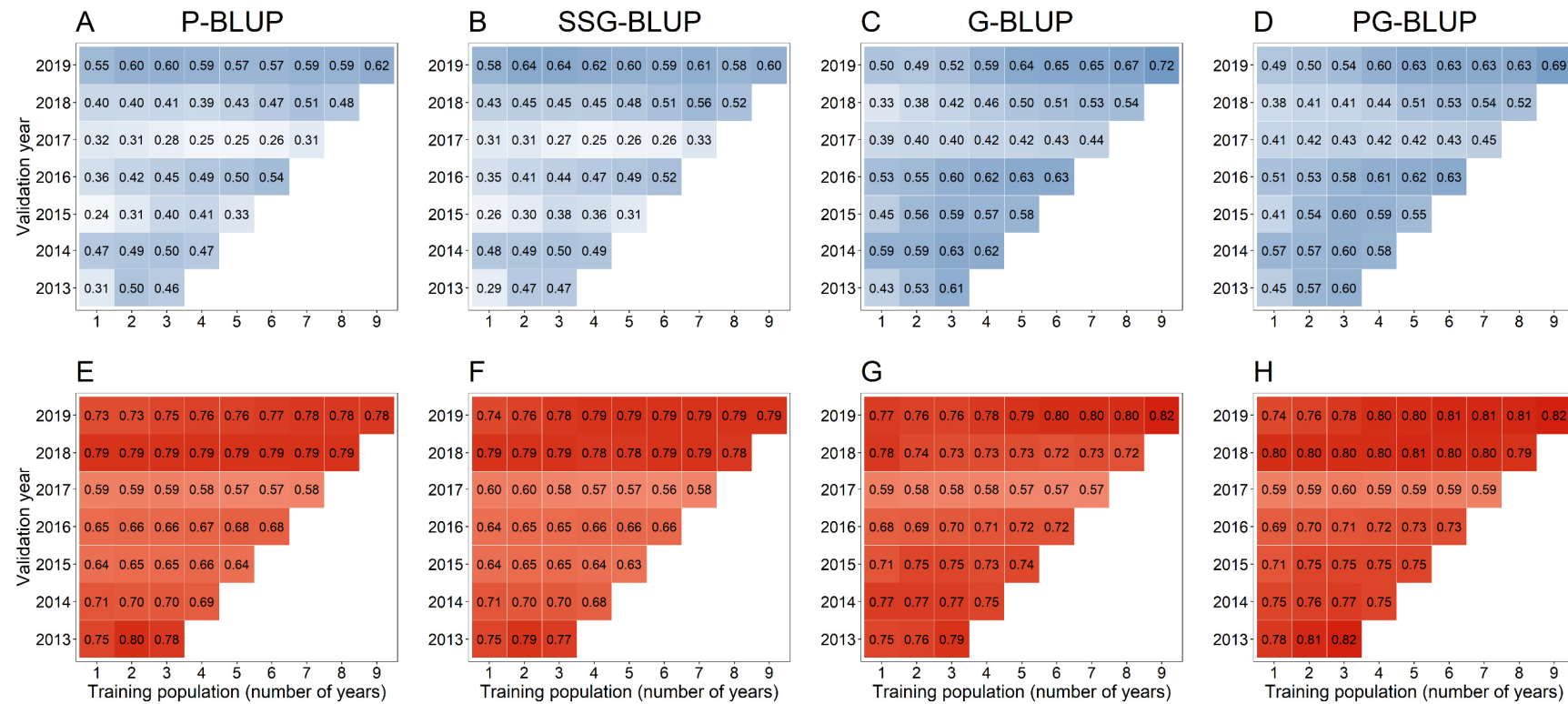e-mail: sebastian.michel@boku.ac.at

Suppl. Figure S1 Principal component analysis of the 4032 recombinant inbred and double haploid breeding lines involved in the study. The lines are coloured according to their membership to the breeding cycles 2010 to 2019 that were designated to the first year of multi-environment testing of a particular cohort.

Suppl. Figure S2 Prediction accuracy for protein yield using pedigree-based (A), single-step-based (B), genomic-based (C), and pedigree-genomic-based (D) forward predictions (blue coloured) as well as pedigree-assisted (E), single-step-assisted (F), genomic-assisted (G), and pedigree-genomic-assisted (H) forward predictions (red coloured) with the latter exploiting pre-existing information from preliminary yield trials. The training population for each prediction model was built by randomly sampling 100 lines from one to nine years preceding the respective validation year.
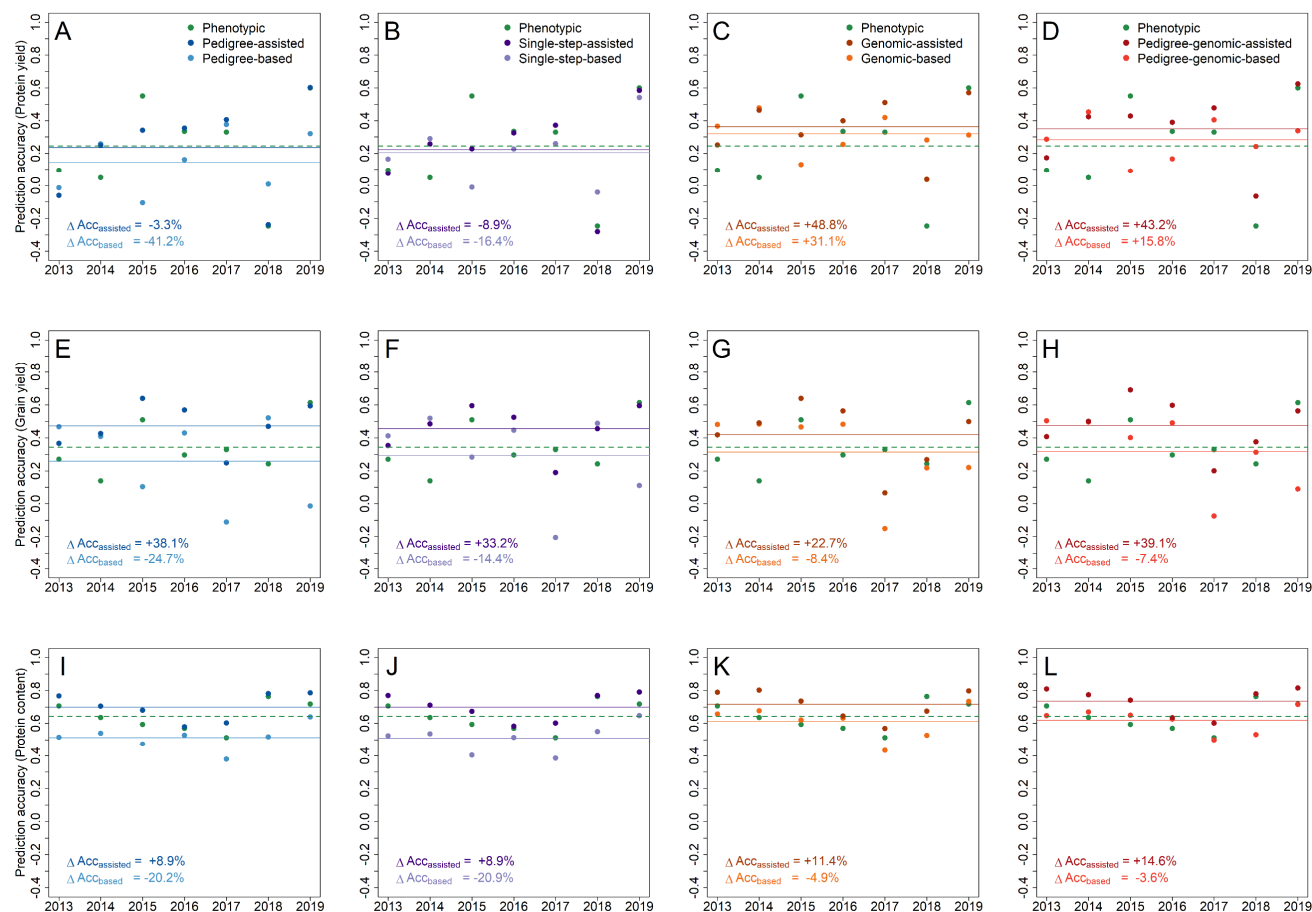
Suppl. Figure S3 Prediction accuracy for grain yield using pedigree-based (A), single-step-based (B), genomic-based (C), and pedigree-genomic-based (D) forward predictions (blue coloured) as well as pedigree-assisted (E), single-step-assisted (F), genomic-assisted (G), and pedigree-genomic-assisted (H) forward predictions (red coloured) with the latter exploiting pre-existing information from preliminary yield trials. The training population for each prediction model was built by randomly sampling 100 lines from one to nine years preceding the respective validation year.

Suppl. Figure S4 Prediction accuracy for the protein content using pedigree-based (A), single-step-based (B), genomic-based (C), and pedigree-genomic-based (D) forward predictions (blue coloured) as well as pedigree-assisted (E), single-step-assisted (F), genomic-assisted (G), and pedigree-genomic-assisted (H) forward predictions (red coloured) with the exploiting pre-existing information from preliminary yield trials. The training population for each prediction model was built by randomly sampling 100 lines from one to nine years preceding the respective validation year.
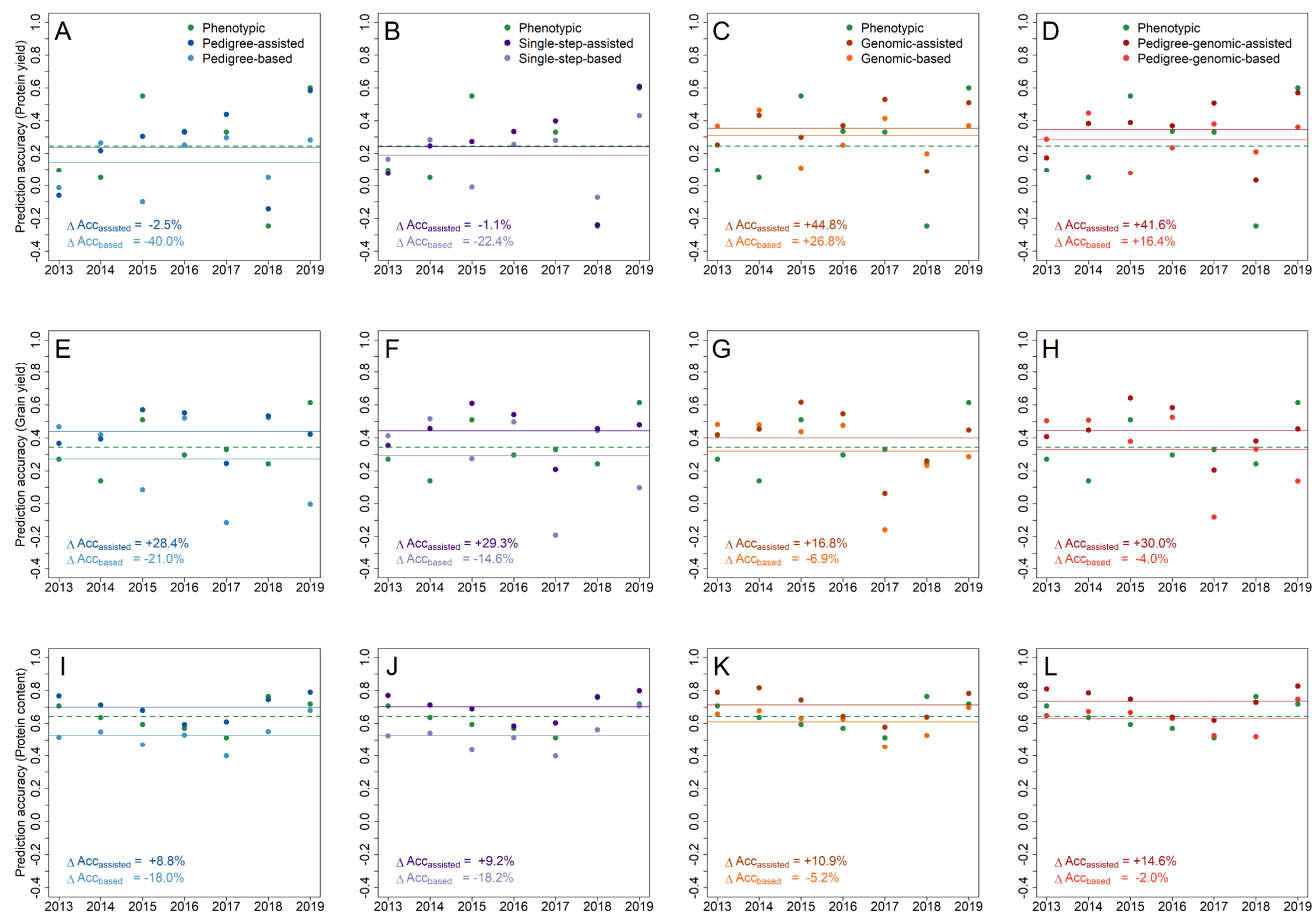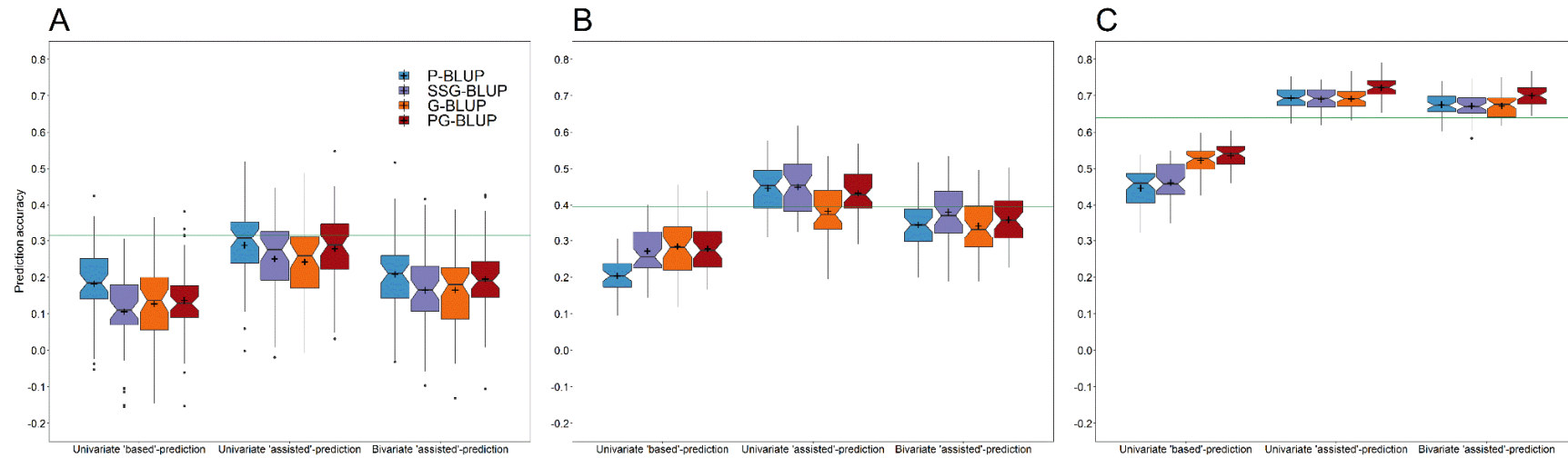
Suppl. Figure S5 Prediction accuracy for protein yield (A-D), grain yield (E-H), and protein content (I-L) using pedigree-based (P-BLUP), single-step-based (SSG-BLUP), genomic-based (G-BLUP), and pedigree-genomic-based (PG-BLUP) forward predictions as well as pedigree-assisted, single-step-assisted, genomic-assisted, and pedigree-genomic-assisted forward predictions exploiting pre-existing information of the validation population from preliminary yield trials. The displayed results were obtained with an accumulating training population of lines tested in multi-environment trials coming from all years preceding the validation years 2013-2019, while only the preliminary yield trial preceding the validation year was used when fitting the prediction models. The average performance of each model (coloured horizontal lines) is compared with the predictive performance of phenotypic selection based on preliminary yield trials (green horizontal dashed line).

Suppl. Figure S6 Prediction accuracy for protein yield (A-D), grain yield (E-H), and protein content (I-L) using pedigree-based (P-BLUP), single-step-based (SSG-BLUP), genomic-based (G-BLUP), and pedigree-genomic-based (PG-BLUP) forward predictions as well as pedigree-assisted, single-step-assisted, genomic-assisted, and pedigree-genomic-assisted forward predictions exploiting pre-existing information of the validation population from preliminary yield trials. The displayed results were obtained with an accumulating training population of lines tested in multi-environment trials coming from all years preceding the validation years 2013-2019, while all preliminary yield trials preceding the validation year were used when fitting the prediction models. The average performance of each model (coloured horizontal lines) is compared with the predictive performance of phenotypic selection based on preliminary yield trials (green horizontal dashed line).

Suppl. Figure S7 Prediction accuracy for protein yield (A), grain yield (B), and protein content (C) using pedigree (P-BLUP), single-step (SSG-BLUP), genomic (G-BLUP), and pedigree-genomic (PG-BLUP) prediction. A random sample of 50 lines tested in multi-environment trials in 2015-2019 constituted one validation population for each of these years, while 100 lines tested in multi-environment trials from each of the three years preceding a respective validation year were sampled into a training population. The training population was subsequently augmented by phenotypic records from preliminary yield trials of lines already present in the training population in order to establish a connection between the multi-environment and preliminary yield trials for a heterogenous compound symmetry model of the form:

$$y_{ijk} = \mu + a_i + g_j + l_k + gl_{jk} + e_{ijk}$$

where μ is the grand mean, y_{ijk} are the BLUEs obtained in the phenotypic analysis for the kth line, a_i the fixed set-by-year effect, g_j the fixed effect of the jth group separating multi-environment and preliminary yield trials, and e_{ijk} the residual effect with $\mathbf{e} \sim N(\mathbf{0}, \mathbf{I}\sigma_e^2)$. The random line effect l_k was modelled as normally distributed with different kinship matrices as described in the main text, while the effect of the kth line nested within the jth group was modelled as random with diagonal variance-covariance structure assigning a unique genetic variance to the multi-environment and preliminary yield trials respectively. Given that only two groups are considered this invoked a variance-covariance structure between the lines that is similar to a bivariate model suggested by [1], and the breeding values were computed as $bv_k = \mu + l_k + gl_{1k}$ where l_k is estimated main additive effect of the kth line and gl_{1k} is the deviation of this main effect in the group of multi-environment trials. All other models were fitted as described in the main text, while in the 'assisted'-predictions pre-existing information of the validation population from preliminary yield trials was exploited in contrast to the 'based'-predictions.

[1] Tsai, H.; Cericola, F.; Edriss, V.; Andersen, J.R.; Id, J.O.; Jensen, J.D.; Jahoor, A.; Janss, L.; Jensen, J. Use of multiple traits genomic prediction , genotype by environment interactions and spatial effect to improve prediction accuracy in yield data. *PLoS One* 2020, *15*, doi:10.1371/journal.pone.0232665.

Suppl. Table S1 Frequency of the models used for spatial correction as in [1] with random row and/or column effects with/without modelling autoregressive variance–covariance structures (AR1) between the plots either in row, in column or in both directions.

Spatial correction	Frequency of chosen model per trait		
	Protein yield	Grain yield	Protein content
1) None	15	19	21
2) AR1 row	4	6	1
3) AR1 column	17	32	11
4) AR1 x AR1	23	31	13
5) row	6	11	7
6) AR1 row + row	3	4	5
7) AR1 column + row	1	4	0
8) AR1 x AR1 + row	5	9	3
9) column	1	1	0
10) AR1 row + column	0	2	1
11) AR1 column + column	8	9	2
12) AR1 x AR1 + column	3	6	3
13) row + column	0	2	0
14) AR1 row + row + column	0	4	0
15) AR1 column + row + column	0	4	0
16) AR1 x AR1 + row + column	1	0	1

[1] Burgueño, J.; Cadena, A.; Crossa, J. *User ' S Guide for Spatial Analysis of Field Variety Trials Using Asreml*; CIMMYT, Mexico, 2000; ISBN 9706480609.

Suppl. Table S2 Mean, range, and heritability (H²) for protein yield, grain yield, and protein yield across the multi-environment trials conducted between 2010 and 2019.

Trait	Year	Trials	Locations	Lines	Min	Mean	Max	H^2
Protein yield (dt ha^{-1})	2010	9	7	127	6.0	7.4	8.8	0.46
	2011	10	6	162	7.5	9.3	10.3	0.23
	2012	15	12	208	6.8	8.2	9.6	0.39
	2013	13	12	193	8.5	10.1	11.7	0.55
	2014	17	12	202	6.9	8.5	10.6	0.70
	2015	4	4	206	7.0	8.8	10.2	0.25
	2016	5	5	192	10.4	12.0	13.6	0.29
	2017	6	6	209	7.9	9.3	10.7	0.48
	2018	2	2	113	7.2	9.2	11.0	0.23
	2019	6	5	114	8.5	9.9	11.3	0.73
Grain yield (dt ha^{-1})	2010	20	16	127	41.5	49.3	56.4	0.52
	2011	20	15	162	59.2	67.6	75.1	0.49
	2012	22	17	208	45.2	52.0	64.6	0.47
	2013	21	17	193	57.9	69.4	84.3	0.60
	2014	24	19	202	50.6	69.2	82.3	0.82
	2015	7	7	206	54.4	64.3	72.7	0.55
	2016	6	6	192	73.3	87.1	100.3	0.59
	2017	9	9	209	55.2	67.3	75.0	0.43
	2018	6	6	113	52.0	67.8	75.2	0.70
	2019	9	8	114	61.7	72.4	82.0	0.67
Protein content (%)	2010	6	5	127	12.7	14.2	15.7	0.62
	2011	12	7	162	11.5	13.1	14.8	0.67
	2012	9	9	208	13.4	15.2	17.2	0.83
	2013	6	6	193	12.4	14.2	15.9	0.70
	2014	7	6	202	12.0	13.9	15.4	0.60
	2015	6	5	206	11.4	12.9	14.7	0.83
	2016	7	6	192	11.8	13.2	15.0	0.80
	2017	7	7	209	12.6	13.9	16.9	0.84
	2018	2	2	113	12.2	14.1	16.4	0.80
	2019	6	5	114	12.2	13.9	16.8	0.93

Suppl. Table S3 Estimated heritability for the pedigree (P-BLUP), genomic (G-BLUP), and pedigree-genomic (PG-BLUP) prediction models with a training population of 1464 lines tested in multi-environment trials between 2010 and 2019.

| | Model | | |
Trait	P-BLUP	G-BLUP	PG-BLUP
Protein yield	0.34	0.27	0.42
Grain yield	0.52	0.37	0.52
Protein content	0.72	0.59	0.75

The heritability was estimated by $h^2 = (\sigma_P^2 - \sigma_e^2)/\sigma_P^2$ where σ_P^2 is the phenotypic variance i.e. the unbiased sample variance computed from the BLUEs of the investigated trait that were adjusted for the estimated year effects of model 3) (see main text), while σ_e^2 was the residual variance estimated with the respective model.