Characterization of the pearl millet cultivation environments in India: status and perspectives enabled by expanded data analytics and digital tools - Supplementary material

Vincent Garin, Sunita Chaudary, Tharanya Murugesan, Sivasakthi Kaliamoorthy, Madina Diancumba, Amir Hajjarpoor, Tara Satyavathi Chellapilla, Shashi Kumar Gupta, Jana Kholova

June 23, 2023

1 Supplementary Figures

Figure S1: Map of the rain pattern

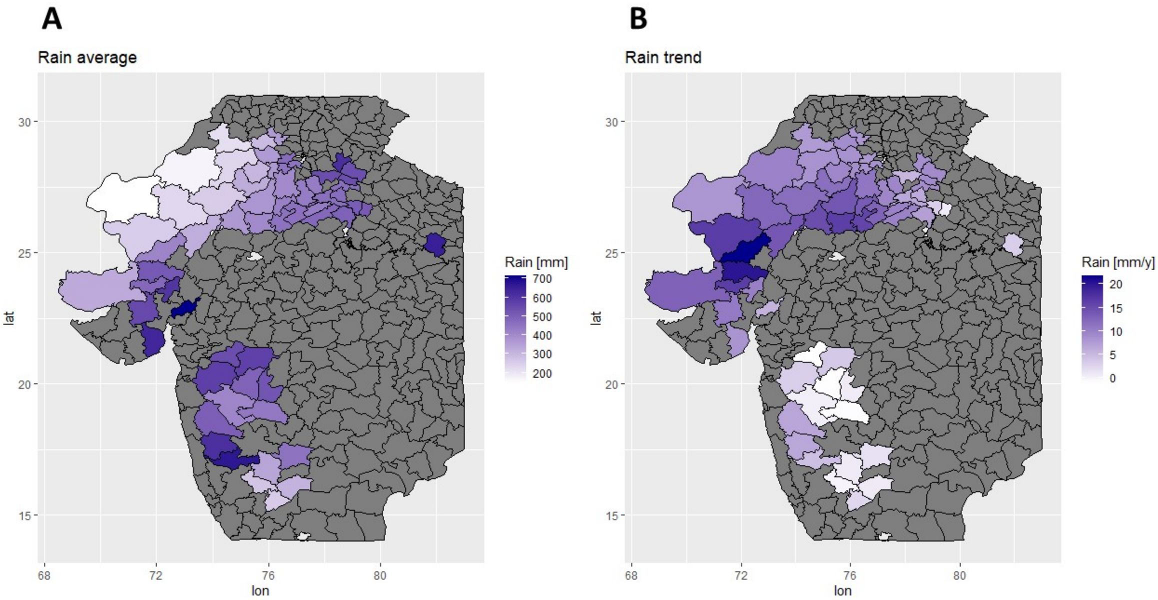

Figure S1: Rain pattern in the selected district representing 90 % of the total Kharif cultivated area between 1998-2017. A) Average rain B) Rain trend

Figure S2: Map of the temperature pattern

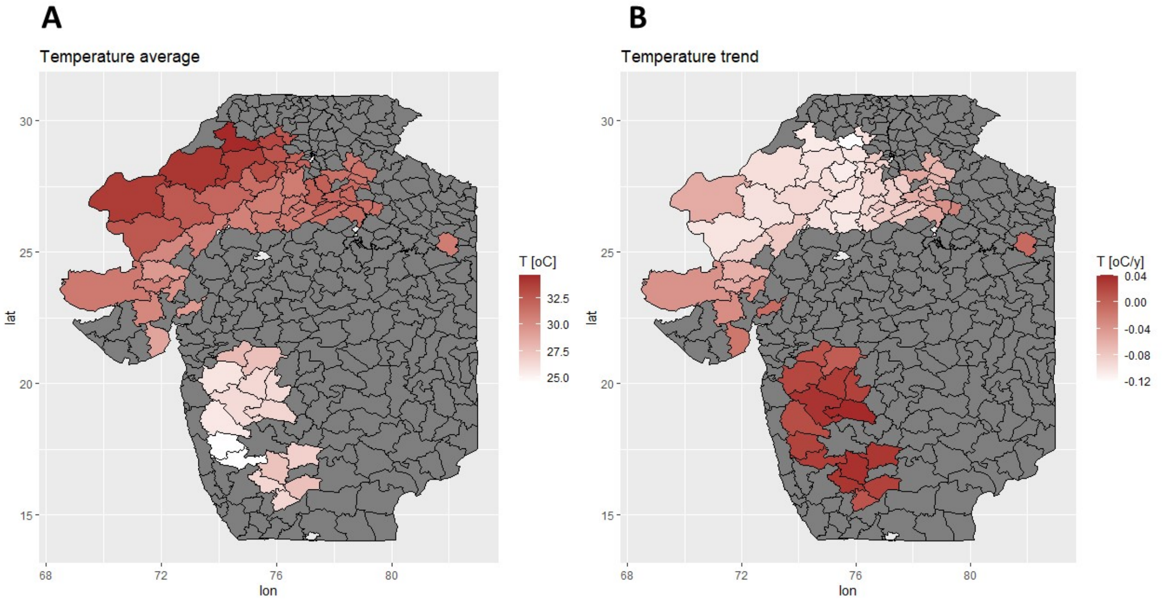

Figure S2: Temperature pattern in the selected district representing 90 % of the total Kharif cultivated area between 1998-2017. A) Average temperature B) Temperature trend

Figure S3: Map of the majority soil type in the pearl millet tract

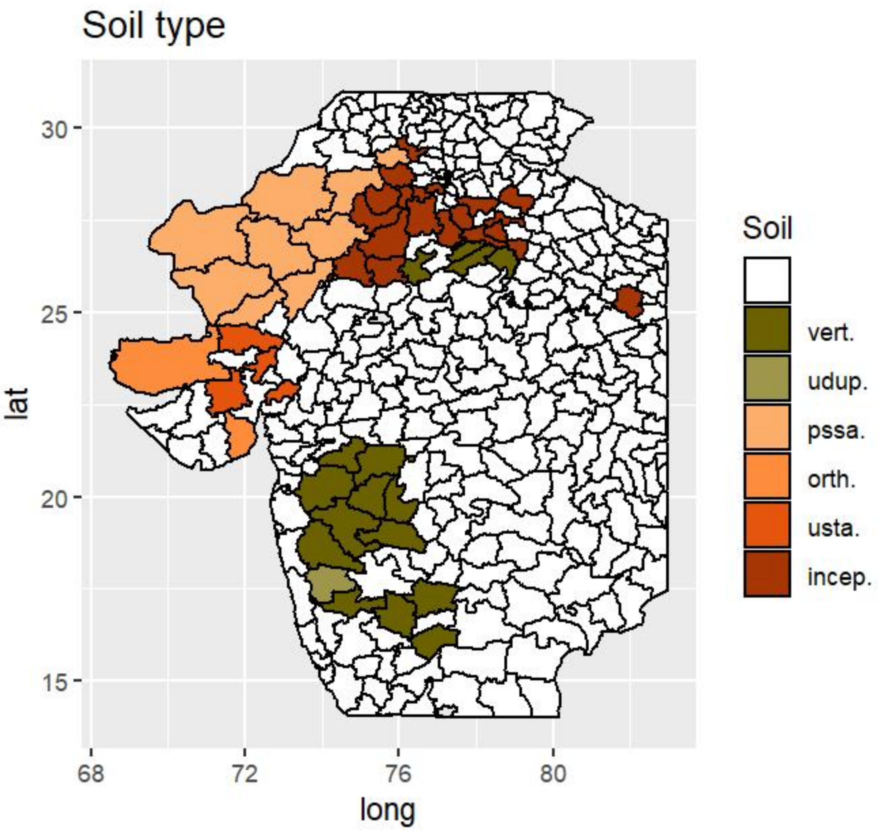

Figure S3: Majority soil type in the selected district representing 90 % of the total Kharif cultivated area between 1998-2017. Soil type: vert. = Vertisol; udup. = Udupt/Udalf; pssa. = Pssament; orth. = Orthid; usta. = Ustalf/Ustoll; incep. = Inceptisol.

Figure S4: Crop model ExM parameter calibration illustration

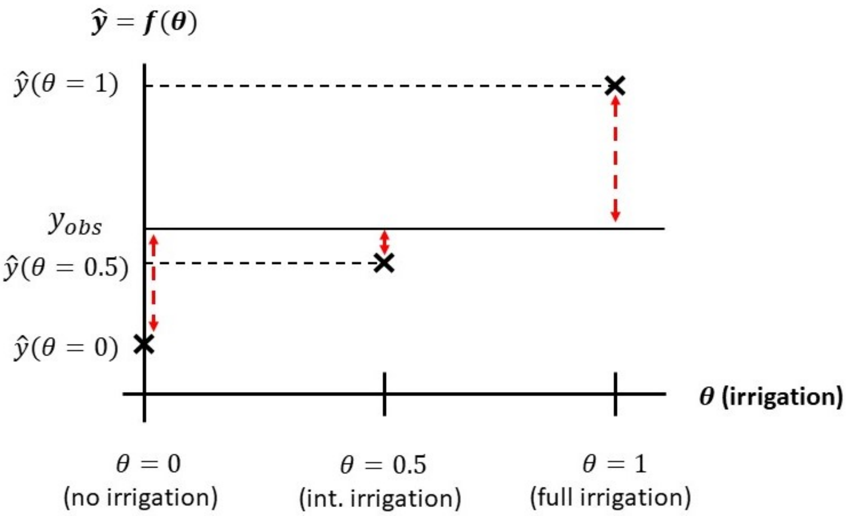

Figure S4: Illustration of the parameter inference strategy by comparing the CM output ($\hat{y}$) obtained for different values of a parameter (θ, here irrigation) with the observed or true value y_{obs}. Here, we can notice that the parameter value ($\theta = 0.5$) produces the output with the lowest difference compared to the true value. Therefore, assuming that an intermediate level of irrigation seems to be the most probable scenario.

Figure S5: B zone area share, area trend and production compared to other crops

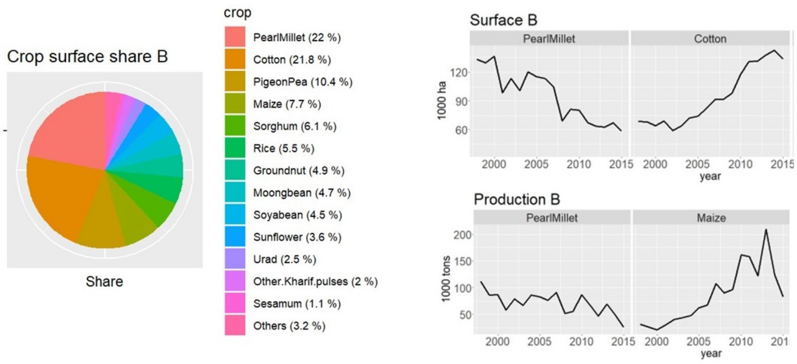

Figure S5: The left part of the figure contains the crop area share of pearl millet and other main crops cultivated in the B zone (1998-2015). The right upper part of the figure illustrates the surface trend of pearl millet and cotton (1998-2015). The lower part is a comparison of the pearl millet production trend compared to maize (1998-2015).

Figure S6: G zone area and production trend compared to other crops

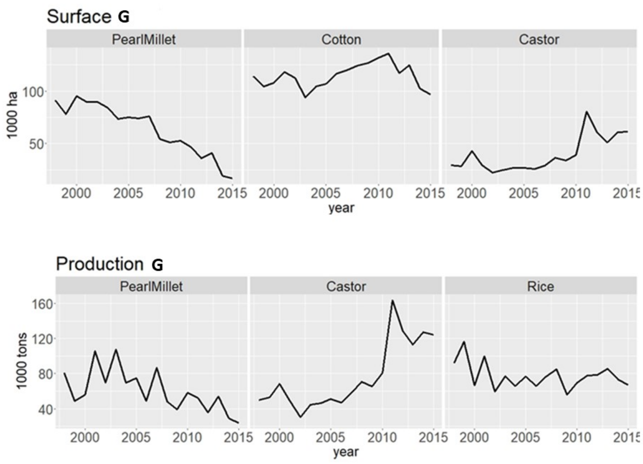

Figure S6: The upper part of the figure represents the area trend of pearl millet, cotton and castor (1998-2015). The lower part is a comparison of the pearl millet production trend compared to castor and rice (1998-2015).

Figure S7: Comparison between observed data and ExM parameters values - soil parameters

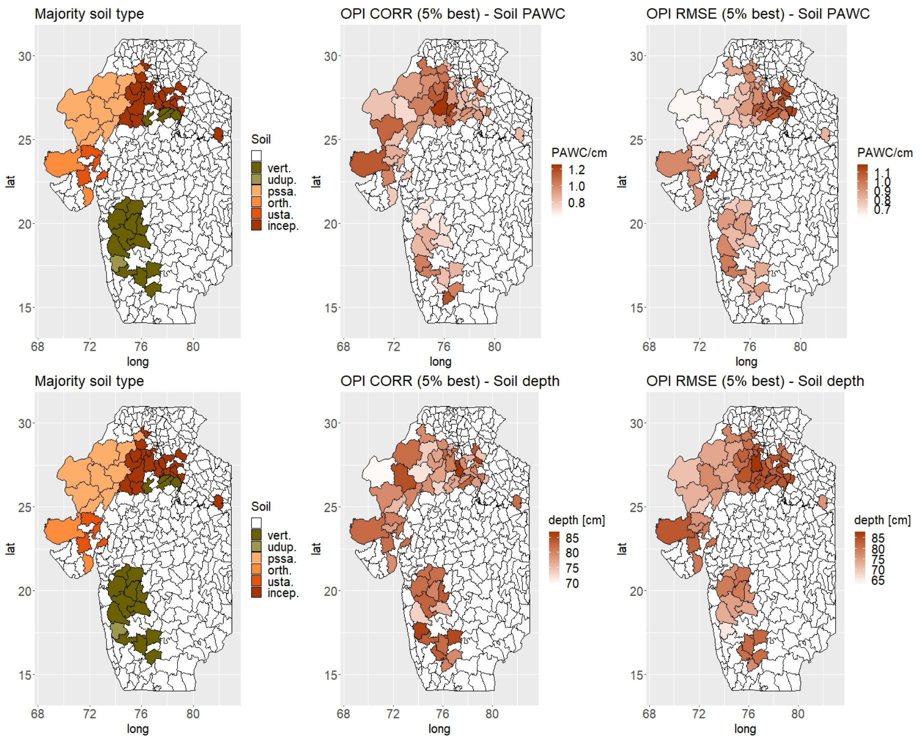

Figure S7: Comparison of the estimated ExM parameter values for soil PAWC and depth with the type of soil. The ExM parameters were the weighted average of the parameter values from the set of parameter with the largest predictive ability according to correlation (5% best results). We also included the results about parameter inferred using the root means square error and the same criteria (average parameter value of the 5% combination with the lowest RMSE). Using the pattern obtained with the RMSE was more consistent with our expectation of soil water content and depth in the observed majority soil types than the one obtained with correlation.

Figure S8: Comparison between observed data and ExM parameters values - management parameters

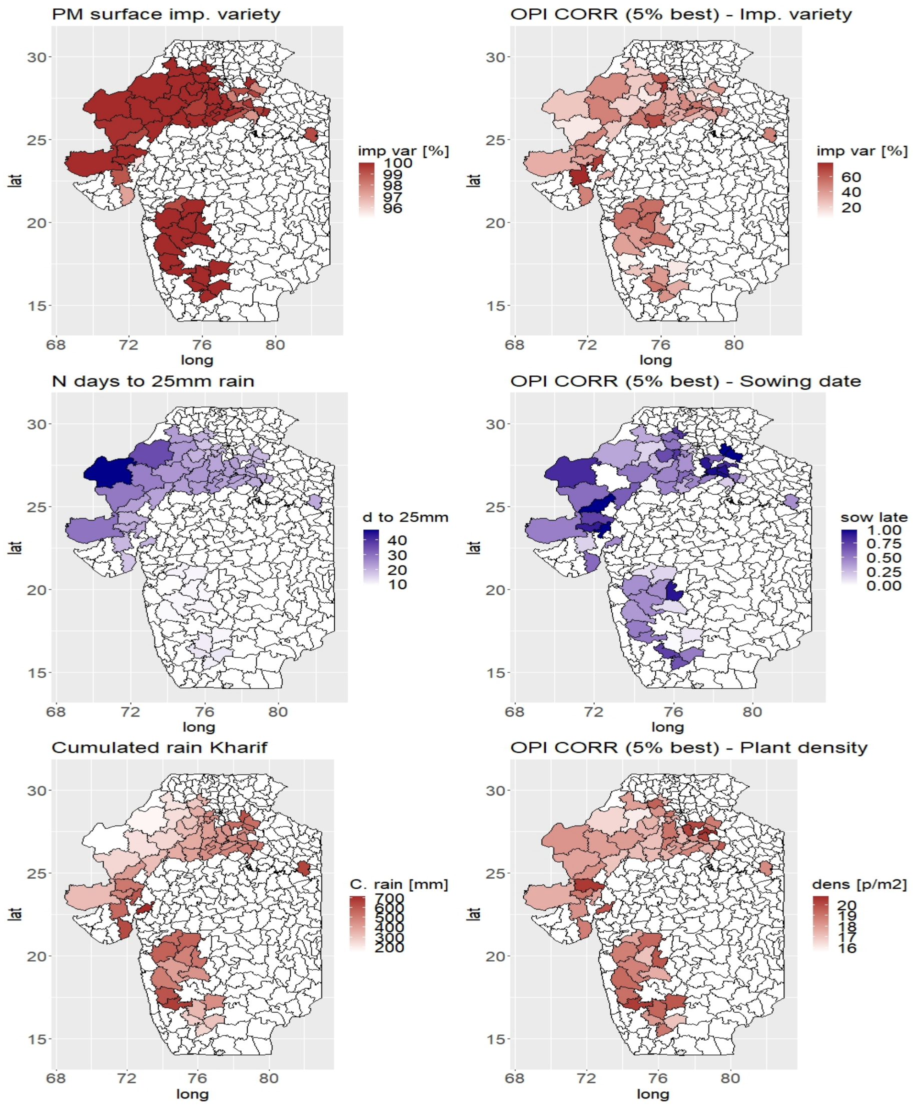

Figure S8: **Upper panel**: comparison of the ExM parameter values for variety with DLD observation about the surface of pearl millet under improved variety. We converted the estimated selected variety using the following scores: landrace (0), HHB67-2 (0.3), 9444 (1). **Middle panel**: comparison of the estimated ExM parameter values for sowing date with the Number of days starting from 1 June to reach 25 mm of rain. **Lower panel**: Comparison of the estimated ExM parameter values for plant density with the average cumulated rain during kharif.

Figure S9: Pearl millet area trend over seasons in the AE2 zone

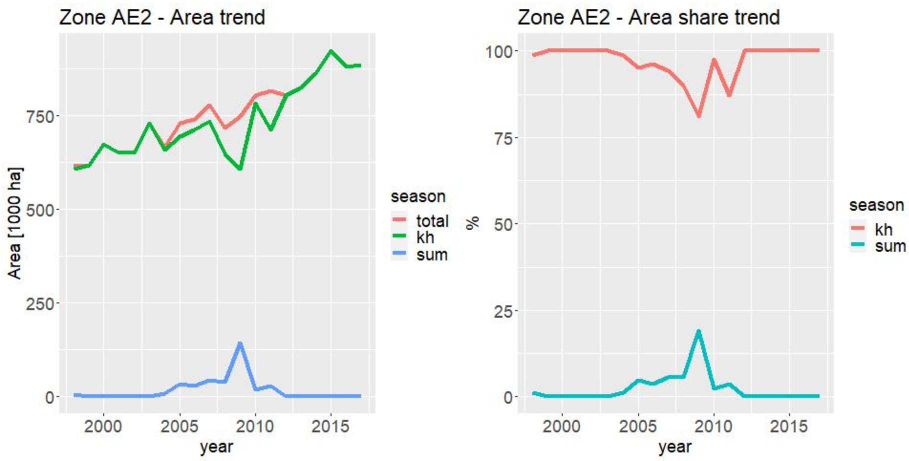

Figure S9: **Left panel**: Absolute value pearl millet cultivated area in the AE2 zone during the kharif, summer and overal (kharif + summer) seasons. **Right panel**: Relative proportion of pearl millet cultivated during the kharif and summer season.

Figure S10: Pearl millet area trend over seasons in the G zone

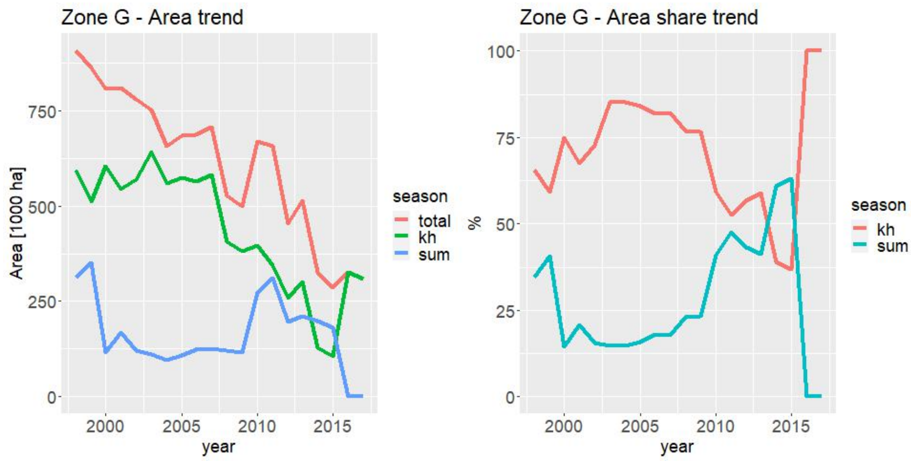

Figure S10: **Left panel**: Absolute value pearl millet cultivated area in the G zone during the kharif, summer and overal (kharif + summer) seasons. **Right panel**: Relative proportion of pearl millet cultivated during the kharif and summer season. We can notice that starting from around 2000, the proportion of area cultivated during summer increase to reach more than 50 %. The sudden drop at the end of the time series is potentially due to incomplete information about the summer season in the last years.

2 Supplementary Tables

Table S1: Reference selection of crop model parameter following the 'informed strategy'

Table S1: Crop model parameters

Parameter	A1 zone	A zone	B zone
Model	New	New	New
Soil water content [mm/cm]	0.6 (sandy)	0.9 (clay/loam)	1.3 (vertisol)
Soil depth [cm]	120 (medium)	120 (medium)	(medium)
Sowing date	1-15 July (av.)	1-15 July (av.)	1-15 July (av.)
Sowing density [pl./m^2]	12	18	18
Genotype (variety)	HHB67-2	9444	9444
Fertilisation [kg N/ha]	DLD data*	DLD data	DLD data
Irrigation	DLD data	DLD data	DLD data

*: For fertilisation and irrigation, we set the parameter value to the one that was the closest to the average level of fertilisation and irrigation as observed in the DLD data over the considered period (1998-2017).

Table S2: Influence of the crop model parameters on the crop model prediction ability

We calculated the effect of the crop model parameter on the correlation (crop model prediction ability) by doing a linear regression of the different parameter options on the correlation between the observed and predicted yield. Those results allow the assessment of which parameter has an influence on the prediction ability, so which parameter should be correctly specified to model the yield trend with a good correlation.

From this analysis, we could notice that the version of the model used had a significant impact on the prediction ability in A1 and AE1 zones. In those zones, the new model performed significantly better. The correct specification of irrigation was also very influential in A1 and AE1. Simulation with intermediate and full irrigation gave less goodness of fit. The soil water content (PAWC) was also an influential parameter, especially in AE1, B, and G. Finally, the variety specification was also influential in AE2, B and G, where the use of improved variety (HHB67 or 9444) seems to be an important parameter for good prediction ability.

Table S2: Crop model parameter influence on model goodness of fit (correlation) in the new TPE as the estimated linear regression coefficient (B) and the R squared statistics over all simulations

	A1		AE1		AE2		B		G	
var.	B	R^2	B	R^2	B	R^2	B	R^2	B	R^2
Model:new	0.00	0.14	0.00	0.12	0.00	0.03	0.00	0.08	0.00	0.06
Model:old	-0.18		-0.19		0.02		0.04		-0.07	
PAWC:1.3 (clay)	0.00	0.03	0.00	0.07	0.00	0.06	0.00	0.09	0.00	0.06
PAWC:0.9 (loam)	0.02		0.12		0.07		0.11		0.11	
PAWC:0.6 (sand)	0.04		0.14		0.06		0.12		0.11	
SoilDepth:180	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SoilDepth:120	0.00		0.00		0.00		0.00		0.00	
SoilDepth:60	0.01		0.03		0.00		0.00		0.00	
Sowing:early	0.00	0.06	0.00	0.04	0.00	0.11	0.00	0.10	0.00	0.05
Sowing:average	0.01		0.00		-0.07		-0.05		0.03	
Sowing:late	-0.01		-0.01		-0.04		-0.07		0.05	
Density:12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Density:18	0.01		0.03		0.00		0.00		0.02	
Density:24	0.02		0.04		0.00		0.00		0.03	
Variety:Landrace	0.00	0.02	0.00	0.03	0.00	0.07	0.00	0.08	0.00	0.09
Variety:HHB67	-0.02		0.02		0.12		0.11		0.09	
Variety:9444	0.01		0.07		0.09		0.13		0.15	
Fertilizer:0	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.01	0.00	0.01
Fertilizer:60	0.00		0.02		0.03		0.03		0.03	
Fertilizer:100	0.01		0.04		0.07		0.06		0.06	
Irrigation:off	0.00	0.17	0.00	0.15	0.00	0.02	0.00	0.04	0.00	0.06
Irrigation:int	-0.06		-0.06		-0.01		0.00		-0.01	
Irrigation:on	-0.20		-0.24		-0.05		-0.08		-0.12	

Table S3: Influence of the crop model parameters on predicted yield - relative sum of squares

Table S3: Crop model parameter influence on grain yield in the new TPEs as the estimated linear regression coefficient (B) and the relative parameter sum of squares compared to the dependent variable (predicted yield) sum of squares.

	A1		AE1		AE2		B		G	
Variable	B	*SS*(%)	B	*SS*(%)	B	*SS*(%)	B	*SS*(%)	B	*SS*(%)
Total rain	3.13	0.24	1.15	0.19	-0.30	0.03	-0.41	0.05	-1.14	0.06
Average max T	-274.66	0.01	-326.16	0.02	-205.19	0.01	-250.04	0.00	-410.39	0.01
Average min T	244.54	0.02	194.69	0.02	101.24	0.01	489.15	0.02	511.27	0.03
Model:new	0.00	0.04	0.00	0.10	0.00	0.21	0.00	0.34	0.00	0.24
Model:old	-310.94		-515.09		-656.03		-1050.82		-810.35	
PAWC:1.3 (clay)	0.00	0.12	0.00	0.10	0.00	0.08	0.00	0.06	0.00	0.09
PAWC:0.9 (loam)	-317.95		-212.49		-54.22		-40.73		-163.03	
PAWC:0.6 (sand)	-719.26		-712.41		-633.46		-600.34		-732.02	
SoilDepth:180	0.00	0.02	0.00	0.02	0.00	0.02	0.00	0.03	0.00	0.02
SoilDepth:120	-10.75		-4.99		-3.05		-5.17		-3.64	
SoilDepth:90	-348.39		-324.55		-297.91		-464.91		-325.53	
Sowing:early	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00
Sowing:average	-33.06		-104.43		-141.86		13.67		-107.31	
Sowing:late	-126.58		-293.17		-422.51		65.67		-179.60	
Density:12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Density:18	2.89		22.94		27.49		59.15		37.77	
Density:24	-21.50		4.50		10.95		63.99		33.25	
Variety:Landrace	0.00	0.02	0.00	0.06	0.00	0.11	0.00	0.11	0.00	0.08
Variety:HHB67	4.25		7.21		81.74		-20.67		-50.24	
Variety:9444	342.73		616.67		928.79		938.29		727.95	
Fertilizer:0	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.01
Fertilizer:60	26.86		68.38		129.68		134.92		113.70	
Fertilizer:100	41.91		117.21		236.16		241.59		208.92	
Irrigation [mm]	4.75	0.17	4.14	0.05	3.08	0.01	5.35	0.02	3.79	0.02

Table S4: Influence of the crop model parameters on predicted yield - relative estimated effect

Table S4: Crop model parameter influence on grain yield in the new TPEs as the estimated linear regression coefficient (B) and the relative range of the parameter effect compared to the predicted yield range ($Brange(\%)$).

	A1		AE1		AE2		B		G	
Variable	B	$Brange(\%)$	B	$Brange(\%)$	B	$Brange(\%)$	B	$Brange(\%)$	B	$Brange(\%)$
Total rain	3.13	0.28	1.15	0.13	-0.30	0.07	-0.41	0.17	-1.14	0.21
Average max T	-274.66	0.41	-326.16	0.53	-205.19	0.37	-250.04	0.20	-410.39	0.50
Average min T	244.54	0.37	194.69	0.26	101.24	0.13	489.15	0.22	511.27	0.35
Model:new	0.00	0.06	0.00	0.10	0.00	0.12	0.00	0.18	0.00	0.16
Model:old	-310.94		-515.09		-656.03		-1050.82		-810.35	
PAWC:1.3 (clay)	0.00	0.14	0.00	0.13	0.00	0.11	0.00	0.10	0.00	0.14
PAWC:0.9 (loam)	-317.95		-212.49		-54.22		-40.73		-163.03	
PAWC:0.6 (sand)	-719.26		-712.41		-633.46		-600.34		-732.02	
SoilDepth:180	0.00	0.07	0.00	0.06	0.00	0.05	0.00	0.08	0.00	0.06
SoilDepth:120	-10.75		-4.99		-3.05		-5.17		-3.64	
SoilDepth:90	-348.39		-324.55		-297.91		-464.91		-325.53	
Sowing:early	0.00	0.02	0.00	0.05	0.00	0.08	0.00	0.03	0.00	0.04
Sowing:average	-33.06		-104.43		-141.86		13.67		-107.31	
Sowing:late	-126.58		-293.17		-422.51		65.67		-179.60	
Density:12	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.01
Density:18	2.89		22.94		27.49		59.15		37.77	
Density:24	-21.50		4.50		10.95		63.99		33.25	
Variety:Landrace	0.00	0.07	0.00	0.11	0.00	0.17	0.00	0.17	0.00	0.14
Variety:HHB67	4.25		7.21		81.74		-20.67		-50.24	
Variety:9444	342.73		616.67		928.79		938.29		727.95	
Fertilizer:0	0.00	0.01	0.00	0.02	0.00	0.04	0.00	0.04	0.00	0.04
Fertilizer:60	26.86		68.38		129.68		134.92		113.70	
Fertilizer:100	41.91		117.21		236.16		241.59		208.92	
Irrigation	4.75	0.45	4.14	0.27	3.08	0.14	5.35	0.19	3.79	0.21