DOI: 10.29050/harranziraat.1572534
Harran Tarım ve Gıda Bilimleri Derg. 2025, 29(2): 237-254

Research Article/Araştırma Makalesi

Modeling quality changes in heat-processed orange juice: a comparative study of artificial neural network and multiple linear regression approaches

Isıl işlem uygulanmış portakal suyunda kalite değişimlerinin modellenmesi: yapay sinir ağı ve çoklu doğrusal regresyon yaklaşımlarının karşılaştırmalı bir çalışması

Asiye AKYILDIZ[1], Tuba ŞİMŞEK MERTOGLU[2], Nuray İNAN ÇINKIR[3*], Erdal AĞCAM[4]

[1,2,4]Cukurova University Agriculture Faculty Department of Food Engineering, Adana/Türkiye
[3]Osmaniye Korkut Ata University Kadirli Applied Sciences Faculty [2]Department of Food Technology, Osmaniye/Türkiye

[1]https://orcid.org/0000-0001-5584-0849; [2]https://orcid.org/0000-0002-9694-211X; [3]https://orcid.org/0000-0002-8878-6794; [4]https://orcid.org/0000-0002-2677-2020

To cite this article:
Akyildiz, A., Simsek Mertoglu, T., Inan Çınkır, N. & Agcam, E. (2025). Modeling quality changes in heat-processed orange juice: a comparative study of artificial neural network and multiple linear regression approaches. Harran Tarım ve Gıda Bilimleri Dergisi, 29(2): 237-254

DOI: 10.29050/harranziraat.1572534

***Address for Correspondence:**
Nuray INAN ÇINKIR
e-mail:
nurayinan@osmaniye.edu.tr

Received Date:
23.10.2024
Accepted Date:
19.04.2025

ABSTRACT

The purpose of the study was to assess the prediction ability of multiple linear regression (MLR) and artificial neural network (ANN) models for the browning index, total carotenoid content, 5-hydroxymethylfurfural (HMF) and ascorbic acid of orange juice during storage after heat processing. For ANN models, the effect of neuron number of the hidden layer, epoch number, training algorithms and transfer functions are investigated using the trial-error method for selecting best design ANN models. The different methods (stepwise and enter) in multiple linear regression models were performed for detecting impact of independent variables on dependent variables. The performance of ANN and MLR models was determined through unseen data by means of statistical analysis. Regarding performance indices of ANN models for test data, overall R and R^2 were recorded as follows: 0.92 and 0.84 (browning index), 0.99 and 0.98 (HMF), 0.92 and 0.86 (ascorbic acid), 0.97 and 0.94 (total carotenoid content), respectively. R and R^2 values of MLR models for test data were 0.79 and 0.68 (browning index), 0.94 and 0.88 (HMF), 0.92 and 0.85 (ascorbic acid), 0.93 and 0.90 (total carotenoid content), respectively. Both models provided accurate predictions. However, the superior predictive power of ANN models is that they can learn directly from examples without calculating the parameters using statistical techniques. The results revealed that ANN models showed greater potential with high R and R^2 value, and the lowest error values when compared with the MLR model, but both ANN and MLR had almost same performance for prediction of ascorbic acid.

Key Words: Artificial neural network, Multiple linear regression, Orange juice, Color, modelling

ÖZ

Çalışmanın amacı, ısıl işlemden sonra depolama sırasında portakal suyunun esmerleşme indeksi, 5-hidroksimetilfurfural (HMF), askorbik asit ve toplam karotenoid içeriği için çoklu doğrusal regresyon (MLR) ve yapay sinir ağı (ANN) modellerinin tahmin yeteneğini değerlendirmektir. En iyi tasarım YSA modellerini seçmek için gizli katmandaki nöron sayısının, epok sayısının, eğitim algoritmalarının ve transfer fonksiyonlarının etkisi deneme-yanılma yöntemi kullanılarak araştırılmıştır. Tüm bağımsız değişkenlerin bağımlı değişkenler üzerindeki etkisini tespit etmek için çoklu doğrusal regresyon modellerindeki farklı yöntemler (stepwise ve enter) uygulanmıştır. ANN ve MLR modellerinin performansı istatistiksel analiz yoluyla test verileri kullanılarak değerlendirilmiştir. Test verileri için ANN modellerinin performans endekslerine ilişkin olarak, genel R ve R^2 sırasıyla; 0,92 ve 0,84 (esmerleşme indeksi), 0,99 ve 0,98 (HMF), 0,92 ve 0,86 (askorbik asit), 0,97 ve 0,94 (toplam karotenoid içeriği) belirlenmiştir. Test verileri için MLR modellerinin R ve R^2 değerleri sırasıyla 0,79 ve 0,68 (esmerleşme indeksi), 0,94 ve 0,88 (HMF), 0,92 ve 0,85 (askorbik asit), 0,93 ve 0,90 (toplam karotenoid içeriği) olarak tespit edilmiştir. YSA modellerinin üstün tahmin gücü, istatistiksel teknikler kullanarak parametreleri hesaplamaya gerek kalmadan doğrudan örneklerden öğrenebilmeleridir. Sonuçlar, ANN modelinin yüksek R ve R^2 değeri, ve düşük hata değerleri ile MLR modellerine kıyasla daha büyük potansiyel göstermiştir. Askorbik asit için hem ANN hem de MLR'nin tahminleme performansı neredeyse aynı düzeyde olduğu belirlenmiştir.

Anahtar Kelimeler: Yapay sinir ağı, Çoklu doğrusal regresyon, Portakal suyu, Renk, Modelleme

Introduction

Orange juice is one of common juice manufactured by beverage industry (Tiwari et al., 2009; Ağçam et al., 2016). Orange juice includes valuable bioactive compounds such as carotenoids, phenolic compounds and ascorbic acid etc. (Lee and Coates, 2003; Zulueta et al., 2013; Ağçam et al., 2014). The color of orange juice is an essential factor of quality. for determining juice quality and its acceptance. Its color is mainly due to carotenoid compounds such as α, β and ζ-carotene, lutein, antheraxanthin, zeinoxanthin, zeaxanthin, luteoxanthin and other carotenoid compounds (Meléndez-Martínez et al., 2007). It is also well known that orange juice is quite sensitive to different processing conditions (Shinoda et al., 2004; Tiwari et al., 2009; Ağçam et al., 2016). During heat treatment and storage, orange juice involves a number of deteriorating processes, including browning, ascorbic acid decomposition, clouding, microbiological spoiling, off-flavor, and viscosity (Zheng et al., 2011). It is stated that color change is a marker of chemical and biochemical reactions (Wibowo et al., 2015a). These changes may occur particularly as a result of the formation of non-browning or degradation of carotenoids during processing and storage. The aerobic as well as anaerobic degradation for ascorbic acid and Maillard reactions are two possible processes of non-enzymatic browning (Fustier et al., 2011;Ağçam et al., 2016). Furfural, 5-hydroxymethylfurfural (HMF), ascorbic acid levels, and furfural are some of the indicators of non-enzymatic browning (Shinoda et al., 2004;Burdurlu et al., 2006;Wibowo et al., 2015; Wibowo et al., 2015a). Heat processing is frequently employed method for long shelf life of orange juice because it inactivates spoilage bacteria and pectin methylesterase (PME), one of the endogenous enzymes that is resistant to heat (Elez-Martinez et al., 2006; Timmermans et al., 2011; Vervoort et al., 2011). Despite these improvements, the studies has been stated that heat processing often causes undesirable changes such as nutrients, flavor and color by affecting the overall juice quality (Yeom et al., 2002; Cortés et al., 2006; Rivas et al., 2006; Vervoort et al., 2011; Ağçam et al., 2014). Therefore, prediction of critical parameters of foods during heat processing and storage conditions by models is very important stages (Agha and Alnahhal, 2012). The artificial neural network (ANN) model is one of the often-employed strategies. The function of ANN is to simulate the neuro-physiological structure of the human brain (Huang et al., 2007; Yasar et al., 2012; Khashei et al., 2012; Torkashvand et al., 2017; Deng et al., 2023). ANN model is commonly accepted as modeling technique to express linear or non-linear relationships of complicated input and output parameters in processes (Tehlah et al., 2016; Bhagya Raj and Dash, 2022). There are several neural network models. The most used of this model is feed-forward multi-layer perceptron (MLP) (Chegini et al., 2008; Tiryaki and Aydın, 2014; Huang et al., 2024). Three layers typically make up this network model: input, output and hidden layers. Numerous interconnected neurons, known as processing elements, are present in each layer (Zheng et al., 2011; Erzin and Cetin, 2013; Tiryaki and Aydin, 2014; Nejatdarabi and Mohebbi, 2023). The construction of these structure determines the success of the ANN (Asilturk and Cunkas, 2011). The statistical method of multiple linear regression (MLR) analysis is used to identify the correlations between variables. This approach has been characterized as a standard prediction technique. The MLR is less complex and allows the simulation of situation by changing the values of independent variables (Kolasa-Wiecek, 2015). MLR models consist of several methods which are backward elimination, enter, forward selection and stepwise regression. To achieve the high coefficient of correlation and lower the mean squared error in linear regression, the stepwise multiple regression model has been utilized to identify the most effective independent variables. To maintain the best possible model, this strategy adds all independent variables and then progressively eliminates them as needed for each phase (Ghani and Ahmad, 2010; Shacham and Brauner, 2014; Kolasa-Wiecek, 2015). In the enter method, all independent variables are simultaneously entered into the equation regardless of whether they are statistically significant, and so the effect of all independent parameters upon dependent variables are assessed (Shacham and Brauner, 2014). Comparative studies of ANN and MLR have shown satisfactory results in several studies, including kiwifruit firmness (Torkashvand et al., 2017), the quality changes during storage of gouda cheese (Stangierski et al., 2019), rehydration process of mushroom powder (Nejatdarabi and Mohebbi 2023), drying of pineapple cubes (Meerasri and Sothornvit 2022). ANN models performed better in systems with nonlinear and complicated interactions. However, MLR had better potential than ANN. Additionally, these models may allow later prediction in similar situations without experimenting.

In the literature, a study about prediction of critical quality parameters of orange juice during storage after heat processing is not found. Therefore, the purpose of the presented study was to identify models based on the ANN and MLR model to predict the loss of ascorbic acid and total carotenoid content (TCC), formation of browning index (BI) and HMF of orange juice during storage after heat processing. We also compared the prediction performance of these models for each output or dependent variables using the common statistical parameters.

Material and Methods

Orange juice

Oranges of the Kozan variety were cleaned, peeled, sliced in half, and then squeezed using an orange press (CANCAN, Turkey). To get rid of the seeds and tough pulps, orange juice was filtered into 1-mm stainless steel sieves. Using heat processing equipment, juice was processed quickly according to Ağçam et al. (2016). Three replications of the investigation were conducted.

Heat pasteurization treatment

Heat pasteurization system for using applications was designed by Ağçam et al. (2014). Heat processing at between 70-90°C for 15-120 s was applied. Following the procedures, the samples were placed in sterile and amber bottles. They stored in darkness at 4 and 25 °C for 180 days before being analyzed. The samples were examined at certain days (0, 60, 120 and 180) (Akyıldız et al., 2021).

Determination of browning Index, HMF, ascorbic acid and total carotenoid content

5 mL of orange juice and 5 mL of 95% ethyl alcohol were mixed with a high-speed vortex in teflon tubes to determine the browning index. The mixture was centrifuged (4000 rpm, 10 minutes, 4°C). The absorbance at 420 nm was measured and supernatant was filtered with a 0.45 µm teflon membrane filter (Tiwari et al., 2008).

HPLC method was utilized to determine the ascorbic acid concentration using extraction process suggested by Lee and Coates (2003). A calibration curve was obtained by comparing retention time with an external standard (R^2= 0.997).

HMF analysis was performed using an HPLC (Shimadzu, Japan) system. HMF was identified at 285 nm by examining the UV spectra and retention time with the standard (Ağçam and Akyildiz, 2014; Akyıldız et al., 2021).

The total carotenoid content (TCC) was calculated with a few adjustments based on Lee and Castle (2001). TCC was determined using Equation (1) based on β-carotene:

$$TCC\ (\frac{mg}{100\ mL}) = \frac{\mathrm{A}*DF}{\mathrm{E}^{1/2}} \times 1000 \tag{1}$$

where A = Absorbance; DF = Dilution factor; $E^{1/2}$ = 2505 (β-carotene extinction coefficient in hexane).

Artificial neural networks analysis

In present paper, the multilayer feed-forward networks structure based on backpropagation algorithm was used to find the best ANN network structure. The multi-layer perceptron (MLP) network is one of the most often used feedforward neural networks. The primary benefit of MLP over other neural networks is its simplicity in construction and estimation of any input and output (Nejatdarabi and Mohebbi, 2023; Gonçalves Neto et al., 2021). This network architecture is consist of combination of various model parameters which have a significant effect on learning ability (Tripathy and Kumar, 2009; Chegini et al., 2008; Aghbashlo et al., 2012; Gonçalves Neto et al., 2021). The structure of neural networks used in this study was presented in Fig. 1.

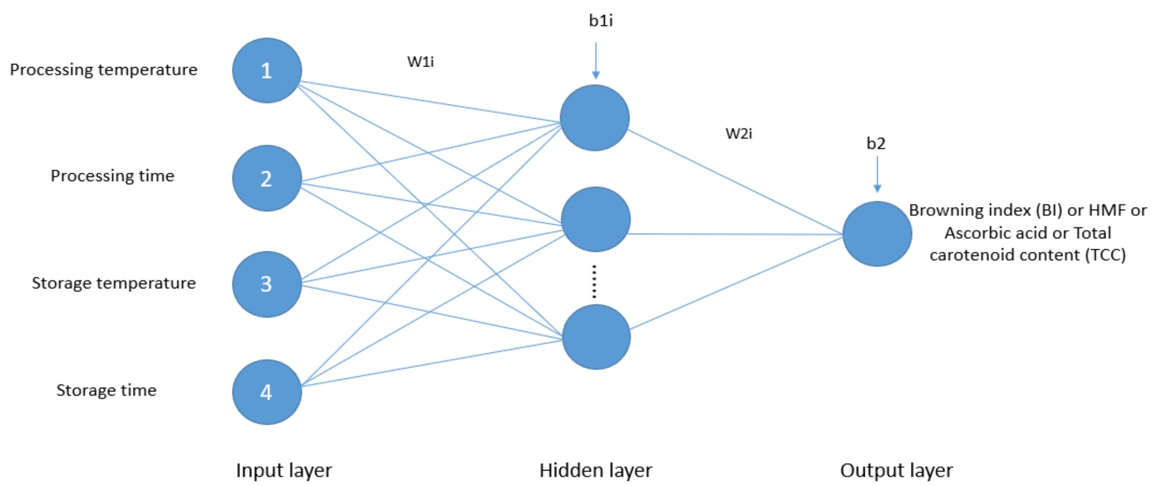

Figure 1. The structure of neural networks

Training of a network is an important point for high or good prediction in neural networks (Zou et al., 2007). At the training process, the weights and biases are initialized (Nikbakht et al., 2014). Training process was used to calculate the gradient and update the weights and biases of network. The training process was run until the validation process can avoid overfitting in the model. It was finished when the validation process resulted in a minimum of the performance goal (Mean Squares Error (MSE)) (Erbil et al., 2018). One of the problems during training is overfitting or memorization. The error of training is a very small value for the network that performed well on the training data, but not well on new data, and so the error of network is large. In other words, the network has memorized the training data, but it has not ability of generalization to unseen data (Sablani and Rahman, 2003; Matworks, 2018). Early stopping or regularization are used to avoid this problem (Matworks, 2018). In this study, early stopping was used for preventing overfitting. The aim of validation process was to verify the efficiency of a network model, and to determine how it perform to a test data known as unseen (outside) data (Bocco et al., 2010). The epoch (iteration) is one training step when the ANN model demonstrates all training data. The error of validation was showed during the training phase. After validation, the models are tested with unseen data about the prediction accuracy of trained model (Sablani and Rahman, 2003; Zheng et al., 2011; Maghrebi et al., 2014). In other words, the basic principle of backpropagation is the use of gradient descent to find the point or points with the lowest degree of error on an error surface. Every iteration consists of two sweeps: forward activation to generate a solution and backward propagation of the calculated error to adjust the weights of the neurons (Nejatdarabi and Mohebbi, 2023).

The inputs selected in the present study are process temperature, process time, storage temperature and storage time while outputs are HMF, BI, TCC and ascorbic acid. Data or variables are divided randomly three parts as training, validation and test. As mentioned above, 85% of all data (70%: training and 15%:validation) was used in multiple linear regression analysis. The last 15% of data was used to compare and test models (Erbil et al., 2018). MSE which is defined in Eq. 4 was used to determine performance of each network.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(xi - yi)^2 \quad (4)$$

where n is overall number of data; yi is the predicted value and xi is the actual value.

To obtain accurate and suitable models, the impact of various back propagation learning algorithms (*trainlm*, *trainscg, trainbr* and *trainrp etc.*) and transfer functions (*logsig, tansig* and *purelin*) were performed for all models (Tripathy and Kumar, 2009). The networks were trained with *trainlm* which accomplished minimization of error (Chia et al., 2012; Xu et al., 2014; Li et al., 2016; Yang et al., 2023; İnan-Çınkır et al., 2024). It provided the best results for all output because the second-degree derivatives of the function are used by the *trainlm* algorithm. Therefore, network output can get closer desired output (Nikbakht et al., 2014). The learning function was used *learngdm.* The transfer function was *logsig* for hidden layer layers, while *purelin* was suited best for the output layer. The learning rate is very important in achieving the minimum of error. The selection of too small or large learning rate may cause slow processing or obtain poor prediction results. The values of learning rate varied from 0.05 to 0.75 (Rai et al., 2005; Yang et al., 2023). The learning rate was selected 0.3 in our study. The other parameters as performance goal and coefficient of momentum were 0.001 and 0.6 for training respectively. The stopping criteria for *trainlm* includes a minimum gradient value of 10^{-7} and a maximum total number of iterations (1000). *Trainlm* use μ as a stopping criterion (maximum value of 10^{10}) when it exceeds the MATLAB standard (Gonçalves Neto et al., 2021).

In neural networks, it is best practice to pre-process (normalization) input and output data before use to prevent impeding the learning process. Normalization is required when there are big differences in the ranges of all variables. It is helpful to facilitate the training of the network, provides memory efficient, improves predictive capabilities of the network, and so yields accurate predictive results (Basheer and Hajmeer, 2000; Bataineh and Marler, 2017). In this study, a total of 200 data were normalized or scaling between -1 and +1 before entering the network so as to avoid the overfitting in training process (İnan-Çınkır et al., 2024). The normalized variable x_N is given by Eq. 5:

$$x_N = \frac{(y_{max}-y_{min})*(x_R-x_{min})}{(x_{max}-x_{min})} + y_{min} \quad (5)$$

where; x_N is normalized data; x_R is actual data, x_{max}, x_{min} are minimum and maximum data, and y_{min}, y_{max} are normalizing limits (respectively (-1, +1).

Multiple linear regression analysis

Multiple linear regression analysis is common method for investigating the best suitable connection between variables (Agha and Alnahhal, 2012; Tiryaki and Aydin, 2014; Erbil et al., 2018; Yu et al., 2018). This study used dependent variables (browning index, HMF, ascorbic acid and total carotenoid content) and independent variables (process temperature (T_p), process time (t_p), storage temperature (T_s) and storage time (t_s)). We were carried out based on stepwise, enter, forward and backward method in the multiple regression model to forecast the dependent variables (Torkashvand et al., 2017). Then, the optimum models, including all independent parameters were chosen with high R and R^2. The analysis was carried out using SPSS version 20 statistical software (IBM Corporation). Multiple regression model can be designed (Eq.2) as;

$$Y = \beta_{0+}\beta_1 X_{1+}\beta_2 X_{2+}\beta_3 X_{3+} \dots\dots\dots.. + \beta_K X_{K+}\varepsilon \qquad (2)$$

where, X_1, X_2, … X_k are independent variables; Y is the dependent variable; β_1, β_2 … β_k are model coefficients; ε is a random element that influences the dependent variable.

The multicollinearity between the independent variables can lead to misleading results in selecting the most important predictors of regression equations. Therefore, a tolerance value (*T*) and variance inflation factor (*VIF*) values were used to determine multicollinearity between independent variables (Eq. 3). There is a multicollinearity issue when the tolerance value is not higher than 0.10 and the VIF is > 10 (Lin, 2008). The Durbin-Watson (DW) test which used to detect the presence of serial correlations (autocorrelation) between errors. The test statistics are changed from 0 to 4, DW values less than 1 or greater than 3 may generally show a problem (Mekanik et al., 2013; Niazian et al., 2018).

$$Tolerans\ (T) = 1 - R^2$$
$$VIF = 1/t \qquad (3)$$

where R^2 is the coefficient of determination,

The data used in regression analysis was 85% (170) of all data. This data was used in training (140: 70% of all data) and validation (30: 15% of all data) of neural nets. The remaining 15% (30) of the data were employed to test and compare the models.

Comparison process

The performances capabilities of all models are compared with statistical methods. To determine prediction accuracy of proposed models, mean square error (*MSE*), root mean square error (*RMSE*), correlation coefficient (*R*), correlation of determination (R^2), mean absolute error (*MAE*), mean absolute percentage error (*MAPE*) and mean percentage error (*MPE%*) were used (Eq. 6-11) (Erbil et al., 2018; İnan-Çınkır et al., 2024). The best trained network models should have the lowest error like *MSE*, *RMSE*, *MAPE* etc. and high *R* and R^2. The statistical parameters are defined as:

$$R = \frac{n(\sum xiyi)-(\sum xi)(\sum yi)}{\sqrt{n\sum xi^2-(\sum xi)^2\ (n\sum yi^2-(\sum yi)^2}} \qquad (6)$$

$$R^2 = 1 - \sum_{i=1}^{n}\frac{(xi-yi)^2}{(xi-yi)} \qquad (7)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(xi-yi)^2} \qquad (8)$$

$$\mathrm{MPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{xi-yi}{xi}\,x100 \qquad (9)$$

$$\mathrm{MAPE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i=1}}^{\mathrm{n}}\left|\frac{xi-yi}{xi}\right|\mathrm{x}\ 100 \qquad (10)$$

$$\mathrm{MAE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i=1}}^{\mathrm{n}}|xi-yi| \qquad (11)$$

where n is the total data number; yi is predicted value for all models while xi is the actual value.

Results and Discussion

According to our results, increasing processing temperature and processing/storage time lead to decrease concentration of ascorbic acid (AA) and carotenoid pigments, and accelerating the formation of HMF and BI (Table 1A). For 0 days, heat processing at 70ºC for 15-120 s (HP1) had little deterioration or best quality compared to other heat treatments. Bull et al. (2004) reported the browning index of pasteurized orange juice as 0.096 abs. Ağçam et al. (2016) determined that BI was 0.223-0.355 in pasteurized orange juice for 60-180 days. Similar results were obtained in this study. Pham et al. (2019) found that HMF was not detected in pasteurized orange juice at the start of storage, its concentration increased noticeably with storage time. HMF was 25.91 mg L^{-1} at 42°C for 15 weeks.

The quality deterioration of orange juice during storage conditions was determined based on AA concentration (Elez-Martinez et al., 2006). For 0 days, HP1 had little deterioration or best quality compared to other heat treatments. HP1 was also suitable for little quality deterioration during the 180 days at 4 and 25ºC storage. The lowest ascorbic acid contents belonged to heat processing at 85ºC for 120 s during the 180 days at 25ºC storage. AA decreased from 666.91±6.95 to 78.69±5.55 mg L^{-1} on 180th day (Table 1B). The European Fruit Juice Association (AIJN) stated that AA as quality criterion has been identified 400-500 mg L^{-1} in fresh orange juice. The amount was defined as 200 mg L^{-1} until the last consumption (Elez-Martinez et al., 2006). Polydera et al. (2003) stated that shelf life of orange juice ends when 50% of ascorbic acid is lost. They determined that processing, temperature of storage, and packaging factors all had a significant impact on the degradation of ascorbic acid. Wibowo et al. (2015b) showed that the amounts of AA and dehydroascorbic acid (DHA) decrease with temperature and storage duration. They observed that over 75% of AA degraded in 32 weeks at 20°C. Gomez Ruiz et al. (2018) determined that degradation of AA caused several undesirable compounds during strorage. pH, oxygen and metal presence, AA concentration, thermal processing and storage temperature, light exposure, and citric acid are the most crucial elements for degradation of AA.

Table 1A. The value ranges of browning index, HMF during storage conditions (4°C and 25°C between 0 and 180 days) after heat processing (70-90°C and 15-120 s).

	BI (abs)				HMF (µg L^{-1})			
	4°C		25°C		4°C		25°C	
Symbol†	0 day	180 day	0 day	180 days	0 day	180 day	0 day	180 day
A	0.179-0.187	0.227-0.261	0.179-0.187	0.288-0.433	5.07-12.80	46.13-181.60	5.07-12.80	1760.80-2500.93
B	0.204-0.228	0.238-0.279	0.203-0.228	0.376-0.493	5.87-14.20	78.93-108.13	5.87-14.20	2217.63-2744.67
C	0.191-0.208	0.257-0.296	0.191-0.208	0.276-0.434	5.64-17.60	82.93-142.13	5.64-17.60	1227.60-2340.80
D	0.181-0.192	0.259-0.288	0.181-0.192	0.281-0.461	6.36-22.13	105.73-127.07	6.36-22.13	977.87-2304.67
E	0.180-0.189	0.239-0.286	0.180-0.189	0.290-0.382	7.33-31.74	88.80-146.27	7.33-31.74	1904.67-2652.93

† A: heat processing at 70°C for 15-120 s, B: heat processing at 75°C for 15-120 s, C: heat processing at 80°C for 15-120 s, D: heat processing at 85°C for 15-120 s E: heat processing at 90°C for 15-120 s, BI: browning index, ± values are standard deviation.

Table 1B. The value ranges of browning index and HMF during storage conditions (4°C and 25°C between 0 and 180 days) after heat processing (70-90°C and 15-120 s).

	AA (mg L^{-1})				TCC (mg L^{-1})			
	4°C		25°C		4°C		25°C	
Symbol†	0 day	180 day	0 day	180 day	0 day	180 day	0 day	180 day
A	635.09-678.59	361.61-427.63	635.09-678.59	257.89-310.12	7.39-8.03	2.72-3.63	7.39-8.03	3.44-4.62
B	577.31-607.27	361.97-427.63	571.33-607.27	195.62-311.81	6.80-6.97	3.51-4.18	6.76-7.43	3.41-4.82
C	585.48-628.81	263.03-370.09	585.48-628.81	159.58-308.69	6.11-6.75	3.00-3.85	6.27-6.67	3.03-3.70
D	608.16-666.91	288.36-385.51	608.16-666.91	78.69-296.25	6.33-6.70	3.66-4.46	6.33-6.71	2.92-3.48
E	611.51-685.07	289.52-401.76	611.51-685.07	134.06-330.15	6.92-7.64	3.62-4.32	7.04-7.34	2.16-2.88

† A: heat processing at 70°C for 15-120 s, B: heat processing at 75°C for 15-120 s, C: heat processing at 80°C for 15-120 s, D: heat processing at 85°C for 15-120 s E: heat processing at 90°C for 15-120 s, TCC: total carotenoid content, ± values are standard deviation.

The orange juice quality during heat processing and storage was imposed by two important changes which are off-flavor and browning. During the storage, BI indicates the color changes of orange juice depending on non-enzymatic browning (Tiwari et al., 2008). More than one of the mechanisms are related to formation of brown pigments. These mechanisms may be Maillard reactions and degradation of the ascorbic acid (Fustier et al., 2011). The brown pigments also consist of the oxidation of phenolic compounds. AA acts as antioxidant and represses browning reaction but it may act formation of furfural and HMF because of easily oxidized and decomposed at high temperatures (Shinoda et al., 2004; Ağçam et al., 2016; Akyıldız et al., 2021).

TCC values were 2.16-8.03 mg L^{-1} for 60-180 days. Gama and Sylos (2007) determined the changes in TCC in orange juice due to thermal pasteurization. They reported that TCC in fresh orange juice, which was 12 mg/L, decreased by approximately 13% to 10.40 mg/L with pasteurization. Cortes et al. (2006) reported that TCC values decreased by %12.6 with pasteurization.

There are many studies about critical parameters of orange juice such as HMF and ascorbic acid treated by heat pasteurization and PEF (pulsed electric fields) Ağçam et al. (2016), the effect of (PEF) and thermal pasteurization on phenolic compounds of orange juice Ağçam et al. (2014), total phenolic content, total carotenoid content, antioxidant capacity, ascorbic acid and turbidity index (Zulueta et al. 2013), color and ascorbic acid in pasteurized blood orange juice (Remini et al., 2015). Researchers have stated that there is a progressive loss of nutritional content and a rise in undesirable elements because of heat processing and storage.

ANN models

Trial and error strategy were used to optimize the ANN models (Saraceno et al., 2012; Ding et al., 2017). The optimum number of neurons in the hidden layer, hidden layer number and epoch number were determined owing to this method. Because trial and error strategy demonstrate that two or more hidden layers and epoch number more than 1000 decrease the prediction efficiency of the non-linear function. Therefore, epoch number fixed at 1000 and one hidden layer was selected for all network models in the our study (Zheng et al., 2011; Torkashvand et al., 2017; Li et al., 2016). The neuron number in the hidden layer ranged from 2 to 20. Our results showed that the network construction included one hidden layer with 4, 3, 8, 4 neurons had the best structure for prediction of BI, HMF, AA and TCC, respectively. All networks had one neuron in the output layer. Selecting optimal neurons number in the hidden layer is very important for the development of the best ANN models. Using a lot of neurons in the network model may prolong the training time and may lead to overfitting. In the same way, using fewer neurons for hidden layer become trapped in a local minimum and so it decrease the performance of model (Erzin et al., 2010).

The network with *logsig* transfer function and *trainlm* algorithm performed the best in terms of obtained minimum errors for each of models. The validation performance plot of the network is given in Fig. 2. All performance goals for validation processes were reached at 74, 52, 15 and 29 epochs for BI, HMF, ascorbic acid and TCC, respectively. According to performance plots, *MSE* converged minimum, and the learning ability for network increased as the number of epochs increased. However, increasing the epoch number may increase the problem of overfitting. To avoid overfitting, epoch numbers should be limited (Youssefi et al., 2009). This figure (Fig. 2) determined that the error of test and validation have same behavior and any significant over fitting did not occur.

Based on Fig. 3, the R values between actual and predicted BI values were higher than 0.90 for training, validation, test and all of data, respectively. For AA, R value of training, validation, test and all data were 0.91, 0.90, 0.92 and 0.91. Predicted values for HMF and TCC were positively correlated to the actual values. The correlation coefficients of training, validation, test and all data of HMF are 0.99, 0.97, 0.99 and 0.99, respectively. The correlation coefficients were 0.95, 0.95, 0.97 and 0.95 for TCC. These results stated that the ANN models were satisfactory, because the output data were as close as possible to the desired output data. For this reason, data points became intense around the line (Fig. 3).

ANN models for predicting and obtaining optimum model were applied by Fazaeli et al. (2013), Torkashvand et al. (2017), Yadav and Chandela (2017), Nikbakht et al. (2014), Anastacio et al. (2016), Deng et al. (2023), Nejatdarabi and Mohebbi (2023). As a result, it is stated that ANN is suitable for obtaining optimum models and calculating satisfactory results.

ANN Formulation

For formulation of ANN models, the following mathematical equations (Eq. 12-17) was used the weights (w) and biases (b) values for selected the best ANN models are given in Table 2.

$E_i = w_{1i} * T_p + w_{2i} * t_p + w_{3i} * T_s + w_{4i} * t_s + b_{1i}$ (12)

where, E_i values are calculated using Eq. 11 includes processing temperature (T_p), processing time (t_p), storage temperature (T_s), storage time (t_s).

$$F_i = \frac{1}{1+ e^{-Ei}} \qquad (13)$$

where, F_i values are detemined using the *logsig* transfer function given in Eq.12.

The network output data for BI, HMF, AA and TCC were given Eq. 14,15,16 and 17, respectively.

$BI = 9,698955*F_1 - 0,12679*F_2 + 0,472012*F_3 + 3,866417*F_4 - 4,74462$ (14)

$\log (HMF) = -3,551471*F_1 + 3,380717*F_2 + 0,537706*F_3 - 0,47159$ (15)

$\log (AA) = -0,34178*F_1 - 0,13789*F_2 + 0,155299*F_3 - 0,10878*F_4 + 0,60083*F_5 + 0,731891*F_6 + 0,144765*F_7 + 0,694978*F_8 - 0,20744$ (16)

$(TCC) = 2,117841*F_1 - 3,84915*F_2 - 1,48123*F_3 - 1,81265*F_4 + 2,639356$ (17)

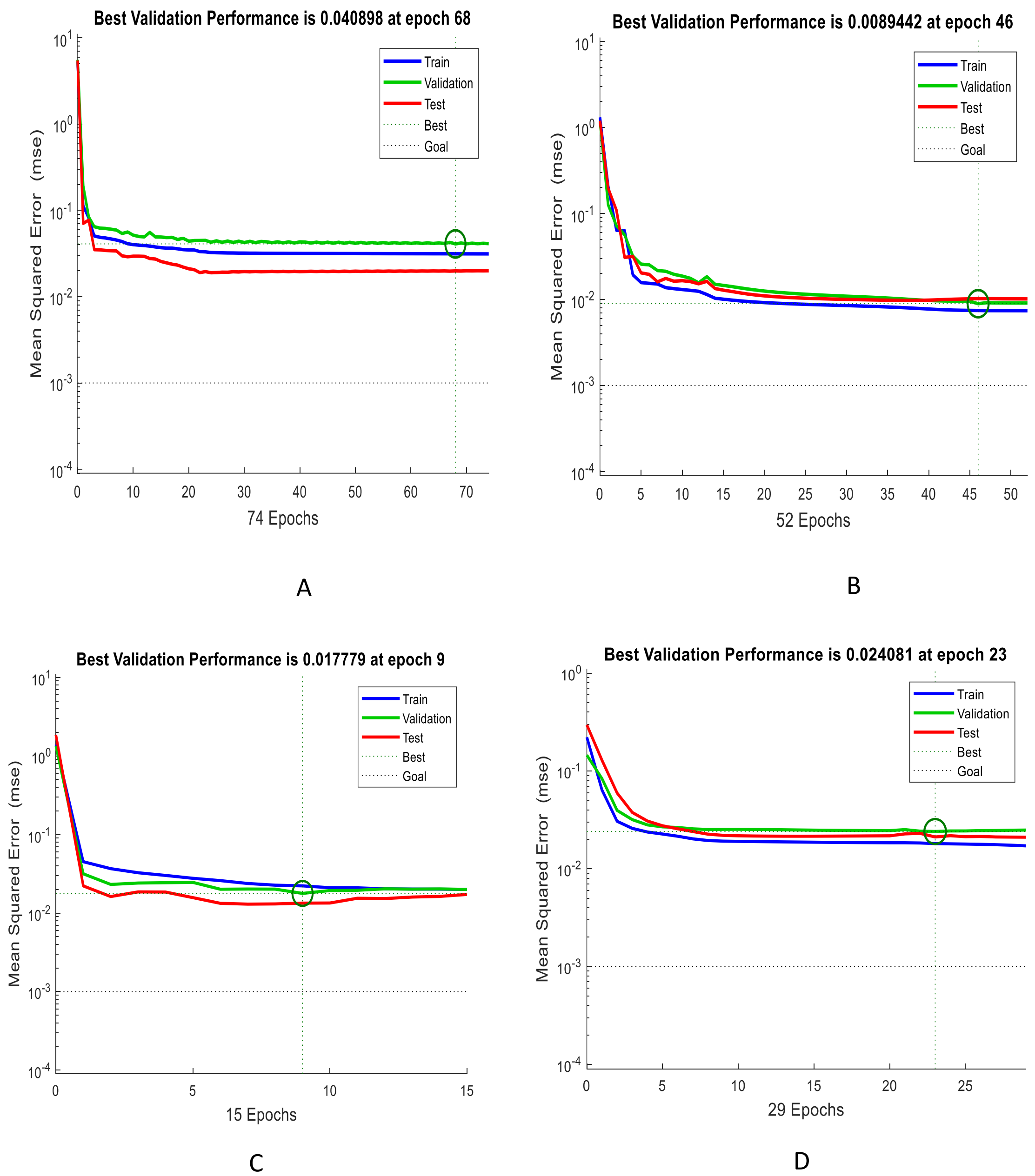

Figure 2. The validation performance plot of the network for A) BI, B) HMF, C) AA and D) TCC.

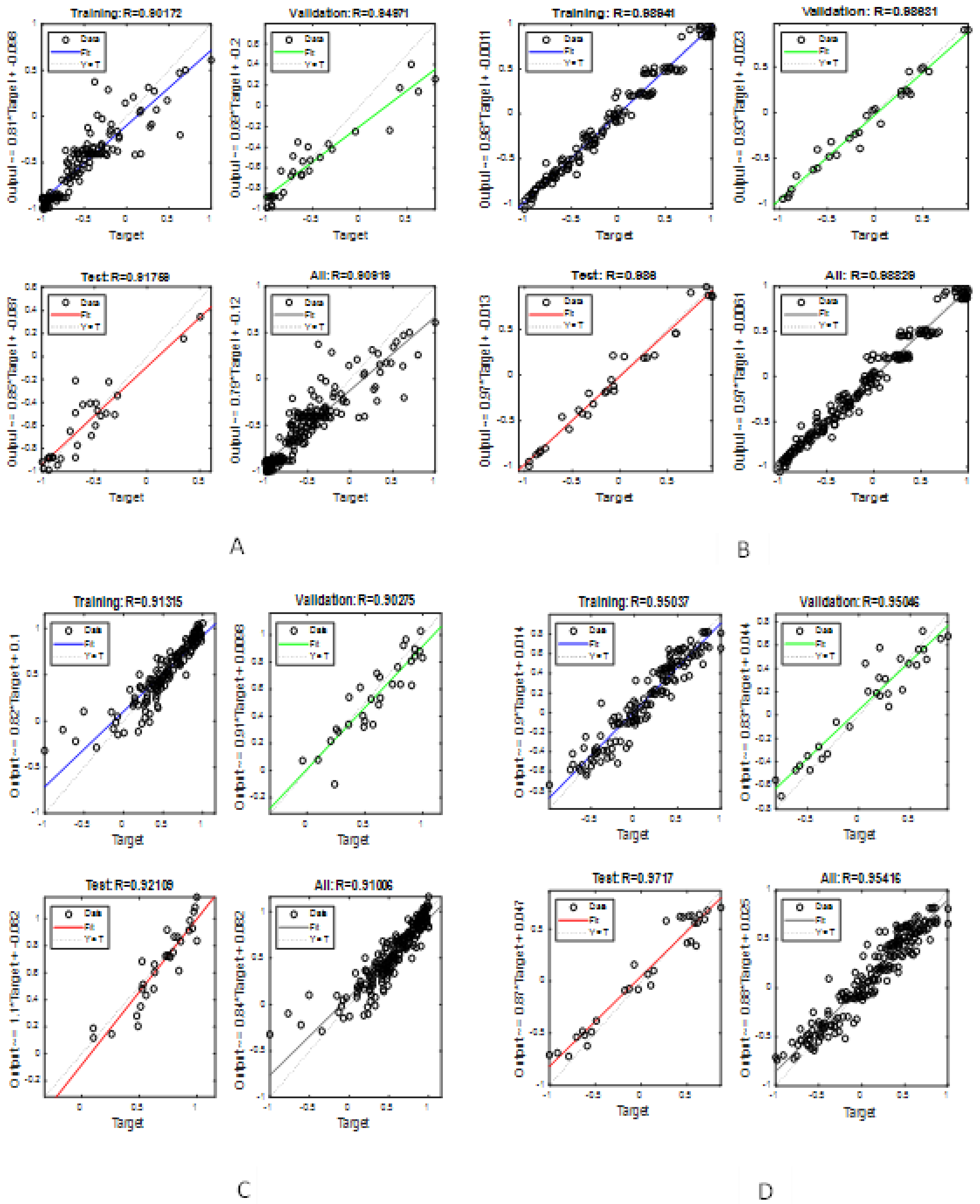

Figure 3. Relationship of the prediction and actual values of BI (A), HMF (B), AA (C) and TCC (D) for training, validation, test and all data, respectively.

Table 2. The weight and bias values of the best ANN models

		Weights					Bias	
		Input layer				Output layer	Hidden layer	Output layer
ANN models	i	w1i	w2i	w3i	w4i	w2j	b1i	b2j
BI	1	-0.3372	0.2235	2.0174	3.8000	9.6990	-8.3483	-4.7446
	2	13.8277	-3.1640	-3.0611	4.7539	-0.1268	-11.3891	
	3	-0.2329	0.7439	14.0288	7.5705	0.4720	3.2407	
	4	10.8565	-0.9146	-1.1401	-0.7883	3.8664	14.6206	
HMF	1	-0.1351	-0.2341	-5.3158	-7.4640	-3.5515	1.0845	-0.4716
	2	-0.0017	0.0870	1.2998	-2.9617	3.3807	0.5673	
	3	0.9656	1.0053	1.6905	3.0620	0.5377	3.9195	
Ascorbic acid	1	-3.6692	0.9634	0.1395	3.4246	-0.3418	5.0797	-0.2074
	2	1.4656	4.8249	2.4622	-2.5826	-0.1379	-3.3427	
	3	1.8019	2.6854	-2.8838	3.7548	0.1553	-0.9538	
	4	-1.9014	-3.1319	2.9213	-2.2445	-0.1088	-0.4960	
	5	-2.7970	0.0238	-2.2907	-1.3944	0.6008	-0.9294	
	6	-2.9236	-2.0830	-3.0205	5.3019	0.7319	-2.3076	
	7	0.6414	3.3304	2.1005	-3.2013	0.1448	3.0403	
	8	4.7526	-0.4884	-1.7770	-5.5692	0.6950	2.5743	
TCC	1	1.6294	0.0096	0.8085	-3.2545	2.1178	-4.0007	2.6394
	2	0.2901	-0.0647	1.0716	-1.3180	-3.8492	-1.3269	
	3	0.0818	0.1795	1.2368	-0.1039	-1.4812	-1.2501	
	4	1.2068	0.0603	-1.2278	2.6437	-1.8127	2.8522	

Multiple linear regression models

MLR models were performed by backward, forward, enter and stepwise regression methods. In statistical analysis, the presence of multi collinearity among independent variables was identified by calculating Variance Inflation Factors (VIF) and tolerance values indicates whether an independent variable has a strong linear relationship with other independent variable (Dormann et al., 2013; Niresh and Velnampy, 2014). All VIF values were well below 10 and tolerance levels were not below 0.1. Thus, there was no multicollinearity between the independent variables. Durbin-Watson values were given Table 3. The presence of autocorrelation between errors was not detected due to these values are between 1 and 3.

In MLR analysis, our aim was detected the effect of processing temperature, processing time, storage temperature and storage time on BI, HMF, ascorbic acid and TCC at confidence level of 1%. MLR equation of each dependent variables were presented in Table 3. We only applied logarithmic modification for HMF and AA data for constituting more reliable predictive equations. For HMF, stepwise regression model was chosen due to high R and R^2. In a stepwise regression model, controllable variables are added and removed based on each step's statistical significance (Ghani and Ahmad, 2010). Therefore, it is found that the effect of processing temperature, processing time, storage temperature and storage time on HMF are statistically significant. According to the results of regression analysis, all models seemed to be good depending on R and R^2 values. However, the MLR models for HMF and TCC had a relatively higher predictability than BI and AA.

MLR is a well-known and practical method for determining a connection between the variables. As a limited approach to creating mathematical models, MLR models have the drawback of only being able to explain linear relationships between variables, ignoring other types of interactions.

Despite being considered unproductive, multiple linear regression models are frequently employed in modeling processes and have been validated with success (Stangierski et al., 2019).

Table 3. Multiple linear regression model equation and statistic values of each output variables

MLR models	Methods	Equation	R	R^2	DW	VIF
BI	Enter	y_1 = 8.728889 - 0.021634* T_p - 0.000798* t_p - 0.017825* t_s + 0.010577* T_s	0.82	0.67	2.13	1.000
log HMF	Stepwise	log (y_2) = -0.378287 + 0.009511* T_p+ 0.002133* t_p + 0.009124* t_s + 0.039685* T_s	0.92	0.84	2.56	1.006
log AA	Enter	Log(y_3) = 2.985761- 0.001271* T_p - 0.000532* t_p - 0.001923* t_s - 0.003716* T_s)	0.83	0.68	1.43	1.001
TCC	Enter	y_4 = 8.728889 - 0.021634* T_p - 0.000798* t_p -0.017825* t_s + 0.010577* T_s	0.92	0.85	1.69	1.007

BI: browning index. TCC: total carotenoid content. VIF: Variance Inflation Factors. DW: Durbin-Watson value

Performance comparison of ANN and MLR models

The prediction accuracy of proposed ANN and MLR models was assessed on basis of various statistical parameters to avoid establishing inconsistent findings on the data from the models. The results are listed in Table 4. It is stated that the models with lowest *MSE*, *RMSE*, *MAE*, *MAPE*, *MPE* and, highest *R* and R^2 are best suited for each condition.

According to Yadav and Chandel (2017), a model is considered to have very good predicting ability if its MAPE value is 10% or less, good prediction performance if it is 10%–20%, acceptable prediction quality if it is 20%–50%, and incorrect prediction efficiency if it is greater than 50%. For ANN models, the MAPE ranged from 1.57% to 8.16%, indicating a larger prediction accuracy range. They range from 1.90% to 17.44% for MLR models. which shows good prediction efficiency.

The comparison of between the actual and prediction values of ANN and MLR for testing data of all output was given Fig. 4. As can be observed from Fig. 4, The curves of predicted values for the MLR and ANN models followed a similar pattern. The ANN model demonstrated a superior match. ANN is a more accurate predictor than MLR for all outputs with the highest *R* and R^2, and the lowest error values. ANN have a unique property that removes errors, making it a more effective prediction tool (Stangierski et al., 2019). MLR had an acceptable prediction performance for HMF, AA and TCC.

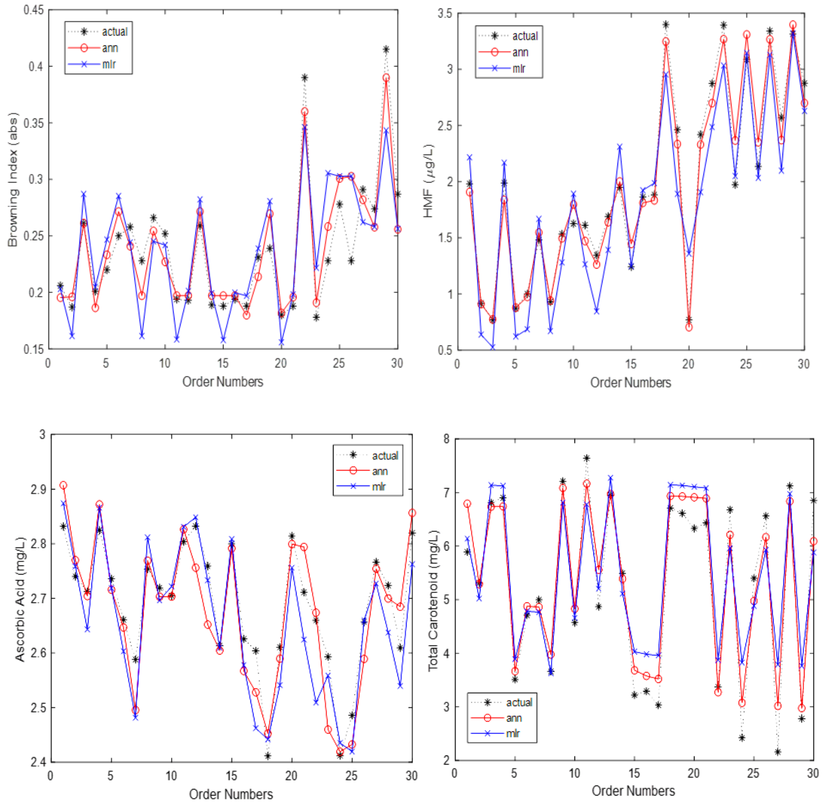

Figure 4. Comparison of the actual and prediction values of ANN and MLR for testing data of browning index (BI), HMF, AA and total carotenoid content (TCC).

ANN models are formed requiring lengthy iterative calculations. whereas the MLR models are generated only one step calculation. Therefore, ANN model has higher predictability because of a greater degree of fault tolerance and robustness. and shows better generalization capacity than the MLR. ANN models seem to be more appropriate method to approximate nonlinearity systems (Erzin and Cetin, 2012; Prakash Maran et al., 2013; Nejatdarabi and Mohebbi, 2023). The comparison of key results/similarities/differences with recent related studies was given in Table 5. ANN had superior performance compared to MLR, but some studies were stated that MLR had greater potential for prediction than ANN.

Table 4. The statistical parameters of MLR and ANN for test data.

Statistical Parameters	BI		HMF		AA		TCC	
	MLR	**ANN**	**MLR**	**ANN**	**MLR**	**ANN**	**MLR**	**ANN**
R	0.79	0.92	0.94	0.99	0.92	0.92	0.93	0.97
R^2	0.68	0.84	0.88	0.98	0.85	0.86	0.9	0.94
MSE	0.00	0.00	0.1	0.02	0.00	0.00	0.44	0.18
RMSE	0.04	0.02	0.31	0.14	0.06	0.05	0.67	0.43
MAE	0.03	0.02	0.27	0.11	0.05	0.04	0.55	0.35
MPE	-0.92	-0.89	6.74	0.39	1.16	0.47	-8.5	-4.49
MAPE (%)	11.85	7.03	17.44	5.47	1.87	1.57	13.88	8.16

Table 5. The comparison of Key results/similarities/differences with recent related studies

Food- Applied Process	Models	Results	References
Orange juice quality parameters-heat process and storage	ANN-MLR	ANN models showed greater potential in comparison with MLR models for all variables, but ANN and MLR performed nearly identically for ascorbic acid prediction (R^2: 0.86-0.98 and RMSE: 0.02-0.43 for ANN; R^2: 0.68-0.94 and RMSE: 0.04-0.67 for MLR).	In this study
Red bayberry juice quality parameters- heat process and storage	ANN- Partial Least-Squares Regression (PLSR)	The PLSR model was sufficient for predicting antioxidant activity, and the ANN model could be used to predict changes in anthocyanin content and ascorbic acid in bayberry juice during storage.	Zheng et al. (2011)
Kiwifruit firmness with different mineral concentration	ANN-MLR	Modeling was made with four different model sets. MLR models evaluated the firmness of the fruit with more accuracy than the ANN models on three datasets (R^2: 0.08-0.724 and RMSE: 0.481-1.37 for ANN; R^2: 0.398-0.551 and RMSE: 0.809-3.11 for MLR).	Torkashvand et al. (2017)
Gouda cheese- quality changes during storage	ANN-MLR	Compared to the MLR models, the ANN model performed moderately better (R^2: 0.98 and RMSE: 1.35 for ANN; R^2: 0.94 and RMSE: 2.48 for MLR).	Stangierski et al. (2019)
Rice- Enzymatic Milling	ANN-MLR	The ANN approach performs better than MLR (R^2: 0.97-0.99 and MSE:0.005-6.13 for ANN; R^2: 0.87–0.90 and SSE: 0.008–8.25 for MLR).	Kothakota et al. (2021)
Pineapple cubes-drying process	ANN-MLR-mathematical models	When compared to other models, ANN models had superior performances (R^2: 0.9975-0.9999 and RMSE: 0.001-0.0008).	Meerasri and Sothornvit (2022)
Mushroom powder - rehydration process	ANN-MLR	ANN showed fewer error values than MLR (R^2: 0.98-0.99 and RMSE: 1.17-3.12 for ANN; R^2: 0.96-0.98 and RMSE:1.53-4.24 for MLR).	Nejatdarabi and Mohebbi (2023)
Physicochemical Properties of Grape-Skin Compost	ANN-MLR-piecewise linear regression (PLR)	MLR and PLR models were suitable for prediction. ANN had potential for predicting some properties like dry-matter and moisture.	Sokač Cvetnić et al. (2024)

Conclusions

In present paper, ANN and MLR models to modeling the browning index, HMF, ascorbic acid and total carotenoid as output variables are examined and analyzed the most common statistical methods to determine performance of models. The developed back propagation ANN models were optimized by changing the number of neurons in the hidden layer and selecting appropriate training algorithm, transfer function and epoch number. The back-propagation training algorithms. The best results were obtained with 4, 3, 8 and 4 neurons in the hidden layer for BI, HMF, AA and TCC, respectively. It is found that both ANN models show good predictions. For MLR model. several methods used to learn relationship between independent variables by eliminating data collinearity. To check the performance of the models, many performance criteria were calculated. Based on results of performance criteria. The ANN models are found to be more suitable with the highest *R* and R^2 value and the lowest *MSE*, *MAE, MAPE* in comparison with MLR models. Lastly, it was found that proposed ANN models may seem a very useful method as an alternative for predictive modelling of quality parameters of orange juice. For quality parameters, using moderate pasteurization standards and selecting a storage temperature below 25 °C are crucial. Due to these results, less undesirable compounds caused by heat processing may be produced, and juice products that are high in nutrients and minimally processed may be obtained. These models also may provide satisfactory knowledge to produce better quality orange juice for researchers and manufacturers. The ANN has higher accuracy, high error tolerance, efficient handling of non-linear functions, shorter calculating time and capability of learning compared to MLR. These advantages of ANN will lead to use more applications for predicting and optimization in food engineering.

Declarations

Acknowledgement

The authors would like to thank Cukurova University (Project No: ZF2012YL3) for financial support.

Conflict of Interest

The authors declare that there is no conflict of interest between them.

Author Contribution

Asiye Akyıldız: Conceptualization, Resources, Project administration, Supervision, Writing–review&editing, **Tuba Şimşek Mertoğlu**: Data curation, Investigation, Visualization, Writing–original draft, **Nuray Inan-Çınkır:** Formal analysis, Conceptualization, Resources, Supervision, Writing–review & editing, **Erdal Ağçam:** Formal analysis, Conceptualization, Resources, Supervision, Writing–review & editing.

References

Agcam, E., Akyildiz, A., & Evrendilek, G. A. (2014). Comparison of phenolic compounds of orange juice processed by pulsed electric fields (PEF) and conventional thermal pasteurisation. Food Chemistry, 143, 354–361. http://doi.org/10.1016/j.foodchem.2013.07.115

Agcam, E., Akyildiz, A., & Akdemir Evrendilek, G. (2016). A comparative assessment of long-term storage stability and quality attributes of orange juice in response to pulsed electric fields and heat treatments. *Food and Bioproducts Processing*, *99*, 90–98. https://doi.org/10.1016/j.fbp.2016.04.006

Akyıldız A., Mertoglu TS., & Ağçam, E. (2021). Kinetic study for ascorbic acid degradation. hydroxymethylfurfural and furfural formations in orange juice. *Journal of Food Composition and Analysis.* 102. 103996. https://doi.org/10.1016/j.jfca.2021.103996

Agha, S. R., & Alnahhal, M. J. (2012). Neural network and multiple linear regression to predict school children dimensions for ergonomic school furniture design. *Applied Ergonomics*, *43*(6), 979–984. https://doi.org/10.1016/j.apergo.2012.01.007.

Aghbashlo, M., Mobli, H., Rafiee, S., & Madadlou, A. (2012). The use of artificial neural network to predict exergetic performance of spray drying process: A preliminary study. Computers and Electronics in Agriculture, 88, 32–43. http://doi.org/10.1016/j.compag.2012.06.007

Anastacio, A., Silva, R., & Carvalho, I.S. (2016). Phenolics extraction from sweet potato peels: modelling and optimization by response surface modelling and artificial neural network. Journal of Food Science and Technology, 53(12), 4117–4125. http://doi.org/10.1007/s13197-016-2354-1

Asiltürk, I., & Cunkas, M. (2011). Modeling and prediction of surface roughness in turning operations using artificial neural network and multiple regression method. Expert Systems with Applications. 38(5), 5826–5832. https://doi.org/10.1016/j.eswa.2010.11.041.

Bataineh, M., & Marler, T. (2017). Neural network for regression problems with reduced training sets. Neural Networks, 95, 1–9. http://doi.org/10.1016/j.neunet.2017.07.018

Bhagya Raj, G.V.S., & Dash, K.K. (2022). Comprehensive study on applications of artificial neural network in food process modeling. Critical Reviews in Food Science and Nutrition. 62(10). 2756–2783. https://doi.org/10.1080/10408398.2020.1858398

Bocco, M., Willington, E., & Arias, M. (2010). Comparison of Regression and Neural Networks Models to Estimate Solar Radiation. Chilean Journal of Agricultural Research, 70(3). http://doi.org/10.4067/S0718-58392010000300010.

Bull, M. K., Zerdin, K., Howe, E., Goicoechea, D., Paramanandhan, P., Stockman, R., Sellahewa, J., Szabo, E. A., Johnson, R. L., & Stewart, C. M., 2004. The Effect of High Pressure Processing on The Microbial, Physical and Chemical Properties of Valencia and Navel Orange Juice. Innovative Food Science and Emerging Technologies, 5:135–149. https://doi.org/10.1016/j.ifset.2003.11.005

Burdurlu, H. S., Koca, N., & Karadeniz, F. (2006). Degradation of vitamin C in citrus juice concentrates during storage. *Journal of Food Engineering*. *74*(2), 211–216. http://doi.org/10.1016/j.jfoodeng.2005.03.026

Chegini, G. R., Khazaei, J., Ghobadian, B., & Goudarzi, A. M. (2008). Prediction of process and product parameters in an orange juice spray dryer using artificial neural networks. Journal of Food Engineering, 84(4), 534–543. http://doi.org/10.1016/j.jfoodeng.2007.06.007

Chia, K. S., Abdul Rahim, H., & Abdul Rahim, R. (2012). Prediction of soluble solids content of pineapple via non-invasive low cost visible and shortwave near infrared spectroscopy and artificial neural network. *Biosystems Engineering*, *113*(2), 158–165. http://doi.org/10.1016/j.biosystemseng.2012.07.003

Cortes, C., Torregrosa, F., Esteve, M. J., & Frígola, A. (2006). Carotenoid profile modification during refrigerated storage in untreated and pasteurized orange juice and orange juice treated with high-intensity pulsed electric fields. *Journal of Agricultural and Food Chemistry*, *54*(17), 6247–6254. https://doi.org/10.1021/jf060995q.

Ding, B., Li, L., & Yang, H. (2017). An artificial neural network approach to estimating the enzymatic hydrolysis of Chinese yam (*Dioscorea opposita* Thunb.) starch. *Journal of Food Processing and Preservation*, (April 2016), e13176. http://doi.org/10.1111/jfpp.13176.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., … Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 027–046. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Elez-Martinez P., Soliva-fortuny, R. C., & Mart, O. (2006). Comparative study on shelf life of orange juice processed by high intensity pulsed electric fields or heat treatment. 321–329. http://doi.org/10.1007/s00217-005-0073-3

Erbil, Y., Babaarslan, O., & Ilhan, İ. (2018). A comparative prediction for tensile properties of ternary blended open-end rotor yarns using regression and neural network models. *The Journal of The Textile Institute*, *109*(4), 560–568. https://doi.org/10.1080/00405000.2017.1361164.

Erzin, Y., Rao, B. H., Patel, A., Gumaste, S. D., & Singh, D. N. (2010). Artificial neural network models for predicting electrical resistivity of soils from their thermal resistivity. International Journal of Thermal Sciences, 49(1), 118–130. http://doi.org/10.1016/j.ijthermalsci.2009.06.008

Erzin, Y., & Cetin, T. (2012). The use of neural networks for the prediction of the critical factor of safety of an artificial slope subjected to earthquake forces. Scientia Iranica, 19(2), 188–194. http://doi.org/10.1016/j.scient.2012.02.008

Fazaeli, M., Emam-Djomeh, Z., Omid, M., & Kalbasi-Ashtari, A. (2013). Prediction of the Physicochemical Properties of Spray-Dried Black Mulberry (Morus nigra) Juice using Artificial Neural Networks. Food and Bioprocess Technology, 6(2), 585–590. https://doi.org/10.1007/s11947-011-0648-x.

Fustier, P., St-Germain, F., Lamarche, F., & Mondor, M. (2011). Non-enzymatic browning and ascorbic acid degradation of orange juice subjected to electroreduction and electro-oxidation treatments. Innovative Food Science and Emerging Technologies, 12(4), 491–498. http://doi.org/10.1016/j.ifset.2011.07.014

Gama, J. J. T. & Sylos, C. M., 2007. Effect of Thermal Pasteurization and Concentration on Carotenoid Composition of Brazilian Valencia Orange Juice. Food Chemistry, 100:1686–1690. https://doi.org/10.1016/j.foodchem.2005.01.062

Ghani, I. M. M., & Ahmad, S. (2010). Stepwise multiple regression method to forecast fish landing. Procedia - Social and Behavioral Sciences, 8(5), 549–554. http://doi.org/10.1016/j.sbspro.2010.12.076

Gomez Ruiz, B., Roux, S., Courtois, F., Bonazzi, C., 2018. Kinetic modelling of ascorbic and dehydroascorbic acids concentrations in a model solution at different temperatures and oxygen contents. Food Res. Int. 106, 901–908.

Gonçalves Neto, J., Vidal Ozorio, L., Campos de Abreu, T. C., Ferreira dos Santos, B., & Pradelle, F. (2021). Modeling of biogas production from food, fruits and vegetables wastes using artificial neural network (ANN). *Fuel*, *285*(August 2020), 119081. https://doi.org/10.1016/j.fuel.2020.119081

Huang, Y., Kangas, L. J., & Rasco, B. A. (2007). Applications of Artificial Neural Networks (ANNs) in food science. *Critical Reviews in Food Science and Nutrition*, *47*(2), 113–126. http://doi.org/10.1080/10408390600626453.

İnan-Çınkır, N., Süfer, Ö., & Pandiselvam, R. (2024). Kinetic and artificial neural network modeling of dried black beauty eggplant (Solanum melongena L.) slices during rehydration. *Journal of Food Processing Engineering*. 47(3), 1–12. https://doi.org/10.1111/jfpe.14567

Kalogirou, S.A. (1999). Applications of artificial neural networks in energy systems. Energy Conversion and Management, 40(10), 1073–1087. http://doi.org/10.1016/S0196-8904(99)00012-6

Khashei. M.. Zeinal Hamadani. A.. & Bijari. M. (2012). A novel hybrid classification model of artificial neural networks and multiple linear regression models. Expert Systems with Applications. *39*(3). 2606–2620. https://doi.org/10.1016/j.eswa.2011.08.116.

Kolasa-Wiecek, A. (2015). ScienceDirect Stepwise multiple regression method of greenhouse gas emission modeling in the energy sector in Poland. *JES*, *30*, 47–54. http://doi.org/10.1016/j.jes.2014.09.037

Kothakota, A., Pandiselvam, R., Siliveru, K., Pandey, J. P., Sagarika, N., Srinivas, C. H. S., Kumar, A., Singh, A., & Prakash, S. D. (2021). Modeling and optimization of process parameters for nutritional enhancement in enzymatic milled rice by multiple linear regression (MLR) and artificial neural network (ANN). Foods, 10(12). https://doi.org/10.3390/foods10122975

Lee, H. S., & Castle, W. S. (2001). Seasonal changes of carotenoid pigments and color in Hamlin, Earlygold, and Budd Blood orange juices. Journal of Agricultural and Food Chemistry, 49(2), 877–882. http://doi.org/10.1021/jf000654r

Lee, H. S., & Coates, G. A. (2003). Effect of thermal pasteurization on Valencia orange juice color and pigments. LWT - Food Science and Technology, 36(1), 153–156. http://doi.org/10.1016/S0023-6438(02)00087-7

Li, S., Hu, Y., Hong, Y., Xu, L., Zhou, M., Fu, C., … Li, D. (2016). Analysis of the Hydrolytic Capacities of Aspergillus oryzae Proteases on Soybean Protein Using Artificial Neural Networks. Journal of Food Processing and Preservation, 40(5), 918–924. http://doi.org/10.1111/jfpp.12670

Lin, F. J. (2008). Solving multicollinearity in the process of fitting regression model using the nested estimate procedure. Quality and Quantity, 42(3), 417–426. http://doi.org/10.1007/s11135-006-9055-1

Maghrebi, M., Sammut, C., & Waller, T. S. (2014). Predicting the Duration of Concrete Operations Via Artificial Neural Network and by Focusing on Supply Chain Parameters. Building Research Journal, 61(1), 1–14. https://doi.org/10.2478/brj-2014-0001.

Matworks, 2018. Improve Shallow Neural Network Generalization and AvoidOverfitting. Retrieved on October 17. 2018 from: https://www.mathworks.com/help/deeplearning/ug/improve-neural-network-generalization-and-avoid overfitting.html;jsessionid=1c721f48297e43c235217 82435ab.ed

Meerasri, J., & Sothornvit, R. (2022). Artificial neural networks (ANNs) and multiple linear regression (MLR) for prediction of moisture content for coated pineapple cubes. *Case Studies in Thermal Engineering*, *33*(March), 101942. https://doi.org/10.1016/j.csite.2022.101942

Mekanik, F., Imteaz, M. A., Gato-Trinidad, S., & Elmahdi, A. (2013). Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes. Journal of Hydrology, 503, 11–21. http://doi.org/10.1016/j.jhydrol.2013.08.035

Meléndez-Martínez, A. J., Vicario, I. M., & Heredia, F. J. (2007). Review: Analysis of carotenoids in orange juice. Journal of Food Composition and Analysis, 20(7), 638–649. http://doi.org/10.1016/j.jfca.2007.04.006

Nejatdarabi S., & Mohebbi, M. (2023). Predicting the rehydration process of mushroom powder by multiple linear regression (MLR) and artificial neural network (ANN) in different rehydration medium. Journal of Food Measurement and Characterization. 17(2), 1962–1973.

Niazian, M., Sadat-Noori, S. A., & Abdipour, M. (2018). Artificial neural network and multiple regression analysis models to predict essential oil content of ajowan (Carum copticum L.). Journal of Applied Research on Medicinal a3nd Aromatic Plants, 9(November 2017), 124–131. http://doi.org/10.1016/j.jarmap.2018.04.001

Nikbakht, A. M., Motevali, A., & 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(2), 81–91. https://doi.org/10.1016/j.jssas.2013.01.005

Niresh, J. A., & Velnampy, T. (2014). Firm Size and Profitability: A Study of Listed Manufacturing Firms ed Manufacturing Firms in Sri Lanka. International Journal of Business and Management, 9(4), 57–64. https://doi.org/10.5539/ijbm.v9n4p57.

Pham, H.T.T., Bazmawe, M., Kebede, B., Buve, C., Hendrickx, M.E., Van Loey, A.M.,2019. Changes in the soluble and insoluble compounds of shelf-stable orange juice in relation to non-enzymatic browning during storage. J. Agric. Food Chem. 67 (46), 12854–12862.

Polydera, A.C., Stoforos, N.G., Taoukis, P.S., 2003. Comparative shelf life study an vitamin C loss kinetics in pasteurised and high pressure processed reconstitute orange juice. J. Food Eng. 60 (1), 21–29.

Prakash Maran, J., Sivakumar, V., Thirugnanasambandham, K., & Sridhar, R. (2013). Artificial neural network and response surface methodology modeling in mass transfer parameters predictions during osmotic dehydration of Carica papaya L. Alexandria Engineering Journal, 52(3), 507–516. http://doi.org/10.1016/j.aej.2013.06.007.

Rai, P., Majumdar, G. C., DasGupta, S., & De, S. (2005). Prediction of the viscosity of clarified fruit juice using artificial neural network: A combined effect of concentration and temperature. Journal of Food Engineering, 68(4), 527–533. http://doi.org/10.1016/j.jfoodeng.2004.07.003

Remini, H., Mertz, C., Belbahi, A., Achir, N., Dornier, M., & Madani, K. (2015). Degradation kinetic modelling of ascorbic acid and colour intensity in pasteurised blood orange juice during storage. Food Chemistry, 173, 665–673. https://doi.org/10.1016/j.foodchem.2014.10.069

Sablani, S. S., & Rahman, M. S. (2003). Using neural networks to predict thermal conductivity of food as a function of moisture content, temperature and apparent porosity. Food Research International, 36(6), 617–623. https://doi.org/10.1016/S0963-9969(03)00012-7

Saraceno, A., Aversa, M., & Curcio, S. (2012). Advanced Modeling of Food Convective Drying: A Comparison Between Artificial Neural Networks and Hybrid Approaches. Food and Bioprocess Technology, 5(5), 1694–1705. https://doi.org/10.1007/s11947-010-0477-3.

Shacham, M., & Brauner, N. (2014). Application of stepwise regression for dynamic parameter estimation. Computers and Chemical Engineering, 69, 26–38. http://doi.org/10.1016/j.compchemeng.2014.06.013

Shinoda, Y., Murata, M., Homma, S., & Komura, H. (2004). Browning and Decomposed Products of Model Orange Juice. Bioscience, Biotechnology, and Biochemistry, 68(3), 529–536. https://doi.org/10.1271/bbb.68.529

Sokač Cvetnić, T., Krog, K., Valinger, D., Gajdoš Kljusurić, J., Benković, M., Jurina, T., Jakovljević, T., Radojčić Redovniković, I., & Jurinjak Tušek, A. (2024). Application of Multivariate Regression and Artificial Neural Network Modelling for Prediction of Physicochemical Properties of Grape-Skin Compost. *Bioengineering*, *11*(3). https://doi.org/10.3390/bioengineering11030285

Stangierski, J., Weiss, D., & Kaczmarek, A. (2019). Multiple regression models and Artificial Neural Network (ANN) as prediction tools of changes in overall quality during the storage of spreadable processed Gouda cheese. *European Food Research and Technology*, *245*(11), 2539–2547. https://doi.org/10.1007/s00217-019-03369-y

Tehlah, N., Kaewpradit, P., & Mujtaba, I. M. (2016). Artificial neural network based modeling and optimization of refined palm oil process. Neurocomputing, 216, 489–501. https://doi.org/10.1016/j.neucom.2016.07.050.

Timmermans, R. A. H., Mastwijk, H. C., Knol, J. J., Quataert, M. C. J., Vervoort, L., Der Plancken, I. Van, … Matser, A. M. (2011). Comparing equivalent thermal, high pressure and pulsed electric field processes for mild pasteurization of orange juice. Part I: Impact on overall quality attributes. Innovative Food Science and Emerging Technologies, 12(3), 235–243. https://doi.org/10.1016/j.ifset.2011.05.001.

Tiryaki, S., & Aydın, A. (2014). An artificial neural network model for predicting compression strength of heat treated woods and comparison with a multiple linear regression model. Construction and Building Materials, 62, 102–108. http://doi.org/10.1016/j.conbuildmat.2014.03.041.

Tiwari, B. K., Muthukumarappan, K., O'Donnell, C. P., & Cullen, P. J. (2008). Effects of sonication on the kinetics of orange juice quality parameters. Journal of Agricultural and Food Chemistry, 56(7), 2423–2428. https://doi.org/10.1021/jf073503y.

Tiwari, B. K., O' Donnell, C. P., Muthukumarappan, K., & Cullen, P. J. (2009). Ascorbic acid degradation kinetics of sonicated orange juice during storage and comparison with thermally pasteurised juice. LWT - Food Science and Technology, 42(3), 700–704. https://doi.org/10.1016/j.lwt.2008.10.009.

Torkashvand, A. M., Ahmadi, A., & Nikravesh, N. L. (2017). Prediction of kiwifruit firmness using fruit mineral nutrient concentration by artificial neural network (ANN) and multiple linear regressions (MLR). Journal of Integrative Agriculture, 16(7), 1634–1644. https://doi.org/10.1016/S2095-3119(16)61546-0.

Tripathy, P. P., & Kumar, S. (2009). Neural network approach for food temperature prediction during solar drying. International Journal of Thermal Sciences, 48(7), 1452–1459. http://doi.org/10.1016/j.ijthermalsci.2008.11.014

Vervoort, L., Van Der Plancken, I., Grauwet, T., Timmermans, R. A. H., Mastwijk, H. C., Matser, A. M., … Van Loey, A. (2011). Comparing equivalent thermal, high pressure and pulsed electric field processes for mild pasteurization of orange juice: Part II: Impact on specific chemical and biochemical quality parameters. Innovative Food Science and Emerging Technologies, 12(4), 466–477. https://doi.org/10.1016/j.ifset.2011.06.003.

Wibowo, S., Grauwet, T., Kebede, B. T., Hendrickx, M., & Van Loey, A. (2015). Study of chemical changes in pasteurised orange juice during shelf-life: A fingerprinting-kinetics evaluation of the volatile fraction. Food Research International, 75, 295–304. https://doi.org/10.1016/j.foodres.2015.06.020.

Wibowo, S., Vervoort, L., Tomic, J., Santiago, J. S., Lemmens, L., Panozzo, A., … Van Loey, A. (2015a). Colour and carotenoid changes of pasteurised orange juice during storage. Food Chemistry, 171, 330–340. http://doi.org/10.1016/j.foodchem.2014.09.007

Wibowo, S., Grauwet, T., Santiago, J.S., Tomic, J., Vervoort, L., Hendrickx, M., Van Loey, A., (2015b). Quality changes of pasteurised orange juice during storage: a kinetic study of specific parameters and their relation to colour instability. Food Chem. 187, 140–151.

Xu, L., Xu, N., Zhu, X., Zhu, Y., Hu, Y., Li, D., & Wang, C. (2014). Modeling of Furfural and 5-Hydroxymethylfurfural Content of Fermented Lotus Root: Artificial Neural Networks and a Genetic Algorithm Approach. International Journal of Food Engineering, 10(4), 757–766. http://doi.org/10.1515/ijfe-2013-0068.

Yadav, A. K., & Chandel, S. S. (2017). Identification of relevant input variables for prediction of 1-minute time-step photovoltaic module power using Artificial Neural Network and Multiple Linear Regression Models. *Renewable and Sustainable Energy Reviews,* 77(December 2016), 955–969. http://doi.org/10.1016/j.rser.2016.12.029

Yang, T., Zheng, X., Vidyarthi, SK., Xiao, H., Yao, X., Li., Zang, Y., & Zhang, J. (2023). Artificial Neural Network Modeling and Genetic Algorithm Multiobjective Optimization of Process of Drying-Assisted Walnut Breaking. Foods, 12(9), 1–21. https://doi.org/10.3390/foods12091897

Yasar, A., Simsek, E., Bilgili, M., Yucel, A., & Ilhan, I. (2012). Estimation of relative humidity based on artificial neural network approach in the Aegean Region of Turkey. Meteorology and Atmospheric Physics, 115(1–2), 81–87. https://doi.org/10.1007/s00703-011-0168-2.

Yeom, H. W., Zhang, Q. H., & Chism, G. W. (2002). Inactivation of pectin methyl esterase in orange juice by pulsed electric fields. Journal of Food Science, 67(6), 2154–2159. http://doi.org/10.1111/j.1365-2621.2002.tb09519.x

Youssefi, S., Emam-Djomeh, Z., & Mousavi, S. M. (2009). Comparison of artificial neural network (ANN) and response surface methodology (RSM) in the prediction of quality parameters of spray-dried pomegranate juice. Drying Technology, 27(7), 910–917. http://doi.org/10.1080/07373930902988247

Yu, P., Low, M. Y., & Zhou, W. (2018). Design of experiments and regression modelling in food flavour and sensory analysis: A review. Trends in Food Science and Technology, 71(November 2017), 202–215. http://doi.org/10.1016/j.tifs.2017.11.013

Zheng, H., Jiang, L., Lou, H., Hu, Y., Kong, X., & Lu, H. (2011). Application of Artificial Neural Network (ANN) and Partial Least-Squares Regression (PLSR) to predict the changes of anthocyanins, ascorbic acid, total phenols, flavonoids, and antioxidant activity during storage of red bayberry juice based on fractal ana. Journal of Agricultural and Food Chemistry, 59(2), 592–600. http://doi.org/10.1021/jf1032476.

Zou, H. F., Xia, G. P., Yang, F. T., & Wang, H. Y. (2007). An investigation and comparison of artificial neural network and time series models for Chinese food grain price forecasting. Neurocomputing, 70(16–18), 2913–2923. http://doi.org/10.1016/j.neucom.2007.01.009.

Zulueta, A., Barba, F. J., Esteve, M. J., & Frígola, A. (2013). Changes in Quality and Nutritional Parameters During Refrigerated Storage of an Orange Juice-Milk Beverage Treated by Equivalent Thermal and Non-thermal Processes for Mild Pasteurization. Food and Bioprocess Technology, 6(8), 2018–2030. http://doi.org/10.1007/s11947-012-0858-x.