Research Article

Received: 31 January 2023 Revised: 6 July 2023 Accepted article published: 5 September 2023 Published online in Wiley Online Library:

(wileyonlinelibrary.com) DOI 10.1002/jsfa.12966

Development and validation of near-infrared spectroscopy procedures for prediction of cassava root dry matter and amylose contents in Ugandan cassava germplasm

Ephraim Nuwamanya,[a,b] Enoch Wembabazi,[a] Michael Kanaabi,[a,b] Fatumah Babirye Namakula,[a] Arnold Katungisa,[a] Ivan Lyatumi,[a] Williams Esuma,[a] Emmanuel Oladeji Alamu,[c] Dominique Dufour,[d] Robert Kawuki[a] and Fabrice Davrieux[d*]

Abstract

BACKGROUND: Cassava utilization for food and/or industrial products depends on inherent properties of root dry matter content (DMC) and the starch fraction of amylose content (AC). Accordingly, in the present study, near-infrared reflectance spectroscopy (NIRS) models were developed to aid breeding and selection of DMC and AC as critical industrial traits taking care of root sample preparation and cassava germplasm diversity available in Uganda.

RESULTS: Upon undertaking calibrations and cross-validations, best models were adopted for validation. DMC in calibration samples ranged from 20 to 45 g 100g^{-1}, whereas, for amylose content, it ranged from 14 to 33 g 100g^{-1}. In the validation set, average DMC was 29.5 g 100g^{-1}, whereas, for amylose content, it was 24.64 g 100g^{-1}. For DMC, a modified partial least square regression model had regression coefficients (R^2) of 0.98 and 0.96, respectively, in the calibration and validation set. These were also associated with low bias (−0.018) and ratio of performance deviation that ranged from 4.7 to 5.0. In addition, standard error of prediction values ranged from 0.9 g 100g^{-1} to 1.06 g 100g^{-1}. For AC, the regression coefficient was 0.91 for the calibration set and 0.94 for the validation set. A bias equivalent to −0.03 and a ratio of performance deviation of 4.23 were observed.

CONCLUSION: These findings confirm the robustness of NIRS in the estimation of dry matter content and amylose content in cassava roots and thus justify its use in routine cassava breeding operations.

Supporting information may be found in the online version of this article.

Keywords: cassava; dry matter content; amylose content; NIRS; selection

INTRODUCTION

Cassava (*Manihot esculenta* Crantz) is one of the most famous crops in Uganda as a result of its wide use in food, feed and industry.[1] These applications are mainly dependent on root properties such as the dry matter content (DMC) and the starch properties. In addition, root properties, which are mainly a function of starch, determine the quality of cassava or its products.[2] It is for these reasons that most cassava breeding programs consider selecting for high DMC as a 'must-have' trait when developing new varieties.[3] Additionally, starch related quality traits such as amylose content are currently being considered for the industrial utilization of cassava.

* *Correspondence to: F Davrieux, UMR Qualisud, University of Montpellier, CIRAD, Montpellier SupAgro, University of Avignon, University of La Reunion, 73 Rue JF Breton, Montpellier 34398, France. E-mail: fabrice.davrieux@cirad.fr*

a *National Crops Resources Research Institute, Kampala, Uganda*

b *Makerere University Kampala, Kampala, Uganda*

d *International Institute of Tropical Agriculture (IITA), Lusaka, Zambia*

e *UMR Qualisud, University of Montpellier, CIRAD, Montpellier SupAgro, University of Avignon, University of La Reunion, Montpellier, France*

Traditionally, most cassava breeding programs have estimated DMC and amylose content using laboratory procedures such as the oven based method for DMC[4] and iodometric method for amylose content.[5] However, these methods are characterized by drudgery and lower precision, which consequently reduces breeding efficiency in the long run.[6] For these reasons, alternative and high throughput methods characterized by speed, versatility and precision are required among breeding programs. Such alternatives include a range of high throughput phenotyping procedures, the most notable of which is near-infrared reflectance spectroscopy (NIRS).[7]

Previous studies have recognized NIRS as a robust technique for estimating DMC in cassava roots.[8-10] Such studies act as a basis for estimating amylose content and the starch yield. As a rapid and low cost technique (in terms of sample preparation and reduced wet chemistry costs), NIRS has the transformative power to measure breeding traits of up to 100 samples per hour. Absorption of wavelengths in the near-infrared electromagnetic region by molecular groups' particularly involving hydrogen bonds (C—H, O—H and N—H) underpins the appropriateness of NIRS. NIRS is therefore important in estimation of primary constituents of organic compounds of plant tissues.[10]

Development of NIRS protocols for DMC, amylose and starch by individual breeding programs is important in the improvement of cassava. This is because of the variability in sample preparation and presentation forms, trait preferences for the target product profiles, and the different working conditions within specific breeding programs. Such differences significantly influence the analysis[9] and hence call for program specific calibrations. Sampling methods are also limited by resource capabilities of different breeding programs, necessitating the development and customization of locally adapted calibrations. Other factors such as vibration mechanisms of spectrum, mathematical and statistical procedures performed, as well as submission and preparation conditions of the sample, also affect predictions.[11]

Given that selection metrics are dependent on proper measurement, the requirement of high throughput phenotyping procedures (HTPPs) is critical. These procedures depend on instrumentation such as NIRS where a trait by trait approach is needed for locally adapting use. For these tools to be used, there is a need for their optimization and for models to be developed for their use. Specifically, there is a need to develop calibrations that take into consideration a range of specific sampling procedures, mindful of the evolution in NIRS technology and equipment. Thus, the present study aimed to optimize NIRS as an HTPP tool for the selection of root quality traits of DMC, amylose content and starch yield.

MATERIALS AND METHODS

Selection of the calibration samples

Different sample sets were used in the development of calibration. For dry matter content, the calibration set was obtained from a diverse set of 300 cassava clones introduced from Latin America for pre-breeding at National Crops Resources Research Institute (NaCRRI) and additional local checks (https://cassavabase.org/breeders/program/164). For the amylose content, the calibration set consisted of a diverse set of 197 clones. The details of the field

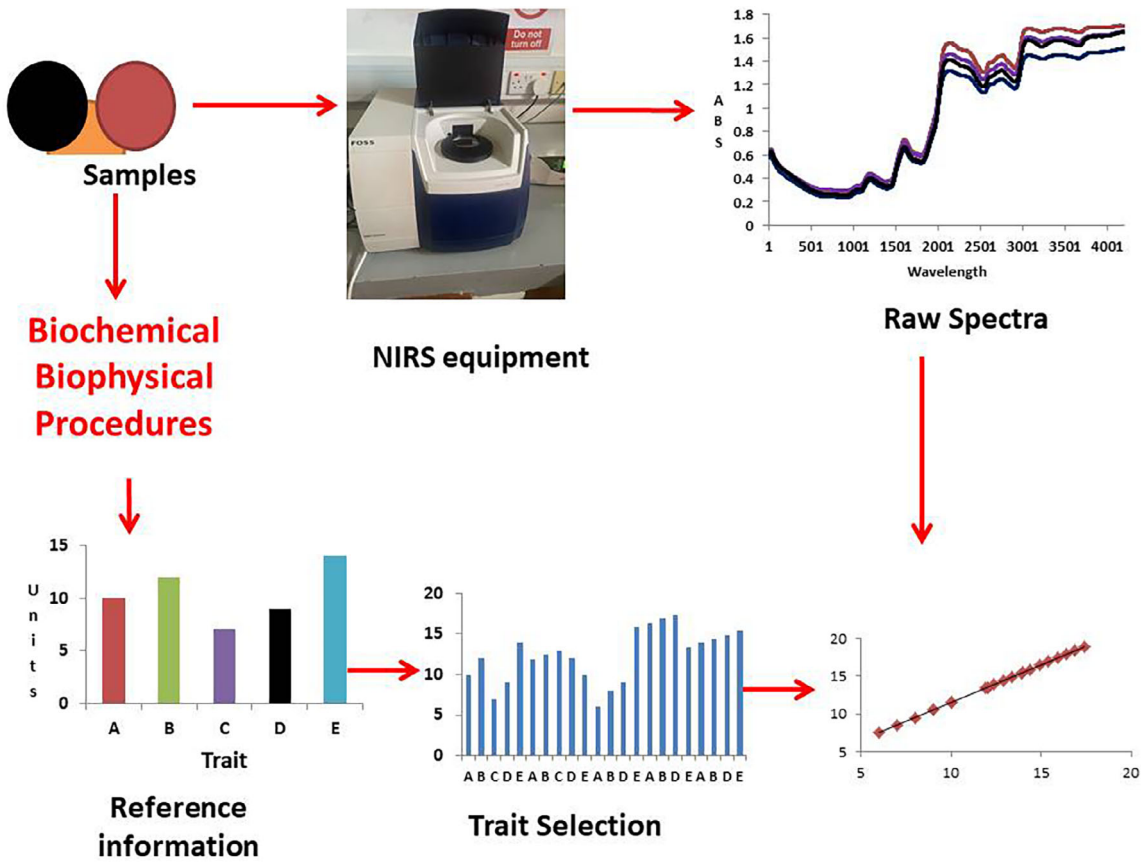

Figure 1. Schematic showing sample handling and analysis of starch/amylose parameters.

experiment is available in an open access data repository at (https://www.cassavabase.org/breeders/trial/4384?format=). The pre-breeding set of germplasms used in the present study contained genotypes that are from diverse backgrounds (from International Institute of Tropical Agriculture (IITA), International Center for Tropical Agriculture (CIAT) and NaCRRI), for which diversity is important in development of NIRS calibrations. They are currently being used to breed for future populations for different traits including DMC and amylose content.

In all cases, the trials were harvested at maturity (12 months after planting), three plants per clone were selected to get at least six to 10 root samples, with each root measuring at least 30 cm in length and 6 cm in diameter. Sampled roots were immediately taken to the laboratory and each root was peeled, washed and dried using a towel. Peeled roots were grated using a kitchen grater to generate approximately 300 g of cassava mash that was placed into an aluminum foil, covered and labeled awaiting spectra acquisition (Fig. 1). The peeled grated cassava root tissue was preferred because of its ability to provide for better spectral acquisition and hence increased spectral reproducibility.[4]

Validation samples

Part of the samples used for calibration was selected and used for constituting a validation set. For dry matter content (100 samples) were selected, whereas 74 samples were selected for amylose content. These samples were sourced from Uganda's cassava breeding program (https://cassavabase.org/breeders/program/164) as described for the calibration set and treated in the same way as the calibration sets before being used for validation. The details of the sample sets used are provided in Table 1 with further description in Figs 2 and 3.

Spectral data acquisition

A benchtop Vis/NIRS device (DS2500; FOSS, Hilleroed, Denmark) was used to acquire spectral data on all grated root samples. Immediately after grating, sample portions (approximately 15 g) were used for spectra acquisition. Samples for spectra acquisition were placed in the small sample cup, which was then placed on the stage and spectral information acquired. Samples were scanned in diffuse reflectance between 400 nm and 2500 nm with a 2-nm step. For each spectrum, a total of two sub-spectra were collected per sample from refilling the sample cup two times with the grated sample and averaged to get the main spectra for that sample. Instrument control was performed with the ISIscan Analysis Software (Infrasoft International LLC, State College, PA, USA).

Determination of DMC

For both calibration and validation sets, DMC was determined from the measurement of the dry weight of the sample after oven-drying and presented as a percentage. Grated fresh root samples (100 g measured to a precision of 0.01 mg) obtained from the same plot were weighed off in triplicate on aluminum dry matter plates. The samples were oven-dried at a constant temperature of 105 °C for 48 h. Thereafter, samples were weighed upon attainment of constant weight.[4,10] The average DMC of the two replications was recorded using the Eqn (1) and used as reference data during model development.

$$\mathrm{DMC} = \frac{W2}{W1} \times 100 \tag{1}$$

where W1 is the weight of the chopped cassava sample before drying and W2 is the weight of the cassava sample after drying.

Methodology for amylose determination

For amylose determination, starch was initially extracted from grated samples. Selected roots were manually peeled with a laboratory knife and the resultant tissue was crushed in a laboratory blender. The slurry was filtered through a muslin cloth. The starch was allowed to settle and the supernatant decanted off. The remaining starch was washed with water for two times and then dried in an air forced oven at 40 °C for 48 h.

The amylose content of the extracted cassava starch was determined using a spectrophotometric procedure based on iodine staining with slight modifications by adjusting the volume of analysis to a total working volume of 50 mL (ISO 6647-2:2020).[5] The amylose standard curve was prepared using potato amylose (Sigma Chemicals, St Louis, MO, USA). Repeated analyses were undertaken where the SD was higher than 5 g 100g^{-1} to ensure consistency. The results were presented as percentage amylose on dry weight basis.

Data analysis

Data pre-processing and model development

Spectra used covered the full wavelength range (400–2500 nm) of visible/NIRS and data in these spectra were pre-treated for improvement and reduction of interferences. Pre-treatment involved undertaking light scatter correction methodologies including the standard normal variate and de-trending (SNVD) and the multiplicative scatter correction (see Supporting information, Tables S1 and S2). For each of these, four derivative and smoothing options were used and compared with no treatment.[5] Based on the different data pre-processing techniques, models were developed using the calibration set and the best model was selected based on: highest coefficient of determination of calibration (R^2c); R^2cv of internal cross-validation; lowest standard error of calibration (SEC) and standard error of cross-validation (SECV); and the smallest difference between SEC and SECV.

Prediction of root traits

The selected model was used to predict DMC of validation sets on either whole root. In all cases, samples were prepared by grating and root portions treated as sample categories. The validation statistics of interest used to understand the quality and performance of the models included coefficient of determination of prediction (R^2p), the bias and the ratio of performance to deviation (RPD).

Table 1. Descriptive statistics for dry matter content and amylose content

Statistic	*N*	Minimum	Maximum	Mean	SD
Dry matter content (g 100g^{-1})	300	10.74	45.00	29.58	5.08
Amylose content (g 100g^{-1})	247	15.17	32.96	24.52	3.37

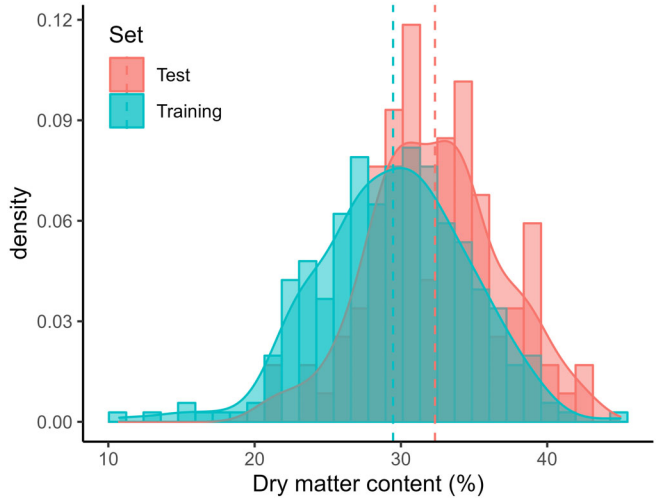

Figure 2. Dry matter content distribution per set.

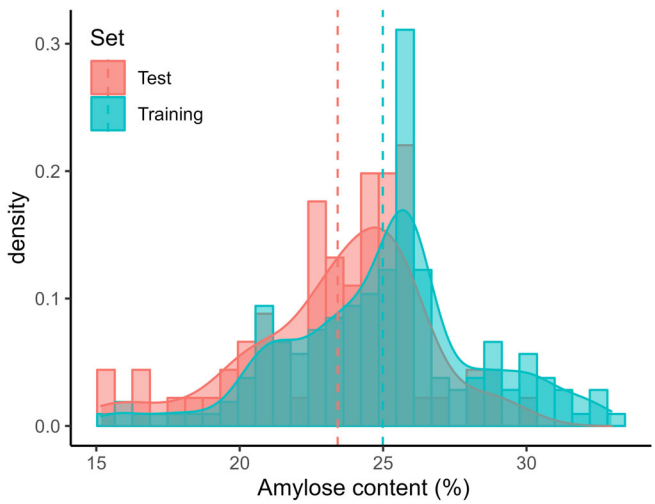

Figure 3. Amylose content distribution per set.

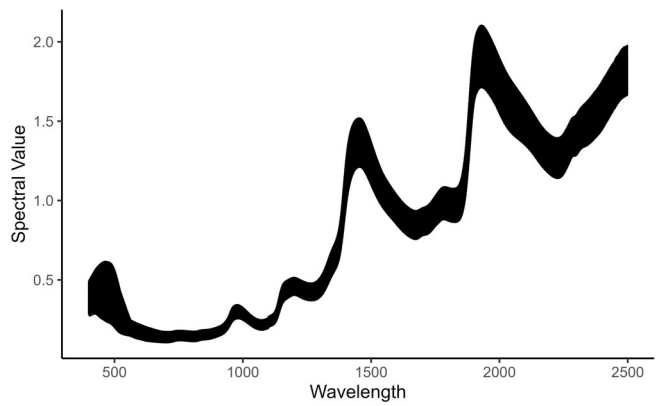

Figure 4. Representative cassava root spectra for amylose calibration with no treatment.

Assessment of the robustness of the prediction models also depended on the standard error of prediction (SEP) and standard error of prediction corrected for bias (SEPC).

Statistical analysis

Statistical analyses were performed using Win-ISI 4.6 software (Infrasoft International and FOSS, Hillerod, Denmark) and R software, version 4.1 (R Foundation, Vienna, Austria). Different pretreatments were tested and the mathematical correction for light scattering using SNVD correction was selected. Partial least squares (PLS) and modified partial least square (MPLS) algorithms were used to develop prediction models. Based on different data pre-processing techniques, models were developed using the calibration set and the best model was selected. Specific factors for each PLS or MPLS model were optimized with WinISI 4.6 software. In the process of calibration development, cross-validation was used to select the optimum number of latent variables and to minimize overfitting the equations. The identification of outlier samples during calibration development was based on the Student's *t*-test. Outlier detection was based on the standardized residuals (= error/SECV) with a cutoff of 2.5. The RPD was calculated by dividing the SD of the reference data from the validation set by the SEP. Scatter plots were used to visualize the relationship between predicted and reference values using Excel (Microspft Corp., Redmond, WA, USA).

RESULTS

Phenotypic variation of root traits

The calibration set consisting of Latin American clones and Ugandan local checks had DMC ranging from 21.5 to 44.5 g 100g^{-1} (Fig. 1). Average DMC was 33.7 g 100g^{-1} with SD of 4.53. On the other hand, the amylose content ranged from 15.2 g 100g^{-1} up to 33 g 100g^{-1} (Fig. 3). Average amylose content was 24.48 g 100g^{-1} with a SD of 3.44. For the validation set, average DMC ranged from 10.7 g 100g^{-1} to 44.3 g 100g^{-1} and on average was 29.46 g 100g^{-1}.

All the cassava genotypes used in the development of the calibration had a definitive spectra with similar patterns observed. There were no atypical spectra observed for both the dry matter and amylose sets used. Spectral variability was observed with differences observed in the water absorption bands (1500 and 1900 nm) (Fig. 4).

Journal of The Science of Food and Agriculture published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.

Table 2. Spectral treatment parameters selected for development of calibrations

Trait	Math treatment	Regression statistic	Scatter	*N*	No. of terms
Dry matter content	SNVD	MPLS	2,4,4,2	262	7
	SNVD	MPLS	2,2,2,2	262	6
Amylose content	None	MPLS	0,0,1,1	197	5
	SNVD	MPLS	2,2,2,2	183	5

Abbreviation: MPLS, modified partial least squares; SNVD, standard normal variate and detrend; N, total number of samples.

Table 3. Calibration statistics for the two traits

Parameter	Scatter	*N*	SEC	SECV	1 − VR	RSQ
Dry matter content	2,2,2,2	262	0.57	0.848	0.963	0.983
	2,4,4,2	262	0.525	0.826	0.965	0.986
Amylose content	0,0,1,1	197	1.267	1.346	0.847	0.865
	2,2,2,2	183	1.510	2.631	0.413	0.779

Abbreviation: SEC, standard error of calibration; SECV, standard error of cross validation; 1 − VR, Statistic 1 − variance ratio; RSQ, coefficient of determination of calibration.

Table 4. Validation statistics for each of the traits (SEP, R^2p, SEPC, RPD and bias)

Parameter	SEP	SEPC	RPD	Bias	R^2p	Predicted average
Dry matter content	1.035	1.037	4.93	0.018	0.962	29.48
Amylose	0.733	0.74	4.23	−0.026	0.943	24.68

Abbreviation: SEP, standard error of prediction; SEPC, standard error of prediction corrected for bias; RPD, Ratio of Perfromance deviation; R^2p, coefficient of prediction.

To compare the effect of mathematical treatment, calibrations based on SNVD were developed using all the spectral segments (visible + NIR). These were compared with no treatments at all and with the main regression method being the MPLS method. Differences were observed in performances of equations developed using SNVD with the MPLS, and the equations which had higher coefficient of determination (R^2c) across all calibrations were selected (Table 2). The number of terms used for this regression to achieve a maximum R^2c value close to 0.98 and minimum difference between SEC and SECV were six[6] terms. The number of terms in each calibration ranged from six to seven for the dry matter content and was five terms for the amylose content.

A different mathematical treatment was used to develop calibration equations (Table 2). For dry matter content, the mathematical treatment SNVD was used across different scatter treatments producing R^2c values ranging from 0.97 to 0.98 for the corresponding MPLS regression. On the other hand, for amylose content, where no specific treatments were used for the MPLS regression, an R^2c value of 0.865 was realized, whereas, with treatments for SNVD, an R^2c of 0.79 was realized. For dry matter content, the SECV values observed for the MPLS regression ranged from 0.826 to 0.848 for both treatments, whereas SEC values ranged from 0.52 to 0.57. On the other hand, the SECV values observed for amylose content was 1.346 for the scatter 0.0.1.1 with a corresponding standard error of calibration of 1.267 and 2.631 for the scatter 2,2,2,2 with a corresponding standard error of calibration of 2.631 (Table 3,4). The squared correlation coefficient (RSQ) for dry matter content ranged between 0.983 and 0.986 for the selected scatter and was high enough for the calibrations to be considered for utilization in selection. Likewise, the RSQ for amylose content ranged between 0.779 and 0.865, explaining a total variation of up to 86.5 g $100g^{-1}$. The RSQ values also were well fitting for the developed calibration models meant for selection at early stages in the breeding program. The low SEC values obtained for the two parameters showed that the selected equations would have acceptable quantitative correlations for the amylose content and the dry matter content predictions.[11] Likewise, the difference between the SEC and SECV was low, which further showed the robustness of the developed equations and their predictive abilities. The internal cross-validation for the selected equations was also high enough (0.961–0.973) to allow for utilization of the calibrations. In addition, the SECV values observed were close to SEC values, and thus showed that the developed models were fair and robustly fitting for DMC estimation.

Calibration equations developed were evaluated to identify their prediction accuracy based on correlation coefficient of actual and predicted values in the external cross-validation set using the R^2p, SEP, the bias and the RPD (Table 4). For dry matter content, the smoothing and derivative options under SNVD 2,4,4,2 had R^2p values ranging from 0.962. For these options, the RPD was 4.93, whereas the bias was 0.018. These

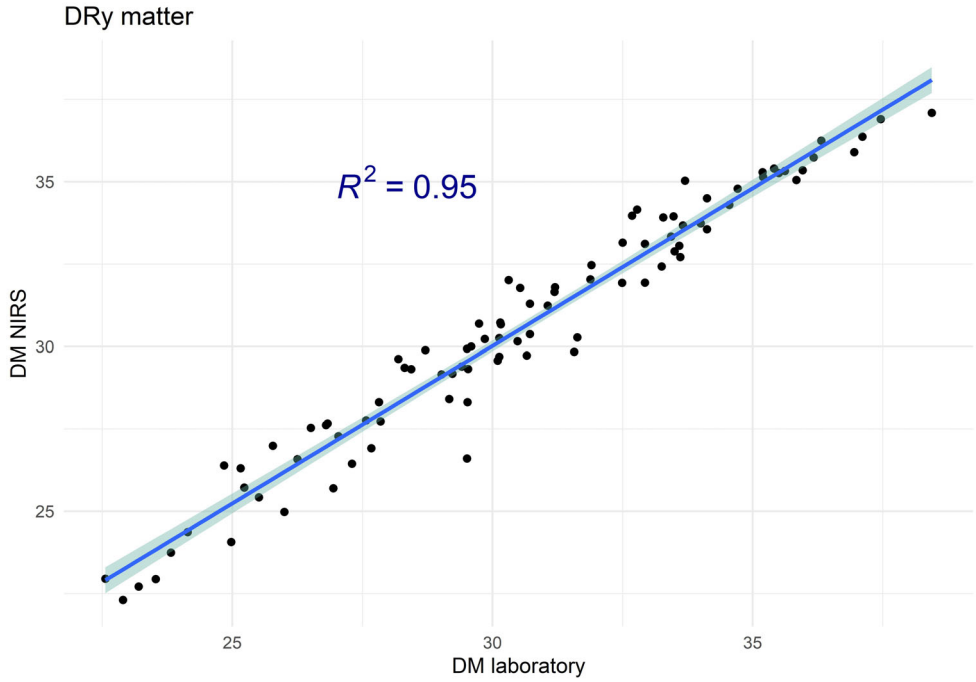

Figure 5. Reference values from oven-based dry matter content (DM laboratory) estimation compared to NIRs predicted dry matter content (DM NIRS) for the dry matter content.

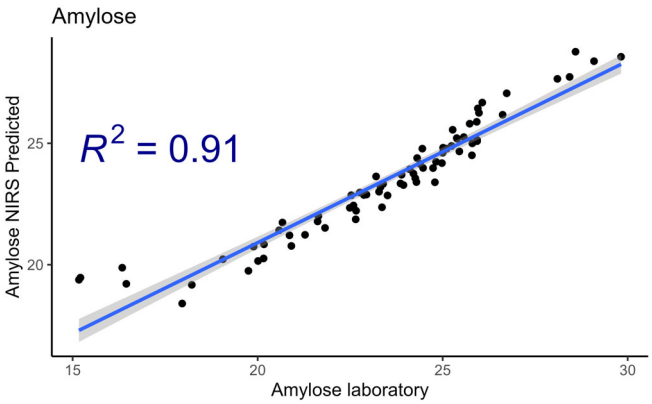

Figure 6. Reference values for amylose content (amylose labarotory) compared to predicted amylose (amylose NIRs predicted).

resulted into a predictive average of a dry matter content of 29.48 g 100g^{-1}, which ranged from 28.9 to 30.6 g 100g^{-1}. On the other hand, the derivative option 0,0,1,1 for amylose content resulted in a SEP of 0.733 and an RPD of 4.23. The bias observed for this option was −0.026, whereas the predicted average was 24.68 g 100g^{-1} (Tables 3,4).

The observed R^2p and SEP were strong enough to allow for evaluation of predictive accuracy of models (Figs 5 and 6).

Validation of the NIRS models

The R^2p for the whole root, averaged at 0.95, with the predicted DMC ranging from 28.86 to 30.6 g 100g^{-1} (Table 4). SEP was 1.04 g 100g^{-1}, 1.07 g 100g^{-1}, 1.06 g 100g^{-1} and 0.97 g 100g^{-1}, respectively. The RPD values ranged from 4.7 to 5.01 (Table 4) and underpinned the robustness of models in the estimation of DMC (Figs 5 and 6).

DISCUSSION

One of the major drawbacks in measurement of root quality traits across cassava breeding programs relates to methods for estimation of such traits and the sampling procedures used.[8,9] The results from the present study highlight the role of NIRS in accurately estimating cassava root quality traits. The high predictive accuracy for root quality traits in in grated root samples shows that NIRS can be used not only for screening large populations, but also in quality control parameters involving determination of root quality traits. Particularly, the observations in the present study showed that minimum preparation procedures involving the increase in surface area though grating provide better root quality traits estimates. Such recommendations were suggested previously.[8] Sample preparation and presentation forms are very important in the successful development of calibrations for particular parameters.[8,11] Thus, in the present study, cassava grates packed in a sample cup used during NIRS scanning provided a good scanning surface to produce spectra with increased homogeneity hence reliable predictions. This is the result of a reduction in noise by reducing light scattering from the sample during spectral acquisition.[8] Thus, sample preparation methodologies that provide for reliable scanning and spectral homogeneity such as grating in an enclosure to reduce heterogeneity in the sampling environment are crucial for achieving spectral homogeneity. Thus the reported sample selection and sample preparation methodologies can be used in estimation of DMC in cassava with minimal cost in addition to saving time.

The selection and utilization of the MPLS regression statistic under SNVD was based on the observed higher coefficients of determination for calibration, which was more than 98 g 100g^{-1} for dry matter content. This showed that oven drying method provides better estimates and thus has increased robustness as a reference method for determination of dry matter content compared to other methods.[4] For amylose content, it was realized

Journal of The Science of Food and Agriculture published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.

that the initial extraction of starch and determination of amylose using the iodometric procedures provided estimates of up to 93 g kg^{-1}. Such estimates are good enough for estimation of amylose content in cassava, especially in routine root based property selections. The selected MPLS regression statistic was also important for identifying outliers and hence providing for a dependable calibration for use in determination of dry matter contents and amylose contents of the cassava root. Because the developed models that explain up to 80 g kg^{-1} of variation (R^2c = 0.8) can be used for screening different germplasm and those that explain up to 96 g kg^{-1} can be used in quality control,[10] our developed models for dry matter content with R^2c values above 0.98 can be used for screening and quality control in all breeding stages for DMC estimation. In addition, the developed model for amylose can be routinely used for screening of cassava clones for amylose content.

The performance of SNVD as a mathematical treatment was better than other mathematical treatments, indicating the relevance of SNVD in developing equations for determination of root quality traits in cassava. It is important to note that calibrations developed in this case were based on the SNVD mainly. In addition, the performance of MPLS as a regression was more satisfactory under the SNVD for the different calibrations developed compared to the PLS.

Validation statistics across different scatter correction on wavelength selection provided a basis for the selection of the best models for prediction of root quality traits in cassava. A high coefficient of determination for prediction ranging from 95.7 to 96.3 g kg^{-1} was observed across validation sets for dry matter content, whereas the range was 77.9–93.4 g kg^{-1} for amylose content. These values were in agreement with previous recommendations[8] on the appropriate prediction parameters for NIRS applications. Higher values for RPD ranging from 4.3 to 4.7 were also observed for the selected model across the two root quality traits of amylose and dry matter content. The values were higher than those reported previously[8] for dry matter content and this could be attributed to the differences in methods used for sample preparation and dry matter determination. The above observations showed that the benchtop NIRS equipment used was useful for providing reliable prediction with respect to dry matter in cassava either for breeding or for quality control.

Overall, it was observed that the developed models are fit for use in breeding programs for estimation of root quality traits. This was further supported by the observation from the plots of reference values against the predicted values, which show a strong relationship between the reference and the predicted values. In particular, the relevance of such models will be useful for the selection of industrial based cassava varieties that are currently required by different industrial cassava consumers.

CONCLUSIONS

The present study demonstrates the capability of NIRS as a tool to accurately predict root quality traits in fresh cassava roots. Based on the datasets generated, three conclusions are apparent: (i) to increase speed and precision at the same time as reducing drudgery in determination of cassava root quality traits, NIRS based procedures that involve efficient sample preparation followed by model development need to be developed by specific cassava breeding programs; (ii) for calibration development, better performance can be achieved using the SNVD mathematical treatment and the MPLS as the regression statistic for models that predict cassava root quality; and (iii) for model selection, equations with SEP values as low as 1 g kg^{-1} and RPD values of more than 5 are robust enough for the prediction of root quality ranging from as low as 10 g kg^{-1} to as high as 45 g kg^{-1} for dry matter content and from as low as 15 g kg^{-1} to as high as 35 g kg^{-1} for amylose content. This range also includes estimation for root quality parameters in pro vitamin A clones, which are currently being improved by the Ugandan cassava breeding program. The admirable predictive ability of NIRS justifies its integration as a tool to enhance cassava breeding as a result of its prediction ability for root quality in fresh cassava roots. Generation of information on cassava traits such as dry matter content using NIRS will also revolutionize the integration of NIRS in the prediction of other traits of importance such as the consumer acceptability traits and industrial traits. The tool will also bridge the gap between phenotyping and genotyping, providing more robust selections for future traits that are presently less considered.

ACKNOWLEDGEMENTS

We are grateful for grant opportunity INV-008567 (formerly OPP1178942): Breeding RTB Products for End User Preferences (RTBfoods) to the French Agricultural Research Center for International Development (CIRAD), Montpellier, France, by the Bill & Melinda Gates Foundation 363 (BMGF): https://rtbfoods.cirad. We acknowledge the support from the various reviewers of the special issue. We acknowledge the contribution of the National Crops Resources Research Institute (NaCRRI) and National Agricultural Research Organisation (NARO) towards the harmonization of the research efforts and their coordinating role. We also recognize the contribution of the RTBfoods family as peers in the research process.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Cassava base at https://www.cassavabase.org/.

SUPPORTING INFORMATION

Supporting information may be found in the online version of this article.

REFERENCES

1 Abass A, Amaza P, Bachwenkizi B, Wanda K, Agona A and Cromme N, The impact of mechanized processing of cassava on farmers' production efficiency in Uganda. *Appl Econ Lett* **24**:102–106 (2017). https://doi.org/10.1080/13504851.2016.1167817.

2 Alamu EO, Nuwamanya E, Cornet D, Meghar K, Adesokan M, Tran T *et al.*, Near-infrared spectroscopy applications for high-throughput phenotyping for cassava and yam: A review. *Int J Food Sci Technol* **1–11**:1491–1501 (2020). https://doi.org/10.1111/ijfs.14773.

3 Esuma W, Kawuki RS, Herselman L and Labuschagne MT, Diallel analysis of provitamin A carotenoid and dry matter content in cassava (Manihot esculenta Crantz). *Breed Sci* **66**:627–635 (2016). https://doi.org/10.1270/jsbbs.15159.

4 Ikeogu UN, Davrieux F, Dufour D, Ceballos H, Egesi CN and Jannink JL, Rapid analyses of dry matter content and carotenoids in fresh cassava roots using a portable visible and near infrared spectrometer (Vis/NIRS). *PloS One* **12**:1–17 (2017). https://doi.org/10.1371/journal.pone.0188918.

5 Karlström A and Karlström A, Evaluation of the Effects of the Waxy Starch Mutation and Environment on Yield and Starch Functional Properties of Cassava (2015).

6 Lebot V, Malapa R and Jung M, Use of NIRS for the rapid prediction of total N, minerals, sugars and starch in tropical root and tuber crops. *N Z J Crop Hortic Sci* **41**:144–153 (2013). https://doi.org/10.1080/01140671.2013.798335.
7 Nuwamanya E, Acheng S, Vuzi P, Muyinza H, Matovu M, Atwijukire E *et al*., Effectiveness of pruning and waxing in reducing postharvest physiological deterioration in Uganda local cassava varieties. *Afr Crop Sci J* **27**:237 (2019). https://doi.org/10.4314/acsj.v27i2.9.
8 Parrini S, Acciaioli A, Franci O, Pugliese C and Bozzi R, Near infrared spectroscopy technology for prediction of chemical composition of natural fresh pastures. *Afr Crop Sci J* **47**:514–520 (2019). https://doi.org/10.1080/09712119.2019.1675669.
9 Sans S, Ferré J, Boqué R, Sabaté J, Casals J and Simó J, Estimating sensory properties with near-infrared spectroscopy: A tool for quality control and breeding of "Calçots" (Allium cepa L). *Agronomy* **10**(6) 828:1–11 (2020). https://doi.org/10.3390/agronomy10060828.
10 Teye E, Asare AP, Amoah RS and Tetteh JP, Determination of the dry matter content of cassava (Manihot esculenta, Crantz) tubers using specific gravity method. *ARPN J Agric Biol Sci* **6**:23–28 (2011).
11 Yi J, Sun Y, Zhu Z, Liu N and Lu J, Near-infrared reflectance spectroscopy for the prediction of chemical composition in walnut kernel. *Int J Food Prop* **20**:1633–1642 (2017). https://doi.org/10.1080/10942912.2016.1217006.

Journal of The Science of Food and Agriculture published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.