Journal of Crop Improvement

ISSN: 1542-7528 (Print) 1542-7536 (Online) Journal homepage: https://www.tandfonline.com/loi/wcim20

Modeling groundnut (*Arachis hypogaea* L.) performance under drought conditions

Richard Oteng-Frimpong, Yussif Baba Kassim, Rukiya Danful, Richard Akromah, Alex Wireko-Kena & Stephen Forson

To cite this article: Richard Oteng-Frimpong, Yussif Baba Kassim, Rukiya Danful, Richard Akromah, Alex Wireko-Kena & Stephen Forson (2019) Modeling groundnut (*Arachis hypogaea* L.) performance under drought conditions, Journal of Crop Improvement, 33:1, 125-144, DOI: 10.1080/15427528.2018.1542363

To link to this article: https://doi.org/10.1080/15427528.2018.1542363

Published online: 27 Nov 2018.

JOURNAL OF CROP IMPROVEMENT
2019, VOL. 33, NO. 1, 125–144
https://doi.org/10.1080/15427528.2018.1542363

Taylor & Francis
Taylor & Francis Group

ARTICLE

OPEN ACCESS Check for updates

Modeling groundnut (*Arachis hypogaea* L.) performance under drought conditions

Richard Oteng-Frimpong [a], Yussif Baba Kassim[b], Rukiya Danful[b], Richard Akromah[b], Alex Wireko-Kena[b], and Stephen Forson[b]

[a]Council for Scientific and Industrial Research-Savanna Agricultural Research Institute (CSIR-SARI), Tamale, Ghana; [b]Department of Crop and Soil Sciences, Kwame Nkrumah University of Science Technology, Kumasi, Ghana

ABSTRACT

Erratic rainfall is often a limiting factor in the semi-arid regions where most groundnut cultivation occurs. As a result, ensuring availability of cultivars that possess inherent tolerance to drought stress has become a priority. Field and box (wooden boxes of 2 m length × 1 m width × 0.3 m depth) experiments were conducted under drought and non-drought conditions to identify physiological and agronomic traits correlated with pod yield (PY). Fifty (50) advanced breeding lines were evaluated. Linear models containing different combinations of total dry matter at maturity, crop growth rate, pod growth rate, partition coefficient, and harvest index were able to predict PY under intermittent drought (adjusted R^2 range: 0.9798–0.9895). The box experiment was more discriminating of genotypes than field experiments, making it a suitable technique for drought tolerance screening using specific leaf area and leaf chlorophyll content. As a result, screening and pre-selection using the seed-box technique before advanced evaluation on the field is recommended.

ARTICLE HISTORY
Received 6 July 2018
Accepted 26 October 2018

KEYWORDS
Intermittent drought; linear models; physiological and agronomic traits; principal component analysis; terminal drought

1. Introduction

Most of the groundnut cultivation occurs in the semi-arid regions where erratic rainfall is often a limiting factor and as a result, mid- and end-of-season droughts are critical as they directly affect pod yield (PY) and quality (Pasupuleti et al. 2016). Development of water-use efficient cultivars has therefore been an important breeding objective in groundnut improvement programs (Pasupuleti et al. 2016).

However, most groundnut breeding programs follow an empirical approach to screening for drought tolerance, which is largely based on kernel yield and traits of local adaptation (Nigam et al. 2005). This slows down breeding progress because of the influence of genotype-by-management interaction on yield. The use of physiological traits in a breeding program either by direct selection or through a surrogate, such as

CONTACT Richard Oteng-Frimpong kotengfrimpong@gmail.com Council for Scientific and Industrial Research-Savanna Agricultural Research Institute (CSIR-SARI), P.O. Box TL 52, Tamale, Ghana

molecular markers, depends on their relative genetic correlation with yield, extent of genetic variation, heritability, and genotype-by-environment interactions (Pandey and Shukla 2016). A large number of studies have identified several physiological traits, whose presence is associated with plant adaptability to drought-prone environments (Akbar et al. 2017; Arunyanark et al. 2009; Cruickshank et al. 2004; Kiniry et al. 2005; Krishnamurthy et al. 2013). These putative traits are more effective in enhancing drought tolerance and contribute to PY under water-stress conditions (Ludlow and Muchow 1990).

On the other hand, the interaction between plants and their environment involves elaborate biological, physical, and chemical processes that affect plants and plant ecosystem development (Měch and Prusinkiewicz 1996). Hence to understand the responses of crops to their environments, models are being used to study both simple and complex aspects of this system (Hoogenboom, Jones, and Boote 1992). Simulation models are very powerful tools for critically assessing the value of putative traits (Ludlow and Muchow 1990). The groundnut model "PNUTGRO" has been used to predict phenology, growth, and yield (Boote, Jones, and Singh 1992; Hoogenboom, Jones, and Boote 1992; Kaur and Hundal 1999; Singh et al. 1994). However, simulating events using this model has reportedly shown deviations for flowering, pegging and physiological maturity (Kaur and Hundal 1999; Singh et al. 1994). Singh et al. (1994) have indicated that this model should work best under stress-free situations, which is an unrealistic proposition. This suggests that new models need to be developed and tested (Ludlow and Muchow 1990). Therefore, the main objectives of this study were: (i) to identify physiological and agronomic traits associated with PY under drought conditions, (ii) to develop a trait-based model that can be used as a tool for predicting PY under drought conditions, and (iii) to propose a technique based on physiological traits that can be used to rapidly screen genotypes for drought tolerance.

2. Materials and methods

2.1. *Design of experiments and management*

The study was conducted at the CSIR-Savanna Agricultural Research Institute (CSIR-SARI), Ghana. The study location falls within the Guinea Savanna Agro-ecological zone of Ghana (N 9°23′38.2″ and W 1°00′18.4″). The study comprised two experiments; one conducted using seed boxes and the other conducted under field conditions. Fifty (50) advanced breeding lines (Table 1), originally selected for diverse traits, were used in this study. All genotypes used were part of the SARI groundnut germplasm obtained from diverse sources.

Table 1. Description of genotypes used in this study.

No.	Cultivar	Growth habit	Branching pattern†	Subspecies	Botanical type	Agronomic attributes‡
1	SARGV 016	Unknown	Unknown	Unknown	Unknown	Unknown
2	NKATIE SARI	Spreading and bunch	Irregular 1	Fastigiata	Spanish	ELS, LLS
3	ICGV 07356	Unknown	Unknown	Unknown	Unknown	DT
4	ICGV 00350	Bunch	Irregular 2	Hypogaea	Virginia	DT
5	ICGV 93305	Spreading	Sequential	Fastigiata	Spanish	SD, AT
6	ICGV 00005	Unknown	Unknown	Unknown	Unknown	SD, ELS, LLS
7	12CS-042	Spreading and bunch	Irregular 1	Fastigiata	Spanish	SD, ELS, LLS
8	ICGV 01276	Bunch	Irregular 2	Hypogaea	Virginia	SD, ELS, LLS
9	ICGV 86124	Bunch	Sequential	Fastigiata	Spanish	DT
10	ICGV 91315	Spreading and bunch	Irregular 1	Fastigiata	Spanish	SD, AT
11	ICGV 94379	Spreading and bunch	Irregular 1	Fastigiata	Spanish	SD, AT, ELS, LLS
12	ICGV 02271	Unknown	Unknown	Unknown	Unknown	DT
13	ICGV 03056	Bunch	Irregular 1	Fastigiata	Spanish	DT
14	ICGV 00068	Unknown	Unknown	Unknown	Unknown	SD, ELS, LLS
15	12CS-008	Spreading and bunch	Irregular 1	Fastigiata	Spanish	SD, ELS, LLS
16	ICGV 91284	Spreading	Alternate	Hypogaea	Virginia	SD, ELS, LLS
17	SARGV 019	Unknown	Unknown	Unknown	Unknown	Unknown
18	ICGV 86015	Bunch	Irregular 2	Hypogaea	Virginia	DT
19	12CS-111	Spreading and bunch	Irregular 2	Hypogaea	Virginia	SD, ELS, LLS
20	ICGV 92302	Spreading	Irregular 1	Fastigiata	Spanish	SD, AT
21	SARGV 008	Unknown	Unknown	Unknown	Unknown	Unknown
22	ICGV 97188	Spreading and bunch	Sequential	Fastigiata	Spanish	DT
23	CHINESE	Erect	Sequential	Fastigiata	Spanish	SD
24	ICGV 91317	Spreading	Sequential	Fastigiata	Spanish	SD, AT
25	SARGV 009	Unknown	Unknown	Unknown	Unknown	Unknown
26	ICGV 89104	Bunch	Irregular 1	Fastigiata	Spanish	SD, ELS, LLS
27	12CS-041	Spreading and bunch	Irregular 1	Fastigiata	Spanish	SD, ELS, LLS
28	ICGV 91328	Spreading	Sequential	Fastigiata	Spanish	SD, AT
29	FLEUR 11	Bunch	Irregular 1	Fastigiata	Spanish	DT
30	SARGV 010	Unknown	Unknown	Unknown	Unknown	Unknown
31	ICGV 89767	Bunch	Irregular 2	Hypogaea	Virginia	SD, ELS, LLS
32	12CS-004	Spreading	Irregular 2	Hypogaea	Virginia	SD, ELS, LLS
33	12CS-058	Spreading and bunch	Irregular 2	Hypogaea	Virginia	SD, ELS, LLS
34	ICGV 91324	Spreading	Sequential	Fastigiata	Spanish	SD, AT
35	CN-94C	Bunch	Irregular 1	Fastigiata	Spanish	DT
36	ICGV 99240	Spreading	Sequential	Fastigiata	Spanish	DT
37	ICGV 91114	Unknown	Unknown	Unknown	Unknown	DT
38	YENYAWOSO	Spreading	Irregular 2	Hypogaea	Virginia	SD
39	KPANIELLI	Erect	Sequential	Fastigiata	Spanish	Unknown
40	ICGV 99247	Unknown	Unknown	Unknown	Unknown	DT
41	ICGV 00064	Unknown	Unknown	Unknown	Unknown	SD, ELS, LLS
42	SUMNUT 23	Bunch	Irregular 1	Fastigiata	Spanish	SD
43	55–437	Bunch	Irregular 2	Hypogaea	Virginia	DT
44	ICGV 91279	Spreading	Sequential	Fastigiata	Spanish	SD, AT
45	ICGV 99241	Spreading	Sequential	Fastigiata	Spanish	DT

(*Continued*)

Table 1. (Continued).

No.	Cultivar	Growth habit	Branching pattern†	Subspecies	Botanical type	Agronomic attributes‡
46	ICGV 93328	Spreading	Sequential	Fastigiata	Spanish	SD, AT
47	ICGV 91,278	Spreading	Sequential	Fastigiata	Spanish	SD, AT
48	ICGV 86024	Bunch	Irregular 1	Fastigiata	Spanish	DT
49	SARGV 018	Unknown	Unknown	Unknown	Unknown	Unknown
50	ICGV 99029	Unknown	Unknown	Unknown	Unknown	SD, ELS, LLS

†The branching pattern of the genotypes were assigned as described by Ibpgr (1992) and Pittman (1995).
‡SD: short duration; AT: aflatoxin tolerance; ELS; early leaf spot tolerance; LLS: late leaf spot tolerance; DT: drought tolerance.

2.1.1. *Seed-box experiment*

Three wooden seed boxes (2 m length × 1 m width × 0.3 m depth) were constructed. The boxes were lined with a black polythene sheet to ensure that water was not lost through any means other than evapotranspiration. Topsoil from SARI experimental field was used to fill the boxes. Seeds were sown in single rows. The inter-row and intra-row spacings were 20 and 10 cm, respectively giving a plant population of 100 plants per box. Four seeds were sown per hill; thinning was done 14 days after planting to keep only one plant per hill. All recommended production practices for groundnut were followed. Water supply was terminated at the flowering stage to simulate terminal drought.

2.1.2. *Field experiment*

Three environments differing in drought scenarios were created using different watering regimes. The first treatment (well irrigated) received regular irrigation from planting till maturity. The second treatment (terminal drought) received regular irrigation from planting to flowering, after which no irrigation was applied. The third treatment (intermittent drought) received irrigation from planting to pod initiation, after which irrigation was given only when there were signs of permanent wilting. To limit the influence of rainfall, the experiment was conducted under the post-rainy season (December 2016–May 2017). The experiment was conducted using rectangular lattice design with a block size of two. Each treatment was replicated three times. One seed per hill was planted in 4 m long single-row plots. Seeds that did not emerge were refilled at 14 days after planting. Inter- and intra-row spacings were 0.5 and 0.1 m, respectively, giving a plant population of 40 plants plot^{-1}. Phosphorus was applied at emergence in the form of triple super phosphate at a rate of 60 kg P_2O_5 ha^{-1}. Plots were further supplemented with ground oyster shells at a rate of 200 kg ha^{-1} to supply calcium.

2.2. *Data collection*

Cumulative thermal time (CTT, °Cd) was used as an objective method of determining growth stages (Rao, Nigam, and Huda 1992). Data collection was

done at 408.00, 932.80, and 2369.70 °Cd, which corresponded with flowering, pod initiation, and physiological maturity, respectively. Data were collected on plant stand, canopy width (cm), canopy height (cm), days to 50% flowering (° Cd), leaf area (LA, cm^2) using a portable leaf area meter (YMJ-B), leaf dry weight (LDW, g) after oven-drying leaves at 60°C for 72 h, specific leaf area (SLA, cm^2g^{-1}), leaf chlorophyll content using a SPAD chlorophyll meter (SPAD-502Plus, Hangzhou Mindfull Technology Co. Ltd, Shunda, China), leaf temperature (°C) using an infra-red meter (Extech 42510A, Agriculture Solutions, Strong ME, USA), dry biomass (DM, tha^{-1}), PY, tha^{-1}, total dry biomass (TDM, tha^{-1}), harvest index (HI, %), crop growth rate (CGR, $g°Cd^{-1}$), pod growth rate (PGR, $g°Cd^{-1}$), partition coefficient (PCt), and percentage yield penalty relative to biomass (YP_{DM}, %) and to pod yield (YP_{PY}, %).

The CTT, SLA, TDM, HI, CGR, PGR, PCt, YP_{DM}, and YP_{PY} were estimated using the equations below, respectively:

$$\mathrm{CTT} = \sum_{\mathrm{Plt}}^{\mathrm{Har}} \frac{T_{max} + T_{min}}{2} - T_{base},$$

$$\mathrm{SLA} = \frac{\mathrm{LA}}{\mathrm{LDW}},$$

$$\mathrm{TDM} = \mathrm{DM} + (\mathrm{PY} \times 1.65),$$

$$\mathrm{HI} = \frac{\mathrm{PY}}{\mathrm{TDM}} \times 100,$$

$$\mathrm{CGR} = \frac{\mathrm{DM} + (\mathrm{PY} \times 1.65)}{T_t},$$

$$\mathrm{PGR} = \frac{\mathrm{PY} \times 1.65}{T_t - T_f},$$

$$\mathrm{PCt} = \frac{\mathrm{PGR}}{\mathrm{CGR}},$$

$$\mathrm{YP_{DM}} = \frac{\mathrm{DM_{ns}} - \mathrm{DM_s}}{\mathrm{DM_{ns}}} \times 100,$$

$$YP_{PY} = \frac{PY_{ns} - PY_s}{PY_{ns}} \times 100,$$

where T_{max} = daily maximum temperature, T_{min} = daily minimum temperature, T_{base} = mean base temperature for groundnut, Plt = start date, and Har = end date, T_t = duration from sowing to harvest in CTT, T_f = duration from sowing to 50% flowering in CTT, DM_{ns} = biomass yield of non-stressed plot, DM_s = biomass yield of stressed plot, PY_{ns} = pod yield of non-stressed plot, and PY_s = pod yield of stressed plot. A base temperature of 13°C (Nigam 2014) was used. Pod weight was multiplied by a correction factor of 1.65 to adjust for the differences in the energy requirement for producing pod dry matter compared with vegetative part (Duncan et al. 1978).

2.3. *Weather and soil parameters*

Meteorological and soil data were recorded using a data logger (Em50 Data Collection System, Decagon Devices, Inc., Pullman, USA.) positioned in the middle of the field experiment. The data logger was set to take readings on an hourly basis. Data were recorded on precipitation (mm), ambient temperature (°C), relative humidity (%), vapor pressure (kPa), volumetric soil moisture content (m^3m^{-3}), and soil water potential (kPa).

2.4. *Statistical analyses*

Environment-specific analysis was done using "agricolae" (Felipe 2017) package of the R statistical software (version 3.4.3) (R Core Team, 2018). The analysis for each trait in an individual environment was based on the following linear additive model:

$$Y_{ijk} = \mu + r_k + b_{jk} + g_i + \varepsilon_{ijk},$$

where Y_{ijk}, μ, r_k, b_{jk}, g_i, and ε_{ijk}, respectively, denote the observation on genotype i in block j of replication k, general mean, effect of replication k, effect of block j within replication k, effect of genotype i, and the residual effect.

Using the mean performance of the genotypes for each trait under the intermittent drought environment, a principal component analysis (PCA) was done on a genotype-by-trait matrix of 25 genotypes (training sample) using FactoMineR (Sebastien, Julie, and Francois 2008) package of the R statistical software, to identify the most important traits related to PY under intermittent drought conditions. Based on the PCA output, models were generated using traits that were in the same quadrant as PY and/or had high correlations with the axis of the PCA plot.

With the "regsubsets ()" function of leaps (Lumley 2017) package of the R statistical software, all subsets regression (traits were combined in all possible ways) was done to obtain linear regression models using the training sample. Genotype-by-trait biplot was used to determine correlations among predictor variables after the data were standardized (mean of zero and standard deviation of one). Multicollinearity between predictor variables was further tested after each regression fit using "vif ()" function of car (Fox and Weisberg 2011) package of the R statistical software, which computes the variance inflation factor (VIF). Models that had predictor variables with high correlations and/or VIF values greater than 2 were dropped (Kabacoff 2015). The assumptions of ordinary least square (OLS) regression of each model were tested using the "Global test of linear model assumptions" function of gvlma (Pena and Slate 2014) package of the R statistical software after each regression fit. The relative importance of predictors in each model was determined using "relweights ()"of the R statistical software. The models with significant p-values and adjusted $R^2 \geq 0.90$ were selected for cross-validation. The validation was done using data from the hold-out sample (the remainder set of 25 genotypes used in this experiment). Further validation was done using data from the well-irrigated environment. XY plots of the "terms of each validated model" (traits found in each model) were used to identify the levels of combination of these traits that would give a high PY ($\geq$3.5 tha^{-1}) under intermittent drought conditions.

To identify the suitability of evaluating genotypes for drought tolerance in one environment (test environment) for another environment (target environment), Pearson correlation was used to estimate the coefficient of relationship between the treatments using agricolae package of the R statistical software. Genotype main effect plus genotype-by-environment (GGE) interaction analysis was conducted to identify the most discriminating environment using GGEBiplotGUI (Frutos, Galindo, and Leiva 2014) package of the R statistical software.

3. Results

3.1. *Weather and soil parameters under field conditions*

The mean daily air temperature ranged between 17.40 and 41.40°C (Figure 1). The maximum precipitation, which occurred during the latter part of the experiment, was 16.40 mm (Figure 1). Mean daily relative humidity ranged between 5.63 and 93.00%, whereas mean daily vapor pressure ranged between 87.59 and 100.05 kPa (Figure 2).

There was a progressive decline in soil moisture from 0.28 to 0.12 m^3m^{-3} at 10 cm soil depth and from 0.33 to 0.10 m^3m^{-3} at 20 cm soil depth (Figure 3). Soil water potential also declined progressively from −6.73 to −455.85 kPa during the experimental period (Figure 3).

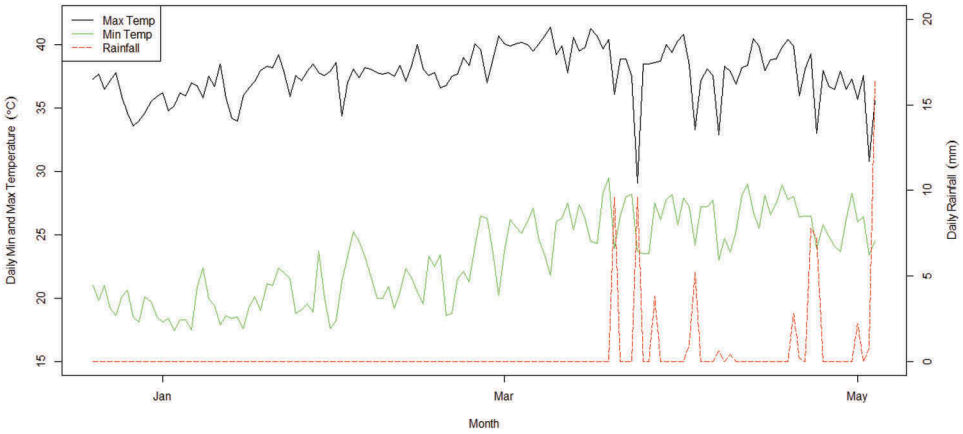

Figure 1. Daily rainfall, minimum and maximum temperatures during the experimental period.

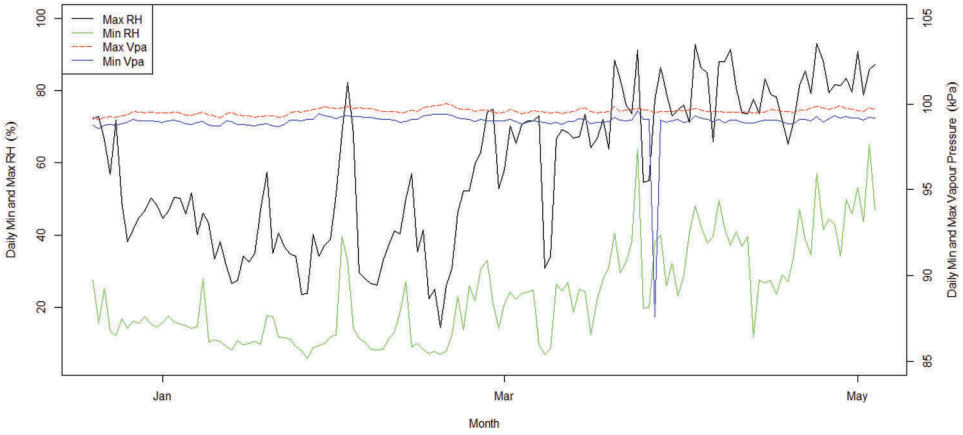

Figure 2. Daily minimum and maximum relative humidity and vapor pressure during the experimental period.

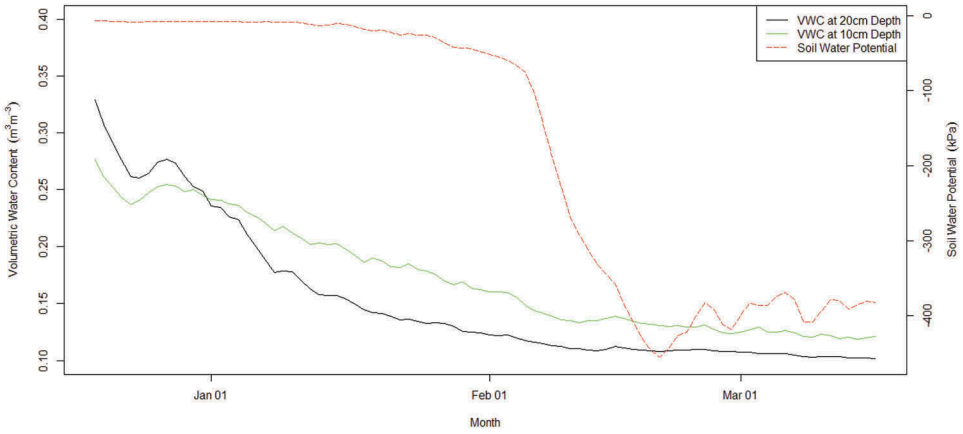

Figure 3. Volumetric moisture content and water potential of the soil during the experimental period.

3.2. *Identification of the most important physiological traits associated with* PY

The first quadrant of the individuals factor map contained genotypes, such as 12 CS-004, 12 CS-041, ICGV 89104, CN-94C, ICGV 03056, and ICGV 91114 (Figure 4). On the other hand, the first quadrant of the variables factor map had PY and other important traits that could be used to select for high PY under intermittent drought conditions. Table 2 contains the traits with significant correlations with dimensions 1 and 2 of Figure 4.

Leaf area at pod initiation had the highest positive correlation, whereas dry matter at maturity had the highest negative correlation with dimension 1 (Table 2). All other traits with significant associations with dimension 1 had coefficients between these two extremes. CGR and PY penalty also had the highest positive and negative correlations, respectively, with dimension 2. All other traits with significant associations with dimension 2 had coefficients between these two extremes.

3.3. *Generation, calibration, and validation of linear models*

All the traits in Table 2 were subjected to all subsets regression, after which a total of 25 models were obtained. However, TDM, CGR, PGR, PCt, and HI gave models with the highest adjusted R^2 values (Figure 5(a)). The model containing all the five traits as predictor variables gave the highest adjusted R^2 (1.00), whereas the model with only TDM as predictor variable gave the lowest adjusted R^2 (0.30) value (Figure 5(a)). All the other models with different combinations of predictor variables were between these two extremes.

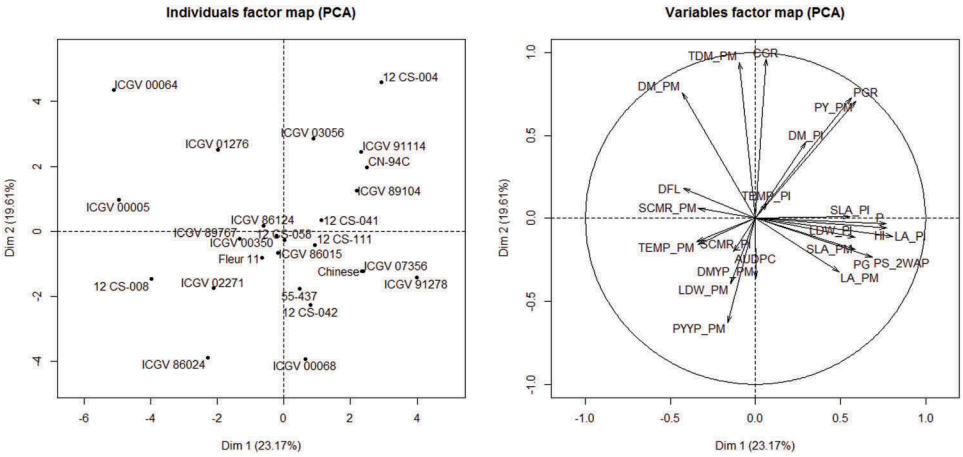

Figure 4. PCA plot showing the most important physiological traits associated with PY under intermittent drought conditions.

Table 2. Traits with high correlations with dimensions 1 and 2 of the PCA plot.

Dim.1			Dim.2		
Trait	Correlation	Significance	Trait	Correlation	Significance
Leaf area at pod initiation	0.81	***	Crop growth rate	0.96	***
Harvest index	0.77	***	Total dry matter at maturity	0.94	***
Partition coefficient	0.77	***	Dry matter at maturity	0.76	***
Plant stand at 14 DAP†	0.69	***	Pod growth rate	0.73	***
Percent germination	0.69	***	Pod yield	0.71	***
Pod yield	0.59	**	Dry matter at pod initiation	0.46	*
SLA‡ at maturity	0.58	**	Pod yield penalty	−0.63	***
Leaf dry weight at pod initiation	0.58	**			
Pod growth rate	0.56	**			
SLA at pod initiation	0.56	**			
Leaf area at maturity	0.49	*			
Degree days to flowering	−0.42	*			
Dry matter at maturity	−0.43	*			

*, **, *** Significant at the 0.05, 0.01, and 0.001 probability level, respectively.
†DAP: days after planting.
‡SLA: specific leaf area.

The genotype-by-trait biplot (Figure 5(b)) showed that PGR had a strong positive correlation with HI, PCt, CGR, and TDM. HI correlated strongly with PCt, and TDM with CGR. Since correlated variables cannot be used as predictor variables in one model because of multicollinearity (Graham 2003; Kabacoff 2015), five models were selected for calibration and validation.

All selected models had significant overall F-test at $p < 0.001$ with adjusted $R^2 > 0.90$ (Table 3). The predictor variables had significant positive regression coefficients ($p \leq 0.001$) in all the models selected (Table 4). None of the indicators of global validation of linear model assumptions was significant in

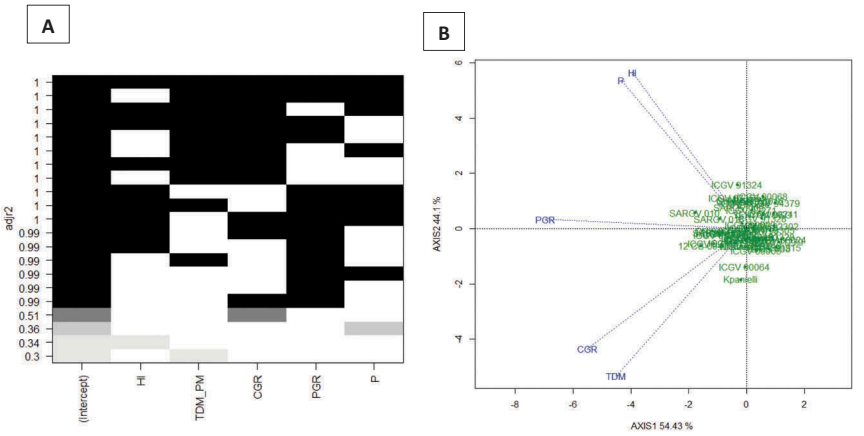

Figure 5. All subsets regression plot (a) and Genotype-by-trait biplot showing associations among selected predictor variables (b).

Table 3. Calibration of models.

No.	Model	*F*-statistic	Pr(> F)	AIC‡‡	Adjusted R^2
1	PY† = TDM‡ + PCt§	510.5	$< 2.2e^{-16}$***	−61.23	0.9798
2	PY = TDM + HI¶	549.8	$< 2.2e^{-16}$***	−62.84	0.9812
3	PY = CGR# + PCt	767.8	$< 2.2e^{-16}$***	−70.08	0.9865
4	PY = CGR + HI	853.2	$< 2.2e^{-16}$***	−72.37	0.9878
5	PY = PGR††	2258	$< 2.2e^{-16}$***	−76.15	0.9895

*, **, *** Significant at the 0.05, 0.01 and 0.001 probability level, respectively.
†PY: pod yield.
‡TDM: total dry matter.
§PCt: partition coefficient.
¶HI: harvest index.
#CGR: crop growth rate.
††PGR: pod growth rate.
‡‡Akaike information criterion.

Table 4. Coefficient estimates and level of significance of the terms of selected models.

Model	Parameter	Estimate	Std. error	*t* Value	Pr(>|t|)
PY† = TDM‡ + PCt§	Intercept	−2.55953	0.14495	−17.66	$3.02e^{-13}$ ***
	TDM	0.34142	0.01387	24.61	$7.10e^{-16}$ ***
	PCt	4.00421	0.14384	27.84	$<2e^{-16}$ ***
PY = TDM + HI¶	Intercept	−2.25906	0.131049	−17.24	$4.66e^{-13}$ ***
	TDM	0.371239	0.013777	26.95	$<2e^{-16}$ ***
	HI	0.060433	0.002091	28.9	$<2e^{-16}$ ***
PY = CGR# + PCt	Intercept	−2.1058	0.1066	−19.76	$3.97e^{-14}$ ***
	CGR	665.2114	21.9927	30.25	$<2e^{-16}$ ***
	PCt	3.2795	0.1113	29.46	$<2e^{-16}$ ***
PY = CGR + HI	Intercept	−1.83056	0.094421	−19.39	$5.6e^{-14}$ ***
	CGR	711.6104	21.1665	33.62	$<2e^{-16}$ ***
	HI	0.048735	0.001568	31.07	$<2e^{-16}$ ***
PY = PGR††	Intercept	−0.02248	0.04486	−0.501	0.621^{ns}
	PGR	1048.853	22.07018	47.524	$< 2e^{-16}$ ***

*, **, *** Significant at the 0.05, 0.01 and 0.001 probability level, respectively.
†PY: pod yield.
‡TDM: total dry matter.
§PCt: partition coefficient.
¶HI: harvest index.
#CGR: crop growth rate.
††PGR: pod growth rate.

any of the models (Table 5), indicating that the data and the selected models met all the statistical assumptions associated with an OLS regression model.

There was no multicollinearity between the predictor variables of any of the selected models, as none of the parameters had a VIF > 2 (Table 6). The TDM contributed 41.72% and 45.05% to the R^2 values of models 1 and 2, respectively, whereas PCt contributed 58.28% and 48.48% to the R^2 values of models 1 and 3, respectively (Table 6). The CGR also contributed 51.52% and 54.83% to the R^2 values of models 3 and 4, respectively, whereas HI contributed 54.95% and 45.17% to the R^2 values of models 2 and 4, respectively.

There were significant correlations ($p \leq 0.001$) between predicted PY and the realized PY (Table 7) under intermittent drought and well-irrigated

Table 5. Global validation of linear model assumptions.

	p-Values				
	Model 1	Model 2	Model 3	Model 4	Model 5
Global stat	0.8146^{ns†}	0.9097 ^{ns}	0.8173 ^{ns}	0.9139 ^{ns}	0.7904 ^{ns}
Skewness	0.4795^{ns}	0.5296^{ns}	0.5137^{ns}	0.5248^{ns}	0.5301^{ns}
Kurtosis	0.3145^{ns}	0.9783^{ns}	0.3102^{ns}	0.9495^{ns}	0.7141^{ns}
Link function	0.8156^{ns}	0.4499^{ns}	0.7618^{ns}	0.4735^{ns}	0.5241^{ns}
Heteroscedasticity	0.9648^{ns}	0.855^{ns}	0.9471^{ns}	0.822^{ns}	0.381^{ns}

†ns: not significant at $p = 0.05$.

Table 6. Multicollinearity and relative importance of predictor variables.

Model	Parameter	Variance inflation factor (VIF)	Contribution of predictor variables (%)
PY† = TDM‡ + PCt§	TDM	1.146997	41.72
	PCt	1.146997	58.28
PY = TDM + HI¶	TDM	1.216868	45.05
	HI	1.216868	54.95
PY = CGR# + PCt	CGR	1.026673	51.52
	PCt	1.026673	48.48
PY = CGR + HI	CGR	1.055516	54.83
	HI	1.055516	45.17
PY = PGR††	PGR	NA‡‡	100

†PY: pod yield.
‡TDM: total dry matter.
§PCt: partition coefficient.
¶HI: harvest index.
#CGR: crop growth rate.
††PGR: pod growth rate.
‡‡NA: not applicable.

Table 7. Validation of the models under drought and non-drought conditions.

		Correlation coefficient	
		Observed pod yield	
Model no.	Simulated pod yield	Intermittent drought	Well-irrigated
1	PY† = TDM‡ + PCt§	0.94***	0.88***
2	PY = TDM + HI¶	0.96***	0.86***
3	PY = CGR# + PCt	0.96***	0.92***
4	PY = CGR + HI	0.97***	0.90***
5	PY = PGR††	0.99***	0.99***

***Significant at the 0.001 probability level.
†PY: pod yield.
‡TDM: total dry matter.
§PCt: partition coefficient.
¶HI: harvest index.
#CGR: crop growth rate.
††PGR: pod growth rate.

conditions. The correlation coefficients ranged from 0.94 to 0.99 under intermittent drought and from 0.86 to 0.99 under well-irrigated conditions, with model 5 showing the highest correlation coefficient under both conditions.

3.4. *Predicting PY using the terms of the models*

For a genotype to attain a PY of 2.36 to 3.56 t ha^{-1} (arced areas in Figure 6) under intermittent drought condition, it must produce a minimum total dry matter of 6.49 t ha^{-1}, with a minimum PCt of 0.65 (Figure 6(a)) or a minimum HI of 35.67% (Figure 6(b)). Similarly, for a genotype to attain a PY of 2.36 to 3.56 t ha^{-1} (arced areas in Figure 6) under intermittent drought condition, it must produce a minimum CGR of 0.0035 g°Cd^{-1} and a minimum PCt of 0.65 (Figure 6(c)) or a minimum HI of 35.67% (Figure 6(d)). A genotype must have a minimum PGR of 0.0023 g°Cd^{-1} to attain a PY of 2.36 to 3.56 tha^{-1} under intermittent drought condition (results not shown).

3.5. *Relationship among environments*

For leaf chlorophyll content, the box environment showed a significant correlation with only the terminal drought environment at 932.80 and 2369.70 °Cd (pod initiation and physiological maturity growth stages, respectively) (Table 8). Intermittent drought environment also had a significant relationship with the well-irrigated environment at 932.80 and 2369.70 °Cd. In addition, for SLA, results from the box environment had significant correlation with only the terminal drought environment at 932.80 and 2369.70 °Cd (Table 8). The SLA at 932.80 °Cd under intermittent drought had a significant correlation with all the environments, except the box environment and well-irrigated environment at 2369.70 °Cd. Leaf chlorophyll content and SLA at 932.80 °Cd were significantly correlated with the respective leaf chlorophyll content and SLA at 2369.70 °Cd under field conditions in all treatments (Table 8). However, the box experiment was

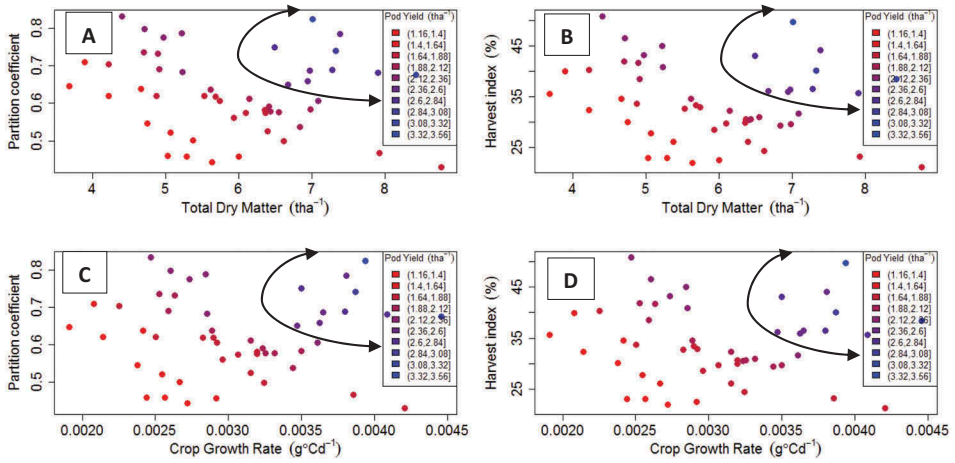

Figure 6. "XY" plots of model terms showing combination of levels leading to an improved PY under intermittent drought conditions.

Table 8. Relationship among the various environments† (BOX, FT1, FT2, and FT3) for drought surrogate traits (leaf chlorophyll content (LCc) and specific leaf area (SLA)) in the upper and lower parts, respectively) [Sample size = 50].

		BOX	FT1		FT2		FT3			
		LCc at PI‡	LCc at PI	LCc at PM§	LCc at PI	LCc at PM	LCc at PI	LCc at PM		
BOX	**SLA at PI**		-0.02^{ns}	0.11^{ns}	0.28*	0.28*	0.14^{ns}	0.18^{ns}	**LCc at PI**	**BOX**
FT1	**SLA at PI**	0.09^{ns}		0.58***	0.12^{ns}	0.11^{ns}	0.37*	0.21^{ns}	**LCc at PI**	**FT1**
	SLA at PM	0.03^{ns}	0.39**		0.11^{ns}	0.39**	0.47**	0.45**	**LCc at PM**	
FT2	**SLA at PI**	0.32*	0.57***	0.13^{ns}		0.46**	0.2^{ns}	0.27^{ns}	**LCc at PI**	**FT2**
	SLA at PM	0.34*	0.57***	0.36*	0.44**		0.26^{ns}	0.39**	**LCc at PM**	
FT3	**SLA at PI**	0.17^{ns}	0.38*	0.12^{ns}	0.54***	0.38*		0.64***	**LCc at PI**	**FT3**
	SLA at PM	0.26^{ns}	0.12^{ns}	0.1^{ns}	0.26^{ns}	0.2^{ns}	0.07^{ns}		**LCc at PM**	
		SLA at PI	**SLA at PI**	**SLA at PM**	**SLA at PI**	**SLA at PM**	**SLA at PI**	**SLA at PM**		
		BOX	**FT1**		**FT2**		**FT3**			

*, **, *** Significant at the 0.05, 0.01 and 0.001 probability level, respectively. ns: not significant.
†Box, FT1, FT2 and FT3 denote box, intermittent drought, terminal drought and well-irrigated treatments, respectively.
‡PI pod initiation.
§PM physiological maturity.

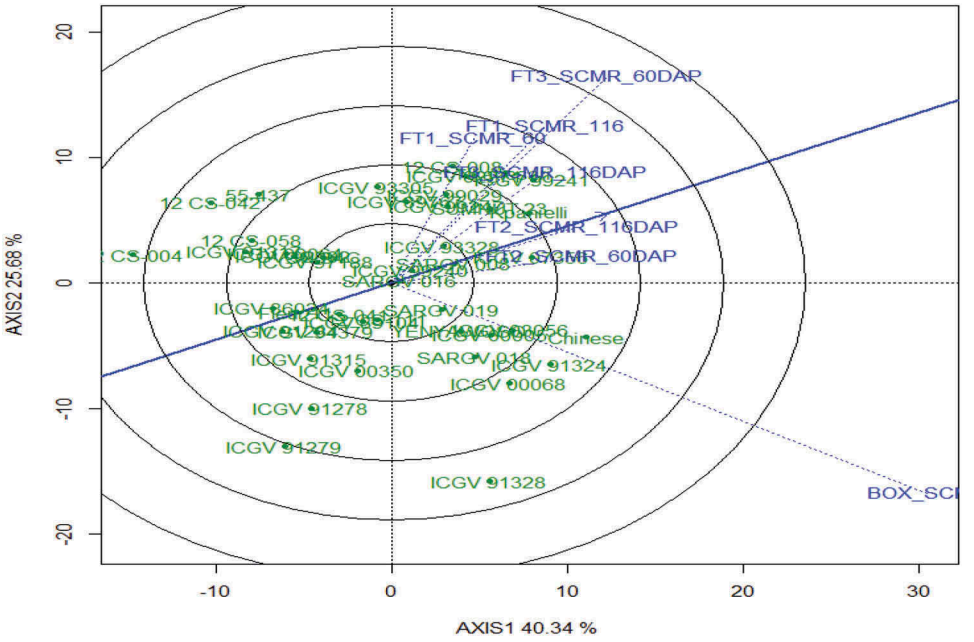

Figure 7. Genotype main effect plus genotype-by-environment interaction (GGE) biplot showing discriminativeness versus representativeness among the environments for leaf chlorophyll content.

the most discriminating among all the treatments studied for the two drought surrogate traits considered (leaf chlorophyll content (Figure 7) and SLA (Figure 8)).

4. Discussion

Figure 4 and Table 2 represent traits with high principal component axis scores under the intermittent drought condition. However, after all subsets regression, TDM, CGR, PGR, PCt, and HI gave models with the highest adjusted R^2 values (Figure 5(a)). This confirms why yield and yield-contributing parameters are the most widely targeted traits in groundnut improvement programs worldwide (Pasupuleti et al. 2013). Our findings support the proposition that PY is a function of CGR, duration of reproductive growth (PGR), and the fraction of CGR partitioned toward PY (PCt) (Pasupuleti et al. 2013).

The ability of a genotype to adapt to a particular water availability level eventually determines the level of "tolerance" of that genotype (Vadez et al. 2013). Therefore, to understand these complex interactions, methods are needed that can be used to analyze as many components as possible such that a comprehensive understanding can be achieved (Vadez et al. 2013). Five linear models were developed using a stepwise regression approach on data from the intermittent drought environment. The intermittent drought data

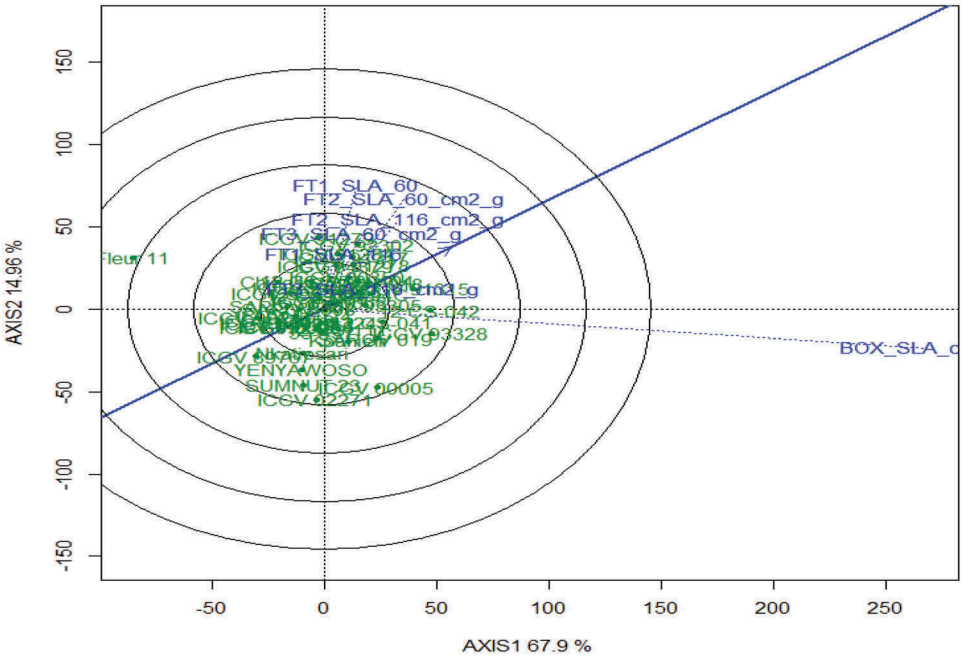

Figure 8. Genotype main effect plus genotype-by-environment interaction (GGE) biplot showing discriminativeness versus representativeness among the environments for SLA.

were used in model development because mid- and end-of-season droughts are critical as they affect PY and quality and predispose genotypes to pre-harvest *Aspergillus* infection (Pasupuleti et al. 2013). All the proposed models had adjusted R-square values >0.90 after calibration (Table 3).

The coefficients of predictor variables in the various models showed that TDM had a positive coefficient in models 1 and 2 (Table 4), confirming the known theory in plant physiology that all plants require a certain minimum amount of above ground biomass for efficient resource capture. It contributed 41.72 and 45.05% to the adjusted R-square values of both models, respectively (Table 6). However, from the XY plots (Figure 6), higher TDM did not necessarily led to higher PY (Figure 6(a,b)). It is therefore an indication that TDM beyond certain minimum or maximum leads to reduced PY under intermittent drought conditions (Figure 6). Similar explanation has been proposed for high TDM under drought conditions; it has been reported that lower rates of vegetative water use leave more water available for grain filling (Vadez et al. 2012, 2013). This is because biomass accumulation is intrinsically linked to transpiration; because of the fact that stomatal aperture and leaf area determine the rates of both processes. There is, therefore, an inherent conflict between biomass accumulation and stress avoidance via reduction of transpiration under drought conditions (Tardieu 2011). The proportion of PY in TDM (HI) was also largely less than 50% (Figure 6(b,d)). Therefore, to attain a unit increase in PY by increasing TDM, other components of the TDM must also be increased. Hence

for a genotype to produce higher PY, the TDM produced must be accompanied with higher PCt (Figure 6(a)) or higher HI (Figure 6(b)).

The CGR, which is the amount of TDM in grams produced per growing degree day, had a positive coefficient in models 3 and 4 (Table 4) and contributed 51.52 and 54.83 % to the adjusted R-square values of both models, respectively (Table 6). However, like TDM, for a genotype to produce higher PY, higher CGR must be accompanied with higher PCt (Figure 6(c)), or higher HI (Figure 6(d)). The CGR in groundnut is largely controlled by temperature (Nigam 2014). Temperature increase within the optimal range variably accelerates CGR and hence the onset of flowering. This is achieved by increasing thermal time accumulation while reducing the length of the growing period by increasing evapotranspiration (Vadez et al. 2012). The increasing evapotranspiration requires that maturity duration should match with the length of growing period available at a given location as conditioned by the soil moisture availability and climatic conditions (mainly temperatures and sunshine hours) (Pasupuleti et al. 2013). Therefore, by choosing environments with the required temperature regimes, CGR can be maximized with a corresponding increase in PY.

The effects of HI, PGR, and PCt were all significant ($p \leq 0.001$) and positive in their respective models. This suggested that separation of yield into its components, including the rate of partitioning and its duration, could provide a better focus on the most relevant traits for yield enhancement (Krishnamurthy et al. 2013). Given that terminal drought tends to curtail the length of the reproductive period, conscious selection for greater partitioning coefficient will confer greater tolerance to abiotic stresses on genotypes (Krishnamurthy et al. 2013). Based on the validation results of the proposed models (Table 7), a selection index based on TDM, CGR, HI, PGR, and PCt might be a useful compromise for selection of superior groundnut genotypes under drought and non-drought conditions.

For the two drought-tolerance surrogate traits used in studying environmental associations, the box environment only had significant correlations with the terminal drought environment in the field (Table 8). This might be attributable to the fact that these environments received the same watering regime, and hence one is representative of the other (Yan 2014). The genotypes' response pattern in these environments was therefore similar (Yan and Rajcan 2002). On the other hand, if groundnut genotypes are to be assessed using leaf chlorophyll content and SLA as surrogates for drought tolerance, the experiment can be terminated at pod initiation. This is because there were significant relationships recorded between the pod initiation and physiological maturity data for leaf chlorophyll content and SLA, respectively, under all environments considered (Table 8). However, significant differences existed between the two growth stages (results not shown).

The box environment was the most discriminating among the environments studied for the two drought-tolerance surrogate traits (Figures 7, 8). This means applying this technique to screen groundnut genotypes using leaf chlorophyll content and SLA should be effective in identifying better performing genotypes. This method also has an additional advantage in that, seedlings from wooden boxes can be saved and transplanted into the field using the bulking procedure for further progeny testing and selection (Singh, Mai-Kodomi, and Terao 1999).

5. Conclusion

The models used conformed to the assumptions of linear models; there was no multicollinearity between predictor variables. Each predictor variable contributed more than 40% to the adjusted R-square values of their respective models. All the proposed models were good predictors of PY under intermittent drought and well-irrigated conditions. When these models are adopted and calibrated in any environment, they should be capable of predicting PY. The models when applied in simulation modeling will give estimates of components (predictor variables) that would produce the required PY. These estimates can assist breeders and policy makers to set breeding objectives and targets.

Screening groundnut genotypes for drought tolerance using wooden boxes was found to be an effective, simple and quick approach. However, because results from the seed-box environment correlated only with the field environment under similar watering regime, the type of drought tolerance being bred for should be simulated in the boxes. Selection can be done at the pod initiation stage because of the strong relationship (r = 0.46–0.64 for leaf chlorophyll, 0.39–0.44 for SLA, $p < 0.05$) between pod initiation and maturity under all environments.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Tropical Legume (TL) III;Alliance for a Green Revolution in Africa (AGRA).

ORCID

Richard Oteng-Frimpong http://orcid.org/0000-0001-6083-1461

References

Akbar, A., S. Singh Manohar, M. Tottekkaad Variath, S. Kurapati, and J. Pasupuleti. 2017. "Efficient Partitioning of Assimilates in Stress-Tolerant Groundnut Genotypes under High-Temperature Stress." *Agronomy* 7: 30. doi:10.3390/agronomy7020030.

Arunyanark, A., S. Jogloy, S. Wongkaew, C. Akkasaeng, N. Vorasoot, G. C. Wright, R. C. Rachaputi, and A. Patanothai. 2009. "Association between Aflatoxin Contamination and Drought Tolerance Traits in Peanut." *Field Crops Researcher* 114: 14–22. doi:10.1016/j.fcr.2009.06.018.

Boote, K. J., J. W. Jones, and P. Singh, 1992. "Modeling Growth and Yield of Groundnut." Presented at the Groundnut – A global Perspective: Proceeding of an International Workshop, 331–43. Patancheru, Andhra Pradesh: ICRISAT Asia Centre.

Cruickshank, A., D. Arnaud, W. Graeme, C. R. NageswaraRao, and N. Shyam. 2004. *Heritability of Drought–Resistance Traits in Peanut*.

Duncan, W. G., D. E. McCloud, R. L. McGraw, and K. J. Boote. 1978. "Physiological Aspects of Peanut Yield Improvement." *Crop Science* 18: 1015–20. doi:10.2135/cropsci1978.0011183X001800060028x.

Felipe, D. M. 2017. *Agricolae: Statistical Procedures for Agricultural Research*. R package version 1.2-8. https://CRAN.R-project.org/package=agricolae

Fox, J., and S. Weisberg. 2011. *An R Companion to Applied Regression*. 2nd ed. Thousand Oaks, CA: Sage.

Frutos, E., M. P. Galindo, and V. Leiva. 2014. "An Interactive Biplot Implementation in R for Modeling Genotype-By-Environment Interaction." *Stoch Environment Researcher Risk Assessment* 28: 1629–41. doi:10.1007/s00477-013-0821-z.

Graham, M. H. 2003. "Confronting Multicollinearity in Ecological Multiple Regression." *Ecology* 84: 2809–15. doi:10.1890/02-3114.

Hoogenboom, G., J. W. Jones, and K. J. Boote. 1992. "Modeling Growth, Development, and Yield of Grain Legumes Using SOYGRO, PNUTGRO, and BEANGRO: A review." *Transactions ASAE* 35: 2043–56. doi:10.13031/2013.28833.

Ibpgr, I. 1992. "Descriptors for Groundnut." Rome Italy Patancheru AP India Int Board Plant Genet. Resour. Int Crops Res Inst Semi-Arid Trop.

Kabacoff, R. I. 2015. *R in Action, 2nd Edition: Data Analysis and Graphics with R. Manning*. Shelter Island, NY: Manning Publications.

Kaur, P., and S. S. Hundal. 1999. "Forecasting Growth and Yield of Groundnut (Arachis Hypogaea) with a Dynamic Simulation Model 'PNUTGRO' under Punjab Conditions." *Journal Agricultural Sciences* 133: 167–73. doi:10.1017/S0021859699006796.

Kiniry, J. R., C. E. Simpson, A. M. Schubert, and J. D. Reed. 2005. "Peanut Leaf Area Index, Light Interception, Radiation Use Efficiency, and Harvest Index at Three Sites in Texas." *Field Crops Researcher* 91: 297–306. doi:10.1016/j.fcr.2004.07.021.

Krishnamurthy, L., J. Kashiwagi, H. D. Upadhyaya, C. L. L. Gowda, P. M. Gaur, S. Singh, R. Purushothaman, and R. K. Varshney. 2013. "Partitioning Coefficient: A Trait that Contributes to Drought Tolerance in Chickpea." *Field Crops Researcher* 149: 354–65. doi:10.1016/j.fcr.2013.05.022.

Ludlow, M. M., and R. C. Muchow. 1990. "A Critical Evaluation of Traits for Improving Crop Yields in Water-Limited Environments." *Advancement Agronomic* 43: 107–53. doi:10.1016/S0065-2113(08)60477-0.

Lumley, T. 2017. *Leaps: Regression Subset Selection*. R package version 3.0. https://CRAN.R-project.org/package=leaps

Měch, R., and P. Prusinkiewicz, 1996. "Visual Models of Plants Interacting with Their Environment." In Proceedings of the 23rd annual conference on computer graphics and

interactive techniques, SIGGRAPH '96, 397–410. New York, NY: ACM. doi:10.1145/237170.237279.

Nigam, S. N. 2014. "Groundnut at a Glance." *ICRISAT*, 121.

Nigam, S. N., S. Chandra, K. R. Sridevi, M. Bhukta, A. G. S. Reddy, N. R. Rachaputi, G. C. Wright, et al. 2005. "Efficiency of Physiological Trait-Based and Empirical Selection Approaches for Drought Tolerance in Groundnut". *The Annals of Applied Biology* 146: 433–39. doi:10.1111/j.1744-7348.2005.040076.x.

Pandey, V., and A. Shukla. 2016. "Improving Crop Yield under Drought Stress through Physiological Breeding." In *Drought Stress Tolerance in Plants*, 1 vol, 331–48. Cham: Springer. doi:10.1007/978-3-319-28899-4_14.

Pasupuleti, J., S. N. Nigam, M. K. Pandey, P. Nagesh, and R. K. Varshney. 2013. "Groundnut Improvement: Use of Genetic and Genomic Tools." *Frontiers Plant Sciences* 4: 23.

Pasupuleti, J., M. T. Variath, M. K. Pandey, H. Desmae, B. N. Motagi, P. Okori, S. S. Manohar, et al. 2016. "Genomic Tools in Groundnut Breeding Program: Status and Perspectives." *Frontiers in Plant Science*, 7: p. 289.

Pena, E. A., and E. H. Slate. 2014. *Gvlma: Global Validation of Linear Models Assumptions*. R package version 1.0.0.2. https://CRAN.R-project.org/package=gvlma

Pittman, R. N., (usda). 1995. *United States Peanut Descriptors*. US: ARS.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rao, M. V., S. N. Nigam, and A. K. S. Huda. 1992. "The Thermal Time Concept as a Selection Criterion for Earliness in Peanut 1." *Peanut Sciences* 19: 7–10. doi:10.3146/i0095-3679-19-1-2.

Sebastien, L., J. Julie, and H. Francois. 2008. "FactoMineR: An R Package for Multivariate Analysis." *Journal of Statistical Software* 25: 1–18. doi:10.18637/jss.v025.i01.

Singh, B. B., Y. Mai-Kodomi, and T. Terao. 1999. "A Simple Screening Method for Drought Tolerance in Cowpea." *Indian Journal Genetics Plant Breeder* 59: 211–20.

Singh, P., K. J. Boote, A. Yogeswara Rao, M. R. Iruthayaraj, A. M. Sheikh, S. S. Hundal, R. S. Narang, and P. Singh. 1994. "Evaluation of the Groundnut Model PNUTGRO for Crop Response to Water Availability, Sowing Dates, and Seasons." *Field Crops Researcher* 39: 147–62. doi:10.1016/0378-4290(94)90017-5.

Tardieu, F. 2011. "Any Trait or Trait-Related Allele Can Confer Drought Tolerance: Just Design the Right Drought Scenario." *Journal of Experimental Botany* 63: 25–31. doi:10.1093/jxb/err269.

Vadez, V., J. D. Berger, T. Warkentin, S. Asseng, P. Ratnakumar, K. P. C. Rao, P. M. Gaur, et al. 2012. "Adaptation of Grain Legumes to Climate Change: A Review". *Agronomic Sustain Developments* 32: 31–44. doi:10.1007/s13593-011-0020-6.

Vadez, V., J. Kholova, M. Zaman-Allah, and N. Belko. 2013. "Water: The Most Important 'Molecular' component of Water Stress Tolerance Research." *Functional Plant Biology: FPB* 40: 1310–22. doi:10.1071/FP13149.

Yan, W. 2014. *Crop Variety Trials: Data Management and Analysis*. Chichester: John Wiley & Sons.

Yan, W., and I. Rajcan. 2002. "Biplot Analysis of Test Sites and Trait Relations of Soybean in Ontario." *Crop Science* 42: 11–20. doi:10.2135/cropsci2002.1100.