International Food Research Journal 24(6): 2347-2352 (December 2017)
Journal homepage: *http://www.ifrj.upm.edu.my*

Prediction of papaw moisture ratio during hot air-drying: GMDH vs. mathematical modeling

Yousefi, A. R. and *Ghasemian, N.

Department of Chemical Engineering, University of Bonab, PO Box 55517-61167, Bonab, Iran

Article history

Received: 25 July 2016
Received in revised form: 25 August 2016
Accepted: 4 October 2016

Keywords

Drying
GMDH
Modeling
Papaw
Thin layer

Abstract

The main objective of this work was prediction of the moisture content of papaw during hot air-drying in a cabinet dryer using both mathematical and GMDH (group method of data handling). The influence of air temperatures (40, 50 and 60°C) and fruit slices thickness (3, 5, 7 mm) on moisture ratio were investigated. Exactly 50% of the data points were used for training and 50% for testing. Furthermore, eight well-known empirical models were subjected to experimental data for modeling of the drying process. The determination coefficient (R^2) and root mean square error *(RMSE)* computed for the GMDH model were 0.9960 and 0.0220. Among the empirical models considered, the Two terms model, was found to be more suitable for predicting drying of papaw fruit slices with the values of R^2=0.9974 and RMSE=0.0123. Thus, it was deduced that the estimation of moisture content of papaw fruit could be modelled by GMDH method as good as the best empirical models.

Introduction

Papaw (*Carica papaw* L.) is a well-known fruit that due to its agreeable flavor and also many pharmacological properties widely consumed (De Oliveira and Vitória, 2011). This fruit has been categorized as a top ranking fruit because of high level of various nutrient compounds such as minerals, vitamins, carotenoids etc. (Liebman, 1992). Based on the FAO report in 2010, the papaw has been ranked third with 11.2 million tons or 15.36 percent of the total tropical fruit production. More food stuffs containing high amount of water which has a direct influence on many physico-chemical and biological changes. Moisture content has a pronounced influence on the quality of food stuffs. Drying is of the most effective operations to diminish the spoilage of agricultural products by reducing the moisture content (Izadifar and Mowla, 2003).

To characterize the parameters involve in drying process, the thin-layer drying procedure was found to be the most feasible tool (Aghdam *et al*., 2015). Different types of models have been used by several researchers to predict the moisture content/drying rate of food materials which finally led to different expression for the prediction (Kingsly and Singh, 2007; Wang *et al*., 2007; Yousefi *et al*., 2013a; Yousefi *et al*., 2013b; Dinani *et al*., 2014; Koukouch *et al*., 2015). Most of these models are mathematical ones which classified to theoretical, semi-theoretical and empirical models (Demirtas *et al*., 1998; Midilli *et al*., 2002). Lately, a new predictive method based on artificial neural networks systems (ANNs) has been used to model the drying process of different food and agricultural products like potato and green pea (Kamiński *et al*., 1998), Echinacea angustifolia (Erenturk *et al*., 2004), grain (Liu *et al*., 2007), tomato (Movagharnejad and Nikzad, 2007), shelled corn (Momenzadeh *et al*., 2011) and pomegranate arils (Nikbakht *et al*., 2014). The ANNs are mostly considered as nonlinear and highly flexible universal approximators (Powell, 1987; Park and Sandberg, 1991). Nonetheless, its main drawback is that the detected dependencies are concealed behind neural network structure (Nariman-Zadeh and Jamali, 2007). Contrarily, the group method of data handling (GMDH) is applied to develop a model which is hidden in the empirical data (Ivakhnenko, 1971). The GMDH method was originated by Ivakhneko in 1966 and it has been improved and evolved over the past 40 years. The GMDH algorithm connects the inputs to outputs with high order polynomial networks which are mainly feed-forward and multi-layered neural networks (Onwubolu, 2009). In this approach, the nodes are hidden units and the activation polynomial coefficients are weights which are estimated by ordinary least square regression (Onwubolu, 2009; Ghanadzadeh *et al*., 2012). In recent years, however,

*Corresponding author.
Email: *n.ghasemian@gmail.com*

the use of such self-organized networks has led to successful application of the GMDH-type algorithm in a wide range of areas in engineering and science (Ahmadi *et al*., 2007; Pazuki and Kakhki, 2013; Abdolrahimi *et al*., 2014; Atashrouz *et al*., 2015; Najafzadeh, 2015).

Based on the literature review, no specific study was found to be associated with the estimation of moisture content of papaw fruit using GMDH. Therefore, the purpose of this work was to undertake a study to investigate the thin-layer drying process of papaw slices in a cabinet drier and modeling of the experimental data using group method of data handling (GMDH) to estimate the moisture content of papaw fruit. In addition to GMDH, eight well-known thin-layer empirical models were employed for the estimation, and finally the estimation quality of both types of models was evaluated and compared.

Materials and Methods

Experimental study

The papaw fruits experimented in this study were purchased from a local market in the Bahookalat region, Iran. After transferring to lab, the fruits stored at 4 ± 1°C before subjecting to any specific process. After that, the fruits were washed and peeled with a sharp knife and then were cut into three thicknesses of 3, 5 and 7 mm. The slices obtained were subjected to hot-air in a cabinet dryer (Model JE10 TECH, F-02G, South Korea) to investigate their drying kinetics. It should be noted that the absolute humidity and the hot-air flow ratio applied for all drying temperatures were 0.6±0.02 g/kg dry air and 1±0.1 m/s, respectively. The initial moisture content attained for the slices (using a laboratory oven dryer at 105°C) was 84.48%±0.05% (w. b.). In each run 3 batches (each batch containing 5 g sample) of thin layer samples were separately placed on the dryer. A programmable balance software recorded the weight of samples at 5-min intervals until the moisture content of them reached to 15±0.02% (w. b.) in the final product. The capacity of dryer was approximately 5-6 kg and all of the experiments were performed in triplicate. Three temperature levels of 40, 50 and 60°C were used for drying process of the samples. The amounts of moisture ratio (MR) which obtained from the Eq. (1) were plotted vs. drying time for various conditions. MR is defined by the equation:

$$MR = \frac{M - M_e}{M_0 - M_e} \qquad (1)$$

Where M and M_0 are the moisture content at any drying time and the initial moisture content, respectively. This equation can be simplified to M/M_0, because the equilibrium moisture content value (M_e) is relatively small compared with that of M or M_0 (Akgun and Doymaz, 2005).

Group method of data handling (GMDH)

The Group method of data handling (GMDH) is a polynomial based model. According to the GMDH approach, each layer can be obtained from a quadratic polynomial function. Thus the input variables are projected to the output variable. The main goal in this method is finding of function, f, that project the input variables to the output variable. Therefore, the output variable (Y_i) can be written from the input variables as the following form:

$$Y_i = f(X_{i1}, X_{i2}, X_{i3}, ..., X_{in})\ i = (1, 2, 3, ..., M) \qquad (2)$$

Where, X s are input variables. The structure of the GMDH can be obtained using the minimization of an objective function. The objective function can be written as:

$$\omega = \sum_{i=1}^{M} [Y(X_{i1}, X_{i2}, ..., X_i) - y_i]^2 \qquad (3)$$

Where, in the above equation y_i is actual data.

The general function between the inputs and the output variables was proposed by Ivakhnekoin the following form (Ivakhnenko, 1968):

$$Y = a_0 + \sum_{i=1}^{n} a_i X_i + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} X_i X_j + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{ijk} X_i X_j X_k + ... \qquad (4)$$

In this work, a quadratic polynomials function with only two variables (neurons) is considered.

$$Y = G(X_i, X_j) = a_0 + a_1 X_i + a_2 X_j + a_3 X_{ij} + a_4 X_i^2 + a_5 X_j^2 \qquad (5)$$

Where, parameters a_0-a_5 can be calculated from the minimization of Eq. (3). The least squares technique from multiple regression analysis is applied to calculate these parameters which obtained from solution of the following matrix:

$$Aa = Y \qquad (6)$$

Where, is the vector of unknown parameters of the quadratic polynomial (Eq. (6)):

$$A = \{a_0, a_1, a_2, a_3, a_4, a_5\} \qquad (7)$$

and

$$y = \{\ y_1, y_2, y_3, ..., y_M\}^T \qquad (8)$$

Table 1. Polynomial equations for prediction of moisture ratio (MR) with GMDH model*

Nod 1	N_1 = 0.923027–Time×0.0071617–Time×Thickness×1.76534e-05+Time^2×1.43249e-05+Tem.×0.00609525+Tem.×Thickness×0.00043174-Tem.^{2}×0.000178618
Node 2	N_2 = 0.059891+Tem.×8.72889e-05+Tem.×Thickness×4.42286e-05+Tem.×N_1×0.0058998-Tem.^{2}×2.32451e-05-Thickness×N_1*0.00692694+N_1×0.549558+N_1^2×0.260731
Output	Moisture ratio= -0.869111+Time×0.00743856+Time×N_1×0.0290015-Time×N_2×0.0415924-Time^2×1.2391e-05+N_2×3.34531–N_2^2×1.46832

*Variables' units (Tim (min), Thickness (mm), Temperature (°C)).

Where, is the vector of the actual data.

$$A=\begin{bmatrix} 1 & X_{1p} & X_{1q} & X_{1p}X_{1q} & X_{1p}{}^2 & X_{1q}{}^2 \\ 1 & X_{2p} & X_{2q} & X_{2p}X_{2q} & X_{2p}{}^2 & X_{2q}{}^2 \\ \vdots & \vdots & \vdots & \vdots\ \ \vdots & \vdots & \vdots \\ 1 & X_{Mp} & X_{Mq} & X_{Mp}X_{Mq} & X_{Mp}{}^2 & X_{Mq}{}^2 \end{bmatrix} \qquad (9)$$

Therefore, the vector of unknown parameter is given as below:

$$a=(A^TA)^{-1}A^TY \qquad (10)$$

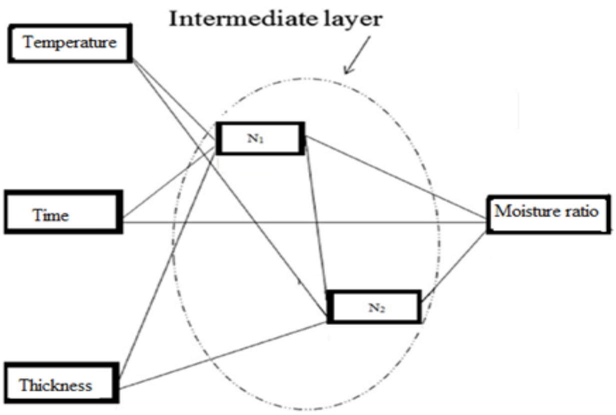

Figure 1. A schematic diagram of the GMDH model.

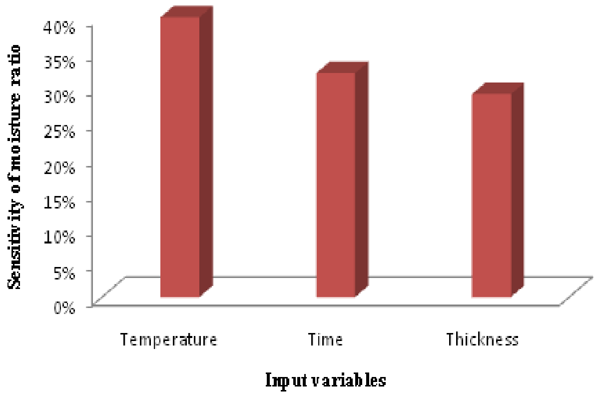

Figure 2. Comparison of moisture ratio sensitivity with input variables.

Results and Discussion

In this work, hybrid GMDH-type neural network was developed for estimation of papaw fruit MR during drying in a cabinet dryer. The experimental data contained 390 points while 50% of these data points were randomly used for training and 50% for testing. To further check for any possibility of over-fitting, different ratios in a range from 1 to 9 with increment of 0.5 are consecutively tested to find the optimum value. No over-fitting and considerably lesser error were observed that can be justified by rough linearity of data set.

Figure 1 shows the optimal structure of GMDH–Neural Network model developed with one hidden layer. As it can be seen from Figure 1, the proposed model has one input layer, one middle layer and one output layer. Generated functions corresponding to each node with total correlation function are reported in Table 1. It is worth meaning that all input variables were accepted by the model. In other words, the GMDH model provided an automated selection of essential input variables and built polynomial equations to model. These polynomial equations showed the quantitative relationship between input and output variables (Table 1).

It should be noted that, the GMDH was modeled with three inputs (temperature (°C), thickness (mm) and time (min)) and three neurons in the hidden layer and one in the output layer (moisture ratio). The performance of the training and testing by the network were estimated by AAD % (Average Absolute Deviations) as bellow:

Table 2. Model statistics GMDH model for predicting moisture ratio

Statistics		Training	Testing
Absolute Fraction of variance (R^2)	$R^2 = 1 - \left[\sum_{i=1}^{N} (Y_i^{model} - Y_i^{actual})^2 / \sum_{i=1}^{N} (Y_i^{actual})^2 \right]$	0.9989	0.9960
Root Mean Square Error (RMSE)	$RMSE = \left[\sum_{i=1}^{N} (Y_i^{model} - Y_i^{actual})^2 / \mathrm{N} \right]^{1/2}$	0. 017	0.022
Mean Square Error (MSE)	$MSE = \sum_{i=1}^{N} (Y_i^{model} - Y_i^{actual})^2 / \mathrm{N}$	0.00029	0.00048
Mean Absolute Deviation (MAD)	$MAD = \sum_{i=1}^{N} \left| Y_i^{model} - Y_i^{actual} \right| / N$	0. 0081	0.0099

Table 3. Statistical analyses for the mathematical models

Model name	Model constants	R^2	RMSE
Newton	k = 0.0089	0.9954	0.023
Page	k = 0.0054, n = 1.0993	0.9961	0.0191
Modified Page	k = 0.0126, n = 0.7074	0.9873	0.0487
Henderson and Pabis	k = 0.0092, a = 1.0407	0.9880	0.050
Two terms	k_0 = 0.0093, k_1 = 0.1962, a = 1.0493, b = -0.0514	**0.9974**	**0.0123**
Exponential two terms	k = 1.1853, a = 0.0074	0.9876	0.0493
Wang and Singh	a = -0.0067, b = 0.00001	0.9918	0.0213
Approximation of diffusion	k = 0.0381, a = -0.1217, b = 0.2579	0.9982	0.0499

R^2: Coefficient of determination; RMSE: Root-mean-square error

$$\%AAD = \frac{100}{N} \sum_{i=1}^{N} \left| \frac{Y_i^{model} - Y_i^{actual}}{Y_i^{actual}} \right| \qquad (11)$$

Where, the superscripts of "model" and "actual" refer to the model and actual results, respectively. The values of Average Absolute Deviations Percent (AAD %) calculated for the test data were within the range of 0.30%-28.11% and for the entire test data was 3.63%. The last value clearly shows the reliability and accuracy of the proposed GMDH model in estimation of moisture ratio.

Some statistical tests can be used for determining the models accuracy and reliability of the GMDH model. These statistical values can be defined as shown in Table 2 and their values were calculated based on the output of the network. The high value of R^2 (0.9960) in addition with the low values of RMSE (0.022), MSE (0.00048) and MAD (0.0099) for GMDH model indicated the high performance of that for estimation of MR. Figure 2 shows the sensitivity of moisture ratio to input variables. It is found that the sensitivity to the temperature was more than other inputs so that sensitivity of this parameter was near 40%. It can be concluded that the temperature has the most important role in this system. In agreement with this result, the high sensitivity of many agricultural crops to drying temperature is reported using activation energy parameter (Park *et al*., 2002; Kaleemullah and Kailappan, 2005).

In addition with the GMDH modelling, the moisture ratio values obtained under various experimental conditions were subjected to eight empirical mathematical models. Calculated R^2 and RMSE indicated that the Two terms model was the best among the mathematical models considered for fitting the experimental data (Table 3). The comparison between R^2 (0.9974) and RMSE (0.0123) of the Two terms and GMDH network models (R^2 = 0.9960, RMSE = 0.022) demonstrated that GMDH predicted close data to the experimental ones almost as good as the Two terms model. Erenturk *et al.* (2004) reported the same results for thin-layer drying of Echinacea Angustifolia root. They reported that the feed-forward neural network based estimation was more concise (R^2 = 0.999) even than the best mathematical model used (modified page) (R^2 = 0.993). For two varieties of green malt, Aghajani *et al*. (2012) found that the estimated moisture ratio by feed-forward back propagation neural network was more accurate than Page's model. Also, similar results which imply the high precision of neural network based modes for prediction of moisture content been reported (Momenzadeh *et al*., 2011; Khazaei *et al*., 2013; Yousefi *et al*., 2013a; Huang and Chen, 2015; Nadian *et al*., 2015). No specific work was found in the case of estimation of moisture content using GMDH-type neural network, but many researchers have reported the remarkable accuracy of this method in other fields (Ahmadi *et al*., 2007; Abdolrahimi *et al*., 2014; Atashrouz *et al*., 2015; Najafzadeh, 2015).

Conclusion

In this study, drying kinetics of thin-layer papaw fruit was investigated experimentally. Besides, a comparative study between a regression analysis and GMDH for estimation of moisture ratio (MR) during drying process was performed. The Two terms model indicated the closest results to the experimental data among the eight thin-layer empirical models considered. Higher R^2 and lower RMSE values calculated for GMDH proved the high performance of GMDH for prediction of moisture content. Altogether, it can be concluded that due to the high precision, GMDH- type neural networks can be applied for on-line state estimation and control of drying processes in industrial operations successfully.

References

Abdolrahimi, S., Nasernejad, B. and Pazuki, G. 2014. Prediction of partition coefficients of alkaloids in ionic liquids based aqueous biphasic systems using hybrid group method of data handling (GMDH) neural network. Journal of Molecular Liquids 191: 79-84.

Aghajani, N., Kashaninejad, M., Dehghani, A. A. and Daraei Garmakhany, A. 2012. Comparison between artificial neural networks and mathematical models for moisture ratio estimation in two varieties of green malt. Quality Assurance and Safety of Crops and Food 4: 93-101.

Aghdam, S., Yousefi, A. R., Mohebbi, M., Razavi, S., Orooji, A. and Akbarzadeh-Totonchi, M. R. 2015. Modeling for drying kinetics of papaw fruit using fuzzy logic table look-up scheme. International Food Research Journal 22: 1234-1239.

Ahmadi, H., Mottaghitalab, M. and Nariman-Zadeh, N. 2007. Group method of data handling-type neural network prediction of broiler performance based on dietary metabolizable energy, methionine, and lysine. The Journal of Applied Poultry Research 16: 494-501.

Akgun, N. A. and Doymaz, I. 2005. Modelling of olive cake thin-layer drying process. Journal of Food Engineering 68: 455-461.

Atashrouz, S., Pazuki, G. and Kakhki, S. S. 2015. A GMDH-type neural network for prediction of water activity in glycol and Poly (ethylene glycol) solutions. Journal of Molecular Liquids 202: 95-100.

De Oliveira, J. G. and Vitória, A. P. 2011. Papaw: Nutritional and pharmacological characterization, and quality loss due to physiological disorders. An overview. Food Research International 44: 1306-1313.

Demirtas, C., Ayhan, T. and Kaygusuz, K. 1998. Drying behaviour of hazelnuts. Journal of the Science of Food and Agriculture 76: 559-564.

Dinani, S. T., Hamdami, N., Shahedi, M. and Havet, M. 2014. Mathematical modeling of hot air/ electrohydrodynamic (EHD) drying kinetics of mushroom slices. Energy Conversion and Management 86: 70-80.

Erenturk, K., Erenturk, S. and Tabil, L. G. 2004. A comparative study for the estimation of dynamical drying behavior of Echinacea angustifolia: regression analysis and neural network. Computers and Electronics in Agriculture 45: 71-90.

Ghanadzadeh, H., Ganji, M. and Fallahi, S. 2012. Mathematical model of liquid–liquid equilibrium for a ternary system using the GMDH-type neural network and genetic algorithm. Applied Mathematical Modelling 36: 4096-4105.

Huang, Y. and Chen, M. 2015. Artificial neural network modeling of thin layer drying behavior of municipal sewage sludge. Measurement 73: 640-648.

Ivakhnenko, A. 1968. The group method of data handling-a rival of the method of stochastic approximation. Soviet Automatic Control 13: 43-55.

Ivakhnenko, A. 1971. Polynomial theory of complex systems. IEEE Transactions on Systems, Man and Cybernetics SMC-1: 364-378.

Izadifar, M. and Mowla, D. 2003. Simulation of a cross-flow continuous fluidized bed dryer for paddy rice. Journal of Food Engineering 58: 325-329.

Kaleemullah, S. and Kailappan, R. 2005. Drying kinetics of red chillies in a rotary dryer. Biosystems Engineering 92: 15-23.

Kamiński, W., Tomczak, E. and Strumill, P. 1998. Neurocomputing approaches to modelling of drying process dynamics. Drying Technology 16: 967-992.

Khazaei, N. B., Tavakoli, T., Ghassemian, H., Khoshtaghaza, M. H. and Banakar, A. 2013. Applied machine vision and artificial neural network for modeling and controlling of the grape drying process. Computers and Electronics in Agriculture 98: 205-213.

Kingsly, A. R. P. and Singh, D. B. 2007. Drying kinetics of pomegranate arils. Journal of Food Engineering 79: 741-744.

Koukouch, A., Idlimam, A., Asbik, M., Sarh, B., Izrar, B., Bah, A. and Ansari, O. 2015. Thermophysical characterization and mathematical modeling of convective solar drying of raw olive pomace. Energy

Conversion and Management 99: 221-230.
Liebman, B. 1992. Nutritional aspects of fruit. Nut Action Newsletter 1: 10-11.
Liu, X., Chen, X., Wu, W. and Peng, G. 2007. A neural network for predicting moisture content of grain drying process using genetic algorithm. Food Control 18: 928-933.
Midilli, A., Kucuk, H. and Yapar, Z. 2002. A new model for single-layer drying. Drying Technology 20: 1503-1513.
Momenzadeh, L., Zomorodian, A. and Mowla, D. 2011. Experimental and theoretical investigation of shelled corn drying in a microwave-assisted fluidized bed dryer using Artificial Neural Network. Food and Bioproducts Processing 89: 15-21.
Movagharnejad, K. and Nikzad, M. 2007. Modeling of tomato drying using artificial neural network. Computers and Electronics in Agriculture 59: 78-85.
Nadian, M. H., Rafiee, S., Aghbashlo, M., Hosseinpour, S. and Mohtasebi, S. S. 2015. Continuous real-time monitoring and neural network modeling of apple slices color changes during hot air drying. Food and Bioproducts Processing 94: 263-274.
Najafzadeh, M. 2015. Neuro-fuzzy GMDH based particle swarm optimization for prediction of scour depth at downstream of grade control structures. Engineering Science and Technology, an International Journal 18: 42-51.
Nariman-Zadeh, N. and Jamali, A. 2007. Pareto genetic design of GMDH-type neural networks for nonlinear systems. In Drchal, J. and Koutnik, J. (Eds). Proceeding of the International Workshop on InductiVe Modelling, p. 96-103. Czech Technical University: Prague, Czech Republic, Citeseer.
Nikbakht, A. M., Motevali, A. and Minaei, S. 2014. Energy and exergy investigation of microwave assisted thin-layer drying of pomegranate arils using artificial neural networks and response surface methodology. Journal of the Saudi Society of Agricultural Sciences 13: 81-91.
Onwubolu, G.C. 2009. Hybrid self-organizing modeling systems, p. 99-138. Berlin: Springer.

Park, J. and Sandberg, I. W. 1991. Universal approximation using radial-basis-function networks. Neural Computation 3: 246-257.
Park, K. J., Vohnikova, Z. and Brod, F. P. R. 2002. Evaluation of drying parameters and desorption isotherms of garden mint leaves (*Mentha crispa* L.). Journal of Food Engineering 51: 193-199.
Pazuki, G. and Kakhki, S. S. 2013. A hybrid GMDH neural network to investigate partition coefficients of Penicillin G Acylase in polymer–salt aqueous two-phase systems. Journal of Molecular Liquids 188: 131-135.
Powell, M.J. 1987. Radial basis functions for multivariable interpolation: a review. In Mason, J. and Cox, M. (Eds). Algorithms for Approximation, p. 143-167, Clarendon Press.
Wang, Z., Sun, J., Liao, X., Chen, F., Zhao, G., Wu, J. and Hu, X. 2007. Mathematical modeling on hot air drying of thin layer apple pomace. Food Research International 40: 39-46.
Yousefi, A. R., Asadi, V., Nassiri, S. M., Niakousari, M. and Khodabakhsh Aghdam, S. 2013a. Comparison of mathematical and neural network models in the estimation of papaw fruit moisture content. The Philippine Agricultural Scientist 95: 186-191.
Yousefi, A. R., Niakousari, M. and Moradi, M. 2013b. Microwave assisted hot air drying of papaw (*Carica papaw* L.) pretreated in osmotic solution. African Journal of Agricultural Research 8: 3229-3235.