in silico Plants Vol. 3, No. 1, pp. 1–13
doi:10.1093/insilicoplants/diab008
Advance Access publication 04 February 2021
Special Issue: Linking Crop/Plant Models and Genetics
Original Article

in silico Plants
OXFORD

Evaluation of soybean (*Glycine max* L.) adaptation to northern European regions under different agro-climatic scenarios

Guénolé Boulch*,§, Chloé Elmerich, Amina Djemel and Bastien Lange§

Institut Polytechnique UniLaSalle, AGHYLE, 60000 Beauvais, France
*Corresponding author's e-mail address: guenole.boulch@unilasalle.fr
§Equally contributing authors.

Guest Editor: Vincent Vadez
Editor-in-Chief: Stephen P. Long

Citation: Boulch G, Elmerich C, Djemel A, Lange B. 2021. Evaluation of soybean (*Glycine max*) adaptation to northern European regions under different agro-climatic scenarios. *In Silico Plants* **2021**: diab008; doi: 10.1093/insilicoplants/diab008

ABSTRACT

Soybean is a candidate crop to increase the independency of Europe in leguminous protein crops. However, its adaptation to northern European regions is not yet well defined due to the lack of long-term references. Herein, we simulated soybean yield potential in northern France and identified the main yield limiting factors under rainfed vs. irrigated conditions. Two cultivars representing maturity groups 000 and 00 were planted within three different trials. Leaf area index, shoot and pod biomass, main phenological stages and yield were recorded to evaluate CROPGRO-soybean model predictability. Adjustment of genetic coefficients was performed prior to simulate yield on 21-years weather database (1999–2018) at Beauvais (France, N 49.46°, E 2.07°) and Estrées-Mons (France, N 49.88°, E 3.01°) under different water regimes and planting dates. Predictions showed that adding irrigation at grain filling period would increase yield potential to the level of non-water limited scenarios. Although simulated yield variability is reduced with irrigation, the remaining variability suggests that water is not the only yield-limiting factor. A tentative explanation is proposed by deriving environmental covariates from the model. The analysis confirmed the importance of precipitation amount (optimum around 200 mm) and duration (optimum around 60 days) of the flowering to physiological maturity period under rainfed conditions. Under irrigated conditions, increasing evapotranspiration and average minimum temperature affected simulated yield positively while increasing the number of days below 10 °C had a negative impact. These results give insights for soybean crop management and bring indications to breeders for adapting the existing genetic material to northern Europe.

KEYWORDS: CROPGRO; crop model; DSSAT; long-term simulation; model calibration; protein crop; soybean breeding; water management.

1. INTRODUCTION

Increasing the production of protein crops in Europe for both food and feed usages is becoming a growing priority for three main reasons: being more independent from imports, decreasing usage of nitrates and pesticides, giving new options to farmers (Plaza-Bonilla *et al.* 2017; Marraccini *et al.* 2020). Among the leguminous crops responding potentially to the demand, soybean (*Glycine max*) is a serious candidate. To date, soybean is very little cultivated in northern Europe, and northern France, despite a rising interest from farmers. The potential of the crop in these regions is not yet well defined and long-term cropping historical data are needed. Several studies have shown that cool temperatures (<15 °C), low solar radiations and low water availability along the crop cycle may affect soybean growth and development (Zhang *et al.* 1996; Krüger *et al.* 2014; Kumagai and Sameshima 2014; Giménez *et al.* 2017).

According to the Köppen-Geiger World climate classification, northern France belongs to a Cfb type of climate: marine temperate without dry seasons, warm summers and mild winters (Peel *et al.* 2007; Beck *et al.* 2018). This cluster includes a large part of France, UK, Belgium, Western Germany, the Netherlands and part of Denmark. Another classification has been proposed for Europe where the region falls under a Maritime zone which includes additional countries like Switzerland, Sweden, Ireland, part of Austria (Bouma 2005). Based on 20 years of Agri4cast data, the average temperature in Northern France is 17.11 °C for soybean growing season (1 May to mid-September) with minimum averages of 12.4 °C and maximum of 21.84 °C, a solar radiation of 17.49 Mj/m² and a cumulative precipitation around 276 mm. The latitude between 48 and 51° North is impacting daylength and the crop is submitted for 2 months at daylength greater than 15.5 h.

The use of Crop Growth Models (CGM) is now well adopted to run simulations, unless there is a minimal calibration effort done in the field, for testing crop in specific pedoclimatic contexts (Boote *et al.* 2016; Maiorano *et al.* 2017). Such *in silico* references can help develop strategies to maximize yield potential (Asseng *et al.* 2003; Chenu *et al.* 2017). Furthermore, analysis of data from CGM simulations can help identify traits of importance to work in a breeding program for improving crop adaptation (Chapman *et al.* 2002; Hammer *et al.* 2002; Chapman 2008; Chenu *et al.* 2011).

Among soybean CGM, CROPGRO has been calibrated for North America under well-watered conditions (Boote *et al.* 1998). In 2001, a calibration of CROPGRO-soybean model under rainfed and cool conditions was performed in Spain (Ruíz-Nogueira *et al.* 2001). In their study, adjusting soil depth, water holding capacity and root elongation rate improved phenology and yield predictability prior to perform long-term simulations under irrigated and rainfed conditions. Then, identification of environmental parameters impacting yield variability is critical to develop agronomic and breeding strategies for adapting soybean to northern European conditions (Board and Kahlon 2011). The objectives of this study are to (i) simulate soybean yield potential in northern France after crop model calibration and to (ii) identify the main soybean yield-limiting factors under rainfed vs. irrigated conditions. The expectations are to help farmers adopting this new crop by providing information that can overcome the lack of long-term soybean cropping references through calibration efforts and simulations under latitudes higher than 49° North.

2. MATERIALS AND METHODS

2.1 Validation of model predictability with experimental data

2.1.1 ***Experimental design.*** Two soybean (*Glycine max* L.) cultivars: RGT Sirelia (MG 000) and ES Mentor (MG 00) were planted within three rainfed trials in 2019, at the experimental farms of the Institut Polytechnique UniLaSalle (Beauvais, France) and ANTEDIS company (Catillon-Fumechon, France). Daily weather data were recorded for the two locations: minimum and maximum temperatures, precipitations and solar radiation. Soil was characterized at each experimental site by analysing 15 cm layer samples along the soil profile (0–90 cm, 3 replicates) (Table 1). Two planting dates were managed at Beauvais (April 24 and May 17) and one at Catillon-Fumechon (April 23). Planting was carried out with a single seed air planter and a density of 70 plants m^{-2} (0.45 m row spacing). Experimental plots (10 m × 2.7 m) were replicated five times and assigned to a randomized complete block design. Before planting, seeds were inoculated with *Bradyrhizobium japonicum* (G49 strain).

Table 1. Location and soil description at experimental sites.

	Beauvais (France)						Catillon-Fumechon (France)					
Coordinates (WGS84 DD)	Lat 49.465 , Long 2.072						Lat 49.544 , Long 2.388					
Elevation (m)	112						135					
Soil depth (cm)	120						120					
Soil layer (cm) (*n* = 3)	**0–15**	**15–30**	**30–45**	**45–60**	**60–75**	**75–90**	**0–15**	**15–30**	**30–45**	**45–60**	**60–75**	**75–90**
Clay (%)	20.0	20.1	18.5	26.3	26.5	21.3	20.4	20.9	22.9	22.5	21.1	21.0
Silt (%)	70.3	70.1	68.2	63.0	61.6	63.1	72.5	68.3	66.0	71.0	71.3	72.2
Initial residual H_20 (%)	1.1	1.0	1.3	1.6	1.7	1.7	1.6	1.4	1.5	1.5	1.5	1.5
Total organic carbon (%)	1.8	1.8	0.5	0.4	0.3	0.3	0.97	0.95	0.41	0.29	0.25	0.26
pH (water)	7.9	7.9	8.4	8.4	8.4	8.5	8.40	8.35	8.40	8.45	8.40	8.40
CEC (cmol kg^{-1})	10.9	11.0	10.9	13.2	14.9	13.5	13.2	13.3	12.8	12.7	12.8	12.1
N-NO3 (mg kg^{-1})	12.6	14.9	6.4	4.3	4.0	3.2	16.7	16.6	9.3	5.7	5.2	6.8
N-NH4 (mg kg^{-1})	51.6	45.6	40.9	43.3	50.4	43.4	66.2	52.4	34.1	33.8	26.0	27.6
N-organic (mg kg^{-1})	1117	1037	613	517	490	440	1100	1055	580	435	390	390
Exchangeable K (mg kg^{-1})	147	157	146	172	180	167	231	256	169	144	136	135
Olsen P (mg kg^{-1})	56.3	57.0	29.7	16.7	17.7	18.0	105	109	59.5	43.0	41.0	46.5

2.1.2 *Growth and phenology measurements.* Plants were sampled (0.45 m^2) at 15- to 20-day intervals throughout the crop life cycle, starting at emergence +20 days. Plants were separated into leaves, stems (petioles included) and pods. Cumulative fresh leaf area was measured in each sample using LI-3100C Area Meter instrument (LI-COR®) and used to determine the Leaf Area Index (LAI). Plant leaves, stems and pods were dried at 65 °C for 72h and weighed. Crop development was observed twice a week to identify growth stages VE, V2, R1, R3 and R7 according to Fehr and Caviness (Fehr and Caviness 1977) staging method. At harvest maturity, seed dry yield, shoot dry biomass and Harvest Index (HI) were measured for each plot after 24 h at 105 °C (Battisti 2016).

2.1.3 *Crop model simulations.* CROPGRO-soybean model – DSSAT v4.7.5 (Decision Support System for Agrotechnology Transfer) (Boote *et al.* 1998, 2003; Jones *et al.* 2003; Hoogenboom *et al.* 2019) was chosen for simulations. The Willmott agreement index (*d*-stat) (Willmott *et al.* 2012) and root mean square error (RMSE) were used to evaluate simulations accuracy (Willmott 1982). A model calibration was performed using the in-season growth and development field measurements. Cultivar coefficients of the two generic cultivars MG 000 and MG 00 available in the model were manually adjusted. The calibration procedure started with phenology first and then growth parameters (Hunt and Boote 1998). The coefficients were kept within the typical range values for soybean (Boote *et al.* 2003) and validated on the basis of maximizing the *d*-stat values and minimizing the RMSE for each time-series growth trait.

2.2 Evaluation of agronomic suitability and yield potential of soybean

To evaluate yield potential and agronomic suitability, two sites representing contrasted environments of Northern France were selected. One was Beauvais (same site used for calibration) and the other one was Estrées-Mons, France (Lat. 49.878, Long. -3.0067). It is expected that our study sites have a prediction power for regions beyond our scope. A weather database including daily minimum and maximum temperature, solar radiation (MJ/m^2), precipitation (mm) and potential evapotranspiration (mm) was assembled from 1999 to 2019 for both sites. Two soils representing each location (deep silt) were selected for simulations. Two agronomic factors were evaluated to assess yield potential of the calibrated cultivars RGT Sirelia (MG 000) and ES Mentor (MG 00): the planting date and water regime. CROPGRO-soybean model has demonstrated its value in evaluating yield response to planting date and irrigation scenarios (Salmerón *et al.* 2017; Yassi *et al.* 2019). A series of planting dates were tested in a first time (April 20, April 30, May 10, May 20 and May 30) to select the optimal and latest possible ones: April 30 and May 20, respectively. The optimal date maximized the average simulated grain yield and minimized its variability across the 21 years of simulations. The latest possible date ensured a harvest maturity date compatible with a secure winter wheat planting. These two dates will be the two planting scenarios used in our study. Two water regimes were tested: rainfed and irrigated. The irrigated regime, compatible with regional agronomic practices, was determined by minimizing the simulated yield difference with the non-water limited simulation option of the model: 3 irrigation amounts of 30 mm at 60, 80 and 100 days after planting. All simulations were performed considering a planting density of 70 seeds per square metre and a row spacing of 0.45 m.

2.3 Identification of environmental covariates

2.3.1 *Environmental variables calculation.* Based on the individual year simulations, 35 environmental covariates were calculated to identify key factors explaining year to year yield variation. The variables were defined around three development stages that are simulated in DSSAT seasonal analysis: sowing to emergence (SEM), emergence to flowering (EMFL), flowering to physiological maturity (FLPM). Four major categories of variables were identified: duration of stages (in number of days), temperature variables (average minimum temperatures, number of days below 10 °C, number of days below 15 °C, average maximum temperatures, number of days above 30 °C, number of days above 34 °C), water variables (cumulative rainfall and irrigation in mm, potential evapotranspiration in mm), solar radiation (cumulative daily solar radiation, average of solar radiation in MJ/m^2), photothermal quotient (defined as the ratio of solar radiation on heat units).

2.3.2 *Statistical analysis.* After generating the environmental variables, two tables were created. The X-table contained the year/location combinations with all the calculated environmental covariates and the Y-table represented the simulated yield values, estimated by the model for each year/location combination. Statistical models are available to sort out the most relevant variables explaining yield variation. Among them, bilinear models such as Partial Least Square Regression (PLS) models seemed particularly appropriate (Crossa *et al.* 2010). One of the advantages of PLS over linear factor regression models is the possibility to evaluate a high number of covariables allowing to integrate as many as possible environmental variables and so, limiting a priori statements on the factors to include or not in the analysis. There are adapted versions of PLS to address different objectives (regression, classification, variable selection, and survival analysis) (Mehmood and Ahmed 2016). We used the variable selection methods accepting that X-matrix contains redundant or irrelevant variables without impacting the results (Mehmood *et al.* 2012). Among these variable selection methods, we used the filter method taking VIP (Variable Important in the Projection) as a selection criterion. Variables with a VIP>1 are the ones that are considered as able to explain the Y-table. The number of components were defined based on the Wold algorithm (Wold et al. 1984, 1987, 2001; Tenenhaus *et al.* 2005). Analysis were performed under R software.

3. RESULTS

3.1 Model performance

Crop stages were well simulated by using generic MG 000 and MG 00 cultivar coefficients, especially for RGT Sirelia (Table 2). For ES Mentor, simulations were site-dependent for phenology and adjustment of phenological trait coefficients did not consistently improve the predictions across sites. For both genotypes, physiological maturity was accurately predicted (with a maximum of 3 days difference). Therefore, the generic values of phenological trait coefficients were

Table 2. Prediction of main phenological stages for RGT Sirelia (MG 000) and ES Mentor (MG 00) using CROPGRO-soybean model generic cultivars. Differences (number of days) between observed and simulated stages (days after planting) are given for each trial and cultivar. Three trials distributed across two sites are considered: Beauvais, France (two planting dates) and Catillon-Fumechon, France.

Experimental sites	Beauvais (France)			Beauvais (France)			Catillon-Fumechon (France)		
Coordinates (WGS84 DD)	Lat 49.465 , Long 2.072			Lat 49.465 , Long 2.072			Lat 49.544 , Long 2.388		
Planting date	**24-04-2019**			**17-05-2019**			**23-04-2019**		
RGT SIRELIA (MG 000)	Simulated	Observed	Difference	Simulated	Observed	Difference	Simulated	Observed	Difference
Emergence day (dap)	17	26	9	10	10	0	16	22	6
Anthesis day (dap)	64	63	-1	49	46	-3	66	64	-2
First pod day (dap)	73	78	5	58	61	3	75	71	-4
First seed day (dap)	87	85	-2	71	77	6	89	88	-1
Physiological maturity (dap)	121	124	3	106	107	1	124	125	1
ES MENTOR (MG 00)									
Emergence day (dap)	17	26	9	10	10	0	16	22	6
Anthesis day (dap)	67	64	-3	53	47	-6	69	63	-6
First pod day (dap)	78	76	-2	63	61	-2	80	71	-9
First seed day (dap)	92	87	-5	76	77	1	93	84	-9
Physiological maturity (dap)	127	129	2	110	110	0	128	130	2

kept for running long-term simulations. For RGT Sirelia and ES Mentor, adjustment of two growth trait coefficients: maximum leaf photosynthesis rate at 30 °C (+0.10 mg CO_2 m^{-2} s^{-1}) and maximum size of full leaf (three leaflets) (+10 cm^2) consistently improved model predictability for time-series growth traits at Beauvais (Table 3). Phenology predictions were not affected by such adjustments (data not shown). Yield prediction was globally satisfactory and improved after calibration, except for the first planting date at Beauvais and for RGT Sirelia at Catillon-Fumechon. The calibration did not necessarily improve model performance for predicting time-series LAI and leaf weight at Catillon-Fumechon for both cultivars. However, values of Willmott agreement index for these two traits remained close to the upper limit of 1 (*d*-stat = 0.951 and 0.922 for RGT Sirelia and *d*-stat = 0.846 and 0.824 for ES Mentor, respectively, for LAI and leaf weight).

3.2 Yield potential under rainfed conditions

At Beauvais site, considering the two planting dates under a rainfed scenario, simulated grain yield reached 2716 ± 652 kg ha^{-1} for RGT Sirelia (MG 000) and 2878 ± 615 kg ha^{-1} for ES Mentor (MG 00) (Fig. 1). At Estrées-Mons site, simulated grain yield reached 2982 ± 402 kg ha^{-1} for RGT Sirelia (MG 000) and 3184 ± 410 kg ha^{-1} for ES Mentor (MG 00). For both maturity group, the planting date had no significant effect on simulated grain yield, neither at Beauvais site nor at Estrées-Mons. Under rainfed conditions at Estrées-Mons site, simulated grain yield was slightly higher for ES Mentor (3272 ± 418 kg ha^{-1}) compared with RGT Sirelia (3006 ± 433 kg ha^{-1}), planted on 30 April (W = 307, $P < 0.05$). No significant differences in simulated grain yield was observed between the two maturity groups in any other treatments.

3.3 Variables driving rainfed yield

PLS analysis sorted the 35 environmental covariates from the highest VIP to the lowest ones relative to simulated yields under rainfed conditions in the two locations, on 21 years (1999–2019) for RGT Sirelia (MG 000) and ES Mentor (MG 00). Table 4 displays the variables that show VIP greater than 1. First of all, the 3 first variables with the highest VIP concerned only the period from flowering to the end of grain filling period. The main variable impacting simulated yields was the amount of precipitation accumulated from flowering to physiological maturity with a very high VIP (2.93) compared with all the other variables. This variable tended to follow a logarithmic function with an R^2 coefficient of 0.45 (Fig. 2). Then, the average minimum temperature in the same period appeared to influence simulated yield with a high VIP of 1.90 and a positive β-coefficient of 0.40, meaning that increasing the average minimum temperature at grain filling has a positive impact on simulated yield level. Then, the duration of the grain filling period was showing a VIP of 1.87 with a positive β-coefficient of 0.30, meaning that an extension of this period has a positive impact on simulated yield level. This variable seems to follow a polynomial trend (R_2 = 0.35) (Fig. 3) with an optimum around 60 days. The other variables showed less impact based on their VIPs.

3.4 Yield potential under irrigation

Yield response to irrigation scenario (3 irrigation paths of 30 mm at 60, 80 and 100 days after planting) was the same for the two cultivars RGT Sirelia (MG 000) and ES Mentor (MG 00) at any location and planting date (data not shown). At Beauvais site, adding irrigation enhanced simulated grain yield by +498 kg ha^{-1} and +467 kg ha^{-1}, respectively, for the planting dates 30 April and 20 May (W = 1287, $P < 0.001$ and W = 1319, $P < 0.001$) (Fig. 4). At Estrées-Mons site, the positive effect

Table 3. **CROPGRO-soybean growth predictability for cultivars RGT Sirelia (MG 000) and ES Mentor (MG 00) after manual adjustment of growth trait coefficients. Three trials distributed across two sites are considered: Beauvais, France (two planting dates) and Catillon-Fumechon, France. Template cultivars were the generic MG 000 and MG 00 available in the model. The calibration consisted in increasing maximum leaf photosynthesis rate at 30 °C (+0.10 mg CO_2 m^{-2} s^{-1}) and maximum size of full leaf (three leaflets) (+10 cm^2) coefficients for both RGT Sirelia (MG 000) and ES Mentor (MG 00). The values of all other coefficients remain the same as the generic. Model performance (prior and after calibration) on time series data prediction is given by the root mean square error (*RMSE*) and Willmott agreement index (*d*-stat) values.**

Experimental sites	**Beauvais (France)**			**Beauvais (France)**			**Catillon-Fumechon (France)**		
Coordinates (WGS84 DD)	Lat 49.465 , Long 2.072			Lat 49.465 , Long 2.072			Lat 49.544 , Long 2.388		
Planting date	**24-04-2019**			**17-05-2019**			**23-04-2019**		
	RGT Sirelia (MG 000)								
	Observed	**Simulated**		**Observed**	**Simulated**		**Observed**	**Simulated**	
		Generic MG 000	**Calibrated cultivar**		**Generic MG 000**	**Calibrated cultivar**		**Generic MG 000**	**Calibrated cultivar**
Yield at harvest maturity (kg [dw] ha^{-1})	**1969**	**1929**	**2278**	**2758**	**2095**	**2339**	**2430**	**2777**	**2873**
Tops weight at maturity (kg [dw] ha^{-1})	**4475**	**3447**	**4077**	**6475**	**3744**	**4223**	**4846**	**5583**	**5850**
Leaf area index, maximum	**3.43**	**2.36**	**2.70**	**3.18**	**2.33**	**2.68**	**3.21**	**3.62**	**3.93**
Harvest index at maturity	**0.44**	**0.56**	**0.56**	**0.43**	**0.56**	**0.55**	**0.50**	**0.50**	**0.49**

	Generic MG 000		**Calibrated cultivar**		**Generic MG 000**		**Calibrated cultivar**		**Generic MG 000**		**Calibrated cultivar**	
Time-series growth traits :	*RMSE*	*d*-stat	*RMSE*	*d*-stat	*RMSE*	*d*-stat	*RMSE*	*d*-stat	*RMSE*	*d*-stat	*RMSE*	*d*-stat
Leaf area index	**0.563**	**0.932**	**0.427**	**0.965**	**0.835**	**0.850**	**0.669**	**0.911**	**0.460**	**0.957**	**0.521**	**0.951**
Leaf weight (kg [dw] ha^{-1})	**250**	**0.947**	**189**	**0.972**	**488**	**0.829**	**423**	**0.879**	**279**	**0.932**	**317**	**0.922**
Tops weight (kg [dw] ha^{-1})	**1301**	**0.866**	**964**	**0.932**	**1776**	**0.861**	**1423**	**0.915**	**382**	**0.995**	**213**	**0.998**
Pod weight (kg [dw] ha^{-1})	**1223**	**0.851**	**864**	**0.932**	**999**	**0.906**	**725**	**0.953**	**607**	**0.982**	**458**	**0.990**
Pod number (no m^{-2})	**251**	**0.086**	**173**	**0.135**	**210**	**0.871**	**138**	**0.943**	**161**	**0.932**	**143**	**0.947**

Table 3. Continued

ES Mentor (MG 00)

	Observed	Simulated		Observed	Simulated		Observed	Simulated	
		Generic MG 00	Calibrated cultivar		Generic MG 00	Calibrated cultivar		Generic MG 00	Calibrated cultivar
Yield at harvest maturity (kg [dw] ha^{-1})	**1630**	**2197**	**2315**	**2631**	**1948**	**2013**	**2432**	**2316**	**2406**
Tops weight at maturity (kg [dw] ha^{-1})	**4017**	**4370**	**4560**	**6253**	**4107**	**4312**	**4626**	**5490**	**5717**
Leaf area index, maximum	**2.79**	**2.91**	**3.00**	**3.39**	**2.92**	**3.07**	**3.15**	**4.37**	**4.70**
Harvest index at maturity	**0.40**	**0.50**	**0.51**	**0.42**	**0.47**	**0.47**	**0.53**	**0.42**	**0.42**

	Generic MG 00		Calibrated cultivar		Generic MG 00		Calibrated cultivar		Generic MG 00		Calibrated cultivar	
Time-series growth traits :	*RMSE*	*d*-stat	*RMSE*	*d*-stat	*RMSE*	*d*-stat	*RMSE*	*d*-stat	*RMSE*	*d*-stat	*RMSE*	*d*-stat
Leaf area index	**0.283**	**0.980**	**0.245**	**0.986**	**0.716**	**0.899**	**0.676**	**0.913**	**0.909**	**0.855**	**0.993**	**0.846**
Leaf weight (kg [dw] ha^{-1})	**155**	**0.974**	**138**	**0.981**	**498**	**0.840**	**461**	**0.868**	**456**	**0.834**	**496**	**0.824**
Tops weight (kg [dw] ha^{-1})	**723**	**0.964**	**618**	**0.974**	**1404**	**0.926**	**1209**	**0.946**	**413**	**0.994**	**476**	**0.993**
Pod weight (kg [dw] ha^{-1})	**706**	**0.962**	**617**	**0.972**	**850**	**0.933**	**777**	**0.945**	**420**	**0.989**	**323**	**0.994**
Pod number (no m^{-2})	**319**	**0.296**	**290**	**0.319**	**314**	**0.810**	**284**	**0.848**	**456**	**0.419**	**453**	**0.424**

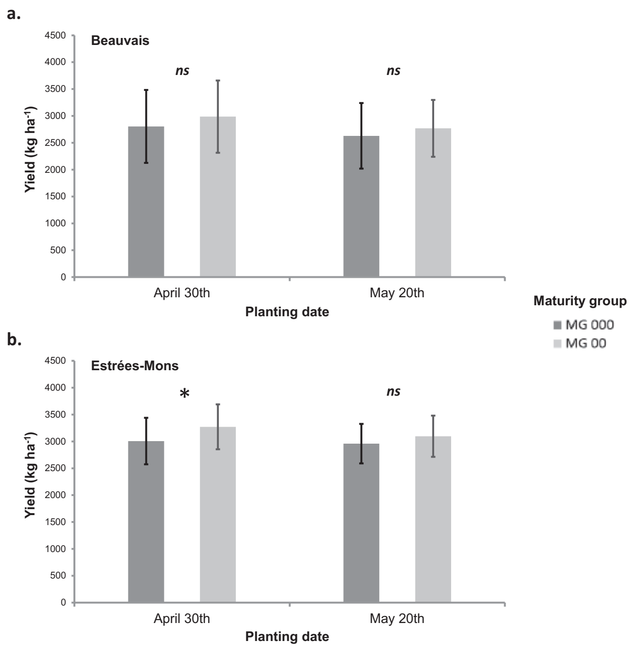

Figure 1. Simulated soybean grain yield under rainfed conditions for two maturity groups considering two locations and two planting dates. (A) Beauvais, France (Lat. 49.465, Long. 2.072). (B) Estrées-Mons, France (Lat. 49.878, Long. -3.0067). Simulations were performed on calibrated genotypes: RGT Sirelia (MG000) and ES Mentor (MG00) using 21 years of weather data recorded at the two locations (1999–2019). Mean (*n* = 21) comparison symbols are result of a two-sample non-parametric Wilcoxon rank test (95 % level of confidence).

of irrigation was only observed for the planting date 30 April (+249 kg ha^{-1}, W = 1147, $P < 0.05$). A decrease of simulated yield variability was obtained by adding irrigation at Beauvais (around 300 kg ha^{-1}) and Estrées-Mons (around 70 kg ha^{-1}) sites for the two planting dates. Considering the 21 years of simulation at Beauvais site, the irrigation scenario raised the lowest yield values from 1126 to 2450 and from 1273 to 2525 kg ha^{-1}, respectively, for the two planting dates.

3.5 Variables driving yield variation under irrigation

PLS analysis sorted the 35 environmental covariates from the highest VIP to the lowest ones in regards of simulated yields under irrigation conditions in the two locations, on 21 years (1999–2019) for RGT Sirelia (MG 000) and ES Mentor (MG 00). Table 5 is displaying the variables that show a VIP greater than 1. The precipitation accumulation during the grain filling period, that was the strongest factor in rainfed conditions, did not appear anymore in the irrigated scenario. Instead, the potential evapotranspiration factor during the same period going from flowering to physiological maturity was showing up as a major driver of yield levels (highest VIP of 1.77), suggesting that the increase of evapotranspiration leads to higher yields (β-coefficient of +0.24). The minimum average temperature in the same period was, as in the rainfed scenario, a key factor (VIP = 1.64). This factor seemed to correlate with the cold stress factor defined as a number of days below 10 °C during the grain filling period. This cold stress variable shown a negative β-coefficient (-0.18) meaning that increasing the number of days below 10 °C during grain filling is detrimental to yield. An additional factor concerning the vegetative period is showing a VIP value of 1.55: heat stress factor defined as the number of days above 34 °C. This factor had a negative β-coefficient (-0.25) meaning that increasing of heat stress has a negative impact on yield potential. It is interesting to note that this factor was also detected in the rainfed scenario but compared with the 3 main factors, its VIP value was much lower.

4. DISCUSSION

4.1 Soybean yield potential and impact of adding irrigation

Soybean is known to be sensitive to water deficit from vegetative to grain filling period (Karam *et al.* 2005; Giménez *et al.* 2017; Nemeskéri and Helyes 2019). Several studies have simulated the positive effect of water supply to soybean crop under tropical and sub-tropical climate (Bhatia *et al.* 2008; Grassini *et al.* 2015; Sharda *et al.* 2019). In our region, yield gap between irrigated and rainfed treatments appeared to be limited globally, although it was higher at Beauvais site (+450 kg ha^{-1}) than at Estrées-Mons (+250 kg ha^{-1}). While the total amount of water received at both sites during the crop cycle was comparable (222 and 228 mm for Beauvais and Estrées-Mons, respectively), it cannot explain the low differences between irrigated and not irrigated scenarios at Estrées-Mons. A more detailed analysis of the water distribution along the cycle shows that the difference in water supply occurs between first pod and physiological maturity (70 and 77 mm for Beauvais and Estrées-Mons, respectively) while between first flower and first pod, the precipitation average is similar for both sites (44 mm). This difference translated into a higher average simulated water stress index calculated at Beauvais (0.50 ± 0.22) compared with Estrées-Mons (0.43 ± 0.23) under rainfed conditions. Additionally, the level of stress can also be linked to the potential evapotranspiration that is higher at Beauvais compared with Estrées-Mons for both emergence to flowering (173 ± 22 mm and 158 ± 21 mm, respectively) and flowering to physiological maturity (230 ± 20 mm and 219 ± 21 mm, respectively) periods. We propose to investigate soybean response to water stress at various levels of irrigation in the region, by implementing multi-year field experiments on MG 000 and MG 00 cultivars.

4.2 Variables impacting soybean yield under rainfed conditions

The PLS results concerning the rainfed simulations underlined the strong impact of water supply during the grain filling period for the yield elaboration under our Northern France conditions. Although, the region is often considered as a humid region, water appears to be the main limiting factor of the production as in general water is not well distributed along the crop cycle. Under Hungarian conditions, it was observed that when whole season precipitation is lower than 300 mm, yield is penalized (Gajić *et al.* 2018; Anda *et al.* 2020) and the drought stress at vegetative stage does not affect yield (Gajić *et al.* 2018). This agrees with our precipitation data calculated on the growing season

Table 4. Environmental variables with a VIP>1 sorted by descending VIP for simulated rainfed yield conditions after PLS analysis. PLS analysis was performed on two genotypes: RGT Sirelia (MG000) and ES Mentor (MG00), considering two locations: Beauvais, France (*n* = 84) and Estrées-Mons, France (*n* = 84) and 21 years of weather data (1999–2019). The statistical model estimated that the 5 components option was the optimal solution based on RMSE and cross-validation. β-regression coefficients are indicating whether the considered variable has a positive or negative effect on yield.

Variables	Description	VIP	β coefficient	Std. err.	t-value	*P*-value
PREDsumFLPM	Sum (mm) of daily precipitation from FLowering to Physiological Maturity	2.93	0.31	0.054	5.72	0.000
TMNMNFLPM	Mean of daily minimum temperature (in °C) from FLowering to Physiological Maturity	1.90	0.40	0.057	6.92	0.000
DurationFLMA	Duration (#days) of FLowering to physiological MAturity period	1.87	0.30	0.053	5.69	0.000
TMX34EMFL	Number of days above 34 °C from EMergence to FLowering	1.51	-0.30	0.068	-4.43	0.000
TMX30EMFL	Number of days above 30 °C from EMergence to FLowering	1.21	0.00	0.079	-0.04	0.965
PTQFLPM	Photothermal Quotient from FLowering to Physiological Maturity	1.07	-0.25	0.044	-5.71	0.000
TMN10FLPM	Number of days below 10 °C from FLowering to Physiological Maturity	1.07	-0.05	0.075	-0.68	0.496
PREDsumEMFL	Sum (mm) of daily precipitation from EMergence to FLowering	1.06	0.30	0.081	3.70	0.000
DurationSEM	Duration (#days) from Sowing to EMergence	1.05	0.13	0.070	1.86	0.064
TMX34FLPM	Number of days above 34 °C from FLowering to Physiological Maturity	1.05	-0.22	0.079	-2.76	0.006

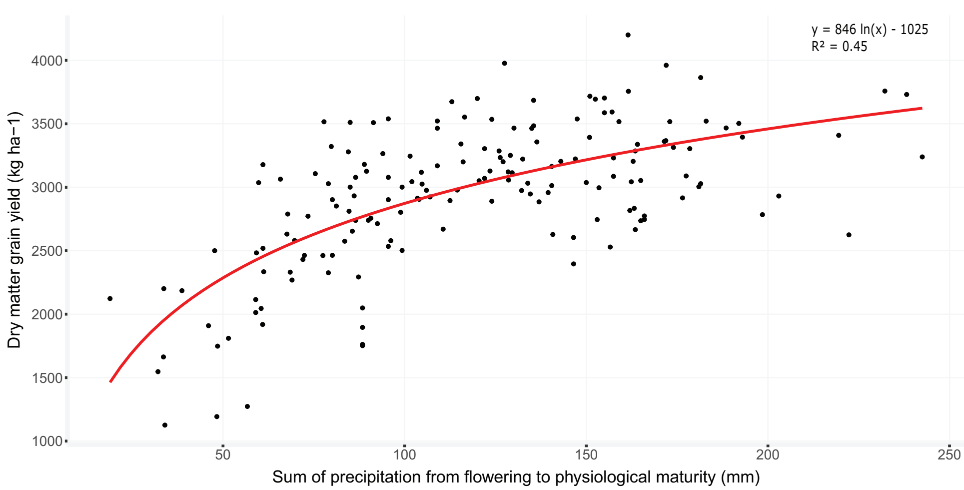

Figure 2. Relationship between simulated soybean grain yield under rainfed conditions and the sum of precipitations received from flowering to physiological maturity. Simulations were performed on two genotypes: RGT Sirelia (MG000) and ES Mentor (MG00), considering two locations: Beauvais, France (*n* = 84) and Estrées-Mons, France (*n* = 84) and 21 years of weather data (1999–2019) registered at each location.

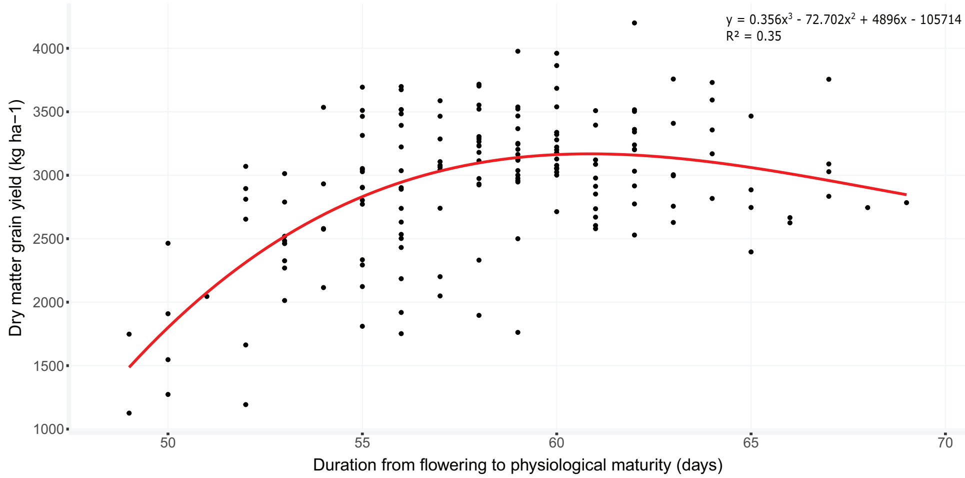

Figure 3. Relationship between simulated soybean grain yield under rainfed conditions and the duration from flowering to physiological maturity. Simulations were performed on two genotypes: RGT Sirelia (MG000) and ES Mentor (MG00), considering two locations: Beauvais, France (n = 84) and Estrées-Mons, France (n = 84) and 21 years of weather data (1999–2019) registered at each location.

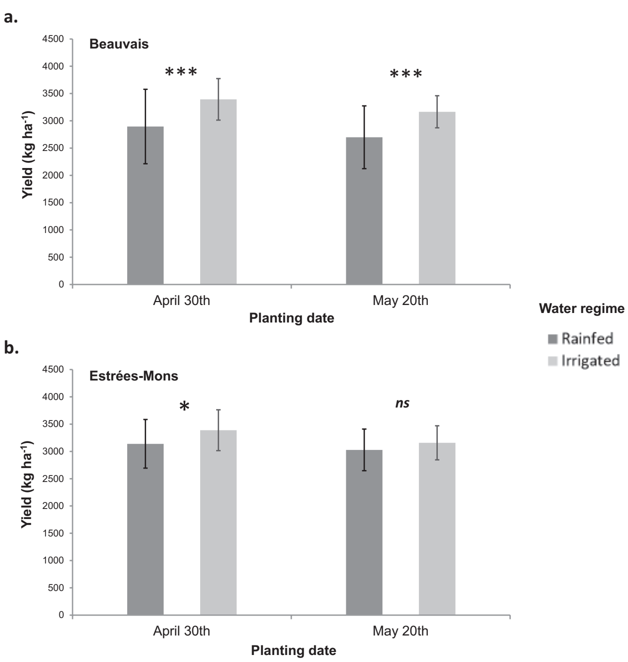

Figure 4. Simulated early maturity soybean grain yield under two contrasted water regimes considering two locations and two planting dates. (A) Beauvais, France (Lat. 49.465, Long. 2.072). (B) Estrées-Mons, France (Lat. 49.878, Long. -3.0067). Simulations were performed on calibrated genotypes: RGT Sirelia (MG000) and ES Mentor (MG00) considered together and 21 years (1999–2019) of weather data were considered for the two locations. The irrigated water regime consisted in 3 irrigation paths of 30 mm at 60, 80 and 100 days after planting. Mean (n = 42) comparison symbols are result of a two-sample non-parametric Wilcoxon rank test (95 % level of confidence).

from emergence to physiological maturity, the 21 years average for the season precipitation is 225 mm for both sites with the following partitioning: 21 mm from planting to crop emergence, 86 mm from emergence to flowering stage and 118 mm from flowering stage to physiological maturity. The factor precipitation alone seems to follow a logarithmic function and the optimum level of precipitation would be around 200 mm, which is close to the average potential evapotranspiration estimated by the model: 224 mm. This result supports the decision to evaluate the irrigation option as a way to increase yield potential in northern regions. Furthermore, this factor underlines the need for improved water stress tolerant germplasm to widely expand the crop.

The second factor affecting yield is the minimum temperature average from flowering to physiological maturity. The average value based on the simulation is around 13 °C (from 11.6 to 15 °C). The β-coefficient shows that the increase of the minimum temperature average is favourable to yield. The average value of 13 °C is below the threshold of 15 °C which is considered as the level below which optimal growing conditions are not reached and where soybean starts to suffer from cool temperatures (Baker *et al.* 1989; Funatsuki *et al.* 2004); as we are testing adaptation for high latitudes, it seems not surprising that this factor is appearing. Evaluating cultivars that are able to maintain a superior level of photosynthetic activity under this level of chilling stress could be an area of investigation to develop better adapted germplasm. On the other hand, the maximum temperature average is not appearing as a critical factor. Maximum temperature average is around 24.6 °C (ranging from 21.5 °C to 28.2 °C) which is not limiting for the crop according to the literature (Caulfield and Bunce 1988; Baker *et al.* 1989).

The third factor concerns the duration of the flowering to maturity period. In our context, the duration of the flowering to physiological maturity period strongly correlates with the following variables: number of days where minimum temperatures are

Table 5. Environmental variables with a VIP>1 sorted by descending VIP for simulated irrigated yield conditions after PLS analysis. PLS analysis was performed on two genotypes: RGT Sirelia (MG000) and ES Mentor (MG00), considering two locations: Beauvais, France (n = 84) and Estrées-Mons, France (n = 84) and 21 years of weather data (1999–2019). The statistical model estimated that the 3 components option was the optimal solution based on RMSE and cross-validation. β-regression coefficients are indicating whether the considered variable has a positive or negative effect on yield.

Variables	Description	VIP	β-coefficient	Std. err.	*t*-value	*P*-value
ETPsumFLPM	Sum (mm) of daily Evapotranspiration from FLowering to Physiological Maturity	1.77	0.24	0.050	4.83	0.000
TMNMNFLPM	Mean of daily minimum temperature (in °C) from FLowering to Physiological Maturity	1.64	0.20	0.038	5.26	0.000
TMN10FLPM	Number of days below 10 °C from FLowering to Physiological Maturity	1.60	-0.18	0.042	-4.35	0.000
TMX34EMFL	Number of days above 34 °C from EMergence to FLowering	1.55	-0.25	0.062	-3.96	0.000
SRADsumEMFL	Sum (MJ/m^2) of daily solar radiation from EMergence to Flowering	1.30	0.16	0.032	5.18	0.000
TMXMNFLPM	Mean of daily maximum temperature (in °C) from FLowering to Physiological Maturity	1.29	0.08	0.033	2.38	0.019
TMX30FLPM	Number of days above 30 °C from FLowering to Physiological Maturity	1.19	0.07	0.035	2.05	0.042

below 15 °C (r = 0.83), maximum temperature average (r = −0.77), precipitations (r = 0.62) as well as water stress index (r = -0.59). The variation of this period is fairly large going from 49 days up to 69 days. It has a positive β-coefficient (0.30) which means that increasing the duration of the period is associated with better yields. Some authors have suggested that a way to increase yields in soybean is increasing this reproductive period length (Boote *et al.* 2001). Moreover, they have shown that the response is not linear but asymptotic. In our case, the simulations showed a similar trend with a tendency to plateau at 60 days before decreasing. From 49 days to 60 days, each additional day brings 154 kg ha^{-1}. Increasing the grain filling period over 60 days does not bring any advantage. This could be related to additional cold stress (r = 0.83) and decrease in maximum temperature average (r = −0.77) that are offsetting the advantage of a longer grain filling period. However, extending the flowering to physiological maturity period to get closer to the optimum can be achieved by earlier planting dates. The simulations are showing that, on average, the planting date of 30 April is leading to a grain filling period of 59 days and an average yield of 3017 kg ha^{-1} when 20 May date leads to a grain filling period of 57 days and an average yield of 2865 kg ha^{-1}. The difference in vegetative period between the two planting dates is 6 days in favour of the earlier planting. Choosing or adapting soybean to earlier planting dates under high latitude environments may be a solution to increase yield potential, as it has an indirect effect on the grain filling duration. It has been suggested that selecting germplasm that can flower earlier under high latitudes (around mid-June) will allow longer grain filling periods and improve yields (Cooper 2003). We can also imagine that earlier planting dates may impact positively root development and in return improve crop capacity to sustain water deficit, by exploring more in depth the soil profile; it may also impact the grain filling duration.

4.3 Variables impacting yield potential under irrigated conditions

The simulations run under a realistic irrigation scenario (three applications of 30 mm) for our region have shown that a substantial yield variation still persists in spite of optimum water supply. PLS-analysis under this scenario has been carried out to define environmental covariates explaining this variation. The results can give indications about the strategies that can be designed to reduce this variability, and as a consequence, stabilize/increase yield potential beyond irrigation application. Again, it appears that the most critical period is the flowering to physiological maturity stage. As expected, precipitations (sum of rainfalls and irrigation) is not the key factor anymore; instead, the potential evapotranspiration appears as the key factor having the most impact on simulated yield. It has been demonstrated that evapotranspiration under irrigated conditions is explaining well the biomass, especially for the earlier maturity groups (Purcell *et al.* 2007). Knowing the strong correlation between biomass and grain yield for soybean, this result seems logical (De Bruin and Pedersen 2009). The positive relationship between yield and evapotranspiration has been demonstrated as a factor explaining maximum yield potential for a corn crop (Cooper *et al.* 2020). It seems also to be relevant for soybean. In our conditions, potential evapotranspiration better correlates to maximum temperature (R_2 = 0.35) rather than solar radiation (R_2 = 0.08) that is low and show little variation (17.48 ± 2.25 MJ m^{-2}). This was already demonstrated in a previous study on temperatures effect on evapotranspiration under controlled conditions (Allen *et al.* 2003). It will be necessary to confirm this information with experimental data.

The second factor is the average minimum temperature during the flowering to physiological maturity phase; this factor has already been discussed in the rainfed data analysis above. Under simulated irrigation, this factor can be related to the cold stress variable (TMN10FLPM) which is showing a negative coefficient meaning a negative impact on yield. It is interesting to underline that the correlation between the cold stress variable and the minimum temperature average variable is not strong ($r = 0.22$), meaning that for a given temperature average, the number of days with temperature below 10 °C can differ substantially. For average temperatures between 12 and 13 °C, number of days below 10 °C varies from 0 to 18 days. So, the two variables need to be treated separately. Concerning the cold stress effect on soybean, one single day of exposure of soybean to a 8 °C night stress is able to reduce by 87 % the photosynthetic activity of soybean (van Heerden and Kruger 2000) or decrease drastically photosynthesis rate at V5 (Wang *et al.* 1997). A long period of cold stress applied after R1 stage has shown to impact harvest index of cultivars when compared with a warm treatment (18/23 °C) (Schmid and Keller 1980). Considering the impact of cold stress, it appears critical to assess more in depth the impact of this factor under more controlled conditions and further investigate if genotypic differences exist, to give breeding directions and eventually develop high throughput screening methods for the trait. The screening methods could be based upon measurements of photosynthetic activity if acceptable correlations with yield are demonstrated.

5. CONCLUSION

Critical environmental variables driving yield potential in Northern France have been defined. The major constraint under rainfed conditions is related to water supply. This result leads to build simulations for optimizing irrigation scenarios. Nevertheless, even if irrigation is able to mitigate substantially the yield penalty, a non-neglectable yield variability still persists in the simulations. This yield variation is related to traits such as duration of grain filling period, adaptation to cool night temperatures and cold stress occurring during the reproductive phase of the crop cycle. Nevertheless, the outcomes of the study are based upon one single crop growth model. It would be worthwhile to run the database with other crop growth models to confirm the relevance of our results. Based on these findings, some interesting directions in both crop management and cultivar improvement can be proposed to increase yield potential under high latitudes. Furthermore, assessing genotypic variability related to those traits needs to be done and if variability exists, the development of high throughput indirect screening methods should be designed to identify the best adapted germplasm for northern European regions.

ACKNOWLEDGEMENTS

The authors are grateful to EURALIS Semences and RAGT Semences for providing trial seeds. We are thankful to Lyes Lakhal (Institut Polytechnique UniLaSalle) for running the PLS analysis and variable selection method. Thanks to Nicolas Latraye (Terre Inovia) for providing weather data from Estrées-Mons station.

SOURCE OF FUNDING

The European Agricultural Fund for Rural Development (FEADER) is acknowledged for financial support.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

CONTRIBUTIONS BY THE AUTHORS

G.B., B.L. and A.D. conducted the field work and data analysis. All authors contributed to writing and revising the manuscript.

LITERATURE CITED

Allen LH, Pan D, Boote KJ, Pickering NB, Jones JW. 2003. Carbon dioxide and temperature effects on evapotranspiration and water use efficiency of soybean. *Agronomy Journal* **95**:1071–1081.

Anda A, Soós G, Menyhárt L, Kucserka T, Simon B. 2020. Yield features of two soybean varieties under different water supplies and field conditions. *Field Crops Research* **245**:107673.

Asseng S, Turner NC, Botwright T, Condon AG. 2003. Evaluating the impact of a trait for increased specific leaf area on wheat yields using a crop simulation model. *Agronomy Journal* **95**:10–19.

Baker JT, Allen LH, Boote KJ, Jones P, Jones JW. 1989. Response of soybean to air temperature and carbon dioxide concentration. *Crop Science* **29**:98–105.

Battisti R. 2016. *Calibration, uncertainties and use of soybean crop simulation models for evaluating strategies to mitigate the effects of climate change in Southern Brazil*. PhD Thesis, Unversidade de São Paulo, Brazil.

Beck HE, Zimmermann NE, McVicar TR, Vergopolan N, Berg A, Wood EF. 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data* **5**:180214.

Bhatia VS, Singh P, Wani SP, Chauhan GS, Rao AK, Mishra AK, Srinivas K. 2008. Analysis of potential yields and yield gaps of rainfed soybean in India using CROPGRO-Soybean model. *Agricultural and Forest Meteorology* **148**:1252–1265.

Board JE, Kahlon CS. 2011. Soybean yield formation: what controls it and how it can be improved. In: El-Shemy H, ed. *Soybean physiology and biochemistry*. Rijeka, Croatia: InTechOpen, 1–36.

Boote KJ, Jones JW, Batchelor WD, Nafziger ED, Myers O. 2003. Genetic coefficients in the CROPGRO–soybean model. *Agronomy Journal* **95**:32–51.

Boote KJ, Jones JW, Hoogenboom G, Pickering NB. 1998. The CROPGRO model for grain legumes. In: Tsuji GY, Hoogenboom G, Thornton PK, eds. *Systems approaches for sustainable agricultural development. Understanding options for agricultural production*. Dordrecht: Springer Netherlands, 99–128.

Boote KJ, Kropff MJ, Bindraban PS. 2001. Physiology and modelling of traits in crop plants: implications for genetic improvement. *Agricultural Systems* **70**:395–420.

Boote KJ, Porter C, Jones JW, Thorburn PJ, Kersebaum KC, Hoogenboom G, White JW, Hatfield JL. 2016. Sentinel site data for crop model improvement—definition and characterization. In: Hatfield JL, Felisher D, eds. *Advances in agricultural systems modeling*. Madison, WI: American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America, Inc., 1–34.

Bouma E. 2005. Development of comparable agro-climatic zones for the international exchange of data on the efficacy and crop safety of plant protection products1. *EPPO Bulletin* **35**:233–238.

Caulfield F, Bunce JA. 1988. Comparative responses fo photosynthesis to growth temperature in soybean, (Glycine max [L.] Merrill) cultivars. *Canadian Journal of Plant Science* **68**:419–425.

Chapman SC. 2008. Use of crop models to understand genotype by environment interactions for drought in real-world and simulated plant breeding trials. *Euphytica* **161**:195–208.

Chapman SC, Cooper M, Hammer GL. 2002. Using crop simulation to generate genotype by environment interaction effects for sorghum in water-limited environments. *Australian Journal of Agricultural Research* **53**:379–389.

Chenu K, Cooper M, Hammer GL, Mathews KL, Dreccer MF, Chapman SC. 2011. Environment characterization as an aid to wheat improvement: interpreting genotype-environment interactions by modelling water-deficit patterns in North-Eastern Australia. *Journal of Experimental Botany* **62**:1743–1755.

Chenu K, Porter JR, Martre P, Basso B, Chapman SC, Ewert F, Bindi M, Asseng S. 2017. Contribution of crop models to adaptation in wheat. *Trends in Plant Science* **22**:472–490.

Cooper RL. 2003. A delayed flowering barrier to higher soybean yields. *Field Crops Research* **82**:27–35.

Cooper M, Tang T, Gho C, Hart T, Hammer G, Messina C. 2020. Integrating genetic gain and gap analysis to predict improvements in crop productivity. *Crop Science* **60**:582–604.

Crossa J, Vargas M, Joshi AK. 2010. Linear, bilinear, and linear-bilinear fixed and mixed models for analyzing genotype × environment interaction in plant breeding and agronomy. *Canadian Journal of Plant Science* **90**:561–574.

De Bruin JL, Pedersen P. 2009. Growth, yield, and yield component changes among old and new soybean cultivars. *Agronomy Journal* **101**:124–130.

Fehr WR, Caviness CE. 1977. Stages of soybean development. *Special Reports Iowa State University* **87**:1–13.

Funatsuki H, Matsuba S, Kawaguchi K, Murakami T, Sato Y. 2004. Methods for evaluation of soybean chilling tolerance at the reproductive stage under artificial climatic conditions. *Plant Breeding* **123**:558–563.

Gajić B, Kresović B, Tapanarova A, Životić L, Todorović M. 2018. Effect of irrigation regime on yield, harvest index and water productivity of soybean grown under different precipitation conditions in a temperate environment. *Agricultural Water Management* **210**:224–231.

Giménez L, Paredes P, Pereira LS. 2017. Water use and yield of soybean under various irrigation regimes and severe water stress. Application of AquaCrop and SIMDualKc models. *Water* **9**:393.

Grassini P, Torrion JA, Yang HS, Rees J, Andersen D, Cassman KG, Specht JE. 2015. Soybean yield gaps and water productivity in the western U.S. Corn Belt. *Field Crops Research* **179**:150–163.

Hammer GL, Kropff MJ, Sinclair TR, Porter JR. 2002. Future contributions of crop modelling—from heuristics and supporting decision making to understanding genetic regulation and aiding crop improvement. *European Journal of Agronomy* **18**:15–31.

Hoogenboom G, Porter CH, Boote KJ, Shelia V, Wilkens PW, Singh U, White JW, Asseng S, Lizaso JI, Moreno LP, Pavan W, Ogoshi R, Hunt LA, Tsuji GY, Jones JW. 2019. The DSSAT crop modeling ecosystem. *In: Burleigh Dodds Series in Agricultural Science*. Cambridge, UK: Burleigh Dodds Science Publishing, 173–216.

Hunt LA, Boote KJ. 1998. Data for model operation, calibration, and evaluation. In: Gordon YT, Hoogenboom G, Thornton PK, eds. *Understanding Options for Agricultural Production.* Springer, Dordrecht, 9–39.

Jones JW, Hoogenboom G, Porter C, Boote K, Batchelor W, Hunt L, Ritchie J. 2003. The DSSAT cropping system model. *European Journal of Agronomy* **18**:235–265.

Karam F, Masaad R, Sfeir T, Mounzer O, Rouphael Y. 2005. Evapotranspiration and seed yield of field grown soybean under deficit irrigation conditions. *Agricultural Water Management* **75**:226–244.

Krüger GHJ, De Villiers MF, Strauss AJ, De Beer M, Van Heerden PDR, Maldonado R, Strasser RJ. 2014. Inhibition of photosystem II activities in soybean (Glycine max) genotypes differing in chilling sensitivity. *South African Journal of Botany* **95**:85–96.

Kumagai E, Sameshima R. 2014. Genotypic differences in soybean yield responses to increasing temperature in a cool climate are related to maturity group. *Agricultural and Forest Meteorology* **198–199**: 265–272.

Maiorano A, Martre P, Asseng S, Ewert F, Müller C, Rötter RP, Ruane AC, Semenov MA, Wallach D, Wang E, Alderman PD, Kassie BT, Biernath C, Basso B, Cammarano D, Challinor AJ, Doltra J, Dumont B, Eyshi Rezaei E, Gayler S, Kersebaum KC, Kimball BA, Koehler AK, Liu B, O'Leary GJ, Olesen JE, Ottman MJ, Priesack E, Reynolds M, Stratonovitch P, Streck T, Thorburn PJ, Waha K, Wall GW, White JW, Zhao Z, Zhu Y. 2017. Crop model improvement reduces the uncertainty of the response to temperature of multi-model ensembles. *Field Crops Research* **202**:5–20.

Marraccini E, Gotor AA, Scheurer O, Leclercq C. 2020. An innovative land suitability method to assess the potential for the introduction of a new crop at a regional level. *Agronomy* **10**:330.

Mehmood T, Ahmed B. 2016. The diversity in the applications of partial least squares: an overview: PLS applications. *Journal of Chemometrics* **30**:4–17.

Mehmood T, Liland KH, Snipen L, Sæbø S. 2012. A review of variable selection methods in Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory Systems* **118**:62–69.

Nemeskéri E, Helyes L. 2019. Physiological responses of selected vegetable crop species to water stress. *Agronomy* **9**:447.

Peel MC, Finlayson BL, Mcmahon TA. 2007. Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences* **11**:1633–1644.

Plaza-Bonilla D, Nolot J-M, Raffaillac D, Justes E. 2017. Innovative cropping systems to reduce N inputs and maintain wheat yields by inserting grain legumes and cover crops in southwestern France. *European Journal of Agronomy* **82**:331–341.

Purcell LC, Edwards JT, Brye KR. 2007. Soybean yield and biomass responses to cumulative transpiration: Questioning widely held beliefs. *Field Crops Research* **101**:10–18.

Ruíz-Nogueira B, Boote KJ, Sau F. 2001. Calibration and use of CROPGRO-soybean model for improving soybean management under rainfed conditions. *Agricultural Systems* **68**:151–173.

Salmerón M, Purcell LC, Vories ED, Shannon G. 2017. Simulation of genotype-by-environment interactions on irrigated soybean yields in the U.S. Midsouth. *Agricultural Systems* **150**:120–129.

Schmid J, Keller ER. 1980. The behavior of three cold-tolerant and a standard soybean variety in relation to the level and the duration of a cold stress. *Canadian Journal of Plant Science* **60**:821–829.

Sharda V, Gowda PH, Marek G, Kisekka I, Ray C, Adhikari P. 2019. Simulating the impacts of irrigation levels on soybean production in Texas high plains to manage diminishing groundwater levels. *JAWRA Journal of the American Water Resources Association* **55**:56–69.

Tenenhaus M, Vinzi VE, Chatelin Y-M, Lauro C. 2005. PLS path modeling. *Computational Statistics & Data Analysis* **48**:159–205.

van Heerden PDR, Kruger GHJ. 2000. Photosynthetic limitation in soybean during cold stress. *South African Journal of Science* **96**:201–206.

Wang Z, Reddy VR, Quebedeaux B. 1997. Growth and photosynthetic responses of soybean to short-term cold temperature. *Environmental and Experimental Botany* **37**:13–24.

Willmott CJ. 1982. Some comments on the evaluation of model performance. *Bulletin of the American Meteorological Society* **63**:1309–1313.

Willmott CJ, Robeson SM, Matsuura K. 2012. A refined index of model performance. *International Journal of Climatology* **32**:2088–2094.

Wold S, Geladi P, Esbensen K, Öhman J. 1987. Multi-way principal components-and PLS-analysis. *Journal of Chemometrics* **1**:41–56.

Wold S, Ruhe A, Wold H, Dunn, III WJ. 1984. The collinearity problem in linear regression. The Partial Least Squares (PLS) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing* **5**:735–743.

Wold S, Sjöström M, Eriksson L. 2001. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* **58**:109–130.

Yassi A, Syaiful SA, Ruslan A, Ridwan I, Andari G. 2019. Simulation and production of soybean plant growth (*Glycine max* (L) Merrill) using the DSSAT model with different scenarios of water supply and compost. *IOP Conference Series: Earth and Environmental Science* **343**:012014.

Zhang F, Charles TC, Pan B, Smith DL. 1996. Inhibition of the expression of Bradyrhizobium japonicum nod genes at low temperatures. *Soil Biology and Biochemistry* **28**:1579–1583.