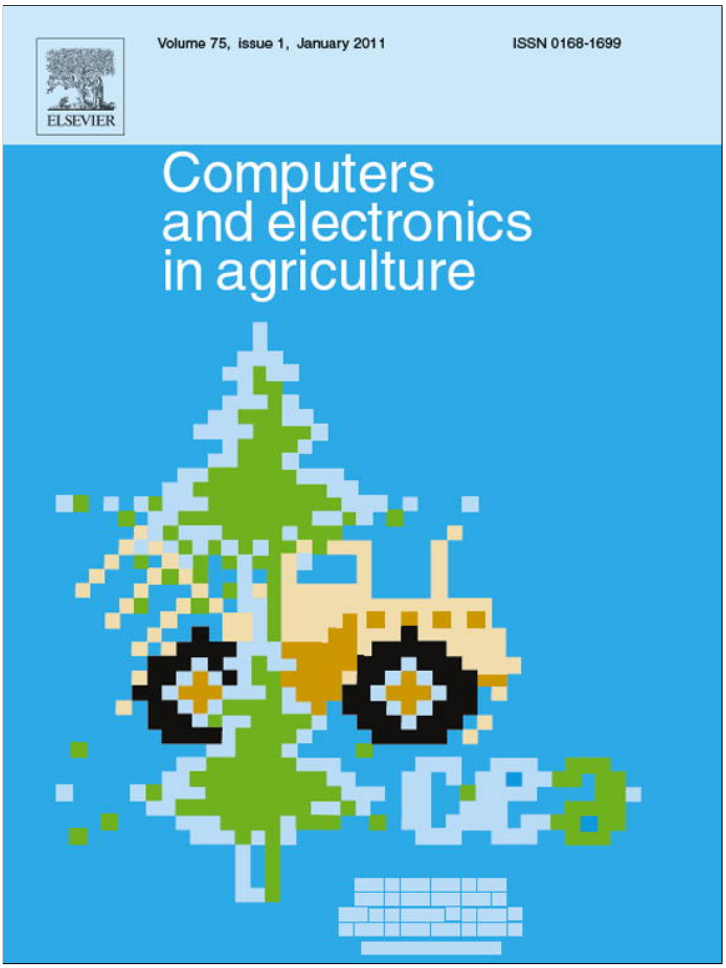

(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

Computers and Electronics in Agriculture 82 (2012) 75–86

Contents lists available at SciVerse ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Predictions of apple bruise volume using artificial neural network

Saeed Zarifneshat [a,*], Abbas Rohani [b], Hamid Reza Ghassemzadeh [c], Morteza Sadeghi [d], Ebrahim Ahmadi [e], Masoud Zarifneshat [f]

[a] Khorasan Razavi Agriculture and Natural Resources Research Center, Mashhad, Iran
[b] Department of Farm Machinery Engineering, College of Agriculture, Shahrood University of Technology, Shahrood, Iran
[c] Department of Agricultural Machinery Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran
[d] Department of Mechanical Engineering, University of Tabriz, Tabriz, Iran
[e] Department of Agricultural Machinery Engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran
[f] Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran

ARTICLE INFO

Article history:
Received 1 July 2011
Received in revised form 22 December 2011
Accepted 26 December 2011

Keywords:
Bruise volume
Golden Delicious apple
Fruit properties
ANN
BDLRF algorithm
Basic Back-propagation

ABSTRACT

Bruise damage is a major cause of fruit quality loss. Bruises occur under dynamic and static loading when stress induced in the fruit exceeds the failure stress of the fruit tissue. In this article the potential of an artificial neural network (ANN) technique has evaluated as an alternative method for the prediction of apple bruise volume. Neural bruise estimation models were constructed to calculate Golden Delicious apple bruise volume with respect to fruit properties. The neural models were built based upon impact force and impact energy as the main input parameters including fruit curvature radius, temperature and acoustical stiffness. Optimal parameters for the network were selected via a trial and error procedure on the available data. In this paper, the performance of Basic Backpropagation (BB) training algorithm was also compared with Backpropagation with Declining Learning Rate Factor algorithm (BDLRF). It was found that BDLRF has a better performance for the prediction of apple bruise volume. It is concluded that ANN represents a promising tool for predicting apple bruise volume in comparison to regression model.

1. Introduction

Apple consumers increasingly demand better quality fruit. Mechanical damage such as bruises, abrasions, cuts and punctures are irreversible and are accumulated damages during the handling process. The inevitable consequence of mechanical damage is low grade and low quality fruits, hence less income to both growers and packers (Timm et al., 1996; Abbott and Lu, 1996). The factors affecting damage severity are fruit fall height, contact energy, the number of contact, the kind of contact surface and the size and ripeness stage of the fruit (Lin and Brusewitz, 1994; Roth et al., 2005). Identifying the impact situations which create bruises is necessary to improve harvesting, transporting, grading procedures and equipments (Lin and Brusewitz, 1994; Ragni and Berardinelli, 2001; Van Linden et al., 2006). Bruising may be intensified by some other factors such as texture, variety, maturity stage, water content, fruit shape, temperature, firmness, size and a series of fruit interior factors such as modulus of elasticity, strength of cell walls and, internal structure and cell shape (Studman et al., 1997; Van Linden et al., 2006).

Detailed information about bruise estimation models for Golden Delicious apple is limited. The statistical models contain contact energy or impact force as the main independent variable (Van Zeebroeck et al., 2007a). The impact force models are advantageous, because they can be generalized to impacts of materials with different material properties and curvature radius. The impact force is the influential factor in explicating the apples bruising when impacted by various materials. The bruise estimation models, containing impact energy that was constructed for metal-apple indenter contact, cannot be applied to apple–apple contact because in both situations the impact energy is the same (Pang et al., 1992; Van Zeebroeck et al., 2007c). Nevertheless, a drawback of exploiting impact force regression model is the high probability of impact force being influenced by fruit properties (e.g., temperature and maturity). In contrast, the energy of impact is not affected by the fruit characteristics, which gives impact energy models an advantage over impact force models. As a result, contact energy models are superior to study the effect of fruit characteristics on bruise damage (Van Zeebroeck et al., 2007a–c).

Bruise estimation models for apples (especially for Golden Delicious apple) reported in the references are limited to the effect of two properties of fruit (either maturity or temperature of fruit) and these models are unstable (Studman et al., 1997). ANNs are being used in a wide variety of applications, including apple bruise

* Corresponding author. Address: P.O. Box 91735-488, Mashhad, Iran. Tel./fax: +98 5113822373.
E-mail addresses: s_zarif@tabrizu.ac.ir, zarifneshat@yahoo.com (S. Zarifneshat).

0168-1699/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.compag.2011.12.015

prediction (Barreiro et al., 1997), fruit grading (Kavdir and Guyer, 2004; Xiaobo et al., 2007; Al-Ohali, 2010; Effendi et al., 2010) and modeling of apple drying process (Khoshhal et al., 2010).

The main objective of this study was to develop apple bruise volume prediction models. The specific objectives were: (1) to investigate the effectiveness of ANN for predicting apple bruise volume; (2) study the variation of model performance with different ANN model parameters; (3) select optimum ANN parameters for accurate prediction apple bruise volume.

2. Materials and methods

2.1. Experimental details

The apples used in this study were Golden Delicious variety. The apples were harvested from "Shabestar" district, Tabriz, Iran. Apples were hand-picked at a commercial orchard to ensure their freshness and avoid damage during harvesting and transporting. Fruits were stored in a cool storage with controlled atmosphere (85% RH, 3 °C). All measurements were performed within maximum 2 days. The fruit was kept at desired temperature for 10 h prior to measurement. Samples at 3 °C were measured during 15 min to reduce apple warming in the measuring room at 20 °C. Apples were placed on a pendulum equipped with a force sensor (type AC20, AP Tech, Netherlands; Sensitivity: 1.87 mV N^{-1}) and an encoder equipment (RON 275, Heidenhain®; Resolution: 0.005°) to measure energy of impact and impact velocity (Van Zeebroeck et al., 2003). The samples then were hit by a spherical metal impact or with radius of 25 mm (Fig. 1).

The dependent variable that used in the bruise estimation models was the bruise damage volume. The bruise volume was recorded 48 h after contact and determined based on:

$$BV = \frac{\pi}{6} dD^2 \quad (1)$$

where BV is the bruise volume (mm^3), d bruise depth (mm) and D bruise diameter (mm), respectively.

Bruise estimation models had either the impact energy (kinetic energy of pendulum rod just before collision) or the impact force as independent variables along with other variables. Used independent variables in the regression models or inputs of neural network consist of: impact energy (E) (J), impact force (F) (N), fruit temperatures (T), curvature radius of apple (R) at the contact location (mm), fruit acoustical stiffness (S) (s^{-2} $kg^{2/3}$).

The applied impact energy levels were chosen above the critical impact level of Golden Delicious apple. The lower limit of applied impact level was based on the measured impact force during handling and transporting but the higher impact level was in apple

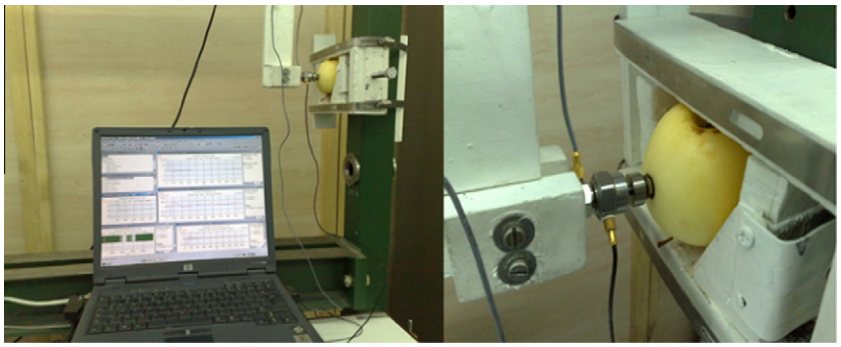

Fig. 1. General view of the pendulum device for measuring impact force and impact velocity of the apple fruit.

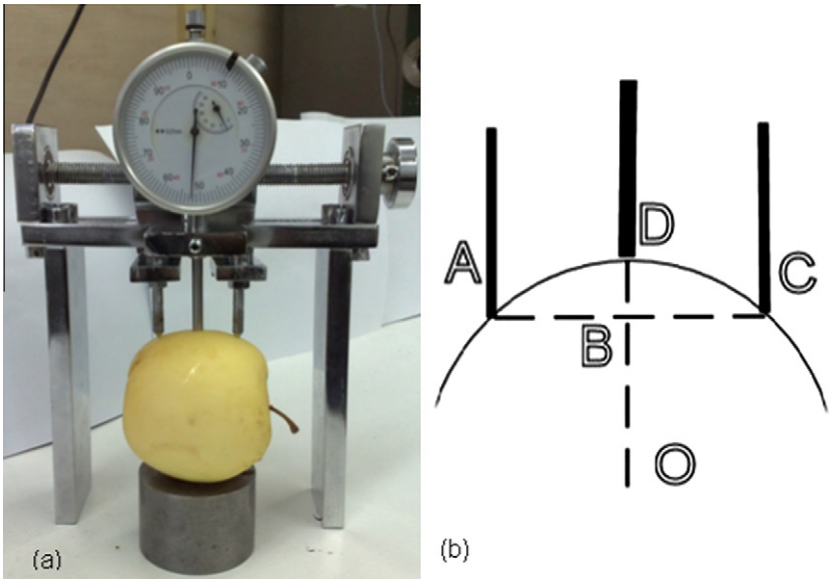

Fig. 2. (a) General view of the radius meter and (b) schematic representation of geometry to calculate the radius of the apple fruit.

mechanical harvesting and sorting. For each impact the exact impact energy and impact force were recorded. The lowest impact level was close to the detection limit of bruise damage and other levels (2 and 3) exposed obviously visible damages and were easily perceivable.

The curvature radius was measured at the fruit contact area by means of a non-commercial radius meter because a proper measuring device was not available. It was, therefore, constructed on an analog height meter base (Fig. 2a).

The curvature radius was determined using the following equation (Mohsenin, 1986):

$$\text{Radius} = \frac{(AC)^2}{8(BD)} + \frac{(BD)}{2} \tag{2}$$

where A, C and D are contact points of analog height meter probes on apple surface and point B is middle of AC line segment (Fig. 2b).

Since an apple cannot be considered to be completely spherical, the harmonic average $(2R_1R_2/(R_1 + R_2))$ was chosen based on circumferential (R_1) and meridian curvature radius (R_2). Based on Hertz theory the use of harmonic mean is more acceptable than the computational mean, due to its accuracy on estimation of smaller curvature radius, which participate more to the maximum contact pressure.

The acoustical stiffness of apple was calculated based on acoustical impulse-response method (Schotte et al., 1999; Van Zeebroeck et al., 2007b). The apple was positioned with the stalk end on a rubber pad. A microphone was fixed on a support at a few millimeters off the apple and was directed upward. The fruit was stimulated by tapping it on the equator at the opposing side of the microphone by a rigid plastic bar. Acoustical measurements were taken with a microphone that records the signal of sound arising from the response vibration. Amplifiers supplied electrical power to the transducer, magnified the signal and give proper output drive signal and permit choosing the proper band-pass filters. The signals of this microphone were collected and processed using a PULSE® program (type 3564, B and K®). Before impact experiments, the mass of apples were measured. The setup was adjusted so that the collision of the apple triggered the measurement. To obtain the signal's frequency spectrum a Fast Fourier Transform (FFT) was performed and consequently, the apples' first resonance frequency was determined. The acoustical stiffness was calculated as:

$$S \cong f^2 m^{2/3} \tag{3}$$

where S is the acoustic stiffness ($s^{-2}\ kg^{2/3}$), f the first resonance frequency (s^{-1}), and m the mass of the apple (kg), respectively.

A total of 120 apples were used for conducting the experiments. These apples were divided into six groups and consequently, 20 apples were tested for each temperature-impact level combination.

2.2. Data preprocessing

Based on these available data, the energy (E), contact force (F), curvature radius (R), temperature (T) and acoustical stiffness (S) were selected as variable inputs. The bruise volume (mm^3) of the Golden Delicious apple was selected as variable output. Prior to any ANN training process with the trend free data, the data must be normalized over the range of [0,1]. This is necessary for the neurons' transfer functions, because a sigmoid function is calculated and consequently these can only be performed over a limited range of values. If the data used with an ANN are not scaled to an appropriate range, the network will not converge on training or it will not produce meaningful results. The method of normalization involves mapping the data nonlinear over a specified range, whereby each value of a variable x is transformed as follows:

$$x_n = \frac{\log(x) - \log(x_{\min})}{\log(x_{\max}) - \log(x_{\min})} \times (r_{\max} - r_{\min}) + r_{\min} \tag{4}$$

where x is the original data, x_n the normalized input or output values, $x_{\max}$ and $x_{\min}$, are the maximum and minimum values of the concerned variable, respectively. $r_{\max}$ and $r_{\min}$ correspond to the desired values of the transformed variable range. A range of 0.1–0.9 is appropriate for the transformation of the variable onto the sensitive range of the sigmoid transfer function.

The data were shuffled and split into two subsets: a training set and a test set. The splitting of samples plays an important role in the evaluation of an ANN performance. The training set is used to estimate model parameters and the test set is used to check the generalization ability of the model. The training set should be a representative of the whole population of input samples. In this study, the training set and the test set includes 96 patterns (80% of total patterns) and 24 patterns (20% of total patterns), respectively. There is no acceptable generalized rule to determine the size of training data for a suitable training; however, the training sample should cover all spectrums of the data available (NeuroDimensions Inc., 2002). The training set can be modified if the performance of the model does not meet the expectations (Zhang and Fuh, 1998). However, by adding new data to the training samples, the network then can be retrained.

2.3. The multilayer perceptron neural network

Fig. 3 shows a MLP with one hidden layer.

The network is in charge of vector mapping, i.e. by inserting the input vector, X^q the network will answer through the vector Z^q in its output (for $q = 1,\ldots,Q$).

In this study, two variants of MLP training algorithm, i.e. Basic Back-propagation (BB) and Back-Propagation with Declining Learning-rate Factor (BDLRF) where employed (Vakil-Baghmisheh and Pavešic, 2001). A computer code was also developed in MATLAB software to implement these ANN models.

2.3.1. BB algorithm

In this algorithm the total sum-squared error (TSSE) is considered as the cost function and can be calculated as

$$\text{TSSE} = \sum_q E_q \tag{5}$$

$$E_q = \sum_k \left(d_k^q - z_k^q\right)^2 \text{ for } (q = 1,\ldots,Q) \tag{6}$$

where d_k^q and z_k^q are the kth components of desired and actual output vectors of the qth input, respectively.

The connection weights between nodes of different layers are updated using the following equations:

$$u_{jk}(n+1) = u_{jk}(n) - \eta \times \frac{\partial E}{\partial u_{jk}} + \alpha\left(u_{jk}(n) - u_{jk}(n-1)\right) \tag{7}$$

$$w_{ij}(n+1) = w_{ij}(n) - \eta \times \frac{\partial E}{\partial w_{ij}} + \alpha\left(w_{ij}(n) - w_{ij}(n-1)\right) \tag{8}$$

where η is the learning rate adjusted between 0 and 1, α is the momentum factor at interval [0,1], w_{ij} is the connection weight between nodes i and j (for $i = 1,\ldots,l_1, j = 1,\ldots,l_2$), and u_{jk} is the connection weight between nodes j and k (for $j = 1,\ldots l_2, k = 1,\ldots l_3$); w_{ij} and u_{jk} are set to small random values $[-0.25, 0.25]$; l_2 and l_3 are the number of neurons in the hidden and output layers. The decision to stop training is based on some test results of the network, which is carried out every N epoch after TSSE becomes smaller than a threshold value. The details could be seen in Vakil-Baghmisheh and Pavešic (2003).

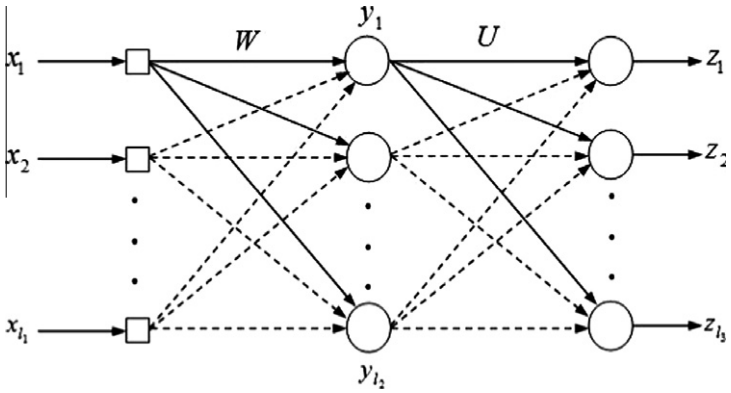

Fig. 3. Configuration of the MLP with one hidden layer (Vakil-Baghmisheh, 2002).

2.3.2. BDLRF algorithm

This training algorithm is started with a relatively constant large step size of learning rate η and momentum term α. Before destabilizing the network or when the convergence is slowed down, these are values are decreased monotonically (Vakil-Baghmisheh and Pavešic, 2001).

2.4. Regression model

The dependent variable was the bruise volume (BD) of apple. The independent variables were impact energy (E), impact force (F), curvature radius at contact location (R), acoustical stiffness (S) and temperature of apple (T). A backward multiple regression method was applied to choose the pertinent independent variables influencing the dependent variable. Furthermore, in order to verify the validity of multiple regression models, a chi-square test was carried out using the predicted and experimental data. SAS software was used for data analysis.

2.5. Performance evaluation criteria

Four criteria were used to evaluate the performance of model. They were mean absolute percentage error (MAPE), root mean-squared error (RMSE), TSSE and the coefficient of determination of the linear regression line between the predicted values from the MLP model and the actual output (R^2).They are defined as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{j=1}^{n}\sum_{i=1}^{m}(d_{ji}-p_{ji})^2}{nm}} \tag{9}$$

$$R^2 = \frac{\left(\sum_{j=1}^{n}(d_j-\bar{d})(p_j-\bar{p})\right)^2}{\sum_{j=1}^{n}(d_j-\bar{d})^2\cdot\sum_{j=1}^{n}(p_j-\bar{p})^2} \tag{10}$$

$$\mathrm{TSSE} = \sum_{j=1}^{n}(d_j-p_j)^2 \tag{11}$$

$$\mathrm{MAPE} = \frac{1}{nm}\sum_{j=1}^{n}\sum_{i=1}^{m}\left|\frac{d_{ji}-p_{ji}}{d_{ji}}\right|\times 100 \tag{12}$$

where d_{ji} is the ith component of the desired (actual) output for the jth pattern; p_{ji} is the ith component of the predicted (fitted) output produced by the network for the jth pattern; $\bar{d}$ and $\bar{p}$ are the average of the desired output and predicted output, respectively; n and m are the number of patterns and the number of variable outputs, respectively. A model with the smallest RMSE, TSSE, MAPE and the largest R^2 is considered to be the best.

3. Results and discussion

Neural networks were developed in order to establish the relationships between (i) bruise volume (mm^3) of the Golden Delicious apple and impact energy (E), curvature radius (R), temperature (T) and acoustical stiffness (S) (Fig 4); (ii) bruise volume (mm^3) of the Golden Delicious apple and contact force (F), curvature radius (R), temperature (T) and acoustical stiffness (S) (Fig 5). All networks were 3-layered feed forward type, trained using both BB and BDLRF training algorithms.

3.1. MLP topology (number of neurons in the hidden layer)

Based on universal approximation theorem, a neural network with a single hidden layer and sufficiently a large number of neurons can well approximate any arbitrary continuous function (Haykin, 1994). Therefore, the ANNs designed in this study are equipped with a single hidden layer. Determination of the number of neurons in the hidden layer is rather an art than science, because it may vary depending on the specific problem under study. In this study, the optimal number of neurons in the hidden layer was selected using a trial-and-error method and keeping the learning rate, momentum term and epoch size constant ($\eta = 0.4$, $\alpha = 0.8$ and epoch = 100,000). The process was repeated several times, one for each set of data. Table 1 shows the effect of number of neurons in the hidden layer on the performance of BB–MLP model. It is observed that the performance of BB–MLP is improved as the number of hidden neurons increased. However, too many neurons in the hidden layer may cause over-fitting problems, which results in good network learning and data memorization, but lack of ability to generalize. On the other hand, if the number of neurons in the hidden layer is not enough, the network may not be able to learn. Considering Table 1, a BB–MLP model with 25 neurons in the hidden layer seems to be appropriate for modeling $f(E,R,T,S)$ and with 30 neurons for modeling $f(F,R,T,S)$.

3.2. Learning rate and momentum term

For the selected topology, several learning processes were performed with different coefficients, ranged from 0 to 0.99 and 0.1 to 0.99 for learning rate and momentum term, respectively. Figs. 6 and 7 show the total sum square error values versus the learning rate and momentum term.

It is observed that the error value is increased and the convergence speed of the learning process is decreased when the momen-

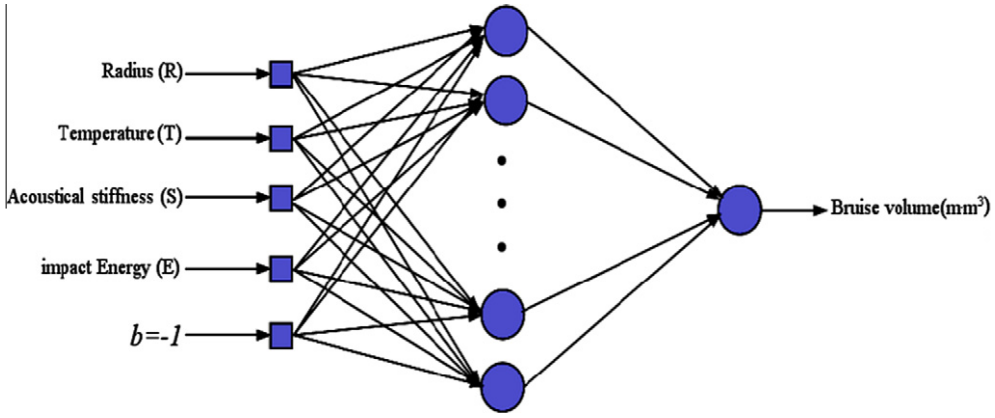

Fig. 4. Multilayer neural network used in the prediction of bruise volume ($f(R,T,S,E)$).

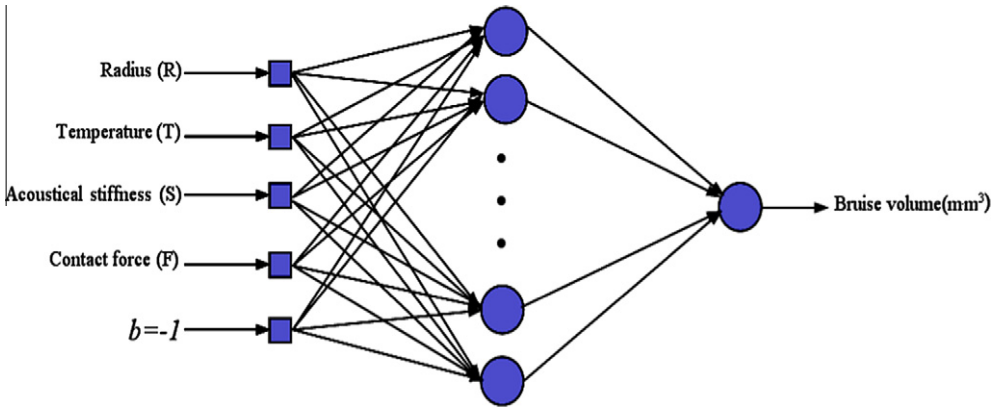

Fig. 5. Multilayer neural network used in the prediction of bruise volume ($f(R,T,S,F)$).

Table 1
Performance variation of a three-layer BB–MLP with different number of neurons in the hidden layer in the training phase.

Model	Criterion	Number of neurons in the hidden layer								
		5	10	15	25	30	35	40	45	50
$f(R,T,S,E)$	MAPE (%)	8.89	6.81	6.73	6.50	6.30	5.85	4.89	3.51	2.69
	RMSE	60.65	44.51	44.10	43.15	41.48	39.65	38.76	36.48	32.38
	TSSE	0.2333	0.1377	0.1356	0.1286	0.1200	0.1129	0.1035	0.0957	0.0841
$f(R,T,S,F)$	MAPE (%)	9.49	8.00	6.49	6.36	6.33	5.27	5.56	4.78	4.58
	RMSE	69.38	57.65	46.79	45.67	45.59	43.32	42.65	41.36	39.87
	TSSE	0.20901	0.1578	0.1182	0.1084	0.1008	0.0987	0.0989	0.0956	0.0932

tum term is zero or close to 1. The results also revealed that the convergence could be faster with a relatively larger learning rate (close to 1). However, with a very high learning rate, the neural network will not converge to its true optimum and the learning process will be instable. Table 2 shows a relative minimum and maximum learning rate and momentum term and minimum epoch associated with BB–MLP.

Table 3 shows the parameters of optimum BDLRF–MLP. It was found that performance of networks in the second approach is better than in the first approach.

3.3. Statistical analysis

3.3.1. Training phase

During training phase the network used the training set. Training was continued until a steady state was reached. The BB and BDLRF algorithms were utilized for model training. Some statistical properties of the sample data used for training process and the prediction values associated with different training algorithms are shown in Table 4. Considering the average values of standard deviation and variance, it can be deduced that the values and the distribution of real and predicted data are analogous. Accordingly, the neural networks have been learned the training set very well, hence the training phase has been completed.

3.3.2. Test phase

Table 4 shows some statistical properties of the data used in test phase and the corresponding prediction values associated with different training algorithms. It can be seen that the differences of statistical values between the measured and predicted data in test phase is more than in training phase for both of training algorithms (Tables 2 and 3).This fact can be justified since these data are completely new for the MLP. On the other hand, the kurtosis, sum and the average values are similar, hence it can be deduced that both series are similar. The predicted values were very close to the desired values and were evenly distributed throughout the entire

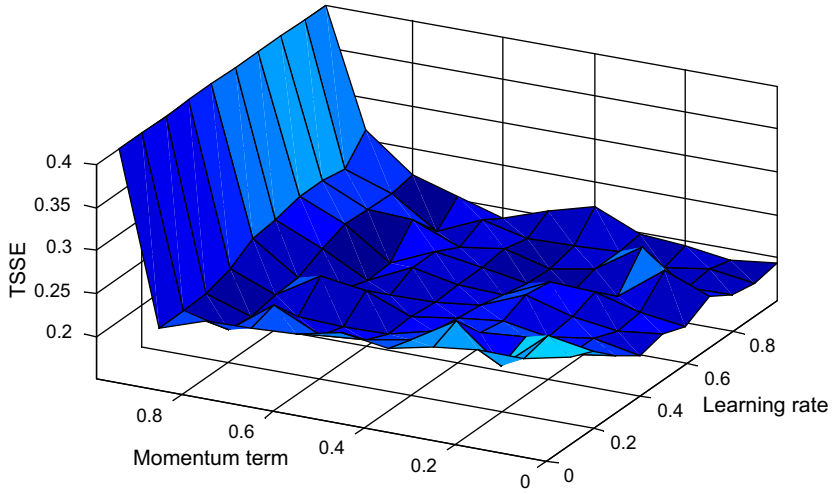

Fig. 6. TSSE profile as a function of learning rate and momentum term for modeling $f(R,T,S,E)$.

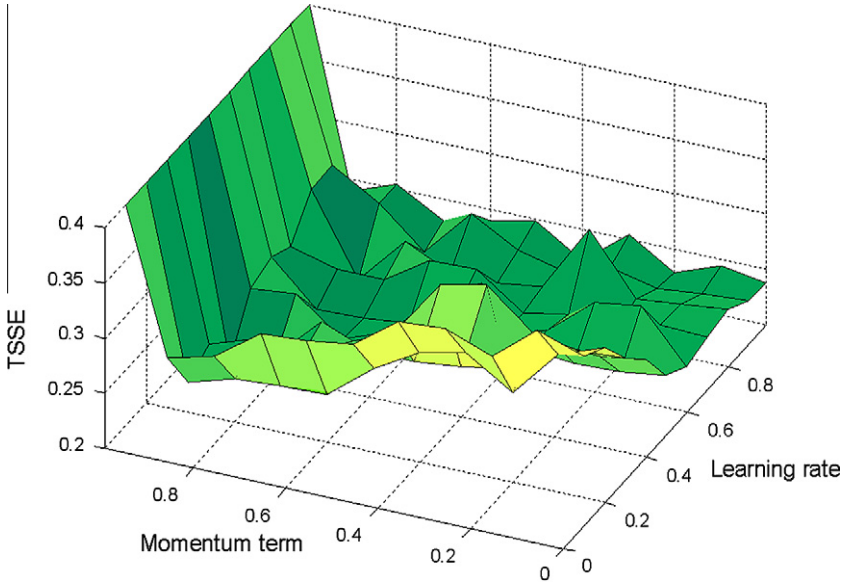

Fig. 7. TSSE profile as a function of learning rate and momentum term for modeling $f(R,T,S,F)$.

Table 2
Optimum parameters of neural network (BB–MLP).

Model	Parameters of neural network					
	Range of learning rate	Learning rate	Range of momentum term	Momentum term	Epoch	Topology
$f(R,T,S,E)$	0.1–0.3	0.25	0.7–0.9	0.8	300,000	5-25-1
$f(R,T,S,F)$	0.3–0.5	0.4	0.85–0.95	0.9	300,000	5-30-1

Table 3
Optimum parameters of neural network (BDLRF–MLP).

Model	Parameters of neural network						
	The start point of BDLRF	First phase(BB)		Second phase(BDLRF)		Epoch	Topology
		η	α	η	α		
$f(R,T,S,E)$	1000	0.8	0.9	0.04	0.05	300,000	5-25-1
$f(R,T,S,F)$	250,000	0.45	0.9	0.02	0.04	300,000	5-30-1

Table 4
Statistical variables of desired and predicted values in training phase (MLP).

Model		Training algorithm	Statistical values							
			Average	Variance	Standard deviation	Minimum	Maximum	Kurtosis	Skewness	Sum
f(*R*,*T*,*S*,*E*)	Desired values	BB and BDLRF	584.61	198098.16	445.08	55.30	1459.00	1.80	0.46	56122.73
	Predicted values	BB	587.32	197835.95	444.79	47.97	1434.75	1.79	0.45	56383.06
		BDLRF	587.01	197603.83	444.53	54.70	1433.33	1.78	0.45	56352.72
f(*R*,*T*,*S*,*F*)	Desired values	BB and BDLRF	577.11	194749.85	441.30	55.30	1459.00	1.84	0.48	54825.73
	Predicted values	BB	574.06	188491.94	434.16	54.81	1454.96	1.82	0.46	54535.29
		BDLRF	576.97	191609.35	437.73	55.29	1457.34	1.81	0.46	54811.79

Table 5
Statistical variables of desired and predicted values (test phase).

Model		Training algorithm	Statistical values							
			Average	Variance	Standard deviation	Minimum	Maximum	Kurtosis	Skewness	Sum
f(*R*,*T*,*S*,*E*)	Desired values	BB and BDLRF	591.79	163167.22	403.94	71.00	1392.00	1.83	0.23	14203.00
	Predicted values	BB	587.18	166777.32	408.38	74.94	1388.78	1.80	0.27	14092.26
		BDLRF	583.92	166019.05	407.45	74.95	1412.54	1.87	0.28	14014.09
f(*R*,*T*,*S*,*F*)	Desired values	BB and BDLRF	620.00	176261.33	419.83	71.00	1392.00	1.76	0.19	15500.00
	Predicted values	BB	622.89	183523.35	428.40	71.76	1355.23	1.68	0.20	15572.28
		BDLRF	617.69	177544.10	421.36	70.63	1338.75	1.70	0.19	15442.24

Table 6
Statistical comparisons of desired and predicted data and the corresponding *p* values.

	Phase	Training algorithm	Analysis types		
			Comparisons of means	Comparisons of variances	Comparisons of distribution
f(*R*,*T*,*S*,*E*)	Train phase	BB	0.966	0.995	0.999
		BDLRF	0.970	0.990	1.000
	Test phase	BB	0.969	0.959	1.000
		BDLRF	0.947	0.967	1.000
f(*R*,*T*,*S*,*F*)	Train phase	BB	0.962	0.875	0.999
		BDLRF	0.998	0.937	1.000
	Test phase	BB	0.981	0.922	1.000
		BDLRF	0.985	0.986	1.000

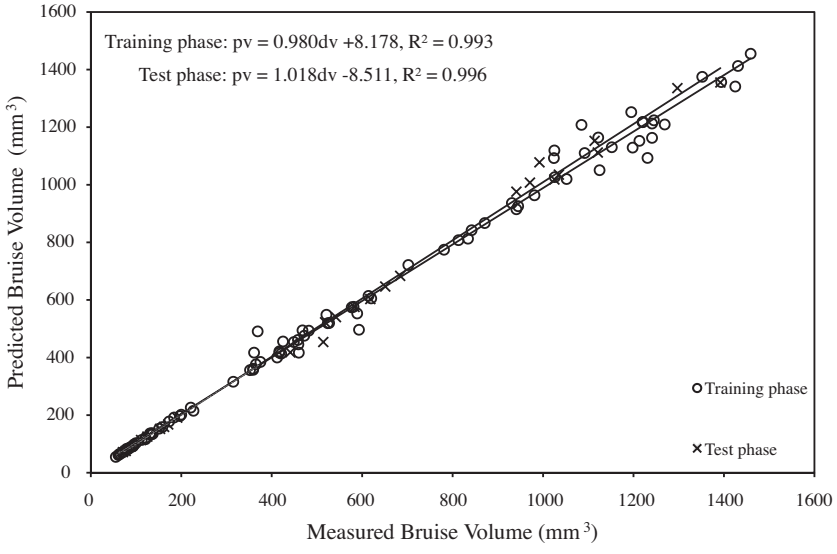

Fig. 8. Predicted values of BB–MLP network versus measured values of bruise volume for *f*(*R*,*T*,*S*,*F*).

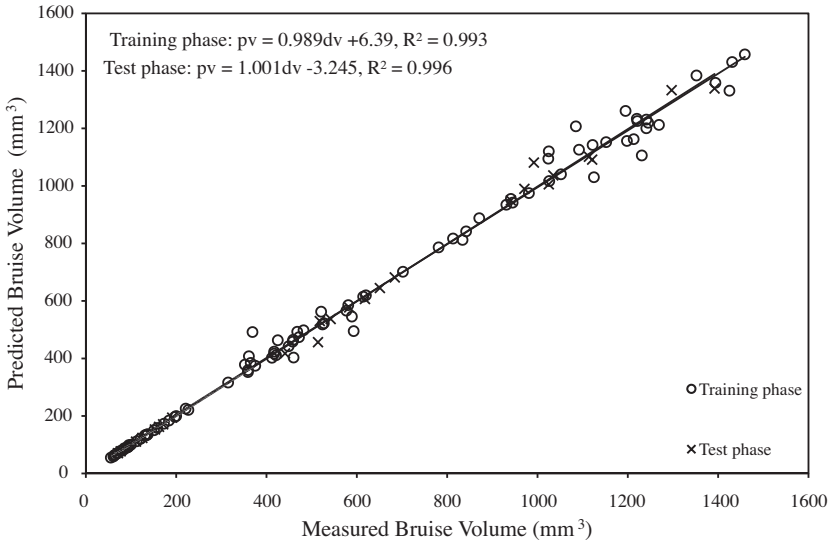

Fig. 9. Predicted values of BDLRF–MLP network versus measured values of bruise volume for *f*(*R*,*T*,*S*,*F*).

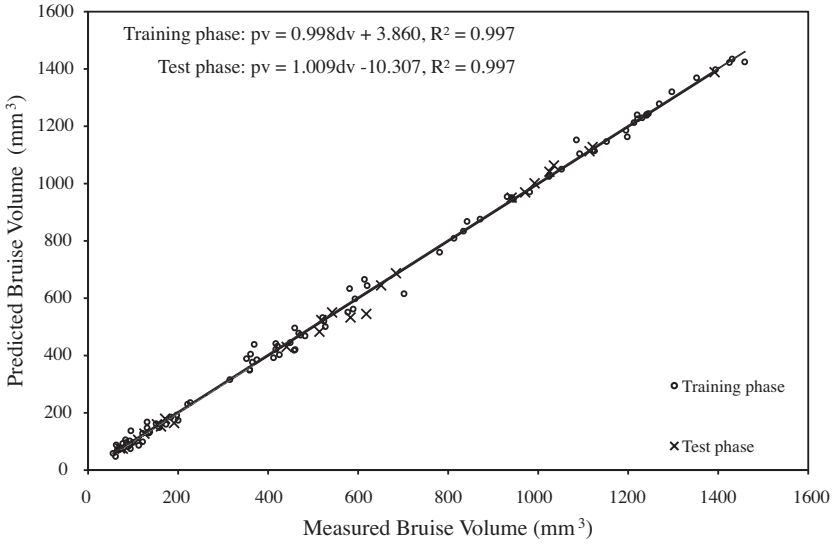

Fig. 10. Predicted values of BB–MLP network versus measured values of bruise volume for *f*(*R*,*T*,*S*,*E*).

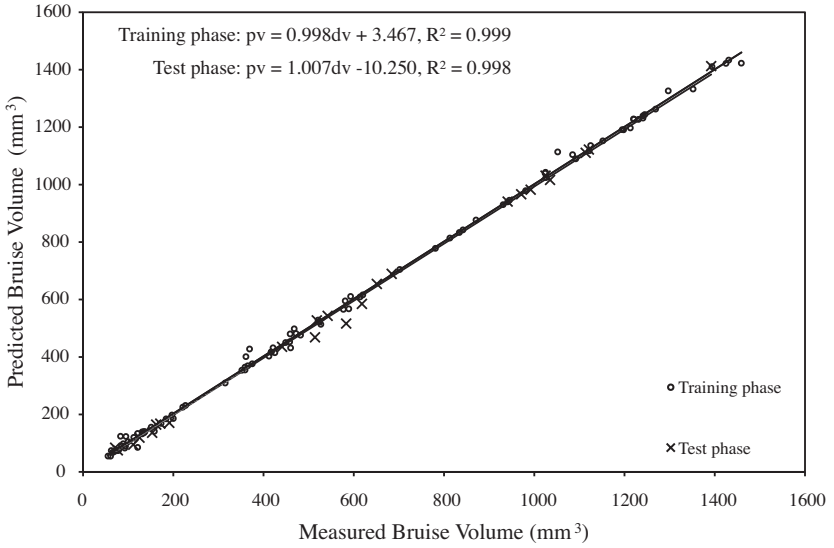

Fig. 11. Predicted values of BDLRF–MLP network versus measured values of bruise volume for *f*(*R*,*T*,*S*,*E*).

Table 7
Performances of two training algorithm in prediction of apple bruise volume.

Model	Phase	Performance criterion			
		Training algorithm	MAPE (%)	RMSE (mm^3)	TSSE $(mm^3)^2$
$f(R,T,S,E)$	Train	BB	5.94	22.77	49757.79
		BDLRF	3.75	15.35	22608.13
	Test	BB	3.55	21.73	11337.70
		BDLRF	4.23	20.32	9905.07
$f(R,T,S,F)$	Train	BB	3.11	36.34	125430.17
		BDLRF	2.80	35.51	119789.88
	Test	BB	2.36	27.49	18891.77
		BDLRF	2.06	26.95	18152.63

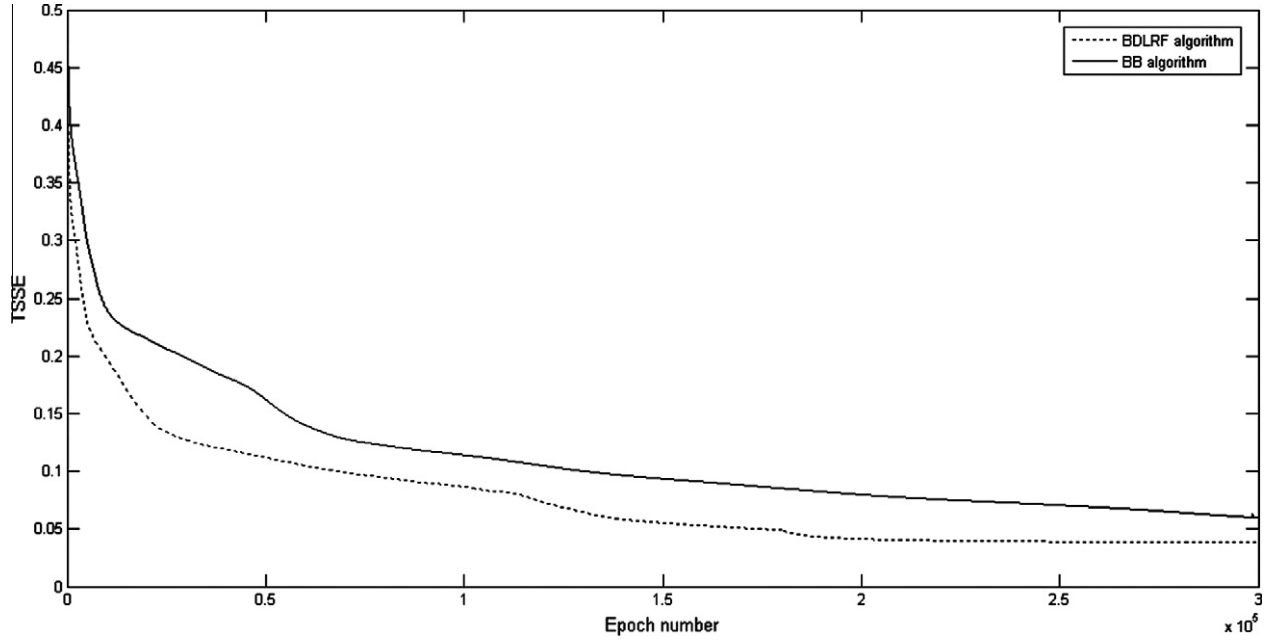

Fig. 12. Convergence diagrams of the MLP network obtained by the BB and the BDLRF algorithms for $f(R,T,S,F)$.

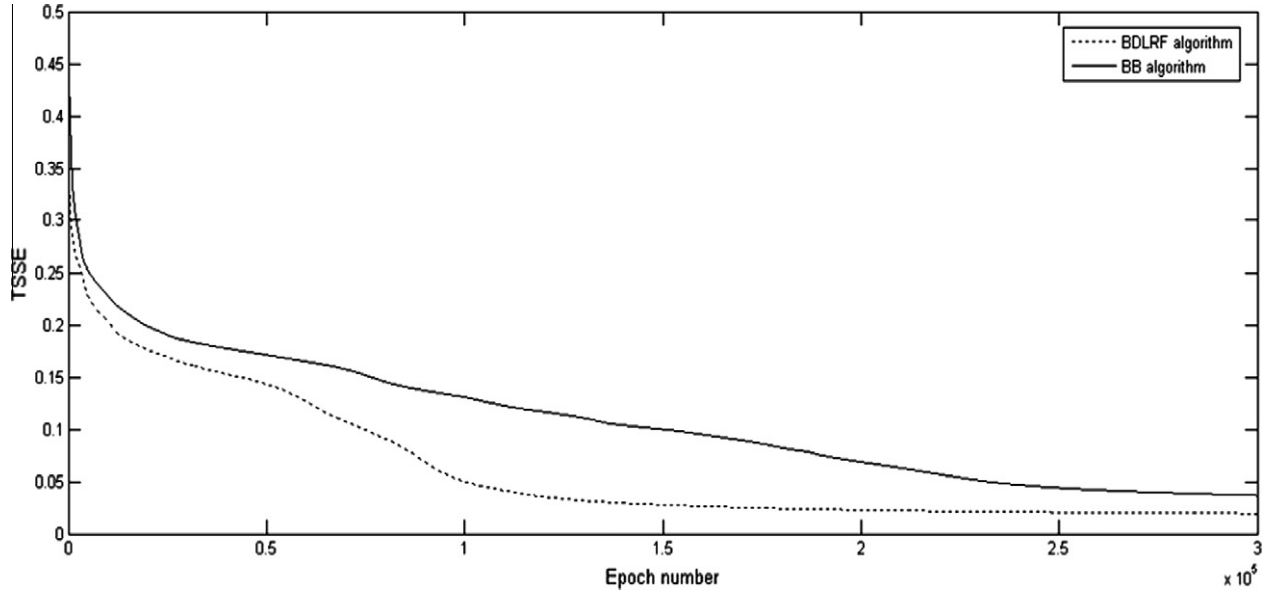

Fig. 13. Convergence diagrams of the MLP network obtained by the BB and the BDLRF algorithms for $f(R,T,S,E)$.

Table 8
Model of regression for bruise volume prediction.

Model	Regression equation	R^2
$f(R,T,S,F)$	$V = -3.25R + 0.97T - 4.5S + 22.9F - 0.23F \times R - 0.097F \times T$	0.93
$f(R,T,S,E)$	$V = -6.97R - 1.48T + 1.94S + 7186.95E + 5.97E \times R - 16.09E \times T$	0.97

Note: Minimum probability threshold $p \leqslant 0.05$.

Table 9
Statistical comparisons of desired and predicted data and the corresponding p values.

	Model type	Analysis types		
		Comparisons of means	Comparisons of variances	Comparisons of distribution
$f(R,T,S,E)$	MLP	0.995	0.997	1.000
	Regression	0.997	0.908	0.999
$f(R,T,S,F)$	MLP	0.992	0.948	1.000
	Regression	0.751	0.445	0.674

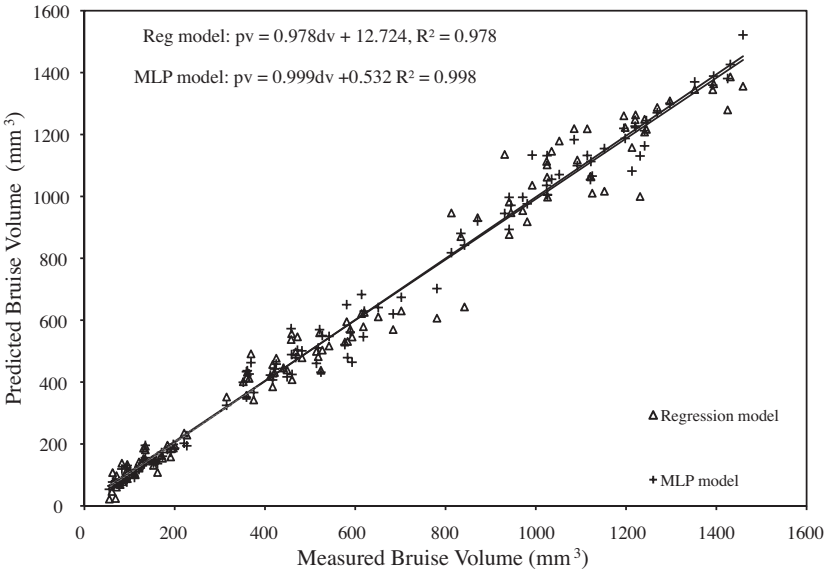

Fig. 14. Predicted values of BDLRF–MLP network and regression model versus measured values of bruise volume for $f(R,T,S,E)$.

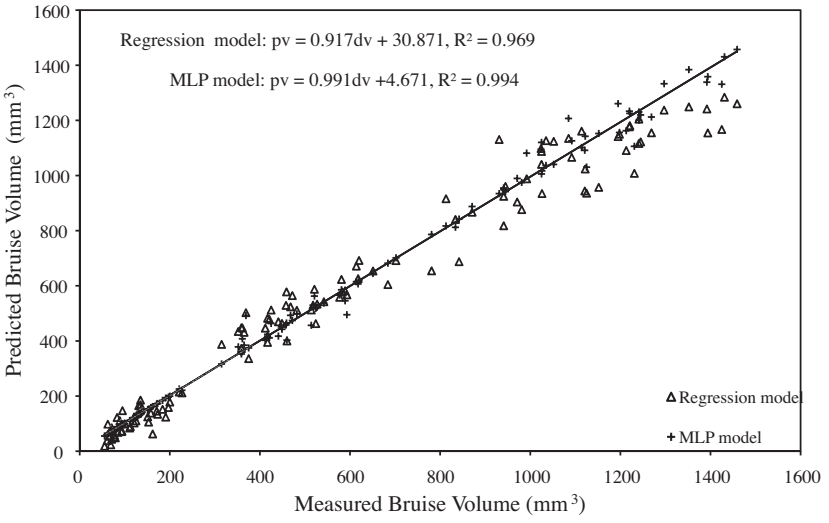

Fig. 15. Predicted values of BDLRF–MLP network and regression model versus measured values of bruise volume for $f(R,T,S,F)$.

range. Although the results of training phase were generally better than the test phase, the latter reveals the capability of neural network to predict the bruise volume with new data.

From statistical point of view, both desired and predicted test data have been analyzed to determine whether there are statistically significant differences between them. The null hypothesis assumes that statistical parameters of both series are equal. p value was used to check each hypothesis. Its threshold value was 0.05. If p value is greater than the threshold, the null hypothesis is then fulfilled. To check the differences between the data series, different tests were performed and p value was calculated for each case. The results are shown in Table 5. The Student's t-test was used to compare the means of both series. It was also assumed that the variance of both samples could be considered equal. The obtained p values were greater than the threshold, hence the null hypothesis cannot be rejected in all cases ($p > 0.94$). The variance was analyzed using the F-test. Here, a normal distribution of samples was assumed. Again, the p values confirm the null hypothesis in all cases ($p > 0.87$). Finally, the Kolmogorov–Smirnov test also confirmed the null hypothesis. From statistical point of view, both desired and predicted test data have a similar distribution for both of training algorithms ($p > 0.99$, Table 6).

Figs. 8–11 show the measured bruise volume versus the predicted ones. It is clear that the regression coefficients of determination between measured and predicted data ($R^2 > 0.9$) are high for the train data sets and test data sets. Since excellent estimation performances were obtained using the trained network, it demonstrates that the trained network was reliable, accurate and hence could be employed for bruise volume prediction. These figures reveal that the bruise volume predictions from BB training algorithm were not as good as fit to measured bruise volume in comparison to BDLRF bruise volume prediction. Comparisons of measured versus predicted bruise volume for BB training algorithm resulted in a least squares linear regression lines with slopes equal to BDLRF, while the BDLRF training algorithm resulted in a lines with y-intercepts lower than BB.

3.4. Comparison of training algorithms

For prediction of bruise volume, several networks with different settings and training algorithms were trained. The performances of the two training algorithm are shown in Table 7. For this specific case study, the comparison of results reveals that both algorithms are capable of generating accurate estimates within the preset range. It can be seen that MAPE, RMSE and TSSE values resulted by BDLRF are much less than or approximately equal those obtained by BB algorithm for training phase and test phase.

Figs. 12 and 13 demonstrates TSSE versus the epoch number (number of learning runs) for the BB and BDLRF algorithm applied. From these figures it can be concluded that the BDLRF training algorithm has achieved much better result, because it results in lower error.

It was quite clear that the BDLRF training algorithm achieved a much better performance than the BB training algorithm. Bearing all the results obtained by this study in mind, the advantages of the BDLRF training algorithm over BB are: faster convergence, lower training time and also it eases the process of parameter adjusting by decreasing the sensitivity to the parameters' values. The results also conforms the findings of Vakil-Baghmisheh and Pavešic (2001) and Rohani et al. (2011).

Table 10
Performances of two methods in prediction of bruise volume.

	Model type	Performance criterion		
		MAPE (%)	RMSE (mm^3)	TSSE $(mm^3)^2$
$f(R,T,S,E)$	MLP	3.84	16.46	22608.13
	Regression	97.31	63.93	490414.75
$f(R,T,S,F)$	MLP	2.65	33.90	137942.51
	Regression	97.31	81.31	788672.81

3.5. Comparison of regression model and MLP model

Any relationship, linear or nonlinear, can be learned and approximated by an ANN such as a three-layer MLP with sufficiently large number of neurons in the hidden layer. Another remarkable advantage of ANN is its capability of modeling the data of multiple inputs and multiple outputs. In contrast, the conventional regression techniques can only be used to learn the relationship between a single output and one or more inputs but cannot be used to model the data of multiple inputs and multiple outputs.

The results of a multiple linear regression analysis between bruise damage volume and series of independent variables (impact energy, contact force, temperature, acoustical stiffness and curvature radius) are presented in Table 8. All main factors in these models had significant effect at 5% probability level.

The analysis of statistical associated with MLP network employing the BDLRF training algorithm and regression model for prediction of bruise volume shown in Table 9. The p values confirm the null hypothesis in all cases ($p > 0.40$). Therefore, from statistical point of view, both measured and predicted bruise volume have a similar means, variances and distribution for both methods.

The plots of predicted bruise volume against measured bruise volume are depicted in Figs. 14 and 15. The results reveal a very good agreement between the predicted and the measured values of bruise volume ($R^2 > 0.9$). Also, these figures reveal that the bruise volume predictions from regression model were not as good as fit to measured bruise volume in comparison to MLP model bruise volume prediction. Comparisons of measured versus predicted bruise volume for MLP model resulted in a least squares linear regression lines with slopes almost equal to regression model, while the MLP model resulted in a lines with y-intercepts very lower than regression model.

Comparing the results generated using MLP network with those generated by the regression model (Table 10), it can be concluded that MLP model has a higher capability of producing accurate predictions in comparison to regression model, because the MLP model had a higher decrease of MAPE, RMSE and TSSE in comparison to regression model.

4. Conclusions

This article focused on the application of MLPNN to predict apple bruise volume. To show the applicability and superiority of the proposed approach, the measured data of apple bruise volume were used. To improve the output, the data were first preprocessed. MLP network was used and applied with the impact energy, contact force, curvature radius, temperature and acoustical stiffness as variable inputs. The network was trained using both BB and BDLRF learning algorithms. Statistical comparisons of measured and predicted test data were applied to the selected ANN. From statistical analysis, it was found that at 95% confidence level (with p-values greater than 0.9) both measured and predicted test data are similar. The results also revealed that, using BDLRF algorithm yields a better performance than BB algorithm. After testing all possible networks with the test data sets, it has been demonstrated that MLP network with 5-30-1 and 5-25-1 instruction and BDLRF algorithm had the best output for force model and energy model, respectively. It is also found that neural network is

particularly suitable for learning nonlinear functional relationships which are not known or cannot be specified.

Because the ANN does not assume any fixed form of dependence between the output and input values, unlike the regression methods, it seems to be more successful in the application under consideration. It could be said that the neural network provides a practical solution to the problem of estimating apple bruise volume in a fast, yet accurate and objective way. It is hoped that the analysis conducted in this article can provide reference for the choice of ANN in such area. Additional research on ANNs is required to make use of these networks more appealing and user-friendly to prediction of fruit bruise volume applications.

References

Abbott, J.A., Lu, R., 1996. Anisotropic mechanical properties of apples. Transactions of the ASAE 39, 1451–1459.

Al-Ohali, Y., 2010. Computer vision based date fruit grading system: design and implementation. Journal of King Saud University – Computer and Information Sciences 23, 29–36.

Barreiro, P., Steinmetz, V., Ruiz-Altisent, M., 1997. Neural bruise prediction models for fruit handling and machinery evaluation. Computers and Electronics in Agriculture 18, 91–103.

Effendi, Z., Ramli, R., Ghani, J.A., 2010. A back propagation neural networks for grading jatrophacurcas fruits maturity. American Journal of Applied Sciences 7 (3), 390–394.

Haykin, S., 1994. Neural Networks: A Comprehensive Foundation. McMillan College Publishing Company, New York.

Kavdir, I., Guyer, D.E., 2004. Comparison of artificial neural networks and statistical classifiers in apple sorting using textural features. Biosystems Engineering 89, 331–334.

Khoshhal, A., Alizadeh, A., Etemadi, A., Zereshki, S., 2010. Artificial neural network modeling of apple drying process. Journal of Food Process Engineering 33, 298–313.

Lin, X., Brusewitz, G.H., 1994. Peach bruise thresholds using the instrumented sphere. Applied Engineering in Agriculture 10, 509–513.

Mohsenin, N.N., 1986. Physical Properties of Plant and Animal Materials. Gordon and Breach science publishers, New York.

NeuroDimensions Inc., 2002. NeuroSolutions Tool for Excel.

Pang, W., Studman, C.J., Ward, G.T., 1992. Bruising damage in apple-to-apple impact. Journal of Agricultural Engineering Research 52, 229–240.

Ragni, L., Berardinelli, A., 2001. Mechanical behavior of apples, and damage during sorting and packaging. Journal of Agricultural Engineering Research 78, 273–279.

Rohani, A., Abbaspour-Fard, M.H., Abdolahpour, S., 2011. Prediction of tractor repair and maintenance costs using artificial neural network. Expert Systems with Applications 38, 8999–9007.

Roth, E., Kovacs, E., Hertog, M., Vanstreels, E., Nicolai, B., 2005. Relationship between physical and biochemical parameters in apple softening. In: Proceeding of the 5th International Postharvest Symposium (PS'05), ISHS, pp. 573–578.

Schotte, S., Belie, N., Baerdemaeker, J., 1999. Acoustic impulse technique for evaluation and modeling of firmness of tomato fruit. Postharvest Biology and Technology 17, 105–115.

Studman, C.J., Brown, G.K., Timm, E.J., Schulte, N.L., Vreede, M.J., 1997. Bruising on blush and nonblush sides in apple-to-apple impacts. Transactions of the ASAE 40, 1655–1663.

Timm, E.J., Brownand, G.K., Armstrong, P.R., 1996. Apple damage in bulk bins during semi-trailer transport. Applied Engineering in Agriculture 12, 369–377.

Vakil-Baghmisheh, M.T., 2002. Farsi Character Recognition Using Artificial Neural Networks. Ph.D. Thesis, Faculty of Electrical Engineering, University of Ljubljana.

Vakil-Baghmisheh, M.T., Pavešic, N., 2001. Back-propagation with declining learning rate. In: Proceeding of the 10th Electrotechnical and Computer Science Conference, Portorož, Slovenia, vol B, pp. 297–300.

Vakil-Baghmisheh, M.T., Pavešic, N., 2003. A fast simplified fuzzy ARTMAP network. Neural Processing Letters 17, 273–301.

Van Linden, V., Scheerlinck, N., Desmet, M., De Baerdemaeker, J., 2006. Factors that affect tomato bruise development as a result of mechanical impact. Postharvest Biology and Technology 42, 260–270.

Van Zeebroeck, M., Tijskens, E., Van Liedekerke, P., Deli, V., De BaerdeMaeker, J., Ramon, H., 2003. Determination of the dynamical behaviour of biological materials during impact using a pendulum device. Journal of Sound and Vibration 266, 465–480.

Van Zeebroeck, M., Van linden, V., Darius, P., de Ketelaere, B., Ramon, H., et al., 2007a. The effect of fruit properties on the bruise susceptibility of tomatoes. Postharvest Biology and Technology 45, 68–75.

Van Zeebroeck, M., Van linden, V., Darius, P., de Ketelaere, B., Ramon, H., et al., 2007b. The effect of fruit factors on the bruise susceptibility of apples. Postharvest Biology and Technology 46, 10–19.

Van Zeebroeck, M., Van Linden, V., Ramon, H., de Baerdemaeker, J., Nicola, B.M., et al., 2007c. Impact damage of apples during transport and handling. Postharvest Biology and Technology 45, 157–167.

Xiaobo, Z., Jiewen, Z., Yanxiao, L., 2007. Apple color grading based on organization feature parameters. Pattern Recognition Letters 28, 2046–2053.

Zhang, Y.F., Fuh, J.Y.H., 1998. A neural network approach for early cost estimation of packaging products. Computers & Industrial Engineering 34, 433–450.