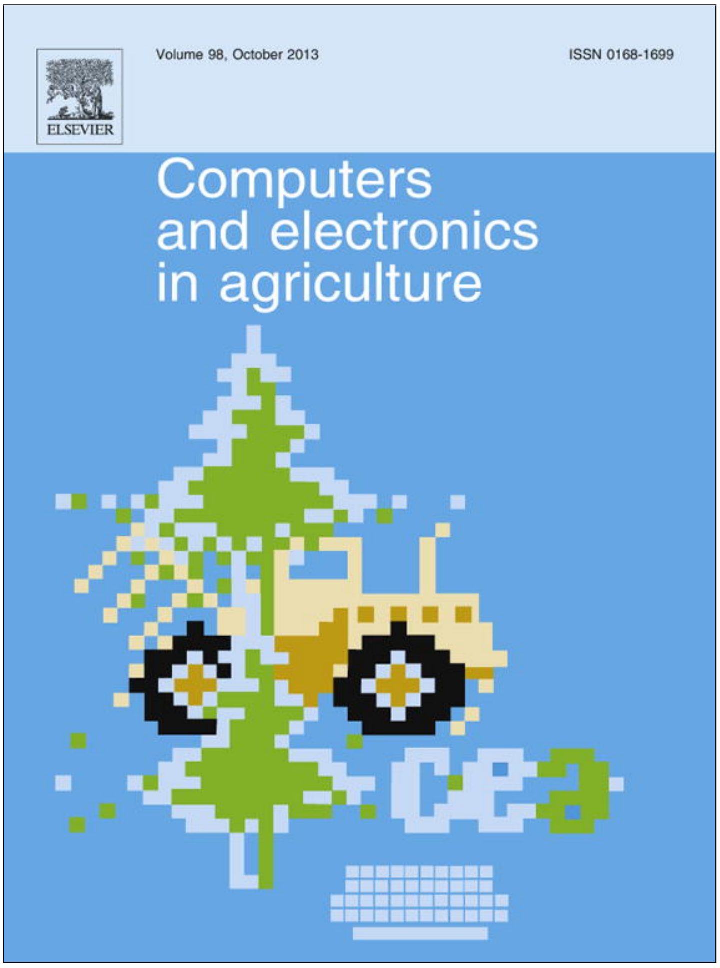
Volume 98, October 2013
ISSN 0168-1699
ELSEVIER
Computers
and electronics
in agriculture

Computers and Electronics in Agriculture 98 (2013) 1–7

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Predictions of viscoelastic behavior of pomegranate using artificial neural network and Maxwell model

Mohammad Hossien Saeidirad [a], Abbas Rohani [b,*], Saeed Zarifneshat [a]

[a] Khorasan Razavi Agriculture and Natural Resources Research Center, Mashhad, Iran
[b] Dept. of Farm Machinery Engineering, College of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

ARTICLE INFO

Article history:
Received 24 May 2012
Received in revised form 28 June 2013
Accepted 16 July 2013

Keywords:
Maxwell model
Stress relaxation
Pomegranate
Viscoelastic behavior
ANN
BDLRF algorithm

ABSTRACT

Stress relaxation is one of the defined tests to characterize the viscoelastic properties of food and agricultural materials. Stress relaxation data are very important because they provide useful and valuable information such as fruit firmness and ripening, food processing and predicting changes in the material during mechanical loading. Viscoelastic behavior of some varieties of pomegranate that are cultivated in Iran has been studied in current research. For this purpose, stress relaxation test was conducted with three cultivars of pomegranate (Ardestani, Shishekap and Malas) for three sizes (small, medium and large). In this article the potential of artificial neural network (ANN) technique is evaluated as an alternative method for Maxwell model to predict the viscoelastic behavior of pomegranate. Neural stress relaxation models were constructed to describe stress relaxation behavior of pomegranate with respect to time. The neural models were built based upon relaxation time as input network and stress relaxation as output network. The results revealed that both ANN model and Maxwell model have high capability of producing accurate and reliable predictions for stress.

1. Introduction

Most of the biological and solid foods behave as viscoelastic materials when they are exposed to small or intermediate levels of deformation. The mechanical response of viscoelastic materials is time-dependent; it depends on not only the current loading level but also the rate and/or history of loading (Kilcast, 2004).

In order to study viscoelastic behavior of food, scientists have utilized the knowledge of rheology which is the science of the deformation and flow of the matter. They have established the rheological modeling of these materials to characterize them and predict their behavior under various conditions of stress or strain. These viscoelastic models contain different combinations of Hookean solid elements (springs) and Newtonian fluid elements (dashpots), and show complex behavior of various food materials (Rao et al., 2005).

One of the defined tests to characterize the viscoelastic properties of food materials is the stress relaxation in which a constant strain is applied and the stress required to maintain the deformation is measured as a function of time. Stress relaxation data are very important because they provide useful and valuable information such as fruit firmness (Blahovec, 1996), food acceptability, food processing, and handling (Bourne, 2002), fruit ripening (Hassan et al., 2005), checking phenomenon (Kim and Okos, 1999), staling of cereal products (Limanond et al., 2002) and predicting changes in the material during mechanical harvesting or handling (Rao et al., 2005).

Hassan et al. (2005) studied the viscoelastic nature of eight date cultivars at their khalal (balah) and rutab stages of maturity by estimating their relaxation parameters from experimental stress relaxation data. Three popular stress relaxation models, namely the generalized Maxwell, Nussinovitch, and Peleg were fitted to experimental data. They pointed out that all three models were valid for quantifying the relaxation behavior of dates; however, the generalized Maxwell was the best in predicting experimental data. Del Nobile et al. (2007) chose five different food matrices including agar gel, meat, ripened cheese, 'mozzarella' cheese and white pan bread. Results showed that the proposed relaxation model perfectly fitted the experimental data. Moreover, they found a substantial difference between the relaxation times distribution curves of the investigated bulky and spongy foods.

Artificial neural networks (ANNs) are being used in a wide variety of applications, including apple bruise prediction (Zarifneshat et al., 2012), prediction of mechanical properties of cumin seed (Saiedirad and Mirsalehi, 2010), fruit grading (Al-Ohali, 2010; Effendi et al., 2010) and modeling of apple drying process (Khoshhal et al., 2010). Recently, several researchers suggested that the implicit constitutive modeling based on artificial neural networks can be used as alternative to traditional explicit constitutive modeling (Daoheng et al., 2000; Huber and Tsakmakis, 2001; Zhang and Friedrich, 2003; Al-Haik et al., 2004; Al-Haik et al., 2006). However,

* Corresponding author. Tel.: +98 9157005167.
E-mail address: abassrohani@yahoo.com (A. Rohani).

0168-1699/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.compag.2013.07.009

all cases were mentioned for the industrial materials. No items were found in the interpretation of viscoelastic behavior of agriculture crops using artificial neural networks.

The main advantages of using neural networks are learning directly from examples without attempting to estimate the statistical parameters. More generally, there is no need for firm assumptions about the statistical distributions of the inputs and generating any continuous nonlinear function of input (universal approximating). (Vakil-Baghmisheh, 2002; Gupta et al., 2003; Rohani et al., 2011; Zarifneshat et al., 2012). Because of these unique characteristics, it also can be employed for prediction of stress relaxation.

The main objective of this study was to develop pomegranate stress relaxation prediction models. The specific objectives were: (1) to investigate the effectiveness of ANN for predicting pomegranate stress relaxation; (2) select optimum ANN parameters for accurate prediction pomegranate stress relaxation; (3) compare the performance of generalized Maxwell model with ANN modeling predicting of pomegranate tress relaxation.

2. Materials and methods

2.1. Sample preparation

The experiment was conducted at Khorasan Razavi Agricultural and Natural Resources Research Centre, Mashhad (Northeast of Iran). For current research, three pomegranate cultivars including: Ardestani, Shishekap and Malaswere selected. Some properties of pomegranate cultivars obtained by sixty replicates are presented in Table 1. Fruits were hand harvested in October 2010 from an experimental orchard and stored for 24 h at 20 °C and 60% relative humidity.

2.2. Stress relaxation test

All stress relaxation tests were carried out by means of Texture Analyzer (CNS Farnell, Model: QTS 25, Hertfordshire, UK) equipped with a 150 mm diameter cylindrical flat disk. Fruits were exposed under uniaxial compression at crosshead speed of 30 mm min^{-1} while the stem-calyx axis was horizontal. The compressive strain imposed on samples was 8%. When the desired strain was reached, the disk plunger was held for 60 s (Blahovec, 1996) and stress required to maintain the deformation was observed as a function of time. Texture analyzer software (TexturePro v2.0, Hertfordshire, UK) collected data and transferred to a computer. All experiments have been carried out in twenty replicates at room temperature (ASAE, 1999).

2.3. Maxwell model

The model most often used to represent stress relaxation behavior of food materials is Maxwell model consisting of a finite number of Maxwell elements (a Hookean spring and a Newtonian dashpot in series) in parallel with each other. In order to consider the equilibrium stress of viscoelastic materials, a spring is added in parallel with the other Maxwell elements. This modified model, which is called generalized Maxwell model, can better describe the stress relaxation of viscoelastic materials since the imposed stress is not alleviated after long periods of time (Mohsenin, 1970). If a generalized Maxwell model is subjected to a fixed strain, the total stress of the system is the sum of the stresses in each Maxwell element that has a different relaxation time (Steffe, 1996). The mathematical representation of this model is as follows:

$$\sigma(t) = \sigma(0) + \sum_{i=1}^{n} C_i \exp\left(\frac{-t}{\tau_i}\right) \quad (1)$$

where $\sigma(t)$ is the magnitude of stress at time t, $\sigma(0)$ represents the residual stress, and c_i and τ_i are the constants of the ith observation. MATLAB (2010) software (Curve Fitting Toolbox 3.0) was utilized to calculate the constants of Maxwell model.

Table 1
Properties of pomegranate cultivars.

Cultivar	Weight mean (g)	Density mean (g/cc)	Geometrical diameter mean (mm)
Ardestani	298.56	0.92	78.75
Shishekap	217.01	1.29	74.95
Malase	250.29	1.2	76.29

2.4. Data preprocessing

Based on these available data, the time (t) was selected as variable input. The stress (Mpa) 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

$$x_n = \frac{\log(x) - \log(x_{\min})}{\log(x_{\max}) - \log(x_{\min})} \times (r_{\max} - r_{\min}) + r_{\min} \quad (2)$$

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 1121 patterns (70% of total patterns) and 480 patterns (30% 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 (Neuro Dimensions Inc., 2002). However, by adding new data to the training samples, the network then can be retrained.

2.5. The multilayer perceptron neural network

Among various ANN models, Multilayer Perceptron (MLP) has maximum practical importance. MLP is a feed-forward layered network with one input layer, one output layer, and some hidden layers. Fig. 1 shows a MLP with one hidden layer. Every node computes a weighted sum of its inputs and passes the sum through a soft nonlinearity. The soft nonlinearity or activity function of neurons should be non-decreasing and differentiable. The most popular function is unipolar sigmoid (Rohani et al., 2011):

$$f(\theta) = \frac{1}{1 + e^{-\theta}} \quad (3)$$

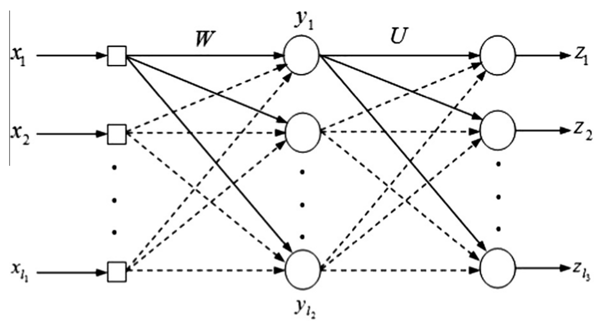

Fig. 1. Configuration of the MLP with one hidden layer (Vakil-Baghmisheh, 2002).

The network will answer through the vector Z^q in its output by inserting X^q as input vector (for $q = 1,\ldots,Q$). The aim is to adapt the parameters of the network in order to bring the actual output Z^q close to corresponding desired output d^q (for $q = 1,\ldots,Q$). The most popular method of MLP training is the back-propagation algorithm, and in literatures there exist many variants of this algorithm.

This algorithm is based on minimization of a suitable error cost function. 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. 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 (Vakil-Baghmisheh and Pavešic, 2001; Rohani et al., 2011; Rohani and Makarian, 2011; Zarifneshat et al., 2012). A computer code was also developed in MATLAB software to implement these ANN models.

2.5.1. BDLRF algorithm

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

$$TSSE = \sum_q E_q \tag{4}$$

$$E_q = \sum_k (d_k^q - z_k^q)^2 \quad for(q = 1,\ldots,Q) \tag{5}$$

where d_k^q and z_k^q are the kth components of desired and actual output vectors of the qth input, respectively. Network learning happens in two phases: forward pass and backward pass. In forward pass an input vector is inserted to the network and the network outputs are computed by proceeding forward through the network, layer by layer:

$$\begin{cases} net_j = \sum_i x_i w_{ij} \\ y_j = \frac{1}{1+e^{-net_j}} \end{cases}, \quad j = 1,\ldots,l_2 \tag{6}$$

$$\begin{cases} net_k = \sum_j y_j u_{jk} \\ z_k = \frac{1}{1+e^{-net_k}} \end{cases}, \quad k = 1,\ldots,l_3 \tag{7}$$

where w_{ij} is the connection weight between nodes i and j, and u_{jk} is the connection weight between nodes j and k; 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. In backward pass the error gradients versus weight values, i.e. $\frac{\partial E}{\partial w_{ij}}$ (for $i = 1,\ldots,l_1$, $j = 1,\ldots,l_2$) and $\frac{\partial E}{\partial u_{jk}}$ (for $j = 1,\ldots,l_2$, $k = 1,\ldots,l_3$), are computed layer by layer starting from the output layer and proceeding backwards. The connection weights between nodes of different layers are updated using the following equations:

$$u_{jk}(n+1) = u_{jk}(n) - \eta \times \frac{\partial E}{\partial u_{jk}} + \alpha(u_{jk}(n) - u_{jk}(n-1)) \tag{8}$$

$$w_{ij}(n+1) = w_{ij}(n) - \eta \times \frac{\partial E}{\partial w_{ij}} + \alpha(w_{ij}(n) - w_{ij}(n-1)) \tag{9}$$

where η is the learning rate adjusted between 0 and 1, α is the momentum factor at interval [0, 1]. Momentum factor is used to speed up the convergence. 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 number of input and output nodes is determined by functional requirements of the ANN.

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, for every T epoch ($3 \leqslant T \leqslant 5$) these values are decreased monotonically by means of arithmetic progression, until they reach to x% (equals to 5) of their initial values. η (and similarly α) was decreased using the following equations:

$$m = \frac{Q - n_1}{T} \tag{10}$$

$$\eta_n = \eta_\circ + n\eta_\circ \frac{x-1}{m} \tag{11}$$

where m,n_1, η_n and $\eta_\circ$ are the total number of arithmetic progression terms, the start point of BDLRF, the learning rate in nth term of arithmetic progression, and the initial learning rate, respectively. The details could be seen in Vakil-Baghmisheh and Pavešic (2003).

2.6. 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:

$$RMSE = \sqrt{\frac{\sum_{j=1}^{n}\sum_{i=1}^{m}(d_{ji} - p_{ji})^2}{nm}} \tag{12}$$

$$R^2 = \frac{\left(\sum_{j=1}^{n}(d_j - \bar{d})(p_j - \bar{p})\right)^2}{\sum_{j=1}^{n}(d_j - \bar{d})^2 \quad . \quad \sum_{j=1}^{n}(p_j - \bar{p})^2} \tag{13}$$

$$TSSE = \sum_{j=1}^{n}(d_j - p_j)^2 \tag{14}$$

$$MAPE = \frac{1}{nm}\sum_{j=1}^{n}\sum_{i=1}^{m}\left|\frac{d_{ji} - p_{ji}}{d_{ji}}\right| \times 100 \tag{15}$$

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 relationship between stress and time (Fig. 2).

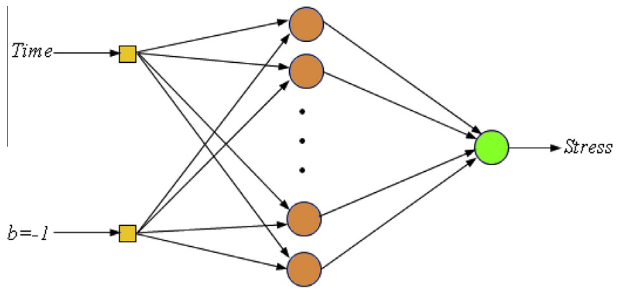

Fig. 2. Diagram of multilayer neural network (including input and output) used in the prediction of stress.

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. MLP model with 5 neurons in the hidden layer seems to be appropriate for modeling stress.

3.2. Learning rate and momentum term

In order to speed up convergence, an extra term called momentum (α) is used to the weights update (Vakil-Baghmisheh, 2002; Gupta et al., 2003; Rohani et al., 2011). The learning rate and momentum factors are only used in the learning process, so the criteria used to optimize them are based on the learning error and the iteration number. When the optimal topology of the neural network was found, the learning rate (η) and momentum term (α) was also optimized throughout a trial–error method. The learning rate and momentum factors have interactive impacts on network training. This makes parameter tuning a difficult task where momentum term is added. It is observed that the error value is increased and the convergence speed of the learning process is decreased when the momentum 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. It is also evident that, the convergence speed of the learning process was improved through an appropriate choice of parameters η and α. In order to improve the behavior of MLP during training, and due to simplicity of adjusting process of network parameters, we used BDLRF algorithm. The results obtained by BDLRF have shown that the best performance of MLP was obtained via a constant momentum term equals to 0.95. The results also conforms the findings of Rohani et al. (2011).

Therefore, when the convergence was slowed down, a point was chosen and η was only decreased using Eq. (11). Table 2 shows the parameters of optimum BDLRF-MLP. For the selected topology, several learning processes were performed with different coefficients, ranged from 0.5 to 0.99 and 0.5 to 0.99 for learning rate and momentum term, respectively.

Table 2
Optimum parameters of neural network (BDLRF-MLP).

Data set		Parameters of neural network						
		The start point of BDLRF	First phase		Second phase		Epoch	Topology
Cultivar	Size		H	α	η	α		
Ardestani	Large	100	0.8	0.95	0.04	0.95	5000	2-5-1
	Small & medium	1000	0.8	0.95	0.04	0.95	5000	2-5-1
Shishekap	Large	1000	0.8	0.95	0.04	0.95	15,000	2-5-1
	Medium	1000	0.9	0.95	0.05	0.95	15,000	2-5-1
	Small	1000	0.8	0.95	0.04	0.95	10,000	2-5-1
Malas	Large	1000	0.8	0.95	0.04	0.95	10,000	2-5-1
	Small & medium	1000	0.8	0.95	0.04	0.95	5000	2-5-1

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. Some statistical properties of the sample data used for training process and the

Table 3
Statistical variables of desired and predicted values in training phase (BDLRF-MLP).

Data set		Values	Statistical values						
Cultivar	Size		Average	Standard deviation	Minimum	Maximum	Kurtosis	Skewness	Sum
Ardestani	Large	Desired	0.076	0.005	0.071	0.099	6.359	1.708	84.723
		Predicted	0.076	0.005	0.71	0.097	6.240	1.698	84.723
	Medium	Desired	0.079	0.006	0.073	0.110	7.829	1.934	88.297
		Predicted	0.079	0.005	0.074	0.110	7.857	1.941	88.300
	Small	Desired	0.114	0.007	0.107	0.159	9.979	2.237	127.526
		Predicted	0.114	0.007	0.107	0.159	10.010	2.240	127.527
Shishekap	Large	Desired	0.099	0.007	0.092	0.137	8.481	2.062	111.358
		Predicted	0.099	0.007	0.093	0.137	8.445	2.060	111.358
	Medium	Desired	0.093	0.006	0.087	0.125	7.612	1.895	104.689
		Predicted	0.093	0.006	0.087	0.123	7.546	1.888	104.689
	Small	Desired	0.140	0.009	0.131	0.188	7.123	1.802	157.064
		Predicted	0.140	0.009	0.131	0.188	7.035	1.795	157.064
Malas	Large	Desired	0.057	0.004	0.053	0.075	5.877	1.598	64.213
		Predicted	0.057	0.004	0.053	0.074	5.866	1.593	64.211
	Medium	Desired	0.059	0.004	0.055	0.081	6.390	1.646	66.165
		Predicted	0.059	0.004	0.055	0.081	6.397	1.651	66.166
	Small	Desired	0.067	0.005	0.062	0.091	7.393	1.884	74.670
		Predicted	0.067	0.005	0.062	0.091	7.341	1.889	74.674

Table 4
Statistical variables of desired and predicted values in test phase (BDLRF-MLP).

Data set		Values	Statistical values						
Cultivar	Size		Average	Standard deviation	Minimum	Maximum	Kurtosis	Skewness	Sum
Ardestani	Large	Desired	0.076	0.005	0.071	0.097	5.334	1.599	36.359
		Predicted	0.076	0.005	0.71	0.096	5.458	1.618	36.359
	Medium	Desired	0.079	0.006	0.073	0.112	8.781	2.207	37.940
		Predicted	0.079	0.006	0.074	0.115	9.419	2.283	37.939
	Small	Desired	0.115	0.008	0.107	0.158	9.165	2.194	55.012
		Predicted	0.115	0.008	0.107	0.157	9.180	2.186	54.995
Shishekap	Large	Desired	0.100	0.007	0.092	0.135	8.085	2.001	47.834
		Predicted	0.100	0.007	0.093	0.133	7.917	1.983	47.847
	Medium	Desired	0.094	0.007	0.087	0.129	8.605	2.189	44.949
		Predicted	0.094	0.007	0.087	0.126	8.142	2.139	44.934
	Small	Desired	0.141	0.010	0.131	0.189	6.469	1.754	67.861
		Predicted	0.141	0.010	0.131	0.184	6.196	1.714	67.866
Malas	Large	Desired	0.057	0.004	0.053	0.075	6.819	1.722	27.376
		Predicted	0.057	0.004	0.05	0.075	6.646	1.688	27.368
	Medium	Desired	0.059	0.004	0.055	0.079	6.067	1.598	28.427
		Predicted	0.059	0.004	0.055	0.079	6.159	1.608	28.428
	Small	Desired	0.066	0.005	0.062	0.089	8.068	2.002	31.917
		Predicted	0.066	0.005	0.062	0.089	8.349	2.053	31.917

Table 5
Statistical comparisons of desired and predicted data and the corresponding p values in test phase.

Data set		Analysis types		
Cultivar	Size	Comparisons of means	Comparisons of variances	Comparisons of distribution
Ardestani	Large	0.996	1.000	0.335
	Medium	0.998	0.931	0.791
	Small	0.944	0.781	0.839
Shishekap	Large	0.953	0.951	0.685
	Medium	0.944	0.877	0.378
	Small	0.987	0.828	0.630
Malas	Large	0.941	0.876	0.791
	Medium	0.996	0.952	0.949
	Small	0.986	0.916	0.791

prediction values associated with different training algorithms are shown in Table 3. Considering the average values of standard deviation, kurtosis, skewnese and sum, 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

In test phase, we used the selected topology with the previously adjusted weights. The objective of this step was to test the network generalization property and to evaluate the competence of the trained network. Therefore, the network was evaluated by data, outside the training set. Table 4 shows some statistical properties of the data used in test phase and the corresponding prediction values. Although these data are completely new for the MLP model, it can be seen that the differences of statistical values between the measured and predicted data in test phase negligible, 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 range.

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 so called 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

Table 6
The constants of generalized Maxwell model for three pomegranate cultivars.

Data set		σ_0	C_1	C_2	τ_1 (s)	τ_2 (s)	R^2	TSSE
Cultivar	Size							
Ardestani	Large	0.069	0.015	0.012	23.68	2.02	0.998	0.0001
	Medium	0.072	0.016	0.019	23.58	1.72	0.996	0.0002
	Small	0.105	0.021	0.029	26.29	1.54	0.996	0.0003
Shishekap	Large	0.089	0.019	0.025	31.73	1.90	0.997	0.0002
	Medium	0.084	0.018	0.022	34.51	2.12	0.997	0.0002
	Small	0.125	0.029	0.029	37.15	2.38	0.997	0.0004
Malas	Large	0.052	0.010	0.012	1.85	28.86	0.998	0.0000
	Medium	0.052	0.012	0.013	1.95	33.71	0.998	0.0000
	Small	0.060	0.015	0.013	2.04	30.42	0.997	0.0001

Table 7
Performances of two generalized Maxwell model and MLP model in prediction of stress.

Data set		Model type	Performance criterion	
Cultivar	Size		MAPE (%)	RMSE(MPa^3)
Ardestani	Large	MLP	0.2222	0.0002
		Maxwell	0.2193	0.0002
	Medium	MLP	0.2435	0.0003
		Maxwell	0.2414	0.0003
	Small	MLP	0.2423	0.0005
		Maxwell	0.2277	0.0005
Shishekap	Large	MLP	0.2558	0.0004
		Maxwell	0.2644	0.0004
	Medium	MLP	0.2582	0.0004
		Maxwell	0.2624	0.0004
	Small	MLP	0.2193	0.0005
		Maxwell	0.1983	0.0005
Malas	Large	MLP	0.2313	0.0002
		Maxwell	0.1982	0.0002
	Medium	MLP	0.1595	0.0001
		Maxwell	0.1711	0.0002
	Small	MLP	0.2717	0.0003
		Maxwell	0.2211	0.0003

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 ($p > 0.78$). 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 ($p > 0.33$).

3.4. Generalized Maxwell model

In order to find the optimum number of the adjustable model parameters, preliminary statistical tests were utilized. Peleg (1979) pointed out that the number of terms for food materials is usually 2–3 which can be varied independently and need to be compared simultaneously among samples. Consequently, the comparison of more number of terms is difficult so it should be kept to a minimum (Peleg, 1979; Bellido and Hatcher, 2009a). The results revealed that two terms is the proper for the current research because it had the higher R^2, RMSE and MAPE.

Table 6 presents the results of the fitting analysis to data for the generalized Maxwell model using MATLAB curve fitting toolbox. It can be seen that the generalized Maxwell model could properly predict the stress relaxation behavior of the pomegranate. The correlation coefficient (R^2) of the Maxwell model for all pomegranate cultivars was larger than 0.99 and TSSE ranged from 0.0000 to 0.0004.

3.5. Comparison of generalized Maxwell model and MLP model

The performances of the two generalized Maxwell model and MLP model are shown in Table 7. For this specific case study, the comparison of results reveals that both models are capable of generating the accurate estimations within the preset range. It can be seen that MAPE and RMSE values resulted by generalized Maxwell model are approximately equal to MLP model. Because the MLP model and Maxwell model had a very low amount of MAPE and RMSE, it can be concluded that both models have high capable of producing the accurate predictions for stress.

The analysis of statistical associated with MLP network employing the BDLRF training algorithm and generalized Maxwell model for prediction of stress shown in Table 8. The *p* values confirm the null hypothesis in all cases ($p > 0.053$), except in four cases

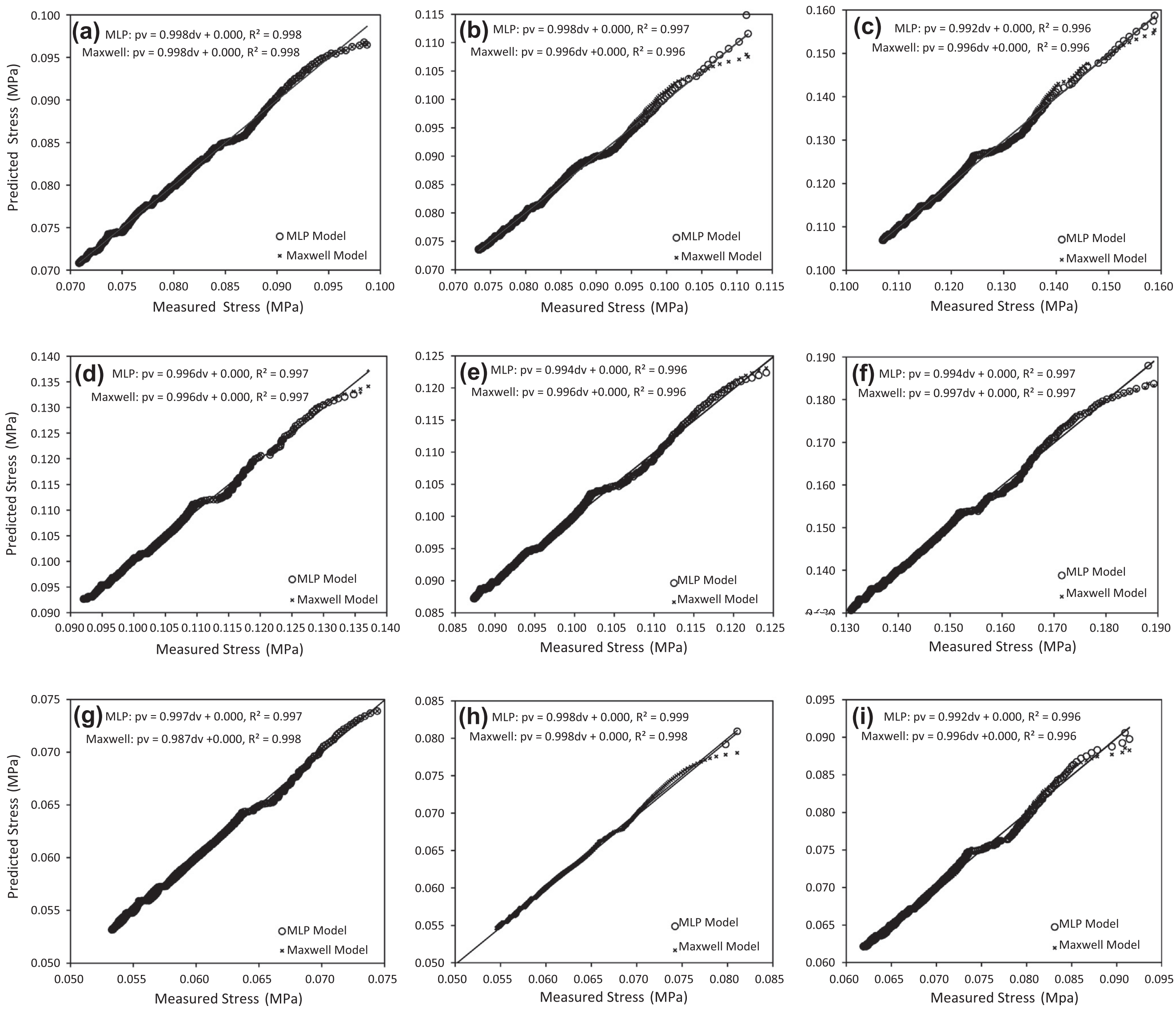

Fig. 3. Predicted values of MLP model and generalized Maxwell model versus measured values of stress forpomegranate cultivars. (a) Large Ardestani. (b) Medium Ardestani. (c) Small Ardestani. (d) Large Shishekap. (e) MmediumShishekap and (f) Small Shishekap. (g) Large Malas. (h) Medium Malas and (i) Small Malas.

Table 8
Statistical comparisons of desired and predicted data and the corresponding p values.

Data set		Model type	Analysis types		
			Comparisons of means	Comparisons of variances	Comparisons of distribution
Ardestani	Large	MLP	0.996	0.964	0.053
		Maxwell	1.000	0.965	0.058
	Medium	MLP	0.993	0.999	0.129
		Maxwell	0.979	0.952	0.035*
	Small	MLP	0.970	0.813	0.151
		Maxwell	0.921	0.944	0.032*
Shishekap	Large	MLP	0.974	0.934	0.083
		Maxwell	0.998	0.950	0.003*
	Medium	MLP	0.965	0.876	0.035*
		Maxwell	0.996	0.940	0.043
	Small	MLP	0.992	0.856	0.129
		Maxwell	0.997	0.943	0.177
Malas	Large	MLP	0.963	0.931	0.223
		Maxwell	0.901	0.644	0.436
	Medium	MLP	0.993	0.953	0.317
		Maxwell	1.000	0.974	0.410
	Small	MLP	0.995	0.878	0.277
		Maxwell	1.000	0.947	0.753

including comparisons of distribution. Therefore, from statistical point of view, both measured and predicted stress has a similar means, variances and distribution for both of methods.

The plots of predicted stress against measured stress are depicted in Fig. 3.The results reveal a very good agreement between the predicted and the measured values of stress ($R^2 > 0.99$). Also, these figures reveal that the stress predictions from Maxwell model were as good as fit to measured stress in comparison to MLP model stress prediction. Comparisons of measured versus predicted stress for MLP model resulted in a least squares linear regression lines with slopes (approximately equal to 1) and y-intercepts (approximately equal to 0) almost equal to Maxwell model.

4. Conclusions

This article focused on the application of MLPNN to predict viscoelastic behavior of pomegranate. To show the applicability and superiority of the proposed approach, the measured data of pomegranate stress test were used. To improve the output, the data were first preprocessed. MLP network was used and applied with the time as variable input. The network trained by BDLRF learning algorithm. 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. After testing all possible networks with the test data sets, it has been demonstrated that MLP network with 2-5-1 instruction had the best output for stress model. 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 dependency in 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 stress 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 stress applications.

References

Al-Haik, M.S., Garmestani, H., Savran, A., 2004. Explicit and implicit viscoplastic models for polymeric composite. International Journal of Plasticity 20 (10), 1875–1907.

Al-Haik, M.S., Hussaini, M.Y., Garmestani, H., 2006. Prediction of nonlinear viscoelastic behavior of polymeric composites using an artificial neural network. International Journalof Plasticity 22, 1367–1392.

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.

ASAE, 1999. Compression test of food materials of convex shape. ASAE Standard S368.3.PP570575.

Bellido, G.G., Hatcher, D.W., 2009. Asian noodles: revisiting Peleg's analysis for presenting stress relaxation data in soft solid foods. Journal of Food Engineering. 92, 29–36.

Blahovec, J., 1996. Stress relaxation in cherry fruit. Biorheology 33, 451–462.

Bourne, M.C., 2002. Food Texture and Viscosity: Concept and Measurement, second ed. Academic Press, London.

Daoheng, S., Qiao, H., Hao, X., 2000. A neuro computing model for the elasto plasticity. Computer Methods in Applied Mechanics and Engineering. 182, 177–186.

Del Nobile, M.A., Chillo, S., Mentana, A., Baiano, A., 2007. Use of the generalized Maxwell model for describing the stress relaxation behavior of solid-like foods. Journal of Food Engineering. 78, 978–983.

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.

Gupta, M.M., Jin, J., Homma, N., 2003. Static and Dynamic Neural Networks: From Fundamentals to Advanced Theory. John Wiley & Sons, Inc., Hoboken, New Jersey.

Hassan, B.H., Alhamdan, A.M., Elansari, A.M., 2005. Stress relaxation of dates at khalal and rutab stages of maturity. Journal of Food Engineering. 6, 439–445.

Haykin, S., 1994. Neural networks: A comprehensive foundation. McMillan College Publishing Company, New York.

Huber, N., Tsakmakis, Ch., 2001. A neural network tool for identifying the material parameters of a finite deformation viscoplasticity model with static recovery. Computer Methods in Applied Mechanics and Engineering. 191, 353–384.

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.

Kilcast, D., 2004. Texture in Food (Volume 2: Solid Foods), 1st ed. Woodhead Publishing Limited, Abington.

Kim, M.H., Okos, M.R., 1999. Some physical, mechanical, and transport properties of crackers related to the checking phenomenon. Journal of Food Engineering 40, 189–198.

Limanond, B., Castell-Perez, M.E., Moreira, R.G., 2002. Modeling the kinetics of corn tortilla staling using stress relaxation data. Journal of Food Engineering 53, 237–247.

MATLAB, 2010. MATLAB Release 2010a. The Math Works, Inc. <http://www.mathworks.com>.

Mohsenin, N.N., 1970. Physical Properties of Plant and Animal Materials, vol. 1. Gordon and Breach Science Publishers, New York.

Neuro Dimensions Inc., 2002. Neuro Solutions Tool for Excel.

Peleg, M., 1979. Characterization of the stress relaxation curves foods. Journal of Food Science 44, 277–281.

Rao, M.A., Rizvi, S.S.H., Datta, A.K., 2005. Engineering Properties of Foods, third ed. Taylor & Francis Group, USA.

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

Rohani, A., Makarian, H., 2011. Making weed management maps by artificial neural networks for using in precision agriculture. Journal of Agricultural Machinery Engineering 1 (2), 74–83.

Saiedirad, M.H., Mirsalehi, M., 2010. Prediction of mechanical properties of cumin seed using artificial neural networks. Journal of Texture Studies 41 (1), 34–48.

Steffe, J.F., 1996. Rheological Methods in Food Process Engineering, second ed. Freeman Press, East Lansing, pp. 299–304.

Vakil-Baghmisheh, M.T., 2002. Farsi Character Recognition Using artificial neural networks. PhD Thesis, Faculty of Electrical Engineering, University of Ljubljana.

Vakil-Baghmisheh, M.T., Pavešic, N., 2001. Back-propagation with declining learning rate. Proceeding of the 10thElectrotechnical 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.

Zarifneshat, S., Rohani, A., Ghassemzadeh, H.R., Sadeghi, M., Ahmadi, E., Zarifneshat, M., 2012. Predictions of apple bruise volume using artificial neural network. Computers and Electronics in Agriculture 82, 75–86.

Zhang, Z., Friedrich, K., 2003. Artificial neural networks applied to polymer composites: a review. Composites Science and Technology 63, 2029–2044.