Artificial Intelligence in Agriculture 4 (2020) 219–228

Contents lists available at ScienceDirect

Artificial Intelligence in Agriculture

journal homepage: http://www.keaipublishing.com/en/journals/artificial-intelligence-in-agriculture/

Succinylation improves the slowly digestible starch fraction of cardaba banana starch. A process parameter optimization study

Babatunde Olawoye [a,*], Oladapo F. Fagbohun [b], Saka O. Gbadamosi [c], Charles T. Akanbi [c]

[a] Department of Food Science and Technology, First Technical University, KM 11, Ibadan-Lagos Expressway, Ibadan, Nigeria
[b] Department of Biomedical Engineering, First Technical University, KM 11, Ibadan-Lagos Expressway, Ibadan, Nigeria
[c] Department of Food Science and Technology, Obafemi Awolowo University Ile-Ife, Nigeria

ARTICLE INFO

Article history:
Received 23 June 2020
Received in revised form 24 September 2020
Accepted 25 September 2020
Available online 5 October 2020

Keywords:
Slowly digestible starch
Low glucose response
Artificial neural network
Partial least square

ABSTRACT

The study investigated the improvement of slowly digestible starch fraction of cardaba banana via octenyl succinic anhydride (OSA) modification process. A nonlinear (Response surface methodology [RSM] and artificial neural network [ANN]) and linear (partial least square [PLS]) models were employed and their predictability was compared. The result revealed that all the modelling techniques were accurate in predicting the experimental process. The optimized RSM values for the production of slowly digestible starch (SDS) fraction were OSA concentration of 4%, reaction time of 47.49 min, and pH of 10 with a predicted SDS value of 44.64%. Among the modelling techniques, ANN was adjudged as the predictive model for improving the SDS yield. The regression coefficient coupled with the variable important in the projection (VIP) values of the PLS model indicated that the OSA concentration was the most important factors responsible for high SDS yield. Finally, a structural comparison of the optimized starch against native starch revealed the formation of high ordered crystalline structure of the starch due to the impregnation of the modifying agent to the hydroxyl group of the cardaba banana starch.

1. Introduction

Green bananas are common herbaceous plant which belongs to the genus *Musa* (Arvanitoyannis et al., 2008). It originated from Southeast Asian, however, many species had been traded many decades ago which had led to the cultivation of bananas in more than 100 countries worldwide(Arvanitoyannis et al., 2008). Among the species of banana cultivated in Africa especially South-Western Nigeria is the Cardaba banana (Musa ABB).

Cardaba bananas (Musa ABB) are majorly utilized in food product development such as fried chips. They are a group of under-utilized bananas which are easily prone to post-harvest spoilage. Their propensity to spoilage is due to their moisture content which is relatively high as well as it metabolic and physiological features (Olawoye and Gbadamosi, 2020a). A report of Ravi and Mustaffa (2013) depicts that cardaba banana contains a high pulp starch content of 83.26%, in which the amylose content is 35.18% while the amylopectin content is 48.08%.. This high starch content makes cardaba banana an important raw material in starch production.

Starch, a major constituent in human diet, is made up of a polymer (homo) of glucose (amylose and amylopectin) and classified into three major categories, which are; the resistant starch, SDS and rapidly digestible starch (RDS) (Englyst et al., 1996). The RDS is the starch fraction that undergoes digestion within 20 min of consumption in the human gastrointestinal tract and foods/starches high in RDS are usually regarded as high glycemic food. The SDS fraction, however, on the other hand, is digested within 20–120 min of consumption in the gastrointestinal tract. The starch digestion is at a slower rate hence, leading to the slower release of glucose in the bloodstream. According to Englyst et al. (1992), SDS promotes satiety; a beneficial effect of athletes who need a prolong and constant supply of glucose into their bloodstream with a low glycemic response. Other benefits of SDS include improving the overall blood glucose control in patient suffering from diabetes mellitus, oxidative stress reduction as well as weaken the cholesterol levels of blood serum in a patient with hypolipoproteinaemia (Fagbohun et al., 2020).

Resistant starch belongs to the starch fraction that resist digestion in the small intestine but undergo fermentation by the colon microorganism to produce short-chain fatty acid (SCFA). This SCFA provides the body with additional energy as well as a high concentration of butyrate; a beneficial chemical that helps to reduce cancer of the colon. Food high in SDS and RS is associated with low glycemic index and hence, the demand for the designing of starch food with low glycemic had increased in recent times. However, starch in their native state contains a high fraction of rapidly digestible starch which in turn resulted in high glucose response and hence, the need to modify the native starch to overcome this challenge. Generally, starch modified using physical,

* Corresponding author.
E-mail address: btolawoye@gmail.com (B. Olawoye).

https://doi.org/10.1016/j.aiia.2020.09.004

chemical, enzymatic and genetic approaches had widely been studied and used in the food industry as an alternative approach to improve the slowly digestive starch fraction (Li et al., 2017).

Octenyl succinic anhydrides are categorized as a safe compound by the Department of Agriculture (FDA), thereby improving it potential for use in the synthesis of SDS fraction. In their studies, Ai et al. (2013) and Simsek et al. (2015) in their separate studies reported that starches modified using OSA reduced enzymatic digestion in comparison with their native. Also, Remya et al. (2018) and Quintero-Castaño et al. (2020), reported that succinylation improved the slow digestion property of banana starch. However, to produce starch with low digestion property, the is need to study the effect of process variables such as reaction time, the concentration of octenyl succinic anhydrides as well as substrate pH on the slow digestion property. Hence, the need to model and optimized the process variables involved in the modification of cardaba banana starch using response surface methodology, partial least square and artificial neural network.

Response surface methodology (RSM), a statistical and mathematical modelling method used for experimental design, building of model, evaluating the inter-correlation between experimental factors and to determine the optimum processing condition for maximization or minimization of experimental response. It is basically an optimization tools used after which the insignificant independent variables had been screened out. RSM a found many applications in food processing operation such as bioactive compound isolation (Olawoye and Kadiri, 2016), starch isolation (Qi et al., 2018), production of high resistant starch (Mutlu et al., 2017), and starch crystallinity (Purohit and Rao, 2017) formation of cross-linked starch.

Partial least square regression analysis (PLS) unlike RSM is a non-linear supervised classification, statistical and mathematical modelling methods used to develop and established the relationship between two experimental data sets. Although, PLS had found little application is food, however, Chen et al. (2012b) and Maulidiani et al. (2013) used it in comparison with artificial neural network and support vector regression to model the antioxidant activity in green tea and Pegaga extract, respectively.

Artificial neural network (ANN) is a machine learning tools that mimic human brain. It is made up of interconnecting neuron and hence, it has the capability to solve complex non-linear process of two experimental data sets (Agu et al., 2020). It is advantageous over other non-linear modelling tools such as RSM in that it requires no prior knowledge of the experimental data set before modelling. Many researchers had employed it along with other modelling tools in modelling their experimental process and found its' superiority over other tools. This study, therefore, aimed at improving the slowly digestible starch fraction of cardaba banana starch through the modelling and optimization of the process variables (OSA concentration, reaction time and substrate pH) using response surface methodology (RSM) as well as comparing the optimized values obtained using RSM with that of partial least square regression analysis (PLS-R) and artificial neural network (ANN).

2. Materials and methods

2.1. Cardaba banana starch isolation

Starch isolation from cardaba banana was carried out using the method described by (Olawoye and Gbadamosi, 2020b). Briefly, the banana was washed and sliced under water to prevent the browning of the banana slices. The sliced cardaba banana was commuted using Stephan machine (Stephan universal machine, Germany). Following commuting, the starch mash was mixed with water and the starch suspension was thereafter sieved using sieve (200um). The starch slurry obtained was along to stand overnight after which the supernatant was decanted. The starch residue was washed three times with distilled water. The starch obtained was dried at 45°C for 12 h. The dried starch was packed in an airtight container and stored at room temperature prior to analyses.

2.2. Production of OSA starch

The isolated cardaba banana starch (native starch) was modified using the methodology of Han and BeMiller (2007) as modified by Olawoye and Gbadamosi (2020b) to produce OSA starch. Native cardaba banana Starch (100 g, dry wt.) was dispersed in 500 mL conical flask containing distilled water with constant stirring on a magnetic stirrer. The pH of starch slurry while stirring on the magnetic stirrer was adjusted to different pH ranging between 8 and 10 as specified by the experimental design which is shown in Table 1. using 1 M NaOH and 1 M HCl. Following the adjustment of the substrate pH, different concentration of Octenyl succinic anhydride (3–5%) was added, and agitation was continued at room temperature ($\approx 26 \pm 2$ °C) while the pH was maintained using 1 M NaOH or HCl at different reaction time ranging from 30 to 60 mins. At the end each reaction time, the starch slurry was neutralized to pH 7.0 with 1 M HCl or 1 M NaOH. The modified starch was collected by centrifugation, washed three times with water and was thereafter, dried at 45 °C for 12 h to obtain OSA starch. The dried starch was kept in an airtight container until further analysis.

2.3. Experimental design

The experimental procedure for the production for improving the slowly digestible starch of cardaba banana was based on a Box-Behnken design (BBD) using three-level factor which generated 17 experimental runs. OSA concentration (3–4%), reaction time (30–60 min) as well as the pH of the (8–10) were the continuous factors while the slowly digestible starch fraction (%) was the response. The modelling using ANN and PLS-R was done by dividing the experimental data obtained from BB design into two sets: testing and training data.

2.4. Determination of OSA starch SDS content

The SDS content of the cardaba banana starch was determined following the method described by Olawoye et al. (2020) with slight modification.

Table 1
Experimental and predicted values of slowly digestible starch.

Exp. Run	Independent variable			Slowly digestible starch (%)			
	Succinate concentration (%)	Time (min)	pH	Actual	RSM predicted	ANN predicted	PLS predicted
1	3	45	10	33.95	33.84	33.77	33.82
2	4	45	9	33.06	33.26	33.26	33.27
3	3	60	9	32.65	32.56	32.65	32.54
4	4	60	10	32.39	32.59	32.39	32.59
5	4	60	8	31.88	31.88	31.88	31.89
6	4	45	9	33.26	33.26	33.35	33.27
7	4	45	9	33.36	33.26	33.26	33.27
8	4	30	10	33.27	33.27	33.27	33.27
9	4	45	9	33.27	33.26	33.26	33.27
10	5	30	9	35.89	35.99	35.89	35.99
11	5	45	8	35.78	35.89	36.03	35.87
12	4	45	9	33.36	33.26	33.26	33.27
13	5	45	10	44.68	44.59	44.68	44.59
14	3	30	9	35.24	35.35	35.24	35.35
15	5	60	9	37.47	37.36	37.47	37.36
16	4	30	8	32.84	32.63	32.77	32.64
17	3	45	8	41.11	41.20	41.11	41.20

2.5. Fourier transform infrared spectroscopy (FT-IR)

The short-range order structure of optimized OSA starch and native starch were identified using Fourier transform infrared spectrometry (FT-IR) equipped with attenuated reflectance (ATR). The starch sample was placed in the FT-IR spectroscopy with the spectrum corrected to by baseline ranging from 1200 to 800 cm^{-1} before deconvolution. After that, the band intensity ratio of 1045/1022 cm^{-1} and 1022/995 cm^{-1} was evaluated.

2.6. Modelling and optimization of SDS

2.6.1. Development of RSM model

The model was fitted to obtain the best polynomial equation using the BBD experimental data. Experimental data was analyzed using Design Expert 12.0.3 (Statease Inc., USA). Regression analysis, response surface plotting and analysis of variance were analytical process employed in the optimization of the processing condition for the isolation of the SDS fraction of the cardaba banana starch. The RSM model was tested for accuracy by comparing the predicted values with the actual values obtained for the experimental design. The data that was derived from the established optimum condition established from the developed mathematical model was employed as the validating set. The quadratic model is given in Eq. (1).

$$Y = b_0 + \sum_{i=1}^{k} b_i X_i + \sum_{i=1}^{k} b_{ii} X^2{}_i + \sum_{i<j}^{k} b_{ij} X_i X_j + e \tag{1}$$

where, Y is the response variable (SDS), $b_{0,}$ $b_{i,}$ b_{ii}, b_{ij} are the regression coefficients that shows the relationship between the responses and the processing conditions, X_i. X_i and X_j are the coded independent variables assigned to the OSA concentration, substrate pH, and reaction time. e represents the random error.

2.6.2. Development of ANN model

The ANN Modelling and optimization was carried out accroding to the methods of Olawoye et al. (2020). To predict the output variable (SDS), two transfer functions were used which were multilayer full feedforward (MFFF) and multilayer normal feedforward (MNFF), while different learning algorithms such as incremental backpropagation (IBP), quickprob (QP), genetic algorithm (GA), batch backpropagation (BBP), and Levenberg-Marquardt algorithm (LM) were used to train the ANN data sets. The ANN architecture consisted of input, output and hidden layer. The transfer function of output and hidden layer was iteratively determined through the expansion of several networks. Optimal network topology was likewise determined. A default stopping criterion of 100,000 iterations was set on the ANN during training.

2.6.3. Development of PLS-R model

Partial least square regression modelling of the experimental data was performed using Simca (v. 14.1, Umetrics, Umea, Sweden). The BBD experimental data (17 observations) were used for the PLS modelling in which the OSA concentration, reaction time and pH were used as the input variables (X), The SDS of the starch was used as the output variable (Y). The BBD experimental data sets were split randomly into two groups in which 11 observations were used for training data sets while 6 observations were used as testing data sets. Models were developed using the training data sets to find the optimal parameters while the validation of the model was done using the testing data sets. The accuracy of the experimental model was done using the R^2 and Q^2. The coefficient of determinant (R^2) indicates the goodness of fit of the experimental model while the goodness of prediction of the experimental model for the Y-variable as well as the significance of the PLS components is depicted by Q^2. The validity and degree of fit of the PLS model were determined using a permutation test. The variable importance in the regression coefficient and projection (VIP) was analyzed using the interaction between the independent (X) and dependent variable (Y). Finally, the score plot and loading of the PLS-R was evaluated for the distribution of the experimental observations.

3. Result and discussions

3.1. Regression model and statistical analysis

The Box-Behnken design (BBD) for the starch digestibility index, as well as their actual and predicted response, is shown in Table 1. The value of the slowly digestible starch varied between 31.88 and 44.68%. The experimental data were fitted and subjected to multiple regression analysis to determine and evaluate the coefficient of estimate and model equation. The model was fitted and the second-order polynomial model equation for the prediction of the response is shown in Eq. (2). As shown in Eq. (2), A, B, C represent the succinate concentration, time and pH, respectively.

$$SDS = 33.26 + 1.36A - 0.36B + 0.33C + 1.04AB + 4.01AC + 0.02BC + 4.17A^2 - 2.12B^2 + 1.45C^2 \tag{2}$$

The multiple regression analysis and ANOVA results of the model is shown in Table 2. The ANOVA result reveals that the quadratic model is significant in predicting the experimental design due to its low *p*-value (<0.0001) as well as its high Fisher test (F value) value (645.80). Among the terms, the linear term of succinate concentration and the reaction time was observed to be significant at $p < 0.01$, while the linear term of the substrate pH was only significant at $p < 0.05$. Of all the terms, only the cross-product term of reaction time and substrate pH was observed not to be significant ($p > 0.05$). In the order of level of importance, the quadratic term of OSA concentration was found to be most important of the terms due to its highest Fishers test value (2308.95) followed by the cross-product between succinate concentration and substrate pH. The lack-of-fit which is a measure of the predictability of the multiple regression model by comparing the variation around the regression model was found to be insignificant ($p > 0.05$) indicating that there was 11.9% chance that the lack of fit is due to noise. The predictability and adequacy of the model were evaluated based on the coefficient of determinant (R^2). According to Guan and Yao (2008), a model is tagged good for fit if the coefficient of determinant (R^2) which is a measure of goodness of fit is greater than 0.80. As it could be seen from the result, the coefficient of determinant, as well as the adjusted R^2, is 0.9980 and 0.9973, respectively. The closeness to one of the coefficients of determinant and adjusted R^2 is a confirmation

Table 2
Regression analysis of slowly digestible starch for succinate starch.

Source	Sum of squares	df	Mean square	F value	*p*-value Prob > F
Slowly digestible starch					
Model	184.18	9	20.46	645.80	<0.0001
A-Succinate Concentration	14.77	1	14.77	466.24	<0.0001
B-Time	1.02	1	1.02	32.11	0.0008
C-pH	0.90	1	0.90	28.25	0.0011
AB	4.34	1	4.34	136.80	<0.0001
AC	64.41	1	64.41	2032.72	<0.0001
BC	1.33×10^{-3}	1	1.33×10^{-3}	0.042	0.8433
A^2	73.17	1	73.17	2308.95	<0.0001
B^2	18.87	1	18.87	595.41	<0.0001
C^2	8.87	1	8.87	279.76	<0.0001
Residual	0.22	7	0.032		
Lack of Fit	0.16	3	0.054	3.71	0.1190
Pure Error	0.059	4	0.015		
Cor Total	184.40	16			
R^2	0.998				
Adjusted R^2	0.9973				

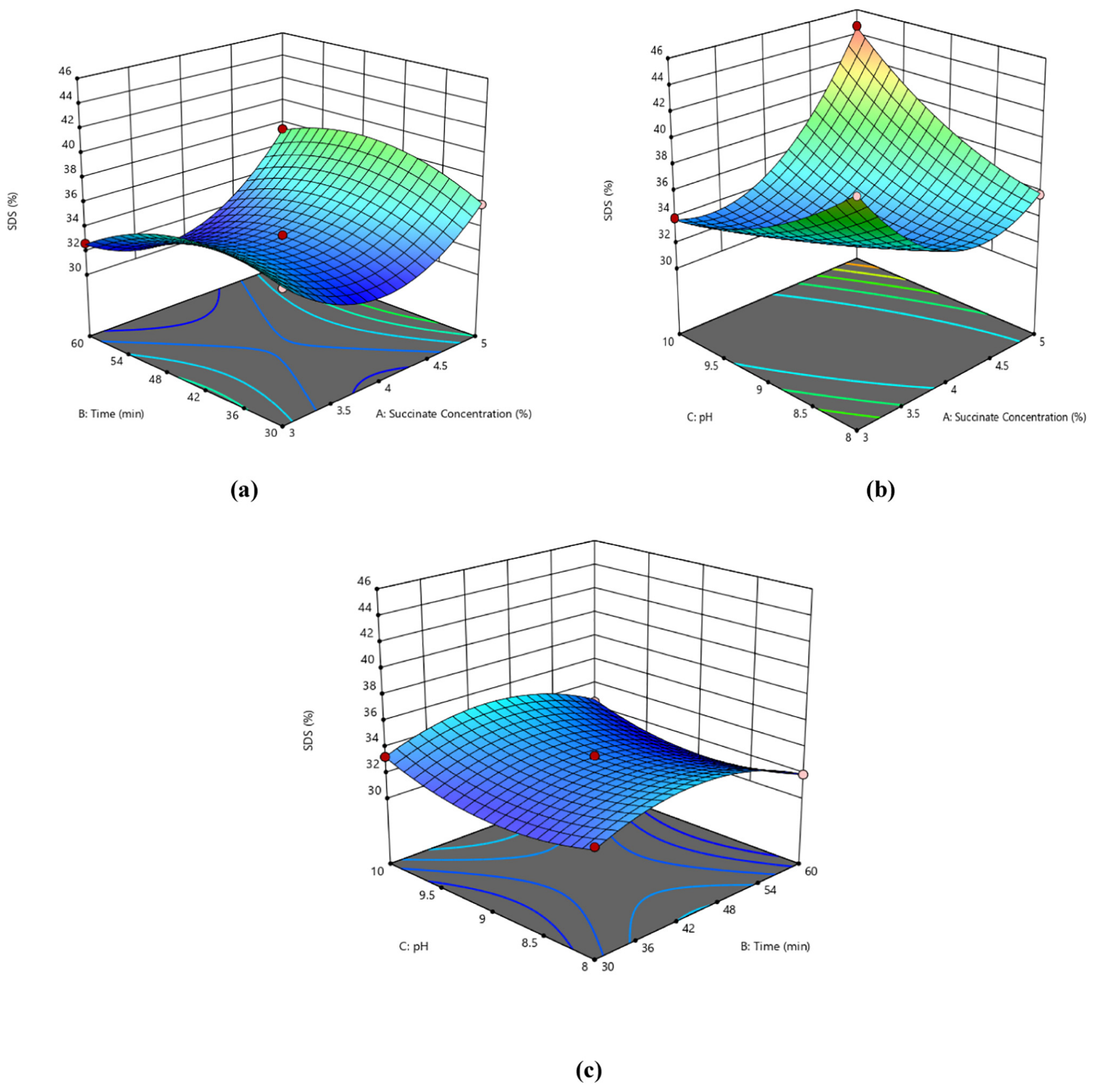


Fig. 1. (a–c): Effects of independent variables on slowly digestible starch using RSM model. (d–f): Effects of independent variables on slowly digestible starch using PLS model.

that the experimental model can predict adequately the experimental process.

Another quality parameter to determine the accuracy of the quadratic model to adequately predict the experimental data is the coefficient of variation which measures the ratio of the standard deviation of estimate to the mean values of the observed response. The low value of the CV (0.51%) denotes the responsibility and suitability of the experimental model.. The synergetic effect and the importance of the experimental model terms on the dependent variable of the slowly digestible starch was evaluated using the regression model and presented in Eq. (2).

Aside reaction time, all linear terms had negative coefficient of estimate as reflect in the equation. These resulted in a negative coefficient of estimate between the model terms and the subsequent slowly digestible starch. This implies that an increase in these terms will result in a decrease in the slowly digestible starch fraction of the succinic anhydride modified cardaba banana starch. The cross-product terms, the quadratic term of succinate concentration and pH had a positive coefficient of estimate and hence positive relationship between the terms and the corresponding slowly digestible starch. This findings also commensurate with the report of Betancur-Ancona et al. (2002) who reported a positive effect of pH variation on Jack bean.

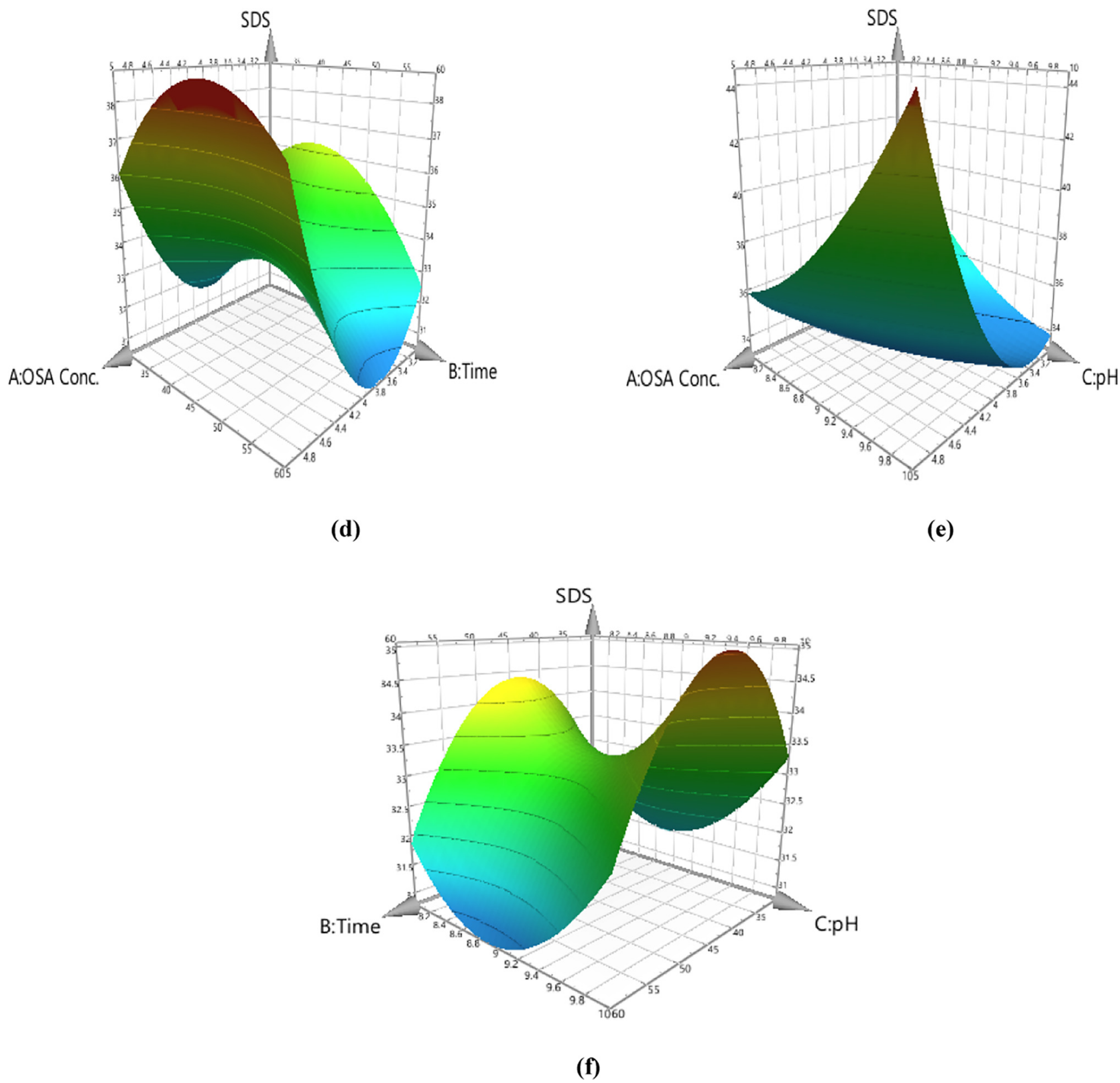


Fig. 1 (*continued*).

Table 3
Effect of different transfer functions on R^2 and RMSE in the determination of slowly digestible starch.

Model	Learning algorithms	Connections types	Output layer transfer function	Input layer transfer function	Training		Testing	
					[f]R^2	[g]RMSE	R^2	RMSE
3-5-1	BBP[a]	MFFF[b]	Hyperbolic Tangent	Hyperbolic Tangent	0.9924	0.193	0.9921	1.78
3-6-1	IBP[c]	MFFF	Hyperbolic Tangent	Sigmoid	0.9948	0.105	0.9948	1.94
3-7-1	IBP	MNFF[d]	Sigmoid	Hyperbolic Tangent	0.9949	0.172	0.9949	2.41
3-7-1	QP[e]	MFFF	Sigmoid	Hyperbolic Tangent	0.9949	0.134	0.9949	3.04
3-9-1	IBP	MFFF	Sigmoid	Sigmoid	0.9989	0.019	0.9991	1.54
3-9-1	IBP	MNFF	Sigmoid	Sigmoid	0.9945	0.073	0.9946	2.79

[a] Batch Back Propagation.
[b] Multilayer Full Feed Forward.
[c] Incremental back propagation.
[d] Multilayer normal Feed Forward.
[e] Quick Propagation.
[f] Coefficient of determination.
[g] Root mean square deviation.

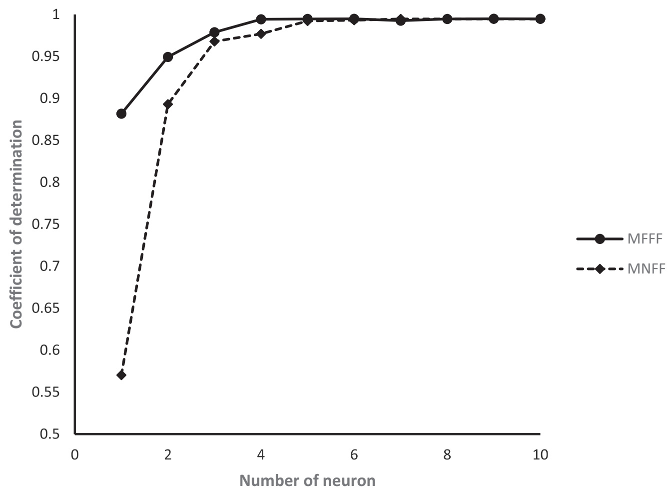


Fig. 2. Optimum Numbers of Hidden Neurons in Determining Slowly Digestible Starch by Comparing IBP-MFFF (circle) and IBP-MNFF (square). *MFFF: Multilayer Full Feed Forward; MNFF: Multilayer normal Feed Forward.

3.2. Interactions of independent variables and process optimization

The 3-D surface plot which shows the impact of the independent variables impacted on the percentage SDS is shown in Fig. 1(a–c). The simultaneous effect of the succinate concentration and reaction time while holding the pH constant is presented in Fig. 1(a). The figure revealed that the initial increase in the succinate concentration resulted in a decrease in the slowly digestible starch yield, however, increasing the concentration above 4.47% resulted in a slight increase in the SDS yield. The initial decrease in SDS value was observed at low reaction time and this could be due to the inability of the functional group within the modify agent (succinic anhydride) to fuse to the hydroxyl group of the starch and hence, the ability of the hydrolyzing enzymes to rapidly digest the starch. The single effect of the reaction time shows that the increase in the reaction time up 46.2 min brought about a slight reduction in the yield of the slowly digestible starch, however, increasing the time further led to a spontaneous decrease in the percentage slowly digestible starch. The increase in the SDS fraction of the starch as the reaction time increases can be attributed to the replacement of the hydroxyl group present in the starch molecule by the functional group of the modifying agents (OSA) (Altuna et al., 2018). For maximum slowly digestible starch yield, the succinic anhydride concentration needs to be at maximum coupled a time of reaction of 48 ± 1 min.

Fig. 1(b) shows how varying concentration of succinic anhydride and pH impacted the percentage of SDS of starches that were modified using the succinylation process while keeping reaction time constant. The result revealed a response surface with a minimum point. The effect of pH on the SDS yield shows that an increase in the pH resulted in a reduction in the yield of the SDS while an increase in the succinate concentration resulted in an initial decrease in the SDS yield. However, a succinate concentration of 4.56% and above slightly led to an increase in the slowly digestible starch. The effect of pH and time of succinylation reveals that the pH of the substrate had a linear effect on the percentage slowly digestible (Fig. 1c). As shown in the figure, the pH at a lower time of reaction had little or no effect on the SDS yield, however, a slight increase in the reaction time led to increase in the percentage SDS at maximum pH. The increase is SDS of the starch at maximum pH of the substrate could be due to the increase in the reaction efficiency as well as the degree of substitution thereby resulting in sufficient activation of the hydroxyl group for the nucleophilic attack of the anhydride moieties (Segura-Campos et al., 2008). Optimal condition values for the production of slowly digestible starch from cardaba banana was established using RSM were OSA concentration of 4%, reaction time of 47.49 min, and pH of 10 which yielded a predicted SDS value of 44.64%, which was validated experimentally as 44.68% SDS.

3.3. SDS modelling using artificial neural network

Artificial neural networks (ANN) are computer-based program which mimics the human brain by modelling the interaction between independent variables and their responses. In the neural network modelling for this study, the input variables were the succinate concentration, pH, and time, while the output variables majorly the slowly digestible starch of the modified cardaba banana starch. Olawoye et al. (2020) in their previous study, reported the occurrence of many learning algorithms hence the difficulty in the pre-selection of the algorithm to be used in advance. Owing to this, several learning algorithms were used to detect the algorithm that best fit the experimental data used in the artificial neural network modelling. This result obtained is presented in Table 3 and the result revealed that the increment backpropagation (IBP) algorithm best fit the experimental data. This was used to build a multilayer full feed-forward neural architecture.

Moreover, the neural network is affected by the type of transfer functions used. This affect both the learning rate as well as the performance of the network. With this respect, several transfer functions were applied to the hidden and output layers and it was observed that the sigmoid-sigmoid transfer functions bring about the acceptable model. In a network architecture, the number of hidden layers to be used in the network topology is important and hence there is a need to carefully select the number of hidden layers. To achieved this, a trial by error method is used in the selection of the number of hidden layers and the result which is based on the best goodness of fit is presented in Fig. 2. Therefore, 3-9-1 network architecture (Fig. 3) was used which indicates three input variables, nine hidden layers and one output layer.

Many input and output transfer functions were used; however, the sigmoid-sigmoid transfer function was chosen in that it gave the highest coefficient of determinant and lowest root mean square of error. The result of the ANN modelling as shown in Table 3 revealed the value of 0.9988 and 0.019 were obtained for the coefficient of determinant (R^2) and RMSE, respectively for the training data set while the result of the testing revealed that the value of the coefficient of determinant (R^2) and RMSE were 0.9991 and 1.54 respectively. The result obtained revealed that the artificial neural network can be used to predict the relationship between the independent (input) variables and slowly digestible starch yield adequately. Optimal condition values for the production of SDS from cardaba banana was established using ANN and the optimal conditions were OSA concentration of 3.69%, a reaction time of 43.75 min, and pH of 9.8 which yield a slowly digestible starch of 44.84%.

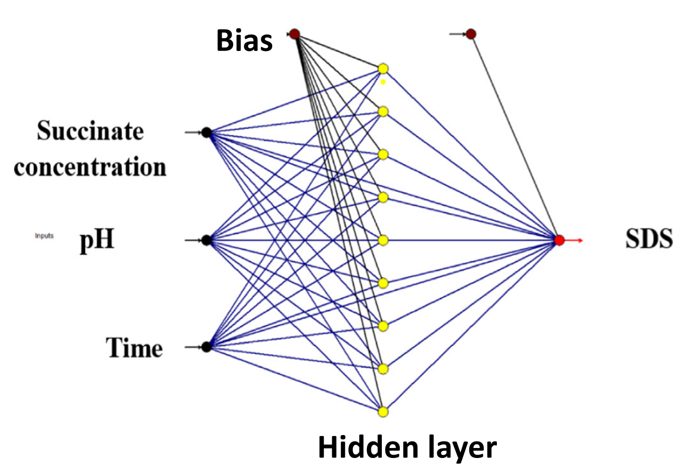
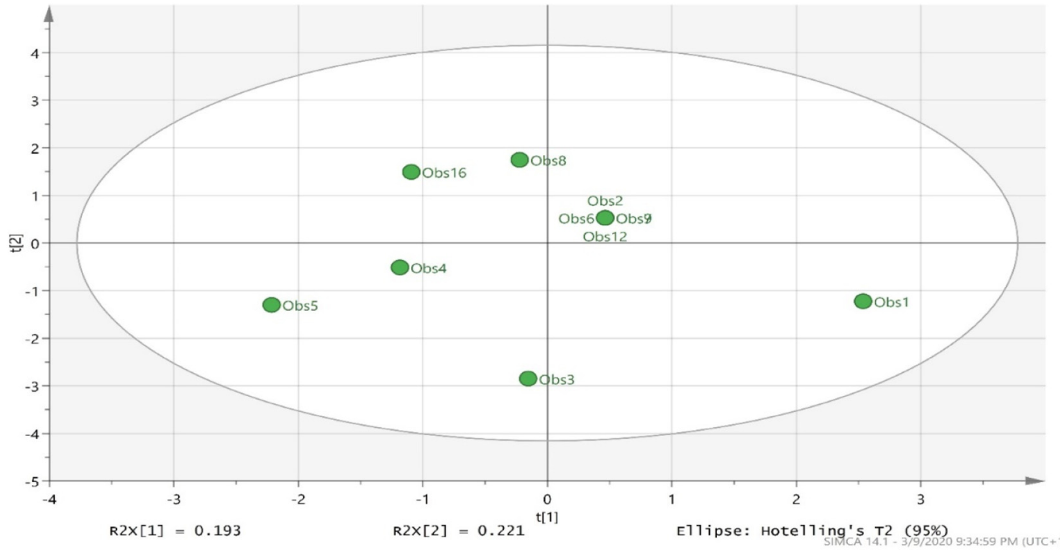
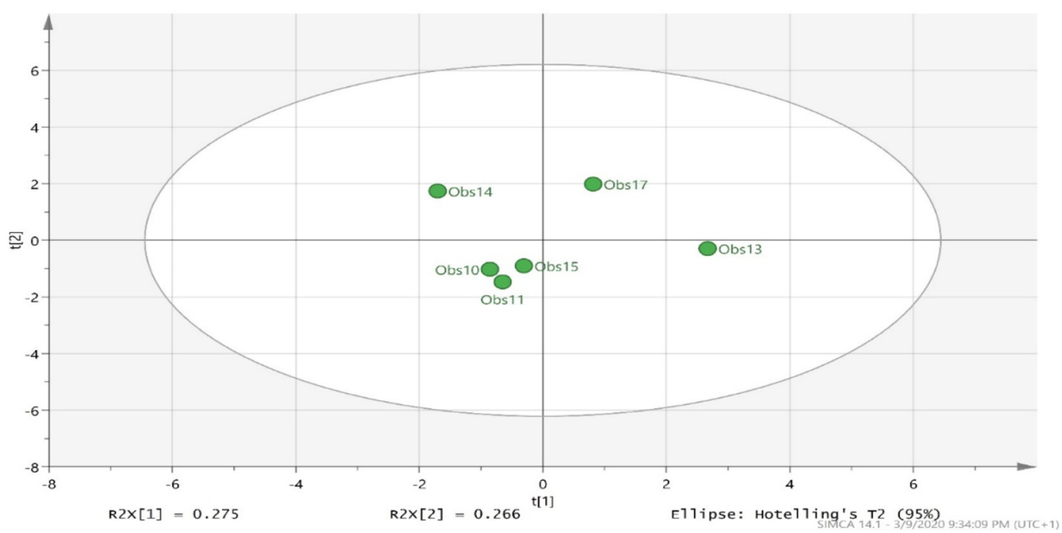


Fig. 3. Neural network topology of SDS.



(a)



(b)

Fig. 4. PLS scatter score plot (t[1] vs t[2]) for (a) training (b) testing.

3.4. SDS modelling using PLS modelling

The result of the PLS analysis revealed that both the first and component contributed 33.33% variance in the experimental data sets, therefore, the two components amounted to 66.67% of the total variation of the experimental data sets. The result of the score scatter plot of t[1] against t[2] for the training and testing data sets from the PLS model is shown in Fig. 4(a & b). The relationship between the independent variables (OSA concentration, reaction time and pH) and the dependent variable (slowly digestible starch) was evaluated using the partial least square (PLS). The accuracy of the PLS model in describing the data sets of the experiment was determined using the coefficient of determinant (R^2), the goodness of prediction Q^2 and root mean square of error (RMSE).

According to Olawoye et al. (2017), a model is said to be accurate in predicting the experimental data sets if the R^2 is higher than 0.8 and the RMSE is as low as possible. The high R^2 (0.999) coupled with low RMSE (0.1263) of the PLS model indicates the goodness of fit of the model, the goodness of predictability of the model for the dependent (Y) variables (Q^2) is 0.599. According to Eriksson et al. (2006), a PLS model with Q^2 above 0.5 indicates good predictability. Table 4, reveals the regression

Table 4
VIP rank, VIP values and regression coefficient of the model terms on PLS model.

Model term	VIP rank	VIP[2]	Var ID	CoeffSC[2]
A^2	1	1.889	1.004	0.630
AC	2	1.783	1.985	0.593
A	3	0.853	2.979	0.284
B^2	4	0.835	3.984	−0.320
C^2	5	0.718	5.000	0.220
AB	6	0.463	5.990	0.154
B	7	0.224	6.998	−0.074
C	8	0.210	8.005	0.070
BC	9	0.005	8.987	0.003

*A: OSA concentration; B: Reaction time; C: pH.

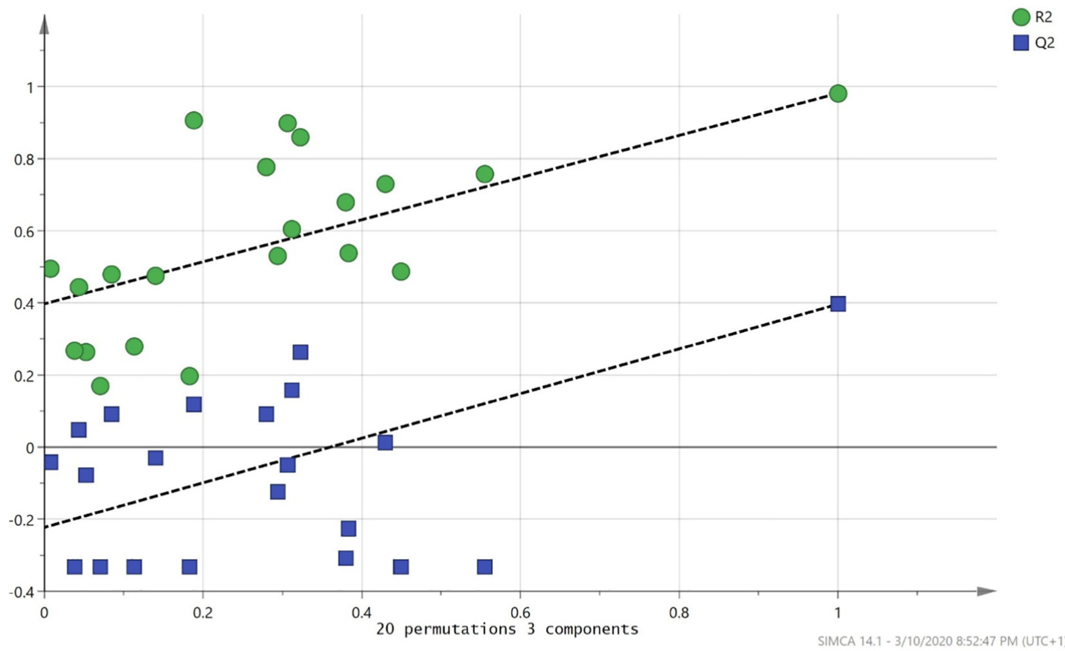


Fig. 5. Permutation test for PLS model. *R^2: Coefficient of determinant; Q^2: Goodness of prediction.

coefficient of the PLS model. As it can be seen from the result, all the model terms had a positive regression coefficient with the slowly digestible starch except for the linear and quadratic term of the reaction time which had a negative regression coefficient. The positive regression coefficient indicates a positive synergy between the model term and the Y variable and vice versa. Owing to the regression coefficient, the prediction regression model is shown in Eq. (3).

The ranking, as well as the variable importance of projection (VIP) of the model terms, is presented in Table 4. According to Cho et al. (2009), a VIP value above one (1) is termed significant while a value below 0.5 is regarded as insignificant ones. As shown in Table 4, the quadratic term of the OSA concentration was ranked first with the highest VIP (1.889) and hence, the most significant model terms. However, the interaction terms of OSA concentration and reaction time, reaction time and pH, as well as the linear terms of reaction time and pH in which their VIP values fell below 0.5, were not significant in predicting the PLS model. A 3D surface plot to show the interactions between the independent variables is shown in Fig. 2(d–f). The result obtained from Fig. 2(d–f) shows a similar trend to that of RSM. Fig. 5. shows the permutation test for the third component of the PLS model. According to the result, the Y-axis intercept of R^2 and Q^2 is 0.297 and −0.223, respectively. In the report of Eriksson et al. (2006), a PLS model is termed valid and obey thumb's rule if the Y-axisintercept of R^2 is less than 0.3 and Q^2 is less than 0.05. Also, the farther the R-line is to the horizontal line, the more valid the PLS model is which is an affirmation of the result obtained for the experimental model.

$$SDS = 10.28 + 0.28A - 0.07B - 0.06C + 0.15AB + 0.59AC + 0.003BC + 0.63A^2 - 0.32B^2 + 0.22C^2 \tag{3}$$

3.5. Comparison between PLS and ANN model

Comparison between PLS and ANN analysis was done by evaluating the conformity of the predicted and actual percentage slowly digestible starch. Fig. 6(a–d) reveals the relationship between the training data sets and testing data sets of the ANN and PLS analysis. It could be seen from the result that both the training and testing data sets of the ANN and PLS models exhibit the same trends. For the PLS model, the R^2 and RMSEE values for the training and testing data sets were 0.981; 0.092 and 0.997; 0.268, respectively. The high coefficient of determinant (R^2) for both the testing and training data set is an indication of the predictability and accuracy of the PLS model. The RMSEE value of the training data set is low which indicates the fit of the observations of the PLS model. Also, the R^2 of the training and testing data set of the ANN model were 0.9945 and 0.9999, respectively. The result obtained from both the ANN and PLS model shows there was an increase in the coefficient of determinant (R^2) of the testing data set over the training data set which indicates good predictability of the models. This observation, however, contradicts the findings of Maulidiani et al. (2012) and Chen et al. (2012a) who reported a decrease in the R^2 from training to testing data sets. This reduction was reported to be due to large numbers of input in the PLS and ANN models.

3.6. Fourier transform infrared spectroscopy (FT-IR)

The external region of the starch granule can be studied using the ATR FT-IR spectrum. According to Warren et al. (2016), the band intensity at 1100–900 cm^{-1} had shown to be sensitive to changes in the structure of starch especially the spectrum band at 1000, 1022 and 1047 cm^{-1}. starch samples with more amorphous region had higher 1022 cm^{-1} band while there exists a definite crystallinity for the band at 1000 and 1047 cm^{-1} for crystalline starch samples. These properties of starch have led to the measurement of the short-range ordered molecular structure through the adoption of the band ratios at 1022:1000 cm^{-1} and 1047:1022 cm^{-1}. Fig. 7 shows the short-range FT-IR spectra of both the native and optimized succinate starch, the figure revealed that both FT-IR spectra were similar. As reported by Warren et al. (2016), starches with similar FT-IR spectra exhibit the same crystalline structure. From Fig. 7, the FT-IR ratios of 1047/1022 and 1022/995 cm^{-1} was calculated and the values were 0.9062 and 1.087, respectively for native starch while the values obtained for optimized succinate starch were 1.0903 and 1.004, respectively. According to Warren et al. (2016), the 1022:1000 cm^{-1} intensity ratio is sensitive to hydration resulting in liquid-crystalline polymeric structure of the native starch and hence the susceptibility of the starch to enzymatic hydrolysis. The 1047/1022 intensity band ratio was found to increase due

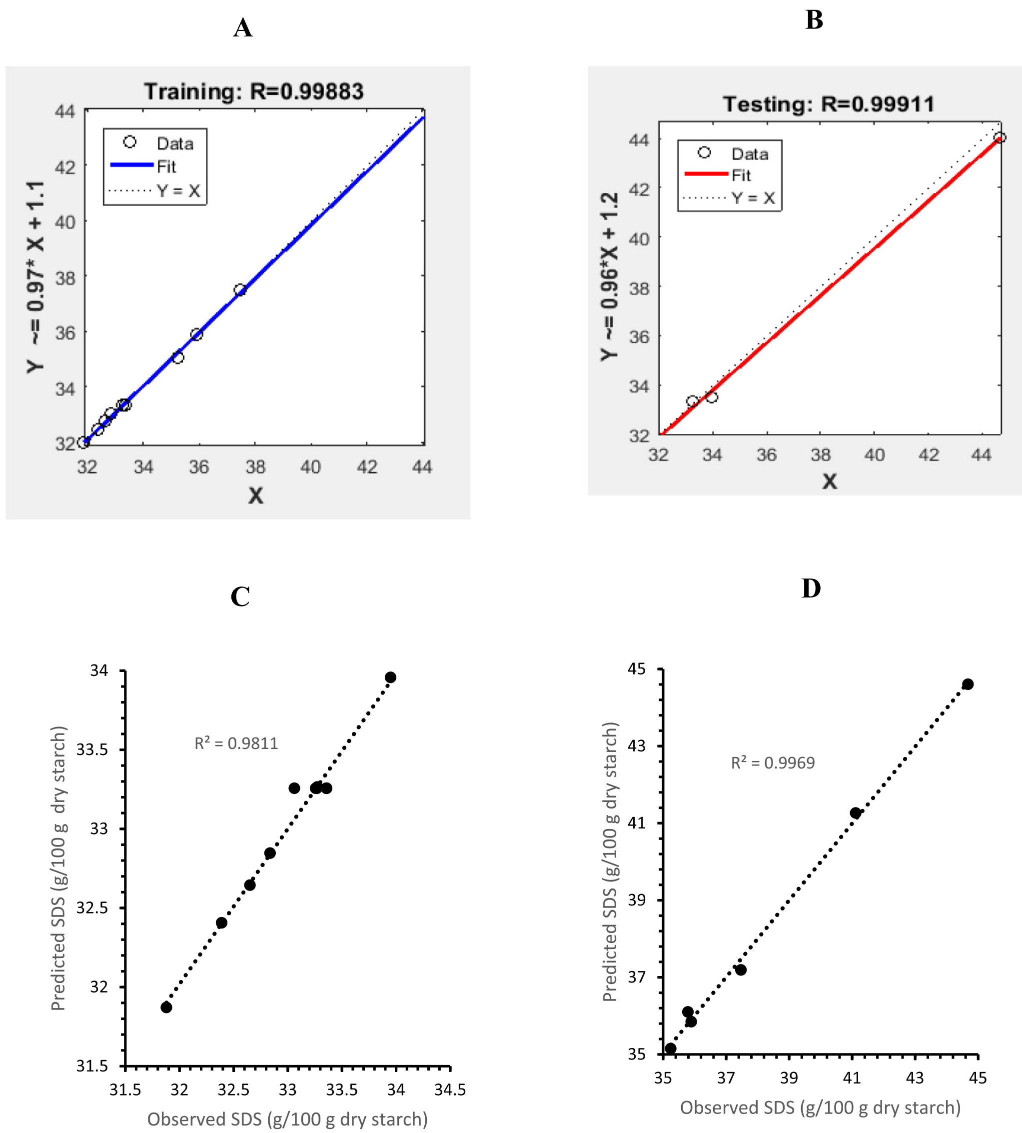


Fig. 6. Relationship between observed and predicted SDS for (a) ANN training (b) ANN testing (c) PLS training (d) PLS testing.

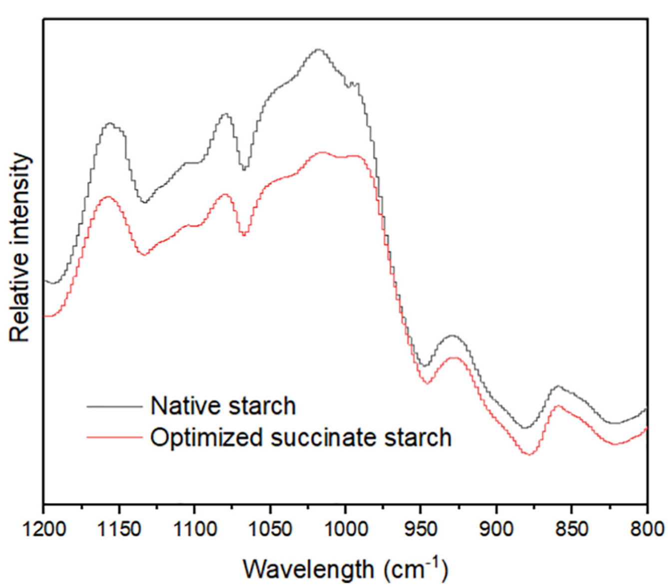


Fig. 7. Short-range Fourier transform infrared spectroscopy.

to succinylation and could be attributed to the ordered helical structure formation leading to the crystallinity of the succinate starch.

4. Conclusion

Cardaba banana starch was extracted and esterified successfully to produce slowly digestible starch using octenyl succinic anhydride. The esterification process yields a slowly digestible starch which ranged between 31.88 and 44.68%. Modelling the SDS using RSM, ANN, and PLS revealed that all the models were significant in predicting accurately the experimental process. Optimal condition values for the production of slowly digestible starch from cardaba banana was established using RSM were OSA concentration of 4%, reaction time of 47.49 min, and pH of 10 which yielded a predicted SDS value of 44.644%, which was validated experimentally as 44.68% SDS. The optimal conditions values for the production of SDS using ANN were OSA concentration of 3.69%, reaction time of 43.75 min, and pH of 9.8 which yield a slowly digestible starch of 44.84%. A comparison between RSM, ANN, and PLS in modelling the SDS fractions revealed that ANN, followed by PLS demonstrated better predictability over RSM. This study revealed that cardaba banana could serve as an alternative and cheap source of slowly digestible starch.

Declaration of Competing Interest

The authors declare that he has no conflict of interest.

References

Agu, C.M., Menkiti, M.C., Ekwe, E.B., Agulanna, A.C., 2020. Modeling and optimization of *Terminalia catappa* L. kernel oil extraction using response surface methodology and artificial neural network. Artific. Intell. Agric. 4, 1–11.

Ai, Y., Nelson, B., Birt, D.F., Jane, J.L., 2013. In vitro and in vivo digestion of octenyl succinic starch. Carbohydr. Polym. 98, 1266–1271.

Altuna, L., Herrera, M.L., Foresti, M.L., 2018. Synthesis and characterization of octenyl succinic anhydride modified starches for food applications. A review of recent literature. Food Hydrocoll. 80, 97–110.

Arvanitoyannis, I.S., Mavromatis, A.G., Grammatikaki-Avgeli, G., Sakellariou, M., 2008. Banana: cultivars, biotechnological approaches and genetic transformation. Int. J. Food Sci. Technol. 43, 1871–1879.

Betancur-Ancona, D., García-Cervera, E., Cañizares-Hernández, E., Chel-Guerrero, L., 2002. Chemical modification of Jack bean (*Canavalia ensiformis*) starch by Succinylation. Starch Stärke 54, 540–546.

Chen, Q., Guo, Z., Zhao, J., Ouyang, Q., 2012b. Comparisons of different regressions tools in measurement of antioxidant activity in green tea using near infrared spectroscopy. J. Pharm. Biomed. Anal. 60, 92–97.

Chen, Q.-F., Chen, X.-Z., Li, G.-Y., Wang, C., Zhang, G.-L., 2012a. Two new 2-phenylbenzofurans from the bark of Styrax perkinsiae. Chin. J. Nat. Med. 10, 92–97.

Cho, S.K., Yang, S.-O., Kim, S.-H., Kim, H., Ko, J.S., Riu, K.Z., Lee, H.-Y., Choi, H.-K., 2009. Classification and prediction of free-radical scavenging activities of dangyuja (*Citrus grandis* Osbeck) fruit extracts using 1H NMR spectroscopy and multivariate statistical analysis. J. Pharm. Biomed. Anal. 49, 567–571.

Englyst, H.N., Kingman, S., Cummings, J., 1992. Classification and measurement of nutritionally important starch fractions. Eur. J. Clin. Nutr. 46, S33–S50.

Englyst, H.N., Veenstra, J., Hudson, G.J., 1996. Measurement of rapidly available glucose (RAG) in plant foods: a potential in vitro predictor of the glycaemic response. Br. J. Nutr. 75, 327–337.

Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J., Wikström, C., Wold, S., 2006. Multi-and Megavariate Data Analysis (Umetrics Sweden).

Fagbohun, O.F., Olawoye, B., Ademakinwa, A.N., Jolayemi, K.A., Msagati, T.A.M., 2020. Metabolome modulatory effects of *Kigelia africana* (Lam.) Benth. fruit extracts on oxidative stress, hyperlipidaemic biomarkers in STZ-induced diabetic rats and antidiabetic effects in 3T3 L1 adipocytes. J. Pharm. Pharmacol.

Guan, X., Yao, H., 2008. Optimization of Viscozyme L-assisted extraction of oat bran protein using response surface methodology. Food Chem. 106, 345–351.

Han, J.-A., BeMiller, J.N., 2007. Preparation and physical characteristics of slowly digesting modified food starches. Carbohydr. Polym. 67, 366–374.

Li, N., Niu, M., Zhang, B., Zhao, S., Xiong, S., Xie, F., 2017. Effects of concurrent ball milling and octenyl succinylation on structure and physicochemical properties of starch. Carbohydr. Polym. 155, 109–116.

Maulidiani F., Abas, Khatib, A., Shitan, M., Shaari, K., Lajis, N.H., 2013. Comparison of partial least squares and artificial neural network for the prediction of antioxidant activity in extract of Pegaga (Centella) varieties from 1H nuclear magnetic resonance spectroscopy. Food Res. Int. 54, 852–860.

Maulidiani, H., Khatib, A., Shaari, K., Abas, F., Shitan, M., Kneer, R., Neto, V., Lajis, N.H., 2012. Discrimination of three Pegaga (Centella) varieties and determination of growth-lighting effects on metabolites content based on the Chemometry of 1H nuclear magnetic resonance spectroscopy. J. Agric. Food Chem. 60, 410–417.

Mutlu, S., Kahraman, K., Öztürk, S., 2017. Optimization of resistant starch formation from high amylose corn starch by microwave irradiation treatments and characterization of starch preparations. Int. J. Biol. Macromol. 95, 635–642.

Olawoye, B., Gbadamosi, S., 2020a. Digestion kinetics of native and modified cardaba banana starch: a biphasic approach. Int. J. Biol. Macromol. 154, 31–38.

Olawoye, B., Gbadamosi, S.O., 2020. Sensory profiling and mapping of gluten-free cookies made from blends Cardaba banana flour and starch. J. Food Process. Preserv. 44 (9), e14643.

Olawoye, B., Kadiri, O., 2016. Optimization and response surface modelling of antioxidant activities of Amaranthus virides seed flour extract. Ann. J. Food Sci. Technol. 17, 114–123.

Olawoye, B., Kadiri, O., Babalola, T., 2017. Modelling of thin-layer drying characteristic of unripe Cardaba banana (Musa ABB) slices. Cogent Food Agric. 3, 1290013.

Olawoye, B., Gbadamosi, S.O., Otemuyiwa, I.O., Akanbi, C.T., 2020. Improving the resistant starch in succinate anhydride-modified cardaba banana starch: a chemometrics approach. J. Food Process. Preserv. 44 (9), e14686.

Purohit, S.R., Rao, P.S., 2017. Optimization of paddy parboiling process for higher starch crystallinity by response surface methodology. Int. J. Biol. Macromol. 104, 1091–1098.

Qi, Y., Du, F., Jiang, Z., Qiu, B., Guan, Q., Liu, J., Xu, T., 2018. Optimization of starch isolation from red sorghum using response surface methodology. LWT 91, 242–248.

Quintero-Castaño, V.D., Castellanos-Galeano, F.J., Álvarez-Barreto, C.I., Lucas-Aguirre, J.C., Bello-Pérez, L.A., Rodríguez-Garcia, M.E., 2020. Starch from two unripe plantains and esterified with octenyl succinic anhydride (OSA): partial characterization. Food Chem. 315, 126241.

Ravi, I., Mustaffa, M.M., 2013. Starch and amylose variability in banana cultivars. Indian J. Plant Physiol. 18, 83–87.

Remya, R., Jyothi, A.N., Sreekumar, J., 2018. Morphological, structural and digestibility properties of RS4 enriched octenyl succinylated sweet potato, banana and lentil starches. Food Hydrocoll. 82, 219–229.

Segura-Campos, M., Chel-Guerrero, L., Betancur-Ancona, D., 2008. Synthesis and partial characterization of octenylsuccinic starch from *Phaseolus lunatus*. Food Hydrocoll. 22, 1467–1474.

Simsek, S., Ovando-Martinez, M., Marefati, A., Sjöö, M., Rayner, M., 2015. Chemical composition, digestibility and emulsification properties of octenyl succinic esters of various starches. Food Res. Int. 75, 41–49.

Warren, F.J., Gidley, M.J., Flanagan, B.M., 2016. Infrared spectroscopy as a tool to characterise starch ordered structure—a joint FTIR–ATR, NMR, XRD and DSC study. Carbohydr. Polym. 139, 35–42.