Options for calibrating CERES-maize genotype specific parameters under data-scarce environments

A. A. Adnan[1,2,3¤]*, J. Diels[2], J. M. Jibrin[3], A. Y. Kamara[4], P. Craufurd[5], A. S. Shaibu[1,3], I. B. Mohammed[4], Z. E. H. Tonnang[4,5]

1 Department of Agronomy, Bayero University Kano, Kano, Nigeria, **2** Department of Earth and Environmental Sciences, Division of Soil and Water Management, KU Leuven, Leuven, Belgium, **3** Centre for Dryland Agriculture (CDA), Bayero University Kano, Kano, Nigeria, **4** International Institute of Tropical Agriculture, Ibadan, Nigeria, Carolyn House, Croydon, United Kingdom, **5** International Maize and Wheat Improvement Center (CIMMYT) World Agroforestry Centre (ICRAF) House United Nations Avenue, Nairobi, Kenya

¤ Current address: Department of Earth and Environmental Sciences, Division of Soil and Water Management, KU Leuven, Leuven, Belgium
* aaadnan.agr@buk.edu.ng

Abstract

Most crop simulation models require the use of Genotype Specific Parameters (GSPs) which provide the Genotype component of G×E×M interactions. Estimation of GSPs is the most difficult aspect of most modelling exercises because it requires expensive and time-consuming field experiments. GSPs could also be estimated using multi-year and multi locational data from breeder evaluation experiments. This research was set up with the following objectives: i) to determine GSPs of 10 newly released maize varieties for the Nigerian Savannas using data from both calibration experiments and by using existing data from breeder varietal evaluation trials; ii) to compare the accuracy of the GSPs generated using experimental and breeder data; and iii) to evaluate CERES-Maize model to simulate grain and tissue nitrogen contents. For experimental evaluation, 8 different experiments were conducted during the rainy and dry seasons of 2016 across the Nigerian Savanna. Breeder evaluation data were also collected for 2 years and 7 locations. The calibrated GSPs were evaluated using data from a 4-year experiment conducted under varying nitrogen rates (0, 60 and 120kg N ha^{-1}). For the model calibration using experimental data, calculated model efficiency (EF) values ranged between 0.88–0.94 and coefficient of determination (d-index) between 0.93–0.98. Calibration of time-series data produced nRMSE below 7% while all prediction deviations were below 10% of the mean. For breeder experiments, EF (0.58–0.88) and d-index (0.56–0.86) ranges were lower. Prediction deviations were below 17% of the means for all measured variables. Model evaluation using both experimental and breeder trials resulted in good agreement (low RMSE, high EF and d-index values) between observed and simulated grain yields, and tissue and grain nitrogen contents. It is concluded that higher calibration accuracy of CERES-Maize model is achieved from detailed experiments. If unavailable, data from breeder experimental trials collected from many locations and planting dates can be used with lower but acceptable accuracy.

1.0 Introduction

Maize has become an important crop in Nigeria in the past decades due to its importance as food for human consumption; feed for animals and as a source of industrial raw material [1]. Despite its importance, yield of maize has remained quite low in the Savannas mostly due to biotic and abiotic constraints [2]. In recent years, new early and extra early maturing maize varieties that are tolerant to most of the biotic and abiotic constraints have been developed for the Nigerian Savannas by the International Institute for Tropical Agriculture (IITA) and its partners. Several agronomic technologies have also been developed to increase the productivity of these varieties with a view to enhancing maize productivity. Before the varieties are released, they are usually grown under multi-locational yield and crop management evaluation trials over several years. Dissemination of such varieties and technologies will require setting up of costly and time-consuming experiments across wide areas. This is needed to adequately evaluate the Genotype × Environment interaction which demonstrates the performance of each variety across diverse environments. Unless this is done, breeders cannot conclusively recommend genotypes for specific environments.

Crop simulation modeling offers an opportunity to explore the potential of new varieties and crop management practices in different environments (soil, climate, management) prior to their release [3]. Recently, use of crop simulation models, particularly DSSAT, is on the increase in Africa through initiatives such as the Agricultural Models Inter-Comparison Project (AgMIP) [4]. In West Africa, the CERES-Maize model has been recently used by McCarthy et al. [3] to evaluate climate-sensitive farm management practices in the Northern Regions of Ghana. Adnan et al. [5,6] used the same model to determine the nitrogen fertilization requirements of early maturing maize in the Sudan Savanna of Nigeria and the optimum planting dates of maize in Northern Nigeria. Iyanda et al. [7] used the CERES-Maize model to identify potential zones for maize production in Nigeria. One of the major requirements for the use of crop simulations is calibration of Genotype Specific Parameters (GSPs). GSPs are sets of parameters that enable crop models to simulate the performance of diverse genotypes under varying soil, weather and management conditions [8]. Like all other parameters in crop simulation models, the GSPs must have a physical or biological meaning [9]. Measuring GSPs directly from real systems (farm and field level) is very complex and impractical, and results in highly inaccurate and uncertain values of estimated variables [10,11]. Direct measurement requires setting up of field or growth chamber studies, collection of many samples, and exposure to different photoperiods where necessary. The most common method for deriving GSPs is from field experiments designed specifically for their estimation [12,13]. This process is quite expensive, time consuming and requires regular sampling of growth, phenology and yield data for each variety following a set of minimum dataset rules [8]. Since the movement of models from research and policy to adoption by farmers and extension, the need for rapid estimation of GSPs for newly released varieties has become more urgent [14]. Several concerns have been raised even in locations where abundant and high-quality data for calibration of GSPs for model uses are available. In a recent publication, Seidel et al. [15] presented various methods for improving the current methods of calibrating crop models.

Since most models have been developed elsewhere in Europe and USA, their use outside their domain of development requires a great deal of data for their calibration and validation. Several approaches for estimating GSPs have been documented. The genetic coefficient calculator (GENCALC) was used by Anothai et al. [16] to determine variety coefficients for new peanut lines in Thailand from standard varietal trials. From their experiments, they were able to successfully calibrate groundnut GSPs using a set of field experiments and yield evaluation experiments using the GENCALC software. Mavromatis et al. [17] successfully generated

GSPs of soybean from crop performance trials in Georgia, USA. Bannayan et al. [18] employed a pattern recognition technique, which is based on similarity measures, to estimate GSPs for maize. In their experiments, pattern recognition was used as an alternative to GENCALC and GLUE in estimation of maize GSPs. The generalized likelihood uncertainty estimation (GLUE) method was used by He et al. [19] to successfully estimate maize GSPs in North Carolina. Welch et al. [14] used data from private-sector variety performance trials to develop soybean GSPs in the soybean belt of the United States of America. Buddhaboon et al. [20] used GENCALC and GLUE to estimate GSPs of deep water rice using CERES-Rice model. Most recently Lamsal et al. [21] used the independent component analysis (ICA) and separate factor approaches to estimate soybean GSPs from large breeding trial datasets in the USA.

With a growing number of researchers using the DSSAT model in the Savannas of Africa, there is need to evaluate the GSP calibration step as it is the aspect that requires the greatest amount of data and expertise. Calibration of GSPs can also be done using secondary data from breeders who routinely conduct multi-location trials. Such datasets are available in Africa where strong breeding programs are present. Because the conventional method of calibrating GSPs is quite expensive and laborious, there is need to utilize secondary breeder trial data for calibrating maize GSPs and to evaluate the accuracy of this approach by comparing it with calibrations done using detailed calibration experiments. The present research compares data generated from conventional experiments and from breeder evaluation trials. This is done to justify the claim that available data from breeder evaluation experiments can potentially be used for generating maize GSPs when setting up conventional experiments is unfeasible.

The objectives of this research were: i) to determine GSPs of 10 newly released open pollinated (OPV) and hybrid maize varieties for the Nigerian Savannas using data from both field experiments specifically designed for this purpose (herein called calibration experiments) and by using data from breeder varietal evaluation trials (herein called breeder evaluation experiments); ii) to compare the accuracy of the GSPs generated using calibration and breeder data; and iii) to evaluate the ability of the GSPs calibrated using the 2 methods to simulate grain yield and tissue/grain nitrogen contents of maize.

2.0 Materials and methods

2.1 Model description

The maize model used in this study is the CSM CERES-Maize model of DSSAT version 4.6. Detailed description of the CERES–maize model of DSSAT can be found in Jones et al [22]. CERES-maize is variety and site specific and operates on a daily time step. It dynamically simulates the development of roots and shoots, the growth and senescence of leaves and stems, biomass accumulation, and the growth of maize grain yield as a function of soil and weather conditions, crop management practices, and variety characteristics. The model uses a standardized system for model inputs and outputs that have been described elsewhere [23,24]. The input system enables the user to select crop genotype (variety), weather, soil, and management data appropriate to experiment being simulated. Required crop genetic inputs for CERES Maize are given in Table 1.

2.2 Field experiments

Three sets of data were used in this study: calibration experiments, breeder evaluation experiments and model validation experiments.

The calibration experiments were conducted during the rainy and dry season of 2016 in four locations in northern Nigeria. The experiments were conducted at the Teaching and Research Farm of the Faculty of Agriculture, Bayero University, Kano (N11.516 E8.516 466m

Table 1. Definition of DSSAT maize genotype specific parameters.

Coefficient	Description
P1 (^{o}C day)	Thermal time from seedling emergence to the end of juvenile phase
P2 (days)	Delay in development for each hour that day-length is above 12.5 hours
P5 (^{o}C day)	Thermal time from silking to time of physiological maturity
G2 (#)	Maximum kernel number per plant
G3 (mg day^{-1})	Kernel growth rate during linear grain filling stage under optimum conditions
PHINT (^{o}C day tip^{-1})	Thermal time between successive leaf tip appearance

https://doi.org/10.1371/journal.pone.0200118.t001

asl), at the Teaching and Research Farm of Audu Bako College of Agriculture Dambatta (N12.333 E8.517 442m asl), at the Irrigation Research Farm of Institute for Agricultural Research (IAR) Samaru, Zaria (N11.187 E7.147 702m asl) and at the Agricultural Research Station of the Kaduna Agricultural Development Project (KADP) in Saminaka, Lere (N10.52 E8.472 786m asl). Eight experiments were used for the calibration spanning over four locations, two seasons and eight planting dates (Table 2). The calibration experiment consisted of 20 varieties, but we focused on 10 varieties that were common to both the on-station calibration experiments and breeder varietal evaluation experiments (Table 3). The calibration experiments were conducted near irrigation facilities so as to maintain optimum moisture by irrigating when the soil moisture is below field capacity. Moisture conditions were monitored using a Time Domain Reflectometry (TDR) Meter 6050X1 TRASE SYSTEM (Soilmoisture Equipment Corp.). Recommended levels of mineral fertilizers for the region were applied (120N:60P:60K kg ha^{-1}); potassium (K) was applied in form of Muriate of Potash, phosphorus in the form of Single Super Phosphate, and Nitrogen was applied in the form of Urea. While all the P and K fertilizers were applied at sowing; only half of the N fertilizer was applied at the time of sowing and the other half applied 21 days later. In addition, poultry manure (approximately NPK 1.1:0.8:0.5) was added to the fields at the rate of 5 tons ha^{-1} to maintain optimum nutrient status. The calibration experiments were laid down in a Randomized Complete Block Design (RCBD) with four replications. The gross plot consisted of six ridges, 0.75 m apart and 3 m long (plot area = 13.5 m^2). The two innermost ridges were used as the net plot for yield assessment and for sampling purposes. A space of 0.5 m was used between plots and 1m between replications. The experimental fields were cleared, harrowed, ridged and thereafter sprayed with a pre-emergence herbicide, Primextra (Atrazine + Metolachlor) at the rate of 4lha^{-1} before planting. The maize was sown at intra-row spacing of 0.25m at two seeds per hole, and later thinned to one plant giving a population of 53, 333 plants ha^{-1}.

For the breeder evaluations, experimental units are one-row plots, each 4 m long with inter-row spacing of 0.75 m and intra-row spacing of 0.40 m. Three seeds planted and later thinned to two per hill at 2 weeks after emergence to give a final plant population density of about 66,666 plants ha^{-1}. Fertilizer is usually applied at the rate of 60 kg ha^{-1} of NPK 15:15:15 at 2 WAP. An additional 60 kg ha^{-1} N using urea is top dressed at 5 WAP. The trials are kept weed free by applying atrazine (1 -chloro-3-ethylamino-5- isopropylamino-2,4,6-triazine) and gramoxone (1,1-dimethyl-4,4-bipyridinium dichloride) as pre- and post-emergence herbicides at 5 L in 220 L of water ha^{-1} and subsequently by hoeing. Grain yield is calculated based on 80% (800 g grain kg^{-1} ear weight) shelling percentage and adjusted to 150 g kg^{-1} moisture content.

For calibration using data from breeder evaluation trials, we collected long-term yield evaluation data from breeders at the International Institute for Tropical Agriculture (IITA), Ibadan and the Institute for Agricultural Research (IAR), Zaria. Data for the 10 maize varieties used in this study were selected. The bulk data was subjected to various quality checks. We

Table 2. Description of sites for field experiments and breeder evaluation data.

Site and Environment	Code	Sowing Date	Ecology*	Dominant Soil Type	Cumulative Rainfall + Irrigation (mm)
On-station Experiments for calibration					
Bayero Uni. Farm Dry Season	BUK DS	16-03-2016	SS	Typic Kandiustalf	843
Bayero Uni. Farm Dry Season	BUK RS	25-07-2016	SS	Typic Kandiustalf	705
Dambatta Dry Season	DBT DS	19-03-2016	SS	Typic Kanhaplustalf	976
Dambatta Rainy Season	DBT RS	26-07-2016	SS	Typic Kanhaplustalf	690
Samaru Dry Season	SMR DS	22-03-2016	NGS	Plinthic Haplustult	840
Samaru Rainy Season	SMR RS	29-07-2016	NGS	Plinthic Haplustult	850
Lere Dry Season	LER DS	17-03-2016	NGS	Plinthic Kandihumult	964
Lere Rainy Season	LER RS	31-07-2016	NGS	Plinthic Kandihumult	1054
Breeder Varietal Evaluation experiments					
Zaria 2012	ZRA 12	12-06-2012	NGS	Typic Kandiustalf	1123
Zaria 2013	ZRA 13	10-06-2013	NGS	Typic Kandiustalf	1222
Mokwa 2012	MKW 12	08-06-2012	SGS	Oxic Haplustult	1346
Mokwa 2013	MKW 13	28-05-2013	SGS	Oxic Haplustult	1402
Bagauda 2012	BGD 12	13-06-2012	SS	Typic Kandiustalf	882
Bagauda 2013	BGD 13	21-06-2013	SS	Typic Kandiustalf	941
Batsari 2012	BTR 12	22-06-2012	SS	Ustoxic Dystropept	806
Batsari 2013	BTR 13	21-06-2013	SS	Ustoxic Dystropept	854
Samaru 2012	SMR 12	11-06-2012	NGS	Typic Plinthiustalfs	1118
Samaru 2013	SMR 13	14-06-2013	NGS	Typic Plinthiustalfs	1241
Minjibir 2012	MJB 12	21-06-2012	SS	Typic Kandiustalfs	791
Minjibir 2013	MJB 13	18-06-2013	SS	Typic Kandiustalfs	824
Kadawa 2012	KDW 12	23-06-2012	SS	Typic Plinthiustalfs	891
Kadawa 2013	KDW 13	19-06-2013	SS	Typic Plinthiustalfs	913
Experiments for Model Evaluation					
BUK 2013	BUK 13	10-06-2013	SS	Typic Kandiustalfs	892
BUK 2014	BUK 14	21-06-2014	SS	Typic Kandiustalfs	967
BUK 2015	BUK 15	29-05-2015	SS	Typic Kandiustalfs	1021
BUK 2016	BUK 16	09-06-2016	SS	Typic Kandiustalfs	972

* SS = Sudan Savanna, NGS = Northern Guinea Savanna, SGS = Southern Guinea Savanna

used data for the 2012 and 2013 seasons from seven locations where weather and soil data were available. Table 2 shows the locations and data used in the calibration with breeder data.

For model validation, field experiments were conducted at the Research Farm (11°59'N, 8°25'E 466m above sea level) of the Faculty of Agriculture, Bayero University, Kano in the rainy seasons between 2013 to 2016 (four seasons). The treatments consisted of three rates of nitrogen (0, 60 and 120 kg N ha^{-1}) and ten maize varieties used in the calibration experiments (Table 2). Treatments were laid out in a split-plot design with three replications. Nitrogen rates were assigned to the main plots while the varieties were assigned to the sub-plot. Although the experiments were conducted in the rainy season, moisture contents were monitored, and supplementary irrigation was provided to ensure no moisture stress. All conventional agronomic cultural practices were followed. The data collected for model evaluation includes grain yield (kg ha^{-1}), total grain nitrogen (kg ha^{-1}), total tissue nitrogen (kg ha^{-1}) and nitrogen harvest index (percentage). Total grain and tissue nitrogen were determined using the Micro Kjeldahl method.

Table 3. Characteristics of maize varieties used in the study.

S/N	Name	Common Name	Type	Maturity	Tolerance
1	2011TZEWDTSTRSYN	Early White	OPV¶	Early	Drought/Striga
2	2013TZEEWPOPDTSTR	E.E White	OPV	Extra Early	Drought/Striga
3	EVDT-W-99STR	Sammaz 32	OPV	Early	Drought
4	EVDT-Y-2000-STR	Sammaz 34	OPV	Early	Drought/Striga
5	OBA SUPER 9	Oba 9	Hybrid	Late	-
6	M0926-8	Seedco White	Hybrid	Late	MSV*
7	TZE124 x TZE125	Sammaz 41	Hybrid	Early	MSV
8	TZEEI29 x TZEEI21	Ife hybrid 5	Hybrid	Extra Early	Drought
9	TZEE-WPOPSTRC5 x TZEEEI6	Ife hybrid 6	Hybrid	Extra Early	Drought/Striga
10	TZEYPOPDTSTRC4 x TZEEI13	Sammaz 42	Hybrid	Extra Early	Drought/Striga

¶ Open pollinated variety

* Maize Streak Virus

https://doi.org/10.1371/journal.pone.0200118.t003

2.3 Plant measurements

Evaluation of crop development was done by observing the phenology of the different maize varieties and recording the length of time (days) it takes to attain each phenological phase. The measurements were then converted to growing degree days (GDD) using a base temperature of 8°C and adopting the relationship:

$$GDD = \sum_{i=1}^{n} \left(\frac{Tmaxi + Tmini}{2} \right) - Tbase \qquad (1)$$

Ten plants were tagged from the center of each plot in each replication for phenological observations. The end of the juvenile stage (i.e. panicle initiation) was determined through destructive sampling and dissection of three plants, followed by observation of apical meristem to check for floral bud development at 2 d intervals starting from 14d after emergence. The end of juvenile stage was recorded when the male flower primordial were visible in 50% of plants examined. Days to 50% tasseling was recorded when tassels were observed on 50% of the tagged plants. Physiological maturity observations were conducted as follows: kernels were removed from the base, middle and distal end of each sampled ear daily, starting when husks begin to show signs of drying. Days to physiological maturity was recorded when 50% of the kernels in each tagged ear had formed a black layer, indicating physiological maturity.

Plant biomass was taken at four different stages: vegetative, anthesis, grain filling and physiological maturity. Five plants within a one-meter strip in a row were cut at the ground level as suggested by Ogoshi et al.[11]. Leaves were separated from the stem, chopped and dried in the shade for three days. Both stems and leaves were oven dried at 70°C for 36–48 hours until the sample had attained constant weight. Yield and yield component measurements were taken at harvest maturity. Plant height was measured from five randomly tagged plants within the net plot using a standard field meter rule. Other variables measured included: the number of seeds per unit area (seed # m^{-2}), dry seed weight (g m^{-2}), dry cob weight (g m^{-2}), dry husks weight (g m^{-2}), grain yield (kg ha^{-1}) and stover weight at harvest (kg ha^{-1}). All yield and yield component measurements were done using procedures and formulae described by Ogoshi et al. [11]. Total grain and tissue nitrogen (measured for the evaluation experiments only) were determined using the Micro–Kjeldahl method.

2.4 Soil and weather data

Detailed soil studies were conducted for each experimental location before planting. Soil pits were dug in each location, and soil samples were taken from each layer. The collected samples were then analyzed for pH, texture, moisture, bulk density, exchangeable potassium (K), organic matter, phosphorus (P), total nitrogen and CEC. For the detailed calibration and validation experiments, daily weather data were collected from weather stations (Watchdog 2000 Series, Spectrum Technologies) adjacent to all experimental sites. All weather stations were less than 5 km away from the experimental sites.

2.5 Initialization of soil and weather parameters

Daily records of minimum and maximum temperature, total solar radiation, and total rainfall are required for the CERES-Maize model weather initialization. The *Weatherman* utility in DSSAT was used to input the weather data to create the weather file used by the CERES- Maize model. The *Weatherman* utility also requires information on name of weather station, latitude, longitude and altitude. Soil data tool (*SBuild*) was used to create the soil database which was used for the general simulation purposes. Name of the country, name of experimental site, site code, site coordinates, soil series and classification were among the data entered in this utility. Initial soil water was set to field capacity for all locations for the calibration experiments, while for the breeder evaluation this condition was not set, leaving the inputted moisture properties of the soils in each location. Measured soil characteristics taken from each profile were used to calculate the soil physical and chemical parameters that are needed to run the model. For calibration experiments, we assumed that N was not limiting while for the breeder evaluation nitrogen was simulated although N stress was not recorded in any of the locations. For the evaluation experiments however, Nitrogen was simulated, and application was done according to treatments. For other simulation options, initial conditions were as reported for each year and location, the Priestly-Taylor/Ritchie method was selected for simulation of evapotranspiration while the Soil Conservation Service (SCS) method was selected for simulation of infiltration. Photosynthesis was simulated using the radiation use efficiency method, while hydrology and soil evaporation were simulated using the Ritchie Water Balance and Suleiman-Ritchie methods respectively. Phosphorus and Potassium were not simulated in all trials and locations.

2.6 Estimating genotype specific parameters

The GENCALC program of the DSSAT (Version 4.6) was used to calibrate the GSPs of the maize varieties. GENCALC is a software package that facilitates the calculation of variety coefficients for use in existing crop models including the CSM-CERES-Maize Model [25]. The CSM–CERES–Maize model has GSPs that define growth and development characteristics or traits of a maize variety (Table 1). Three parameters (P1, P2 and P5) define the life cycle development characteristics, two coefficients (G2 and G3) define growth and yield characteristics and one coefficient, PHINT, defines leaf tip appearances [10]. All the candidate genetic coefficients were selected and calculated using GENCALC except P2 because all the varieties used were day-neutral. Conventionally day-neutral varieties should have constant P2 value, ideally the value should be zero which means that the variety does not generate delays when photoperiods exceed 12.5 hours. In our calibration procedure, a small positive number (0.01) was used as P2 for all varieties so that computer arithmetic problems like division by zero are prevented.

The varieties used in the trials were representative of all the maturity groups, i.e. extra early to late maturity. The default values in DSSAT were therefore used as initial coefficients for the extra-early, early and late maturity classes. Variety coefficient values for each variety are then varied, relative to each simulated and observed measurement. The model algorithm then

searches the output file and uses the difference between simulated and observed variables to decide whether to increase or decrease the value of the coefficient that is being estimated. When GENCALC finds a good fit for each observation, it averages the coefficients and calculates the root mean square error (RMSE) [26]. According to each genetic parameter, the process is repeated until the best fit is selected. An interactive procedure is used by GENCALC where the user changes the variety coefficient step to minimize the errors and speed-up the convergence of the algorithm. The search finishes when the user accepts the parameters providing the lowest RMSE for a single target trait.

For calibration of maize genotypes using the experimental data, four variables connected to four out of the six coefficients were directly measured (P1, P5, G2 and PHINT), while P2 was not estimated because all the varieties used in the experiments were day-neutral. G3 of the initial genotypes were first selected and later adjusted using a set of truncated rules in the GENCALC2.rul file until a good fit is observed. For calibration using the breeder data, five out of the six coefficients (except P2) were estimated following an optimization procedure (Fig 1) similar to that used by Anothai et al. [16]. This approach has not been reported for maize, especially in Sub-Saharan Africa. At each step of the calibration process in GENCALC, measured number of days from emergence to flowering was compared with days to anthesis (ADAP), measured number of days from emergence to physiological maturity was compared to days to physiological maturity (MDAP), measured grain yield at harvest was compared with harvest weight at maturity (HWAM), measured overall biomass at maturity was compared with tops weight at maturity (CWAM), measured maximum leaf area index was compared with maximum leaf area index (LAIX), while measured harvest index was compared with harvest index at maturity (HIAM).The generated coefficients were then used to run sensitivity analysis, using various iterations (not less than 6000 for each coefficient) to confirm the accuracy of the sequential approach. The adjustment for each target coefficient was done while all other non-target coefficients were kept constant. Despite the sensitivity analysis conducted, there is a possibility that pathologies associated with staged optimizations like GENCALC will occur thereby influencing the goodness of fit [27].

2.7 Model evaluation

The model was calibrated using data from conventional experiments or breeder evaluation trials. Model evaluation was done using data from the nitrogen trials (Table 2). The data sets used for model evaluation were of two types; single measured data and time series data. For single measured data, we used r^2 and root mean squared error (RMSE) (Eq 1) to evaluate the agreements between simulated and observed values. Normalized Root Mean Squared Error (nRMSE, Eq 2) and the index of agreement (d, Eq 3) [28] were used to evaluate the time series data. We used nRMSE for time series data because RMSE varies with growth over time as the magnitude of the growth variables increase. The d-statistic was used because it gives a single index of model performance, which covers bias and variability; it also indicates 1:1 prediction better than R^2. A low value for nRMSE (expressed in percent) is desired to define a good fit. The d statistic has values between zero and one, with one being the best fit. The modeling efficiency, EF [29] was employed to test modeling efficiency (Eq 4). EF has no dimension and an EF = 1 corresponds to a perfect match between observed and simulated data. When EF < 0, the simulated values are worse than simply using the observed mean. R^2, RMSE, RMSEn, d-index and EF are shown in Eqs 1, 2, 3 and 4 respectively.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(m_i - Si)^2}{\sum_{i=1}^{n}(m_i - \bar{m})^2} \qquad (2)$$

Fig 1. Order sequence of optimizations for calibrating the cultivar coefficients using GENCALC.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (m_i - s_i)^2}{n}} \tag{3}$$

$$RMSEn = \frac{RMSE \times 100}{\bar{m}} \tag{4}$$

$$d = 1 - \frac{\sum_{i=1}^{n} (m_i - S_i)^2}{\sum_{i=1}^{n} (|S_i| + |m_i|)^2} \tag{5}$$

Table 4. Generated genotype specific parameters (GSPs) using experimental and breeder data.

Variety	P1		P2		P5		G2		G3		PHINT	
	Experiment	Breeder	Expt.	Breeder	Experiment	Breeder	Experiment	Breeder	Expt.	Breeder	Experiment	Breeder
Ife hybrid 6	223.6 (11.16)*	247.4	0.01	0.01	520.7 (6.77)	518.3	706.7 (13.89)	663.7	7.09	6.98	36.90 (0.18)	35.70
Sammaz 41	233.6 (9.77)	263.2	0.01	0.01	550.7 (9.12)	540.4	806.9 (16.33)	782.1	7.76	7.59	37.00 (0.19)	39.66
Ife hybrid 5	213.7 (10.83)	221.6	0.01	0.01	511.6 (6.31)	502.7	518.7 (9.17)	533.7	7.47	6.99	40.00 (0.21)	39.03
Sammaz 42	230.0 (5.75)	244.3	0.01	0.01	683.4 (5.16)	679.2	786.7 (16.44)	806.4	7.59	7.72	45.50 (0.23)	39.98
OBA 9	293.1 (8.33)	288.6	0.01	0.01	768.1 (7.11)	772.9	828.7 (12.88)	830.7	7.83	7.80	45.00 (0.25)	45.00
Seedco White	289.8 (6.98)	284.1	0.01	0.01	781.8 (7.32)	778.8	834.1 (11.13)	829.6	8.42	8.51	41.20 (0.19)	42.90
Sammaz 34	287.0 (8.11)	283.7	0.01	0.01	596.0 (5.12)	589.7	827.0 (9.22)	822.6	6.77	6.39	40.00 (0.21)	40.00
Sammaz 32	282.0 (7.29)	233.9	0.01	0.01	601.0 (4.61)	692.7	822.0 (8.76)	788.1	6.55	6.62	45.04 (0.27)	43.21
IITA E White	270.0 (8.91)	221.6	0.01	0.01	614.3 (5.33)	622.2	713.4 (12.13)	759.7	6.58	7.07	45.00 (0.26)	50.20
IITA EE White	183.6 (9.51)	192.3	0.01	0.01	601.0 (6.19)	627.8	523.3 (10.16)	614.3	6.91	7.32	42.10 (0.21)	44.35

*Numbers in parenthesis are Standard Errors (SEM) for the measured experiment values

$$EF = \frac{\sum_{i=1}^{n}(m_i - \bar{m})^2 - \sum_{i=1}^{n}(s_i - m_i)^2}{\sum_{i=1}^{n}(m_i - \bar{m})^2} \tag{6}$$

Where n is the number of observations, S_i is the simulated data, m_i is the measured data, and $\bar{m}$ is the mean of the measured data.

3.0 Results

3.1 Calibration and breeder evaluation experiments

3.1.1 Genotype specific parameters. The values of GSPs generated using data from calibration experiments and breeder evaluation data are shown in Table 4. The highest degree days from emergence to end of Juvenile stage (P1) was recorded for OBA 9 in both experimental and breeder data. For number of days from silking to end of physiological maturity (P5), the highest values were recorded for Seedco white in both the experimental and breeder data. The lowest P1 values were recorded for IITA EE white using both experimental and breeder data. The variety Seedco white produced the largest number of maximum possible kernels (G2) for experimental data while OBA 9 had the highest values for breeder data. The value of G3 (kernel filling rate) ranged between 6.55 and 8.42 for the experimental data, and between 6.39 and 8.51 for the breeder data. Phyllochron interval (PHINT) values ranged from 36.9 and 45.5°Cd for the experimental data and between 35.7 and 50.2° Cd for the breeder data. The results show that about half of the GENCALC estimates are near to or beyond two SEMs away from measured values. Majority of these estimates are for the phenology parameters P1, P5 and PHINT.

3.1.2 Phenology and growth. Evaluation of CERES-Maize for grain yield, number of days to anthesis, number of days to physiological maturity and plant height and using both calibration experiments and breeder evaluation is shown in Fig 2 for two varieties. Calibration of number of days to anthesis, and plant height, were more accurate when experimental data were used compared with breeder data for both varieties. Calibration of both variables using experimental data resulted in d-index values in the range of 0.85–0.96 for the trial data. For the breeder data however, d-index values ranged from 0.49 to 0.89. Days to anthesis was calibrated with higher accuracy than plant height for all varieties. Number of leaves per plant and plant
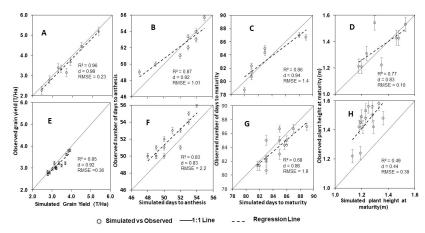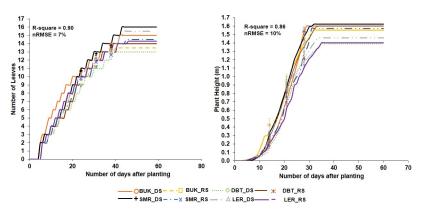
Fig 2. Comparisons between simulated and observed grain yield, days to anthesis, days to maturity and plant height at harvest for SAMMAZ 32 using experiment (A, B, C, D) and breeder (E, F, G, H) data. Solid lines = 1:1 lines; dashed lines = regression lines. Error bars denote Standard Error of Mean (SEM).
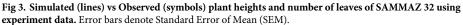
https://doi.org/10.1371/journal.pone.0200118.g002

height were measured for the experimental data at different time intervals. The simulated values for both plant height and number of leaves were accurate at all sampling periods (Fig 3).
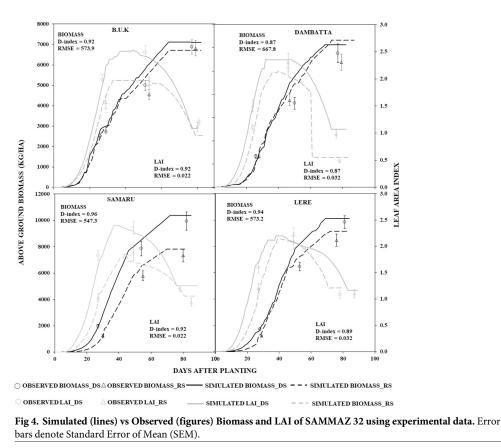
3.1.3 Biomass and leaf area index. Biomass and LAI were measured at juvenile stage, at anthesis, and at physiological maturity for the calibration data only. Fig 4 shows the result of simulation of above-ground biomass and LAI for Sammaz 32 across the trial locations. Good agreements were found between simulated and observed variables for all other varieties. Biomass was simulated with higher accuracy than LAI across all locations. Simulation of both biomass and LAI were most accurate using data from Samaru (d-index = 0.96, RMSE = 547.3 for biomass and d-index 0.92, RMSE 0.022 for LAI). Calibration of both variables had the lowest accuracy at Dambatta. Agreements between observed and simulated LAI were closer for the earliest measurement (juvenile stage), followed by measurement at anthesis, and physiological maturity in all locations except at Samaru where the reverse was observed. For biomass however, measurement at physiological maturity produced the closest agreements between observed and simulated values, while measurement at anthesis produced the lowest agreement between observed and simulated variables.

3.1.4 Yield and yield attributes. Yield and yield attributes were well calibrated for all varieties in both calibration and breeder datasets. Table 5 shows the result of comparisons between

Fig 3. Simulated (lines) vs Observed (symbols) plant heights and number of leaves of SAMMAZ 32 using experiment data. Error bars denote Standard Error of Mean (SEM).

https://doi.org/10.1371/journal.pone.0200118.g003

Fig 4. Simulated (lines) vs Observed (figures) Biomass and LAI of SAMMAZ 32 using experimental data. Error bars denote Standard Error of Mean (SEM).

observed and simulated mean grain yields of all varieties across different locations. Calibration of grain yield using experimental data was more accurate, as evidenced by low percentage prediction deviations (3.1 to 12.9). Values for model statistics were also good for the experimental data (RMSE = 264.6 kg ha^{-1}, nRMSE = 11.1%, and d-index = 0.97). For the breeder data however, prediction deviations of up to 24.7% were observed, with higher RMSE (510.1 kg ha^{-1}) and nRMSE (16.1%). Negative prediction deviation which indicate under simulation was only observed in one location (BGD 13) for the breeder evaluation data, while in all instances positive prediction deviations were observed.

3.2 Model validation experiments

Grain and tissue nitrogen, as well as grain yield, at harvest were simulated using independent datasets from trials conducted at BUK during the rainy seasons between 2013 and 2016. Simulations were done using GSPs generated from both experimental and breeder data. Table 6 shows the comparison between observed and simulated grain yields with accompanying model statistics for the two datasets taking SAMMAZ 32 and EE-White as examples. Grain yield was well simulated for both varieties using both datasets, although better fits were observed for GSPs from the calibration data. Nonetheless, low values of RMSE (below 2% of mean for experimental and 4.5% for breeder), high values of d index (0.99 for experimental and 0.96 for breeder) and good EF values (slightly less than 1 for both datasets) were observed. Tables 7 and 8 shows comparisons of simulated grain and stover nitrogen using GSPs generated from calibration and breeder evaluation experiments. Better agreements between observed and simulated grain and stover Nitrogen were observed at high Nitrogen (120 and 60

Table 5. Observed and simulated mean grain yields (kg ha^{-1}) of all varieties across different locations.

Data Type	Observed*	Simulated	PD%#
Experiment Data			
BUK_DS	3828	4080	6.6
BUK_RS	3209	3489	8.7
DBT_DS	2758	2866	3.9
DBT_RS	2628	2709	3.1
SMR_DS	5030	5259	4.6
SMR_RS	3536	3887	9.9
LERE_DS	4561	4723	3.6
LERE_RS	3452	3896	12.9
RMSE (kg/ha)		264.6	
nRMSE (%)		11.4	
EF		0.91	
d-index		0.97	
Breeder Data			
ZRA 12	2958	3345	5.4
ZRA 13	2969	3625	24.7
MKW 12	3214	3866	15.0
MKW 13	3042	3213	2.5
BGD 12	3812	3913	5.5
BGD 13	2782	2885	-3.3
BTR 12	3226	3110	8.9
BTR 13	3112	3487	13.1
SMR 12	3214	3863	18.0
SMR 13	3779	4329	6.3
MJB 12	3612	3746	7.5
MJB 13	2831	2470	5.7
KDW 12	2711	2779	7.5
KDW 13	3017	3956	22.2
RMSE (kg/ha)		510.1	
nRMSE (%)		16.1	
EF		0.52	
d-index		0.78	

* Mean for all varieties

Percentage prediction deviation

https://doi.org/10.1371/journal.pone.0200118.t005

Kg N) for both calibration and breeder evaluation experiments. At zero nitrogen application however, the agreements between observed and simulated values where low as evidenced by higher RMSE and lower d-index values.

4.0 Discussion

Calibrated GSPs from the on-station experiments and breeder evaluation experiments were similar to GSPs reported for related varieties in West and Southern Africa [30–32] with respect to yield and yield attributes. For growth and phenology however, data from our experiments produced better calibration of growth and phenology than earlier reported experiments in the Nigerian Savannas. For calibration using both experimental and breeder data, we set the values of P2 to 0.01 to simulate the day-neutral characteristics of all the varieties used. Recent

Table 6. Simulated vs Observed grain yields of Sammaz 32 and EE White in the model validation experiments, under different nitrogen levels using GSPs derived from calibration experiment and breeder evaluation experiment.

Treatment	Observed	Simulated (GSPs < Calibration experiment)	Simulated (GSPs < Breeder evaluation trials)
		Sammaz 32	
0 kg N	1245	1291	1177
60 kg N	2648	2573	2592
120 kg N	3255	3308	2983
SE±	57.3		
RMSE		36.3	101.1
D-Index		0.99	0.97
EF		0.92	0.91
		EE White	
0 kg N	979	953	1024
60 kg N	2177	2062	2333
120 kg N	3092	3129	3291
SE±	60.6		
RMSE		43.6	90.8
D-Index		0.99	0.98
EF		0.96	0.91

Table 7. Comparison of simulated and observed grain nitrogen (kg N ha^{-1}) of SAMMAZ 32 for GSPs generated using calibration experiments and breeder evaluation experiments.

	SIM (Calibration Experiments)	SIM (Breeder Evaluation Expts.)	OBS
	120 kg N ha^{-1}		
BUK 13	42.9	44.8	42.1
BUK 14	45.8	46.9	44.3
BUK 15	44.4	45.3	42.2
BUK 16	42.3	45.1	43.3
SE±			0.81
RMSE	1.48	2.59	
d-index	0.67	0.47	
	60 kg N ha^{-1}		
BUK 13	44.0	45.2	43.2
BUK 14	44.9	42.4	43.3
BUK 15	40.6	42.3	41.0
BUK 16	42.7	46.8	43.6
SE±			0.79
RMSE	1.02	2.05	
d-index	0.87	0.59	
	0 kg N ha^{-1}		
BUK 13	12.9	10.7	14.3
BUK 14	20.1	22.3	21.8
BUK 15	21.4	26.8	20.6
BUK 16	7.8	11.6	10.2
SE±			0.94
RMSE	1.68	3.66	
d-index	0.98	0.88	

Table 8. Comparison of simulated and observed stover nitrogen (kg N ha^{-1}) of SAMMAZ 32 for GSPs generated using data from calibration experiments and breeder evaluation experiments.

	SIM (Calibration Experiments)	SIM (Breeder Evaluation Expts.)	OBS
120 kg N ha^{-1}			
BUK 13	79.6	83.2	78.7
BUK 14	74.7	89.2	76.5
BUK 15	74.6	80.5	72.6
BUK 16	80.4	92.6	76.3
SE±			17.8
RMSE	2.49	11.3	
d-index	0.88	0.31	
60 kg N ha^{-1}			
BUK 13	64.9	73.8	67.8
BUK 14	81.4	88.5	77.4
BUK 15	86.7	81.3	78.8
BUK 16	70.8	70.0	62.6
SE±			4.6
RMSE	6.2	7.4	
d-index	0.89	0.80	
0 kg N ha^{-1}			
BUK 13	21.7	27.3	23.2
BUK 14	26.2	32.2	27.4
BUK 15	32.3	40.1	30.6
BUK 16	15.7	16.6	14.3
SE±			2.47
RMSE	1.46	5.8	
d-index	0.97	0.81	

publications by Lamsal et al. [21,33] highlighted the need to test for possible pathologies while estimating GSPs in crop simulation models. Pathologies like expressivity failure (occur when a model cannot reproduce some observations for any combination of GSP values due to how the models' mathematical structure is set up), equifinality (multiple GSP combinations producing exactly same model predictions), and environmental hypersensitivity (GSP estimates depend on the environments used in generating them) should be checked especially if GSPs generated are to be used for genetic mapping. The presence of equifinality in our results is suggested by the closeness of the predictions generated from the two sets of GSPs even though 50% of the estimates differ by close to two SEMs or more. However, while more detailed testing for this phenomenon might be useful future work, our goal was to assess the degree of alignment between model predictions and observations given different sources of calibration data and this has been shown to be adequate.

A high percentage (75%) of the GENCALC estimates that are near to or beyond two SEMs of the measured values were recorded for coefficients that determine phenology and therefore dependent on accurate measurement of developmental events (in observed days) and subsequent conversion to degree days. This high percentage shows that phenological events like number of days to flowering and number of days to maturity were not accurately measured in the breeder experiments due to small sample sizes and because they are not the traits of interest in the breeding program. This is evidenced for example by the under-simulation of days to flowering by 2.2 days and over simulation of days to maturity by 1.8 days for SAMMAZ 32

shown in Fig 2. The implication of poor phenology measurements is seen by a slight over-estimation of yield and yield attributes thereby confirming assertions made by Kumudini et al. [34] who suggested that accurate prediction of phenology is fundamental to determining crop adaptation and yield potential.

Calibration of the GSPs using the on-station experiment datasets produced better model fits than the breeder evaluation data as expected. The closeness of fit observed for the on-station data could be attributed to better experimental sites (soils with higher fertility and better moisture retention), better crop management (timely weeding, fertilizer application etc.) and higher experimental precision. This is evidenced by the breeder data having higher experimental errors for all measured variables when compared to the evaluation experiments. The evaluation experiments were also done on larger plot sizes and no missing plants were recorded at harvest, while in the breeder data smaller plots were used and there were no considerations for missing plants during yield calculations. In addition, for the experimental datasets more plant-related variables were measured compared to the breeder evaluation experiment data where only grain yield, days to flowering, plant height and days to physiological maturity were measured. For the breeder evaluation experiment, the closeness between observed and simulated plant heights was low. This could be attributed to the fact that most breeder trials are conducted under water limited conditions, thus rainfall variability may affect crop performance and data quality. Although the model can properly simulate water stress, no stress was observed in any of the breeder evaluation sites and years. Grain yield and days to anthesis were simulated more accurately than plant height for the breeder evaluation experiment. This can be attributed to the high number of datasets used (7 locations and 2 seasons). Anothai et al. [23] suggested that more accurate predictions of yield and phenology are observed when data is collected from many locations and seasons. For the on-station experiment, plant height, number of leaves, leaf area index, biomass, number of grains per meter square and grain yields were well calibrated as the differences between observed and simulated values were very minimal.

According to literature [23,35] when many years and locations are available, GSPs calibrated using breeder evaluation experiments produced very accurate comparisons between observed and simulated growth, yield and phenology of maize. As suggested by Fensterseifer et al. [35], uncertainties exist in the reliability of model based simulations of growth, yield and phenology when calibrations are done using data from trials conducted under few environmental conditions. Also, several researchers [36,37] reported that the major factors determining the success of a model calibration process, which determines the applicability of the model on a larger scale is dependent on the wide variability of data used during the calibration process. Thorp et al. [38] suggested that for accurate calibration of crop models, integration of time variation using different planting dates and seasons, and spatial differences using different locations of datasets should be adopted for calibration of crop models using datasets from yield/breeder evaluation trials. To further verify these claims, we re-ran a couple of contrasting varieties under both on-station experiments and breeder evaluation experiments using different number of trials and data sets. For the on-station experiments, we first reduced the number of experiments by subtracting 2 stations concurrently (i.e. reducing from 8–6, 6–4 and 4–2). With every decrease in number of experiments, a subsequent decrease in model efficiency and increase in prediction error were recorded. The higher the number of trials the better the model fitted the observations, also reducing the number of experiments to 4 led to EF and d-index values below 0.4, while further reduction to 2 reduced the model efficiency to 0.25 and increased the prediction error to 55%. Using 4 experiments and all measured data produced the lowest level of acceptable model statistics (d-index $\geq$ 0.50, nRMSE $\leq$ 16% and EF $\geq$0.4). For the breeder evaluation experiments, every reduction in number of experiments led to a

Table 9. Model statistics values for reduction in number of experimental sites for both calibration experiments and breeder evaluation experiments.

No. Sites	Calibration Experiments	Breeder Evaluation Expts
8	EF = 0.93	EF = 0.88
	nRMSE = 6.9%	nRMSE = 8.9%
6 (8–2)	EF = 0.79	EF = 0.67
	nRMSE = 10.4%	nRMSE = 12.6%
4 (6–2)	EF = 0.51	EF = 0.44
	nRMSE = 16.4%	nRMSE = 18.9%
2 (4–2)	EF = 0.44	EF = 0.41

https://doi.org/10.1371/journal.pone.0200118.t009

decrease in model efficiency and an increase in prediction error. We also reduced the number of datasets from the evaluation experiments to the same that was used in the breeder evaluation experiments. This resulted in decrease in model efficiency (0.89 to 0.81 for Sammaz 32 and from 0.94 to 0.88 for Seedco white). This shows that the number of experimental sites is more important than amount of calibration data if the minimum data sets (MDS) are collected as shown in Table 9. This view is supported by [35] and [31]. When many locations and planting dates are available, data from breeder evaluation experiments in the SSA can be used to make good calibration of calibrations with lower RMSE & nRMSE and higher d-index & EF values.

Although the calibration experiments provided more accurate GSPs, they are still very expensive and laborious and thus are nearly impossible to carry out especially in Sub-Saharan Africa where expertise and resources are limiting. Breeder evaluation data could also be used for calibration of GSPs where such data is available. As shown earlier, very accurate GSPs could be generated if large amount of data from many years (also planting dates) and various locations are available. This will go a long way in providing model users with cheap and easy ways of calibrating GSPs of existing and newly released varieties to their locations.

Evaluating the generated GSPs for simulation of grain yield, tissue nitrogen and grain nitrogen using independent datasets resulted in good agreements between observed and simulated values. For grain yield, comparisons of measured and simulated values using both GSPs generated from experimental and breeder data showed very close agreements under medium and high nitrogen applications. For comparisons under nitrogen stressed conditions however, poor agreements existed between observed and simulated grain yields for both GSPs. This is a common occurrence with simulations of grain yield and yield attributes under low nitrogen fertilizer applications. Gungula et al. [39], reported that the CERES-Maize model poorly predicts performance of maize under low nitrogen conditions in the tropics. The agreements between observed and simulated grain and stover nitrogen for both GSPs under high fertilizer applications is an indication that CERES-model still performs best under high nitrogen applications especially on tropical soils.

The CERES-Maize model has been shown over the years to be an important tool in evaluating crop management [3], climate change impacts [40], fertilizer recommendations [5,39] and yield forecasting [41]. Calibrating the newly released maize varieties currently recommended for the Nigerian maize belts will provide an important input requirement for using crop models to evaluate major production constraints including optimum stand density (OSD), appropriate varietal selection (targeting/stability analysis), choice of major partner crop (in case of mixed cropping) and fertilizer (especially N and P) managements. The availability of accurate GSPs for all major varieties will also increase the applicability of the model on a wider scale and for broader applications.

5.0 Conclusion

Financial as well as time constraints coupled with frequent release of new varieties makes it difficult for model users to conveniently calibrate GSPs of crop models using detailed calibration experiments. Large numbers of evaluation trials are conducted across multiple locations under diverse planting dates by breeders and other growers prior to varietal release. Availability of such datasets, especially from evaluation trials conducted under minimal stress (moisture and nutrient) conditions provides an opportunity for efficient and rapid means of generating GSPs of newly released maize varieties. A systematic approach (as proposed in this study) as well as availability of large datasets from different locations and planting dates provide opportunities for estimation of accurate GSPs. Although it is possible to generate GSPs from breeder evaluation data, care must be taken to collect data from trials conducted under optimal conditions and not too far away from weather stations. Also, breeder data to be used for calibration of crop models must be collected from sites where detailed soil data is available. Additionally, appropriate tests must be conducted to ensure that pathologies such as equifinality, expressivity failures and environmental hypersensitivity are minimized especially when the objective is to generate GSPs for genetic mapping or for application under many environments where the estimation was not conducted. Availability of GSPs of new varieties as soon as they are released will help farmers and growers to make improved site-specific decision support tools (DST). Also, researchers will be provided with new ways to making variety groupings as well as studying complex Genotype, Environment, Management (G×E×M) interactions. Model users should endeavor to join breeding units/teams to ensure collection of robust data needed for model calibrations that are not traditionally collected by breeders.

Acknowledgments

Author Contributions

Conceptualization: A. A. Adnan, J. Diels, J. M. Jibrin, A. Y. Kamara, P. Craufurd, A. S. Shaibu.

Data curation: A. A. Adnan, Z. E. H. Tonnang.

Investigation: A. A. Adnan, Z. E. H. Tonnang.

Project administration: P. Craufurd, I. B. Mohammed.

Supervision: J. Diels, J. M. Jibrin, A. Y. Kamara.

Writing – original draft: A. A. Adnan.

Writing – review & editing: J. Diels, J. M. Jibrin, A. Y. Kamara, P. Craufurd, A. S. Shaibu, I. B. Mohammed, Z. E. H. Tonnang.

References

1. Badu-Apraku B, Fakorede MAB, Fontem LA, Akinwale RO. Improvement of Yield and other Traits of Extra-Early Maize under Stress and Nonstress Environments Development of Hybrid Maize for the Rainforest Agro-climatic Zone of SW Nigeria View project Molecular Characterization and Combining Ability of Early Maturing Maize Inbred Lines under Drought, low Soil Nitrogen, Striga Infestation and Optimal Growing Conditions View project. 2009; https://doi.org/10.2134/agronj2008.0089x

2. FAOSTAT. FAOSTAT statistics database [Internet]. 2017. Available: https://search.library.wisc.edu/catalog/999882363002121

3. MacCarthy DS, Adiku SGK, Freduah BS, Gbefo F, Kamara AY. Using CERES-Maize and ENSO as Decision Support Tools to Evaluate Climate-Sensitive Farm Management Practices for Maize Production in the Northern Regions of Ghana. Front Plant Sci. 2017; 8. https://doi.org/10.3389/fpls.2017.00031 PMID: 28184227

4. Zinyengere N, Crespo O, Hachigonta S, Tadross M. Crop model usefulness in drylands of southern Africa: an application of DSSAT. South African J Plant Soil. 2015; 32: 95–104. https://doi.org/10.1080/02571862.2015.1006271

5. Adnan A. A.; Jibrin J. M.; Kamara A. Y.; Abdulrahman B. L.; Shaibu AS. Using CERES-Maize model to determine the nitrogen fertilization requirements of early maturing maize in the Sudan Savanna of Nigeria. J Plant Nutr. 2017; 40: 1066–1082. https://doi.org/10.1080/01904167.2016.1263330

6. Adnan AA, Jibrin JM, Kamara AY, Abdulrahman BL, Shaibu AS, Garba II. CERES–Maize Model for Determining the Optimum Planting Dates of Early Maturing Maize Varieties in Northern Nigeria. Front Plant Sci. 2017; 8. https://doi.org/10.3389/fpls.2017.01118 PMID: 28702039

7. Iyanda R. A., Pranuthi G., Dubey S.K., and Tripathi SK. Use of dssat ceres maize model as a tool of identifying potential zones for maize production in Nigeria. Int J Agric Policy Res. 2014; 2: 69–75.

8. IBSNAT. Decision Support System for Agrotechnology Transfer (DSSAT). University of Hawaii, Honolulu. HI; 1994. p. 1776.

9. Román-Paoli E, Welch SM, Vanderlip RL. Comparing genetic coefficient estimation methods using the CERES-Maize model. Agric Syst. 2000; 65: 29–41. https://doi.org/10.1016/S0308-521X(00)00024-X

10. Jones CA, Kiniry JR (James R, Dyke PT. CERES-Maize: a simulation model of maize growth and development. Texas A & M University Press; 1986.

11. Ogoshi RM, Cagauan BG T. Field and laboratory methods for collection of minimum data sets. IBSNAT-ICASA, University of Hawaii, HI, USA.; 1999.

12. Suriharn B, Patanothai A, Pannangpetch K, Jogloy S, Hoogenboom G. Determination of Cultivar Coefficients of Peanut Lines for Breeding Applications of the CSM-CROPGRO-Peanut Model. Crops. 2007; 621: 607–621. https://doi.org/10.2135/cropsci2006.01.0050

13. du Toit AS. Comparisons of using fitted (calculated) and determined (measured) genetic coefficient G2 in CERES-Maize. South African J Plant Soil. 2002; 19: 208–210. https://doi.org/10.1080/02571862.2002.10634467

14. Welch SM, Wilkerson G, Whiting K, Sun N, Vagts T, Buol G, et al. Estimating soybean model genetic coefficients from private–sector variety performance trial data. Trans ASAE. American Society of Agricultural and Biological Engineers; 2002; 45: 1163. https://doi.org/10.13031/2013.9925

15. Seidel SJ, Palosuo T, Thorburn P, Wallach D. Towards improved calibration of crop models–Where are we now and where should we go? Eur J Agron. Elsevier; 2018; 94: 25–35. https://doi.org/10.1016/j.eja.2018.01.006

16. Anothai J, Patanothai A, Jogloy S, Pannangpetch K, Boote KJ, Hoogenboom G. A sequential approach for determining the cultivar coefficients of peanut lines using end-of-season data of crop performance trials. F Crop Res. 2008; 108: 169–178. https://doi.org/10.1016/j.fcr.2008.04.012

17. Mavromatis T, Boote KJ, Jones JW, Irmak a, Shinde D, Hoogenboom. Developing genetic for crop simulation models with data from crop. Crop sci Madison, Wis Crop Sci Soc Am 1961 Jan/Feb. 1998; 40–51. https://doi.org/10.2135/cropsci2001.41140x

18. Bannayan M, Hoogenboom G. Using pattern recognition for estimating cultivar coefficients of a crop simulation model. F Crop Res. Elsevier; 2009; 111: 290–302. https://doi.org/10.1016/J.FCR.2009.01.007

19. He J, Jones JW, Graham WD, Dukes MD. Influence of likelihood function choice for estimating crop model parameters using the generalized likelihood uncertainty estimation method. Agric Syst. Elsevier; 2010; 103: 256–264. https://doi.org/10.1016/J.AGSY.2010.01.006

20. Buddhaboon C, Jintrawet A, Hoogenboom G. Methodology to estimate rice genetic coefficients for the CSM-CERES-Rice model using GENCALC and GLUE genetic coefficient estimators. J Agric Sci. Cambridge University Press; 2018; 156: 482–492. https://doi.org/10.1017/S0021859618000527

21. Lamsal A, Welch SM, Jones JW, Boote KJ, Asebedo A, Crain J, et al. Efficient crop model parameter estimation and site characterization using large breeding trial data sets. Agric Syst. Elsevier; 2017; 157: 170–184. https://doi.org/10.1016/j.agsy.2017.07.016

22. Jones JW, Hoogenboom G, Porter CH, Boote KJ, Batchelor WD, Hunt LA, et al. The DSSAT cropping system model. European Journal of Agronomy. 2003. https://doi.org/10.1016/S1161-0301(02)00107-7

23. Anothai J, Patanothai A, Pannangpetch K, Jogloy S, Boote KJ, Hoogenboom G. Reduction in data collection for determination of cultivar coefficients for breeding applications. Agric Syst. Elsevier; 2008; 96: 195–206. https://doi.org/10.1016/J.AGSY.2007.08.002

24. Batchelor WD, Basso B, Paz JO. Examples of strategies to analyze spatial and temporal yield variability using crop models. Eur J Agron. Elsevier; 2002; 18: 141–158. https://doi.org/10.1016/S1161-0301(02)00101-6

25. Hunt LA, Pararajasingham S, Jones JW, Hoogenboom G, Imamura DT, Ogoshi RM. GENCALC: Software to facilitate the use of crop models for analyzing field experiments. Agron J. 1993; 85: 1090–1094. https://doi.org/10.2134/agronj1993.00021962008500050025x

26. Wallach D, Goffinet B. Mean squared error of prediction as a criterion for evaluating and comparing system models. Ecol Modell. 1989; 44: 299–306. https://doi.org/10.1016/0304-3800(89)90035-5

27. Welch SM, Zhang J, Sun N, Mak TY. Efficient estimation of genetic coefficients for crop models. In: Bowen WT, Malagamba P, Quiroz R, Holle M, White J, Leon-Velarde C, et al., editors. Proceedings—The Third International Symposium on Systems Approaches for Agricultural Development. International Potato Center, Lima, Peru.; 2001.

28. Willmott CJ, Willmott CJ. Some Comments on the Evaluation of Model Performance. Bull Am Meteorol Soc. American Meteorological Society; 1982; 63: 1309–1313. https://doi.org/10.1175/1520-0477(1982)063<1309:SCOTEO>2.0.CO;2

29. Loague K, Green RE. Statistical and graphical methods for evaluating solute transport models: Overview and application. J Contam Hydrol. 1991; 7: 51–73. https://doi.org/10.1016/0169-7722(91)90038-3

30. Jagtap SS, Abamu FJ, Kling JG. Long-term assessment of nitrogen and variety technologies on attainable maize yields in Nigeria using CERES-maize. Agric Syst. Elsevier; 1999; 60: 77–86. https://doi.org/10.1016/S0308-521X(99)00019-0

31. Jibrin JM, Kamara AY, Ekeleme F. Simulating planting date and cultivar effects on dryland maize production using CERES-maize model. African J Agric Res. 2012; 7: 5530–5536. https://doi.org/10.5897/AJAR12.1303

32. Chisanga Charles Bwalya, Phiri Elijah CS& HS. Evaluating CERES-Maize Model Using Planting Dates and Nitrogen Fertilizer in Zambia. J Agric Sci. 2015; 7: 1–19. Available: https://www.researchgate.net/publication/269695853

33. Lamsal A, Welch SM, White JW, Thorp KR, Bello NM. Estimating parametric phenotypes that determine anthesis date in Zea mays: Challenges in combining ecophysiological models with genetics. PLoS One. 2018; 13: 1–23. https://doi.org/10.1371/journal.pone.0195841 PMID: 29672629

34. Kumudini S, Andrade FH, Boote KJ, Brown GA, Dzotsi KA, Edmeades GO, et al. Predicting maize phenology: Intercomparison of functions for developmental response to temperature. Agron J. 2014; 106: 2087–2097. https://doi.org/10.2134/agronj14.0200

35. Fensterseifer CA, Streck NA, Baigorria GA, Timilsina AP, Zanon AJ, Cera JC, et al. On the number of experiments required to calibrate a cultivar in a crop model: The case of CROPGRO-soybean. F Crop Res. Elsevier; 2017; 204: 146–152. https://doi.org/10.1016/J.FCR.2017.01.007

36. Ruíz-Nogueira B, Boote KJ, Sau F. Calibration and use of CROPGRO-soybean model for improving soybean management under rainfed conditions. Agric Syst. 2001; 68: 151–173. https://doi.org/10.1016/S0308-521X(01)00008-7

37. Xiong W, Balkovič J, van der Velde M, Zhang X, Izaurralde RC, Skalský R, et al. A calibration procedure to improve global rice yield simulations with EPIC. Ecol Modell. Elsevier; 2014; 273: 128–139. https://doi.org/10.1016/J.ECOLMODEL.2013.10.026

38. Thorp KR, DeJonge KC, Kaleita AL, Batchelor WD, Paz JO. Methodology for the use of DSSAT models for precision agriculture decision support. Comput Electron Agric. 2008; 64: 276–285. https://doi.org/10.1016/j.compag.2008.05.022

39. Gungula DT, Kling JG, Togun AO. CERES-Maize Predictions of Maize Phenology under Nitrogen-Stressed Conditions in Nigeria. Agron J. American Society of Agronomy; 2003; 95: 892–899. https://doi.org/10.2134/AGRONJ2003.8920

40. Angulo C, Rötter R, Lock R, Enders A, Fronzek S, Ewert F. Implication of crop model calibration strategies for assessing regional impacts of climate change in Europe. Agric For Meteorol. Elsevier; 2013; 170: 32–46. https://doi.org/10.1016/J.AGRFORMET.2012.11.017

41. Soler CMT, Sentelhas PC, Hoogenboom G. Application of the CSM-CERES-Maize model for planting date evaluation and yield forecasting for maize grown off-season in a subtropical environment. Eur J Agron. 2007; 27: 165–177. https://doi.org/10.1016/j.eja.2007.03.002