PREDICTING SOYBEAN DEVELOPMENT WITH A SIMPLE PHOTOTHERMAL DYNAMIC ALGORITHM

AD Severini[1*], S Álvarez Prado[23], ME Otegui[123], CRC Vega[4], S Zuil[5], M Kavanová[6], S Ceretta[6], M Acreche[37], RF Scholz Drodowski[8], RA Serrago[23] and DJ Miralles[23]

[1*]INTA. Centro Regional Buenos Aires Norte. Estación Experimental Agropecuaria Pergamino. Argentina. severini.alan@inta.gob.ar; [2]Facultad de Agronomía, Universidad de Buenos Aires, Argentina; [3]CONICET; [4]INTA. Centro Regional Córdoba. Estación Experimental Agropecuaria Manfredi. Argentina; [5]INTA. Centro Regional Santa Fe. Estación Experimental Agropecuaria Reconquista. Argentina; [6]INIA La Estanzuela. Uruguay; [7]INTA. Centro Regional Salta Jujuy. Estación Experimental Agropecuaria Salta. Argentina; [8]IPTA Capitán Miranda. Paraguay.

INTRODUCTION

Predicting the occurrence of the critical period for soybean's yield determination is important for farmers to decide on variety and sowing date with the objective to expose this period (during which yield is mainly determined) to the best environmental conditions. Simulation models like APSIM (Keating et al., 2003) and DSSAT (Jones et al., 2003) are extremely useful to predict yield under different environments. However, in these models the parameterization of genetic coefficients for simulating phenology and yield is a complex and time-consuming process limiting the number of genotypes available for simulation. Moreover, the modeled genotypes may not be representative of the broad range of genetic material grown by farmers. We created a simple, dynamic model based on photoperiod and temperature to predict flowering initiation (R1), start of grain filling (R5) and physiological maturity (R7, Fehr and Caviness (1977)) in a wide number of commercial soybean varieties ranging from maturity group (MG) II to VI used by farmers in Argentina, Uruguay, and Paraguay. Our aim is to build a model that is simple enough to be calibrated, yet able to predict these stages with reasonable accuracy. The "CRONOSOJA" model outputs will be freely available in the near future on an interactive website (http://soja.cronos.agro.uba.ar).

MATERIALS AND METHODS

Soybean varieties, field experiments and measurements

We selected 34 soybean commercial varieties from major seed companies of Argentina, Uruguay and Paraguay, covering MG II to VI (Table 2). By previous expert consultation, we ensured that these varieties were representative of the regional soybean seed market and widely used by farmers. To explore a broad range of temperature and photoperiod conditions, we sowed those varieties from October to February, with a gap of around one month between successive sowing dates, during three seasons in the locations listed in Table 1. We determined the phenological stage of each plot every 2-3 days using the scale of Fehr and Caviness (1977).

Table 1: Number of different sowing dates tested in each experimental site during three growing seasons. Bold numbers denote datasets that were used for validating the model, while those in italics show datasets that were not included yet (but they will be included soon), neither for calibration nor for validation of the model. The rest of the datasets were used for the initial calibration of the model.

				Season		
Country	Site	Latitude (°S)	Longitude (°W)	2016/17	2017/18	2018/19
Argentina	CABA	34.6	58.5	1	*1*	0
Argentina	Chascomús	35.6	58.0	3	0	0
Argentina	Manfredi	31.8	63.7	0	**4**	*3*
Argentina	Pergamino	33.9	60.6	4	3	*3*
Argentina	Reconquista	29.2	59.9	3	3	*3*
Argentina	Salta	24.9	65.5	0	0	**3**
Paraguay	Capitán Miranda	27.2	55.8	0	0	**1**
Uruguay	La Estanzuela	34.3	57.7	0	2	*4*

Model development

For model fitting, we transformed soybean developmental stages into a continuous numerical scale: emergence (EM) = 0, R1 = 1, R5 = 2 and R7 = 3. In its current state, the model only simulates the latter three stages but R3 (beginning of pod development) and R6 (full seed size) will be included soon as temperature-corrected calendar-day deviations from R1 and R5, respectively. The equation describing the developmental stage (R) at a specified time (t, in days after EM) is

$$R_t = \begin{cases} R_{t-1} + \dfrac{1}{D_{\mathrm{EM-R1}}} f(T_t) f(P_t), & \text{if } R_{t-1} < 1 \\ R_{t-1} + \dfrac{1}{D_{\mathrm{R1-R5}}} f(T_t) f(P_t), & \text{if } 1 \leq R_{t-1} < 2 \\ R_{t-1} + \dfrac{1}{D_{\mathrm{R5-R7}}} f(T_t) f(P_t), & \text{otherwise} \end{cases}$$

where $D_{\mathrm{EM-R1}}$, $D_{\mathrm{R1-R5}}$, and $D_{\mathrm{R5-R7}}$ are the photothermal days for phases EM-R1, R1-R5, and R5-R7; and $f(T_t)$ and $f(P_t)$ are functions that describe the temperature and photoperiod multipliers, which can adopt values between 0 and 1. Therefore, $f(T_t)$ and $f(P_t)$ penalize the maximum developmental rates $1/D_{\mathrm{EM-R1}}$, $1/D_{\mathrm{R1-R5}}$ and $1/D_{\mathrm{R5-R7}}$.

The temperature multiplier, $f(T_t)$, is defined as

$$f(T_t) = \begin{cases} \dfrac{2(T_t - T_{\min})^{\alpha}(T_{\mathrm{opt}} - T_{\min})^{\alpha} - (T_t - T_{\min})^{2\alpha}}{(T_{\mathrm{opt}} - T_{\min})^{2\alpha}}, & \text{if } T_{\min} < T_t < T_{\max} \\ 0, & \text{otherwise} \end{cases}$$

where T_t is the mean air temperature of day t; $T_{\min}$ (also known as T_{base}), T_{opt} and $T_{\max}$ are the cardinal temperatures; and α is a shape parameter, which in turn is defined as

$$\alpha = \frac{\log(2)}{\log(\frac{T_{\max} - T_{\min}}{T_{\mathrm{opt}} - T_{\min}})}$$

And the photoperiod multiplier, $f(P_t)$, is defined as

$$f(P_t) = \begin{cases} 1 - P_{\mathrm{sen}}(P_t - P_{\mathrm{opt}}), & \text{if } P_t > P_{\mathrm{opt}} \\ 1, & \text{otherwise} \end{cases}$$

whereP_t is the photoperiod of day t; P_{sen} is the photoperiod sensitivity; and P_{opt} is the optimum photoperiod.

Given the colinearity among the parameters of the model, not all parameters could be simultaneously calibrated from the data. For this reason, some parameters were fixed while others were calibrated. After sensitivity analysis, we calibrated those parameters to which the model was most sensitive and fixed the less-sensitive parameters in accordance to values reported in the literature. For instance, the cardinal temperatures $T_{\min}$ and $T_{\max}$ were set to 0 and 45°C, respectively, whereas T_{opt} was set to 32, 28 and 25 °C for phases EM-R1, R1-R5 and R5-R7 (Setiyono et al., 2007). The photoperiod sensitivity P_{sen} was set to 0.25 h^{-1} after averaging all photoperiod sensitivities for MGs between III and VI described in Archontoulis et al. (2014). The calibrated parameters were the photothermal days from emergence to R1 ($D_{\mathrm{EM-R1}}$), from R1 to R5 ($D_{\mathrm{R1-R5}}$) and from R5 to R7 ($D_{\mathrm{R5-R7}}$), and the optimum photoperiod (P_{opt}), below which the rate of development is maximum.

The model was fitted by means of Markov chain Monte Carlo (MCMC) algorithm implemented by the function modMCMC from the package FME (Soetaert and Petzoldt, 2010) in R (R Core Team, 2019). The MCMC iterated across combinations of parameters $D_{\mathrm{EM-R1}}$, $D_{\mathrm{R1-R5}}$, $D_{\mathrm{R5-R7}}$ and P_{opt} while trying to minimize the root mean squared error (RMSE) between modeled and observed days from emergence to R1, R5 and R7 stages. The model was run simultaneously across all locations and sowing dates where a particular soybean variety was tested. The number of iterations was set to 10000 and the burn-in period was set to 5000 iterations with 3 chains. Convergence was tested by comparing chains with the $\widehat{R}$ diagnostic, which was allowed not to be greater than 1.1 (Gelman et al., 2013).

RESULTS

The calibrated parameters were within ranges reported previously (Archontoulis et al., 2014; Setiyono et al., 2007) (Table 2). As expected, P_{opt} was inversely related to MG, with values around 13.3 h for MG II and III, and 12.4 h for MG VI (Table 2). The overall RMSE averaged across all varieties for the calibration dataset was 4, 6 and 8 days for R1, R5 and R7, respectively; with some varieties showing a better fit than others (Table 2). In turn, the averaged RMSE of the model used to predict R1, R5 and R7 in the validation data set was 5, 6 and 10 days, respectively. The RMSE values were similar or even lower than other widely used models, such as CERES and APSIM (Fig. 1).

Table 2: Calibrated parameters P_{opt}, $D_{\mathrm{EM-R1}}$, $D_{\mathrm{R1-R5}}$ and $D_{\mathrm{R5-R7}}$, and root mean square errors (RMSE) for stages R1, R5 and R7 for all evaluated varieties.

Company	Variety	$\boldsymbol{P}_{\mathbf{opt}}$ (h/day)	$\boldsymbol{D}_{\mathbf{EM-R1}}$ (days)	$\boldsymbol{D}_{\mathbf{R1-R5}}$ (days)	$\boldsymbol{D}_{\mathbf{R5-R7}}$ (days)	RMSE (days)		
						R1	R5	R7
Don Mario	DM 2200	13.3	26.3	17.1	38.8	4.2	7.4	11.5
Don Mario	DM 3312	13.3	26.7	18.8	40.7	3.8	6.6	12.3
Nidera	NS 2632	13.3	24.6	19.3	38.3	4.2	5.3	11.2
Nidera	NS 3220	13.2	26.1	22.2	28.1	3.0	4.3	4.2
Bayer	FN 4.35	13.1	27.5	24.0	31.0	2.2	4.7	7.0
Bayer	CZ 4505	13.0	30.9	22.1	35.9	5.2	6.5	10.6
Asgrow	AW 4326	12.9	26.1	24.7	29.0	1.8	5.1	6.2
Bayer	CZ 4.97	12.9	27.8	24.7	33.7	2.6	5.9	11.0
Bayer	FN 3.85	12.9	26.3	20.8	28.6	3.4	3.9	4.9
Nidera	NS 5258	12.9	30.7	21.4	35.6	3.4	5.9	7.7
Santa Rosa	RA 349	12.9	24.8	18.3	39.8	3.8	6.3	11.0
Don Mario	DM 3815	12.8	24.6	18.7	34.3	4.2	5.1	8.9
Don Mario	DM 40R16	12.8	25.0	20.8	37.2	4.4	6.7	9.5
Don Mario	DM 4612	12.8	24.5	21.6	36.8	5.1	5.8	9.3
INIA	Genesis 5501	12.8	36.4	19.5	38.1	5.5	9.0	8.0
INIA	Genesis 5602	12.8	37.0	18.3	40.0	4.1	7.6	9.5
Nidera	NS 3809	12.8	25.6	16.6	38.2	4.5	3.9	11.7
Nidera	NS 4955	12.8	26.0	21.1	39.3	3.8	6.6	7.4
INIA	Genesis 5601	12.7	36.4	18.9	37.1	5.4	6.9	5.8
Santa Rosa	RA 550	12.7	24.1	23.5	36.8	4.4	8.4	8.7
Bayer	CZ 5905	12.6	35.9	18.4	36.7	5.8	6.2	8.1
Don Mario	DM 6.2i	12.6	35.0	15.4	38.1	4.2	7.8	8.0
Asgrow	AW 5714	12.5	35.3	17.4	35.1	6.8	7.8	7.6
Don Mario	DM 50i17	12.5	25.9	16.4	39.2	3.3	8.9	6.6
Don Mario	DM 53i53	12.5	28.2	20.6	38.3	4.4	5.0	10.2
INIA	Genesis 6201	12.5	33.8	18.2	39.1	2.6	2.0	2.9
INTA	INTA Parana 5500	12.5	35.0	19.2	34.2	6.6	5.3	8.8
Asgrow	AW 5815	12.4	31.0	16.3	37.8	5.4	10.1	6.4
Asgrow	AW M6410	12.4	36.4	13.7	41.9	7.0	9.9	9.3
Bayer	CZ 6505	12.4	35.7	18.6	38.2	5.0	10.9	10.3
INTA	INTA Parana 6200	12.4	34.7	17.0	38.1	4.2	6.0	9.9
Nidera	NS 6248	12.4	33.6	16.8	40.7	6.2	7.4	9.0
Santa Rosa	RA 549	12.4	33.1	17.7	38.3	5.2	7.1	7.7
Santa Rosa	RA 655	12.1	29.4	21.9	40.5	4.3	8.5	10.2

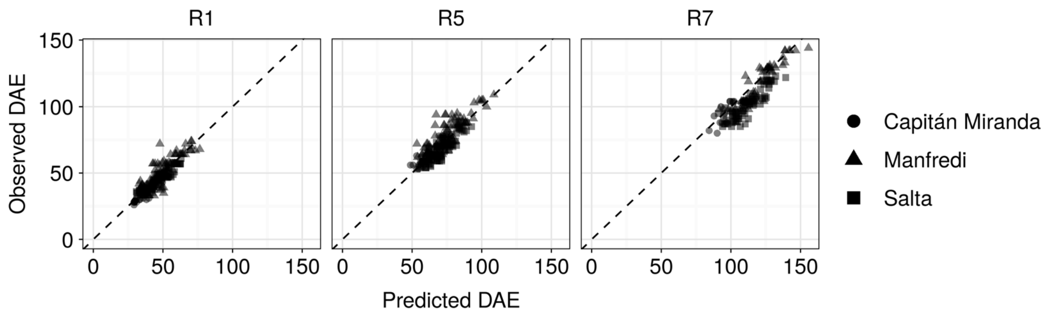

Figure 1: Observed vs. predicted days from emergence (DAE) to R1, R5 and R7 stages for all varieties listed in Table 2 in the validation dataset. The dashed line represents 1:1. The root mean squared error, averaged across all varieties, for R1, R5 and R7 stage was 5, 6 and 10 days, respectively.

CONCLUSIONS

The model performed well when tested to predict an independent validation data set (Fig. 1). Its ability to accurately predict the occurrence of pivotal stages of soybean development along an ample latitudinal range highlights the importance of formalizing models based on ecophysiological knowledge rather than empirical relationships. The model will be made freely available on a website (http://soja.cronos.agro.uba.ar) in the near future, providing an interactive tool for farmers to select the best combination of sowing date and variety for a given location.

ACKNOWLEDGEMENTS

This work was funded by the project "Bases fisiológicas y genéticas de las respuestas de trigo y soja a limitantes bióticas y abióticas: estudios orientados al mejoramiento genético y al manejo de los cultivos en el Cono Sur" from PROCISUR.

REFERENCES

- ARCHONTOULIS, S.V., MIGUEZ, F.E., MOORE, K.J., 2014. A methodology and an optimization tool to calibrate phenology of short-day species included in the APSIM PLANT model: Application to soybean. Environmental Modelling & Software 62, 465–477.
- FEHR, W.R., CAVINESS, C.E., 1977. Stages of soybean development (No. 87).
- GELMAN, A., CARLIN, J.B., STERN, H.S., DUNSON, D.B., VEHTARI, A., Rubin, D.B., 2013. Bayesian data analysis. Chapman and Hall/CRC.
- JONES, J.W., HOOGENBOOM, G., PORTER, C.H., BOOTE, K.J., BATCHELOR, W.D., HUNT, L.A., WILKENS, P.W., SINGH, U., GIJSMAN, A.J., RITCHIE, J.T., 2003. The DSSAT cropping system model. European Journal of Agronomy, Modelling Cropping Systems: Science, computer software and Applications 18, 235–265.
- KEATING, B.A., CARBERRY, P.S., HAMMER, G.L., PROBERT, M.E., ROBERTSON, M.J., HOLZWORTH, D., HUTH, N.I., HARGREAVES, J.N., MEINKE, H., HOCHMAN, Z., 2003. An overview of APSIM, a model designed for farming systems simulation. European Journal of Agronomy 18, 267–288.
- R CORE TEAM, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- SETIYONO, T.D., WEISS, A., SPECHT, J., BASTIDAS, A.M., CASSMAN, K.G., DOBERMANN, A., 2007. Understanding and modeling the effect of temperature and daylength on soybean phenology under high-yield conditions. Field Crops Research 100, 257–271.
- SOETAERT, K., PETZOLDT, T., 2010. Inverse Modelling, Sensitivity and Monte Carlo Analysis in R Using Package FME. Journal of Statistical Software 33, 1–28.