Journal of Food Engineering 109 (2012) 553–560

Contents lists available at SciVerse ScienceDirect

Journal of Food Engineering

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

Artificial neural network model for prediction of cold spot temperature in retort sterilization of starch-based foods

Yvan Antonio Llave, Tomoaki Hagiwara, Takaharu Sakiyama *

Department of Food Science and Technology, Tokyo University of Marine Science and Technology, 4-5-7 Konan, Minato-ku, Tokyo 108-8477, Japan

ARTICLE INFO

Article history:
Received 30 April 2011
Received in revised form 18 August 2011
Accepted 22 October 2011
Available online 30 October 2011

Keywords:
Artificial neural network
Retort processing
Starch dispersion
Cold spot temperature
F_0 value

ABSTRACT

An artificial neural network (ANN) model was developed for prediction of the cold spot temperature profile during retort processing using starch dispersion (STD) as a model food. STDs of different concentrations were prepared by mixing corn starch powder with distilled water at 90 °C for 30 min. Each of the partially gelatinized STDs thus prepared was filled in retort pouches and processed in a retort under various combinations of holding temperature, holding time, and rotational speed. Thermocouples were inserted into selected pouches one by one to monitor the cold spot temperature at regular intervals. The profiles of cold spot temperature together with retort temperature thus obtained were served to ANN modeling as training or validation data. Back-propagation network was chosen as the network model. Input variables for the model were current and past temperatures of the cold spot (T_n, T_{n-1}, and T_{n-2}) and current retort temperature θ_n and current time t_n. Output was the temperature of the cold spot at the next time step T_{n+1}. A model with 2 hidden layers, which contained 11 and 15 nodes, respectively, was the best among the models tested. Using the model developed, prediction of a whole profile of the cold spot temperature was tested, starting from temperature data of the first three time steps with a whole profile of retort temperature monitored. The results showed very good performance of the model, relative errors for F_0 value prediction being less than 2%.

1. Introduction

Thermal processing is a major technique for producing safe and shelf-stable foods. Heating enhances bacterial lethality but usually lowers product quality. Thus the technique is still desired to keep evolving for better product quality with better energy utilization, and more efficient production.

Ideal approach to the control of retort processing would be based upon accurate prediction of the product cold spot temperature in response to dynamic retort temperature profile. Researchers have developed computer models capable of simulating thermal processing based on principles of heat transfer (Teixeira and Manson, 1982; Tandon and Bhowmik, 1986; Tucker, 1991; Tucker and Holdsworth, 1991; Noronha et al., 1995). Those models can predict the product cold spot temperature in response to dynamic temperature variations brought about during a retort process. However, they require several parameters including thermo-physical properties of materials (heat capacity, density, thermal conductivity, etc.) and heat transfer rate constant. For precise simulation, we need precise values of the parameters. Moreover, the parameters cannot be always constant. In fact, because starch dispersion may suffer from viscosity change during heating, some models (Yang and Rao, 1998; Tattiyakul et al., 2001, 2002) have included viscosity change as a function of temperature to improve the accuracy of the predictions. From practical point of view, however, it is often hard to obtain the precise temperature dependence of all the parameters, which makes proper modeling even harder.

In some cases, when starch-based food is thermally processed under intermittent agitation, broken-heating behavior is known to occur (Stumbo, 1973; Tattiyakul and Rao, 2000; Holdsworth and Simpson, 2008). Broken-heating behavior is characterized by more than one distinct straight portion in a heat penetration curve, which is a plot of the logarithmic difference between retort temperature and product temperature versus processing time. More than one straight portion in a heat penetration curve indicates transition of overall heat transfer rate. Broken-heating behaviors in many starch-based foods have been attributed mostly to starch gelatinization (Noronha et al., 1995; Yang and Rao, 1998; Tattiyakul et al., 2001, 2002), which causes change of the major heat penetration mechanism inside the product from convection to conduction. For accurate prediction, we should know whether and when broken-heating occurs. Most of the traditional models have a weak point to cope with such change in heat penetration mechanism.

Recently artificial neural network (ANN) gained popularity as a modeling tool in many disciplines of engineering and science. The

* Corresponding author. Tel.: +81 3 5463 0619; fax: +81 3 5463 0699.
E-mail address: sakiyama@kaiyodai.ac.jp (T. Sakiyama).

0260-8774/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jfoodeng.2011.10.024

Nomenclature

f_h	heating rate factor (min)
j_h	heating lag factor
h	number of hidden layers (h = 1 or 2)
k_i	number of nodes in layer i (i = 0,1,...,h + 1)
t	time (min)
t_n	time at the nth monitoring step (min)
T	cold spot temperature (°C)
T_n	cold spot temperature at the nth monitoring step (°C)
$\tilde{T}_n$	cold spot temperature predicted for the nth monitoring step (°C)
$w^{[i]}_{k,l,m}$	weight for $x^{[i-m]}_l$ connected to node k in layer i ($i \geqslant 1$, m = 1,...,i)
$w^{[i]}_{k,1,0}$	bias for node k in layer i ($i \geqslant 1$)
$x^{[i]}_k$	output from node k in layer i ($i \geqslant 0$, k = 1,...,k_i)
θ	retort temperature (°C)
θ_n	retort temperature at the nth monitoring step (°C)

most attractive feature of ANN may be its ability to find and learn a complex relationship hidden among a number of variables. Some recent applications of ANN to modeling of thermal process are following: prediction of optimal processing conditions for conduction-heated foods in cans (Sablani et al., 1995), prediction of heat transfer coefficients associated with liquid/particle mixtures in cans (Sablani et al., 1997), prediction of thermal process lethality for a range of can sizes and operating conditions (Afagui et al., 2001), prediction of heating rate parameters in sterilization process (Mittal and Zhang, 2002), modeling and optimization of thermal processing with variable retort temperature for conduction heated foods (Chen and Ramaswamy, 2002), prediction of the cold spot temperature of a liver paste filled in cans (Gonçalves et al., 2005), modeling and optimization of the multiple ramp-variable retort temperature control for thermal processing (Chen and Ramaswamy, 2004), and modeling of heat transfer to canned particulate fluids under axial rotation processing (Dwivedi and Ramaswamy, 2010).

In this study, an ANN model was developed for prediction of the cold spot temperature profile during retort processing using starch dispersion (STD) as a model food. STD is selected because it shows broken-heating behavior depending on conditions. The ANN model was trained with various temperature profiles including broken-heating behaviors to make it applicable to any situation encountered in retort processing of STDs.

2. Materials and methods

2.1. Sample preparation

Corn starch powder (Sigma–Aldrich, MO, USA), containing 73% amylopectin and 27% amylose, was used to prepare 3%, 10%, and 20% STDs by mixing thoroughly with distilled water in a jacketed kettle mixer (STM-5, Shinagawa Machinery Works Co. Ltd., Nara, Japan). During the mixing, the mixture was heated at 90 °C for 30 min for partial gelatinization, while evaporation loss was compensated for by water addition. After the partially gelatinized STD was cooled to room temperature, laminated retort pouches (PET12/AL9/NY15/CP60; HR-Type, Meiwa Sanshou Co. Ltd., Osaka, Japan) were filled with 250 g portion of STD and sealed under vacuum (Old Rivers FVC-II, Furukawa Mfg. Co. Ltd., Tokyo, Japan). The size of the pouch was 130 × 180 mm in most experimental runs. In a few runs for validation purpose, pouches of a slightly larger size of 130 × 230 mm were also used.

2.2. Experimental procedure

Retort processing was carried out using a pilot scale retort (RCS-40RTGN, Hisaka Works Ltd., Japan) in its stationary or rotary mode. Ten trays, each of which was loaded with four pouches, were piled up in the retort. Preliminary experiments showed that the slowest heating occurred in a pouch on the second tray adjacent to the bottom one. Therefore, a thermocouple was inserted into the geometric center of each pouch loaded on the second tray to obtain heat penetration data. Temperature readings of the thermocouple were computer-recorded at 2-s intervals via a data logger. The temperature history with the slowest rise was chosen as the cold spot temperature. We employed a range of retort processing conditions as listed in Table 1. Each run started with pre-heating by holding retort temperature at 90 °C for 3 min. Further heating schedule was designed for F_0 value to reach 6.0 min at least. F_0 value, a measure of accumulated lethality, denotes the equivalent exposure time at 121.1 °C calculated for a microorganism with a Z-value equal to 10 °C, as shown in Eq. (1).

$$F_0 = \int_0^t 10^{(T-121.1)/10} dt \tag{1}$$

where T (°C) is the cold spot temperature at time t (min). At the end of the sterilization, cooling was initiated with air pressurization to avoid inflation of the pouches.

2.3. Development of ANN model

The main feature of the model was prediction of the cold spot temperature step by step from the following data: cold spot temperatures monitored at the latest three time steps (T_n, T_{n-1}, and T_{n-2}), retort temperature (θ_n), and time (t_n) at the step. The output variable was the cold spot temperature at the next time step (T_{n+1}). The time step width was the same as the interval of temperature monitoring. Thus the model was designed to calculate the cold spot temperature shortly ahead on the basis of the short temperature history already known.

As shown in Fig. 1, ANN model architecture employed here had h + 2 layers, where h was the number of hidden layer (h = 1 or 2). Layer i (i = 0,...,h + 1) was composed of k_i nodes. Layer 0 was the input layer having five nodes (k_0 = 5), each of which took the pre-treated value of T_n, T_{n-1}, T_{n-2}, θ_n, or t_n as a single input and put it out as it was to the hidden layer(s). The pre-treatment of input data was performed by applying logarithmic function to the base 10 twice, as shown in Eq. (2), in which T_n is taken as an example of input data.

$$T'_n = \log_{10}(\log_{10}(T_n)) \tag{2}$$

where T'_n is the pre-treated value of T_n. Hereafter the output from each node of layer 0 is referred to as $x^{[0]}_k$ (k = 1,...,k_0) for a time step. As for the hidden layer(s), node k (k = 1,...,k_i; $k_i \leqslant 15$) in layer i ($1 \leqslant i \leqslant h$) accepts inputs, $x^{[i-m]}_l$ (l = 1,...,k_{i-m}; m = 1,...,i), from all the prior layers and makes an output, $x^{[i]}_k$, according to Eq. (3).

$$x^{[i]}_k = f\left(\sum_{m=1}^{i}\sum_{l=1}^{k_{i-m}} w^{[i]}_{k,l,m}\cdot x^{[i-m]}_l + w^{[i]}_{k,1,0}\right) \tag{3}$$

where $w^{[i]}_{k,l,m}$ and $w^{[i]}_{k,1,0}$ (k = 1,...,k_i) represent connection weight for each input and bias, respectively. Transfer function, indicated as f in

Table 1
Retort processing conditions employed for acquisition of training and testing data.

Parameter	Experimental range
Rotational speed (rpm)	0, 5, 10, 15, 20, 25
Holding temperature (°C)	118, 120, 122
Holding time (min)	12, 15, 20, 22

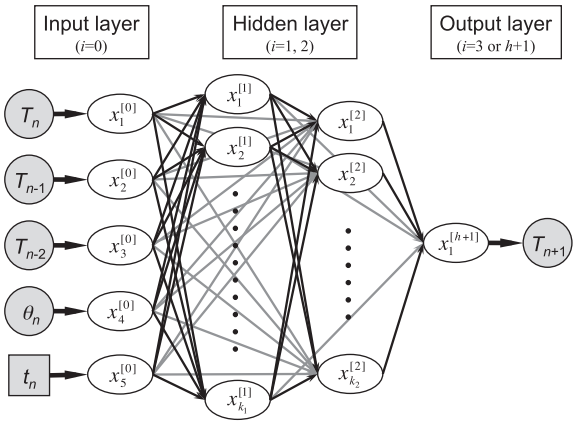

Fig. 1. Network architecture employed for ANN model in this study. A model with two hidden layers is illustrated. The input layer (layer 0) had five nodes. Each node in hidden and output layers was connected with nodes in prior layers.

Eq. (3), is usually selected from s-shaped functions. In the model finally established, hyperbolic tangent function was employed as the transfer function. The highest layer $h+1$ was the output layer having a single node ($k_{h+1}=1$) to give a single output $x_1^{[h+1]}$ according to Eq. (4).

$$x_1^{[h+1]} = f\left(\sum_{m=1}^{h}\sum_{l=1}^{k_{i-m}} w_{1,l,m}^{[h+1]} \cdot x_l^{[h+1-m]} + w_{1,1,0}^{[h+1]}\right) \tag{4}$$

The prediction candidate for T_{n+1}, referred to as $\tilde{T}_{n+1}$ hereafter, was obtained by post-treatment of the output $x_1^{[h+1]}$ using the inverse transformations of Eq. (2).

Because prediction error depends on a set of weights and bias, $w_{k,l,m}^{[i]}$ and $w_{k,1,0}^{[i]}$, we can improve the weight set by a well-known procedure called back-propagation, to obtain prediction candidates closer to the corresponding monitored values (training data). Iterations of back propagation according to Extended Delta–Bar-Delta training rule, with testing processes at appropriate intervals using sets of data (testing data) preliminarily excluded from the training data, gave the trained model with the optimum weight set for the layer configuration. The testing stage is important to avoid overtraining of the network. The iterations of back propagation and testing at appropriate intervals were conducted with a commercial software, NeuralWorks Professional II/Plus (Neuralware Inc., Pittsburg, PA, USA). Hereafter the model thus trained was denoted as BP-5-k_1-1 or BP-5-k_1-k_2-1 on the basis of the number of nodes contained in each layer.

The optimal layer configuration for the best predictive performance was selected on a trial-and-error basis among 240 possible configurations, from BP-5-1-1 to BP-5-15-15-1. The lower prediction error was sought by increasing the number of nodes in each hidden layer. As measures of the prediction error, we employed the coefficient of determination, R^2, together with the mean and standard deviation of absolute values of relative error, $\overline{|\mathrm{RE}|}$ and $\mathrm{SD}_{|\mathrm{RE}|}$. The measure R^2 is defined as follows:

$$R^2 = 1 - \frac{\sum_{n=1}^{N}(T_n - \tilde{T}_n)^2}{\sum_{n=1}^{N}(T_n - \bar{T})^2} \tag{5}$$

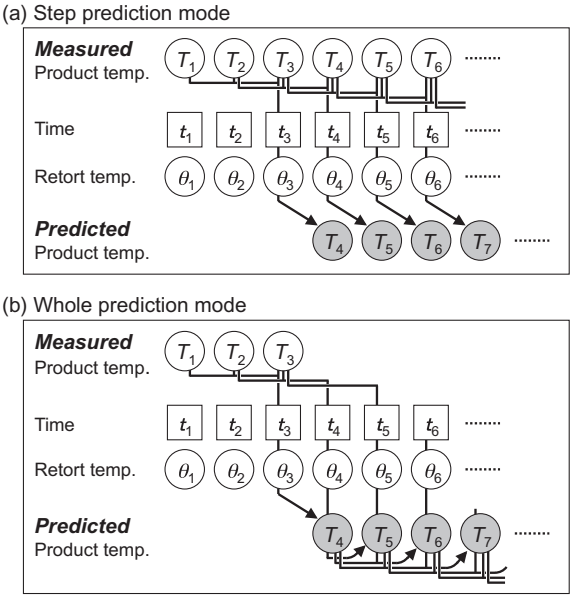

Fig. 2. Prediction modes for validation of the model constructed: (a) step prediction mode and (b) whole prediction mode. Five input variables in each prediction step are connected with each other by an arrow pointing to an output variable indicated as a shaded circle.

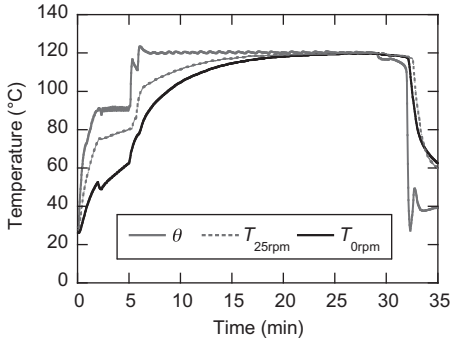

Fig. 3. Comparison of cold spot temperature profiles of 20% STD processed with and without rotation (25 rpm) under the same retort temperature program. For the sake of simplicity, only the retort temperature profile monitored with rotation is illustrated.

where $\tilde{T}_n$ is the cold spot temperature predicted by ANN model for the nth time step, $\bar{T}_n$ is the mean of the cold spot temperatures monitored through the experiment, and N is the total number of time step. The measures $\overline{|\mathrm{RE}|}$ and $\mathrm{SD}_{|\mathrm{RE}|}$ are defined as follows:

$$|\mathrm{RE}|_n = \frac{|T_n - \tilde{T}_n|}{T_n} \tag{6}$$

$$\overline{|\mathrm{RE}|} = \frac{1}{N}\sum_{n=1}^{N}|\mathrm{RE}|_n \tag{7}$$

$$\mathrm{SD}_{|\mathrm{RE}|} = \sqrt{\frac{\sum_{n=1}^{N}(|\mathrm{RE}|_n - \overline{|\mathrm{RE}|})^2}{N-1}} \tag{8}$$

The values, R^2 and $\overline{|\mathrm{RE}|}$, have been often employed as measures of predictive accuracy for ANN models (Sablani et al., 1997; Afagui et al., 2001; Chen and Ramaswamy, 2002, 2004, 2006a,b; Mittal and Zhang, 2002; Sablani and Rahman, 2003; Dwivedi and Ramaswamy, 2010).

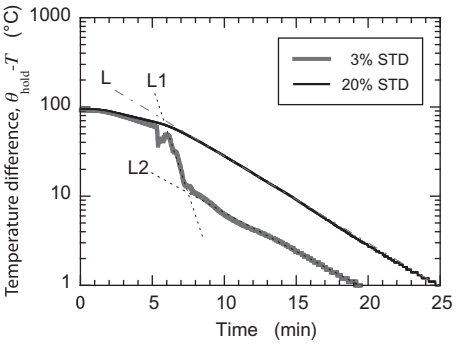

Fig. 4. Typical examples of heat penetration curve for STD. STDs of 3% and 20% were processed with rotation at 25 rpm and 5 rpm, respectively. Difference between the holding temperature and the cold spot temperature was plotted in logarithmic scale versus processing time. The curve for 20% STD has a single asymptotic line (L), while that for 3% STD has two straight portions (L1 and L2).

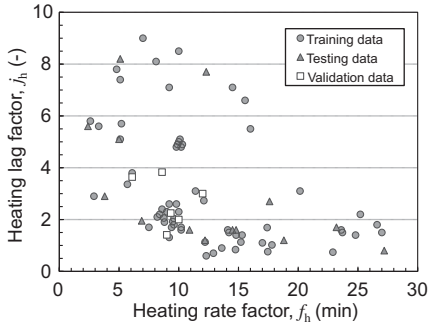

Fig. 5. Distribution of the heat penetration curves used for training, testing, and validation of ANN model as expressed by combination of heating rate factor f_h and heating lag factor j_h.

Table 2
Comparison of prediction errors for time–temperature profile produced at ANN-testing stage by trained networks with best architectures.

Network architecture		R^2	$\overline{\lvert RE \rvert}$[a]
Type I	BP-5-7-12-1	0.9955	0.0103 ± 0.0092
	BP-5-11-15-1	0.9955	0.0050 ± 0.0068
	BP-5-13-14-1	0.9955	0.0054 ± 0.0062
Type II	BP-5-5-0-1	0.9951	0.0219 ± 0.0115
	BP-5-7-1-1	0.9951	0.0234 ± 0.0158
	BP-5-4-14-1	0.9936	0.0220 ± 0.0114

[a] Mean ± standard deviation of absolute values of relative error for all the time steps.

2.4. Validation of predictive performance of ANN model

Predictive performance of the model obtained as mentioned above was checked for temperature profiles not previously encountered by the model. Four new temperature profiles of STD during retort processing were obtained for this purpose. The ANN model developed above predicts the next step temperature from the latest three temperatures already known. The first check was based on such step-by-step mode of prediction. On the other hand, we can take an output temperature as the input at the next time step. Thus, even with the model developed, it is possible to calculate a whole temperature profile starting from temperatures at the first three time steps together with a whole course of retort temperature. Hereafter, this prediction mode is referred to as whole prediction mode, while the original one is referred to as step prediction mode. Input data employed at initial time steps for both prediction modes are shown schematically in Fig. 2. Calculation by the whole prediction mode was performed on a spreadsheet, as was suggested by Hajmeer et al. (2006), based on Eqs. (2)–(4) with all the weights determined through training of the ANN model.

3. Results and discussion

3.1. Characteristics of heat penetration

Major factors affecting retort process of STDs considered here, in addition to retort temperature program, are the starch concentration and the rotational speed of retort. Fig. 3 shows typical temperature histories of 20% STD with and without rotation at 25 rpm during retort processing. Retort temperature programs in the two cases were the same, resulting in similar retort temperature histories. Thus, for the sake of simplicity, only the retort temperature history for the case with rotation is shown in Fig. 3. As already mentioned, the retort temperature was initially controlled to keep 90 °C for 3 min, thereafter it was raised to a holding temperature (120 °C in these cases). Comparison of these two cases shows that rotation definitely increased heat penetration rate.

Fig. 4 shows typical heat penetration curves, logarithm of difference between holding temperature and product temperature as a function of processing time, for 3% and 20% STDs processed with rotation. For 20% STD, the curve gave a single straight line after several minutes of initial lag period. The reciprocal of slope of the asymptote, called heating rate factor, f_h, which means time for one log change in ordinate, was determined to be 10.0 min in this case. Similar heat penetration curve with a single straight portion was obtained in every run for 20% STD and also for 10% STD. On the other hand, the curve for 3% STD after an initial lag period was approximated by two straight lines, indicating two stages of heat penetration, recognized as a broken-heating behavior. The value of f_h was initially as low as 2.6 min. But after the break point, it became 11.0 min, a comparable value to the case of 20% STD. The significant increase in f_h value indicates decrease in heat penetration rate, which is probably attributed to sol–gel transition caused by starch gelatinization, as known for many starch-based foods (Noronha et al., 1995; Yang and Rao, 1998; Tattiyakul et al., 2002). It has been recognized that low starch concentrations with intermittent agitation are apt to bring about broken-heating

Table 3-1
Weights of connections, $w^{[1]}_{k,l,m}$, and biases, $w^{[1]}_{k,1,0}$ ($m = 0$, $l = 1$), to node k on hidden layer 1 of BP-5-11-15-1.

m	l	k										
		1	2	3	4	5	6	7	8	9	10	11
1	1	−0.07216	0.064064	−0.116733	0.24463	−0.092895	0.035416	−0.170997	0.304539	−0.372617	−0.318406	0.408249
1	2	−0.11744	0.033343	0.005981	0.016803	−0.040465	−0.230484	−0.029974	0.090115	0.346221	−0.194951	−0.101963
1	3	−0.347027	0.199926	0.303523	0.901079	0.395982	−0.202817	−0.522875	0.086771	0.028165	0.343548	−0.435095
1	4	0.111007	0.120675	0.185015	0.221771	−0.286663	−0.007606	−0.847707	−0.088135	0.230044	0.211095	−0.047548
1	5	−0.085652	1.014736	−0.261032	−0.540751	−0.400346	0.200062	−0.71432	0.036225	−0.149112	0.013083	0.239624
0	1	0.241929	0.189456	−0.139505	−0.701701	−0.051167	−0.254199	1.01424	0.125151	−0.261349	−0.256727	−0.136715

Table 3-2
Weights of connections, $w^{[2]}_{k,l,m}$, and biases, $w^{[2]}_{k,1,0}$ ($m = 0$, $l = 1$), to node k on hidden layer 2 of BP-5-11-15-1.

m	l	k														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	1	−0.031643	0.268168	0.168603	0.248962	0.182196	0.08535	0.138324	−0.125173	−0.004945	−0.02707	−0.158068	0.14659	−0.195059	0.095242	−0.145672
2	2	0.075335	−0.132403	−0.296853	0.232809	0.052149	−0.138666	0.190902	−0.084293	0.284716	−0.201624	−0.160931	−0.067657	0.171612	−0.293111	−0.013684
2	3	−0.272312	0.236833	−0.160158	−0.196662	−0.454523	−0.015697	0.217596	0.111918	−0.030054	−0.082568	0.201232	−0.244598	0.185659	0.02687	0.349619
2	4	−0.450007	0.282049	−0.297276	0.136865	−0.127584	−0.048257	0.044251	0.024161	−0.164288	−0.188638	0.241774	−0.17017	0.218022	0.039188	0.274366
2	5	−0.398197	−0.10053	−0.30111	−0.077689	−0.264343	−0.150105	0.301611	0.075896	0.206774	−0.094224	−0.191393	−0.06418	−0.05102	−0.091652	−0.2766
1	1	−0.056397	0.227936	0.03134	0.101068	−0.211128	−0.069342	0.18879	0.142026	0.295746	0.05874	0.002087	0.240806	−0.285312	−0.221821	0.22932
1	2	−0.363586	−0.128663	0.357585	0.113023	0.180462	0.028305	0.284123	0.099488	−0.245165	0.163219	−0.127417	−0.101289	−0.191396	0.02668	−0.126713
1	3	0.141329	−0.085174	−0.160031	−0.128489	−0.245488	−0.176747	−0.2708	0.037567	−0.224799	−0.237952	−0.240095	−0.308018	−0.157652	0.146542	0.042299
1	4	0.16313	0.231103	−0.228775	−0.173623	−0.024903	−0.054978	0.302359	−0.135789	0.206813	−0.274528	−0.002782	0.06588	0.136742	−0.047953	−0.053933
1	5	0.394024	−0.132319	−0.176392	0.070889	−0.08551	−0.056266	0.287695	−0.068245	0.145451	−0.049712	−0.01534	−0.129885	0.184563	−0.243525	−0.171872
1	6	0.186804	−0.166311	−0.153209	0.308781	0.076135	0.129882	−0.188236	−0.191796	0.274788	0.11877	0.092271	−0.169223	−0.016318	0.008879	0.161107
1	7	0.026769	−0.132224	−0.370843	0.011681	−0.593005	−0.365663	0.262126	−0.254861	0.073139	−0.316793	−0.058268	0.311439	0.149565	0.082591	0.088738
1	8	−0.079351	−0.22662	0.323933	−0.199316	−0.170161	0.228698	0.139915	0.040362	0.095019	−0.270028	0.002804	−0.072309	−0.225061	0.005462	−0.015282
1	9	−0.202552	0.084916	−0.133355	0.268127	0.034366	0.020566	−0.181881	−0.007984	0.163932	0.215014	0.10515	−0.039174	−0.207798	−0.137021	0.140354
1	10	0.118788	−0.230704	0.202039	0.024539	0.263404	−0.192983	0.184582	0.051856	−0.236634	−0.266906	−0.256187	0.184324	0.142783	−0.075569	−0.022533
1	11	−0.144736	0.243941	−0.282637	0.052219	−0.212628	0.236061	−0.073631	−0.254831	0.226597	−0.278086	−0.266696	−0.268707	−0.011429	0.089278	0.056053
0	1	−0.271853	0.085521	−0.031239	−0.215167	−0.251758	−0.07275	−0.022524	0.075	−0.146561	0.005728	−0.237796	−0.222658	−0.111075	0.12146	0.185313

Table 3-3
Weights of connections, $w^{[3]}_{1,l,m}$, and bias, $w^{[3]}_{1,1,0}$ ($m = 0$, $l = 1$), to the node on output layer of BP-5-11-15-1.

m	l															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	−0.199688	0.056659	−0.032404	0.140934	−0.094153	−0.108568	−0.023172	0.171969	−0.031515	0.179055	−0.205381	–	–	–	–
1	0.438569	0.015969	−0.268172	−0.160161	−0.366338	−0.145059	−0.017018	−0.009095	0.071753	−0.023777	0.089336	0.072486	0.202592	−0.094482	0.138529
0	0.044207	–	–	–	–	–	–	–	–	–	–	–	–	–	–

Table 4
Comparison of prediction errors by BP−5-11−15-1 for validation runs.

Run	Conditions					Step prediction			Whole prediction		
	STD (%)	Pouch (mm × mm)	Rot[a] (rpm)	Temp.[b] (°C)	Time[c] (min)	Temperature profile		F_0	Temperature profile		F_0
						R^2	$\overline{\lvert RE \rvert}$[d]	RE[e]	R^2	$\overline{\lvert RE \rvert}$[d]	RE[e]
A	3	130 × 230	25	120	15	0.9992	0.0052 ± 0.073	−0.0351	0.9979	0.0164 ± 0.0188	−0.0195
B	3	130 × 230	0	120	20	0.9983	0.0042 ± 0.0049	−0.0320	0.9981	0.0143 ± 0.0192	−0.0046
C	10	130 × 180	15	120	20	0.9981	0.0045 ± 0.0055	0.0073	0.9983	0.0132 ± 0.0199	0.0043
D	20	130 × 180	5	120	22	0.9993	0.0060 ± 0.0069	0.0333	0.9983	0.0155 ± 0.0154	−0.0151

[a] Rotation speed.
[b] Holding temperature.
[c] Holding time.
[d] Mean ± standard deviation of absolute values of relative error for all the time steps.
[e] Relative error of the final F_0 value.

behaviors (Stumbo, 1973; Tattiyakul and Rao, 2000; Holdsworth and Simpson, 2008).

From each set of data monitored, a heat penetration curve was drawn to determine its heating rate factor, f_h, together with another characteristic parameter called heating lag factor, j_h, relating to the intercept of the asymptote of heat penetration curve on the vertical axis defining the beginning of the heating period (Holdsworth and Simpson, 2008).

3.2. Construction of ANN model

Seventy heat penetration curves (temperature profiles of the cold spot and the retort) obtained through retort processing of STDs were supplied for network training and testing. They were divided into a training subset composed of 55 curves (58,997 time steps) and a testing subset composed of 15 curves (15,268 time steps). The heat penetration curves included in each subset are indicated in Fig. 5 by their characteristic parameters of heating rate index f_h and heating lag factor j_h.

Although we finally employed hyperbolic tangent as the transfer function and Extended Delta–Bar-Delta as the learning rule to search for the best layer configuration, another combination of transfer function and learning rule, standard sigmoid transfer function and Delta learning rule, was also tested for comparison. The two transfer functions are quite similar to each other as to their s-shape, though their output ranges are different. The latter combination, a rather familiar combination for ANN modeling, is referred to as type II network, while the former is referred to as type I network.

Table 2 lists three best layer configurations of each network type selected on the basis of low levels of prediction error in step prediction mode. For type I network, the best layer configuration, which produced the highest R^2 value and the lowest $\overline{|RE|}$ value, was composed of two hidden layers with 11 nodes on the first layer and 15 on the second layer. The value of $\overline{|RE|}$ for this configuration was 0.50% with a standard deviation of 0.68%. For type II network, the best layer configuration consisted of one hidden layer with five nodes, attaining $\overline{|RE|}$ value of 2.19% with a standard deviation of 1.15%. Thus, type I network showed the better predictive performance.

We decided to employ BP-5-11-15-1 of type I network, which gave the best prediction among the models tested. For further improvement of its predictive performance, a variety of values of tuning parameters for training, namely learning rate η ($0 < \eta \leqslant 1$; 0.5 as the default value of the software) and momentum μ($0 \leqslant \mu \leqslant 1$; 0.4 as the default value), were tried. We finally set $\eta = 0.095$ and $\mu = 0$, because this combination yielded the highest R^2 value and the lowest $\overline{|RE|}$ value. The set of connection weights and biases thus obtained finally for BP-5-11-15-1 of type I network are listed in Tables 3-1, 3-2 and 3-3.

3.3. Predictive performance of the model constructed

BP-5-11-15-1 of type I network, which was found to show the best performance among tested, was subjected to its validation. Four experimental runs A–D were newly performed to provide validation data. Their experimental conditions are listed in Table 4. Runs A and B with 3% STD were typical cases of broken-heating. Runs C and D with concentrated STDs (10% and 20%) were cases producing gel-like samples during retort processing, thus showing heat penetration behaviors different from broken-heating. Not only time–temperature profile predicted, but also F_0 value, calculated based on the predicted time–temperature profile, were compared with experimental ones. The prediction was performed in step prediction mode as well as in whole prediction mode.

In Fig. 6, the cold spot temperatures predicted by BP-5-11-15-1 in step prediction mode are compared with experimental ones

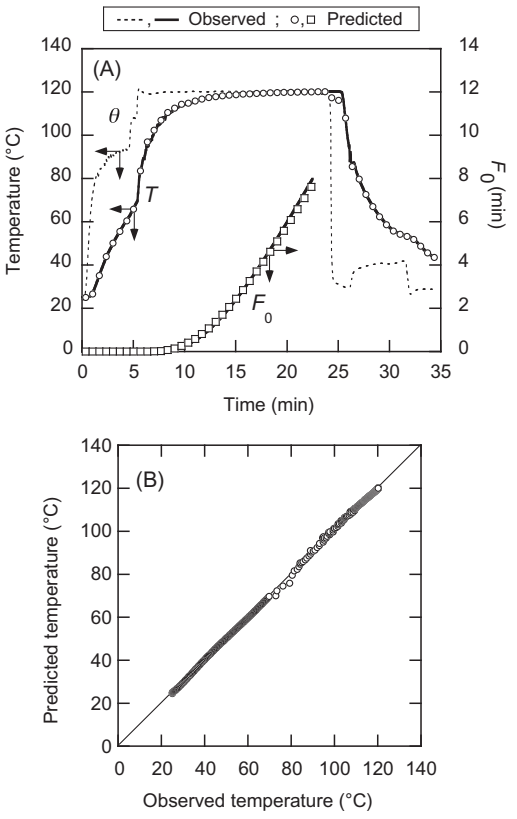

Fig. 6. Prediction by the model BP-5-11-15-1 in step prediction mode for validation run A. (a) Predicted time–temperature profile and F_0 value in comparison with experimental ones. (b) Correlation between predicted and observed temperatures at the cold spot.

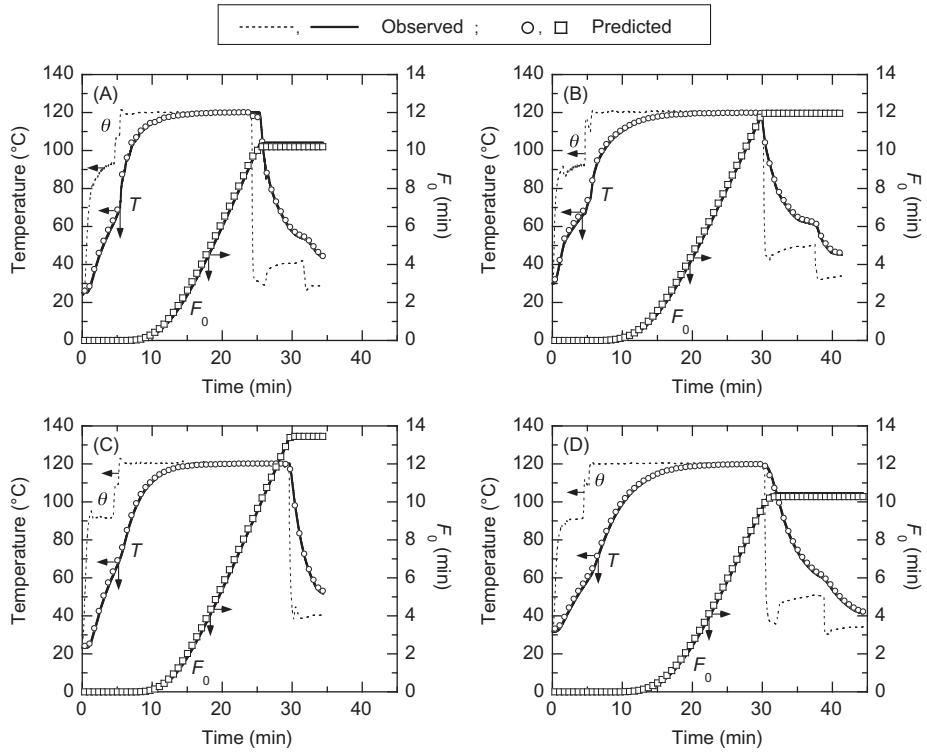

Fig. 7. Time–temperature profiles and F_0 values of runs A–D predicted by BP-5-11-15-1 in whole prediction mode in comparison with experimental ones.

during a retort process for Run A (3% STD with rotation at 25 rpm and holding at 120 °C for 15 min), for example. A very high degree of coincidence was observed between them ($R^2 > 0.999$), indicating an excellent adjustment of ANN model to the retort process behavior.

Table 4 compares prediction errors of the both prediction modes for each validation run. The values of $\overline{|RE|}$ for the time–temperature profile were found slightly higher in whole prediction mode than in step prediction mode. On the contrary, whole prediction mode gave lower relative errors to F_0 value than step prediction mode. This can be because whole prediction mode gave large $|RE|_n$ values for the cold spot temperature at beginning, where temperature was too low to affect F_0 value. At higher temperatures, whole prediction mode gave rather good prediction, which lowered the relative error of F_0 value. To illustrate this, time–temperature profiles and F_0 values for runs A–D predicted by BP-5-11-15-1 in whole prediction mode are shown in Fig. 7 in comparison with corresponding experimental ones. Good similarity between the experimental values and the predicted ones were found especially at a high temperature region for all of the four runs.

Successful results were obtained by whole prediction mode even though only three data of cold spot temperature at the very beginning of the heating were used. This may be because the model distinguishes the sample type on the basis of tiny increases in cold spot temperature during the initial time steps. The average increase in cold spot temperature in 4 s of initial heating was 0.9 °C for 3% STD, 0.5 °C for 10% STD, and 0.2 °C for 20% STD. The common pre-heating schedule and initial rapid elevation of retort temperature probably help the successful estimation of sample type by the model from cold spot temperatures at the very beginning of heating. Alternatively, the common pre-heating program and initial rapid elevation of retort temperature may be requisite conditions for the applicability of the model developed.

Using a similar ANN model architecture, namely taking known temperatures at three time steps as input, Gonçalves et al. (2005) reported that relative prediction errors of F_0 value were lower than 2.6% for food cans which underwent purely conductive heat transfer. Although the experimental situations and ANN model construction were not the same, we obtained similar, yet better, accuracy of prediction. Moreover, several types of heat transfer mode were included in the ANN model constructed in this study: conduction, convection, and their transition (broken-heating). Thus the model constructed in this study expanded the scope of the prediction by ANN to retort foods with non-conductive heat transfers, including broken-heating.

ANN model is recognized as a black box regression model (Sablani et al., 1997; Dwivedi and Ramaswamy, 2010), not based on heat transfer principles. Thus, in general, ANN model is much less reliable outside the range of training data set. This may be an obstacle for ANN model to extend the applicable range of conditions. Further investigation on the applicable conditions or limitations of the model is desired.

4. Conclusions

ANN model BP-5-11-15-1 developed by training with heat penetration data for STDs was found successful to predict time–temperature profile and thermal lethality with high accuracy under variety of processing conditions, irrespective of heat transfer modes. Its excellent performance for STDs suggests that the model may be useful for prediction and control of retort process of foods containing starch as a thickening agent. Further study is desired to extend the range of target foods for the model developed.

Acknowledgement

This work was supported in part by a research grant from the Iijima Memorial Foundation for the Promotion of Food Science and Technology.

References

Afagui, M., Ramaswamy, H.S., Prasher, S.O., 2001. Thermal process calculations using artificial neural network models. Food Research International 34 (1), 55–65.

Chen, C.R., Ramaswamy, H.S., 2002. Modeling and optimization of variable retort temperature (VRT) thermal processing using coupled neural networks and genetic algorithms. Journal of Food Engineering 53 (3), 209–220.

Chen, C.R., Ramaswamy, H.S., 2004. Multiple ramp-variable retort temperature control for optimal thermal processing. Food and Bioproducts Processing 82 (1), 78–88.

Chen, C.R., Ramaswamy, H.S., 2006a. Visual Basics computer simulation package for thermal process calculations. Chemical Engineering and Processing 46, 603–613.

Chen, C.R., Ramaswamy, H.S., 2006b. Modeling food thermal processes using artificial neural networks. In: Sun, D.-W. (Ed.), Thermal Food Processing: New Technologies and Quality Issues. CRC Press, Taylor & Francis Group, Boca Raton, FL, pp. 107–131.

Dwivedi, M., Ramaswamy, H.S., 2010. Artificial neural network modeling of heat transfer to canned particulate fluids under axial rotation processing. International Journal of Food Engineering 6 (3), article 5.

Gonçalves, E.C., Minim, L.A., Coimbra, J.S.R., Minim, V.P.R., 2005. Modeling sterilization process of canned foods using artificial neural networks. Chemical Engineering and Processing 44 (12), 1269–1276.

Hajmeer, M., Basheer, I., Cliver, D.O., 2006. Survival curves of *Lysteria monocytogenes* in chorizos modeled with artificial neural networks. Food Microbiology 23 (6), 561–570.

Holdsworth, D., Simpson, R., 2008. Thermal Processing of Packaged Foods, second ed. Springer, New York.

Mittal, G.S., Zhang, J., 2002. Prediction of food thermal process evaluation parameters using neural networks. International Journal of Food Microbiology 79 (3), 153–159.

Noronha, J., Hendrickx, A., Van Loey, A., Tobback, P., 1995. New semi-empirical approach to handle time-variable boundary conditions during sterilization of non-conductive heating foods. Journal of Food Engineering 24 (2), 249–268.

Sablani, S.S., Ramaswamy, H.S., Prasher, S.O., 1995. A neural network approach for thermal processing applications. Journal of Food Processing and Preservation 19 (4), 283–301.

Sablani, S.S., Ramaswamy, H.S., Prasher, S.O., 1997. A neural network modeling of heat transfer to liquid particle mixtures in cans subjected to end-over-end applications. Food Research International 30 (2), 105–116.

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.

Stumbo, C.R., 1973. Thermobacteriology in Food Processing, second ed. Academic Press, New York.

Tandon, S., Bhowmik, S.R., 1986. Evaluation of thermal processing of retortable pouches filled with conduction heated foods considering their actual shapes. Journal of Food Science 51 (3), 709–714.

Tattiyakul, J., Rao, M.A., 2000. Rheological behavior of cross-linked waxy maize starch dispersions during and after heating. Carbohydrate Polymers 43 (3), 215–222.

Tattiyakul, J., Rao, M.A., Datta, A.K., 2001. Simulation of heat transfer to a canned corn starch dispersion subjected to axial rotation. Chemical Engineering and Processing 40 (4), 391–399.

Tattiyakul, J., Rao, M.A., Datta, A.K., 2002. Heat transfer to a canned corn starch dispersion under intermittent agitation. Journal of Food Engineering 54 (4), 321–329.

Teixeira, A.A., Manson, J.E., 1982. Computer control of batch retort operations with on-line correction of process deviations. Food Technology 36 (4), 85–90.

Tucker, G.S., 1991. Development and use of numerical techniques for improved thermal process calculations and control. Food Control 2 (1), 15–19.

Tucker, G.S., Holdsworth, S.D., 1991. Mathematical modeling of sterilization and cooking processes for heat preserved foods, applications of a new heat transfer model. Transactions of IChemE 69C, 5–12.

Yang, W.H., Rao, M.A., 1998. Numerical study of parameters affecting broken-heating curve. Journal of Food Engineering 37 (1), 43–61.