

Accelerating Cassava Genetic Improvement through NDVI-Based High-Throughput Phenotyping

Abiodun Fatai Olayinka

`aolayinka@wacci.ug.edu.gh`

University of Ghana

Adesike Oladoyin Olayinka

Ladoke Akintola University of Technology

Nkouaya Gaby Edwige Mbanjo

International Institute of Tropical Agriculture (IITA)

Daniel Kwadjo Dzidzienyo

University of Ghana

Samuel Kwame Offei

University of Ghana

Pangirayi Bernard Tongoona

University of Ghana

Chiedozie Egesi

National Root Crops Research Institute (NRCRI)

Ismail Yusuf Rabbi

International Institute of Tropical Agriculture (IITA)

Research Article

Keywords: NDVI, High-Throughput Phenotyping, Plant Architecture

Posted Date: November 28th, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-8118741/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Accelerating Cassava Genetic Improvement through NDVI-Based High-Throughput Phenotyping

Abiodun Fatai Olayinka ^{1,2*}, Adesike Oladoyin Olayinka ^{1,5}, Nkouaya Gaby Edwige Mbanjo ², Daniel Kwadjo Dzidzienyo ^{1,3}, Samuel Kwame Offei ¹, Pangirayi Bernard Tongoona ¹, Chiedozie Egesi ^{4,5} and Ismail Yusuf Rabbi ².

1 West Africa Centre for Crop Improvement (WACCI), College of Basic and Applied Sciences, University of Ghana, Accra, Ghana,

2 International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria,

3 Biotechnology Centre, College of Basic and Applied Sciences, University of Ghana, Accra, Ghana,

4 Cassava Research Program, National Root Crops Research Institute (NRCRI), Umudike, Nigeria,

5 Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria.

aolayinka@wacci.ug.edu.gh

aokolawole@lautech.edu.ng

ddzidzienyo@wacci.ug.edu.gh

e.mbanjo@cgiar.org

soffe@wacci.ug.edu.gh

ptongoona@wacci.ug.edu.gh

cne22@cornell.edu

i.rabbi@cgiar.org

*Corresponding author: Abiodun Fatai Olayinka aolayinka@wacci.ug.edu.gh

Abstract

Cassava (*Manihot esculenta* Crantz) is an important food security crop in sub-Saharan Africa and other tropical regions, but its genetic improvement is hindered by long breeding cycles and labour-intensive phenotyping procedures. This study aimed to develop a rapid phenotyping protocol and assess its predictive capacity for yield and plant architecture traits in cassava using Normalized Difference Vegetation Index (NDVI) data obtained with affordable handheld sensor (Trimble GreenSeeker). A diverse panel of 453 cassava accessions was evaluated across two contrasting agroecological zones in Nigeria; Mokwa (Southern Guinea Savannah) and Onne (Humid Forest) during the 2021/2022 planting season. NDVI data collected at 3, 6, and 9 months after planting (MAP) were integrated with ground truth phenotypic measurements of 26 agronomic traits.

Genetic parameters including broad-sense heritability and genotype-by-environment interactions were estimated. Results showed moderate to high heritability for important traits such as fresh root yield (FYLD), dry matter content (DM), and harvest index (HI). NDVI data, especially at 6 months after planting, demonstrated strong predictive power (R^2 up to 0.9) for yield components, with prediction accuracy varying across locations. Significant negative correlations between lodging (LODG) and yield traits highlighted the influence of plant architecture on productivity in cassava. These findings affirm the applicability of handheld NDVI sensors as cost-effective tools for enhanced phenotyping and selection in cassava breeding programs for rapid genetic gains and varietal development under diverse field conditions.

Keywords: NDVI, High-Throughput Phenotyping, Plant Architecture

Background

Cassava (*Manihot esculenta* Crantz) is a an important staple crop that is widely cultivated in sub-Saharan Africa, Latin America, and parts of Asia for domestic consumption

and industrial raw materials (Parmar *et al.*, 2017; Tize *et al.*, 2021). Its resilience to drought, poor soils, and erratic rainfall makes it a strategic food security crop in the face of climate change (Jarvis *et al.*, 2012; EPA, 2016; Amelework *et al.*, 2021; Dordley, 2022; IPCC, 2023). However, the turnaround time for genetic improvement in cassava through conventional breeding technique is limited by its heterozygous nature, asynchronous flowering and long breeding cycles (Maxmen, 2019; Ceballos *et al.*, 2020). Conventional cassava breeding relies on phenotypic evaluation of traits such as yield, dry matter content and plant architecture, which is labour-intensive, time-consuming and costly (Ceballos *et al.*, 2020; Conceição *et al.*, 2023; Fei & Mahama, 2023). These limitations hinder the scalability of breeding programs and delay the release of improved varieties. Breeding gains can be accelerated by employing high-throughput phenotyping tools that enable rapid, non-destructive assessments at scale. Remote sensing, spectral vegetation indices have been effectively deployed in monitoring crop growth and predictively relating canopy vigour to yield (Dash, 2021; Machwitz *et al.*, 2021; Marzougui *et al.*, 2023; Tattaris *et al.*, 2016). Normalized difference vegetation index (NDVI) is the most widely used of all the vegetation data for predicting crop yield (Huang *et al.*, 2014; Panek & Gozdowski, 2020; Ji *et al.*, 2021; Roznik *et al.*, 2022). Satellite and Unmanned Aerial Vehicle (UAV) platforms are effective in capturing NDVI data, their cost and complexity limit adoption in resource-constrained research institutes. Handheld sensors such as the Trimble GreenSeeker provide a low-cost, user-friendly solution for NDVI data collection (Trimble, 2023). Prior studies have demonstrated that NDVI correlates with cassava yield and biomass traits, making it suitable for use in breeding trials (Ayu Purnamasari *et al.*, 2019; Rattanasopa *et al.*, 2022). However, variability due to environmental factors and genotype-by-environment interactions (GEI) necessitates multi-location validation and the development of robust predictive models. This study aims to (i) develop a rapid phenotyping protocol by integrating the Trimble GreenSeeker NDVI data with ground truth phenotypic data, (ii) quantify genetic parameters including heritability and G×E interactions for key traits and (iii) evaluate NDVI's predictive capacity for cassava yield and plant architecture in contrasting agroecological zones. The outcomes will inform cost-effective breeding strategies to enhance cassava productivity.

Methods

A total of 453 cassava accessions which include 448 obtained from clonal evaluation trials and five commercial checks were evaluated across two locations in Nigeria, Mokwa (Southern

Guinea Savannah) and Onne (Humid Forest). Trials were conducted during the 2021/2022 planting season using an augmented block design with two replications per location. Each plot consisted of three rows with three plants planted per row and a spacing of 1M between ridges and 0.8M within rows.

Phenotypic Data Collection

Traits measured included shoot weight (SHTWT), root weight (RTWT), dry matter content (DM), starch content (STARCH), fresh root yield (FYLD), dry yield (DYLD), top yield (TYLD), harvest index (HI), NDVI data at 3 months after planting (NDVI3), NDVI data at 6 months after planting (NDVI6), NDVI data at 9 months after planting (NDVI9), number of lodged plants per plot (LODG), plant height at 3 months after planting (PLTHT3), plant height at 6 months after planting (PLTHT6), plant height at 9 months after planting (PLTHT9), angle of branching (ANGBR), stem diameter (STMDI), number of harvested roots (RTNO), number of plants per stand (PPSTD9), height at first branch at 9 months after planting (BRNHT9), height at first branch at 6 months after planting (BRNHT6), number of harvested plants per plot (NOHAV), height at first branch at 3 months after planting (BRNHT3), level of branching at 3 months after planting (BRNLEV3), level of branching at 6 months after planting (BRNLEV6), level of branching at 9 months after planting (BRNLEV9), chlorophyll content at 3 months after planting (CHL3), chlorophyll content at 6 months after planting (CHL6) and chlorophyll content at 9 months after planting (CHL9). Measurements were taken at 3, 6, 9, and 12 months after planting (MAP) using standard protocols described by Fukuda *et al.* (1998) and in CassavaBase (www.cassavabase.org).

Collection of NDVI Data

NDVI Data Collection with Trimble handheld sensor

The Trimble GreenSeeker (Figure 1), a vegetation sensor manufactured by the agricultural division of the Trimble Inc USA, was used to capture the spectral properties of the cassava canopy. The sensor is made up of electroluminescent diodes (LEDs) that emit radiation in the red and near-infrared (NIR) regions. NDVI values were calculated based on the relationship between the red and near-infrared bands: $(NIR - R) / (NIR + R)$, where NIR and R refer to the near-infrared and red reflectance, respectively (Farias *et al.*, 2023). The normalized difference vegetation index (NDVI) values were measured using the Trimble GreenSeeker handheld NDVI meter device. The device emits red light at wavelengths within 660 nm, 25 nm FWHM (full width at half maximum), and near-infrared light at 780 nm and 25 nm FWHM. The device

has two FOV (field of view) options: 25 cm at 60 cm: This means that when the device is positioned 60 cm away from the target, it captures an area with a width of 25 cm.

50 cm at 122 cm: Alternatively, when placed 122 cm away, it covers a wider area with a width of 50 cm. (Sahin, 2022).

The sensor was held at a height of 1 m above the canopy and pointed at a 45° angle towards the plants. The sensor was triggered at three points across the plants in the middle row within each plot, and the average NDVI value was recorded. The NDVI values can vary numerically from -1 to +1, with positive values indicating the vegetative vigour of the cassava plant, while negative values indicate the absence of vegetation or soil exposure. The data were collected at 3 months intervals, viz: 3, 6 and 9 months after planting.

The plant height, height at first branch, and other plant architecture traits data were also collected at the same intervals, using the conventional methods.

Figure 1| Trimble GreenSeeker, a vegetation sensor

175 **Data Analysis**

176 **Data Pre-processing:**

177 The data was split into training (80%) and testing (20%) sets using the 'createDataPartition'
178 function from the 'caret' package (Kuhn, 2019) of the (R Core Team, 2020). The seed was set
179 to 123 for reproducibility. Reproducibility is crucial in data analysis, machine learning,
180 simulations, and computational experiments to ensure reliability and validity.

181 **Model Building and Calibration:**

182 Two regression models (linear regression and polynomial regression model) were used
183 for this analysis. A linear regression model ('lm') was fitted with ground truth data of each of
184 the yield and plant architecture traits used as the dependent variable and NDVI data obtained
185 at 3, 6 and 9 months after planting used as the independent variable using the training data.
186 This was done using the 'lm' function from the 'rms' package (Harrell, 2023) of the (R Core
187 Team, 2020). The mathematical formula of the linear regression model is herein stated:

$$188 \text{ FYLD} = \beta_0 + \beta_1 * \text{NDVI} + \varepsilon$$

189 FYLD – is the response variable, which is the fresh root yield

190 NDVI – is the predictor variable, which is the NDVI data taken at 3, 6 or 9 months after planting

191 β_0 – is the intercept term

192 β_1 – is the coefficient of the predictor variable NDVI

193 ε - is the error term

194 A polynomial regression model was developed to predict the cassava fresh root yield (FYLD)
195 based on several predictor variables. This was done using the 'lm' function from the 'leaps'
196 package, version 3.1 (Miller, 2022) of the (R Core Team, 2020). The model included second-
197 degree polynomial terms for each predictor variable to capture potential non-linear
198 relationships. The predictors used in the model were: Normalized Difference Vegetation Index
199 (NDVI), plant height (PLTHT), chlorophyll content (CHL), stem diameter (STMDI), angle of
200 branching (ANGBR), number of lodged plants per plot (LODG), number of harvested plants
201 per plot (NOHAV), and number of plants per stand (PPSTD). The mathematical formula of the
202 polynomial regression model is herein stated:

$$\begin{aligned} 203 \text{ FYLD} = & \beta_0 + \beta_{11} * \text{NDVI} + \beta_{12} * \text{NDVI}^2 + \beta_{21} * \text{PLTHT} + \beta_{22} * \text{PLTHT}^2 + \beta_{31} * \text{CHL} + \beta_{32} \\ 204 & * \text{CHL}^2 + \beta_{41} * \text{STMDI} + \beta_{42} * \text{STMDI}^2 + \beta_{51} * \text{ANGBR} + \beta_{52} * \text{ANGBR}^2 + \beta_{61} * \text{LODG} + \\ 205 & \beta_{62} * \text{LODG}^2 + \beta_{71} * \text{NOHAV} + \beta_{72} * \text{NOHAV}^2 + \beta_{81} * \text{PPSTD} + \beta_{82} * \text{PPSTD}^2 + \varepsilon \end{aligned}$$

Each predictor variable X_t (for $t = 1, 2, \dots, 12$) is modelled with both its linear and quadratic terms to capture any non-linear relationships between the predictor variables and the response variable

Where:

FYLD – is the dependent variable, which is the fresh root yield

β_0 – is the model intercept term

β_t ($t = 1$ to $n+1$) – these are the model coefficients for each independent variable and its squared term. There are a total number of $n+1$ coefficients, where n is the number of independent variables ($n=8$)

ε - is the error term

NDVI – normalized difference vegetation index

PLTHT – plant height

CHL – chlorophyll content

STMDI – stem diameter

ANGBR – angle of branching

LODG – number of lodged plants per plot

NOHAV – number of harvested plants per plot

PPSTD – number of plants per stand

All these traits represent the independent variables used in developing the second-order polynomial regression model. For these three traits, normalized difference vegetation index, plant height, and chlorophyll content, measurements were taken at 3, 6 and 9 months after planting. Their corresponding values were used for the model development at these growth stages.

Model performance was evaluated using summary statistics and diagnostic plots generated by the 'lm' function.

The model was calibrated using bootstrapping with 100 iterations using the calibrate function from the 'rms' package (Harrell, 2023) of the (R Core Team, 2020). The calibrate function in the 'rms' package enhances resampling model calibration. It produces estimates of predicted versus observed values that have been corrected for bias or overfitting. The calibration results were printed and plotted using the print and plot functions.

Model Validation and Evaluation:

The calibrated model was used to predict the values of plant architecture and yield traits for the testing data set. This was done using the `predict` function from the `rms` package.

Model performance was evaluated using the following statistical metrics:

R-squared (R^2) calculated using the `cor` function from the `base` package. It measures the proportion of variance in the plant architecture and yield traits explained by the NDVI predictor.

Root Mean Squared Error (RMSE) calculated using basic mathematical operations. It quantifies the average deviation between predicted and observed values of plant architecture and yield traits.

Mean Absolute Error (MAE) calculated using basic mathematical operations. It provides an average of the absolute errors between predicted and observed values of plant architecture and yield traits.

Validation results were presented for R^2 , RMSE, and MAE.

Analysis of variance of the plant architecture and yield traits was performed using the `statgenGxE` package of the R software (R Core Team, 2020). The mathematical formula of the adopted model is herein stated:

The mathematical formular for the ANOVA model (linear mixed model)

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk}$$

Y_{ijk} is the response for the i -th genotype in the j -th environment and the k -th replicate.

μ is the overall mean

α_i is the fixed effect of the i -th genotype.

β_j is the random effect of the j -th environment.

$(\alpha\beta)_{ij}$ is the random interaction effect of the i -th genotype and the j -th environment.

e_{ijk} is the residual error.

The random effects β_j , and $(\alpha\beta)_{ij}$ are assumed to be independent and normally distributed with mean 0 and variance components σ^2_a , and σ^2_{ea} , respectively. The residual error e_{ijk} is also assumed to be independent and normally distributed with mean 0 and variance component σ^2_e .

The broad sense heritability (H^2) values were calculated as follows:

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{ge}^2}{e} + \frac{\sigma_e^2}{er}}$$

Where H^2 is the broad sense heritability estimate, σ^2_g is the genetic variance component of the accession effect, σ^2_e is the variance component of the residual error, and r is the number of replicates.

The estimation of genetic advance as a percentage of mean (GAM) was performed using the formula described in (Johnson *et al.*, 1955).

$$GAM = \frac{K \times \sigma_p \times H}{100}$$

Where:

$K = 2.06$ at 5% selection intensity.

H = Heritability.

σ_p = Phenotypic standard deviation.

Genotypic correlation analysis for plant architecture and yield traits were performed using the statistical software Meta-R Version 6.0 (Alvarado *et al.*, 2020).

Using the following model:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_k + e_{ijk}$$

y_{ijk} is the observation for the i -th genotype in the j -th block and the k -th replication

μ is the overall mean

α_i is the fixed effect of the i -th genotype

β_j is the random effect of the j -th block

$(\alpha\beta)_{ij}$ is the random interaction effect of the i -th genotype and the j -th block

γ_k is the random effect of the k -th replication within block

e_{ijk} is the residual error.

The estimates of the genetic and phenotypic correlations were obtained using the variance components from the linear mixed model:

301 $y = X\beta + Zu + e$
 302 y is the vector of observations
 303 β is the vector of fixed effects
 304 u is the vector of random effects
 305 e is the vector of residuals
 306 X and Z are the design matrices.

307
 308 The genetic correlation between two traits is calculated as:

309
 310
 311
$$rg = \frac{\sigma_{g1g2}}{\sqrt{\sigma_{g1}^2 \sigma_{g2}^2}}$$

312
 313 σ_{g1g2} is the covariance between the random effects of the two traits

314 σ_{g1}^2 and σ_{g2}^2 are the variances of the random effects of the two traits

315 Data visualization was performed with the R package ‘corrplot’ version 0.92 (Wei & Simko,
 316 2021).

317 Path Analysis was performed with the R package ‘lavaan’ version 0.6-17 (Rosseel, 2023).

318 Data visualization was performed with the R package ‘semPlot’ version 1.1.6 (Epskamp,
 319 2022).

320

321 **Results**

322 The summary statistics of the NDVI data, plant architecture and yield traits for Mokwa
 323 trial (Table 1) and Onne trial (Table 2) showed that the accessions had differentially expressed
 324 performance for NDVI, plant architecture (angle of branching, stem diameter, shoot weight,
 325 number of lodged plants per plot, plant height at 3 months after planting, plant height at 6
 326 months after planting, and plant height at 9 months after planting) and yield traits (root weight,
 327 dry matter content, starch content, fresh root yield, dry yield, top yield, and harvest index) in
 328 respect to each trial environments. Based on the mean values of these traits, higher values were
 329 reported from the trial conducted in Onne compared to that of Mokwa for all the traits except
 330 fresh root yield (FYLD) which was (24.08 t/ha) in Mokwa compared to (23.45 t/ha) in Onne,

top yield (TYLD) 35.89 t/ha compared to 18.41 t/ha, NDVI at 3 months after planting 0.85 to 0.79, and plant height at 3 months after planting (PLTHT3) 110.58cm to 109.22cm. The error variance was lower than the genotypic variance for all the traits except for NDVI at 6 months after planting and levels of branching at 3 and 9 months after planting (Table 3). Broad sense heritability values ranged from 0.0 for NDVI data at 3 months after planting to 0.87 for height at the first branch at 3 months after planting. Moderately high values of broad sense heritability were recorded for important yield and quality traits like starch and dry matter content (0.61) as well as harvest index (0.66). The phenotypic coefficient of variation ranged from 0% for NDVI data at 3 months after planting to 69.88% for the number of lodged plants per plot, while genotypic coefficient of variation values ranged from 0% to 49.21% for these two traits respectively. The highest value of genetic advance as a percentage of the mean (71.57) was obtained for the number of lodged plants per plot while the least value of (0) was obtained for NDVI data at 3 months after planting. The analysis of variance (Table 4) for the NDVI, plant architecture (angle of branching, stem diameter, shoot weight, number of lodged plants per plot, plant height at 3 months after planting, plant height at 6 months after planting, and plant height at 9 months after planting) and yield traits (root weight, dry matter content, starch content, fresh root yield, dry yield, top yield, and harvest index) traits across the two environments showed that the accessions performed differently from one another for all the traits except for NDVI value at 6 months after planting.

Table 1| Summary statistics of the NDVI data, plant architecture and yield traits for the Mokwa trial

Trait	Min	Max	Mean	Variance	Coefficient of Variation
-------	-----	-----	------	----------	--------------------------------

SHTWT	1.74	68.55	26.92	113.88	39.64
RTWT	0.55	46.25	18.06	70.1	46.36
DM	21.21	38.7	30.21	7.86	9.28
STARCH	3.93	27.23	15.92	13.93	23.44
FYLD	0.73	61.67	24.08	124.63	46.36
DYLD	0.97	18.61	7.63	11.22	43.9
TYLD	2.32	91.4	35.89	202.47	39.65
HI	0.05	0.81	0.4	0.01	25
NDVI3	0.37	0.92	0.85	0	5.26
NDVI6	0.24	0.8	0.6	0.01	14.91
NDVI9	0	0.79	0.61	0.01	13.72
LODG	0	9	1.35	2.75	122.84
PLTHT3	21	202.4	110.58	779.71	25.25
PLTHT6	52	231	130.44	802.12	21.71
PLTHT9	65	258	144.18	848.11	20.2
ANGBR	0	117.05	69.3	594.4	35.18
STMDI	9.82	33.14	21.47	10.67	15.21

SHTWT – shoot weight, RTWT – root weight, DM – dry matter content, STARCH – starch content, FYLD – fresh root yield, DYLD – dry yield, TYLD – top yield, HI – harvest index, NDVI3 – NDVI data at 3 months after planting, NDVI6 - NDVI data at 6 months after planting, NDVI9 - NDVI data at 9 months after planting, LODG – number of lodged plants per plot, PLTHT3 – plant height at 3 months after planting, PLTHT6 - plant height at 6 months after planting, PLTHT9 - plant height at 9 months after planting, ANGBR – angle of branching, and STMDI – stem diameter.

Table 2 | Summary statistics of the NDVI data, plant architecture and yield traits for the Onne trial

Trait	Min	Max	Mean	Variance	Coefficient of Variation
SHTWT	1.02	66.2	18.41	93.76	52.6

RTWT	1.15	63.25	23.45	115.35	45.8
DM	25.1	43.75	33.62	8.16	8.5
STARCH	9.12	33.96	20.47	14.47	18.58
FYLD	1.15	63.25	23.45	115.35	45.79
DYLD	1.03	22.36	8.27	13.22	43.97
TYLD	1.02	66.2	18.41	93.73	52.59
HI	0.09	0.94	0.56	0.01	17.86
NDVI3	0.52	0.88	0.79	0.002	5.66
NDVI6	0.5	0.91	0.86	0.06	28.48
NDVI9	0.5	0.93	0.84	0.002	5.32
LODG	1	12	2.24	3.2	79.86
PLTHT3	48	186	109.22	385.7	40.92
PLTHT6	75	284	158.29	766.7	17.49
PLTHT9	106	366	203.71	1183.16	16.89
ANGBR	0	125.17	72.68	1295.6	49.52
STMDI	12.43	41.14	24.62	22.42	19.23

SHTWT – shoot weight, RTWT – root weight, DM – dry matter content, STARCH – starch content, FYLD – fresh root yield, DYLD – dry yield, TYLD – top yield, HI – harvest index, NDVI3 – NDVI data at 3 months after planting, NDVI6 - NDVI data at 6 months after planting, NDVI9 - NDVI data at 9 months after planting, LODG – number of lodged plants per plot, PLTHT3 – plant height at 3 months after planting, PLTHT6 - plant height at 6 months after planting, PLTHT9 - plant height at 9 months after planting, ANGBR – angle of branching, and STMDI – stem diameter.

Table 3 | Estimates of variance components, broad sense heritability, PCV, GCV, and genetic advance for NDVI data, plant architecture and yield traits for all the cassava accessions evaluated across two environments in Nigeria.

Trait	Mean	σ^2_g	$\sigma^2_{G \times E}$	σ^2_e	σ^2_p	H ²	PCV(%)	GCV(%)	GA	GAM
DM	31.90	3.02	1.37	0.59	4.98	0.61	6.99	5.45	2.80	8.76
NOHAV	6.50	1.07	0.48	0.81	2.35	0.46	23.58	15.91	1.44	22.17
RTNO	48.00	130.37	46.82	73.05	250.24	0.52	32.96	23.79	17.02	35.46
SHTWT	22.65	26.28	9.50	15.06	50.84	0.52	31.48	22.63	7.61	33.60

RTWT	20.77	26.11	10.53	11.28	47.92	0.54	33.33	24.60	7.79	37.50
STARCH	18.20	5.36	2.42	1.05	8.83	0.61	16.32	12.72	3.73	20.47
FYLD	23.76	34.97	11.83	15.23	62.03	0.56	33.15	24.89	9.17	38.59
DYLD	7.95	2.62	1.46	1.66	5.74	0.46	30.12	20.36	2.26	28.42
TYLD	27.12	34.90	12.59	22.45	69.94	0.50	30.84	21.78	8.62	31.78
HI	0.48	0.0048	0.24	0.0012	0.2412	0.66	17.7	14.434	0.12	24.29
NDVI3	0.82	0.0002900	0.00	0.0018200	0.00	0.00	0.00	0.00	0.00	0.00
NDVI6	0.73	0.0000000	0.00	0.0100000	0.01	0.00	13.70	0.00	0.00	0.00
NDVI9	0.73	0.0000000	0.00	0.0000000	0.00	0.00	0.00	0.00	0.00	0.00
LODG	1.60	0.62	0.12	0.52	1.25	0.50	69.88	49.21	1.15	71.57
PLTH3	109.87	202.38	14.70	88.40	305.48	0.66	15.91	12.95	23.91	21.76
BRNHT3	55.85	304.95	3.71	41.48	350.14	0.87	33.50	31.27	33.65	60.26
BRNLEV3	1.60	0.01	0.04	0.05	0.09	0.11	18.75	6.25	0.07	4.30
PLTHT6	144.60	340.40	36.59	96.98	473.97	0.72	15.06	12.76	32.29	22.33
BRNHT6	68.10	95.33	206.51	41.31	343.15	0.28	27.20	14.34	10.63	15.60
BRNLEV6	2.00	0.13	0.09	0.05	0.27	0.49	25.74	18.03	0.52	26.07
PLTHT9	174.60	402.33	62.29	127.01	591.62	0.68	13.93	11.49	34.16	19.56
BRNHT9	85.10	473.15	102.22	60.97	636.34	0.74	29.64	25.56	38.73	45.51
BRNLEV9	2.50	0.02	0.12	0.04	0.18	0.11	16.73	5.66	0.10	3.95
CHL3	49.00	4.77	0.07	1.30	6.13	0.78	5.05	4.46	3.98	8.12
CHL6	47.89	3.02	0.51	1.28	4.81	0.63	4.58	3.63	2.84	5.94
CHL9	45.60	5.16	2.97	1.85	9.98	0.52	6.93	4.98	3.37	7.40
ANGBR	171.00	492.45	99.22	50.86	642.52	0.77	14.82	12.98	40.12	23.46
STMDI	23.10	6.34	1.42	2.02	9.78	0.65	13.54	10.90	4.19	18.13
PPSTD	2.00	0.10	0.02	0.08	0.20	0.51	22.22	15.81	0.46	23.23

STMDI - stem diameter, FYLD - fresh root yield, DYLD - dry yield, STARCH – starch content, DM – dry matter content, RTWT – fresh root weight, SHTWT – shoot weight, PLTHT9 – plant height at 9 months after planting, LODG – number of lodged plants per plot, HI – harvest index, ANGBR9 – angle of branching, RTNO – number of harvested roots, PPSTD9 – number of plants per stand, BRNHT9 – height at the first branch at 9 months after planting, PLTHT6 – plant height at 6 months after planting, TYLD – top yield, BRNHT6 – height at the first branch at 6 months after planting, NOHAV – number of harvested plants per plot, NDVI3 – NDVI at 3 months after planting, NDVI6 - NDVI at 6 months after planting, NDVI9 - NDVI at 9 months after planting, PLTHT3 – plant height at 3 months after planting, BRNHT3 – height at the first branch at 3 months after planting, BRNLEV3 – level of branching at 3 months after planting, BRNLEV6 – level of branching at 6 months after planting, BRNLEV9 – level of branching at 9 months after planting, CHL3 – chlorophyll content at 3 months after planting, CHL6 – chlorophyll content at 6 months after planting, CHL9 – chlorophyll content at 9 months after planting, σ^2_g – genotypic variance, σ^2_e – residual variance, σ^2_p – phenotypic variance, H^2 - broad sense heritability, PCV, phenotypic coefficient of variation, GCV- genotypic coefficient of variation, GA - Genetic advance, GAM - genetic advance as a percentage of mean.

405 **Table 4 | Analysis of variance (ANOVA) of the NDVI data, plant architecture and yield traits across two environments**

The exploratory analysis revealed the interrelationships among the plant architecture and yield traits. Based on the Pearson correlation coefficient values (r), correlation coefficient plots were used to visualize and quantify these relationships.

Based on the output of the genetic correlation analysis of the data obtained from the trial conducted in Mokwa (Figure 2), there exists a strong and significant negative relationship between HI and LODG ($r = -0.72$, $p < 0.001$). In addition to this, moderate, significant negative relationships were observed between LODG and FYLD ($r = -0.56$, $p < 0.001$), RTWT ($r = -0.56$, $p < 0.001$), and RTNO ($r = -0.50$, $p < 0.001$) respectively. This shows that the higher the number of lodged plants per plot, the lesser the values of the key yield traits such as FYLD, RTNO, RTWT and HI. Also, there exists high positive and significant relationships between LODG and PLTHT3 ($r = 0.75$, $p < 0.001$), PLTHT6 ($r = 0.87$, $p < 0.001$), and PLTHT9 ($r = 0.79$, $p < 0.001$). This shows that the taller the plant grows, the more it is prone to lodging.

The output of the genotypic correlation among yield and plant architecture traits in the Onne trial (Figure 4) showed weaker, negative, and significant relationships between LODG and HI ($r = -0.4$, $p < 0.001$) compared to the value ($r = -0.72$, $p < 0.001$) obtained in Mokwa trial. Contrary to the positive relationships recorded between LODG and plant height in Mokwa, low to moderate, PLTHT3 ($r = -0.1$, $p < 0.001$), and PLTHT9 ($r = -0.49$, $p < 0.001$) respectively were observed. Unlike in Mokwa, this means that the taller plants in Onne grew, the lesser their propensity for lodging.

The path coefficient analysis revealed the direct and indirect contributions of the plant architecture traits to fresh root yield and root weight in cassava.

For the trial at Mokwa (Figure 3), RTWT had an absolute (1.0) positive contribution to FYLD. However, HI had a strong positive contribution to RTWT (0.67), while the contributions of PLTHT9 (0.41) and PLTHT3 (0.17) were also positive. Contrary to these, negative relationships were observed for LODG (-0.12) and PLTHT6 (-0.18) in respect of their contributions to RTWT.

Likewise, RTWT had absolute contributions to FYLD in the path coefficient analysis conducted on the data from the Onne trial (Figure 5). PLTHT6 had strong positive contributions (1.06) to RTWT while the reverse is the case for ANGBR with negative contribution (-0.30) and PLTHT6 (-0.21).

440

441 **Figure 2| Genotypic correlation coefficient plot of plant architecture and yield traits in**
442 **Mokwa trial**

443 NOHAV - number of harvested plants per plot, RTNO – number of harvested roots per plot, SHTWT – shoot
444 weight, RTWT – root weight, DM – dry matter content, STARCH – starch content, FYLD – fresh root yield,
445 DYLD – dry yield, TYLD – top yield, HI – harvest index, LODG – number of lodged plants per plot, PLTHT6 –
446 plant height at 6 months after planting, BRNHT6 – height at first branch at 6 months after planting, PLTHT9 –
447 plant height at 9 months after planting, BRNHT9 – height at first branch at 9 months after planting, BRNLEV9 -
448 level of branching at 9 months after planting, BRNHB9 – branching habit at 9 months after planting, ANGBR9
449 – angle of branching, and STMDI9 – stem diameter.

450

451

452 **Figure 3| Path coefficient analysis plot of plant architecture and yield traits in Mokwa trial**

453 NOH - number of harvested plants per plot, RTN – number of harvested roots per plot, RTW – root weight, STA
454 – starch content, FYL – fresh root yield, HI – harvest index, LOD – number of lodged plants per plot, PLTHT6 –
455 plant height at 6 months after planting, PLTHT9 – plant height at 9 months after planting, ANG – angle of
456 branching, and STM – stem diameter.

457

458

459 **Figure 4| Genotypic correlation coefficient plot of plant architecture and yield traits in**
460 **Onne trial**

461 NOHAV - number of harvested plants per plot, RTNO – number of harvested roots per plot, SHTWT – shoot
462 weight, RTWT – root weight, DM – dry matter content, STARCH – starch content, FYLD – fresh root yield,
463 DYLD – dry yield, TYLD – top yield, HI – harvest index, LODG – number of lodged plants per plot, PLTHT6 –
464 plant height at 6 months after planting, BRNHT6 – height at first branch at 6 months after planting, PLTHT9 –
465 plant height at 9 months after planting, BRNHT9 – height at first branch at 9 months after planting, BRNLEV9 -
466 level of branching at 9 months after planting, BRNHB9 – branching habit at 9 months after planting, ANGBR9
467 – angle of branching, and STMDI9 – stem diameter.

468

469

470 **Figure 5| Path coefficient analysis plot of plant architecture and yield traits in Onne trial**

471 NOH - number of harvested plants per plot, RTN – number of harvested roots per plot, RTW – root weight, STA
472 – starch content, FYL – fresh root yield, HI – harvest index, LOD – number of lodged plants per plot, PLTHT6 –
473 plant height at 6 months after planting, PLTHT9 – plant height at 9 months after planting, ANG – angle of
474 branching, and STM – stem diameter.

475

476

477 The linear regression model prediction statistics (Table 5) for the prediction of plant
 478 architecture and yield traits using the NDVI data showed that the value of NDVI3 predicted
 479 the PLTHT3 with high R^2 values (0.94) and NDVI6 value for PLTHT6 (0.7) respectively but
 480 with high root mean squared error (RMSE) values of 29.3 and 13.2 respectively, and mean
 481 absolute error (MAE) values of 27.3 and 11.7 respectively. Meanwhile, high R^2 value of (0.9)
 482 was recorded for the prediction of FYLD, DYLD and RTWT using the NDVI6 data, with
 483 relatively low values of RMSE (8.3, 2.2 and 6.2) respectively and MAE values of (7.4, 1.9 and
 484 5.5) respectively for these traits. In the same vein, high R^2 value of (0.7) was recorded for HI
 485 for the prediction made using the NDVI6 and NDVI9 data with very low RMSE values of (0.09
 486 and 0.04) respectively and MAE values of (0.09 and 0.04) respectively. Similar result was
 487 reported from the prediction of STMDI from the NDVI3 data with R^2 value of (0.7) and
 488 relatively higher values of RMSE (1.9) and MAE (1.6).
 489 Moderate values of R^2 (0.6, 0.4, 0.4 and 0.4) were reported for DYLD and RTWT both
 490 predicted from NDVI9 data, STARCH and DM predicted from NDVI3 data, respectively.
 491 These traits also had relatively low values of RMSE (6.2, 4.7, 2.2 and 1.6) as well as MAE
 492 values of (4.5, 3.4, 1.8 and 1.3) respectively.
 493 The prediction metrics obtained from the trial at Onne (Table 6) revealed lower R^2 values for
 494 the predicted traits compared to the values obtained from Mokwa trial.
 495 Moderate R^2 values of (0.5) were obtained on the prediction of FYLD, DYLD, and RTWT,
 496 using the NDVI9 data. However, high values of RMSE (15.3, 5.3 and 15.3) were reported for
 497 these traits respectively as well as high values of MAE (13, 4.7 and 13).
 498 However, moderate R^2 value of (0.5) was reported for the prediction of HI, using the NDVI3
 499 data with low RMSE value (0.1) and MAE value (0.07).
 500 Low R^2 values of (0.3, 0.3 and 0.4) were reported for STARCH, DM and STMDI respectively
 501 from the NDVI3 data. Their RMSE values (3.8, 2.8 and 3.1) were relatively low as well as
 502 MAE values (3.1, 2.3 and 2.9).
 503 The polynomial prediction statistics (Table 7) for the prediction of FYLD from the NDVI data,
 504 PLTHT, CHL, STMDI, ANGBR, LODG, NOHAV, and PPSTD showed the values of the
 505 model prediction statistics were quite similar across environment and the growth stage at which
 506 the data were collected. The R^2 value 0.39 was obtained for the data collected from Mokwa
 507 trial at 3 and 6 months after planting and Onne at 6 months after planting while R^2 value of 0.4
 508 was recorded on data collected from Onne at 3 and 9 months after planting and Mokwa at 9
 509 months after planting. Similarly, MAE values of 6.3 was obtained for the data collected from
 510 Mokwa and Onne at 3 and 6 months after planting, with values of 6.2 and 6.4 for data obtained

from Mokwa and Onne respectively at 9 months after planting. The RMSE values ranged from 7.8 to 8.2 for data obtained from Mokwa and Onne respectively at 9 months after planting.

542 **Table** Error! No text of specified style in document. | **Linear regression model prediction**
543 **statistics for plant architecture and yield traits from NDVI data at MOKWA / 2021-2022**
544 **Trial**

Trait	NDVI MOKWA 3MAP			NDVI MOKWA 6MAP			NDVI MOKWA 9MAP		
	R ²	RMSE	MAE	R ²	RMSE	MAE	R ²	RMSE	MAE
Plant Height	0.94	29.3	27.3	0.7	13.2	11.7	0.1	20.3	15
Fresh Root Yield	0.08	13.4	11.9	0.9	8.3	7.4	0.4	6.2	4.5
Dry Yield	0.1	4.1	3.5	0.9	2.2	1.9	0.6	1.4	1.1
Root Weight	0.07	10.1	8.9	0.9	6.2	5.5	0.4	4.7	3.4
Starch Content	0.4	2.2	1.8	0.2	1.5	1.2	0.3	2.2	1.4
Dry Matter Content	0.4	1.6	1.3	0.2	1.2	0.9	0.3	1.6	1
Shoot Weight	0.08	9.5	8.8	0.2	8.2	7.4	0	10.7	8.4
Harvest Index	0.01	0.1	0.1	0.7	0.09	0.09	0.7	0.04	0.04
Number of lodged plants per plot	0	0.7	0.6	0.1	0.9	0.9	0.2	0.9	0.7
Stem Diameter	0.7	1.9	1.6	0.4	3.1	3	0.2	3	2.3
Angle of Branching	0.3	6.5	5.6	0.1	6	5.2	0.2	7.3	6.1

545
546 R-squared (R²), root mean squared error (RMSE) and mean absolute error (MAE)Months after planting (MAP).
547
548
549
550

551 **Table 6 | Linear regression model prediction statistics for plant architecture and yield**
552 **traits from NDVI data at ONNE / 2021-2022 Trial**

Trait	NDVI ONNE 3MAP			NDVI ONNE 6MAP			NDVI ONNE 9MAP		
	R ²	RMSE	MAE	R ²	RMSE	MAE	R ²	RMSE	MAE
Plant Height	0.02	15.4	10.8	0.01	25	24.8	0.2	27.6	24
Fresh Root Yield	0.2	10.4	8.2	0.4	10.9	9.2	0.5	15.3	13
Dry Yield	0.2	3.8	2.9	0.3	4.2	3.4	0.5	5.3	4.7
Root Weight	0.2	10.4	8.2	0.4	11	9.2	0.5	15.3	13
Starch Content	0.3	3.8	3.1	0	3	2.7	0.2	3.9	2.5
Dry Matter Content	0.3	2.8	2.3	0	2.2	2	0.2	2.9	1.9
Shoot Weight	0.01	6.4	5.3	0.2	18.5	11	0.4	7.6	6.6
Harvest Index	0.5	0.1	0.07	0.01	0.07	0.04	0.3	0.09	0.1
Number of lodged plants per plot	0.1	1.5	1	0.04	1	0.8	0.2	1.4	1.1
Stem Diameter	0.4	3.1	2.9	0	2.1	1.9	0	2.8	2.8
Angle of Branching	0.04	35.7	31.5	0.1	50.5	48	0.2	42.2	40

553
554 R-squared (R²), root mean squared error (RMSE) and mean absolute error (MAE)Months after planting (MAP).
555

556
557
558
559

Table 7 | Polynomial regression model prediction statistics for fresh root yield at MOKWA & ONNE / 2021-2022 Trials

MOKWA									
Trait	3MAP			R ²	6MAP			9MAP	
	R ²	RMSE	MAE		RMSE	MAE	R ²	RMSE	MAE
FYLD	0.39	7.9	6.3	0.39	7.9	6.3	0.4	7.8	6.2

ONNE									
Trait	3MAP			R ²	6MAP			9MAP	
	R ²	RMSE	MAE		RMSE	MAE	R ²	RMSE	MAE
FYLD	0.4	8	6.3	0.39	8.1	6.3	0.4	8.2	6.4

R-squared (R²), root mean squared error (RMSE) and mean absolute error (MAE)Months after planting (MAP).

Discussion

The cassava accessions were planted out in two distinct environments: the humid forest zone and the southern Guinea Savannah zone. There was a strong genotype-by-environment interaction (GEI) for all the plant architecture and yield traits at ($p < 0.001$) except for the number of lodged plants per plot and plant height at 3 months after planting at ($p < 0.01$). This agreed with the findings of Bakare *et al.* (2022) who reported the presence of a strong interaction between cassava genotypes and test environments with respect to fresh root yield, dry yield, top yield, harvest index and dry matter content. Bakare *et al.* (2022) planted in five agroecological zones in Nigeria which included both the humid forest and the southern Guinea Savannah zone. Similar results were reported by Amelework *et al.* (2022) on trials conducted in three distinct locations that encompass both tropical and subtropical agro-ecological zones in South Africa. Amelework *et al.* (2022) observed significant variation among the test environments, genotype, and genotype-by-environment among the cassava accessions in respect to number of roots, root weight, fresh root yield, dry matter content, plant height, shoot weight and harvest index. In another trial, Ogwuche *et al.* (2023) reported similar results for the following traits; root number, fresh root yield, dry yield, dry matter content at ($p < 0.01$). Ogwuche *et al.* (2023) conducted their trials in eight locations in Nigeria that cut across four agroecological zones in which both the humid forest and the southern Guinea Savannah zones were represented.

The presence of genotype-by-environment interaction (GEI) for plant architecture and yield traits makes it difficult for cassava breeders to adequately select superior cassava genotypes for advancement in breeding programs. With the presence of GEI, effective selection of genotypes is rarely achieved due to disparity in the rankings of genotypes across different environments and introduction of bias that diminishes selection gains (Begna, 2022; Sampaio Filho *et al.*, 2023).

Rattanasopa *et al.* (2022) used normalized difference vegetation index (NDVI), simple ratio vegetation index (RVI), and chlorophyll vegetation index (CIRedEdge) to predict the fresh root yield and height of cassava plants. The authors found out NDVI with R^2 value of 0.58 for fresh root yield and R^2 value of 0.87 for plant height produced the best prediction accuracy for fresh root yield. This 0.58 R^2 value was quite low, compared to the 0.9 R^2 value for the prediction of fresh root yield recorded at the trial in Mokwa using NDVI data collected at 6 months after planting. However, it was close to 0.5 R^2 value from the data collected at Onne trial at 9 months after planting. Unlike the R^2 value obtained for yield prediction, the

plant height values were closer to both values obtained from the data collected at Mokwa trial which were 0.94 and 0.7 for plant height at 3 and 6 months after planting respectively. Although, the models have high RMSE and MAE values which require revalidation on a new set of accessions and trials. In a similar study conducted by Ayu Purnamasari *et al.* (2019), on the use of vegetation indices data such as Normalized Difference Vegetation Index (NDVI), Soil Adjusted Vegetation Index (SAVI) and Infrared Emission Chlorophyll Index (IRECI) in predicting cassava yield, the authors reported that NDVI had the highest accuracy for cassava yield prediction with R^2 value of 0.62.

According to Grados & Schrevens (2020), the NDVI values taken from the cassava fields between 2 – 10 months after planting ranged between 0.3 – 0.4. These values were quite low compared to the lowest mean value of 0.60 recorded at 6 months after planting on the trial in Mokwa, and 0.79 at 3 months after planting recorded on the trial in Onne.

In the same vein the prediction accuracy of starch content and dry matter content using the NDVI data collected from Mokwa has the R^2 value of 0.4, though a little below average (0.5). The model is characterized by very low values of RMSE which ranged from 1.6 – 2.2 and RAE values which ranged from 1.3 – 1.8. These give much credence to the reliability of this model. This makes it possible to have a fair estimate of these important traits right from the field and allows for on-the-spot assessment and initial selection.

According to the research conducted by León *et al.* (2021), a mean value of dry matter content as low as 27.9 was reported with the highest mean value of 41. However, in the trial conducted at Mokwa and Onne, mean values of 30.21 and 33.62 were recorded respectively. This is an indication that the accessions have the inherent potential for exploration of genetic gains for dry matter content. Dry matter content is not only an important determinant of yield in cassava, but also highly correlated with starch content (Peprah *et al.*, 2020; Silva *et al.*, 2023; Malai *et al.*, 2024), which is an important indicator for cassava varieties that are suitable for industrial use. This was confirmed with the genetic correlation coefficient value of 0.99. By implication, effective selection for starch content could be carried out using the dry matter content. Given this moderately high estimate of broad sense heritability (0.61), for starch and dry matter content, there is significant potential for genetic improvement of these traits in cassava through selection. Johnson *et al.* (1955) grouped GAM into 3 categories as follows: GAM values that were less than 10% were low, values between 10-20% were classified as moderate while values that were more than 20% were high. The moderately high value of broad sense heritability (0.61) recorded for these traits and moderate genetic advance as a percentage

of mean (GAM) value of 20% for starch content indicate that the trait would respond favourably to selection, thereby resulting into enhanced improvement.

For example, selection for starch content, which has a heritability of 0.61 and genetic advance as a percentage of mean (GAM) value of 20% implies that imposing selection pressure on the top 5% of the base population could result in significant gains of 20% in starch content over time. This will enhance the improvement of the quality of cassava roots for industrial processing and consumption.

Plant height and stem diameter are also important plant architecture traits. Impressive prediction accuracy was obtained for these traits with plant height at 3 months after planting using the NDVI data collected in Mokwa at 3 months after planting with an R^2 value of (0.94) but relatively high values of RMSE (29.3) and MAE (27.3) and R^2 (0.7), RMSE (13.2) and MAE (11.7) for NDVI data collected at 6 months after planting in Mokwa trial. For stem diameter, using the NDVI data collected at 3 months after planting, a high R^2 value of (0.7) with very low RMSE (1.9) and low MAE (1.6) were realized; this suggests that this model is highly reliable for the estimation of this trait and it will facilitate rapid on-field assessment of this trait for real-time decision making.

Amarullah (2021), reported high values of 277.78 cm for plant height and 30 cm for stem diameter. However, in this trial, mean values of plant height ranged between 110.58cm to 203.71cm while the stem diameter ranged from 21.47 cm to 24.62 cm. There exist a positive and significant genotypic relationship between plant height and stem diameter. These traits were also positively correlated to the number of lodged plants per plot. However, all the three traits were negatively and significantly correlated with an important yield indicator, which is harvest index. This necessitates the need to strike a balance between these traits for enhanced productivity. With the high values of broad sense heritability reported for plant height, 0.68 – 0.72, and stem diameter, 0.65 it is possible to manipulate these traits to develop plants with ideal architecture and improved yield.

The genotypic correlation coefficients showed that traits like root dry yield, root weight, harvest index and number of roots were positively and significantly correlated with the principal yield determinant, fresh root yield. This is in agreement with the findings of Edet *et al.* (2015). This was also confirmed in the output of the path coefficient analysis which showed that root weight had an absolute contribution to the fresh root yield. Meanwhile, traits like harvest index, plant height at 3 and 9 months after planting had positive contributions to fresh root yield through the root weight while number of lodged plants per plot and plant height at 6 months after planting made negative contributions to the fresh root yield through the root

weight. Edet *et al.* (2015), also observed a negative relationship between plant height and fresh root yield at 1 month after planting.

Mehari *et al.* (2015) ranked broad sense heritability as follows: low ($h^2 < 0.30$), moderate ($0.30 \leq h^2 \leq 0.60$) and high ($h^2 > 0.60$). Based on these classifications, broad sense heritability for both plant architecture and yield traits in this study across environments ranged from low to high: NDVI at 3 and 9 months after planting (0.0), levels of branching at 3 and 9 months after planting (0.11), stem diameter (0.65), fresh root yield (0.56), root weight (0.54), plant height at 9 months after planting (0.68), harvest index (0.66), starch content (0.61), dry matter content (0.61) and angle of branching (0.77). Similar results were reported by Ewa *et al.* (2017) on dry matter content (0.47) and fresh root yield (0.25). The range of the broad sense heritability values reported in this study are also consistent with the findings of Peprah *et al.* (2020), who reported broad sense heritability value of (0.41) for number of harvested roots, as well as Adu *et al.* (2020b) who reported harvest index value of (0.58), and fresh root weight (0.45). Harvest index (HI) is a measure of the efficiency of a crop in converting assimilated photosynthate (source) into economic yield (sink) (Alemu, 2021). The presence of moderately high values of broad sense heritability for harvest index 0.49 and high genetic advance as a percentage of mean (GAM) value of 21% shows that there is huge potential for desirable response to selection and optimization of cassava plants for this trait through recombination of favourable alleles.

However, broad sense heritability for NDVI data ranged from (0.008) to (0.13), so this shows that the environmental component of the phenotypic variance had a larger impact on the values of the NDVI data and this could explain the disparity observed in the degree of accuracy of the linear regression models developed using the data from the two trial environments.

Conclusions

In this research, two separate models (linear regression model and polynomial regression model) were used in developing phenotyping protocols that could be used in predicting the values of yield and plant architecture traits in cassava, using the Normalized Difference Vegetation Index (NDVI) data collected with the handheld Trimble GreenSeeker.

The results that were obtained using the polynomial regression model to predict fresh root yield from a number of variables which include: NDVI data, plant height, chlorophyll content, stem diameter, angle of branching, number of lodged plants per plot, number of harvested plants per plot and number of plants per stand showed that the prediction accuracy for the fresh root yield

across the two environments were more or less the same at 3, 6 and 9 months after planting. However, the values of (R^2) that signify the reliability of the prediction models were quite low.

Unlike the polynomial regression model where several variables were used in predicting the fresh root yield, the linear regression model entails the use of NDVI data in predicting the yield or the plant architecture trait directly and one trait at a time. Based on the results obtained from the prediction models that were developed using the linear regression model, there exist huge disparity between the data obtained from Mokwa and Onne trials with respect to the accuracy of prediction of the yield and plant architecture traits. Generally, higher coefficient of determination (R^2) values, which signify more reliable prediction accuracy and lower values of root mean squared error (RMSE) and mean absolute error (MAE) were observed for the plant architecture and yield traits using the NDVI data collected from Mokwa compared to that of Onne. This suggests that the disparity in prevailing environmental conditions could influence the accuracy of model development. This implies that the findings of this experiment should be subjected to further validation on a new set of cassava accessions and in different environments.

Ethics approval and consent to participate: Not applicable. This study is not a clinical trial and does not involve human participants, human data or human tissue. There are no necessary registration details to be provided.

Consent for publication: Not applicable. This manuscript does not contain any individual person's data in any form.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors without undue reservation.

Author Contributions

Abiodun Fatai Olayinka was responsible for conceptualization, data management, formal analysis, methodology, visualization and both drafting and revising the manuscript. Adesike Oladoyin Olayinka contributed to the conceptualization, formal analysis, methodology, validation, visualization, manuscript review and editing. Edwige Gaby Nkouaya Mbanjo provided supervision, methodology and participated in writing, reviewing and editing. Daniel Kwadjo Dzidzienyo provided supervision, methodology and participated in writing, reviewing and editing. Pangirayi B. Tongoona offered supervision and was involved in writing, reviewing and editing. Samuel Kwame Offei also provided supervision and took part in writing, reviewing and editing. Chiedozie Egesi was involved in conceptualization, supervision, writing, reviewing and editing. Ismail Yusuf Rabbi took part in conceptualization, formal analysis, supervision and manuscript review and editing.

Funding

Abiodun Fatai OLAYINKA carried out this research as part of his PhD studies in Plant Breeding at WACCI, University of Ghana, Legon, Accra, Ghana.

Acknowledgments

We appreciate the efforts and support received from the efficient and industrious members of the staff of the cassava breeding program at IITA, Ibadan during field evaluation and laboratory analysis.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References:

- Parmar A, Sturm B, Hensel O. Crops that feed the world: Production and improvement of cassava for food, feed, and industrial uses. *Food Security*. 2017;9(5):907-27.
- Tize I, Fotso AK, Nukenine EN, Masso C, Ngome FA, Suh C, Lenzemo VW, Nchoutnji I, Manga G, Parkes E, Kulakow P. New cassava germplasm for food and nutritional security in Central Africa. *Scientific Reports*. 2021;11(1):7394.
- Amelework AB, Bairu MW, Maema O, Venter SL, Laing M. Adoption and promotion of resilient crops for climate risk mitigation and import substitution: A case analysis of cassava for South African agriculture. *Frontiers in Sustainable Food Systems*. 2021;5:617783.
- Dordley L. Cassava, a staple crop that sustains a continent. *Food For Afrika*. 2022. <https://foodforafrika.com/2022/01/11/cassava-a-staple-crop-that-sustains-a-continent/>. Accessed 17 Sep 2023.
- EPA. Climate Impacts on Agriculture and Food Supply | Climate Change Impacts | US EPA. 2016. <https://climatechange.chicago.gov/climate-impacts/climate-impacts-agriculture-and-food-supply>. Accessed 17 Sep 2023.
- IPCC. Chapter 5 : Food Security — Special Report on Climate Change and Land. 2023. <https://www.ipcc.ch/srccl/chapter/chapter-5/>. Accessed 17 Sep 2023.
- Maxmen A. How African scientists are improving cassava to help feed the world. *Nature*. 2019;565(7738):144–6.

Ceballos H, Rojanaridpiched C, Phumichai C. Excellence in Cassava Breeding: Perspectives for the Future. *Crop Breeding, Genetics and Genomics*. 2020;1–31.

Conceição LV da, Cortes DFM, Klauser D, Robinson M, Oliveira EJ de. New protocol for rapid cassava multiplication in field conditions: a perspective on speed breeding. *Frontiers in Plant Science*. 2023;14. Available from: <https://www.frontiersin.org/articles/10.3389/fpls.2023.1258101>

Fei SZ, Mahama AA. Chapter 15: Cassava Breeding. 2023 Jun 14 [cited 2023 Sep 14]; Available from: <https://iastate.pressbooks.pub/cropimprovement/chapter/cassava-breeding/>

AiDash M. Remote sensing: How to collect a world of data. AiDash. 2021: <https://www.aidash.com/remote-sensing-the-art-behind-geospatial-data-collection-for-non-experts/>. Accessed 17 Sep 2023.

Machwitz M, Pieruschka R, Berger K, Schlerf M, Aasen H, Fahrner S, et al. Bridging the Gap Between Remote Sensing and Plant Phenotyping—Challenges and Opportunities for the Next Generation of Sustainable Agriculture. *Frontiers in Plant Science*. 2021;12. <https://www.frontiersin.org/articles/10.3389/fpls.2021.749374>

Marzougui A, McGee RJ, Van Vleet S, Sankaran S. Remote sensing for field pea yield estimation: A study of multi-scale data fusion approaches in phenomics. *Front Plant Sci*. 2023;14:1111575.

Tattaris M, Reynolds MP, Chapman SC. A Direct Comparison of Remote Sensing Approaches for High-Throughput Phenotyping in Plant Breeding. *Frontiers in Plant Science*. 2016;7. <https://www.frontiersin.org/articles/10.3389/fpls.2016.01131>

Huang J, Wang H, Dai Q, Han D. Analysis of NDVI data for crop identification and yield estimation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2014;7(11):4374–84.

Ji Z, Pan Y, Zhu X, Wang J, Li Q. Prediction of crop yield using phenological information extracted from remote sensing vegetation index. *Sensors (Switzerland)*. 2021;21(4):1–17.

Panek E, Gozdowski D. Analysis of relationship between cereal yield and NDVI for selected regions of Central Europe based on MODIS satellite data. *Remote Sensing Applications: Society and Environment*. 2020;100286.

Roznik M, Boyd M, Porth L. Improving crop yield estimation by applying higher resolution satellite NDVI imagery and high-resolution cropland masks. *Remote Sensing Applications: Society and Environment*. 2022;100693.

Trimble. Trimble Agriculture. GreenSeeker Handheld Crop Sensor.2023.
<https://www2.agriculture.trimble.com/product/greenseeker-handheld-crop-sensor/>. Accessed 23 May 2024.

Ayu Purnamasari R, Noguchi R, Ahamed T. Land suitability assessments for yield prediction of cassava using geospatial fuzzy expert systems and remote sensing. *Computers and Electronics in Agriculture*. 2019;166.

Rattanasopa K, Saengprachatanarug K, Wongpichet S, Posom J, Saikaew K, Ungsathittavorn K, et al. UAV-Based Multispectral Imagery for Estimating Cassava Tuber Yields. *Engineering in Agriculture, Environment and Food*. 2022;15(1):1–12.

Fukuda WM, Guevara CL, Kawuki R, Ferguson ME. Selected morphological and agronomic descriptors for the characterization of cassava. IITA; 1998.

Farias GD, Bremm C, Bredemeier C, de Lima Menezes J, Alves LA, Tiecher T, et al. Normalized Difference Vegetation Index (NDVI) for soybean biomass and nutrient uptake estimation in response to production systems and fertilization strategies. *Frontiers in Sustainable Food Systems*. 2023;6.
<https://www.frontiersin.org/articles/10.3389/fsufs.2022.959681>

Sahin H. Investigation of the effectiveness of AC/DC electric current as a weed control method using NDVI technique. *Adv Weed Sci.* 2022;21;40:e020220112.

Kuhn M. The caret Package. 2019. <https://topepo.github.io/caret/index.html>. Accessed 14 Feb, 2024.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 1. 2020.

Harrell F. Regression Modeling Strategies (RMS). 2023. <https://hbiostat.org/r/rms/>. Accessed 14 Feb, 2024.

Miller TL based on F code by A. leaps: Regression Subset Selection. 2022. <https://cran.r-project.org/web/packages/leaps/index.html>. Accessed 4 Jun, 2024.

Johnson HW, Robinson HF, Comstock RE. Estimates of genetic and environmental variability in soybeans. *Agronomy Journal.* 1955;47(314–18).

Alvarado G, Rodríguez FM, Pacheco A, Burgueño J, Crossa J, Vargas M, et al. META-R: A software to analyze data from multi-environment plant breeding trials. *Crop Journal.* 2020;8(5):745–56.

Wei T, Simko V. An Introduction to Corrplot Package. 2021. Available from: <https://cran.r-project.org/web/packages/corrplot/vignettes/corrplot-intro.html>. Accessed 15 Jan, 2024.

Rosseel Y. lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software.* 2023;48:1–36.

Epskamp S. semPlot: Path Diagrams and Visual Analysis of Various SEM Packages' Output version 1.1.6 from CRAN. 2022. <https://rdrr.io/cran/semPlot/>. Accessed 15 Jan, 2024.

Bakare MA, Kayondo SI, Aghogho CI, Wolfe MD, Parkes EY, Kulakow P, et al. Exploring genotype by environment interaction on cassava yield and yield related traits using classical statistical methods. PLOS ONE. 2022;17(7):e0268189.

Amelework AB, Bairu MW, Marx R, Owoeye L, Laing M, Venter SL. On-Farm Multi-Environment Evaluation of Selected Cassava (*Manihot esculenta* Crantz) Cultivars in South Africa. Plants. 2022;11(23):1–13.

Ogwuche TO, Diebiru-Ojo ME, Najimu A, Ossai CO, Ekanem U, Adegbite B, et al. Performance and Stability of Improved Cassava (*Manihot esculenta* Crantz) Clones in Demand Creation Trials in Nigeria. Crops. 2023;3(3):209–19.

Begna T. Application of genotype by environmental interaction in crop plant enhancement. Int J Res. 2022;8(2):1-2.

Sampaio Filho JS, Olivoto T, Campos M de S, de Oliveira EJ. Multi-trait selection in multi-environments for performance and stability in cassava genotypes. Front Plant Sci. 2023;14:1282221.

Grados D, Schrevers E. Cassava NDVI Analysis : A Nonlinear Mixed Model Approach Based on UAV - Imagery. PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science. 2020;(0123456789).
<https://doi.org/10.1007/s41064-020-00116-x>

León R, Rosero A, García JL, Morelo J, Orozco A, Silva G, et al. Multi-trait selection indices for identifying new cassava varieties adapted to the caribbean region of Colombia. Agronomy. 2021;11(9).

Malai C, Maraphum K, Saengprachatanarug K, Wongpichet S, Phuphaphud A, Posom J. Effective measurement of starch and dry matter content in fresh cassava tubers using interactance Vis/NIR spectra. Journal of Food Composition and Analysis. 2024;125:105783.

Peprah BB, Parkes E, Manu-Aduening J, Kulakow P, van Biljon A, Labuschagne M. Genetic variability, stability and heritability for quality and yield characteristics in provitamin A cassava varieties. *Euphytica*. 2020;216(2):1–13.

Silva RM da, Fernandes AM, Leonel M, Pelvine RA, Figueiredo RT de, Rangel MAS, et al. Measurement of Dry Matter and Starch in Modern Cassava Genotypes during Long Harvest Cycles. *Horticulturae*. 2023;9(7):733.

Amarullah. Morphological, Physiological and Agronomic Characteristics of Cassava Superior Variety of Coastal Land - IOPscience. 2021.
<https://iopscience.iop.org/article/10.1088/1755-1315/748/1/012030>.

Edet MA, Tijani-Eniola H, Lagoke STO, Tarawali G. Relationship of Cassava Growth Parameters with Yield, Yield Related Components and Harvest Time in Ibadan, Southwestern Nigeria. 2015;

Mehari G, Amsalu N, Tewodros M. Estimates of Genetic Components for Yield and Quality of Cassava (*Manihot esculenta* Crantz) Genotypes at Jimma, Southwest Ethiopia. 2015. <https://scialert.net/abstract/?doi=ijpbg.2015.1.12>

Ewa F, Nwofia E, Egesi C, Olasanmi B, Okogbenin E. Genetic variability, heritability and variance components of some yield and yield related traits in second backcross population (BC2) of cassava. *AJPS*. 2017;11(6):185–9.

Adu MO, Asare PA, Yawson DO, Nyarko MA, Razak AA, Kusi AK, et al. The search for yield predictors for mature field-grown plants from juvenile pot-grown cassava (*Manihot esculenta* Crantz). *PLoS ONE*. 2020;15(5):1–27.

Alemu G. The role of harvest index in improving crop productivity. *Journal of Natural Science Research*. 2019;9(6):24-8.

941 **Table 4 |**

	NDVI3	NDVI6	NDVI9	HI	LODG	PLTHT3	PLTHT6	PLTHT9	ANGBR9	STMDI	TYLD	DYLD	FYLD	STARCH	DM	RTWT	SHTWT
ENV	1.72***	29.72***	23.14***	11.8***	23.14***	1573**	247743.74***	247743.74***	3137.82***	2819.71***	125111.94***	118.92***	192.25*	6572.41***	3706.58***	11576.81***	29827.92***
BLK(ENV*REP)	0.006***	0.07	0.006***	0.009***	0.006***	384.34*	518.46***	518.46***	259.01*	11.04***	117.07***	12.09***	117.15***	6.13***	3.46**	84.63***	79.10**
REP(ENV)	0.17***	0.05	0.05***	0.01	0.05***	15686.45***	20057.5***	20057.5***	316.12	195.71***	1563.36***	80.66***	659.64***	14.49**	8.19*	441.03***	890.61***
GEN	0.003***	0.01	0.005***	0.03***	0.005***	1121.88***	1672.35***	1672.35***	2312.22***	32.66***	275.73***	22.16***	247.44***	33.69***	19.0***	190.93***	199.92***
GEN * ENV	0.00	0.01	0.005***	0.009***	0.005***	393.4***	489.39***	489.39***	654.54***	12.12***	141.63***	11.95***	105.99***	13.67***	7.71***	85.0***	99.43***
Residuals	0.00	0.06	0.00	0.00	0.00	312.35	326.15	326.15	200.059	6.99	83.44	5.94	54.68	3.95	2.23	40.85	56.33

942
943 ***, **, * significant at: 0.1%, 1% and 5% respectively

944 NDVI3 – NDVI at 3 months after planting, NDVI6 - NDVI at 6 months after planting, NDVI9 - NDVI at 9 months after planting, HI – harvest index, LODG – number of lodged plants per plot, PLTHT3 – plant height
945 at 3 months after planting, PLTHT6 – plant height at 6 months after planting, PLTHT9 – plant height at 9 months after planting, ANGBR9 – angle of branching, STMDI - stem diameter, TYLD – top yield, DYLD - dry
946 yield, FYLD - fresh root yield, STARCH – starch content, DM – dry matter content, RTWT – fresh root weight, SHTWT – shoot weight

947
948
949

Figures



Figure 1

Trimble GreenSeeker, a vegetation sensor

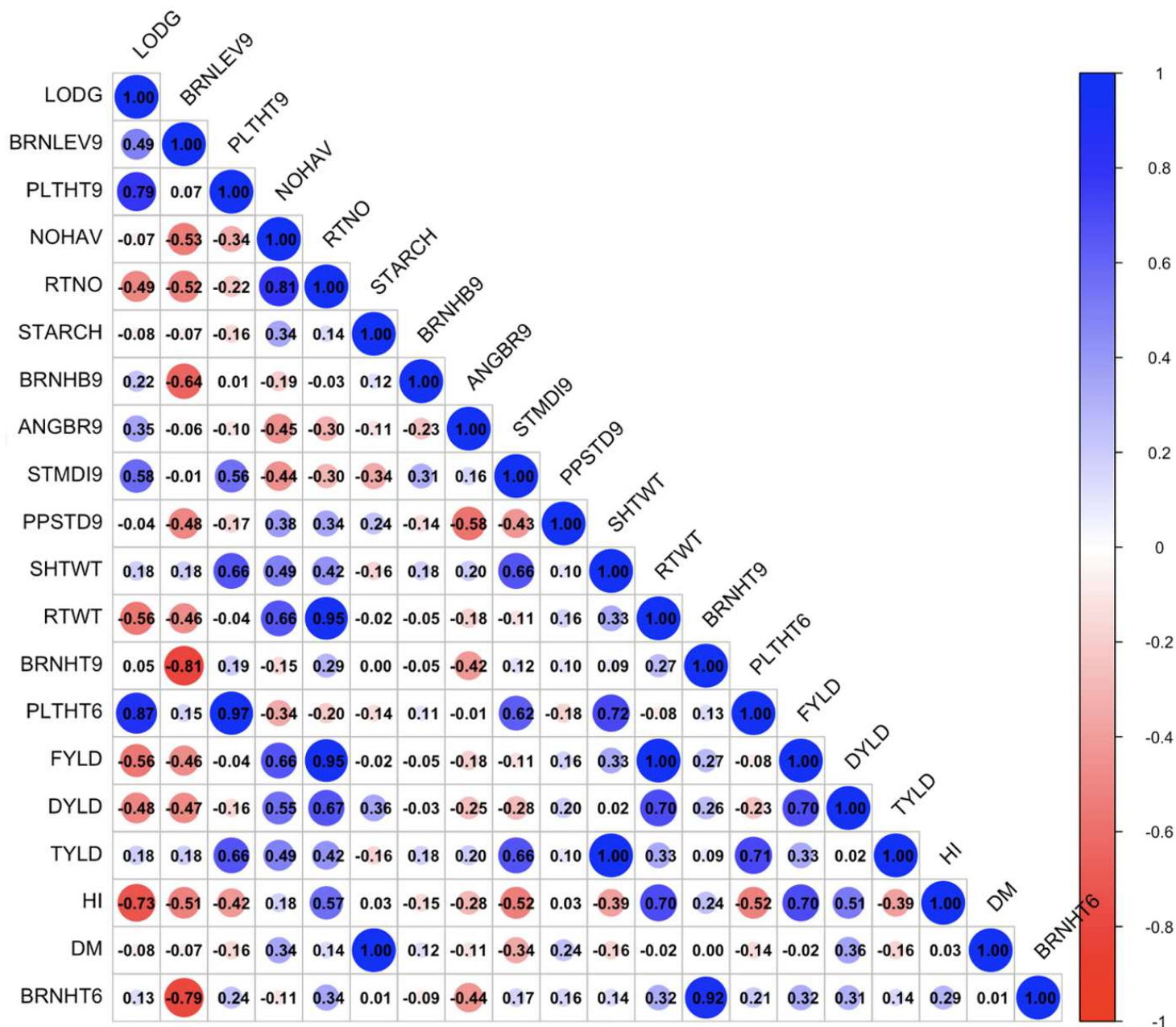


Figure 2

Genotypic correlation coefficient plot of plant architecture and yield traits in Mokwa trial

NOHAV - number of harvested plants per plot, RTNO – number of harvested roots per plot, SHTWT – shoot weight, RTWT – root weight, DM – dry matter content, STARCH – starch content, FYLD – fresh root yield, DYLD – dry yield, TYLD – top yield, HI – harvest index, LODG – number of lodged plants per plot, PLTHT6 – plant height at 6 months after planting, BRNHT6 – height at first branch at 6 months after planting, PLTHT9 – plant height at 9 months after planting, BRNHT9 – height at first branch at 9 months after planting, BRNLEV9 - level of branching at 9 months after planting, BRNHB9 – branching habit at 9 months after planting, ANGBR9 – angle of branching, and STMDI9 – stem diameter.

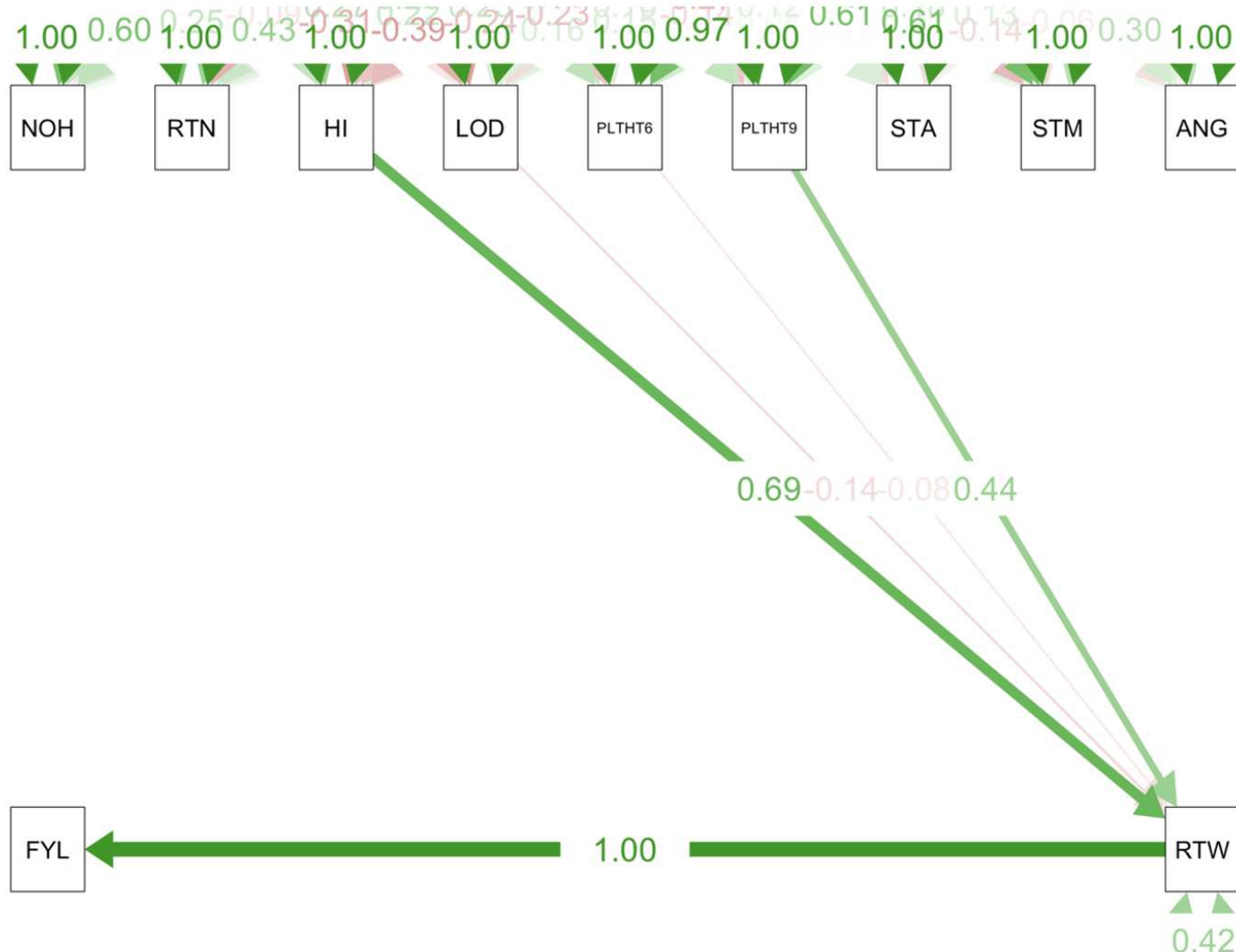


Figure 3

Path coefficient analysis plot of plant architecture and yield traits in Mokwa trial

NOH - number of harvested plants per plot, RTN – number of harvested roots per plot, RTW – root weight, STA – starch content, FYL – fresh root yield, HI – harvest index, LOD – number of lodged plants per plot, PLTH6 – plant height at 6 months after planting, PLTH9 – plant height at 9 months after planting, ANG – angle of branching, and STM – stem diameter.

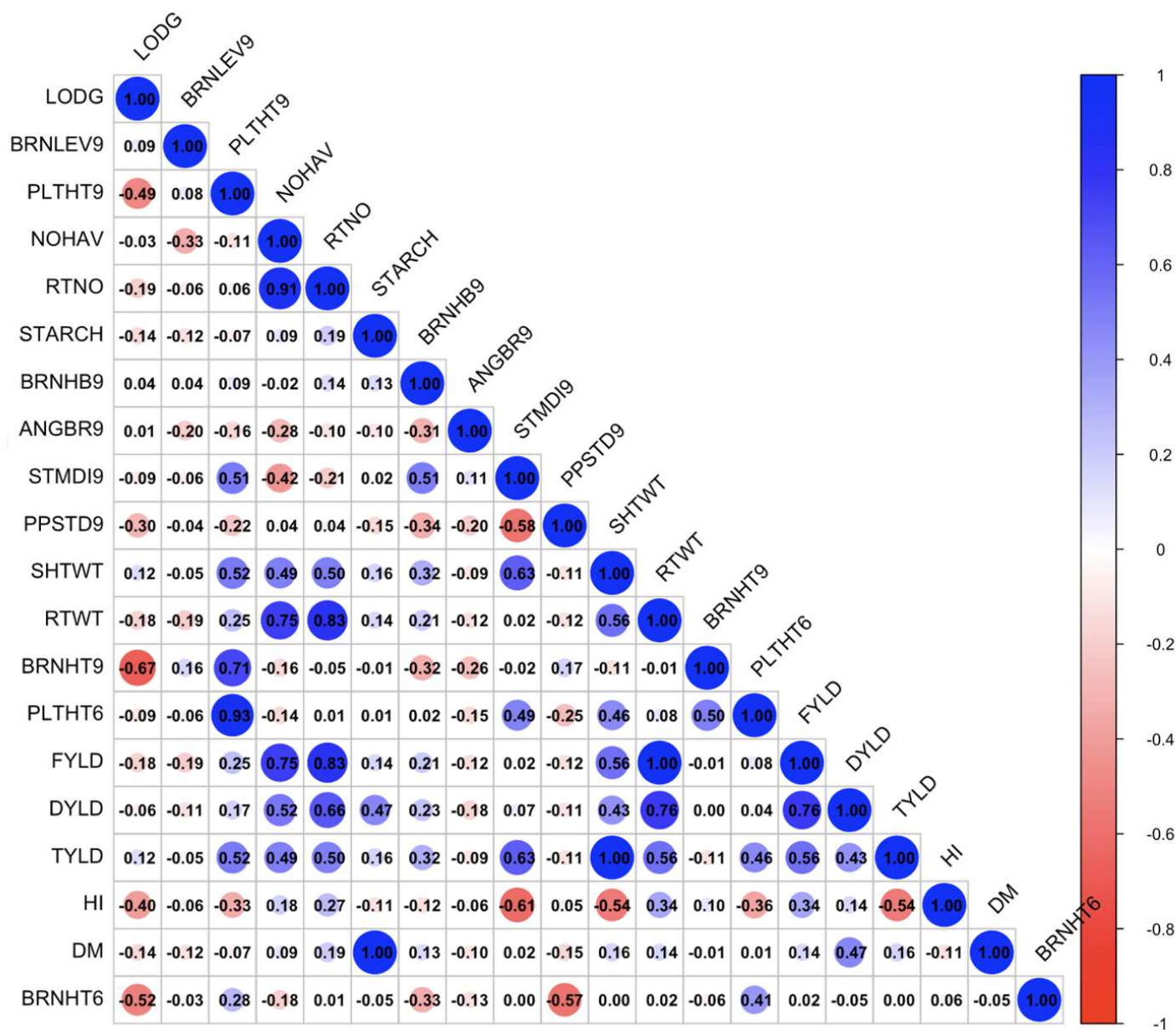


Figure 4

Genotypic correlation coefficient plot of plant architecture and yield traits in Onne trial

NOHAV - number of harvested plants per plot, RTNO – number of harvested roots per plot, SHTWT – shoot weight, RTWT – root weight, DM – dry matter content, STARCH – starch content, FYLD – fresh root yield, DYLD – dry yield, TYLD – top yield, HI – harvest index, LODG – number of lodged plants per plot, PLTHT6 – plant height at 6 months after planting, BRNHT6 – height at first branch at 6 months after planting, PLTHT9 – plant height at 9 months after planting, BRNHT9 – height at first branch at 9 months after planting, BRNLEV9 - level of branching at 9 months after planting, BRNHB9 – branching habit at 9 months after planting, ANGBR9 – angle of branching, and STMDI9 – stem diameter.

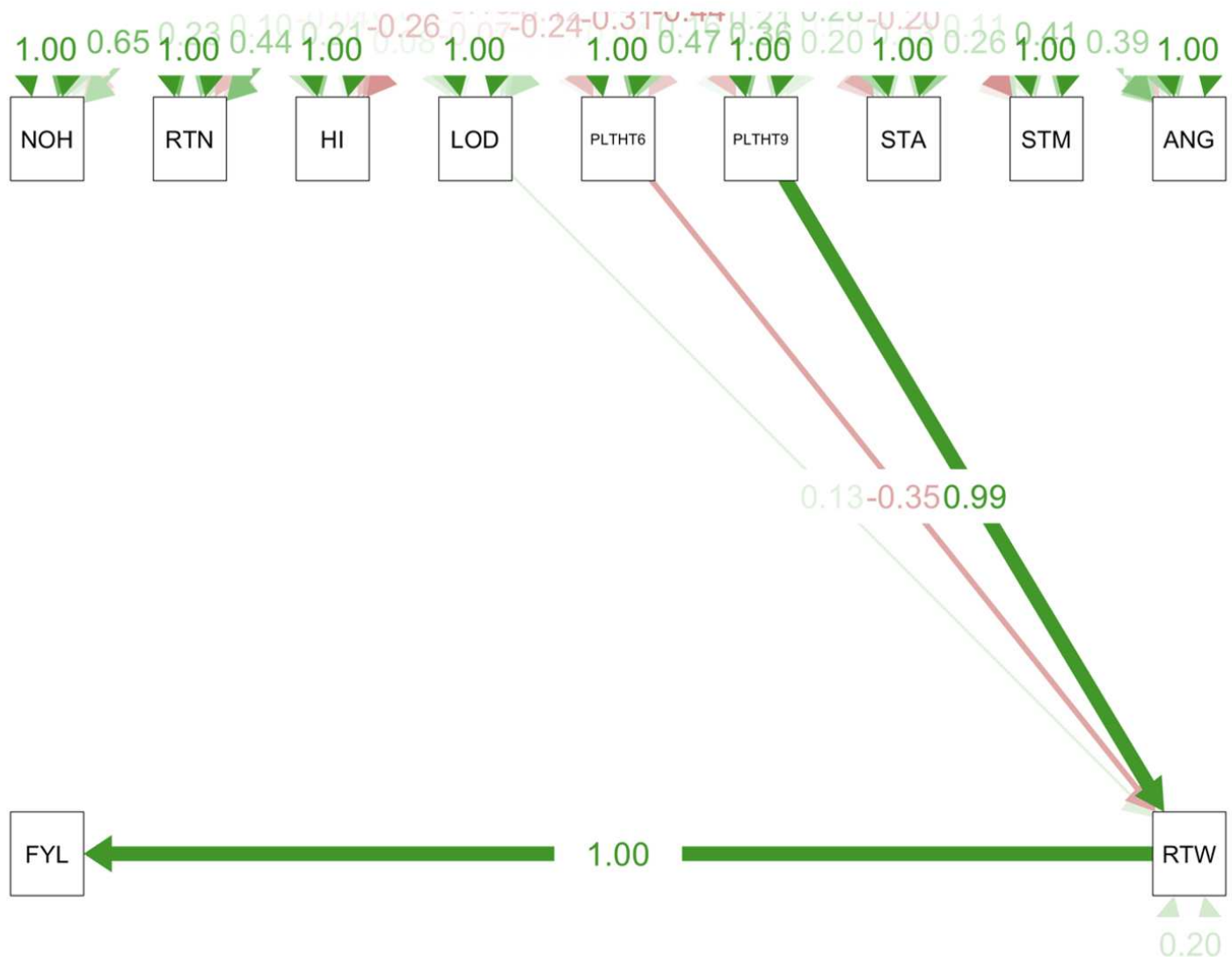


Figure 5

Path coefficient analysis plot of plant architecture and yield traits in Onne trial

NOH - number of harvested plants per plot, RTN – number of harvested roots per plot, RTW – root weight, STA – starch content, FYL – fresh root yield, HI – harvest index, LOD – number of lodged plants per plot, PLTH6 – plant height at 6 months after planting, PLTH9 – plant height at 9 months after planting, ANG – angle of branching, and STM – stem diameter.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1.docx](#)
- [TableS2.docx](#)
- [TableS3.docx](#)